



(51) International Patent Classification:

A61K 38/46 (2006.01) C12N 15/11 (2006.01)

C12N 9/22 (2006.01) C12N 15/87 (2006.01)

C12N 15/10 (2006.01)

(21) International Application Number:

PCT/US2021/056361

(22) International Filing Date:

22 October 2021 (22.10.2021)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/105,177 23 October 2020 (23.10.2020) US

63/105,191 23 October 2020 (23.10.2020) US

63/156,857 04 March 2021 (04.03.2021) US

63/195,659 01 June 2021 (01.06.2021) US

63/235,583 20 August 2021 (20.08.2021) US

(71) Applicants: THE BROAD INSTITUTE, INC. [US/US];

415 Main Street, Cambridge, Massachusetts 02142 (US).

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

[US/US]; 77 Massachusetts Avenue, Cambridge, Massa-

chusetts 02139 (US).

(72) Inventor; and

(71) Applicant: ZHANG, Feng [US/US]; c/o 415 Main Street, Cambridge, Massachusetts 02142 (US).

(72) Inventors: ALTAE-TRAN, Han; c/o 77 Massachusetts

Avenue, Cambridge, Massachusetts 02139 (US). KAN-

NAN, Soumya; c/o 77 Massachusetts Avenue, Cambridge,

Massachusetts 02139 (US). DEMIRCI OGLU, Esra; c/o 415 Main Street, Cambridge, Massachusetts 02142 (US).

(74) Agent: REEVES, Andrew R. et al.; Johnson, Marcou, Isaacs & Nix, LLC, P.O. Box 691, Hoschton, Georgia 30548 (US).

(81) Designated States (unless otherwise indicated, for every

kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every

kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(54) Title: REPROGRAMMABLE ISCB NUCLEASES AND USES THEREOF

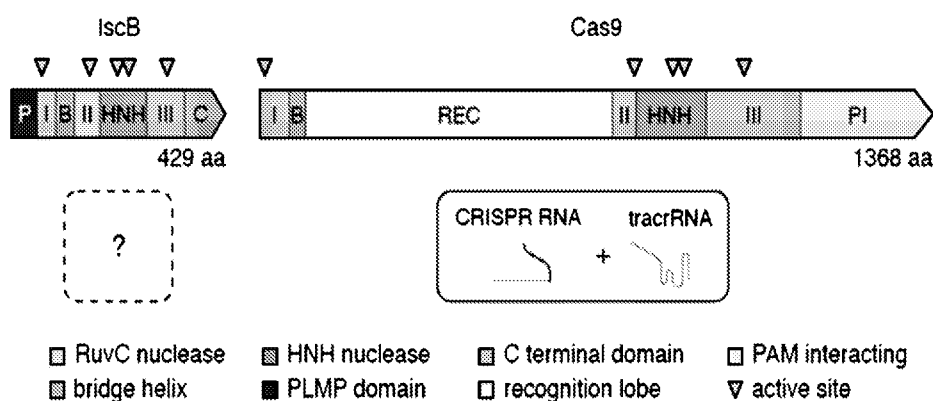


FIG. 9A

(57) **Abstract:** Systems, methods and compositions for targeting polynucleotides are detailed herein. In particular, engineered DNA-targeting systems comprising IscB polypeptides, novel IscB nucleases and reprogrammable targeting nucleic acid components and methods and application of use are provided.

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
- *with sequence listing part of description (Rule 5.2(a))*

REPROGRAMMABLE ISCB NUCLEASES AND USES THEREOF

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of and priority to U.S. Provisional Application No. 63/105,191, filed October 23, 2020, entitled “Reprogrammable IscB Polypeptide Nucleases and Use Thereof”, U.S. Provisional No. 63/105,177, filed October 23, 2020, entitled “Nucleic Acid-Guided Nucleases and Use Thereof,” U.S. Provisional Application No. 63/156,857, filed March 4, 2021, entitled “Reprogrammable IscB Polypeptide Nucleases and Use Thereof”, U.S. Provisional Application No. 63/195,659, filed June 1, 2021, entitled “Reprogrammable IscB Polypeptide Nucleases and Use Thereof”, and U.S. Provisional Application No. 63/235,583, filed August 20, 2021, entitled “Reprogrammable IscB Polypeptide Nucleases and Use Thereof”, the contents of which are incorporated by reference in their entirety herein.

STATEMENT AS TO FEDERALLY SPONSORED RESEARCH

[0002] This invention was made with government support under Grant Nos. HL141201 and HG09761 awarded by The National Institutes of Health. The government has certain rights in the invention.

REFERENCE TO AN ELECTRONIC SEQUENCE LISTING

[0003] The contents of the electronic sequence listing ("BROD-5290WP_ST25.txt"; Size is 3,078,736 bytes and it was created on October 22, 2021) is herein incorporated by reference in its entirety.

TECHNICAL FIELD

[0004] The subject matter disclosed herein is generally directed to systems, methods and compositions used for targeted gene modification and nucleic acid editing utilizing systems comprising Isc polypeptides. In particular, the present disclosure provides DNA or RNA-targeting compositions comprising novel DNA or RNA-targeting nucleases and at least one targeting nucleic acid component.

BACKGROUND

[0005] While there are genome-editing techniques available for producing targeted genome perturbations, there remains a pressing need for new and alternative genome engineering technologies that employ robust novel strategies and molecular mechanisms and are affordable, easy to set up, scalable, and amenable to targeting multiple positions within the genome. The CRISPR-Cas systems of bacterial and archaeal adaptive immunity are some such systems that show extreme diversity of protein composition and genomic loci architecture. These additional desirable tools in genome engineering and biotechnology would further advance the art.

[0006] Citation or identification of any document in this application is not an admission that such a document is available as prior art to the present invention.

SUMMARY

[0007] In certain example embodiments, non-naturally occurring, engineered compositions comprising an IscB polypeptide comprising a split Ruv-C nuclease domain comprising Ruv-C I, Ruv-CII, and Ruv-CIII subdomains, an HNH domain or both and b) an ω RNA molecule comprising a scaffold and a reprogrammable spacer sequence, the ω RNA molecule capable of forming a complex with the IscB polypeptide and directing the IscB polypeptide to a target polynucleotide.

[0008] The IscB polypeptides may further comprise a N-terminal PLMP domain and/or a conserved C-terminal domain.

[0009] In one embodiment, the IscB polypeptides comprise both a HNH and a split RuvC domain. The HNH domain is located between the Ruv-C II and RuvC-III subdomains. In other embodiments, the IscB polypeptide comprises a split RuvC domain but no HNH domain. In yet other embodiments, the IscB polypeptide comprises a split RuvC domain and no HNH domain.

[0010] In embodiments, the IscB polypeptide comprises about 200 to about 1000 amino acids. The composition may comprise a reprogrammable spacer sequence of 10 nucleotides to 150 nucleotides in length, more preferably about 15 to 45 nucleotides in length. In embodiments, the TAM sequence is 3' of the target polynucleotide.

[0011] In embodiments, the target polynucleotide is DNA. In an aspect, the ω RNA further comprises an aptamer. In an embodiment, the ω RNA molecule further comprises an extension to add an RNA template.

[0012] In embodiments, the composition of may comprising a functional domain associated with the IscB protein. In an aspect, the functional domain has transposase activity, methylase activity, demethylase activity, translation activation activity, translation repression activity, transcription activation activity, transcription repression activity, transcription release factor activity, chromatin modifying or remodeling activity, histone modification activity, nuclease activity, single-strand RNA cleavage activity, double-strand RNA cleavage activity, single-strand DNA cleavage activity, double-strand DNA cleavage activity, nucleic acid binding activity, detectable activity, or any combination thereof.

[0013] In an embodiment, the composition may further comprise a homologous recombination donor template comprising a donor sequence for insertion into a target polynucleotide.

[0014] A vector system is also provided and may comprise one or more vectors encoding the Isc polypeptide and the ω RNA compositions as detailed herein.

[0015] In embodiments, an engineered cell comprising the composition as detailed herein is provided.

[0016] Methods of modifying a target polynucleotide sequence in a cell, comprising introducing to the cell any one of the compositions as described herein are provided. In an aspect, the polypeptide and/or nucleic acid components are provided via one or more polynucleotides encoding the polypeptides and/or nucleic acid component(s), and wherein the one or more polynucleotides are operably configured to express the IscB polypeptide and/or the ω RNA molecule. In an embodiment, the method introduces one or more mutations include substitutions, deletions, and insertions.

[0017] In an aspect, the composition provides site-specific modification that may comprise cleaving a DNA polynucleotide. In an aspect, the cleaving results in a 5' overhang on a DNA molecule.

[0018] In one aspect, the present disclosure provides an engineered, non-naturally occurring composition comprising a IscB protein, wherein the IscB protein comprises an N-terminal X domain, a RuvC domain, a Bridge Helix domain, and a C-terminal Y domain.

[0019] In an embodiment, the X domain has an amino acid sequence that share at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with X domains in Table 1. In one embodiment, the Y domain has an amino acid sequence that share at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with Y domains in Table 2. In one embodiment, the IscB protein shares at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with a IscB protein selected from Tables 2 and 3.

[0020] In an embodiment, the N-terminal X domain is no more than 50 amino acids in length. In an embodiment, the composition further comprises an HNH domain. In one embodiment, the RuvC domain comprises a RuvC I subdomain, a RuvC II subdomain and a RuvC III subdomain, and the HNH is located between the RuvC II and RuvC III subdomains of the RuvC domain. In one embodiment, the IscB protein is no more than 500, no more than 600, no more than 700, or no more than 800 amino acids in length.

[0021] In one embodiment, the composition further comprises a first and second nucleic acid molecules, the first and second nucleic acid molecules capable of forming a duplex, the duplex capable of forming a complex with the IscB protein, wherein the second nucleic acid molecule is a recombinant molecule comprising a heterologous guide sequence capable of directing site-specific binding of the complex to a target sequence of a target polynucleotide. In one embodiment, the composition comprises a single guide molecule capable of forming a complex with the IscB protein and directing site-specific binding of the complex to a target sequence of a target polynucleotide.

[0022] In one embodiment, the IscB protein targets DNA. In one embodiment, the nuclease domains of the IscB protein are catalytically inactive. In one embodiment, the nuclease domain has nickase activity or is engineered to have nickase activity. In one embodiment, the composition comprises a functional domain associated with the IscB protein.

[0023] In one embodiment, the functional domain has transposase activity, methylase activity, demethylase activity, translation activation activity, translation repression activity, transcription activation activity, transcription repression activity, transcription release factor activity, chromatin modifying or remodeling activity, histone modification activity, nuclease activity, single-strand RNA cleavage activity, double-strand RNA cleavage activity, single-

strand DNA cleavage activity, double-strand DNA cleavage activity, nucleic acid binding activity, detectable activity, or any combination thereof. In one embodiment, the composition comprises a homologous recombination donor template comprising a donor sequence for insertion into a target polynucleotide. In one embodiment, the target sequence comprises a PAM of NGG or NAC, where N is A, C, G, or T.

[0024] In another aspect, the present disclosure provides one or more polynucleotides encoding one or more components of the composition herein. In another aspect, the present disclosure provides one or more vectors comprising the one or more polynucleotides herein. In another aspect, the present disclosure provides a cell or progeny thereof genetically engineered to express one or more components of the compositions herein. In another aspect, the present disclosure provides a method of targeting a polynucleotide, comprising contacting a sample that comprises a target polynucleotide with the composition herein, or the one or more polynucleotides or one or more vectors of herein.

[0025] In one embodiment, contacting results in modification of a gene product or modification of the amount or expression of a gene product. In one embodiment, the target sequence of the polynucleotide is a disease-associated target sequence.

[0026] In another aspect, the present disclosure provides an engineered, non-naturally occurring composition comprising: the IscB protein herein, wherein the IscB protein is catalytically inactive, a nucleotide deaminase associated with or otherwise capable of forming a complex with the IscB protein, and a single guide molecule capable of forming a complex with the IscB protein and directing site-specific binding at a target sequence. In one embodiment, the nucleotide deaminase is an adenosine deaminase or a cytidine deaminase.

[0027] In another aspect, the present disclosure provides one or more polynucleotides encoding one or more components of the composition herein. In another aspect, the present disclosure provides one or more vectors encoding the one or more polynucleotides herein. In another aspect, the present disclosure provides a cell or progeny thereof genetically engineered to express one or more components of the composition herein.

[0028] In another aspect, the present disclosure provides a method of editing nucleic acids in target polynucleotides comprising delivering the composition herein, the one or more polynucleotides herein, or one or more vectors herein to a cell or population of cells comprising the target polynucleotides. In one embodiment, the target polynucleotides are target sequences

within genomic DNA. In one embodiment, the target polynucleotide is edited at one or more bases to introduce a G→A or C→T mutation.

[0029] In another aspect, the present disclosure provides an isolated cell or progeny thereof comprising one or more base edits made using the method herein. In another aspect, the present disclosure provides an engineered, non-naturally occurring composition comprising: the IscB protein herein, wherein the IscB is catalytically inactive, a reverse transcriptase associated with or otherwise capable of forming a complex with the IscB protein, and a guide molecule capable of forming a complex with the IscB protein and directing site-specific binding of the complex to a target sequence of a target polynucleotide, the guide molecule further comprising a donor sequence for insertion into the target polynucleotide.

[0030] In another aspect, the present disclosure provides one or more polynucleotides encoding one or more components of the composition herein. In another aspect, the present disclosure provides one or more vectors encoding the one or more polynucleotides herein. In another aspect, the present disclosure provides a method of modifying target polynucleotides comprising: delivering the composition herein, the one or more polynucleotides herein, or one or more vectors herein to a cell or population of cells comprising the target polynucleotides, wherein the complex directs the reverse transcriptase to the target sequence and the reverse transcriptase facilitates insertion of the donor sequence from the guide molecule into the target polynucleotide.

[0031] In one embodiment, insertion of the donor sequence: introduces one or more base edits; corrects or introduces a premature stop codon; disrupts a splice site; inserts or restores a splice site; inserts a gene or gene fragment at one or both alleles of the target polynucleotide; or; a combination thereof.

[0032] In another aspect, the present disclosure provides an isolated cell or progeny thereof comprising the modifications made using the method herein. In another aspect, the present disclosure provides an engineered, non-naturally occurring composition comprising: the IscB protein herein, a non-LTR retrotransposon protein associated with or otherwise capable of forming a complex with the IscB protein; a single guide molecule capable of forming a complex with the IscB protein and directing site-specific binding to a target sequence of a target polynucleotide; and a donor construct comprising a donor polynucleotide for insertion to the target polynucleotide and located between two binding elements capable of forming a complex with the non-LTR retrotransposon protein.

[0033] In one embodiment, the IscB protein is fused to the N-terminus of the non-LTR retrotransposon protein. In one embodiment, the IscB protein is engineered to have nickase activity. In one embodiment, the guides directs the fusion protein to a target sequence 5' of the targeted insertion site, and wherein the IscB protein generates a double-strand break at the targeted insertion site. In one embodiment, the guides directs the fusion protein to a target sequence 3' of the targeted insertion site, and wherein the IscB protein generates a double-strand break at the targeted insertion site. In one embodiment, the donor polynucleotide further comprises a polymerase processing element to facilitate 3' end processing of the donor polynucleotide sequence. In one embodiment, the donor polynucleotide further comprises a homology region to the target sequence on the 5' end of the donor construct, the 3' end of the donor construct, or both. In one embodiment, the homology region is from 8 to 25 base pairs.

[0034] In another aspect, the present disclosure provides one or more polynucleotides encoding one or more components of the composition herein. In another aspect, the present disclosure provides one or more vectors comprising the one or more polynucleotides herein. In another aspect, the present disclosure provides a method of modifying target polynucleotides comprising: delivering the composition herein, the one or more polynucleotides herein, or one or more vectors herei to a cell or population of cells comprising the target polynucleotides, wherein the complex directs the non-LTR retrotransposon protein to the target sequence and the non-LTR retrotransposon protein facilitates insertion of the donor polynucleotide sequence from the donor construct into the target polynucleotide.

[0035] In one embodiment, insertion of the donor sequence: introduces one or more base edits; corrects or introduces a premature stop codon; disrupts a splice site; inserts or restores a splice site; inserts a gene or gene fragment at one or both alleles of the target polynucleotide; or; a combination thereof.

[0036] In another aspect, the present disclosure provides an isolated cell or progeny thereof comprising the modifications made using the method herein.

[0037] These and other aspects, objects, features, and advantages of the example embodiments will become apparent to those having ordinary skill in the art upon consideration of the following detailed description of illustrated example embodiments.

BRIEF DESCRIPTION OF THE DRAWINGS

[0038] An understanding of the features and advantages of the present invention will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention may be utilized, and the accompanying drawings of which:

[0039] **FIG. 1** – IscB is reprogrammable and cleaves dsDNA in a target and target adjacent motif (TAM)-specific manner. Left panel shows cleaving of endogenous spacer, right panel, cleavage of engineered spacer.

[0040] **FIG. 2** – TAM weblogo shows 3' TAM base preference of K racemifer IscB system.

[0041] **FIG. 3** – IscB sequence logo of the N-terminal domain from sequence alignment of polypeptides in Table 1, with conserved motifs boxed and annotated.

[0042] **FIG. 4-1-4-46** – includes a sequence alignment of representative IscB loci from clusters of IscB at 60% identity and 70% coverage.

[0043] **FIG. 5** – Consensus sequence from representative IscB loci of Table 1.

[0044] **FIG. 6A-6C** – **(6A)** TAM weblogo of exemplary IscB from OGEU010000025.1 **(6B)** Indel frequency compared to negative control condition at VEGFA site 2 using exemplary IscB system in HEK293 cells **(6C)** Representative indels at VEGFA site 2 from IscB mediated editing, a 20 nt guide is identified.

[0045] **FIG. 7A-7B** – **(7A)** HNH domain amino acid sequence of IscB identified in this study (OGEU010000025.1, 494 aa). **(7B)** ω RNA scaffold nucleotide sequence of IscB identified in this study (OGEU010000025.1_ ω RNA).

[0046] **FIG. 8A-8B** – **(8A)** Design of guide RNA expression plasmid, pHS0812_Isc_large_27, in backbone of pHS0728 pcDNA3.1 (+) CM. **(8B)** Design of IscB expression plasmid, pHS0810_Isc_large_27, in backbone of pHS0728 pcDNA3.1 (+) CM.

[0047] **FIG. 9A-9G** – IscBs are associated with ncRNAs of unknown function. **(9A)** Comparison of IscB and Cas9 domains and previously described ncRNAs. **(9B)** Phylogenetic analysis of the RuvC, bridge helix, and HNH domains of Cas9 and IscB clusters. Genomic association shows 15/603 IscB clusters have strong association to CRISPR, occurring independently in multiple clades. **(9C)** Small RNA-seq of a heterologously expressed locus (top) and additionally after an RNP pulldown (bottom). **(9D)** Weblogo of 3' PAMs depleted

more than 5 standard deviations relative to a non-targeting control. **(9E)** *In vitro* cleavage by IscB-single guide RNA RNP complex. **(9F)** (Top) Conservation analysis of regions upstream of *N*=563 non-redundant IscB loci. (Bottom) Small RNA-seq of an IscB locus in *K. racemifer*. **(9G)** Secondary structure predictions of CRISPR-associated IscB ncRNA and IscB ω RNA. Guiding function of ω RNAs was inferred by comparison of the two structures. TE: transposon end.

[0048] FIG. 10 – PLMP Domain. Weblogo of PLMP domain found in IscB and IsrB proteins immediately upstream of the RuvC-I domain.

[0049] FIG. 11 – Non-coding region IscB RNA examples. Associated IscB non-coding region examples folded as RNA via ViennaRNA at 55°C. Black arrows indicate GU pairs characteristic of RNA structure.

[0050] FIG. 12 – Small RNA-seq of IscB loci in *K. racemifer*. Small RNA-seq reads greater than 200 bp mapped to the 49 IscB loci present in *K. racemifer*. 38 of 49 loci contains an expressed ncRNA transcript corresponding to a guide and ω RNA scaffold upstream of the IscB ORF. Loci with low or undetectable levels of ω RNA are annotated based on computational prediction of the ω RNA scaffold but the guide is not annotated.

[0051] FIG. 13A-13C - Characterization of KraIscB-1 reprogramming and cleavage. **(13A)** Small RNA-seq of recombinantly purified KraIscB-1 in the presence of its endogenous locus. The predicted ω RNA scaffold along with an upstream region co-purified with KraIscB-1 protein, indicating physical interaction of the ω RNA with KraIscB-1. **(13B)** KraIscB-1 is a reprogrammable dsDNA nuclease. IVTT reactions with KraIscB-1 and ω RNAs with endogenous or reprogrammed guide sequences incubated with cognate or incorrect targets demonstrates TAM and target-dependent cleavage. Reactions were run on native PAGE gels and imaged in IR800 and IR700 channels to capture target strand (TS) and non-target strand (NTS) cleavage products, respectively. **(13C)** Substrate cleavage by wild-type and nuclease domain mutants of KraIscB-1 demonstrates strand-specific cleavage by each nuclease domain.

[0052] FIG. 14A-14B – CRISPR-associated IscB ncRNA pseudoknot plays a necessary role in target cleavage. **(14A)** CRISPR-associated IscB ncRNA variants tested. The leftmost sequence is the endogenous sequence. Middle sequence (ncRNA 1) is mutated (blue) on the nexus-adjacent region to abolish predicted base-pairing interactions in the pseudoknot. The rightmost sequence (ncRNA 2) contains mutations in both strands of the pseudoknot (blue) such that predicted base pairing is retained. **(14B)** IVTT cleavage assays with CRISPR-

associated IscB and ncRNA variants shows that mutations which abolish the predicted base-pairing (ncRNA 1) also abolish activity, whereas compensatory mutations that retain the predicted base-pairing interaction (ncRNA 2) allow for target cleavage, implying that the pseudoknot structure plays a necessary functional role in CRISPR-associated IscB-mediated target cleavage.

[0053] FIG. 15A-15G - IscB is an RNA-guided DNA endonuclease. (15A) Design of an IVTT-based TAM screen. **(15B)** KraIscB-1 endogenous target and reprogrammed target sequences used in IVTT TAM screens. **(15C)** KraIscB-1 cleaves DNA in an ω RNA-dependent manner with an ATAAA 3' TAM. **(15D)** AwaIscB cleaves DNA with an ATGA 3' TAM. **(15E)** *In vitro*-reconstituted AwaIscB- ω RNA RNP cleavage of dsDNA substrates in the presence or absence of a target and/or TAM. **(15F)** *In vitro* cleavage of AwaIscB with selectively inactivated nuclease domains. **(15G)** Sequencing of cleavage products generated by AwaIscB.

[0054] FIG. 16A-16D - Guide-encoding mechanisms of IscB. (16A) Example loci for each major mechanism of encoding multiple guides. top to bottom: 1) ω RNAs duplicate or insert into CRISPRs, 2) entire ω RNAs arrays associate with IscB, 3) transposition expansion results in multiple nearly identical loci in each expressing different guides, 4) standalone *trans*-acting ω RNAs form independently of adjacent IscBs. **(16B)** *K. racemifer* encodes 48 IscB loci with *cis* ω RNAs and 10 standalone *trans*-acting ω RNAs. **(16C)** Expression of standalone ω RNAs in *K. racemifer*. **(16D)** KraIscB-1, in complex with *cis* or *trans* ω RNAs with the same guide sequence, mediate cleavage of dsDNA in a TAM and target-dependent manner. Reactions were performed in IVTT using 5' strand-specific labeled linear targets.

[0055] FIG. 17A-17G – Biochemical properties of AwaIscB. (17A) Target cleavage by AwaIscB at various temperatures. Reactions were performed at the indicated temperature for 1 hour, run on native PAGE gels and stained with SYBR Gold for imaging. Optimal cleavage activity is observed between 35-40°C. **(17B)** Kinetics of AwaIscB target cleavage. Reactions were performed at 37°C and stopped by addition of EDTA at the indicated times, run on a native PAGE gel and stained with SYBR Gold for imaging. Cleavage activity is saturated after 60 min. **(17C)** Target cleavage by AwaIscB in the presence of various divalent metal ions. AwaIscB requires Mg^{2+} for optimal activity, but can mediate target cleavage in the presence of Ca^{2+} . **(17D)** Guide length optimization for AwaIscB. Cleavage activity is supported with 11-12 nt guides, but at least 17-18 nt guides are required for robust activity. In C and D, all

reactions were performed at 37°C for 1 hour, run on native PAGE gels and stained with SYBR Gold for imaging. **(17E)** Cy5.5-labeled ssDNA cleavage by AwaIscB wild type and nuclease domain catalytic mutants. Reactions were performed at 37°C for 1 hour, run on denaturing PAGE gels and imaged in the IR700 channel. AwaIscB exhibits weak TAM-independent but target-dependent activity, with specific cleavage products generated by each nuclease domain. Cleavage activity of the HNH domain is enhanced in RuvC-inactivated AwaIscB in a TAM-dependent manner. Cleavage activity is abolished upon mutation of both nuclease domains. **(17F)** Cy5-labeled ssRNA cleavage by AwaIscB wild type and nuclease domain catalytic mutants. Reactions were performed at 37°C for 1 hour, run on denaturing PAGE gels and imaged in the Cy5 channel. No cleavage activity is observed on ssRNA substrates by AwaIscB. **(17G)** Collateral activity of AwaIscB. Wild-type or RuvC-inactivated AwaIscB were incubated with unlabeled dsDNA or ssDNA targets and a Cy5.5-labeled collateral ssDNA substrate for 3 hours at 37°C. Reactions were run on denaturing PAGE gel and imaged in the IR700 channel to capture cleavage of the collateral substrate. No collateral activity is observed.

[0056] FIG. 18A-18B – Target cleavage site mapping of awaiscb nickase mutants. Sequencing of cleavage products from **(18A)** Awaiscb RuvC-II (e157a) and **(18B)** hnh (h212a) catalytic mutants demonstrates strand-specific nicking of targeted strand by the hnh domain 3 nt downstream of the tam and non-targeted strand by the ruvc domain 8-16 nt upstream of the TAM.

[0057] FIG. 19A-19E – Exonuclease III footprinting of dAwaIscB ternary complex. **(19A)** Schematic of Exonuclease III (ExoIII) footprinting experiment. Catalytically inactivated AwaIscB (dAwaIscB)-ωRNA complex bound to a target dsDNA substrate is digested with Exo III. ExoIII is sterically hindered when the dAwaIscB RNP complex is reached. Quenched reactions are subjected to ligation of adapters for next-generation sequencing, and position of adapter ligation allows for inference of the position of ExoIII hindrance, indicating protection by the dAwaIscB RNP complex. **(19B-C)** 3' adapter ligation position after ExoIII treatment of dAwaIscB with and without ωRNA, respectively. Specific protection of the target strand 19 nt upstream of the TAM and the non-target strand 6 nt downstream of the target sequence is observed in the ωRNA condition, in contrast to a low level of non-specific adapter ligation when the ωRNA is not present. **(19D-E)** dSpCas9 with or without a corresponding sgRNA, respectively, was assayed as a positive control. Results shown in **(19D)** replicate previously reported results using a gel-based readout.

[0058] **FIG. 20** – Distribution of loci counts.

[0059] **FIG. 21** – Small RNA-seq of standalone ω RNAs in *K. racemifer*. Small RNA-seq reads greater than 200 bp mapped to standalone ω RNA loci in *K. racemifer*. 9 of the 10 loci contain an expressed ncRNA transcript corresponding to a guide and ω RNA scaffold. The ω RNA scaffold that is not expressed belongs to a group associated primarily with IsrB (G1c group - see FIG. 40).

[0060] **FIG. 22A-22B** – Likelihood mapping of main alignments. **(22A)** Likelihood mapping analysis for the main alignments used in this study performed using IQ Tree 2. The PLMP aa alignment displays high star-like behavior due to the presence of many divergent sequences. **(22B)** Results for statistical analysis assessing whether or not phylogenetic assumptions hold for the main alignments used in this study. 3 types of tests were performed using IQ Tree 2: symmetry (sym), marginal symmetry (mar), and internal symmetry (sym). P-values indicating severe violations ($p < 0.01$) are shown in bold. The RuvC/BH/HNH aa alignment containing IscBs and Cas9s had a significant p-value for the marginal symmetry test, indicating it likely violates the stationarity assumption of typical phylogenetic analysis. Similarly, the hi-res full CDS DNA alignment of early Cas9s violates the stationarity assumption. No alignments had significant p-values for the internal symmetry test, suggesting that they might not violate the homogeneity assumption.

[0061] **FIG. 23** – Complete RuvC/BH phylogenetic analysis with IQ Tree 2. Maximum likelihood phylogenetic analysis of all IsrB, IscB and Cas9 RuvC/BH domains using IQ Tree 2. The LG substitution model with Gamma rates with 4 categories was used with 5000 ultra fast bootstraps (with hill-climbing nearest neighbor change for each bootstrap tree). The tree was rooted on the IsrB family. Associations are calculated for each cluster based on non-redundant loci (at 90% sequence identity of the main locus ORF). Ga-Gi refer to the IscB/IsrB main ω RNA profiles in FIG. 40. HNH domain associations are shown with 3 colors, with cyan indicating the HNH domain has the H, N, and H catalytic residues, magenta indicating the HNH domain has the H, N, and N catalytic residues, and grey indicating that the HNH domain has an H, N, and not H/N catalytic residues. Sizes of REC-like insertion in the representative protein sequence for each cluster are shown as determined by the number of amino acids between the BH and the RuvC-II in the alignment. Total size of the representative protein sequence for each cluster is shown on the outer ring.

[0062] FIG. 24 – Complete RuvC/BH/HNH phylogenetic (IQ Tree 2) x 5000 UFbs tree with associations. Maximum likelihood phylogenetic analysis of all IscB and Cas9 RuvC/BH/HNH domains using IQ Tree 2. The LG substitution model with Gamma rates with 4 categories was used with 5000 ultra fast bootstraps (with hill-climbing nearest neighbor change for each bootstrap tree). Tree is rooted using cluster 34777, which include some of the most ancestral IscBs as determined by the RuvC/BH phylogenetic analyses. Associations are calculated for each cluster based on non-redundant loci (at 90% sequence identity of the main locus ORF). Ga-Gi refer to the IscB/IsrB main ω RNA profiles in FIG. 38A. HNH domain associations are shown with 3 colors, with cyan indicating the HNH domain has the H, N, and H catalytic residues, magenta indicating the HNH domain has the H, N, and N catalytic residues, and grey indicating that the HNH domain has an H, N, and not H/N catalytic residues. Sizes of REC-like insertion in the representative protein sequence for each cluster are shown as determined by the number of amino acids between the BH and the RuvC-II in the alignment. Total size of the representative protein sequence

[0063] FIG. 25 – Complete RuvC/BH/HNH phylogenetic (RAxML) x 2000 bs. Maximum likelihood phylogenetic analysis of all IscB and Cas9 RuvC/BH/HNH domains using RAxML. The PROTGAMMALG model was used with 2000 rapid bootstraps. Tree is rooted using cluster 34777, which include some of the most ancestral IscBs as determined by the RuvC/BH phylogenetic analyses. Associations are calculated for each cluster based on non-redundant loci (at 90% sequence identity of the main locus ORF). Ga-Gi refer to the IscB/IsrB main ω RNA profiles in FIG. 40. HNH domain associations are shown with 3 colors, with cyan indicating the HNH domain has the H, N, and H catalytic residues, magenta indicating the HNH domain has the H, N, and N catalytic residues, and grey indicating that the HNH domain has an H, N, and not H/N catalytic residues. Sizes of REC-like insertion in the representative protein sequence for each cluster are shown as determined by the number of amino acids between the BH and the RuvC-II in the alignment. Total size of the representative protein sequence for each cluster is shown on the outer ring.

[0064] FIG. 26 – Complete RuvC/BH/HNH phylogenetic (mrBayes) x 10M iterations. Bayesian phylogenetic analysis of IscB and early Cas9 RuvC/BH/HNH domains using MrBayes with random starting trees. The LG substitution model was used with Gamma rates with 4 categories. 4 independent runs were run with 16 chains per with a delta temperature of 0.025 per chain for a total of 10M generations. 1000 swaps were attempted each generation,

and tree samples were collected every 50 generations. The average standard deviation of split frequencies was 0.057890 at the final generation. Associations are calculated for each cluster based on non-redundant loci (at 90% sequence identity of the main locus ORF). Ga-Gi refer to the IscB/IsrB main ω RNA profiles in FIG. 40. HNH domain associations are shown with 3 colors, with cyan indicating the HNH domain has the H, N, and H catalytic residues, magenta indicating the HNH domain has the H, N, and N catalytic residues, and grey indicating that the HNH domain has an H, N, and not H/N catalytic residues. Sizes of REC-like insertion in the representative protein sequence for each cluster are shown as determined by the number of amino acids between the BH and the RuvC-II in the alignment. Total size of the representative protein sequence for each cluster is shown on the outer ring.

[0065] FIG. 27 – Same phylogenetic tree as FIG. 26 with a focus on early Cas9 evolution. Bayesian posterior probabilities for each branch are shown along with the standard deviation of the posterior across all 4 runs.

[0066] FIG. 28 – High resolution early Cas9 evolution tree (aa model) (IQ Tree 2). Maximum likelihood phylogenetic analysis of early Cas9 evolution complete protein sequences (excluding large portions of Cas9 specific REC-like insertions) using IQ Tree 2. The WAG substitution model with empirical amino acid frequencies, invariant sites, and Gamma rates with 4 categories was used with 5000 ultra fast bootstraps (with hill-climbing nearest neighbor change for each bootstrap tree). Tree is rooted using a representative from cluster 18054, which is more distantly related to the other sequences as determined by RuvC/BH/HNH trees. Support values are shown above each branch.

[0067] FIG. 29 – Complete RuvC/BH phylogenetic analysis with IQ Tree 2. Maximum likelihood phylogenetic analysis of all IsrB, IscB and Cas9 RuvC/BH domains using IQ Tree 2. The LG substitution model with Gamma rates with 4 categories was used with 5000 ultra fast bootstraps (with hill-climbing nearest neighbor change for each bootstrap tree). The tree was rooted on the IsrB family. Associations are calculated for each cluster based on non-redundant loci (at 90% sequence identity of the main locus ORF). Ga-Gi refer to the IscB/IsrB main ω RNA profiles in FIG. 40. HNH domain associations are shown with 3 colors, with cyan indicating the HNH domain has the H, N, and H catalytic residues, magenta indicating the HNH domain has the H, N, and N catalytic residues, and grey indicating that the HNH domain has an H, N, and not H/N catalytic residues. Sizes of REC-like insertion in the representative protein sequence for each cluster are shown as determined by the number of amino acids

between the BH and the RuvC-II in the alignment. Total size of the representative protein sequence for each cluster is shown on the outer ring.

[0068] FIG. 30 – IscB/IsrB ω RNA phylogenetic analysis focused on Cas9 evolution. Same phylogenetic tree is FIG. 39 but focused on the early Cas9 evolution with the CRISPR-associated IscB cluster 2089. Support values for each branching are shown above the branches. Not all clusters included in other phylogenetic analyses could not be included in this analysis due to lack of a completely alignable ω RNA. For example, clusters 57212 and 50962 were not included. Clusters 2964, 21041, 57212, and 50962 were inferred as ancestral relative to the CRISPR-associated IscB cluster 2089 for the RuvC/BH/HNH amino acid phylogenetic analyses with RAxML (FIG. 37).

[0069] FIG. 31A-31C – Diversity and evolution of IscB. **(31A)** Phylogenetic tree of IsrB, IscB and Cas9. Associations with IS200/605 TnpA, ω RNA, CRISPR arrays, anti-repeats (where applicable), and *Cas* acquisition genes. ORF size of cluster representative is shown on the outermost ring. Positions of evolutionary events described in **(31A)** are marked by colored circles/squares. **(31B)** Inferred evolutionary timeline linking IsrB to Cas9 with exemplifying loci. **(31C)** Structural diversity and evolution of ω RNAs in IsrB and IscB systems.

[0070] FIG. 32A-32B – High resolution early Cas9 evolution tree (aa model) (IQ Tree 2). **(32A)** Maximum likelihood phylogenetic analysis of early Cas9 evolution complete protein sequences (excluding large portions of Cas9 specific REC-like insertions) using IQ Tree 2. The WAG substitution model with empirical amino acid frequencies, invariant sites, and Gamma rates with 4 categories was used with 5000 ultra fast bootstraps (with hill-climbing nearest neighbor change for each bootstrap tree). Tree is rooted using a representative from cluster 18054, which is more distantly related to the other sequences as determined by RuvC/BH/HNH trees. Support values are shown above each branch. **(32B)** Bayesian phylogenetic analysis of the high resolution early Cas9 amino acid alignment. MrBayes was run with 8 independent runs with 16 chains and a temperature delta of 0.025 per chain for 1M generations with 1000 swaps attempted each generation. The model parameters were LG substitution model and Gamma rates with 4 categories. MCMC samples were collected from each cold chain every 50 generations. The average standard deviation of split frequencies was 0.005069 at the final generation. Each leaf in the tree corresponds to an individual locus with the cluster id preceding the contig accession number separated by an underscore. Taxon 18054_CP026721.1 was included as a more distant IscB in the alignment and selected as the outgroup. Posterior branch

probabilities (percentages) are displayed along with the standard deviation computed across all 8 runs with branch colors ranging from red (probability 0.7) to black (probability 1.0).

[0071] FIG. 33A-33C – High resolution early Cas9 evolution tree (dna model). **(33A)** Maximum likelihood phylogenetic analysis of early Cas9 evolution CDS DNA sequences using IQ-Tree 2. The GTR substitution model with empirical amino acid frequencies, invariant sites, and Gamma rates with 4 categories was used with 5000 ultrafast bootstraps (with hill-climbing nearest neighbor change for each bootstrap tree). Tree is rooted using a representative from cluster 18054, which is more distantly related to the other sequences as determined by RuvC/BH/HNH trees. Support values are shown above each branch. **(33B)** Same as (A) except with the GHOST heterotachy mixture model with 2 mixture classes in place of Gamma rates (Crotty, S. et al. (2020), *Syst. Biol.* **69**, 249–264). Bootstrap support values are shown for each branch, followed by the corresponding branch length for each mixture tree separated by a backslash. Taxa are shown at each leaf followed by the corresponding branch length for each mixture tree. **(33C)** Bayesian phylogenetic analysis of the high resolution early Cas9 DNA alignment. MrBayes was run with 8 independent runs with random starting trees, and 16 chains with a temperature delta of 0.01 per chain for 2M generations with 1000 swaps attempted each generation. The model parameters were GTR substitution model and Gamma rates with 4 categories. MCMC samples were collected from each cold chain every 50 generations. The average standard deviation of split frequencies was 0.043215 at the final generation. Each leaf in the tree corresponds to an individual locus with the cluster id preceding the contig accession number separated by an underscore. Taxon 18054_CP026721.1 was included as a more distant IscB in the alignment and selected as the outgroup. Posterior branch probabilities (percentages) are displayed along with the standard deviation computed across all 8 runs with branch colors ranging from red (probability 0.7) to black (probability 1.0).

[0072] FIG. 34A-34B – Early Cas9 phylogeny using maximum likelihood Phylogenetic analysis of the RuvC/BH/HNH domains of early Cas9s and all IscBs using IQ Tree 2. Each tree is the best scoring ML tree of 5 independent runs. Bootstrap supports were computed with 5000 ultrafast bootstraps. **(34A)** Phylogenetic analysis using the LG substitution model with gamma rates (4 categories). **(34B)** Phylogenetic analysis using the LG substitution model with invariant sites and gamma rates (4 categories).

[0073] FIG. 35A-35D – Sensitivity analysis for inferred Cas9 ancestor **(35A)** RAxML maximum likelihood phylogenetic tree of the RuvC/BH/HNH alignment with 2000 rapid boot

straps for computing support values. Only sections of the tree relevant to the early evolution of Cas9 are shown. **(35B)** BLOSUM62 similarity comparison of the RuvC-I, RuvC-II, RuvC-III, and HNH core regions (with alignment trimming, alignments provided in supplementary file XXX) for early Cas9 II-D (clusters Cas9_1261, Cas9_665, Cas9_1079), a typical Cas9 (cluster Cas9_758), the putative Cas9 ancestor (2089), and example IscBs. **(35C-35D)** random taxon dropout analysis using FastTree2. Sample size for each dropout percentage category was calculated such that each taxon is retained on average for 1000 bootstrap samples. Clusters 2089, Cas9_1079, Cas9_665, and Cas9_1261 were retained in all samples. Error bars were calculated using 2000 bootstraps from the final samples. **(35C)** proportion of trees supporting CRISPR-associated IscB 2089 as the direct ancestor of all Cas9s as a function of taxa dropout rate. **(35D)** proportion of trees supporting mono/paraphyletic topologies involving Cas9, IsrB, or early II-D Cas9s as a function of the taxa dropout rate.

[0074] FIG. 36 – Comparison of early Cas9 tracrRNAs to conserved ω RNAs from IscB and IsrB. ω RNA from the putative ancestor of all Cas9s (2089) is shown as well. Conserved region shared by the tracrRNA and IscB/IsrB ω RNAs corresponds to the nexus pseudoknot hairpin. Alignment was generated using MAFFT-ginsi. Additional, less conserved regions are not shown for this alignment. Specifically, the 5' end is not conserved between tracrRNA and IscB ω RNAs.

[0075] FIG. 37 – IscB/IsrB ω RNA phylogenetic analysis using IQ Tree 2. Maximum likelihood phylogenetic tree inference for the DNA alignment of ω RNA from IscB/IsrBs using IQ Tree 2. This tree was built using the best likelihood scoring tree of 200 independent runs as the starting tree with 5000 ultra fast bootstraps (with hill-climbing nearest neighbor change for each bootstrap tree) under the GTR substitution model, using empirical DNA frequencies from the alignment, ascertainment bias correction, and Gamma rates with 4 categories.

[0076] FIG. 38A-38B – Diverse ω RNAs associated with *isrB* and *iscB*. Secondary structure predictions for the main groups of ω RNA scaffolds associated with *iscBs* and *iscBs*. **(38A)** G1a, G1d, G1e, G1f, G1g, and G1i are associated with *iscB* while **(38B)** G1b, G1c, G1h are associated with *isrB*. G1a, G1b, G1c, G1d, G1h, and G1i secondary structures were predicted using R-scape while G1e, G1f, G1g were computed using consensus secondary structures with ViennaRNA due to the smaller sample sizes. While pseudoknots were not identified de novo for G1e, G2f, G1g, potential pseudoknots in similar locations to the other

iscB/iscB ω RNAs can be found. Guide locations for all *iscB/iscB* ω RNAs would be predicted to be immediately upstream from each ω RNA scaffold where the 5' label is located.

[0077] FIG. 39A-39J – Exploration of the diversity of IS200/605 superfamily nucleases. **(39A)** Evolution between IS200/605 transposon superfamily-encoded nucleases and associated RNAs. Dashed lines reflect tentative/unknown relationships. **(39B)** Locations of *IscB* loci and fragments in the *I. tetrasporus* genome. Intact locus is labeled as “ChlorIscB.” **(39C)** Small RNA-seq of *I. tetrasporus*. **(39D)** Weblogo of ChlorIscB cleavage TAM using a reprogrammed guide in an IVTT TAM screen. **(39E)** Weblogo of OgeuIscB TAM using a reprogrammed guide in an IVTT TAM screen. **(39F)** Targeted OgeuIscB mediated indel formation in HEK293FT cells ordered by abundance, with indel size on the left. **(39G)** OgeuIscB mediated indel formation at multiple sites in HEK293T cells (* indicates $p < 0.05$). **(39H)** Native expression of *IsrB* ω RNA in *K. racemifer*. **(39I)** Weblogo of *Desulfovibrio thermocuniculi* (DthIscB) TAM using a reprogrammed guide in an IVTT TAM screen. **(39J)** DthIscB mediates ω RNA-guided non-target strand nicking in a TAM- and target-dependent manner in an IVTT cleavage assay using 5' strand-specific labeled targets.

[0078] FIG. 40A-40C – Genome editing in human cells with OgeuIscB. **(40A)** Schematic of experiment to screen large *IscB* proteins for indel-generating activity in HEK293FT cells. Plasmids expressing the protein of interest were co-transfected with a mini-library of 12 ω RNAs targeting various loci in the human genome. After approximately 3 days, genomic DNA was harvested and amplicons containing loci targeted by each ω RNA in the sample were amplified and sequenced to determine indel rates. **(40B)** Targeting OgeuIscB to 3 human genomic loci in HEK293FT cells with ω RNAs containing guides of various lengths shows that a 16 nt guide generally mediates optimal indel formation. NT: non-targeting ω RNA. Statistical significance was assessed using a two-tailed T-test with the non-targeting ω RNA as the null condition, * $p < 0.05$. **(40C)** Additional genomic loci targeted by OgeuIscB using ω RNAs with 16 nt guides. Statistical significance was assessed using a two-tailed T-test with the non-targeting ω RNA as the null condition, * $p < 0.05$.

[0079] FIG. 41 – Small RNA-seq of *IsrB* loci from *K. racemifer* shows expressed associated ω RNAs. Small RNA-seq reads greater than 200 bp mapped to the 5 *IsrB* loci present in *K. racemifer*. Each locus contains an expressed ncRNA transcript corresponding to a guide and ω RNA scaffold upstream of the *IsrB* ORF.

[0080] FIG. 42A-42C – IsrB nicks dsDNA in a target and TAM-dependent manner. (42A) Target cleavage by DthIsrB at various temperatures from 40 C to 70 C at 5 C increments. All cleavage reactions were performed using RNP complexes produced by IVTT reactions for 1 hour at the indicated temperatures, run on denaturing PAGE gels, and imaged in the IR800 and IR700 channels. Optimal temperature for nicking activity is approximately 60 C. Additionally, double-stranded cleavage was not observed at any temperature. **(42B)** Target cleavage by DchIsrB at various temperatures from 30 C to 60 C at 5 C increments. All cleavage reactions were performed using NP complexes produced by IVTT reactions for 1 hour at the indicated temperatures for 1 hour at the indicated temperatures, run on denaturing PAGE gels, and imaged in the IR700 and IR800 channels. Optimal temperature for nicking activity is approximately 45°C. Double-stranded cleavage was not observed at any temperature. **(42C)** Target cleavage by DthIsrB, DchIsrB, and KraIscB-1 performed at optimal temperatures (60°C, 45 °C, and 37°C respectively). All cleavage reactions were performed using RNP complexes produced by IVTT and incubated for 1 hour at their respective temperatures. Products were run on native PAGE and denaturing PAGE gels and imaged in the IR800 and IR700 channels. DthIsrB and DchIsrB perform non-target strand dsDNA nicking with no detectable double-stranded cleavage compared to KraIscB-

[0081] FIG. 43 – Phylogenetic distribution. Distribution of IscB, IsrB, and Cas9 across archaeal and bacterial phyla. Heatmap displays percentages of genomes containing a specific system.

[0082] FIG. 44 – Examples of Type II-E Cas9 loci. ITRs are found in multiple loci, though ITRs within the same loci may not be identical. Black rectangles represent CRISPR direct repeats.

[0083] FIG. 45 – Naturally-occurring RNA-guided DNA-targeting systems. Comparison of Ω (OMEGA) systems with other known RNA-guided systems. In contrast to CRISPR systems, which capture spacer sequences and store them within the CRISPR array, in the locus, Ω systems transpose their loci (or *trans*-acting loci) into target sequences, apparently, converting targets into ω RNA guides in a process that can be called guide conscription.

[0084] FIG. 46A-46C – Activity of individual spacers from a CRISPR-associated IscB locus

(46A) Schematic of CRISPR-associated IscB locus from Chesapeake Bay sample containing three spacers flanked by four DRs in a CRISPR array. **(46B)** Spacer and corresponding 8N

PAM library targets for each spacer in the CRISPR array. PSP3 (Fn) is reprogrammed from the sequence endogenously present in the locus to the Fn spacer. **(46C)** Weblogos of 3' PAMs depleted more than 5 standard deviations relative to a non-targeting control for each protospacer library.

[0085] FIG. 47A-47B – CRISPR-associated IscB ncRNA pseudoknot plays a necessary role in target cleavage. **(47A)** CRISPR-associated IscB ncRNA nexus pseudoknot mutants tested. The leftmost sequence is the endogenous sequence. Middle sequence (ncRNA mutant 1) is mutated (blue) on the nexus- adjacent region to abolish predicted base-pairing interactions in the pseudoknot. The rightmost sequence (ncRNA mutant 2) contains mutations in both strands of the pseudoknot (blue) such that predicted base pairing is retained. **(47B)** IVTT cleavage assays with CRISPR-associated IscB and ncRNA variants shows that mutations which abolish the predicted base-pairing (ncRNA 1) also abolish activity, whereas compensatory mutations that retain the predicted base-pairing interaction (ncRNA 2) allow for target cleavage, implying that the pseudoknot structure plays a necessary functional role in CRISPR-associated IscB-mediated target cleavage.

[0086] FIG. 48 – TAMs of active IscB proteins. TAMs of active IscB proteins determined by *in vitro* plasmid cleavage assays. 57/86 of tested IscBs were found to mediate RNA-guided cleavage activity as assessed by the detection of a TAM. All tested protein sequences and accession of source contig are listed in Table 9.

[0087] FIG. 49 – PLMP domain is essential for RNA-guided cleavage function. Cell-free transcription translation cleavage assays with AwaIscB successively truncated at single aa resolution from the N-terminal end guided to labeled Fn target with an ATGAGATC 3' TAM. *In vitro* transcription/translation cleavage assays were performed as described, run on a 6% TBE- Urea gel and imaged in the Cy3 and Cy5 channels. Truncating more than 4 aa from the N- terminal PLMP domain abolished cleavage activity.

[0088] FIG. 50 – Targets of IscB/IsrB guides. Same as Fig. 52A with results of target search mapped on the second outermost ring. Notable groups are shown as labeled arcs on the outermost ring.

[0089] FIG. 51A-51C – Examples of *iscB*-containing IS200/605 insertions. **(51A)** Full view of alignment of contigs with uninserted (top) versus IS200/605 inserted (bottom) sequences. **(51B)** 5' end of alignment of uninserted (top) and inserted (bottom) locus. The inferred ω RNA guide (light gray), perfectly matches the target (dark gray), with the alignment

gap beginning at the immediate 5' end of the ω RNA scaffold. **(53C)** 3' end of alignment of uninserted (top) and inserted (bottom) locus. ATAAA, a common IscB TAM (Fig. 50), is present at the junction.

[0090] FIG. 52A-52B – Complete RuvC/BH phylogenetic analysis. **(52A)** Maximum likelihood phylogenetic analysis of all IscrB, IscrB and Cas9 RuvC/BH domains using IQ-Tree 2. The LG substitution model with Gamma rates with 4 categories was used with 5000 ultrafast bootstraps (with hill-climbing nearest neighbor change for each bootstrap tree). **(52B)** Maximum likelihood phylogenetic analysis of all IscrB, IscrB and Cas9 RuvC/BH domains using RAxML. The PROTGAMMALG model was used with 2000 rapid bootstraps. For both (52A) and (52B), the tree was rooted on the IscrB family. Associations are calculated for each cluster based on non-redundant loci (at 90% sequence identity of the main locus ORF). Ga-Gi refer to the IscrB/IscrB main ω RNA profiles in FIG. 38A. HNH domain associations are shown with 3 colors, with cyan indicating the HNH domain has the H, N, and H catalytic residues, magenta indicating the HNH domain has the H, N, and N catalytic residues, and grey indicating that the HNH domain has an H, N, and not H/N catalytic residues. Sizes of REC-like insertion in the representative protein sequence for each cluster are shown as determined by the number of amino acids between the BH and the RuvC-II in the alignment. Total size of the representative protein sequence for each cluster is shown on the second outermost ring. Notable groups are shown as labeled colored arcs on the outermost ring.

[0091] FIG. 53A-53B – Complete RuvC/BH/HNH phylogenetic analysis. **(53A)** Maximum likelihood phylogenetic analysis of all IscrB and Cas9 RuvC/BH/HNH domains using IQ-Tree 2. The LG substitution model with Gamma rates with 4 categories was used with 5000 ultrafast bootstraps (with hill-climbing nearest neighbor change for each bootstrap tree). **(53B)** Maximum likelihood phylogenetic analysis of all IscrB and Cas9 RuvC/BH/HNH domains using RAxML. The PROTGAMMALG model was used with 2000 rapid bootstraps. For both (A) and (B), tree is rooted using cluster 34777, which include some of the most ancestral IscrBs as determined by the RuvC/BH phylogenetic analyses. Associations are calculated for each cluster based on non-redundant loci (at 90% sequence identity of the main locus ORF). Ga-Gi refer to the IscrB/IscrB main ω RNA profiles in FIG. 38A. HNH domain associations are shown with 3 colors, with cyan indicating the HNH domain has the H, N, and H catalytic residues, magenta indicating the HNH domain has the H, N, and N catalytic residues, and grey indicating that the HNH domain has an H, N, and not H/N catalytic residues.

Sizes of REC-like insertion in the representative protein sequence for each cluster are shown as determined by the number of amino acids between the BH and the RuvC-II in the alignment. Total size of the representative protein sequence for each cluster is shown on the second outermost ring. Notable groups are shown as labeled colored arcs on the outermost ring.

[0092] FIG. 54A-54D – Complete RuvC/BH/HNH phylogenetic analysis of early Cas9 evolution. **(54A)** Bayesian phylogenetic analysis of IscB and early Cas9 RuvC/BH/HNH domains using MrBayes with random starting trees. The LG substitution model was used with Gamma rates with 4 categories. 4 independent runs were run with 16 chains per with a delta temperature of 0.025 per chain for a total of 10M generations on a GPU for ~10 days. 1000 swaps were attempted each generation, and tree samples were collected every 50 generations. The average standard deviation of split frequencies was 0.057890 at the final generation. Associations are calculated for each cluster based on non-redundant loci (at 90% sequence identity of the main locus ORF). Ga-Gi refer to the IscB/IsrB main ω RNA profiles in FIG. 38A. HNH domain associations are shown with 3 colors, with cyan indicating the HNH domain has the H, N, and H catalytic residues, magenta indicating the HNH domain has the H, N, and N catalytic residues, and grey indicating that the HNH domain has an H, N, and not H/N catalytic residues. Sizes of REC-like insertion in the representative protein sequence for each cluster are shown as determined by the number of amino acids between the BH and the RuvC-II in the alignment. Total size of the representative protein sequence for each cluster is shown on the second outermost ring. Notable groups are shown as labeled colored arcs on the outermost ring. **(54B)** Same phylogenetic tree as (A) with a focus on early Cas9 evolution. Bayesian posterior probabilities for each branch are shown along with the standard deviation of the posterior across all 4 runs. **(54C)-(54D)** Phylogenetic analysis of the RuvC/BH/HNH domains of early Cas9s and all IscBs using IQ-Tree 2. Each tree is the best scoring ML tree of 5 independent runs. Bootstrap supports were computed with 5000 ultrafast bootstraps. **(54C)** Phylogenetic analysis using the LG substitution model with gamma rates (4 categories). **(54D)** Phylogenetic analysis using the LG substitution model with invariant sites and gamma rates (4 categories).

[0093] FIG. 55A-55C – IscB/IsrB ω RNA phylogenetic analysis. **(55A)** Maximum likelihood phylogenetic tree inference for the DNA alignment of ω RNA from IscB/IsrBs using IQ-Tree 2. This tree was built using the best likelihood scoring tree of 200 independent runs as

the starting tree with 5000 ultrafast bootstraps (with hill-climbing nearest neighbor change for each bootstrap tree) under the GTR substitution model, using empirical DNA frequencies from the alignment, ascertainment bias correction, and Gamma rates with 4 categories. **(55B)** Same phylogenetic tree as **(55A)** but focused on the early Cas9 evolution with the CRISPR-associated IscB cluster 2089. Support values for each branching are shown above the branches. Not all clusters included in other phylogenetic analyses could not be included in this analysis due to lack of a completely alignable ω RNA. For example, clusters 57212 and 50962 were not included. Clusters 2964, 21041, 57212, and 50962 were inferred as ancestral relative to the CRISPR-associated IscB cluster 2089 for the RuvC/BH/HNH amino acid phylogenetic analyses with RAxML (Fig. 35). **(55C)** Bayesian phylogenetic analysis of tracrRNA like ω RNAs. TracrRNAs from the early Cas9 clusters Cas9_1261 and Cas9_1665 were joined with their respective DRs and separated by a 4 bp poly-A tetraloop. 23 ω RNAs sharing alignment homology to all structural regions from the two tracrRNAs were identified. The resulting 25 RNAs were then aligned with MAFFT-ginsi and manually curated to reduce gappiness. Bayesian phylogenetic analysis of the resulting alignment was performed using MrBayes with 2 chains at a delta temperature of 0.025 with 8 independent runs for 5M generations. A standard GTR model with gamma rates and 4 categories was used. Trees were sampled every 50 generations. The average standard deviation of split frequencies was 0.005966 at the final generation. Bayesian posterior probabilities for each branching are shown above the branch, along with the average standard deviation across the 8 runs. The analysis suggests that the putative modern IscB ancestor of Cas9 (IscB cluster 2089) has an ω RNA descending from the same lineage of ω RNAs that likely resulted in the DR/tracrRNA (Bayesian posterior probability 89%).

[0094] FIG. 56 – Full protein phylogenetic analysis of IscB + earliest Cas9s. Maximum likelihood phylogenetic inference of all IscBs plus earliest Cas9s (Cas9_1261, Cas9_665) with complete protein alignments excluding the PLMP domain and C terminal domain. Tree was inferred using IQ-Tree 2 with the LG substitution model and Gamma rates with 4 categories. Support values for 5000 ultrafast bootstraps are shown above each branch.

[0095] FIG. 57A-57D – Comparison of IsrB, IscB and Cas9 subtype features. **(57A)** Comparison of protein lengths between IsrB, IscB, IscB (large) and Cas9 subtypes identified in this study. The II-D Cas9 group contains members which are substantially smaller than other Cas9 subtypes, while *tnpA*-associated II-C encompasses some substantially larger members.

(57B) P-values resulting from t-tests of pairwise comparison of length distributions shown in (A). (57C) Comparison of median DR lengths for CRISPR arrays associated with IsrB, IscB, IscB (large), where CRISPR-associated, and Cas9 subtypes. Some *tnpA*-associated II-C loci contain substantially longer DRs (46-47 bp). (57D) Rate of *tnpA* association with IsrB, IscB, IscB (large) and Cas9 subtypes. 1/545 (0.2%) of unique IsrB loci, 56/2811 (2.0%) of unique IscB loci, including both IscB and IscB (large), and 115/1918 (6.0%) of unique II-C (TnpA) loci are associated with *tnpA*.

[0096] **FIG. 58** – Alignment of IscBs and early Cas9s. Alignment of early Cas9s with the founding IscB (cluster 2089) and other various IscB. Domains and conserved motifs are annotated by red arrows below the consensus alignment.

[0097] **FIG. 59A-59C** – IscB loci in *I. tetrasporus* UTEX B 2012. (59A) Alignment of IscB loci in *I. tetrasporus* UTEX B 2012 chloroplast genome. Experimentally characterized active *iscB* CDS is shown in dark red, with fragmented *iscB* CDS shown in lighter red. Top row represents consensus sequence. Second row represents percent identity over a 5 bp sliding window. (59B) Codon usage distribution of *iscB* (red bars) vs non-*iscB* (black points) CDS. (59C) Kullback-Leibler divergence of codon usage distribution in each CDS in the *I. tetrasporus* UTEX B 2012 chloroplast genome relative to the average distribution across all CDS. Experimentally characterized active *iscB* CDS is shown in red.

[0098] **FIG. 60** – *TnpB* locus conservation analysis. Conservation of the 3' end of *tnpB* loci that share the KraIscB-1 transposon end. The conserved region on the 3' region of the *tnpB* loci corresponds to the 5' region of the ω RNA of *iscB*. The conservation of the *tnpB* loci outside of the ORF on the 3' end suggests the presence of a ncRNA that may function similarly to the ω RNA of *iscB*.

[0099] **FIG. 61A-61F** – Characterization of TnpB ω RNA-guided cleavage. (61A) Small RNA-seq of *A. lobatus* DSM 43150 TnpB-2 recombinantly purified in the presence of the downstream predicted ω RNA and guide. The predicted ω RNA scaffold and a downstream region constituting the putative guide co-purified with the *A. lobatus* TnpB-2 protein, suggesting interaction of the protein with the ω RNA transcript. Contig accession and start codon information is available in Tables 11 and 13. NCBI contig accession of original locus: JACHNC010000001.1; *tnpB* start coordinate: 25000. (61B) TAM screens of additional TnpB loci. (61C) Target cleavage by AmaTnpB at various temperatures. Reactions were performed at the indicated temperature for 1 hour and subsequently run on 2% agarose gels stained with

SYBR Gold for imaging. Optimal cleavage activity is observed from 50-60°C. **(61D)** Kinetics of AmaTnpB target cleavage. Reactions were performed at 60°C, terminated by addition of EDTA at the indicated times, and subsequently run on a 2% agarose gel stained with SYBR Gold for imaging. Cleavage activity is saturated after 30 min. **(61E)** Sanger sequencing traces of AmaTnpB- digested dsDNA targets show 5' staggered overhangs. The non-templated addition of a final base is an artifact of the polymerase used in sequencing (which manifests as a terminal Adenine in the TS trace and a terminal Thymine in the NTS trace). The trace for the NTS cleavage product is reverse complemented so that both traces illustrate the sequence of the NTS. Cleavage sites are indicated by red triangles. TS: target strand; NTS: non-target strand. **(61F)** Cleavage of Cy5- labeled ssRNA by AmaTnpB. Reactions were performed at 60°C for 1 hour, run on denaturing PAGE gels and imaged in the Cy5 channel. No cleavage of RNA substrates is observed.

[0100] **FIG. 62** shows reclustering IscB at 60% sequence identity revealed novel IscB proteins.

[0101] **FIG. 63A-63C** show that the identified IscB proteins from 00644 cluster were functional with an NAC PAM sequence. **(63A)** Best fit curve and Weblogo for locus 1: JGI accession Gaa0099850_1002913; **(63B)** Best fit curve and Weblogo for Locus 2: (JGI Accession Ga0348337_018242). **(63C)** Best fit curve and Weblogo for Locus 2: (JGI Accession Ga0208542_1002724).

[0102] **FIG. 64.** PLMP domain is essential for RNA-guided cleavage function. Cell-free transcription translation cleavage assays with AwaIscB successively truncated at single aa resolution from the N-terminal end guided to labeled Fn target with an ATGAGATC 3' TAM. *In vitro* transcription/translation cleavage assays were performed as described, run on a 6% TBE- Urea gel and imaged in the Cy3 and Cy5 channels. Panels show that truncations up to 70 aa including deletion of the PLMP domain abolishes activity. For reference, the RuvC-I active aspartate is at residue 57.

[0103] The figures herein are for illustrative purposes only and are not necessarily drawn to scale.

DETAILED DESCRIPTION OF THE EXAMPLE EMBODIMENTS

General Definitions

[0104] Unless defined otherwise, technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure pertains. Definitions of common terms and techniques in molecular biology may be found in Molecular Cloning: A Laboratory Manual, 2nd edition (1989) (Sambrook, Fritsch, and Maniatis); Molecular Cloning: A Laboratory Manual, 4th edition (2012) (Green and Sambrook); Current Protocols in Molecular Biology (1987) (F.M. Ausubel et al. eds.); the series Methods in Enzymology (Academic Press, Inc.); PCR 2: A Practical Approach (1995) (M.J. MacPherson, B.D. Hames, and G.R. Taylor eds.); Antibodies, A Laboratory Manual (1988) (Harlow and Lane, eds.); Antibodies A Laboratory Manual, 2nd edition 2013 (E.A. Greenfield ed.); Animal Cell Culture (1987) (R.I. Freshney, ed.); Benjamin Lewin, Genes IX, published by Jones and Bartlet, 2008 (ISBN 0763752223); Kendrew *et al.* (eds.), The Encyclopedia of Molecular Biology, published by Blackwell Science Ltd., 1994 (ISBN 0632021829); Robert A. Meyers (ed.), Molecular Biology and Biotechnology: a Comprehensive Desk Reference, published by VCH Publishers, Inc., 1995 (ISBN 9780471185710); Singleton *et al.*, Dictionary of Microbiology and Molecular Biology 2nd ed., J. Wiley & Sons (New York, N.Y. 1994), March, Advanced Organic Chemistry Reactions, Mechanisms and Structure 4th ed., John Wiley & Sons (New York, N.Y. 1992); and Marten H. Hofker and Jan van Deursen, Transgenic Mouse Methods and Protocols, 2nd edition (2011).

[0105] As used herein, the singular forms “a”, “an”, and “the” include both singular and plural referents unless the context clearly dictates otherwise.

[0106] The term “optional” or “optionally” means that the subsequent described event, circumstance or substituent may or may not occur, and that the description includes instances where the event or circumstance occurs and instances where it does not.

[0107] The recitation of numerical ranges by endpoints includes all numbers and fractions subsumed within the respective ranges, as well as the recited endpoints.

[0108] The terms “about” or “approximately” as used herein when referring to a measurable value such as a parameter, an amount, a temporal duration, and the like, are meant to encompass variations of and from the specified value, such as variations of +/-10% or less, +/-5% or less, +/-1% or less, and +/-0.1% or less of and from the specified value, insofar such variations are appropriate to perform in the disclosed invention. For example, the amount

“about 10” includes 10 and any amounts from 9 to 11. For example, the term “about” in relation to a reference numerical value can also include a range of values plus or minus 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, or 1% from that value. It is to be understood that the value to which the modifier “about” or “approximately” refers is itself also specifically, and preferably, disclosed.

[0109] The term “about” as used herein when describing an amino acid sequence length or size or a range or ranges of amino acid sequence lengths or sizes are meant to encompass variations of and from the specified value, such as variations in amino acid length or size of +/- 5 amino acids.

[0110] As used herein, a “biological sample” may contain whole cells and/or live cells and/or cell debris. The biological sample may contain (or be derived from) a “bodily fluid”. The present invention encompasses embodiments wherein the bodily fluid is selected from amniotic fluid, aqueous humour, vitreous humour, bile, blood serum, breast milk, cerebrospinal fluid, cerumen (earwax), chyle, chyme, endolymph, perilymph, exudates, feces, female ejaculate, gastric acid, gastric juice, lymph, mucus (including nasal drainage and phlegm), pericardial fluid, peritoneal fluid, pleural fluid, pus, rheum, saliva, sebum (skin oil), semen, sputum, synovial fluid, sweat, tears, urine, vaginal secretion, vomit and mixtures of one or more thereof. Biological samples include cell cultures, bodily fluids, cell cultures from bodily fluids. Bodily fluids may be obtained from a mammal organism, for example by puncture, or other collecting or sampling procedures.

[0111] The terms “subject,” “individual,” and “patient” are used interchangeably herein to refer to a vertebrate, preferably a mammal, more preferably a human. Mammals include, but are not limited to, murines, simians, humans, farm animals, sport animals, and pets. Tissues, cells and their progeny of a biological entity obtained in vivo or cultured in vitro are also encompassed.

[0112] The term “exemplary” is used herein to mean serving as an example, instance, or illustration. Any aspect or design described herein as “exemplary” is not necessarily to be construed as preferred or advantageous over other aspects or designs. Rather, use of the word exemplary is intended to present concepts in a concrete fashion.

[0113] A protein or nucleic acid derived from a species means that the protein or nucleic acid has a sequence identical to an endogenous protein or nucleic acid or a portion thereof in the species. The protein or nucleic acid derived from the species may be directly obtained from

an organism of the species (e.g., by isolation), or may be produced, e.g., by recombination production or chemical synthesis.

[0114] Various embodiments are described hereinafter. It should be noted that the specific embodiments are not intended as an exhaustive description or as a limitation to the broader aspects discussed herein. One aspect described in conjunction with a particular embodiment is not necessarily limited to that embodiment and can be practiced with any other embodiment(s). Reference throughout this specification to “one embodiment”, “an embodiment,” “an example embodiment,” means that a particular feature, structure or characteristic described in connection with the embodiment is included in at least one embodiment of the present invention. Thus, appearances of the phrases “in one embodiment,” “in an embodiment,” or “an example embodiment” in various places throughout this specification are not necessarily all referring to the same embodiment but may. Furthermore, the particular features, structures or characteristics may be combined in any suitable manner, as would be apparent to a person skilled in the art from this disclosure, in one or more embodiments. Furthermore, while some embodiments described herein include some but not other features included in other embodiments, combinations of features of different embodiments are meant to be within the scope of the invention. For example, in the appended claims, any of the claimed embodiments can be used in any combination.

[0115] All publications, published patent documents, and patent applications cited herein are hereby incorporated by reference to the same extent as though each individual publication, published patent document, or patent application was specifically and individually indicated as being incorporated by reference.

OVERVIEW

[0116] Embodiments disclosed herein provide IscB systems that function as re-programmable nucleases. An IscB system comprises a IscB polypeptide and a nucleic acid component capable of forming a complex with the IscB polypeptide and directing the complex to a target polynucleotide. The IscB systems include homologs thereof including IsrB and IshB systems that collectively, along with TnpB systems, may be referred to as OMEGA (Obligate Mobile Element Guided Activity) systems or complexes, or Ω systems or complexes. The nucleic acid component may also be referred to herein as a ω RNA or hRNA. IscB polypeptides, and homologs thereof, are considerably smaller than other RNA-guide nucleases. As such, IscB

polypeptide represent a novel class of RNA-guided nucleases that do not suffer from the delivery size limitations of other larger single-effector, RNA-guided nucleases, such as Type II and Type V CRISPR-Cas systems. Due to their smaller size, IscBs may be combined with other functional domains, such as nucleobase deaminases, reverse transcriptases, transposases, ligases, topoisomerases, and serine and threonine recombinases and still be packaged in conventional delivery systems, like certain adenoviruses and lentiviral based viral vectors. Thus, among other improvements, the IscB system disclosed herein allows more flexible and effective strategies to manipulate and modify target polynucleotides.

[0117] In another aspect, embodiments disclosed herein include applications of the IscB systems, including diagnostics, therapeutics, and methods of detection. Delivery of the proteins and systems disclosed is also provided, including to a variety of cells and via a variety of particles and vectors.

[0118] In another aspect, embodiments disclosed herein provide additional, alternative CRISPR-associated IscB systems that function as reprogrammable nucleases comprising nucleic acid-guided nuclease compositions and methods of use thereof. In general, the compositions may comprise nucleic acid-guided nuclease with small sizes that allow more flexible and effective strategies to manipulate specific polynucleotides. In one aspect, the present disclosure provides compositions comprising nucleic acid-guided nuclease proteins (e.g., IscB proteins) that comprises an N-terminal X domain, a RuvC domain, a Bridge Helix domain, and a C-terminal Y domain. In one embodiment, the X domain may be no more than 50 amino acids in length.

[0119] In one embodiment, the CRISPR-associated IscB composition may further comprise one or more functional domains associated with the nucleic acid-guided nuclease protein, enabling various modifications of target polynucleotides. In some examples, the functional domain may be a nucleotide deaminase, e.g., for modifying a single nucleotide or base pair in a target polynucleotide. In some examples, the functional domain may be a reverse transcriptase, or a non-LTR retrotransposon, e.g., for inserting a donor polynucleotide into a desired location of a target polynucleotide, and/or replacing an existing sequence in the target polynucleotide.

[0120] In another aspect, embodiments disclosed herein include applications of the CRISPR-associated IscB compositions herein, including diagnostics, therapeutics, and

methods of detection. Delivery of the proteins and systems disclosed is also provided, including to a variety of cells and via a variety of particles, vesicles and vectors.

ISCB POLYPEPTIDES

[0121] Unless indicated otherwise, the term “IscB polypeptide” will be intended to include IscB, IsrB, and IshB. In one embodiment, IscB polypeptides of the present invention may comprise a split RuvC nuclease domain comprising RuvC-1, RuvC-II, and RuvC-III subdomains. Some IscB proteins may further comprise a HNH endonuclease domain. In one example embodiment, the RuvC endonuclease domain is split by the insertion of a bridge helix, a HNH domain, or both. However, unlike Cas9, IscB polypeptides do not contain a Rec domain. In addition, IscB polypeptides may further comprise a conserved N-terminal domain (also referred to herein as a PLMP domain), which is not present in Cas9 proteins. IscB proteins may also further comprise a conserved C-terminal domain.

[0122] In one embodiment, IscB nucleic acid-guided polypeptides may comprise CRISPR-associated IscB polypeptides. In one embodiment, the IscB polypeptides are CRISPR-associated proteins, e.g., the loci of the nucleases are associated with an CRISPR array. In one embodiment the IscBs may be referred to as Cas IscBs.

[0123] The Cas IscB nucleic acid-guided nuclease may comprise one or more domains, e.g., one or more of a X domain (e.g., at N-terminus), a RuvC domain, a Bridge Helix domain, and a Y domain (e.g., at C-terminus).

IscB

[0124] In one example embodiment, an IscB polypeptide comprises, moving from the N- to C-terminus, a PLMP domain, a RuvC-I subdomain, a bridge helix, a RuvC-II subdomain, a HNH domain, a RuvC-III subdomain, and a C terminal domain.

[0125] In certain example embodiments, the IscB polypeptides are between 180 and 800 amino acids in size, between 200 and 790 amino acids in size, between 200 and 780 amino acids in size, between 200 and 770 amino acids in size, between 200 and 760 amino acids in size, between 200 and 750 amino acids in size, between 200 and 740 amino acids in size, between 200 and 730 amino acids in size, between 200 and 720 amino acids in size, between 200 and 710 amino acids in size, between 200 and 700 amino acids in size, between 200 and 690 amino acids in size, between 200 and 680 amino acids in size, between 200 and 670 amino acids in size, between 200 and 660 amino acids in size, between 200 and 650 amino acids in size, between 200 and 640 amino acids in size,

between 200 and 630 amino acids in size, between 200 and 620 amino acids in size, between 200 and 610 amino acids in size, between 200 and 600 amino acids in size, between 200 and 590 amino acids in size, between 200 and 580 amino acids in size, between 200 and 570 amino acids in size, between 200 and 560 amino acid, between 200 between 550 amino acids, between 200 and 540 amino acids, between 200 and 530 amino acids, between 200 and 520 amino acids, between 200 and 510 amino acids, between 200 and 500 amino acids, between 200 and 490 amino acids, between 200 and 480 amino acids, between 200 and 470 amino acids, between 200 and 460 amino acids, between 200 and 450 amino acids, between 200 and 440 amino acids, between 200 and 430 amino acids, between 200 and 420 amino acids, between 200 and 410 amino acids, between 200 and 400 amino acids, between 300 and 400 amino acids. between 300 and 500 amino acids, between 300 and 600 amino acids, between 400 and 500 amino acids, or between 500-600 amino acids. In one example embodiment, the polypeptide may range in size from 400-500 amino acids, 400-490 amino acids, 400-480 amino acids, 400-470 amino acids, 400-460 amino acids, 400-450 amino acids, 400-440 amino acids, 400-430 amino acids. Size variation may be dependent, in part, on the particular domain architecture of the IscB or its homolog.

[0126] The IscB polypeptides may be derived from a naturally occurring protein, a modified naturally occurring protein, functional fragment or truncated version thereof, or a non-naturally occurring protein. In one example embodiments, the IscB polypeptide may comprise one or more domains originating from other IscB polypeptide nucleases, more particularly originating from different organisms. In an embodiment, the IscB polypeptide nucleases may be designed by *in silico* approaches. Examples of *in silico* protein design have been described in the art and are therefore known to a skilled person. In particular embodiments, the IscB polypeptide loci is not associated with a CRISPR array.

[0127] The IscB polypeptides may also encompasses homologs or orthologs of IscB polypeptides whose sequences are specifically described herein. The terms “ortholog” and “homolog” are well known in the art. By means of further guidance, a “homolog” refers to two genes that share a common ancestral gene. Homologous proteins may but need not be structurally related or are only partially structurally related. An “ortholog” are two genes that share common ancestral gene but occur in different species. Orthologous proteins may but need not be structurally related or are only partially structurally related. In one embodiment, the homolog or ortholog of a IscB polypeptide nucleases such as referred to herein has a sequence

homology or identity of at least 80%, at least 85%, at least 90%, at least 95% with a IscB polypeptide nuclease. In further embodiments, the homolog or ortholog of a IscB polypeptide nuclease has a sequence identity of at least 80%, at least 85%, at least 90%, or at least 95% with a wildtype IscB polypeptide nuclease, in particular embodiment the IscB sequence identified in Table 1 and Table 12.

RuvC domain

[0128] The RuvC domain may comprise multiple subdomains, e.g., RuvC-I, RuvC-II and RuvC-III. The subdomains may be separated by interval sequences on the amino acid sequence of the protein.

[0129] Examples of RuvC domains include any polypeptides having a structural similarity and/or sequence similarity to a RuvC domain described in the art. For example, the RuvC domain may share a structural similarity and/or sequence similarity to a RuvC of Cas9. In some examples, the RuvC domain may have an amino acid sequence that share at least 50%, at least 55%, at least 60%, at least 5%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with RuvC domains.

[0130] In some examples, the RuvC domain comprise RuvC-I polypeptide, RuvC-II polypeptide, and RuvC-III polypeptide. Examples of the RuvC-I domain also include any polypeptides having a structural similarity and/or sequence similarity to a RuvC-I domain described in the art. For example, the RuvC-I domain may share a structural similarity and/or sequence similarity to a RuvC-I of Cas9. In some examples, the RuvC domain may have an amino acid sequence that share at least 50%, at least 55%, at least 60%, at least 5%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with RuvC-I domain. The RuvC-II domain also include any polypeptides a structural similarity and/or sequence similarity to a RuvC-II domain described in the art. For example, the RuvC-II domain may share a structural similarity and/or sequence similarity to a RuvC-II of Cas9. In some examples, the RuvC domain may have an amino acid sequence that share at least 50%, at least 55%, at least 60%, at least 5%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with RuvC-II domains. The RuvC-III domain also include any polypeptides a structural similarity and/or sequence similarity to a RuvC-III domain described in the art. For example, the RuvC-III domains may share a structural similarity and/or sequence similarity to a RuvC-III of Cas9. In some examples, the RuvC domain may have an amino acid sequence that share at least 50%,

at least 55%, at least 60%, at least 5%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with RuvC-III domains.

[0131] For example, and as described in the art (e.g., Crystal structure of Cas9 in complex with guide RNA and target DNA, Nishimasu et al. Cell, 2014) the RuvC domain of Cas9 consists of a six-stranded mixed β -sheet ($\beta 1$, $\beta 2$, $\beta 5$, $\beta 11$, $\beta 14$ and $\beta 17$) flanked by α -helices ($\alpha 33$, $\alpha 34$ and $\alpha 39$ – $\alpha 45$) and two additional two-stranded antiparallel β -sheets ($\beta 3/\beta 4$ and $\beta 15/\beta 16$). It has been described that the RuvC domain of Cas9 shares structural similarity with the retroviral integrase superfamily members characterized by an RNase H fold, such as Escherichia coli RuvC (PDB code 1HJR, 14% identity, root-mean-square deviation (rmsd) of 3.6 Å for 126 equivalent C α atoms) and Thermus thermophilus RuvC (PDB code 4LD0, 12% identity, rmsd of 3.4 Å for 131 equivalent C α atoms). E. coli RuvC is a 3-layer alpha-beta sandwich containing a 5-stranded beta-sheet sandwiched between 5 alpha-helices. RuvC nucleases have four catalytic residues (e.g., Asp7, Glu70, His143 and Asp146 in T. thermophilus RuvC), and cleave Holliday junctions (or structurally analogous cruciform junctions) through a two-metal mechanism. Asp10 (Ala), Glu762, His983 and Asp986 of the Cas9 RuvC domain are located at positions similar to those of the catalytic residues of T. thermophilus RuvC.

[0132] In an example embodiment, split Ruv-C domain of the IscB proteins may have an HNH domain located between the Ruv-C II and Ruv-C III subdomains as described in more detail below. For example, the IscB protein domain architecture is comprised of the PLMP (P) domain, RuvC-I-II-III domains, a bridge domain (B), an HNH domain and a 3' terminal carboxyl (C) domain spanning 494 amino acids in the schematic shown in FIG. 9A. The bridge domain is located between the RuvC-I and RuvC-II domains and the HNH domain is located between the RuvC-II and RuvC-III domains (FIG. 9A). Size ranges of the individual subdomains described above range from

HNH domain

[0133] HNH domain comprise two antiparallel β strands connected with a variable length loop, an alpha helix, with a metal binding site between the two. The HNH conserved sites are conserved across the HNH superfamily, with HNH conservation throughout bacteria. In Cas9 proteins, for example, the HNH domain comprises a two-stranded antiparallel β -sheet ($\beta 12$ and $\beta 13$) flanked by four α -helices ($\alpha 35$ – $\alpha 38$). It shares structural similarity with the HNH

endonucleases characterized by a $\beta\beta\alpha$ -metal fold, such as phage T4 endonuclease VII (Endo VII) (PDB code 2QNC, 20% identity, rmsd of 2.7 Å for 61 equivalent Ca atoms) and *Vibrio vulnificus* nuclease (PDB code 1OUP, 8% identity, rmsd of 2.7 Å for 77 equivalent Ca atoms). HNH nucleases have three catalytic residues (e.g., Asp40, His41, and Asn62 in Endo VII), and cleave nucleic acid substrates through a single-metal mechanism. In the structure of the Endo VII N62D mutant in complex with a Holliday junction, a Mg²⁺ ion is coordinated by Asp40, Asp62, and the oxygen atoms of the scissile phosphate group of the substrate, while His41 acts as a general base to activate a water molecule for catalysis. Asp839, His840, and Asn863 of the Cas9 HNH domain correspond to Asp40, His41, and Asn62 of Endo VII, respectively, consistent with the observation that His840 is critical for the cleavage of the complementary DNA strand. The N863A mutant functions as a nickase, indicating that Asn863 participates in catalysis. The Cas9 HNH domain may cleave the complementary strand of the target DNA through a single-metal mechanism, as observed for other HNH superfamily nucleases. Although the Cas9 HNH domain shares a $\beta\beta\alpha$ -metal fold with other HNH endonucleases, their overall structures are distinct, consistent with the differences in their substrate specificities. Accordingly, IscB polypeptides of the present invention may comprise similar HNH domains in terms of sequence and/or function and may likewise comprise mutations analogous to those described above for Cas9 which convert the IscB polypeptide to a nickase. In an exemplary embodiment, a mutation to catalytic RuvC-II residue corresponding to E157A in corresponding to the sequence numbering of AwaIscB in an IscB polypeptide can be performed to abolish or significantly reduce the nucleolytic activity on the non-target DNA strand.

PLMP Domain

The IscB polypeptides comprise a conserved N-terminal domain, which is referred to herein as a PLMP domain or an X domain. In embodiments, the N-terminal X domain may have one or more conserved residues and/or motifs as identified in FIG. 3 and FIG. 10; see also FIG. 4-3 for PLMP motif alignment. In one embodiment, the PLMP domain comprises a conserved PLMP (SEQ ID NO:2372) amino acid motif. The PLMP motif can be located at or near the N terminus of the IscB polypeptide, including, for example at amino acids 12-15 of AwaIscB, or amino acids corresponding to *A. warmingii* IscB.

[0134] In some examples, the PLMP domain may be no more than 10, no more than 20, no more than 30, no more than 40, no more than 50, no more than 60, no more than 70, no more than 80, no more than 90, or no more than 100 amino acids in length. For example, the

PLMP domain may be no more than 70 amino acids in length, such as comprising 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, or 70 amino acids in length. An example PLMP domain can be as identified in, e.g. FIG. 58. PLMP domains may be found upstream of the RuvC-I domain and/or Bridge Helix, where present, of an IscB polypeptide. In one embodiment, the PLMP domain is located within 150, 140, 130, 120, 110, 100, 90, 80, 70, 60, 50, 40, 30, 20 or 10 amino acids upstream of the RuvC-I domain. See, e.g. FIG. 58.

[0135] In an aspect, truncation of the N-terminus domain of an IscB polypeptide, including more than 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15 amino acids, up to 70 amino acids of the N terminus, i.e. truncation of the PLMP domain, abolishes activity of the IscB polypeptide. In an aspect, more than 4 amino acids PLMP domain may reduce or abolish IscB activity. C-terminal domain

[0136] The C-terminal domain (also referred to herein as a Y domain) may comprise one or more conserved residues or motifs as shown in FIG. 3. See also, FIGs. 4, 58; The C-terminal domain may be no more than 10, no more than 20, no more than 30, no more than 40, no more than 50, no more than 60, no more than 70, no more than 80, no more than 90, or no more than 100 amino acids in length. For example, the Y domain may be no more than 70 amino acids in length, such as comprising 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, or 70 amino acids in length.

[0137] In an aspect, the IscB polypeptide, comprises a C-terminal domain that is structurally homologous to a tudor domain. See, e.g. Ren et al., Cell Res. (2014) 24:1146-1149. Tudor domains typically comprise a barrel-shaped beta strand fold and range in size around 50 and 60 amino acids. See, e.g. Kawale, A.A. & Burmann, B.M. Inherent backbone dynamics fine-tune the functional plasticity of Tudor domains. Structure (2021), incorporated herein by reference; see, in particular, Figure 1 showing exemplary tudor domain structure.

Protein modifications

[0138] The IscB polypeptide nucleases may comprise one or more modifications. As used herein, the term “modified” with regard to a IscB polypeptide nuclease generally refers to a

IscB polypeptide nuclease having one or more modifications or mutations (including point mutations, truncations, insertions, deletions, chimeras, fusion proteins, etc.) compared to the wild type counterpart from which it is derived. By derived is meant that the derived enzyme is largely based, in the sense of having a high degree of sequence homology with, a wildtype enzyme, but that it has been mutated (modified) in some way as known in the art or as described herein.

[0139] The modified proteins, e.g., modified IscB polypeptide nuclease may be catalytically inactive (also referred as dead). As used herein, a catalytically inactive or dead nuclease may have reduced or no nuclease activity compared to a wildtype counterpart nuclease. In some cases, a catalytically inactive or dead nuclease may have nickase activity. In some cases, a catalytically inactive or dead nuclease may not have nickase activity. Such a catalytically inactive or dead nuclease may not make either double-strand or single-strand break on a target polynucleotide, but may still bind or otherwise form complex with the target polynucleotide.

[0140] In an embodiment, the IscB comprises one or more mutation in the HNH domain of the polypeptide, or in the RuvC-II of the polypeptide. In an embodiment, the IscB polypeptide comprises a mutation of the catalytic RuvC-II residue corresponding to E157 to alanine (E157A) in *A. warmingii*. In an aspect, the mutation of a catalytic RuvC-II residue abolishes the nucleolytic activity on the non-target DNA strand. In an embodiment, the IscB polypeptide comprises a mutation of the catalytic HNH residue corresponding to H212 to alanine (H212A) in *A. warmingii*. In one embodiment, the mutation of the catalytic HNH residue abolishes nucleolytic activity on the target DNA strand. In an aspect, the IscB comprises a mutation corresponding to both E157A and H212A of *A. warmingii*, or corresponding to the positions according to consensus sequence numbering relative to *A. warmingii*. In an embodiment, mutation at both an HNH domain and RuvC abolishes all dsDNA nucleolytic activity, providing a dead IscB polypeptide (dIscB).

[0141] In one embodiment, the modifications of the IscB polypeptide may or may not cause an altered functionality. By means of example, modifications which do not result in an altered functionality include for instance codon optimization for expression into a particular host, or providing the nuclease with a particular marker (e.g., for visualization). Modifications which may result in altered functionality may also include mutations, including point mutations, insertions, deletions, truncations (including split nucleases), etc., as well as chimeric nucleases

(e.g., comprising domains from different orthologues or homologues) or fusion proteins. A chimeric enzyme can comprise a first fragment and a second fragment, and the fragments can be of IscB polypeptide nuclease orthologs of organisms of a genus or of a species, e.g., the fragments are from IscB polypeptide nuclease orthologs of different species. Fusion proteins may without limitation include, for instance, fusions with heterologous domains or functional domains (e.g., localization signals, catalytic domains, etc.). In an embodiment, various different modifications may be combined (e.g., a mutated nuclease which is catalytically inactive and which further is fused to a functional domain, such as for instance to induce DNA methylation or another nucleic acid modification, such as including without limitation, a break (e.g. by a different nuclease (domain)), a mutation, a deletion, an insertion, a replacement, a ligation, a digestion, a break or a recombination). As used herein, “altered functionality” includes without limitation an altered specificity (e.g., altered target recognition, increased (e.g., “enhanced” IscB polypeptide nuclease) or decreased specificity, or altered TAM recognition), altered activity (e.g., increased or decreased catalytic activity, including catalytically inactive nucleases or nickases), and/or altered stability (e.g., fusions with destabilization domains). Examples of all these modifications are known in the art. It will be understood that a “modified” nuclease as referred to herein, and in particular a “modified” IscB polypeptide nuclease or system or complex preferably still has the capacity to interact with or bind to the polynucleic acid (e.g., in complex with the ω RNA molecule). Such modified IscB polypeptide nuclease can be combined with the deaminase protein or active domain thereof as described herein.

[0142] In one embodiment, unmodified IscB polypeptide nucleases may have cleavage activity. In one embodiment, the IscB polypeptide nucleases may direct cleavage of one or both DNA strands at the location of or near a target sequence, such as within the target sequence and/or within the complement of the target sequence or at sequences associated with the target sequence. In one embodiment, the IscB polypeptide nucleases may direct cleavage of one or both strands within about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 50, 100, 200, 500, or more base pairs or nucleotides from the first or last nucleotide of a target sequence. In one embodiment, the cleavage may be staggered, i.e., generating sticky ends. In one embodiment, the cleavage is a staggered cut with a 5' overhang. In one embodiment, the cleavage is a staggered cut with a 5' overhang of 1 to 15 nucleotides, preferably of 4 or 9 nucleotides.

[0143] In one embodiment, the cleavage site is distant from the Target Adjacent Motif (TAM), which is used interchangeably with the term PAM herein, e.g., the cleavage occurs after the nth nucleotide on the non-target strand and after the nucleotide on the targeted strand. In one embodiment, the cleavage site occurs after an identified nucleotide (counted from the TAM) on the non-target strand and after the further identified nucleotide (counted from the TAM) on the targeted strand. In one embodiment, a vector encodes a nucleic acid-targeting effector protein that may be mutated with respect to a corresponding wild-type enzyme such that the mutated nucleic acid-targeting effector protein lacks the ability to cleave one or both DNA and RNA strands of a target polynucleotide containing a target sequence. As a further example, two or more catalytic domains of a IscB polypeptide nuclease (e.g., RuvC I, RuvC II, and RuvC III or the HNH domain) may be mutated to produce a mutated IscB polypeptide nuclease substantially lacking all DNA cleavage activity. As described herein, corresponding catalytic domains of a IscB polypeptide nuclease may also be mutated to produce a mutated IscB polypeptide nuclease lacking all DNA cleavage activity or having substantially reduced DNA cleavage activity. In one embodiment, a IscB polypeptide nuclease may be considered to substantially lack all polynucleotide cleavage activity when the polynucleotide cleavage activity of the mutated enzyme is no more than 25%, no more than 10%, no more than 5%, no more than 1%, no more than 0.1%, no more than 0.01% of the nucleic acid cleavage activity of the non-mutated form of the enzyme; an example can be when the nucleic acid cleavage activity of the mutated form is nil or negligible as compared with the non-mutated form. A IscB polypeptide nuclease may be identified with reference to the general class of enzymes that share homology to the biggest nuclease with multiple nuclease domains from the Type I, II, III, IV, V, or VI CRISPR systems.

[0144] TAM identification and specificity may be identified, for example, using the methods disclosed in the Examples section below.

[0145] In an embodiment, the nuclease domains of the IscB polypeptide nuclease are catalytically inactive, or modified to be catalytically inactive, or when the protein is a nickase. In an embodiment, both nuclease domains are catalytically inactive.

[0146] In an embodiment, the IscB polypeptide nuclease may comprise one or more modifications resulting in enhanced activity and/or specificity, such as including mutating residues that stabilize the targeted or non-targeted strand. In an embodiment, the altered or modified activity of the engineered IscB polypeptide nuclease comprises increased targeting

efficiency or decreased off-target binding. In an embodiment, the altered activity of the engineered IscB polypeptide nuclease comprises modified cleavage activity. In an embodiment, the altered activity comprises increased cleavage activity as to the target polynucleotide loci. In an embodiment, the altered activity comprises decreased cleavage activity as to the target polynucleotide loci. In an embodiment, the altered activity comprises decreased cleavage activity as to off-target polynucleotide loci. In an embodiment, the altered or modified activity of the modified nuclease comprises altered helicase kinetics. In an embodiment, the modified nuclease comprises a modification that alters association of the protein with the nucleic acid molecule comprising RNA, or a strand of the target polynucleotide loci, or a strand of off-target polynucleotide loci. In an aspect of the invention, the engineered IscB polypeptide nuclease comprises a modification that alters formation of the IscB polypeptide nuclease and related complex. In an embodiment, the altered activity comprises increased cleavage activity as to off-target polynucleotide loci. Accordingly, in an embodiment, there is increased specificity for target polynucleotide loci as compared to off-target polynucleotide loci. In other embodiments, there is reduced specificity for target polynucleotide loci as compared to off-target polynucleotide loci. In an embodiment, the mutations result in decreased off-target effects (e.g., cleavage or binding properties, activity, or kinetics), such as in case for IscB polypeptide nuclease for instance resulting in a lower tolerance for mismatches between target and mRNA. Other mutations may lead to increased off-target effects (e.g., cleavage or binding properties, activity, or kinetics). Other mutations may lead to increased or decreased on-target effects (e.g., cleavage or binding properties, activity, or kinetics). In an embodiment, the mutations result in altered (e.g., increased or decreased) helicase activity, association or formation of the functional nuclease complex. In an embodiment, the mutations result in an altered TAM recognition, i.e., a different TAM may be (in addition or in the alternative) be recognized, compared to the unmodified IscB polypeptide nuclease. Examples mutations include positively charged residues and/or (evolutionary) conserved residues, such as conserved positively charged residues, in order to enhance specificity. In an embodiment, such residues may be mutated to uncharged residues, such as alanine.

ωRNA Molecules

[0147] The systems herein may further comprise one or more ωRNA molecules, which are referred to herein interchangeably as ωRNA. The ωRNA complex can comprise a guide sequence and a scaffold that interacts with the IscB polypeptide. An ωRNA molecule may form a complex with IscB polypeptide nuclease or IscB polypeptide, and direct the complex to bind with a target sequence. In certain example embodiments, the ωRNA molecule is a single molecule comprising a scaffold sequence and a spacer sequence. In certain example embodiments, the spacer is 5' of the scaffold sequence. In certain example embodiments, the ωRNA molecule may further comprise a conserved nucleic acid sequence between the scaffold and spacer portions.

[0148] In certain example embodiments, the ωRNA scaffold comprises a spacer sequence and a conserved nucleotide sequence. The ωRNA scaffold typically comprises conserved regions, with the scaffold comprising 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 105, 115, 125, 135, 145, 155, 165, 175, 185, 195, 205, 215, 225, 235, 245, 255, 265, 275, 285, 295, 305, 315, 325, 335, 345, or 355 or more nt. In an aspect, the ωRNA scaffold comprises one conserved nucleotide sequence. In embodiments, the conserved nucleotide sequence is on or near a 5' end of the scaffold. In embodiments, the scaffold may comprise a short 3-4 base pair nexus, a conserved nexus hairpin and a large multi-stem loop region that may consist of two interconnected multi-stem loops. In an aspect, an IscrB associated scaffold may comprise a spacer, which can be re-programmed to direct site-specific binding to a target sequence of a target polynucleotide. The spacer may also be referred to herein as part of the ωRNA scaffold or as gRNA, and may comprise an engineered heterologous sequence. In an embodiment the scaffold may comprise a sequence from Table 1.

[0149] In an embodiment, the spacer length of the ωRNA is from 10 to 150 nt. In an embodiment, the spacer length of the guide RNA is at least 15 nucleotides. In an embodiment, the spacer length is from 15 to 17 nt, e.g., 15, 16, or 17 nt, from 17 to 20 nt, e.g., 17, 18, 19, or 20 nt, from 20 to 24 nt, e.g., 20, 21, 22, 23, or 24 nt, from 23 to 25 nt, e.g., 23, 24, or 25 nt, from 24 to 27 nt, e.g., 24, 25, 26, or 27 nt, from 27 to 30 nt, e.g., 27, 28, 29, or 30 nt, from 30

to 35 nt, e.g., 30, 31, 32, 33, 34, or 35 nt, or 35 nt or longer. In certain example embodiment, the guide sequence is 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149 or 150 nt.

[0150] In an embodiment, the ω RNA spacer length is from 15 to 50 nt. In an embodiment, the spacer length of the ω RNA is at least 15 nucleotides. In an embodiment, the spacer length is from 15 to 50 nt, e.g., 15, 16, or 17 nt, from 17 to 20 nt, e.g., 17, 18, 19, or 20 nt, from 20 to 24 nt, e.g., 20, 21, 22, 23, or 24 nt, from 23 to 25 nt, e.g., 23, 24, or 25 nt, from 24 to 27 nt, e.g., 24, 25, 26, or 27 nt, from 27 to 30 nt, e.g., 27, 28, 29, or 30 nt, from 30 to 35 nt, e.g., 30, 31, 32, 33, 34, or 35 nt, or 35 nt, from 34 to 40 nt, e.g., 34, 35, 36, 37, 38, 39, 40, from 35 to 39, from 36 to 38 nt long, about 37 nt, or longer.

[0151] In one embodiment, the sequence of the ω RNA molecule is selected to reduce the degree secondary structure within the ω RNA molecule. In one embodiment, about or less than about 75%, 50%, 40%, 30%, 25%, 20%, 15%, 10%, 5%, 1%, or fewer of the nucleotides of the nucleic acid-targeting ω RNA participate in self-complementary base pairing when optimally folded. Optimal folding may be determined by any suitable polynucleotide folding algorithm. Some programs are based on calculating the minimal Gibbs free energy. An example of one such algorithm is mFold, as described by Zuker and Stiegler (Nucleic Acids Res. 9 (1981), 133-148). Another example of a folding algorithm is the online webserver RNAfold, developed at Institute for Theoretical Chemistry at the University of Vienna, using the centroid structure prediction algorithm (see e.g., A.R. Gruber et al., 2008, Cell 106(1): 23-24; and PA Carr and GM Church, 2009, Nature Biotechnology 27(12): 1151-62).

[0152] As used herein, a heterologous ω RNA molecule is an ω RNA molecule that is not derived from the same species as the IscB polypeptide nuclease, or comprises a portion of the molecule, e.g., spacer, that is not derived from the same species as the IscB polypeptide nuclease, e.g. IscB protein. For example, a heterologous ω RNA molecule of a IscB polypeptide

nuclease derived from species A comprises a polynucleotide derived from a species different from species A, or an artificial polynucleotide.

[0153] In a particular embodiment, the ω RNA comprises a guide sequence linked to a conserved nucleotide sequence, wherein the conserved nucleotide sequence may comprise one or more stem loops or optimized secondary structures. In an embodiment, the conserved nucleotide sequence has a minimum length of 16 nts and a single stem loop. In further embodiments the conserved nucleotide sequence has a length longer than 16 nts, preferably more than 17 nts, and has more than one stem loops or optimized secondary structures. In one embodiment, the guide sequence may be linked to all or part of the natural conserved nucleotide sequence. In one embodiment, certain aspects of the guide architecture can be modified, for example by addition, subtraction, or substitution of features, whereas certain other aspects of guide architecture are maintained. Preferred locations for engineered guide modifications, including but not limited to insertions, deletions, and substitutions include guide termini and regions of the guide that are exposed when complexed with IscB polypeptide nuclease and/or target, for example the tetraloop and/or loop2.

[0154] In one embodiment, a loop in the guide RNA is provided. This may be a stem loop or a tetra loop. The loop is preferably GAAA, but it is not limited to this sequence or indeed to being only 4bp in length. Indeed, preferred loop forming sequences for use in hairpin structures are four nucleotides in length, and most preferably have the sequence GAAA. However, longer or shorter loop sequences may be used, as may alternative sequences. The sequences preferably include a nucleotide triplet (for example, AAA), and an additional nucleotide (for example C or G). Examples of loop forming sequences include CAAA and AAAG.

[0155] In one embodiment, the ω RNA forms a stemloop with a separate non-covalently linked sequence, which can be DNA or RNA. In an embodiment, the sequences forming the guide are first synthesized using the standard phosphoramidite synthetic protocol (Herdewijn, P., ed., *Methods in Molecular Biology* Vol 288, *Oligonucleotide Synthesis: Methods and Applications*, Humana Press, New Jersey (2012)). In one embodiment, these sequences can be functionalized to contain an appropriate functional group for ligation using the standard protocol known in the art (Hermanson, G. T., *Bioconjugate Techniques*, Academic Press (2013)). Examples of functional groups include, but are not limited to, hydroxyl, amine, carboxylic acid, carboxylic acid halide, carboxylic acid active ester, aldehyde, carbonyl,

chlorocarbonyl, imidazolylcarbonyl, hydrozide, semicarbazide, thio semicarbazide, thiol, maleimide, haloalkyl, sulfonyl, ally, propargyl, diene, alkyne, and azide. Once this sequence is functionalized, a covalent chemical bond or linkage can be formed between this sequence and the conserved nucleotide sequence. Examples of chemical bonds include, but are not limited to, those based on carbamates, ethers, esters, amides, imines, amidines, aminotrizines, hydrozone, disulfides, thioethers, thioesters, phosphorothioates, phosphorodithioates, sulfonamides, sulfonates, fulfones, sulfoxides, ureas, thioureas, hydrazide, oxime, triazole, photolabile linkages, C-C bond forming groups such as Diels-Alder cyclo-addition pairs or ring-closing metathesis pairs, and Michael reaction pairs.

[0156] In one embodiment, these stem-loop forming sequences can be chemically synthesized. In one embodiment, the chemical synthesis uses automated, solid-phase oligonucleotide synthesis machines with 2'-acetoxyethyl orthoester (2'-ACE) (Scaringe et al., J. Am. Chem. Soc. (1998) 120: 11820-11821; Scaringe, Methods Enzymol. (2000) 317: 3-18) or 2'-thionocarbamate (2'-TC) chemistry (Dellinger et al., J. Am. Chem. Soc. (2011) 133: 11540-11546; Hendel et al., Nat. Biotechnol. (2015) 33:985-989).

[0157] The repeat:anti repeat duplex will be apparent from the secondary structure of the ω RNA. It may be typically a first complimentary stretch after (in 5' to 3' direction) the poly U tract and before the tetraloop; and a second complimentary stretch after (in 5' to 3' direction) the tetraloop and before the poly A tract. The first complimentary stretch (the "repeat") is complimentary to the second complimentary stretch (the "anti-repeat"). As such, they Watson-Crick base pair to form a duplex of dsRNA when folded back on one another. As such, the anti-repeat sequence is the complimentary sequence of the repeat and in terms to A-U or C-G base pairing, but also in terms of the fact that the anti-repeat is in the reverse orientation due to the tetraloop.

[0158] In an embodiment of the invention, modification of guide architecture comprises replacing bases in stemloop 2. For example, In one embodiment, "actt" ("acuu" in RNA) and "aagt" ("aagu" in RNA) bases in stemloop2 are replaced with "cgcc" and "gcgg". In one embodiment, "actt" and "aagt" bases in stemloop2 are replaced with complimentary GC-rich regions of 4 nucleotides. In one embodiment, the complimentary GC-rich regions of 4 nucleotides are "cgcc" and "gcgg" (both in 5' to 3' direction). In one embodiment, the complimentary GC-rich regions of 4 nucleotides are "gcgg" and "cgcc" (both in 5' to 3'

direction). Other combination of C and G in the complimentary GC-rich regions of 4 nucleotides will be apparent including CCCC and GGGG.

[0159] In one aspect, the stemloop 2, e.g., “ACTTgtttAAGT” (SEQ ID NO: 1) can be replaced by any “XXXXgtttYYYY” (SEQ ID NO: 2), e.g., where XXXX and YYYY represent any complementary sets of nucleotides that together will base pair to each other to create a stem.

[0160] As used herein, the term “spacer” may also be referred to as a “guide sequence.” In one embodiment, the degree of complementarity of the guide quence to a given target sequence, when optimally aligned using a suitable alignment algorithm, is about or more than 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99%, or more. In certain example embodiments, the ωRNA molecule comprises a guide sequence that may be designed to have at least one mismatch with the target sequence, such that a RNA duplex formed between the sequence and the target sequence. Accordingly, the degree of complementarity is less than 99%. For instance, where the guide sequence consists of 24 nucleotides, the degree of complementarity is more particularly about 96% or less. In one embodiment, the guide sequence is designed to have a stretch of two or more adjacent mismatching nucleotides, such that the degree of complementarity over the entire sequence is further reduced. For instance, where the guide sequence consists of 24 nucleotides, the degree of complementarity is more particularly about 96% or less, more particularly, about 92% or less, more particularly about 88% or less, more particularly about 84% or less, more particularly about 80% or less, more particularly about 76% or less, more particularly about 72% or less, depending on whether the stretch of two or more mismatching nucleotides encompasses 2, 3, 4, 5, 6 or 7 nucleotides, etc. In one embodiment, aside from the stretch of one or more mismatching nucleotides, the degree of complementarity, when optimally aligned using a suitable alignment algorithm, is about or more than about 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99%, or more. Optimal alignment may be determined with the use of any suitable algorithm for aligning sequences, non-limiting example of which include the Smith-Waterman algorithm, the Needleman-Wunsch algorithm, algorithms based on the Burrows-Wheeler Transform (e.g., the Burrows Wheeler Aligner), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies; available at novocraft.com), ELAND (Illumina, San Diego, CA), SOAP (available at soap.genomics.org.cn), and Maq (available at maq.sourceforge.net). The ability of a sequence

(within a nucleic acid-targeting guide sequence) to direct sequence-specific binding of a nucleic acid -targeting complex to a target nucleic acid sequence may be assessed by any suitable assay. For example, the components of a ω RNA system sufficient to form a nucleic acid-targeting complex, including the guide sequence to be tested, may be provided to a host cell having the corresponding target nucleic acid sequence, such as by transfection with vectors encoding the components of the nucleic acid-targeting complex, followed by an assessment of preferential targeting (e.g., cleavage) within the target nucleic acid sequence, such as by Surveyor assay as described herein. Similarly, cleavage of a target nucleic acid sequence (or a sequence in the vicinity thereof) may be evaluated in a test tube by providing the target nucleic acid sequence, components of a nucleic acid-targeting complex, including the sequence to be tested and a control sequence different from the test guide sequence, and comparing binding or rate of cleavage at or in the vicinity of the target sequence between the test and control guide sequence reactions. Other assays are possible, and will occur to those skilled in the art. A guide sequence, and hence a nucleic acid-targeting ω RNA may be selected to target any target nucleic acid sequence.

[0161] A ω RNA sequence, and hence a nucleic acid-targeting guide, may be selected to target any target nucleic acid sequence. The target sequence may be DNA. The target sequence may be any RNA sequence. In one embodiment, the target sequence may be a sequence within a RNA molecule selected from the group consisting of messenger RNA (mRNA), pre-mRNA, ribosomal RNA (rRNA), transfer RNA (tRNA), micro-RNA (miRNA), small interfering RNA (siRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), double stranded RNA (dsRNA), non-coding RNA (ncRNA), long non-coding RNA (lncRNA), and small cytoplasmatic RNA (scRNA). In some preferred embodiments, the target sequence may be a sequence within a RNA molecule selected from the group consisting of mRNA, pre-mRNA, and rRNA. In some preferred embodiments, the target sequence may be a sequence within a RNA molecule selected from the group consisting of ncRNA, and lncRNA. In some more preferred embodiments, the target sequence may be a sequence within an mRNA molecule or a pre-mRNA molecule.

[0162] In one embodiment, the ω RNA molecule forms a stemloop with a separate non-covalently linked sequence, which can be DNA or RNA. In one embodiment, the sequences forming the ω RNA are first synthesized using the standard phosphoramidite synthetic protocol

(Herdewijn, P., ed., *Methods in Molecular Biology* Col 288, *Oligonucleotide Synthesis: Methods and Applications*, Humana Press, New Jersey (2012)). In one embodiment, these sequences can be functionalized to contain an appropriate functional group for ligation using the standard protocol known in the art (Hermanson, G. T., *Bioconjugate Techniques*, Academic Press (2013)). Examples of functional groups include, but are not limited to, hydroxyl, amine, carboxylic acid, carboxylic acid halide, carboxylic acid active ester, aldehyde, carbonyl, chlorocarbonyl, imidazolylcarbonyl, hydrozide, semicarbazide, thio semicarbazide, thiol, maleimide, haloalkyl, sulfonyl, ally, propargyl, diene, alkyne, and azide. Once this sequence is functionalized, a covalent chemical bond or linkage can be formed between this sequence and the conserved nucleotide sequence. Examples of chemical bonds include, but are not limited to, those based on carbamates, ethers, esters, amides, imines, amidines, aminotrizines, hydrozone, disulfides, thioethers, thioesters, phosphorothioates, phosphorodithioates, sulfonamides, sulfonates, fulfones, sulfoxides, ureas, thioureas, hydrazide, oxime, triazole, photolabile linkages, C-C bond forming groups such as Diels-Alder cyclo-addition pairs or ring-closing metathesis pairs, and Michael reaction pairs.

[0163] In one embodiment, these stem-loop forming sequences can be chemically synthesized. In one embodiment, the chemical synthesis uses automated, solid-phase oligonucleotide synthesis machines with 2'-acetoxyethyl orthoester (2'-ACE) (Scaringe et al., *J. Am. Chem. Soc.* (1998) 120: 11820-11821; Scaringe, *Methods Enzymol.* (2000) 317: 3-18) or 2'-thionocarbamate (2'-TC) chemistry (Dellinger et al., *J. Am. Chem. Soc.* (2011) 133: 11540-11546; Hendel et al., *Nat. Biotechnol.* (2015) 33:985-989).

[0164] In an embodiment, the ω RNA molecule comprises non-naturally occurring nucleic acids and/or non-naturally occurring nucleotides and/or nucleotide analogs, and/or chemically modifications. Preferably, these non-naturally occurring nucleic acids and non-naturally occurring nucleotides are located outside the ω RNA sequence. Non-naturally occurring nucleic acids can include, for example, mixtures of naturally and non-naturally occurring nucleotides. Non-naturally occurring nucleotides and/or nucleotide analogs may be modified at the ribose, phosphate, and/or base moiety. In an embodiment of the invention, a ω RNA nucleic acid comprises ribonucleotides and non-ribonucleotides. In one such embodiment, a ω RNA comprises one or more ribonucleotides and one or more deoxyribonucleotides. In an embodiment of the invention, the ω RNA comprises one or more non-naturally occurring

nucleotide or nucleotide analog such as a nucleotide with phosphorothioate linkage, a locked nucleic acid (LNA) nucleotides comprising a methylene bridge between the 2' and 4' carbons of the ribose ring, or bridged nucleic acids (BNA). Other examples of modified nucleotides include 2'-O-methyl analogs, 2'-deoxy analogs, or 2'-fluoro analogs. Further examples of modified bases include, but are not limited to, 2-aminopurine, 5-bromo-uridine, pseudouridine, inosine, 7-methylguanosine. Examples of ω RNA chemical modifications include, without limitation, incorporation of 2'-O-methyl (M), 2'-O-methyl 3'phosphorothioate (MS), S-constrained ethyl(cEt), or 2'-O-methyl 3'thioPACE (MSP) at one or more terminal nucleotides. Such chemically modified ω RNA can comprise increased stability and increased activity as compared to unmodified ω RNA, though on-target vs. off-target specificity is not predictable. (See, Hendel, 2015, Nat Biotechnol. 33(9):985-9, doi: 10.1038/nbt.3290, published online 29 June 2015 Ragdarm et al., 0215, PNAS, E7110-E7111; Allerson et al., J. Med. Chem. 2005, 48:901-904; Bramsen et al., Front. Genet., 2012, 3:154; Deng et al., PNAS, 2015, 112:11870-11875; Sharma et al., MedChemComm., 2014, 5:1454-1471; Hendel et al., Nat. Biotechnol. (2015) 33(9): 985-989; Li et al., Nature Biomedical Engineering, 2017, 1, 0066 DOI:10.1038/s41551-017-0066). In one embodiment, the 5' and/or 3' end of a ω RNA is modified by a variety of functional moieties including fluorescent dyes, polyethylene glycol, cholesterol, proteins, or detection tags. (See Kelly et al., 2016, J. Biotech. 233:74-83). In an embodiment, a ω RNA comprises ribonucleotides in a region that binds to a target sequence and one or more deoxyribonucleotides and/or nucleotide analogs in a region that binds to the IscB polypeptide nuclease. In an embodiment, deoxyribonucleotides and/or nucleotide analogs are incorporated in engineered hRNA structures. In one embodiment, 3-5 nucleotides at either the 3' or the 5' end of a hRNA is chemically modified. In one embodiment, only minor modifications are introduced in the seed region, such as 2'-F modifications. In one embodiment, 2'-F modification is introduced at the 3' end of a hRNA. In an embodiment, three to five nucleotides at the 5' and/or the 3' end of the hRNA are chemically modified with 2'-O-methyl (M), 2'-O-methyl 3' phosphorothioate (MS), S-constrained ethyl(cEt), or 2'-O-methyl 3' thioPACE (MSP). Such modification can enhance genome editing efficiency (see Hendel et al., Nat. Biotechnol. (2015) 33(9): 985-989). In an embodiment, all of the phosphodiester bonds of a hRNA are substituted with phosphorothioates (PS) for enhancing levels of gene disruption. In an embodiment, more than five nucleotides at the 5' and/or the 3' end of the hRNA are

chemically modified with 2'-O-Me, 2'-F or S-constrained ethyl(cEt). Such chemically modified hRNA can mediate enhanced levels of gene disruption (see Ragdarm et al., 0215, PNAS, E7110-E7111). In an embodiment of the invention, a hRNA is modified to comprise a chemical moiety at its 3' and/or 5' end. Such moieties include, but are not limited to amine, azide, alkyne, thio, dibenzocyclooctyne (DBCO), or Rhodamine. In certain embodiment, the chemical moiety is conjugated to the hRNA by a linker, such as an alkyl chain. In an embodiment, the chemical moiety of the modified hRNA can be used to attach the hRNA to another molecule, such as DNA, RNA, protein, or nanoparticles. Such chemically modified hRNA can be used to identify or enrich cells generically edited by a IscB polypeptide nuclease and related systems (see Lee et al., eLife, 2017, 6:e25312, DOI:10.7554).

[0165] In a particular embodiment, the conserved nucleotide sequence may be modified to comprise one or more protein-binding RNA aptamers. In a particular embodiment, one or more aptamers may be included such as part of optimized secondary structure. Such aptamers may be capable of binding a bacteriophage coat protein as detailed further herein.

[0166] In embodiments, the IscB polypeptide utilizes the hRNA scaffold comprising a polynucleotide sequence that facilitates the interaction with the IscB protein, allowing for sequence specific binding and/or targeting of the guide sequence with the target polynucleotide. Chemical synthesis of the hRNA scaffold is contemplated, using covalent linkage using various bioconjugation reactions, loops, bridges, and non-nucleotide links via modifications of sugar, internucleotide phosphodiester bonds, purine and pyrimidine residues. Sletten et al., *Angew. Chem. Int. Ed.* (2009) 48:6974-6998; Manoharan, M. *Curr. Opin. Chem. Biol.* (2004) 8: 570-9; Behlke et al., *Oligonucleotides* (2008) 18: 305-19; Watts, et al., *Drug. Discov. Today* (2008) 13: 842-55; Shukla, et al., *ChemMedChem* (2010) 5: 328-49; chemical synthesis using automated, solid-phase oligonucleotide synthesis machines with 2'-acetoxyethyl orthoester (2'-ACE) (Scaringe et al., *J. Am. Chem. Soc.* (1998) 120: 11820-11821; Scaringe, *Methods Enzymol.* (2000) 317: 3-18) or 2'-thionocarbamate (2'-TC) chemistry (Dellinger et al., *J. Am. Chem. Soc.* (2011) 133: 11540-11546; Hendel et al., *Nat. Biotechnol.* (2015) 33:985-989).

[0167] In certain example embodiments, the scaffold and spacer may be designed as two separate molecules that can hybridize or covalently joined into a single molecule. Covalent linkage can be via a linker (e.g., a non-nucleotide loop) that comprises a moiety such as spacers, attachments, bioconjugates, chromophores, reporter groups, dye labeled RNAs, and non-

naturally occurring nucleotide analogues. More specifically, suitable spacers for purposes of this invention include, but are not limited to, polyethers (e.g., polyethylene glycols, polyalcohols, polypropylene glycol or mixtures of ethylene and propylene glycols), polyamines group (e.g., spennine, spermidine and polymeric derivatives thereof), polyesters (e.g., poly(ethyl acrylate)), polyphosphodiesteres, alkylenes, and combinations thereof. Suitable attachments include any moiety that can be added to the linker to add additional properties to the linker, such as but not limited to, fluorescent labels. Suitable bioconjugates include, but are not limited to, peptides, glycosides, lipids, cholesterol, phospholipids, diacyl glycerols and dialkyl glycerols, fatty acids, hydrocarbons, enzyme substrates, steroids, biotin, digoxigenin, carbohydrates, polysaccharides. Suitable chromophores, reporter groups, and dye-labeled RNAs include, but are not limited to, fluorescent dyes such as fluorescein and rhodamine, chemiluminescent, electrochemiluminescent, and bioluminescent marker compounds. The design of example linkers conjugating two RNA components are also described in WO 2004/015075.

[0168] The linker (e.g., a non-nucleotide loop) can be of any length. In one embodiment, the linker has a length equivalent to about 0-16 nucleotides. In one embodiment, the linker has a length equivalent to about 0-8 nucleotides. In one embodiment, the linker has a length equivalent to about 0-4 nucleotides. In one embodiment, the linker has a length equivalent to about 2 nucleotides. Example linker design is also described in International Patent Publication No. WO 2011/008730.

Escorted ω RNA molecules

[0169] In one embodiment, the compositions or complexes have a hRNA molecule with a functional structure designed to improve hRNA molecule structure, architecture, stability, genetic expression, or any combination thereof. Such a structure can include an aptamer.

[0170] Aptamers are biomolecules that can be designed or selected to bind tightly to other ligands, for example using a technique called systematic evolution of ligands by exponential enrichment (SELEX; Tuerk C, Gold L: "Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase." *Science* 1990, 249:505-510). Nucleic acid aptamers can for example be selected from pools of random-sequence oligonucleotides, with high binding affinities and specificities for a wide range of biomedically relevant targets, suggesting a wide range of therapeutic utilities for aptamers (Keefe, Anthony

D., Supriya Pai, and Andrew Ellington. "Aptamers as therapeutics." *Nature Reviews Drug Discovery* 9.7 (2010): 537-550). These characteristics also suggest a wide range of uses for aptamers as drug delivery vehicles (Levy-Nissenbaum, Etgar, et al. "Nanotechnology and aptamers: applications in drug delivery." *Trends in biotechnology* 26.8 (2008): 442-449; and Hicke BJ, Stephens AW. "Escort aptamers: a delivery service for diagnosis and therapy." *J Clin Invest* 2000, 106:923-928.). Aptamers may also be constructed that function as molecular switches, responding to a cue by changing properties, such as RNA aptamers that bind fluorophores to mimic the activity of green fluorescent protein (Paige, Jeremy S., Karen Y. Wu, and Samie R. Jaffrey. "RNA mimics of green fluorescent protein." *Science* 333.6042 (2011): 642-646). It has also been suggested that aptamers may be used as components of targeted siRNA therapeutic delivery systems, for example targeting cell surface proteins (Zhou, Jiehua, and John J. Rossi. "Aptamer-targeted cell-specific RNA interference." *Silence* 1.1 (2010): 4).

[0171] Accordingly, in one embodiment, the hRNA molecule is modified, e.g., by one or more aptamer(s) designed to improve hRNA molecule delivery, including delivery across the cellular membrane, to intracellular compartments, or into the nucleus. Such a structure can include, either in addition to the one or more aptamer(s) or without such one or more aptamer(s), moiety(ies) so as to render the hRNA molecule deliverable, inducible or responsive to a selected effector. The invention accordingly comprehends a hRNA molecule that responds to normal or pathological physiological conditions, including without limitation pH, hypoxia, O₂ concentration, temperature, protein concentration, enzymatic concentration, lipid structure, light exposure, mechanical disruption (e.g., ultrasound waves), magnetic fields, electric fields, or electromagnetic radiation.

[0172] Light responsiveness of an inducible system may be achieved via the activation and binding of cryptochrome-2 and CIB1. Blue light stimulation induces an activating conformational change in cryptochrome-2, resulting in recruitment of its binding partner CIB1. This binding is fast and reversible, achieving saturation in <15 sec following pulsed stimulation and returning to baseline <15 min after the end of stimulation. These rapid binding kinetics result in a system temporally bound only by the speed of transcription/translation and transcript/protein degradation, rather than uptake and clearance of inducing agents. Cryptochrome-2 activation is also highly sensitive, allowing for the use of low light intensity stimulation and mitigating the risks of phototoxicity. Further, in a context such as the intact

mammalian brain, variable light intensity may be used to control the size of a stimulated region, allowing for greater precision than vector delivery alone may offer.

[0173] Energy sources such as electromagnetic radiation, sound energy or thermal energy may induce the guide. Advantageously, the electromagnetic radiation is a component of visible light. In a preferred embodiment, the light is a blue light with a wavelength of about 450 to about 495 nm. In an especially preferred embodiment, the wavelength is about 488 nm. In another preferred embodiment, the light stimulation is via pulses. The light power may range from about 0-9 mW/cm². In a preferred embodiment, a stimulation paradigm of as low as 0.25 sec every 15 sec should result in maximal activation.

[0174] The chemical or energy sensitive hRNA may undergo a conformational change upon induction by the binding of a chemical source or by the energy allowing it act as a hRNA and have the IscB polypeptide nuclease system or complex function. The invention can involve applying the chemical source or energy so as to have the hRNA function and the IscB polypeptide nuclease system or complex function; and optionally further determining that the expression of the genomic locus is altered.

[0175] There are several different designs of this chemical inducible system: 1. ABI-PYL based system inducible by Absciscic Acid (ABA) (see, e.g., stke.sciencemag.org/cgi/content/abstract/sigtrans;4/164/rs2), 2. FKBP-FRB based system inducible by rapamycin (or related chemicals based on rapamycin) (see, e.g., www.nature.com/nmeth/journal/v2/n6/full/nmeth763.html), 3. GID1-GAI based system inducible by Gibberellin (GA) (see, e.g., www.nature.com/nchembio/journal/v8/n5/full/nchembio.922.html).

[0176] A chemical inducible system can be an estrogen receptor (ER) based system inducible by 4-hydroxytamoxifen (4OHT) (see, e.g., www.pnas.org/content/104/3/1027.abstract). A mutated ligand-binding domain of the estrogen receptor called ERT2 translocates into the nucleus of cells upon binding of 4-hydroxytamoxifen. In further embodiments of the invention any naturally occurring or engineered derivative of any nuclear receptor, thyroid hormone receptor, retinoic acid receptor, estrogen receptor, estrogen-related receptor, glucocorticoid receptor, progesterone receptor, androgen receptor may be used in inducible systems analogous to the ER based inducible system.

[0177] Another inducible system is based on the design using Transient receptor potential (TRP) ion channel-based system inducible by energy, heat or radio-wave (see, e.g., www.sciencemag.org/content/336/6081/604). These TRP family proteins respond to different stimuli, including light and heat. When this protein is activated by light or heat, the ion channel will open and allow the entering of ions such as calcium into the plasma membrane. This influx of ions will bind to intracellular ion interacting partners linked to a polypeptide including the hRNA and the other components of the IscB polypeptide nuclease/hRNA molecule complex or system, and the binding will induce the change of sub-cellular localization of the polypeptide, leading to the entire polypeptide entering the nucleus of cells. Once inside the nucleus, the hRNA protein and the other components of the IscB polypeptide nuclease/hRNA molecule complex will be active and modulating target gene expression in cells.

[0178] While light activation may be an advantageous embodiment, sometimes it may be disadvantageous especially for *in vivo* applications in which the light may not penetrate the skin or other organs. In this instance, other methods of energy activation are contemplated, in particular, electric field energy and/or ultrasound which have a similar effect.

[0179] Electric field energy is preferably administered substantially as described in the art, using one or more electric pulses of from about 1 Volt/cm to about 10 kVolts/cm under *in vivo* conditions. Instead of or in addition to the pulses, the electric field may be delivered in a continuous manner. The electric pulse may be applied for between 1 μ s and 500 milliseconds, preferably between 1 μ s and 100 milliseconds. The electric field may be applied continuously or in a pulsed manner for 5 about minutes.

[0180] As used herein, ‘electric field energy’ is the electrical energy to which a cell is exposed. Preferably the electric field has a strength of from about 1 Volt/cm to about 10 kVolts/cm or more under *in vivo* conditions (see WO97/49450).

[0181] As used herein, the term “electric field” includes one or more pulses at variable capacitance and voltage and including exponential and/or square wave and/or modulated wave and/or modulated square wave forms. References to electric fields and electricity should be taken to include reference the presence of an electric potential difference in the environment of a cell. Such an environment may be set up by way of static electricity, alternating current (AC), direct current (DC), etc., as known in the art. The electric field may be uniform, non-uniform or otherwise, and may vary in strength and/or direction in a time dependent manner.

[0182] Single or multiple applications of electric field, as well as single or multiple applications of ultrasound are also possible, in any order and in any combination. The ultrasound and/or the electric field may be delivered as single or multiple continuous applications, or as pulses (pulsatile delivery).

[0183] Electroporation has been used in both *in vitro* and *in vivo* procedures to introduce foreign material into living cells. With *in vitro* applications, a sample of live cells is first mixed with the agent of interest and placed between electrodes such as parallel plates. Then, the electrodes apply an electrical field to the cell/implant mixture. Examples of systems that perform *in vitro* electroporation include the Electro Cell Manipulator ECM600 product, and the Electro Square Porator T820, both made by the BTX Division of Genetronics, Inc (see U.S. Pat. No 5,869,326).

[0184] The known electroporation techniques (both *in vitro* and *in vivo*) function by applying a brief high voltage pulse to electrodes positioned around the treatment region. The electric field generated between the electrodes causes the cell membranes to temporarily become porous, whereupon molecules of the agent of interest enter the cells. In known electroporation applications, this electric field comprises a single square wave pulse on the order of 1000 V/cm, of about 100 .mu.s duration. Such a pulse may be generated, for example, in known applications of the Electro Square Porator T820.

[0185] Preferably, the electric field has a strength of from about 1 V/cm to about 10 kV/cm under *in vitro* conditions. Thus, the electric field may have a strength of 1 V/cm, 2 V/cm, 3 V/cm, 4 V/cm, 5 V/cm, 6 V/cm, 7 V/cm, 8 V/cm, 9 V/cm, 10 V/cm, 20 V/cm, 50 V/cm, 100 V/cm, 200 V/cm, 300 V/cm, 400 V/cm, 500 V/cm, 600 V/cm, 700 V/cm, 800 V/cm, 900 V/cm, 1 kV/cm, 2 kV/cm, 5 kV/cm, 10 kV/cm, 20 kV/cm, 50 kV/cm or more. More preferably from about 0.5 kV/cm to about 4.0 kV/cm under *in vitro* conditions. Preferably the electric field has a strength of from about 1 V/cm to about 10 kV/cm under *in vivo* conditions. However, the electric field strengths may be lowered where the number of pulses delivered to the target site are increased. Thus, pulsatile delivery of electric fields at lower field strengths is envisaged.

[0186] Preferably, the application of the electric field is in the form of multiple pulses such as double pulses of the same strength and capacitance or sequential pulses of varying strength and/or capacitance. As used herein, the term “pulse” includes one or more electric pulses at variable capacitance and voltage and including exponential and/or square wave and/or modulated wave/square wave forms.

[0187] Preferably, the electric pulse is delivered as a waveform selected from an exponential wave form, a square wave form, a modulated wave form and a modulated square wave form.

[0188] A preferred embodiment employs direct current at low voltage. Thus, Applicants disclose the use of an electric field which is applied to the cell, tissue, or tissue mass at a field strength of between 1V/cm and 20V/cm, for a period of 100 milliseconds or more, preferably 15 minutes or more.

[0189] Ultrasound is advantageously administered at a power level of from about 0.05 W/cm² to about 100 W/cm². Diagnostic or therapeutic ultrasound may be used, or combinations thereof.

[0190] As used herein, the term "ultrasound" refers to a form of energy which consists of mechanical vibrations the frequencies of which are so high they are above the range of human hearing. Lower frequency limit of the ultrasonic spectrum may generally be taken as about 20 kHz. Most diagnostic applications of ultrasound employ frequencies in the range 1 and 15 MHz' (From Ultrasonics in Clinical Diagnosis, P. N. T. Wells, ed., 2nd. Edition, Publ. Churchill Livingstone [Edinburgh, London & NY, 1977]).

[0191] Ultrasound has been used in both diagnostic and therapeutic applications. When used as a diagnostic tool ("diagnostic ultrasound"), ultrasound is typically used in an energy density range of up to about 100 mW/cm² (FDA recommendation), although energy densities of up to 750 mW/cm² have been used. In physiotherapy, ultrasound is typically used as an energy source in a range up to about 3 to 4 W/cm² (WHO recommendation). In other therapeutic applications, higher intensities of ultrasound may be employed, for example, HIFU at 100 W/cm up to 1 kW/cm² (or even higher) for short periods of time. The term "ultrasound" as used in this specification is intended to encompass diagnostic, therapeutic, and focused ultrasound.

[0192] Focused ultrasound (FUS) allows thermal energy to be delivered without an invasive probe (see Morocz et al 1998 Journal of Magnetic Resonance Imaging Vol.8, No. 1, pp.136-142. Another form of focused ultrasound is high intensity focused ultrasound (HIFU) which is reviewed by Moussatov et al in Ultrasonics (1998) Vol.36, No.8, pp.893-900 and TranHuuHue et al in Acustica (1997) Vol.83, No.6, pp.1103-1106.

[0193] Preferably, a combination of diagnostic ultrasound and a therapeutic ultrasound is employed. This combination is not intended to be limiting, however, and the skilled reader will

appreciate that any variety of combinations of ultrasound may be used. Additionally, the energy density, frequency of ultrasound, and period of exposure may be varied.

[0194] Preferably, the exposure to an ultrasound energy source is at a power density of from about 0.05 to about 100 Wcm⁻². Even more preferably, the exposure to an ultrasound energy source is at a power density of from about 1 to about 15 Wcm⁻².

[0195] Preferably, the exposure to an ultrasound energy source is at a frequency of from about 0.015 to about 10.0 MHz. More preferably the exposure to an ultrasound energy source is at a frequency of from about 0.02 to about 5.0 MHz or about 6.0 MHz. Most preferably, the ultrasound is applied at a frequency of 3 MHz.

[0196] Preferably the exposure is for periods of from about 10 milliseconds to about 60 minutes. Preferably the exposure is for periods of from about 1 second to about 5 minutes. More preferably, the ultrasound is applied for about 2 minutes. Depending on the particular target cell to be disrupted, however, the exposure may be for a longer duration, for example, for 15 minutes.

[0197] Advantageously, the target tissue is exposed to an ultrasound energy source at an acoustic power density of from about 0.05 Wcm⁻² to about 10 Wcm⁻² with a frequency ranging from about 0.015 to about 10 MHz (see WO 98/52609). However, alternatives are also possible, for example, exposure to an ultrasound energy source at an acoustic power density of above 100 Wcm⁻², but for reduced periods of time, for example, 1000 Wcm⁻² for periods in the millisecond range or less.

[0198] Preferably, the application of the ultrasound is in the form of multiple pulses; thus, both continuous wave and pulsed wave (pulsatile delivery of ultrasound) may be employed in any combination. For example, continuous wave ultrasound may be applied, followed by pulsed wave ultrasound, or vice versa. This may be repeated any number of times, in any order and combination. The pulsed wave ultrasound may be applied against a background of continuous wave ultrasound, and any number of pulses may be used in any number of groups.

[0199] Preferably, the ultrasound may comprise pulsed wave ultrasound. In a highly preferred embodiment, the ultrasound is applied at a power density of 0.7 Wcm⁻² or 1.25 Wcm⁻² as a continuous wave. Higher power densities may be employed if pulsed wave ultrasound is used.

[0200] Use of ultrasound is advantageous as, like light, it may be focused accurately on a target. Moreover, ultrasound is advantageous as it may be focused more deeply into tissues

unlike light. It is therefore better suited to whole-tissue penetration (such as but not limited to a lobe of the liver) or whole organ (such as but not limited to the entire liver or an entire muscle, such as the heart) therapy. Another important advantage is that ultrasound is a non-invasive stimulus which is used in a wide variety of diagnostic and therapeutic applications. By way of example, ultrasound is well known in medical imaging techniques and, additionally, in orthopedic therapy. Furthermore, instruments suitable for the application of ultrasound to a subject vertebrate are widely available and their use is well known in the art.

[0201] In one embodiment, the hRNA molecule is modified by a secondary structure to increase the specificity of the IscB polypeptide nuclease and related system and the secondary structure can protect against exonuclease activity and allow for 5' additions to the hRNA sequence also referred to herein as a protected hRNA molecule.

[0202] In one aspect, the invention provides for hybridizing a “protector RNA” to a sequence of the hRNA molecule, wherein the “protector RNA” is an RNA strand complementary to the 3' end of the hRNA molecule to thereby generate a partially double-stranded hRNA. In an embodiment of the invention, protecting mismatched bases (i.e., the bases of the hRNA molecule which do not form part of the hRNA sequence) with a perfectly complementary protector sequence decreases the likelihood of target DNA binding to the mismatched basepairs at the 3' end. In one embodiment of the invention, additional sequences comprising an extended length may also be present within the hRNA molecule such that the hRNA comprises a protector sequence within the hRNA molecule. This “protector sequence” ensures that the hRNA molecule comprises a “protected sequence” in addition to an “exposed sequence” (comprising the part of the hRNA sequence hybridizing to the target sequence). In one embodiment, the hRNA molecule is modified by the presence of the protector hRNA to comprise a secondary structure such as a hairpin. Advantageously there are three or four to thirty or more, e.g., about 10 or more, contiguous base pairs having complementarity to the protected sequence, the hRNA sequence or both. It is advantageous that the protected portion does not impede thermodynamics of the IscB polypeptide nuclease and related system interacting with its target. By providing such an extension including a partially double stranded hRNA molecule, the hRNA molecule is considered protected and results in improved specific binding of the IscB polypeptide nuclease/hRNA molecule complex, while maintaining specific activity.

[0203] In one embodiment, use is made of a truncated hRNA (tru-hRNA), i.e., a hRNA molecule which comprises a hRNA sequence which is truncated in length with respect to the canonical hRNA sequence length. As described by Nowak et al. (Nucleic Acids Res (2016) 44 (20): 9555-9564), such guides may allow catalytically active IscB polypeptide nuclease to bind its target without cleaving the target DNA. In one embodiment, a truncated hRNA is used which allows the binding of the target but retains only nickase activity of the IscB polypeptide nuclease.

[0204] In one embodiment, conjugation of triantennary N-acetyl galactosamine (GalNAc) to oligonucleotide components may be used to improve delivery, for example delivery to select cell types, for example hepatocytes (see International Patent Publication No. WO 2014/118272 incorporated herein by reference; Nair, JK et al., 2014, Journal of the American Chemical Society 136 (49), 16958-16961). This is considered to be a sugar-based particle and further details on other particle delivery systems and/or formulations are provided herein. GalNAc can therefore be considered to be a particle in the sense of the other particles described herein, such that general uses and other considerations, for instance delivery of said particles, apply to GalNAc particles as well. A solution-phase conjugation strategy may for example be used to attach triantennary GalNAc clusters (mol. wt. ~2000) activated as PFP (pentafluorophenyl) esters onto 5'-hexylamino modified oligonucleotides (5'-HA ASOs, mol. wt. ~8000 Da; Østergaard et al., Bioconjugate Chem., 2015, 26 (8), pp 1451–1455). Similarly, poly(acrylate) polymers have been described for *in vivo* nucleic acid delivery (see WO2013158141 incorporated herein by reference). In further alternative embodiments, pre-mixing IscB polypeptide nuclease nanoparticles (or protein complexes) with naturally occurring serum proteins may be used in order to improve delivery (Akinc A et al, 2010, Molecular Therapy vol. 18 no. 7, 1357–1364).

[0205] Screening techniques are available to identify delivery enhancers, for example by screening chemical libraries (Gilleron J. et al., 2015, Nucl. Acids Res. 43 (16): 7984-8001). Approaches have also been described for assessing the efficiency of delivery vehicles, such as lipid nanoparticles, which may be employed to identify effective delivery vehicles for components (see Sahay G. et al., 2013, Nature Biotechnology 31, 653–658).

Large IscBs (CRISPR-associated IscB)

[0206] One embodiment of IscB may comprise a bridge helix domain that is split in two by REC-like insertions. The REC-like insertions can be inserted between the RuvC-I and

RuvC-II domains. Such IscB polypeptides are referred to herein as large IscB polypeptides and or CRISPR-associated IscB polypeptides which contain a hybrid CRISPR omega RNA that consists of a CRISPR array preceding a partial ω RNA. Such large IscB polypeptides can generate insertions/deletions (indels) in the eukaryotic genome (See, e.g., Figs. 31A, 39A,G, 40A-C and Table 11.

IsrBs

[0207] As noted above IsrBs are homologs of IscB polypeptides. IsrB polypeptides comprise the PLMP and RuvC domains but do not comprise a HNH domain. IsrB polypeptides may be from about 200 to about 500 amino acids in length, from about 250 to about 450 amino acids in length, from about 300 to about 400 amino acids in length. In one embodiment, the IsrB polypeptide comprises a PLMP domain and a split RuvC but lacks the HNH domain present between the RuvC-II and III subdomains in IscB polypeptides. In one embodiment, the IsrB is an ω RNA guided nickase. In one embodiment, the ω RNA guided IsrB nicks a DNA target. In one embodiment, the DNA target is a dsDNA and the nicks occurs on the non-target strand of the dsDNA target. In one embodiment, the IsrB nicks the dsDNA in a guide and TAM specific manner. Accordingly, applications where a nickase is utilized can be used with the IsrB polypeptides detailed herein in a manner functionally similar to an IscB that has been inactivated at the HNH domain.

IshB

[0208] As noted above IshBs are IscB homologs and may be referred to herein as an Insertion sequence HNH-like OrfB (IshB) polypeptide. IshB polypeptides are generally smaller than IsrB or IscB polypeptides and contain only the PLMP and HNH domain, but no RuVC domain. The IshB polypeptide may be about 150 to about 235 amino acids in length, about 160 to about 220 amino acids in length, about 170 to about 200 amino acids in length, about 170 to about 190 amino acids in length, or about 175 to 185 amino acids in length. In one embodiment, the IshB, or IscB homolog, comprises a PLMP domain and an HNH domain, but does not comprise a RuvC domain.

[0209] Some IshB polypeptides may be part of the IS605 OrfB family of transposases. In an embodiment, the IshB polypeptide is from *Actinoplanes lobatus* and has the Genbank accession number MBB4752409. In an embodiment, the RefSeq database accession number for the polypeptide with accession number MBB4752409 is WP_188124268 and the INSDC number is GGN95087. In an embodiment the protein sequence is 383 amino acids in length. In

an embodiment the amino acid sequence corresponding to accession number MBB4752409 is in the table below.

Accession Number	Sequence (aa)
MBB4752409	MKLVVQVKLQ PTAEQASMLE ATLRACNTAA NEVAQVARRA RVYRNYDLRK HVYAGIKADHRLGSQAAQHV IKKVCDAYKT LTSNLRAGNY GPPDAKRYRR VSTEPVRFRW QAAQPYDARMLSWQHDARTV SIWTVAGRMK NIAYTGSPDQ LKAVAELPVG ECDLVHRDGM WLLYATVEIAEATPVEPAGF LGVDLGIVQI ATDSDGTVYA GEQLNRYRRR QIRLRAKLQA KKTESARRLLVKRARRESRH ATNVNHVISK SIVAEAERTS RGIAVEDLTG IRARVRLRKP QRAALHSWSFAQLGGFLTYK ARRAGIPLVQ VDPRYTSQTC SACGHRDKRN RPDQATFICR SCGVVAHADVNAAVNIAARG VDVWGAVSRPYVA (SEQ ID NO: 3)

Specialty IscB systems

[0210] In one embodiment, the system is a IscB-based system that is capable of performing a specialized function or activity. For example, the IscB protein may be fused, operably coupled to, or otherwise associated with one or more functional domains. In certain example embodiments, the IscB protein may be a catalytically dead IscB protein and/or have nickase activity. A nickase is an IscB protein that cuts only one strand of a double stranded target. In such embodiments, the catalytically inactive IscB or nickase provide a sequence specific targeting functionality via the hRNA that delivers the functional domain to or proximate a target sequence. In an embodiment, the catalytically inactive Cas IscB or nickase provide a sequence specific targeting functionality via the guide RNA that delivers the functional domain to or proximate a target sequence. Example functional domains that may be fused to, operably coupled to, or otherwise associated with an IscB protein can be or include, but are not limited

to a nuclear localization signal (NLS) domain, a nuclear export signal (NES) domain, a translational activation domain, a transcriptional activation domain (e.g. VP64, p65, MyoD1, HSF1, RTA, and SET7/9), a translation initiation domain, a transcriptional repression domain (e.g., a KRAB domain, NuE domain, NcoR domain, and a SID domain such as a SID4X domain), a nuclease domain (e.g., FokI), a histone modification domain (e.g., a histone acetyltransferase), a light inducible/controllable domain, a chemically inducible/controllable domain, a transposase domain, a homologous recombination machinery domain, a recombinase domain, an integrase domain, and combinations thereof. Methods for generating catalytically dead IscB or a nickase IscB can be adapted from approaches in Cas9 proteins, see, for example, WO 2014/204725, Ran et al. *Cell*. 2013 Sept 12; 154(6):1380-1389, known in the art and incorporated herein by reference. Briefly, one or more mutations in the catalytic domain of the RuvC domain and/or the HNH domain of the IscB protein can be introduced that may reduce or abolish NHEJ activity. In an aspect, at least one mutation in the RuvC domain and at least one mutation in the HNH domain is provided.

[0211] In one embodiment, the functional domains can have one or more of the following activities: methylase activity, demethylase activity, translation activation activity, translation initiation activity, translation repression activity, transcription activation activity, transcription repression activity, transcription release factor activity, histone modification activity, nuclease activity (e.g. VirD2), single-strand RNA cleavage activity, double-strand RNA cleavage activity, single-strand DNA cleavage activity, double-strand DNA cleavage activity, molecular switch activity, chemical inducibility, light inducibility, and nucleic acid binding activity. In one embodiment, the one or more functional domains may comprise epitope tags or reporters. Non-limiting examples of epitope tags include histidine (His) tags, V5 tags, FLAG tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Examples of reporters include, but are not limited to, glutathione-S-transferase (GST), horseradish peroxidase (HRP), chloramphenicol acetyltransferase (CAT) beta-galactosidase, beta-glucuronidase, luciferase, green fluorescent protein (GFP), HcRed, DsRed, cyan fluorescent protein (CFP), yellow fluorescent protein (YFP), and auto-fluorescent proteins including blue fluorescent protein (BFP).

[0212] The one or more functional domain(s) may be positioned at, near, and/or in proximity to a terminus of the effector protein (e.g., a IscB protein). In embodiments having two or more functional domains, each of the two can be positioned at or near or in proximity

to a terminus of the effector protein (e.g., a IscB protein). In one embodiment, such as those where the functional domain is operably coupled to the effector protein, the one or more functional domains can be tethered or linked via a suitable linker (including, but not limited to, GlySer linkers) to the effector protein (e.g., a IscB protein). When there is more than one functional domain, the functional domains can be same or different. In one embodiment, all the functional domains are the same. In one embodiment, all of the functional domains are different from each other. In one embodiment, at least two of the functional domains are different from each other. In one embodiment, at least two of the functional domains are the same as each other.

[0213] Other suitable functional domains can be found, for example, in International Application Publication No. WO 2019/018423, for example, at [0678]-[0692], incorporated herein by reference.

Functional Domains Modifications

[0214] The IscB polypeptide (including variants such as a catalytically inactive form) may be associated with one or more functional domains (e.g., via fusion protein or suitable linkers). In an embodiment, the IscB polypeptide nuclease, or an ortholog or homolog thereof, may be used as a generic nucleic acid binding protein with fusion to or being operably linked to one or more functional domains. In one example, the functional domain is a deaminase. In another example, the functional domain is a transposase. In another example, the functional domain is a reverse transcriptase. In some cases, a functional domain may be associate with (e.g., fuse to) the IscB polypeptide nuclease. In some cases, a functional domain may be a protein different from the IscB polypeptide nuclease. In such cases, a functional domain and the IscB polypeptide nuclease may form a protein complex.

[0215] It is also envisaged that the IscB polypeptide nuclease-hRNA molecule complex or Cas IscB polypeptide nuclease-guide RNA molecule complex as a whole may be associated with two or more functional domains. For example, there may be two or more functional domains associated with the IscB polypeptide nuclease, or there may be two or more functional domains associated with the hRNA (via one or more adaptor proteins), or there may be one or more functional domains associated with the RNA-targeting effector protein and one or more functional domains associated with the hRNA (via one or more adaptor proteins) or one or

more functional domains associated with the guide RNA molecule (via one or more adaptor proteins).

[0216] In one embodiment, the IscB polypeptide nuclease is associated with one or more functional domains. The association can be by direct linkage of the effector protein to the functional domain, or by association with the crRNA. In a non-limiting example, the crRNA comprises an added or inserted sequence that can be associated with a functional domain of interest, including, for example, an aptamer or a nucleotide that binds to a nucleic acid binding adapter protein. The functional domain may be a functional heterologous domain.

[0217] In one embodiment, the invention also provides for the one or more heterologous functional domains to have one or more of the following activities: methylase activity, demethylase activity, transcription activation activity, transcription repression activity, transcription release factor activity, histone modification activity, nuclease activity, single-strand RNA cleavage activity, double-strand RNA cleavage activity, single-strand DNA cleavage activity, double-strand DNA cleavage activity and nucleic acid binding activity. At least one or more heterologous functional domains may be at or near the amino-terminus of the effector protein and/or wherein at least one or more heterologous functional domains is at or near the carboxy-terminus of the effector protein. The one or more heterologous functional domains may be fused to the effector protein. The one or more heterologous functional domains may be tethered to the effector protein. The one or more heterologous functional domains may be linked to the effector protein by a linker moiety.

[0218] In an embodiment, the IscB polypeptide nuclease or an ortholog or homolog thereof, may be used as a generic nucleic acid binding protein with fusion to or being operably linked to a functional domain. Exemplary functional domains may include but are not limited to translational initiator, translational activator, translational repressor, nucleases, in particular ribonucleases, a spliceosome, beads, a light inducible/controllable domain or a chemically inducible/controllable domain. In an embodiment, the one or more functional domains are controllable, e.g., inducible.

[0219] In one embodiment, one or more functional domains are associated with a IscB polypeptide nuclease via an adaptor protein, for example as used with the modified guides of Konnerman et al. (Nature 517, 583–588, 29 January 2015).

[0220] In one embodiment, the one or more functional domains is attached to the adaptor protein so that upon binding of the IscB polypeptide nuclease to the hRNA molecule and target,

the functional domain is in a spatial orientation allowing for the functional domain to function in its attributed function.

[0221] In one embodiment, the one or more functional domains is attached to the adaptor protein so that upon binding of the Cas IscB polypeptide nuclease to the guide RNA molecule and target, the functional domain is in a spatial orientation allowing for the functional domain to function in its attributed function.

[0222] In one embodiment, one or more functional domains are associated with a dead hRNA molecule. In one embodiment, a hRNA complex with active IscB polypeptide nuclease directs gene regulation by a functional domain at one gene locus while an hRNA directs DNA cleavage by the active IscB polypeptide nuclease at another locus, for example as described analogously in CRISPR-Cas systems by Dahlman et al., 'Orthogonal gene control with a catalytically active Cas9 nuclease'. In one embodiment, hRNAs are selected to maximize selectivity of regulation for a gene locus of interest compared to off-target regulation. In one embodiment, hRNAs are selected to maximize target gene regulation and minimize target cleavage.

[0223] In one embodiment, one or more functional domains are associated with a dead guide RNA molecule. In one embodiment, a guide RNA complex with active Cas IscB polypeptide nuclease directs gene regulation by a functional domain at one gene locus while an hRNA directs DNA cleavage by the active IscB polypeptide nuclease at another locus, for example as described analogously in CRISPR-Cas systems by Dahlman et al., 'Orthogonal gene control with a catalytically active Cas9 nuclease'. In one embodiment, hRNAs are selected to maximize selectivity of regulation for a gene locus of interest compared to off-target regulation. In one embodiment, hRNAs are selected to maximize target gene regulation and minimize target cleavage

[0224] For the purposes of the following discussion, reference to a functional domain could be a functional domain associated with the IscB polypeptide nuclease or a functional domain associated with the adaptor protein. In one embodiment, the one or more functional domains is attached to the adaptor protein so that upon binding of the IscB polypeptide nuclease to the hRNA molecule and target, the functional domain is in a spatial orientation allowing for the functional domain to function in its attributed function.

[0225] In the practice of the invention, loops of the hRNA may be extended, without colliding with the IscB polypeptide nuclease by the insertion of distinct RNA loop(s) or distinct

sequence(s) that may recruit adaptor proteins that can bind to the distinct RNA loop(s) or distinct sequence(s). The adaptor proteins may include but are not limited to orthogonal RNA-binding protein / aptamer combinations that exist within the diversity of bacteriophage coat proteins. A list of such coat proteins includes, but is not limited to: Q β , F2, GA, fr, JP501, M12, R17, BZ13, JP34, JP500, KU1, M11, MX1, TW18, VK, SP, FI, ID2, NL95, TW19, AP205, ϕ Cb5, ϕ Cb8r, ϕ Cb12r, ϕ Cb23r, 7s and PRR1. These adaptor proteins or orthogonal RNA binding proteins can further recruit effector proteins or fusions which comprise one or more functional domains.

[0226] Examples of functional domains include deaminase domain, transposase domain (e.g. helitron), reverse transcriptase domain, integrase domain, recombinase domain, resolvase domain, invertase domain, protease domain, DNA methyltransferase domain, DNA hydroxymethylase domain, RNA polymerase domains, DNA demethylase domain, histone acetylase domain, histone deacetylases domain, nuclease domain (e.g. VirD2 domain), repressor domain, activator domain, nuclear-localization signal domains, transcription-regulatory protein (or transcription complex recruiting) domain, cellular uptake activity associated domain, nucleic acid binding domain, antibody presentation domain, histone modifying enzymes, recruiter of histone modifying enzymes; inhibitor of histone modifying enzymes, histone methyltransferase, histone demethylase, histone kinase, histone phosphatase, histone ribosylase, histone deribosylase, histone ubiquitinase, histone deubiquitinase, histone biotinase and histone tail protease. In some preferred embodiments, the functional domain is a transcriptional activation domain, such as, without limitation, VP64, p65, MyoD1, HSF1, RTA, SET7/9 or a histone acetyltransferase. In one embodiment, the functional domain is a transcription repression domain, preferably KRAB. In one embodiment, the transcription repression domain is SID, or concatemers of SID (e.g. SID4X). In one embodiment, the functional domain is an epigenetic modifying domain, such that an epigenetic modifying enzyme is provided. In one embodiment, the functional domain is an activation domain, which may be the P65 activation domain.

[0227] In some examples, the IscB polypeptide nuclease is associated with a ligase or functional fragment thereof. The ligase may ligate a single-strand break (a nick) generated by the IscB polypeptide nuclease. In certain cases, the ligase may ligate a double-strand break generated by the IscB polypeptide nuclease. In certain examples, the IscB polypeptide nuclease is associated with a reverse transcriptase or functional fragment thereof.

[0228] In one embodiment, the one or more functional domains is a transcriptional repressor domain. In one embodiment, the transcriptional repressor domain is a KRAB domain. In one embodiment, the transcriptional repressor domain is a NuE domain, NcoR domain, SID domain or a SID4X domain.

[0229] In one embodiment, the one or more functional domains have one or more activities, e.g., one or more of transposase activity, methylase activity, demethylase activity, translation activation activity, translation repression activity, transcription activation activity, transcription repression activity, transcription release factor activity, chromatin modifying or remodeling activity, histone modification activity, nuclease activity, single-strand RNA cleavage activity, double-strand RNA cleavage activity, single-strand DNA cleavage activity, double-strand DNA cleavage activity, nucleic acid binding activity, and detectable activity.

[0230] Histone modifying domains are also preferred In one embodiment. Exemplary histone modifying domains are discussed below. Transposase domains, HR (Homologous Recombination) machinery domains, recombinase domains, and/or integrase domains are also preferred as the present functional domains. In one embodiment, DNA integration activity includes HR machinery domains, integrase domains, recombinase domains and/or transposase domains.

[0231] In one embodiment, the DNA cleavage activity is due to a nuclease. In one embodiment, the nuclease comprises a FokI nuclease. See, “Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing”, Shengdar Q. Tsai, Nicolas Wyvekens, Cyd Khayter, Jennifer A. Foden, Vishal Thapar, Deepak Reyon, Mathew J. Goodwin, Martin J. Aryee, J. Keith Joung *Nature Biotechnology* 32(6): 569-77 (2014), relates to dimeric RNA-guided FokI Nucleases that recognize extended sequences and can edit endogenous genes with high efficiencies in human cells.

[0232] In one embodiment, the one or more functional domains is attached to the IscB polypeptide nuclease so that upon binding to the sgRNA and target the functional domain is in a spatial orientation allowing for the functional domain to function in its attributed function.

[0233] In one embodiment, the IscB polypeptide nuclease comprise one or more heterologous functional domains. As used herein, a heterologous functional domain is a polypeptide that is not derived from the same species as the IscB polypeptide nuclease. For example, a heterologous functional domain of a IscB polypeptide nuclease derived from species A is a polypeptide derived from a species different from species A, or an artificial

polypeptide. The one or more heterologous functional domains may comprise one or more nuclear localization signal (NLS) domains. The one or more heterologous functional domains may comprise at least two or more NLSs. The one or more heterologous functional domains may comprise one or more transcriptional activation domains. A transcriptional activation domain may comprise VP64. The one or more heterologous functional domains may comprise one or more transcriptional repression domains. A transcriptional repression domain may comprise a KRAB domain or a SID domain. The one or more heterologous functional domain may comprise one or more nuclease domains. The one or more nuclease domains may comprise FokI.

[0234] Functional domains may be used to regulate transcription, e.g., transcriptional repression. Transcriptional repression is often mediated by chromatin modifying enzymes such as histone methyltransferases (HMTs) and deacetylases (HDACs). Repressive histone effector domains are known and an exemplary list is provided below. In the exemplary table, preference was given to proteins and functional truncations of small size to facilitate efficient viral packaging (for instance via AAV). In general, however, the domains may include HDACs, histone methyltransferases (HMTs), and histone acetyltransferase (HAT) inhibitors, as well as HDAC and HMT recruiting proteins. The functional domain may be or include, In one embodiment, HDAC Effector Domains, HDAC Recruiter Effector Domains, Histone Methyltransferase (HMT) Effector Domains, Histone Methyltransferase (HMT) Recruiter Effector Domains, or Histone Acetyltransferase Inhibitor Effector Domains.

[0235] In one embodiment, the functional domain may be a Methyltransferase (HMT) Effector Domain. Preferred examples include NUE, vSET, EHMT2/G9A, SUV39H1, dim-5, KYP, SUVR4, SET4, SET1, SETD8, and TgSET8. NUE is exemplified in the present Examples and, although preferred, it is envisaged that others in the class will also be useful.

[0236] In one embodiment, the functional domain may be a Histone Methyltransferase (HMT) Recruiter Effector Domain. Preferred examples include Hpl1a, PHF19, and NIPP1.

[0237] In one embodiment, the functional domain may be Histone Acetyltransferase Inhibitor Effector Domain. Preferred examples include SET/TAF-1 β .

[0238] In some cases, the target endogenous (regulatory) control elements (such as enhancers and silencers) in addition to a promoter or promoter-proximal elements. Thus, the invention can also be used to target endogenous control elements (including enhancers and silencers) in addition to targeting of the promoter. These control elements can be located

upstream and downstream of the transcriptional start site (TSS), starting from 200bp from the TSS to 100kb away. Targeting of known control elements can be used to activate or repress the gene of interest. In some cases, a single control element can influence the transcription of multiple target genes. Targeting of a single control element could therefore be used to control the transcription of multiple genes simultaneously.

[0239] Targeting of putative control elements on the other hand (e.g. by tiling the region of the putative control element as well as 200bp up to 100kB around the element) can be used as a means to verify such elements (by measuring the transcription of the gene of interest) or to detect novel control elements (e.g. by tiling 100kb upstream and downstream of the TSS of the gene of interest). In addition, targeting of putative control elements can be useful in the context of understanding genetic causes of disease. Many mutations and common SNP variants associated with disease phenotypes are located outside coding regions. Targeting of such regions with either the activation or repression systems described herein can be followed by readout of transcription of either a) a set of putative targets (e.g. a set of genes located in closest proximity to the control element) or b) whole-transcriptome readout by e.g. RNAseq or microarray. This would allow for the identification of likely candidate genes involved in the disease phenotype. Such candidate genes could be useful as novel drug targets.

[0240] In one embodiment is for the one or more functional domains to comprise an acetyltransferase, preferably a histone acetyltransferase. These are useful in the field of epigenomics, for example in methods of interrogating the epigenome. Methods of interrogating the epigenome may include, for example, targeting epigenomic sequences. Targeting epigenomic sequences may include the hRNA being directed to an epigenomic target sequence. Epigenomic target sequence may include, In one embodiment, include a promoter, silencer or an enhancer sequence.

[0241] The functional domains may be acetyltransferases domains. Examples of acetyltransferases are known but may include, In one embodiment, histone acetyltransferases. In one embodiment, the histone acetyltransferase may comprise the catalytic core of the human acetyltransferase p300 (Gerbasch & Reddy, Nature Biotech 6th April 2015).

Example IscB Systems

[0242] Example IscB polypeptide and ω RNAs that may be used in the composition embodiments disclosed herein are set forth in Table 1.

Table 1. IscB and ω RNAs

SEQ ID NO:	Name	IscB sequence	ω RNAs
4-5	Ga0207030_1011 GENOME_ACESSION: Alteio_BWMinCont_147_SF_3300020741 GENOME_ID: 25586 CONTIG_ID: 10 SOURCE: JGI	atgctctgttttgttagacaaaaagaaaacccattaatgccatgttctgaaaagcgtgcg cggctattgctacaacgtggcggtgcaagttgtcattcgaagtgatccattaccattcgtttg aaagatcagagtggtggcggtgaaacacagggattacgttttaggcattgatccaggagtaaa acaacagggttaattgctcgcgcgggagtgtagaaaattgattcagagagcggcgaaat taagcgaaactcgttaacagcgcggggtttccctcgcggatatttgctgcgtaaaaaattc gttaattgatttcaaacaggagacatggttattgctgatatcaaaaagggtttaaagt gcgtacatgttggagagtggaatacgttcattctgttatttataatcgaacaaaa atgttccaaggtgtatcacatcgtcactgcaaaatgatgcagcagctgatggatggt tattcaattgtagcgcaaaaaagaggttaagtgggaattggggcacgcttcgcgcgc gcgctatccatccccggcatgaatgctgaggtttcccgcaattga	gtcaatcaccoccccccctaaaggacggagct tgtaaagcataagccaaacagcttgaggtt gaccagggaagcggtaaccaaccgcta agttgtaaacaggtaaacaccaccagcg aatgcttccctcagttcgtctctggaagttg cagatgcagacaaggttcgggttaactacga aacggctctgtgactaggtagacatactgaag gcccgtttacaacattcccaggggagata caccgcaaggtgtgcgtcacaaagaccgta agggtatttttagaaaatggttatgtct
6-7	Moorea sp. SIO3B2 3B2_NODE_173, whole genome shotgun sequence GENOME_ACESSION: GCA_010692525.1_AS M1069252v1_genomic GENOME_ID: 325542 CONTIG_ID: 100 SOURCE: NCBI_Prokaryotes	atgcgagttttgtactgtataaaaaacaaaacattagaccatgccagcctgctagag ctagaatcttgcacaacaggaagggttaaggcttttaggagataccgttcacaatcat cataatgcgatttggaggaattggaattgtgtacacataatcatcagatcaaaacttgaccctg gttcacaaacacaggttttagcaattgttcaagaaaagggtgtgtttgggtgcggaacta accacagaggtcttcaaatcagagatgtttaaaccctcagtagaaaagttgagaagttctc gtagaacccgcaaaactgttatcgcagcgcacgtttcttaaccgttaacgtcctgatgg gtggttagctccagcttaattgtcagaggttcataatttctacatgggttaaaaaacttac tcgattttgtccgataactggcatttctcaggaattgtaggttcgatactcagaagatgga aaacctgaaaatttctgaacagagatcagcagggcacactctacggatatgaagttag agaatatctactttccaatggaatcgtaaatgtgtctactgtgtgtgactgatactcaacta gaggtcgaacacatcaaaactaaataaaaggcggttctgaccgagtaagtaacctagc aagtgaggtatggttccagtgatgagtaagagtgcaactatccccgttactgatttgggt cgtggcagggaataatgactgtgggtctag	gtcaggagcctgggtagcacaacccac ggataaatccgggggctgtgaatcattacc tgaccagcccaagtcctcactagggtactacg ttatcagcaagagtttaagttcttacccttga atgcgacgtagtttcaagctctagaaccgg attgttaacatctttatcagggttaaggagtg caacgcgcgaaggtaccgactgataacatt ggcgacgcaactttaccctagcaatagga gaatactcacattatgcga
8-9	contig_4171967 GENOME_ACESSION: mgm4547164.3 GENOME_ID: 237431 CONTIG_ID: 388754 SOURCE: MG-RAST DATE: 2019-01-27.	gtgatctggatagctgtatcgaagctatggaattgtaataccacaggggaagtgcttacag gcggtgcccctgtggaacaggggaagatttttggaaaaggagcatcggcagatgctcata tatacatagacgcatacaggaacccgtgatgccagctgaaagaaaaggacataattaa gagacttttgaatcgtgcaaggcacgaattgtcactaagggtccattcacggtagacgtg aaatatgagacaccaggtcatcacacagccgttatcagcgggactgacccggtagaac aaacctcggggaagctgtgattgatgaaaatggtactgtgatatataatcacatgtcga acacgaagcaagataccctaagctaatgtctgagagaaaacagcacaggcaggcgt cccgcaggggagaacgtctgagaagaacagcgcgtgctcgaaatgcggtacaaacta aagtatttactgaaggacgcacgataccaggatgcgaaaaacccgtcatgatgaaggat atcatcaatctgaaagcagatttgcacaaggaagcgttctgaaaaactggataacaccg acaacgaggcaactcatacagacgcacctcaatattgtgcgaaaaatattgacgtgctt ccggtaacagactggactcttgaatcaaccgttttgccttcataagatggatgatggcg gtatcagaggtttgatttccagaacgcaggctcaaggctacagcagctcacatgagt acatatagcctatgcaggatggcaaatgtgcctgctgtgataataacatcgacattatcac catctgacaccacgttcaaaaggcgggtctgataccggaaaaccttataggcctctgt gtttcatgccataaaaagatacactactgtgacctgaagcttgaataatcgactcagaa agaaatcagctgactaagcgtactgaacaggcaataaccgtttattattgatggactgta gagataattcgtgaaaaacacgtccatttctgcgaaggctataatactcagatgataagg aaagcctctcataggaagatcatccggaggatgcggtatgcacgcacttcttgag tgggtatccaaaaggtaacggataacacgggaagccttgaagtgatgcagttccgaagg cataatcgagcaataataataatcagcgtgagcgtacctattatctgatggttaaggccgt ttgtaagaatcgtcacaagcgttttgagcagaaggagactctcttgaggaaatttagtaaga aacatccgtctgacgtgaaacgtttgactgttaaaaggagcacacgtttctataatgcta aacagattgatccctggagtagtatttactacaaggggcagaggtatgtaattgtcaggac agcttacaggaggaaaatacctcgtgtgctggagacaaaagacaaactatccggtta cacaatgctatacttaaatgaaaacacaggccttgtgtattttaa	gtcaataaccccgcccgtgaagggcgga ggcttgaanaagcctaactgtgattttgagc agagctatgatacagcagcaaacgccac cacatctgattccaaggtatgtatctggata gctgtatcgaagctatggaattgtaataccac agggaagtgctttacagccggtg
10-11	rumenHiSeq_NODE_386 1232_len_211451_cov_5 332990 GENOME_ACESSION:	atgaaccacaaacataaaatcaaaaggaggaagaatgcttaagcagaactagtttt gtagtgtataaagaagacgtttcttgatcctgtacacctaagaatagcaagattattgctt aaacaatgtaagcaaggatatacaaacacgcttacacaggtatttcttatcaaaccta attatattccttcaagtgttaccctcagaagaatagattggagttgatactggaagtaata	gtaaactactgtagtctaaagacatagcag cttccttaaatgttaccagactagttctttagt aggactacgataaagaagtcatacacctta ggttgcgcctcagacctatgctctgctgtg

SEQ ID NO:	Name	IscB sequence	ωRNAs
	IMG_2061766007_\$F_2 061766007 GENOME_ID: 27375 CONTIG_ID: 59183 SOURCE: JGI	catcgggtgttcagttgtaaggaattgacaaaaatcaagaaaaagtttctagaagtgacac acctttatgaagtttaactaagagggtgacgaaataactaagaacatcgaacagcgaagga tgtacagaagaataagaaggaataggaagacaaggatcgttaaacctagatttttaacc gtaagaactctagacgggaaggactaaaaatccacaatgatataaatttgaaactc attgtaaggtttaatactctacaatccttactccaaaaacaaagttatcttcgaagttg aaactttgacccacatttgatgaagaatgaagaaaaagcatttaacagacattggggtatc aacgtggggtaaattatgattcgtactcgttaagccttatgtattatgtagagataattaca catgtcaacaatgtaagaaaaaagattgtagcattacattgtaccataatgatttatcgtttaa cggcgggaagtgacgatgaatcaacttgataactctttgtgaggattgtaccacaaactt catcaaggaaagataaagttaaagaaatcctacgtcaaggtaaaaaagaaagcacttaaa gatgcaactcaaatgaactatcttaagggtctttgatagaacattatcccaaggctagaata acttgggttatatactaaagaaaaccgtcaataccttaactatctaaagagcattattttg atgctctggtgattcgtctaaaggtaagaaagttaaagggtgaaacaaatcaagttacca atgatgcaaaagttgctaaagtgactatcaactttcaaaaggaaagctgtatgcaaaagatg ctacctaaaggtaagttatggattagaagattgataaagttaaagtatttggtaaatgaat atttcattaaaggaaagaaacatctggaacaggagaattaatgatgtagaaggaata aaattgactttttatattccggaagggtattagaacaccttcaatgaagaatatgataagatt ttcatcaagaaggtctttgatagttaaaaacgtttcacctttaa	atagttaaacaaggcgggaattaaatgaacc tagtgctatcaacataaaaagctttctatctt gtcgagaggaaatccgaagtcctgtaataa ggcagtaggtacactgtgttgagaacaaca aggatggacatcaagtttctatcaataatct ctaaagcaaggagagttaaaaaatgaacc aaaacataaaatcaaaaggagggaagaatg cttaaag
12-13	Ga0334820_006144 GENOME_ACESSION: 713E2X2metaG_FD_\$F_ 3300037089 GENOME_ID: 312159 CONTIG_ID: 6143 SOURCE: JGI	atgcaggatgcttctcaagtcctgctcccttggttagtcagtcctgaagagagaattgcc ccgtaaggagataagaaggttaactctttgtcaataagaagttccgggtgttcgtcaaacg gaattccattgatgccattggaaccagcaagagcgagagaattgtaagagtgatagg gcgattaggagattcaagacaggaattttctatataaattgtagaagagaaaacggag atgtacagaaagtaacatgcggaaatagatcctgggaagtaagcgagaaaggttcacagta aagtcagccaatcatacatataaattgtactttctaattgacagtgatagcgttaaaagataa aattaaaaaagaatgatgatgagaagaatagagaacaaaagaaaccccttatagaaa aataaatgtaatcgaacacatgggggtttgcccttcaaccaggggccagggtggcaag caaaftaagaataaataacataftaaacacattatccaataagttttatattgtagaagat ataaaagtcagaacaataccaggaaaaacaaattggaacaaatcatttcatcattagaaa ctggaaaaaattgtttttttgaactgagaaaaattgtagtgttatataaaacaaagttat gaaacaaaagaatatagagataaacttgatttaataaaacaaaagctaaattagagaat gtttttcagcacacaatgtagatagttgggttttgctaafttgaatttgaccagaaggtc cagataacactaatttttagaatgattccattaaatttcacagaagcgagttgtattattac aaacagaaaagaatgaataaaaaagaccgtttgtgggtactataagtcacggaaattactaa aggaaactgttataaaacataataatattggttatattataggtgaaaaatataaagatag agtttctgggcattgtaagtagacggggcgagagtttatgctaatttaaacaccattgaagt agataactatataatgattaaagtgaggttttgattagtcctaaattcttaa	atataattatcatataaaatattggactgagta agtaggaatgccatttccgaaagaaacagt taaacagagattagagtgcccaagatcatt ggacgcaagcgtgaagtcgtcccaagggt cgtttgaaggaaactgaactaacaccgtg gcaggttgaaacaatagcaggtgcttctca agtcctgcttccctgttgtagtcagtcgaa gagagaaattgccccgtaaggagataag aaggttaactctttgtcaataaga
14-15	Ga0210025_1000354 GENOME_ACESSION: CryGey0102SPAdes_\$F_ 3300025126 GENOME_ID: 25008 CONTIG_ID: 353 SOURCE: JGI	atgcagcagtggttgattgtagacaaacataaaaaatccattaatgccatgcacaccgcaa gggcaaggatattattgtctcaacataaagctgcagtggttaagatctatccattacattat tttgaaaaacaaaacccaatacactcaaccaatacaagttaaaatagatcaagggttc aaaaatacagcattgcactggtgctgagtttaagcaaggttaaacgggtgatttgcgtc taacttacacataaaggccattgcttctgaaaactggctagccgtgctgctcgcgt agatcaagacgcagtagaaaaactcgtatcgaacccgaaatggacatgctatgtct aaaaagcaattggtcatataatcaacgtcgaagaggtgtttccctcatctgttacttca agagtgataagtgtaacgaatttagtaataagtttaacaaagattaacctattatgctctat ggcgttgaaaatgtgcgttttgacactcaattaatggaataaaagcataagttggatcg aatatcaacaggtacattgtttgaaaaagaaagtaaaagaaatatttgcattattccattat aaatgtgcctattgtaagtgattaaagatccattttgaaaaagagcatattattccta gatcaaaaaaggttcaacagaatagataattatccitggcatgtcactctgcaatcaag ctaaaaataattgttgccacaagaatggtaattggtgcttaaaaaatcaacgtcaaaaatta ataaaagcgagtaaaacgctttaccatatacattaaaggcattaaaccgtcattgcgtgat gcggcagtcagtaaacgccattcgtataaattagtagacgcaattaaagggtatttattaccg attgaattaggcagtggtgggttaaccaaaattcaatagaaccaatcaacaattaccaaaag atcattggatcgatgcagcgtgtattggacaagcggaataatattattatccctaaaaat ctaataccattataaataaggccgttggggcgtggttcaaggcaaatgtccttgatgaataa atttggtttccacgaacaaagccgaagaagaaaggcattgttattttaaaccagggtg atctcattaaaggctatagtcactaaagggttaaaaaaggaaggaacgtatattgtagattagc ggtaagaacaagtggttaatttcaatatcaacaaaaaatggcatagtgcaaggtattaaact accgttattgtcattgtcttcaaaaaaacgatggctatcattatagcaatcatataatcaatt aattactaa	gtcaacaaattaaggtagatcgtaaagattat gggtgacccgactcagtggtttgcat tactactgtaagtaagaatatatagtgagtc ggaatactactccagttccacaaactcagga tctacgttaaacagtcattttggtaagaag cagtcgtagtgcataaaccttatttaacatt gtctaggagattttaccctcgtaaagggtgt cccggtatgggaaaaattcaaaaaaggaaat a
16-17	Ga0256405_10001455 GENOME_ACESSION: RumenRJG_02DNA_2_\$ F_3300028048 GENOME_ID: 34937	atgtatttagtatactgtaaaatctaaaatggaaccccttaattgectactaacgggtggtg agggttagaagctgtttaaaaggaggaaagcaaaaagtatatgtataatgtccttttacga ttaaactctttatgacagactgagtagactcaaaaactgtctttgggtattgattggttca tcacatctcgttctgccgtgttgaatgaaaaggcgatagctgtttatggctgaacacaca atcaaaaatgacataaaagataaaaatggacaaagacggatgtataggagacacagacg ttcaagaaaaactcgttatagaaagcgtatgattctaaatagaagaatagtactaaaaag	gtaaaactacccaccccatgaagggtgtg gctttcattaaataacagtttaccagactaagt ttagagaaatctaaactacgatgttaaggcca tgctacctttgttgacgcacagatcaatgc tctagcatcttaaaaaactgataaggtgaag taataaatgcttactgataagatgacaagc

SEQ ID NO:	Name	IscB sequence	ωRNAs
	CONTIG_ID: 1454 SOURCE: JGI	gtagattaccacactctttaattagtaaaattcattctcatgtaaaagagattgaatttgttaaat ctatttctccggtaacagatgatgatctaatftttgaaacagctaaattcgatatgcatttattg aaaaatccataattgcataatgaaaaatagacattttggtatcaaaaaggatatactttatg ggatgctaattgcccgtgaatcacgttttagaacgagataatcatgaatgtagatattgtta aaaaagaagggtataaacgtaaaaatggtattcgttagaacaacatcatattgtttatagaa gtcgggggtggaagtgacgatccaaggaafttaattactgtatgtccagatgcccaccaa aatacatgatgtgtaaaatacaatagacattaaaggaaftcccttttgggttttaagacatg atactcatatgaatataatctcaaacgattagtgtgaccgttaccctaaatgcaatagaacat acgggtatataactaaacaaaatcgtttgaagctaaattgcctaaacgacattatatagatg catgtataatagctaattggtggaccagatgttaattttaaattcgtatatagtatataataaag aagtgttactaaagggtattatagacagactaacggaaaacggtctgaaaagcgaatgaa tagaggtaaaagtaaacggttttaagacgtttatgataaagttcaatataaagggtatataacttt ataaaaaggattgatagcaaaagggttgcgactttaatggatataaacaataaaacaattaa atttccagatgcgcctaaaagataaaaacacctaaattatctaaaataaaaaggataact gcaagaacacatgcttaatagatataaaaagggttcacattgcatatactcgttaa	cttaatgctattgctgagatgaagtcggatttt aattgtgtaacagcaattaaatcgcataac taagctttaggacttgaattatattctaagga ataattat
18-19	Ga0224415_10000689 GENOME_ACESSION: 3300021399 GENOME_ID: 241587 CONTIG_ID: 297 SOURCE: JGI_Dark_Harvest	atgccttctgctgaatccaaagaaggtacgaaagcttctgaaagaaggacgtgccgtcatc gccggatataagccgtttaccattcagcttacatcacgaatccgggaaggaaaaacagccct gtgaaatgtccatagacgccggcgacagccagctagggttaattccatcaaatcagaaaa gcatgaattcattcatgaacagcgtgacctgcttaaaagtggaacaaaaacatgatgat cagagaagactacgcaggacacgccgttaacagaaagcgtgacagaaaacccaggttt gacaacagacgtatccctgaagcgtgctgtcccgatcagaaacagaaaaacat ccatgcatgttatagacacgtacagaaaaagtaattcctataacggatgtttttattgaaac cgttctgtttgacacaaacgcgctgcatctcagagcaaggcgtcccgacactgaaga aacagattatcagatggctcccgttcggatagacacacgtcgggaagcagattttac agggtatcatcatcctgtcagatctgtgtgtctacgacgtcagataaaaaagaaagatg ggctctttaaaccaggcgaagttatttaaggatgcatcacataggatgacgaacaggaga ccggaccgatcgtatgctaatctccttaccagtatgacaaggtgcccatacgcctaataatc ataaacccggcgggaagttatgatctggaacctgtgacacgaaccatccggagcc gcttcatgaatacagatcagatgtatgtattcaatgagataaaagccattgacagcgtgt acatgttcatatgacatatggatcggttaacaaacgtgaagggctgtcaaggcgtatcaga aagacgcacgcaaatgacgcgtactgtataggatattcagaccgaacataaagcttct gaggaaatcttcagaagatccgccgtaataacagggtgccttgaataattctatgacgtg tatatatcagacccgtacaggaagagcttccggcgatcattaaagctgcggaaga acaacagatcagatcagaacacgtgatacaataaacgaatataatcagcagaaaa cgtacaaaagatacagaataatccgagaaaacgtatccctgcaggcaggagaca aagtaatttccaagggaataaacacagtaaaaggccagcaggacaagatataccaaa acacatggtttcatgaaaccgttgaaactaaaagaataccaaaagaacatcttttagaaga gataaaactggttagtcacataccggctggaaaaaagtcacagccgcatcataa	ttcggattactctttatgaatctctgtaagccc gggttattagccctaaatgacgaattgggtgc tctgacatcacaccgtcaattatcagagaag gcatttatgccagatctacgtttgatccaat aatatagtcacgtcagaatgtccacaggttc tgtccatgagcctgtatattaaaaatttga gggtaggaaacgtgtatcagagatgatacc gaaacccgtttaccggttttgatattcga cgggtgtgaacattggcgtgtgtgcccactc ccgagagggaagataccgtacatgcttttaa tgtacgggaaggcgtgaagccattaaaagaa aggaggcataagccctgagctatatt
20-21	Ga0210041_1000150 GENOME_ACESSION: CryGey2003SPAdes_\$F_ 3300025034 GENOME_ID: 25011 CONTIG_ID: 149 SOURCE: JGI	gtgtctgtattgtgatcgacaagcgaagaacccgtgatgcgtgtgagcgagaacg agcaagaaggttctgtgttctggccgtgcgcgtatctacgcgtcattccgtttgcgatac gactgattgatcgtgaagttgccgggttgcgatttgcagccgtttaaatacaagattgacc gggtagcaagtacaccgcatgtccgtgtgctgcgaatcaaaaggaagtggtgttgc ctggcgagatcagcgttaccgcatgtactgaatattttgagtactgcatgcggccgt cagatcagcgaagccttaacgtcgcgcggcgaatgcgcgcgttagcgtgtggtgtt gcgctaccgagcggccaggtttcttaacgcggtaacaaagcgaaggcgtggttagcgc caagcctgcaacaccgtgtcgatcagctgttgcattgggtgaaccgttccggcaattgg ctccgatccagatgctcaggagttggtgctgtcagatgcagcggatggagaatc cggagatttcgggtgtcgaatatcagcaaggtgcgtgattggttaccgaagtgcgcgagt atttcttggaataatgggtgcgaagtgcgcgtactgtgacgccaaggtgtgcccgtg aatctggatcacatccagccaaaagccagcgggtggaagcaaccgtgtttcaatttgacc ctggcctgtgtccgtgcaaccagaagaagggcgacaggtgttagcgtgttcttggc gaaagaccccaagcgattggaatcgaatcgtgacgcaagcgaagcgtctttaaaggatg ctgctgcgtcaatgcgacacgctggcggtgtttcaacgcgtcaaggccactggcggttc ctgtatcaacagggtcaggtggcttaacgaagtataaccgtgctcgtattggtatccaaaa gcgcacgcgtttgacgctgtgtgttggcctggtgagtgccgtgacagcgtgggaaa gccgacgctggcaatcaaggctaccgggctggggaagctaccagcgcacgcgttaga tgcgttccgggtttccgcgcgttacctgacacgagaaaagcggattcaaggattccagac tggcgatatggcagcgtatgttccaacggcgaagattggttttatgtgtgctgag tggcggttcgcgcacccgtgatttcaattatcagaccgcgaaggagctgataagggc atatccatagcagcgttgaaggtgattcagcgtgctgtagtatgttattcaatcgtggc aaaaatgtag	gtcaactaccccgccctaaccggcgagct tggaaaggcagtaagaagctcgttgacca gcctaagttacggagaacgtaactacgttg tagcgaagtacaagaccagctgtggatgc ttcctcagtcacacacctctgaagcggcagg agtagacaagcgaagggttaagcacgaac agctgtgcgaagtcaaaagaccgaactg aagctgcgttacaacattgctgaggggagc gaaccgaaagggttctgtcacaagggtgcta agcacatttaggagaacaagctgtct
22-23	Ga0163150_10004194 GENOME_ACESSION: OliLakLVI9P2IB_2_\$F_ 3300020195	atgttagcttcaattttaaattctcatgtgcaaacctctgatgccttgaagccgcagaaagcc cgtgttcttttaaaagacaacaaagcaaaagtgatcaaacgactgccgttcaccattaaagct caaatgtgagctcaggttttaagcaggagttaacggcagcgatgactctggcagtaaa agtcattggcaccgctgtagtaactagaagtgcaaaagtactgtatcaggctgaacacgac	gtcaattacccccacctaaggagggggctt gtgaggtgactcataagggtactagttgac cagaccactaacaagaaggaagcaaaaa attgttagtaaacgttaaagcagaatatcaca

SEQ ID NO:	Name	IscB sequence	ωRNAs
	GENOME_ID: 33727 CONTIG_ID: 4193 SOURCE: JGI	tcctcagggtgaagagattaaaagtaagatggccaccaggcgatgatcggcggaata gaagaggcaggaagaccagataccgacaaccaagattctcaatagaagagccctcgac ttcccttaaccggcttccccctcgacaaaacacaaaggttgagtcacatttagccgaaaag aagtataattgagtcacatcctgccgatcaaaattggcgctagaattagcttcttcgacatc cacgccttaagcaatccagaagfatccaaagcccgatggtagcctatcaacgtggcgga aatgtatggcttcagaaatcttaagcaatagtcttaagtcgagatacctatactgtcagac ctgtaaaaagaaaaccgaagcagaacgttgagcttcatgtccaccatattcattcaaatcaa atgggggggctgataccaaaacacctgattacctgtgttaaaccttggccgacaaagc ttcacagcataaaaaatgctcaacagcattctctaaaattaaagccaaagcgacaaatc caaacatgcaactgagattacatcgtcgccgctcaactaagaaaaagtgattggaacttc gcagagacttttggctttattacaaaagttacgattggcacaaagtttggcgaaaagac acttcatcgtatgccgagtcatacctcccaaggactggatgccgagctttaaaaagaac gatcattagaagacgtgtagccaaagcgattaccagcagacaaaaggttctagatctga gaaaaccataccgacgggaaaacttttggctaaagaaatcgaattatcaaaaacacct aagtgtactggctttatcaaaagcgaagatcgtcaggtttttgcccattagcgacatacat ggccattcgtatctgaactccgtcaaaagttaaagccgactgtaccgcatcacagcgga acgacaacccctgcttatgtggagaatttcgaggtcgacgcatcgtatcgtcgccgaggg cttcaccattctcctccacctaa	ccctcggatgccgcctcagtcctgttgcctgt ggcctactattaaacagagttgaaagactca gtgtagtagcgtaaaaaagtgttttaactgg tcgggaggaagctgtgatatcagattcaacc ttttaagagctgatatcgtcgtcactgttag aatttagcaggag
24-25	Ga0209777_10000113 GENOME_ACESSION: FrelakHBHBSA Des_ \$F _3300027896 GENOME_ID: 26210 CONTIG_ID: 112 SOURCE: JGI	atgaagaaagtaattgctctgtagtagacaaaaataataagccacttatgccaactagtgt ttggagagctccaagtggaattagtagaaaagctactccattttggaacatggtatatt ttgtgtacgattaaacgtagaaccttctgccagaataatgcaacctatagcagtaggtattg atccaggtagcaaaaagagaagcatttactgttaaatctaaagaataacctatgtaaatatt ctgacacatgccgttactcatgtttaaagatgtgtaggagatagaaaaagtgctcgtaaaa gtcgccgatttagaaaaacacatatagaaaacaaagaaaaatagaaaaaggggtga ctaccaccttctactaaagctagatggcagttgaacatttagggttatataagcttattaaa attttccaatttcaattttagtagaaaacattaaagccgttaacactggacaaaaaga tgggattctaatgttagtccattaggggctggcaagaatgttttatggcgagcttagaaa aattgctccagttagcttatgcaaggttgggaaactgttaattaaagaaatcgaactggact tgagaaaacttatcaaaattggatgaaaaatttccgtacataatgttgatagttggacttta gcttgagtggggttggcggaagaaaaaccagataatgaacttttaatttttagtttca ctaagatttcatcgaagacaattacattttagtaatttttaagggggggaacgacgtttat atggagggaacaagaagtttaggttaaaacgggggaagtttagtagatcattcagaatag gactttgttatgttgggtgatcatctaaaggtagataagtttaccactgttagatgttgc agaatttatcaaaaagttaaattgaaagattataaatttaaattgtatagtagtttttagatttata cagagaagataattactgaccttttggcttaa	gtcaatggccagtttctaaatgtttgaaag tcctccaaaggcgcaactttcaacgtaaagca ttgactagacacagattggaagggaggttaa agacaaaaagaaattccaactcaactgaac aatctaccctgggaaggggtgaaacacgttg cgaatgctaccctagtttgcacctctctgtgatt gtcaggtgcgaaggggtgtatacatgataga gttttcttgtaattgcctaaaacgattcaagga attttggagaagcaaatgaagaaagtaattg
26-27	uncultured Clostridiales bacterium isolate RUG10138 genome assembly, contig: RUG10138_asm_5, whole genome shotgun sequence GENOME_ACESSION: GCA_902756305.1_Rum en_uncultured_genome_ RUG10138_genomic GENOME_ID: 344355 CONTIG_ID: 4 SOURCE: NCBI_Prokaryotes	atgacgactttcttataatggcagaagggcgaaaagctgatgccgacaacaaatataaaa aaaatccgaaaagctctccgatcaggaagagccaagatcgtgaagcacgcgccgttca ccgtgcagctctgtacgagagcggaacgccgtgcagccaatcgagttacggaagat accggctatcagtatatcggagctctctcaaatcgagaaagcacgaatcgtgagcgcc gagtatacgtctttaaagaaagagaagcagcatcacgacgagcgtcgaggttaag acgcccgcgtcgcaacagaaaagcgtaccgcaagggcgcttcgacacagaagaaa gtctgaagcgctgctcgtccgtcgctcagaaacaaagcggaaccgcatgtagatactt caagatgtactcgcaggtatgccgataaagcgctacgctcagattggacagttcga cccggcggtcgtcgtcgtgtagagcaagcgcaagcccttccggaagcggtgtattacc aatacggcgcggttacggatagacacgctgtagagaagcagcttctgcaagagacgg ttacagatgcctctgctcggttaaatcagttatagaagacggcaccgctcaggtccat catgtggcgcttcgcacggcgacagatcgaacaggtcggcaacctcgcacatgtagt cgagaaatgccacagcccgaagaacataaaccggcggaactgtgggatatggc acctctaaaggcacggcatccgcagcgtacatgaacatcgtcaggtggcatactacga gggagtcaaggcggttggcgtgacacacataaacctatggtgccgtaacgaagcgca caagacgtgatctcaacatcggaagtcacacgcgaatgatgcctactgcatcggcagc atgagaccgaacgccgcacgaagacgcaatactttgaaaagcgccgccgcaacaac cgcatatcgaagaagtctacgacgcgaagtatctggatattcgtgatggcaaaacaaag aaggcgcgagcgtcgggtgtaatcgacacctgagatccataccgcgaagcaatccgc agaacgagcgctcttccgcggggagaaggtgtcgaacggcagacgcagtatagcg ctcagaggtatctgtaccagcggacgacattatcatcttcggcggaagaaagcgcatgg tcaagggtacgcataaagggctcaggtgtcagcttattggaggagcgatatttgc cgagaaaataaaactcatcaccacgctggtgctggagacaagtagtttaa	gtcaactaccaccgcttagactgacgtcttg aagcgggggcttgtgaaaacagtcagttg attagcctgagtgctcagggcactacgttatc tgcgaatatacaggtaccgggagtgctcca caagtcgccggcactacggatagctgttaaa catcgggtgaggtagccgaagtgcgtatat catcaaacgcggaataacattggcgatgtg gacaccgcccttcggggcgaggatccgtat acacgcctcttcggagacaggtatgcgga aagcgtaagcgagcatgttagcactaaaa aaacagaagaggagacatgacg
28-29	Ga0190327_1001504 GENOME_ACESSION: 48720401_MG_2_ \$F_33 00021505 GENOME_ID: 20975 CONTIG_ID: 1503 SOURCE: JGI	atgattcacaaggttaattggacatccaacacgaaaaggtagaatgttcggaagattat tgaagaaaggtaagcaaaagtaagtagctgggtgtgtaagaaaggcagccattgttaa tacaattgctgtgataaagtaattgataaactaagactatagatacagaatttctgtattggtatt gattcaggctataaatattggttatgtttatttaaatatataaaaatcatattgaactatta ctatctggtgaagtagaaactgtacatctgaagttactaaaacttaagtgatcggaagat gtatcgaagcttagcatcaatataagacgaagaatgttaaacgtaagtttagtaagccta agtttagacatcctaaatggaaaaacagagcctaaccatgctttcaacctactatagacatt	ttttaaaaaacaggaggagaaaatttatgaaa gtcgtgcaggcattaaagtaactttaaatttta ataatttaaggtgacacagcagctgactattg gttcaccctgggaatgccactccagttccag gctctgaaggtgttagtaattatattacttagtc ctcattgcatataatttcttaataataacgtatt accaatagtcagatcgggggagaccaacct

SEQ ID NO:	Name	IscB sequence	ωRNAs
		tgattacaagtcattataatattaaaatgggtatttfaaaagagtaccaaaggatcaatgtga ggtacatttagaataatagtaaaattgatgtcaaaaaatgatcaatccaggattcacagtgg caatatacaacgtggctcctaataatgggtttgaaaatgtgaagcatataatcgtgatagat aattatacttgcagatatgtagaacaattatgctaataaaaaaagaagtacatcatatta taccagaatctaaagggtgcatagatagaccggataatttgattctttatgtcaaaattgtca taaaaagtacataccggaaaagtcttgcacatcaaatgtgattcacagtaaaattcgtga cgtggtgtattgaattcttgcataaatggatgtttgataacttttagtaagaaatttcattag taaaaacttttgatataattacaagactgttagtaataatagtaattggaagatttcaagac acatgctcatgatccatgataattgcatttgcaatgaaatggctcgaactaaattcaa aacataataaattatgatcatcataaaccgttaattttaacaataatcgtagacataattgat ctggacacaaagattagaagatcgaataattatttgaattgattgaaatttaaaaaatg cgtagctcataatagaagaagagctacagcacaataaaaaatctcctgtcttgaaga atatttagaaaaatctaatatgcaataatagcaaaagcctgggtgacggattatgaaca aagcgaagaataatgaatttgcggaggtgatataaataatgtccaaaaggattgtcta ctgtgcaattttgaacttagacataaagaattactactgcacaatttgtagaattagag aaagagttgtactaagattttaataatgtcgaattgtctatcgttag	aattaattaagctgagcttttaattagctttatc atgaagtattttttaaagcaagtgaattgga gacaaattttaacatgatt
30-31	Ga0376669_0068582 GENOME_ACESSION: CalSwaMetagenome_3_F D_SF_3300036830 GENOME_ID: 312381 CONTIG_ID: 68581 SOURCE: JGI	ttgtcaggccccattacctgggcaaccagagactccgaaagggtgaatcgtgagcaa agtattcatgttagacaccaatttcagacaaactaatcccgctccatcctggggaagccaga aaattgctttcagcaggaaaagcagcgggtaccgtcgtatccttttaccatcatcctcaa aagggctgtcaagacacgtgtgaacactacgggtcaagatcgatccaggctcaaaaa ccacggggatcgccgtcgtcaacgatcagcaggagaagctgtgtttgagcagaattg tctcatcaggagagatcatcaagaagcgttagacaaaacggcgtggagtgcgtagaag cagacgttaacggcgtaccgtatcgccaaagcagatggagaaaaccgacgcaacaag aaaaagggtggtgtccaccatcttgcagagcaggatccaacagtgacggcgcat ggcagcacgtgcgacaggtctcattatcatcaaccaagggcaagacgggtccaagga attggataccagattgcaagctctccatcggtcagatggatagctaccagaaggga gcgaccattgtgctgtcccgccggaaacagagacggtagtgccttccctgaacgtgtt gtctag	Gtcaggggacccacgcatgaatgcgggg gcttcagtgatccggcaatggattagtgcg gcttcacgtgaccagactcagtgtaggctt gtctctagcactacgttaggagcgaatatg tacgttggatgctgtggccagctcccaacc gctacgggtgcaggttaaacaggattacgag ggttagtgccagtgctacacgctaaaccg cataacctgtcaggccccattacctggg caaccagagactccgaaagggtgaatcgtg agcaaa
32-33	Ga0394878_0025107 GENOME_ACESSION: SIS_S4_MG_FD_SF_330 0037310 GENOME_ID: 312643 CONTIG_ID: 25106 SOURCE: JGI	atgtccaaagtatttcaatcgacactgatttaccaccattagaccccggtcatccagcaca gcaagacaattattaaagagaaaaagcagcagatttaggcgggttccctttacattat ctgaaaggagtcacgcctaatctctgttttcagttcacccctcagattaaaaaattgatcct gtagctaaagttagcagaattgccttagttaacgattctaatggcgagggtgtctgtgct gatttaagcacagagatttgaattcgagacgttttaattttagaagacaattacgacg tagtagaagaaccgcaaaacccgtaccgtcaaccaagattcttaacagaaacagac ctaaaagctggttagtccaaagcttacaaagctgagtgataacatcaaaacttgggtgaa agactaagaaaaatagccccaatcaagctatcagtcagagtttagtactgttttgatacac agctaattcgtaacccaacatccaaggtaacgagatcaacaaggaaactcaatcaaca cctgttcttaaaataaaggaggttaaacattgttaattacagccaattggacatggttctaga caatcatgccgtactgataatgatgatttccatctgttatgttccaggttttaatttgttaa ggttttcaactggtgacatcattaaatctattgtcacgaatggcaaaaagattggcgttat attggacgaattgtctgtccgtcgacaggagtttaacatcgcaacttctcaaggattagt acaaggcatttaattacaatactgtaaccaattcacaacaaagacgggtattctctatgcaa catag	gtcaacaaccaccgattatcgggggcttg aaatacagccctagttagcagattcaggta gcaatacctacgtttaagggaagaggttaact cctacctggaaatgcgaagctagtccaagc tctagaactcaaaaggttaaacagccatctgg gaggcagtgcttttgatagtagtaccgacct aaacattcttaaacattatcaagctaacatt accggaaggagggcgagaattgtccaaa
34-35	Ga0163150_10002834 GENOME_ACESSION: OliLakLV19P2IB_2_SF_3300020195 GENOME_ID: 33727 CONTIG_ID: 2833 SOURCE: JGI	atgcccaccgaaaggcatggatgggtgctgcccgtctatatacgggtgggcaagtcaaaag tcgtcaagaggttcccttcacgattcaattgacttcaaatcgaatcgaagacgagttcagcct ttgactctaggtcaggatattgatttgggacgggttgagtcagcgttaactccgaactaaa ggaagctttgtcgtccgagtcacaaatccgacacgtttcagaaaaagtaactgaaag aagatcctaccgcagaactagcggagcgaataagaccgttacaggccagcgaaggttt gacaatcgaaagagaaaagcgttgacgcttcgatcaagcaaaaagtcgagagtcagta gcagataatcaaaaatttgaacgatcctgccaatctcaaatgtgattatcgaagccaata actttgataatgccaataatcaataagcccaacatatacaggacgagattatcagaattgtga acaaaaaggcttctataatgtcaaacagatgttctcgcagagatggctatcctgtcagg cagggaaaaaaggctgctgtggacaacttcacgtccaccatctcacattcaaaaagccag ggcgggaagcgtatcccccttccaacctctgacgtttgtgaaaaacatcatgccgatctc atgctggcaagcttcaagttaccatcaaaaagcataagacactcaaaaccgccacatga tgaatctgtgcgagccagctactatcacggaatccagggtttacggaacttttgatgat gagacaaaatttgaacgagaactcgttgaaactgcaaaaacccatcacaatgatgccttc gtcatagctggcggcagatgtcaaaaggcagccagggttcatttcatcacgcaaaaagcg caagaataatcgggcaattcagatgaaccgaaaaggccaagcaccagccaagaaggag caaagatacaaaattcagcccaagacatcacaatgagcaggaaaaaagtatttgc ggggggtatgcaaaaacaaagggttacctatgttttggagtgattgtaaggaaaagtatg ttaaacgattgcgcagatcaaaatcttttcatcaaaagagctatgttctcagcgcgatc aa	gtcaatccccaccacccatagcgggtat gggtagggcttggaagaaaggaaacaattt tcaagctcttgatttagtactcagccctg gcaacagggtcagcttactggcgaatgaca taggcaccttgggatacttctagtccagg ctctgcgcgagagcggttaaacagttctgatg ggtaggacagtgctttctcgaaaaaccgcc ggataactttgtcgaaggaatccacagccg aaaggctgccttaccaggaacccctaacggg gaagaagaagaatagatg
36-37	Ga0208279_1011523 GENOME_ACESSION:	atggtgcgcagaggacgcagaactcgaagaccggtaccccaaccacgcttcgaca accggcgagcggaggaagatgcggcgacccatctctatgtctcgcgtgcagaatgt	ggcaaccgctcattttcagcagatttctg gatccgggtgtgacggccccaggaccg

SEQ ID NO:	Name	IscB sequence	ωRNAs
	AntAce02UKSPAdes_7_ \$F_3300025649 GENOME_ID: 22962 CONTIG_ID: 11522 SOURCE: JGI	gcgccattggcgagatcgttggcgctctgtaccactcagtcagattcggttagagac tggttccttcgacaccaactcatggagaaccaagtatcctgagtgtagtaccagca aggcgaaactggcaggctacgaggtgcgcgagtagctactggagaagtgggcgcaa gtgtgctactgtggcgcgagaaatgtgccacttgagtagagcacatccgtccgaagtc gcggcgtggaagcaatcgcgtcagcaaacactcatttcgtccatgcctgcaaccagg ccaaggggagtcgggatgtgcgagagtcctggcacatgccagcacgattgaaag gatactggcgcaagccagacagccgcttcggatgcggcggggtgaacgccaccg atacgcgattggg	gtgcctcgcgtggcgctcgaagtgtatattc cttccccctggcgactcagtgcccttttga ctgtcggcgcgagtgccatctctcggg ctattcgctgatgctctgttttcatgaaaga aagctcgtcggatgtttctcgttcgtccacct ttatcatctctccctccgtatcatagattgt tctcactgccataatagcagagatcggtt caggtgtgctctgtttcagtgatcaaaatccg gtcaggtgtgctccctttatgttatcaaaaaca caaggcgctcgtggacaagagg
38-39	Ga0247609_10000157 GENOME_ACESSION: 172DNAGHGlowgp2_2_ \$F_3300028888 GENOME_ID: 279758 CONTIG_ID: 156 SOURCE: JGI	atggtttacgtaatagacaaaaaatcatcctctaatgcctacgcacgttacggaaaagt acgcagaatgcttaagaaggaagctaaagtcgtaaacgagaaacctttacgattaa actgctgtacgattcatgtattttgtcagccactagtctcggattgacacaggaagcg aatacataggaaactgccctgttctgacaaagggaatggaaagttaattttgtattcatcc atcgttaaatctgtacagacatcaagactaaatgaacaacgtagaagatcgttccg cacgtcgtggaaggttactgtgcagaccaccaaggttcttaacacgttttgcgtacactaaa aaagacatcatggacaactgaactctgtggaatctgtcaagcaaacgggttattaaag acggcaaaagagtcacaactcagtggttaagaatacaacaagaagcagatacagcatt tgaacgatgcctgtcatgtgcaccagacaaactcctgactcacagacgacaactctgtc tcctacgatacgaagcaagataggttcgcatgtaagagaagttgtgataatagcaggata cttctgtcagtcgtttatgtgctgaagtcggcagttcgacacacatgctttgaaaaatcca gaactgaaagacatgtcgaatcctgacattcgtcagtggtttatcagcacggaccacaac tatgtttatgaaagcacaagaacacatgttttagctagggaatagtacaatgcagaatat gcgggtgttcttcaaggacgttcagggaatacactgcatgttcatcacattgtattccgaa gtaatggcggaagtgtgaccttgaataatgtttacgctatgtgaaagactgtcaccacac actgcataatgatttcaagacgatgacaaggacgcatttgaatgtgaaatacaagaccttg ggttctggaacgcataagatgaaagtgttaaaacacgcagctcacatgaatgtcata cgtgcacaactttacagaacgccccaattcagtcacaaatggctgtgaaagattcatcag aacaagaacatattatctcgtttatgtagaatgcaaccgaaactttgggtacattacaaaag cgaacaggcaatggcatggaatcgaagaagaccatcatctggatgcgtgtatgattgcat ctgtgttctaaattcacgggtgaccagatacaccgggtttcaagaacgccatattgaa tctggaaattatcaggtatgcgtccgtaagacagcaatcgggatattatggctgagtcaa gaaagcactgaaaaagggaaggtagagaacttggggatgctcaagaatgaatacaag cgggttaagaagcagctagagagcgtttaaagctgaatttgatcttcaatagtgataa tgcgggaacagctcaaaaatgtaaacgcagatggattcagacggatgtgtcctgtcaac acgcgacaaggtcatggatgtcggctgtattacaaagtcaattactacgggtcagctggat tctcgtgtcgttcggactcaaaagtgcgggtatcttttgatgtgttcggtaaaaaattg gaaatctcagaacacagaatacattgaagaaaaagacagtgactgcaaatcgccttcc aataaaaaggatgttctgtgtcgtcggcacggcatactacgttatgcgaccatgaagttc caaaaagactgagtcatacaacgattgtacgtctgtttctgaggatga	acgagtgttataccagggtgagtggtgata tggcactgattatgctcagaggatgacattt gttatgagttaatacagcaaacgtctga ctcaacggatacgtggtctcagaggatgac atttgaacgggtgttaatactctgccgtac cgactgtctcagaggatgacatttattga gtgtaatacctctaaaacgcctgtacaatg tcatgctaaggggaacagactcagctcttcg gagctatgcatcaactttaacgaaacgttcg atgcctcctcagtcagttcaattctgtcgtgc ttgtcagggggaagtcgtatcgcccttacg ggctgtacgcattacatcggtctcatgctga gtgcttacgcaacatcttaacgtggaggttc taaaatg
40-41	Ga0116227_10015234 GENOME_ACESSION: SITSphgellanicum_5_ \$F_3300009709 GENOME_ID: 34953 CONTIG_ID: 15233 SOURCE: JGI	atgtccaaagtactagtaattgacaccaacaaaaaacgcaaaatccaattcatccctccc tggcgagacagcttttaataacggttaaagccgcaatatttagaaaaataccgtttacagg ctatgaaacgcgagaatacctgttagaaaagtggaacagaaaatgtgcttactgcgatac aaagatgccgttgcagatagaacacattcctcagagcacaaggagggcacaatcg catttctaactaactttaagttgtgagaatgcaatttttgcacaaaacgatcagtag	gtcagtaactcaggcttaaaagcgtgagct tgtctcagagttttgaccacactagctgacc gatgcgagtagctcgtgtactcgttatcggt aaggttttaagttcctacgtgagatgcattcg ccagtttcaagctcagaactgggtgattaaa caggtttatattgagttaacactcctgttggc cagatagtagctccgataacttggcattag cgcagcgttagcagtagctcgcagcgtcggc taacatcaccagtaattgagggaatttatgt ccaaa
42-43	wastewater metagenome genome assembly, contig: NODE_3741_length_625 8_cov_6.828954, whole genome shotgun sequence GENOME_ACESSION: OVQ001.1 GENOME_ID: 11626 CONTIG_ID: 3740 SOURCE: WGS	ttggcgaagggaagtaacttaaaagattttgacttaggtacatatgttttaagagtttatct ctacttcaaaaaggagttttatgattaaaatacaaatgttccagtttaagttgtagtga aagccattaatgcttacaacacaagcatatgctagaataattacttaaaagaaataaagcaa aagtgttgatactgaaatttctactataaacttttgaagaaaataaacctagtgctaaaa aacacgaagtaacattaggttagatgttggattttctaataattggaatttcagttcgtaatga cgagcaagagttcattgggtgggacgcttgaatgttagaaggaatgtctgaacgcattact aagaaagcgatgcatcggaagcaacgaaatagcagaaaaagatactgtgctgtagatt tgataatagaaaaaagaacgtgtgtgttggctccaacttcaacataaaaacaataca catgataagataattgagaagcttgaatcttttctgtactgttaaaagtgtttagattggt aatttcgatactaaaaaatgtgtgctccaaacatgattgtattaattacaagaagcatac ttagaatataaaggcaatctacgtaattcaatttttctgtgattatactgtgcaattgtgt aataagaatttttaaggatgctaatgtttcattgcaagttcatcatatttttgaatcaaaa gggtgtacaataaaacgaagtaatttaataactttgtgtacaaaatgtcacactcaagcaaa tcataaagaaggtggagttatgataatgagatgataaagcgaagaaagtcacaaaag cttaaaagaagcaactcatgataatttttagctctagattaaaggaagaaacatgatttata	gttaagtttctactaccctaaagggtgagag cttattgagtttttgataagttacttaaatagct taagcatatcgaatatatgctactgttactaaa gaatatataggtacttcagggtttgtatcagc acaaactagatttgactctagctcgaactcta cgggtataaagcctaactctcagcagaagag gagaagtgcatatatacaaaaacatttagat aacattggcgaagggaagtaacttaaaagat ttgtacttaggtacatagttttaagagatttta tctctacttcaaaaaggagttttatgattaaa atacaaat

SEQ ID NO:	Name	IscB sequence	ωRNAs
		caggtttttgatatcaaacaggtgttaaacgaaagaatctaggattggaaaaatcacatag taatgatgcttttgaattagtgtgggacacgcaaatagcagcaccacatcaaaaaatt tcaagcaaaaacgtaggcataaacgttagctttagaaactttcaagacgcaaaatattga tattagaagcgcggaaaagcgtagacgcatcagtttaattgtggcgtagaagaagaaa taaaaatttaaacactgagaacttaagagtatttcgaggaaagaaatgaggaaagggtga atctctagaagattaaaggacttataaattgaatcgtgtgataaggttagatataaaggaaa ctttatactgttcgtgtaatcagaaccaaggcactagaatgaaatttaaaagaaacagggtga taagacatatttggftaaagatgtgttttgcataaagcaaatgtaggaaacatatgtcgacta g	
44-45	Ga0207421_10008055 GENOME_ACESSION: SodLak8KL_SPAdes_\$F _3300027784 GENOME_ID: 35608 CONTIG_ID: 8054 SOURCE: JGI	atgcaacgagttttgtacaagggttaccgggggaacgactgatccgtgccatccagcc cgtgctcgacaactacttcggtctggggggggcgggtaattcgtcccgcccttcacc atcgagatgatcgaccgagccgaggggcacggtgcagccggtgcgcttgaagattgacc ccggcgacgaggtcaccggcattgccctagtgtccgagggcgaagaccggagccggg ttgttggccggggaaactggtcaccgtagtcaggctatccgcaaggctctcggcgacc ggcgctctatcggcgacgacggggggggcgcaagtccgcaatcgccccacgttt caacaaccgagtcgggtaccgggtgtgtccgagccgtcccaaccgggtagac accaccgggacatgggttggcgactactatccgggttccgggtacccgtgtggtgtg gagacggtgcgcttcgatgtgcgtggtccggcgggcgccgctctccagcgtgtg agtaccagcaagcaccctgcacggcgtcgagctacgtgaatacctacttcaacggat ggctatgctgtctattgccgagggcgagccatgctgtgttggaggttgaccatgt gcaaccgagcagccggggcgagcaaccgaaccgcaatctggttaaccagtgcac cacctgtatcaggcaaaaaaacaccgactgccgaggagtggggcgggcgactggct ggcagttactccagactgaccgacccggggcagaacggggggggaagatccaagcc ggatggagcccgggctgcgggatgccggcgtatgaacgccagccgctatgccatcg gacgcgcgttaaaagaaccggcctgccgtgacgttcgccagcgggcgggcgacca aacacaaccgcagtaccagcattatccgaaggcacactgattgacgcggtgctgc ggcgaatccggcgaaaaggtgaaactggacccaaaaacccgaccttcatatcgaag cccagagtcgcgggcagcgctagtctccgggtggtatgggttgggttccccgcacg gctcccggcggtgaagcgggtgcatgctttcaaacggcgacgtatgctgctctaaa ccagccccggggaagatcagagccagcacaccggggcgctagccgggtattcgtgc ccgggctctatggatcttcgacacttcttgacagaaatcagtgccctgcaccaacat atgcgtctgctccggcgatcgacgggtattgctacggaggacgggcatgctag	gtcgccgttttagccctgccatccccggcct ccgtgcccggatgatcaggcgagactcaacc agcccagagcctctttgaggggtacggcc gacaggaatagatagtcactccgggtgctt ctccagccccggaccatgagcggtgtgt catgctgcccgttgggtgaaggcggaag gctccgggtgaaaaccctgtcgatcatgg gagggagacttgaccggccctgcggggc cgagactgagaggggtaacgaaatgcaac ga
46-47	k141_555391 flag=1 multi=18.2163 len=111838 GENOME_ACESSION: mgm4781941.3 GENOME_ID: 238320 CONTIG_ID: 63078 SOURCE: MG-RAST	atggtcgtatttgcataaaacatcatggcggaagcactaatgccgtgtcagaagaaaag ctcgcttatttattagagataaaaaagggaattgggttatttcaggattttctgaaaaa cgcaatgttatgtgaaaaacattttgggtattatgtcagatgttcgcaaaatacaaaatggt agggtcaaaaaactaaaccgattagcgcacaacaataattggatcacgcaaaagaacac cagcgtacgctggctaa	gtcaataaccaccacttaaaatattttatga ttttgaagtgggggcttcaaaagcttgatt gactagcatcagttcttcgagaactcgttcg ttttgtatcataccgttggtgattcccaagc cgtcggctctatgctggctctgtaaaatcct gagaggttaaggacggtcaaccagatgtga cagtcacactgttaagcgttcgaacatttg cgaagggaacaactccaaaggaggaaa tacgatattggtc
48-49	Candidatus Anoxychlamydiales bacterium isolate K940_chlam_1 K940chlam1_4, whole genome shotgun sequence GENOME_ACESSION: GCA_011064745.1_AS M1106474v1_genomic GENOME_ID: 327239 CONTIG_ID: 90 SOURCE: NCBI_Prokaryotes	atgttagcctttgttttaacaaattaaaaaacctttgatgccttctctcggctaaggctaa aagggtactaaaaaagggttggccaaagtcatacaaaaaacatttacaataaagtgc tttttggctctagcggacataaaagaagtatttagtgaatggatcgggttcaaaaacc attggcatagcagctattccaatggaaaaattttgatcaagtcacaaacaaactaagag gcgaagaaattaaaaaagatggatcaaaagaaggatgtacagaaggataggagga gcagaaaacttcgttatcgaacacctagatttttaataagagcaagtacagctattaat aggctcgtcctagtgttaacataaactctctctcatcaagagaaaaaaatttatcga atcgattctcctgtaagtgtggtgttagaaactgcaagttttgacatccacaaaatcac aaatccgaaagggtttcaaaagcttttagttaaaggccgactatcaaaaggcgaggat gttagattttataatgtaaaacaatcgttctcaatcagataagatcaatccaagtgtgt aaaaagaaaaataatcttaattgcattgcatcattcaatttagatcaaacggcggtct aattcacctgacaactagtaattctttgaaactgtcatgacaaactccataagctaaaa aaagaagaagctgaaaaatcttcaaaaaatacaaaaaagcgcacaaaaacagactaa acatgcaacagagctctcaattctcagatcacaactttgtaaacattttaaagctgaaag ttctcaagtatttgaagaaacttttggctatatatacaaaatttaacagagagagctctcttc ccaaatcgcatatataagtgcaatctgtatcgctagccgggggaagatacctgagatgca tattcaaaataacacctcagatctttttagaagatgtgttcaaaaggcgattataacaa agaagaggaaattgttcggaactgaaaatacaaccggaaaactttttgattaaaaaatt tgatctgttaagactccaaggaggtgtgattgttaaggggaaaaagagctccggttctt tgccatttcagatacaatggaaactctatctgatagttaaacatttaagaaaaacatagg ccggatacaagctaggaaaggcgttttaacttgagggtcgcaattctccctgacctaaag gacagggttcttcgagaaaaaagatga	gtcaatccccctccataaggaaaggcgtt gaaccgtgaggattaaaggcaactggttga ccagaagacaaactcaaggatttttattatg tttgaaccgttaagacagagaaagacaac agaccaacgagtgccacctcagctgtgtgct ctctgctgtgcaattaaacaaagttcaaa ctctgagagggaggactgacgatctattg gttctccaaccggctgagctgtgagcata accggttgttgtaccgggaaggctttatgtt a
50-51	Ga0172380_10022713 GENOME_ACESSION:	atgtcaagtaagaaattgctttgtgtagatattaaagtaagaaattgacacctacacgtg ataataatgcatgtatttaattagaagaaggcaggcgctaaattattacaaaaattccctatgg	aacagtaatttcttaagtgacacatagcttt agaaattagttgttgatgcttctagcctaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
	Leawe162metaG_2_\$F_3 300014205 GENOME_ID: 31497 CONTIG_ID: 22712 SOURCE: JGI	ttattgaattacaagagaaataccaaaagaacaattagataattcagaataatttttggtat agacgatggttctaaacatgctgggtatttctatttcagaaatgtaataaagactaaaact atatttaaggaacttttagaacaagaagaatgactgtaccaaatfaataagtttaagaagatct tatagacggcatagaagaggggcaaaaagataccgctccagttagggttaataatagaagtt catctaaaagaaaaggtagaatacctccatctatttaagcaaaagaagatgctatttttaaga gtagtaaaataaattaaattcattacttccctctatataataaatttttagaagatgtagcaata gataatcgtgtattacaagatggtaaaaagttatatggaggccaataaaaatttctaataga ttagacgagaatattcgtaaaagctgtattgattagagataaaaattctgtatgaactgtgta gatctaattgtaaaactgaaatccatcatataaattcctaaaagaatgcaaggaaataacta tggataatcttataatctctgtgagaaatgtcacaagaaggtaacaggacaagatttaga ttataataaagtttaacaataatcaaaaggttaaaacattcgtttgattatgctcaacatgtt atgcaaggaaagacttatttaagaataatttaagagatatagcagaagtaacttaactac tggaggagatacggcttaacaaaagaataaatttaaatattgaaaaatctcattctaagatg cagtagtaattgctggaggttaagaaattacatttatgattgggttataaacctctacgaa aaaaatcaaaaagtaaaagggtatataaattatggttttagatgtagacattgtaaaatata ctaaaaaaatggagagtcataaaaggctataacacaaagcttgcacaaaaagaataac ttgcaattattactacattaatggtatgcaattaataagatggaataaaaagattgatctga tagatcgtccaaaaaatgtaatatgatttaa	gtgagtgtaaaaatcgaagatttttagcac gatgaactcgttagcggcaaaaggcaaaag acacacctttagatgtatccagcagcttaag ctctgtgatttaggaaagaacaaftttaa gtcctgaagttaatcaccaaaaatacatgtc ctcccgctgacattggcaagcaggaaaattc ctaataagtggtgacagagatgtcaagtaa gaa
52-53	Ga0310695_10007302 GENOME_ACESSION: RumRJGDNav2_\$F_330 0032007 GENOME_ID: 280889 CONTIG_ID: 7301 SOURCE: JGI	atgggttatgtattagataagtaggaaacccctctatgcccaacaaacgggtatgggaagg taagacatcttctaaaggatggcaaggcaggtgtgtcagacgtgacccgtttacgatacag gcttaccatagacacggggaaacacacacagcccgtaagccctggcgttgatgcaggaa gtaagcatatcgctctatctgctacaactgaaaggaaagagctgctttctgcacaagtaga tttaagacaggatctcagtaagttgctcatgctaggagagaaacacggcgtagcagac gtagcagaagacacgttacagggaagccacgtttcagaaccgtgctcatagcaagcag aaaggatggcttccccatctgttcaggcaaaagtgcatcacacgttaaaagctgtaaaag gatgtatgtagatactccctgtaacaactataacgattgagatggcacccttttgatacgcg gaaactgaaggcagacatccttggtatgaagactcctcgggaactgactatcagcatgg ggaagcagaaggccttgataatataaagccttatgttaaatggcgtgatggttataagtg ccgtatgctggcagaaacacgttcaactacaggtacaccacaaaagcaacgcaaggat ggcgggtacggatagccggcaaatctataaccgtctgtgctgactgccacaaggcatac cagcaggaaactcttactggcagaaatccgaggtcatgagacctgacacaagataaa gactatgcaggatgcttgcctatggcctatgagatggcgtgctggaacaggctcaa agcactcgggtatcccgtccacatgacgtacggttataaagcggctgaaagcgtaagc aatgtgatttgcgaaagatcaccgattgatgcaagggtgataaaggcgtaccctgacgt agaacccgctggtgaatggcttctgttaagaaggtgcatgcccacaacaggcaaataca taaggctcaagacactaaaggcgacatccgtaaacgcaatcaggcagaacacgagatt aaaggcttcagactgttcgataaagtc aaatgtaacagtagggatgctttatcttcggaag acgttctgccgatacatggtatgctgacgcttgatggaacaaaataaattgaggtatc agttataaaaaactcaaatgtttaaccctgcaaaacatttactatagaaggagggtgct ctctcccgcgactaa	gtcaatgacctactaagtcacgagctt gtaatcaaaagctcattgactagcctaagc cttgaataaaggctacgttagatgttatca cacccgtggcgctagcacctaacctacggc tctgtcgccggcctgttaaacagcggtgaga gggtaaaccgcagtcacccaagatgacc gagtagcggaaagcatttctaactggtgaa gggtgacaactgatatggtttatcatctgaa attcagaagttaagactgtatgaaggcttact gtttaagtacagtagaaaggtaacttatg
54-55	contig_1148181 GENOME_ACESSION: mgm4547164.3 GENOME_ID: 237431 CONTIG_ID: 176019 SOURCE: MG-RAST DATE: 2019-01-27.	atgaaacattcagctgttcttgaatagccatggacggggacaaggctgatgccatgcct gaatccaagaagaagtagctgattgctgaaaagcaggaaaagcgttcattgcaaaatgat accgggttcagctcattacatataaagcactaatgaagtacagccgggtgagctgtc cgtggatacaggtgaacaacatattggcatalccatgaaatcagaaaaacatgaattcgtt catgaacaagagacgtcgttgcggatgagaatcacaccacgatgaccagaggcgct atcgcaaggcgaagcgtaaccggaaacgttacaggaaagccgcttttgataacaggcg cattccaaaaggctggctggcaccgagccgtgaccacaaaaagaaacacatccgg ctttgcgaaaaatagctgctgtagcaccgggtcagccatctgctggaagccggaca attcgatacaacggcgtgctgacattgacggaacaaggcgtgccagcgctatggtgacgg attaccagcagggaccgctgggttgctatgataacctcgggaagccgttttacaggg atggtcatacctgccaggtatgctggttcgactatcggaagattaaagaccaaggaaaggt ataagtcaggtacggtcattctcgaatgcaccatattgggtataggaccggtgaccatac agacaggatgagcaatctcctgactgttgcaccaggtgtcacaccctggcaaatcaca gcctggcggcgctttatgacttaagccaaaagccaaaactctgaaaggcgccgctt catgaacacggctcaggtggtacatagtttaattgcttttaacctaa	gtcaataacccccactaaatcgtcggcg attttgagggggcgtgacaggaattcaatga attgcctttgggtgaatcattgacctgtt aagcccgattgattagcccttctgtgctg caacagtacggaagagagcttaccgttaa ataacatagtcacgccagggtgcttcacaag ctctggccactgagcctgttacattaaacatc cctgagggttaggggaagtgtgtacagggt tacgtgaagctgttctccgggacatggctta tgtatataaaaactcgggataacaaggcgct tgtgaaccacgtccctacgggtagatgct gtacatgctctttatgtacagaaggcgtaagc caatttttcagggaaggagcagccatg aaacattcagctgtt
56-57	contig_421708 GENOME_ACESSION: mgm4547164.3 GENOME_ID: 237431 CONTIG_ID: 76265 SOURCE: MG-RAST	atgacaacaatgtttatattatgaacaagaatggaacaccccttaattgccaacagaacgac acggaaaagtcgaaaacttttaaacgaaaataaagcgattgttgaagaaaattcctttc acaataaaattgctttatgacaaacggaatttacacaacccatcagctcgggtattgacac aggtgctaaacatattggtatttctgcacacagagaatcaaaagaattatagtcaggagaa cataatctctgtgtaggaaaaaagttgtaacatcattatgacaaaagagaagaacattaa gacgagcaagcgcacagaaaaactcgttatagaaaagccgatttgaataatcgcataa agaaaacctacagatggattcgacaaatggttacccaacaacacgcacacagattgct gggtcatgaacatggtatataagaagtatacaagattctccaataacacatcatagttgaa actgctcgtttgatacgaacattttaagaatccagaatcaaggaattgaatccaaca aggagaaatgagtgattggcagctaaacttaagagaatatatacttgcgagagacaattac	ataaaactcaacaatagattcaaacacat attatttaacaataatcataaagattatacaaaa caagactaagttccaataaaggaaactacgtt atgtgagaatataatgtacctgtcagatagaa gccaagcttcagccttaaggctttagtttaa cagctcatttaggtaggagacaggttaagaa catataaaacctcaataaactgtgcatgg ctaccattacgagttatgcatctcgcttaca gcaataaataatgataataacaagaaa ggaatgacaacaata

SEQ ID NO:	Name	IscB sequence	ωRNAs
		acatgtcaatggtgcaaaaaagtctatcccaacatgatttgggttccaaacgcatcatata cagtttagaagtaattggaggttctaataagaccggacaacctcataacgctgtgttggactg tcatcaaaaactccacaccattacaagaactggaaaaattcctattgatttaagaaat cgccaaatcttaaatgtgcagcatattcatctattatgaaatattgaatctttaatatggcaaa aaagtatgataaaaacgcaaacatgacatttgggtataaaaacagaacacgaatcga cacaaatcgaaaattggatfacatctccgaacaacattatattgatcaagatgtataac cggaaatcctacagcaaaacctctgggtcaaatatacgttcagaacaaagacgatgcat aatcggctcctcttgacacagatcattaggatacctaaaaaaccaaatgaaaaagcact tatcaaaaataactcatatttcagaccccaaatagtcacggatattcctggatttagaga cggcgatataatagaagctgataactctatacgtgtaaaaagacgcggagatttaaaa acgcgtgtgacactcggatgtattcgttgaatagcgtgacactacactagaacaatct cctcaacaaagtaaagctctgtgcaagaataaagatagaattgtaattctaaaaa tattggaggataa	
58-59	Ga0255336_100223 GENOME_ACESSION: MeOH1_35cm_T3_17_2 _SF_3300028024 GENOME_ID: 32475 CONTIG_ID: 222 SOURCE: JGI	atggcagattatcgtactgtatgcagcgcggcagacattatgcccgtgtcggagaacgg gctcggctgatgcttgagcgtcagcagcagcgcgttcacgctgatgccgtttgtcattc gtctaccgaccggctcctcagggaaagcattgcagccggtacgcattagattgaat cggggagcaagataacgggctcgcactgggtgcagaggatggcgagtcgaccgccg ttcttaacctgtttgaattgatactcggggctgtcagatgcagggcctatctgcccg agcaacttccggcagcagcgcggcgctaacctcgggtaccgggccccacactatg ccaatcccgcacgaacaagtgttgggtcccccagcgttgagacatcggatcgacacg acgaaatcatgggtgacgcgttttaccggtgggacacagtgagcggactgttagcgaa ctggtgcgtacacctcgcagcggcagaagggtgacgagagacccctcgacaataag gagcatgctgaaatcgccgccaggagatcgcgctatctcttgataagtggggacg ccaatgcgtttatgcataccgatcgccaacatctgcagattgacccgatcgtcttctgta cggcagcgcagccggcgaaggcacaatccggtttgatcctgacctgtacttctgcatcg cgcggcgagctggacgtgacattggcgagttcgtaccgatcggcgccggctcgagcg gctcggaactggacaaaacgcccgcagagagatcgcgcgctgtagatggcgcgcg ttggccatcgcgcaattctcgaacgacggggctccctgtcagctgtccagtggcg gacgcacgaaatggaaccgtacacgcctgtcttcccagtcgagcgtctggtatcg gtatgtgtggcgagtcacggctgtcgtcgtgctgtattccgtcatcaagataaagt gtacggcgaggccgataccagcgaacactgtgaacgcgtatggcttccccgtgcc cacctgatcgcaacaagcgcatacataacgttcagaccgggtgatcggttcgcgccac cgtcccgaccggcaagaaagctggtatccatacggggcgagtgccggtccgttcgcgc ggctattttgatgtgcacaagccagacgggcccgttactgggattcaccatcgatactcg tcgttttgacggcggtgatggatacagttattcgtatga	gtcaatcgctccggcctaacggcgagggct tgcaaaagcctgaggttgaccagccagagc cctgcaaggggctacgttgtagagaattca agaccacacccgcgagcttctcagttgcg gacactggaacctgcggcagtagacacgct tcgagcaagcacgaaacgggtcgccgggt aacgctgctgtacaacattggcgaggggag cggagcccgaaatgctcccgtacaaggtt ccgtaagggacaattgagaacatggca
60-61	Okeania sp. SIO315 315_NODE_11, whole genome shotgun sequence GENOME_ACESSION: GCA_010672085.1_AS M1067208v1_genomic GENOME_ID: 325455 CONTIG_ID: 12 SOURCE: NCBI_Prokaryotes	atgtccaccaattatgtattgtactgtatgcaataaaaaagcctttagcacatgcaaggca ggaatggcaaggtcattatfaaaagcaggtgaagccaaggtttttaggcgctatccattca gcataattctcaataaattgtagcagaaaaacatcaggtattacagctaaaaattgacct gggcttaagcaaacagggttttctttagtaactcaagaaggggaaattttggcgatgg tcttaatccatcgaggtcaacaataaaaaacgctatgcttagcgcgcaaatcttcgcgcg aggtcgccgccgtcgaacaaacccgtatcgacaaccgcgttttcttaacagaagcgtaa aaaagatggctgcccacaaagtttaagtcacagagttttaactgtagaagcttgggttaac aaattatgctttagcaccatatttccctgaatggaattgtaaaagtcgatacccaa aaagttgaaaatcccgaatctctggaattgaataccagcaaggaactttatcgggacag atcttcgcgaatattgtcttaaaagttaacgcaaatgtttattggcgcaaaaagatac tcgcttagaattgagcatacaactccccgttcaaaaggtggaagcaatagggtgtcgaat ctagtattgtctgcatcaatgcaacaagataagggtgtatgataatagaaattttga aggataaacctctcttttgaacgagttttgaacaagccaaaacgccattaaaagatgc agcagcagttaatgcaactcgttggaaaatcttgaactttaaaaaagacagaactaccc gtgttaccgggagtggtgcgaacaaagtaacatcgtcgtcgtattggaattacaaaag aacattggatagatgcagcgtcgttgggaagtagaaaaattacaattcttacggctca accattgattgtacagcaatgggacacggttcagggcagatggtacaaatgataaataat gggtttccccgcaaagggtataaggtacaaaacgttccctggttggaaaactggagata ttatcaatgtcgtaaaaggccaaaatttgactaaaaggagtcaggattaaaactgtaag aagtaaggaattttgatatccgacatcaagatgaaattttgtctgtttctgaaaccatac caacccgttcacagcagagatgatacaattactgttttga	ataaagactaaccacaaatctaatcaaaacg cttatttgtgttaaatggttagtcagcctact ttttaaggtaaacgttattttggtcacaatcag tcggaatgcgcggccagttccagcctctatt gctcgccattaaacaggcaagagacttga aaagctagtgtgtcgagcctaacaagccaa aataacccggcgaggccaactttacatttcta taggagggacgcaaatgtccaccaattat
62-63	JGI11783J13700_101611 3 GENOME_ACESSION: ElksIonscriptome_47_SF _3300001214 GENOME_ID: 25904 CONTIG_ID: 483 SOURCE: JGI	atgttacgagtaccagttctatcaaatgcaggcaaaccttaattgccgactaaacctagtc gtgctaaccagacaccagagacacacgctgttgatgtgtgagcctagcggcggttgagt ttattcaatggcgggagcgtatatttaagaatgccaagcatggcaactggagggtgtgtg ttcatataacaccagcggctttacggtaatccgcagaccacttaaacggccggcaggt gcacattttgttccatccagggtggcagcagcaaaagtatggcgccacggttactcgt tcatagcttcagaaaaggagacaagggtgtgactgtaaaaagccggaaaaacctatcgg gctggtgttctggagacacaaaacacaggtttctgttttaacggaaactggaacgaat tggcgagtttactgccaaaaagtcaggtgtgtgcagcgaagcacgggttaaatcgtcgtg ccttcaactggattgtcaaatcttaccactcaagtgtgaagatttga	ttaataacccccacttacttctgtgaagtgg ggattgccggacaacggactgttaaac gttgaataccacatagagctcgtttgttac aaactcggatgctcctatgctcggattac atttaagcctttttgtcagtcgtgttaagt aggacatcttaaccgagatgggtgggaagg acatagtgctagtgctagtgctgattatc agtggctagtaattgctgataatgataacc

SEQ ID NO:	Name	IscB sequence	ωRNAs
			gatacctgtgaactgataactgataactaac tcgtgaggtttatcacccatgttacga
64-65	Symploca sp. SIO2C1 2C1_NODE_137, whole genome shotgun sequence GENOME_ACESSION: GCA_010692635.1_AS M1069263v1_genomic GENOME_ID: 325548 CONTIG_ID: 42 SOURCE: NCBI_Prokaryotes	atgcgagtttctgttagacaaaaaccattatctcttgaccctgccaccagcacggg cgagagaactactacactcaaggagggttaagatatacagacgtatccgtttactattgtt ttgcaagacagaacgggtgaagaatctacaactcattctcatcggattaaaatgatcctgg tagcaagggtactgggttgcagttgtccaaagaagacagggacgtatcacaaacgcct tgaagtctctatagaggacaacaaatcaaaattctcttgagtctctagagctttaagga gaggtcgtcgttaaccgtaaaaccgttaccgcaagtctctgtttcttaaccgcactcgtaaa caaggctggctaccaccatcacttgagagccgaatttccaatttgaaacttgggtcagga gaattagaaaattgtgtccaatcactgcaatttctcaagagctagtcagggttgatttgcac agalttaaggtagcggcaaaaggacacggaaccagacaacgatgcctctgataagttt ggattcccaaaagctcatgctccaaaggctaagttttccaaggctttcaaaactggcgat cgttaaacccgatgttcaaaaaggtaagtttaccgggtcaatatgttggtcgaattgccattc gggtcagaccgagtttgcctgcactaccgaatcaaaagtgtgatgttaccaccaatac ctaaaaactatccacaggaatgacggctatgaataccaattctga	gtcaacaactcaccgttaagttctagegaact ataacggggagcttgaataaagctctagttg accagactaagggttgaataaaacctacgtt cagagtaagagttaaagttcctaccctaaat acgtgccagttttaagctctagaactagaaa gttaaacatctgtagttgagtttaaggaggtgc tttctagatgtaccgacttgaacattgtcag gctaactttaccagaaatgagagtcctcaag aatcgca
66-67	Burkholderiaceae bacterium 16 contig55, whole genome shotgun sequence GENOME_ACESSION: GCA_000955785.1_AS M95578v1_genomic GENOME_ID: 71753 CONTIG_ID: 209 SOURCE: NCBI_Prokaryotes	ttgcagacaagcggcagggtaagcacgaaacggatccgggcacatcgcgggttctcaa catgccgaggggagacctcccgaaggagggtgttccggcccgtaacaatgaggg ggattgccatggcagtcattgtgttgacagaagcagccggcgtgatccgtgcagc gaggagcggagcccgccagctctccgagcggcgtcgcgctgtctacaccgctgatg ccgttcgccatccgagatcgcggcggagcggactgcacatcccgccgatgcgcacaa actgatccaggctgcaatgtcactggcatggcgtcgtgcgcggcgtgatccaatg gcgctgcaccgagcggccatgccgtcgcgtattctctgttgagccgatccaccaggc atgcggatccgaccgaatctaagtgtgcgagggccatgcaccaaccgcgcgcagc tgcaagccccgttatcgcgcgctcacttcgacaatcagcggccggccgcaggtggct gggtccccgcctgcagcaccacatcgacacgacaatggcctgggtgcgcgcaccaagt acaaccggcggcgccctgatctgccaaagaccgcgccatcgatgcactgtgcattgg cgctgtcgtctatccagcattggcaggtaccgggtgcacccatcaaatgcgtgcacg cgactcggcaccgtcgggttccggcacggcgacccatgcgcagcaagcaggtga	gctgtcaacccctccctgacggatggagctc taaggcagctccgctacgttgcacacagg gcaagaccaacgttgggatgcttactcagt cccaacctccgcaagccccgggtgcagaca agcggcagggtaagcacgaaacggatccg ggcacatcgcgggttctcaacatgccagg ggagaccctccgggaaggaggtgtttccg ggcgttaacaatgaggggaggttgcattggc a
68-69	filamentous cyanobacterium ESFC-1 genomic scaffold A3MYDRAFT_scaffold1 .1, whole genome shotgun sequence GENOME_ACESSION: GCA_000380225.1_AS M38022v1_genomic GENOME_ID: 49861 CONTIG_ID: 0 SOURCE: NCBI_Prokaryotes	atgcaacgtatccagatataaaacccctgacggcacaaccgctatgccacaagcgcac ccgcgctgaaaaattgggtaaaacagggtaagcccagtggtgcaaaaccgacctccga atcaaaagctgtccgctacgcactgttaccgcacagcgttccgaaaaaggagattgttc gggctgaaatggcagacgcacttccatgggctatgtgagtggtgacaccgccgtcag gtgtctgtatctgatttcaattggaagcgaattggcaggttaccgcttcaaaagttcaattac tcgatcgcgcacgggcatttctgttacttgcgcgaagatattgttagtcaacggggcaac ccccccgactaacgctctatccctctcactcgtcgcgcgaggtgggggtatcaaaagg agttttttcatga	ttcaatgacctccaccgactacgcgctgtat gggggagcgtacggagataaccggctcgt agctgcgaataggctattcaggttagtgcgg caacacttccagccgttctctaggttggatt atctgtaagccctggaatctcaggaggtg gtatcaaacagacacaccgtactagctgg cttaagagacctgtattggggaattatcttc catgcaac
70-71	Erysipelotrichaceae bacterium SG0102 DNA, complete genome GENOME_ACESSION: GCA_003925875.1_AS M392587v1_genomic GENOME_ID: 208690 CONTIG_ID: 0 SOURCE: NCBI_Prokaryotes	atgagcgtatgcgtttagcgcgggataaaaaagccgctgatccgacaagcgagtaccg tgcaagaaagctgctgaaaaagcggaagcgtgttatcttaagtataagccgttcacgac atgcttacaagaatcgtaagcgagaatgacagccgatcagtagtctgctgcgatacggga tacaacataattgtgtatccatcaaatcagaaaaacagagatatttgcactgccaatttgac atgcttcaggatgagaagaagcggcacgatgaccgtcggaaaaatgcgcgcgcaaga agaacagggcttcgtacagaaagccccgttttgataatcgacagcttcaaaaaggaa ggatggctggcggcgtcgtgagaacatccgtgatcagcatatccgtattttttagcgggt tccttgaggctatgccgatagtctcagccaggttgaaatgggctcgtttgacgtccacgc atgcatgagtttaggcaacaggcaccgtgctttaaaggcgatgattatcagaaggggcc gcgatacggcatgaacacgctgagaaaagccgttttctatcgtgacaattacacatgcc gggtgtcgggggaaaccgcggatgaaggcgctattctagagtgcatcatatcggttcg aaacaggcgatcatcagaccgcatgagcaatctgctgacccgtctgcacaaaatgccat acttcgcaaccacaagccggcggaagctgtacgacctgaagcctaggacgaag ccgtttaaaggcgagccctcatgaacgctgtcagatggcagatgttcagaacgctgaaa agcaccaccctgattagaatggcagatgacatagcgctgctacgcaggaggcaag aagagtcctacaccttgaaaagtcgcatgccaatgacgcctacgcatgggagaaattcca tccaagacgcaggacgcctttatgcattttcagaagctgagacggaataaccgcatcctt gaaaaattctttgatcaaaataggtgatgcgcgagacggcaagacaaaagagggtgc agagctgtcatgcggacgcacagacagaagcgagtcgaagacactccgagaagaacctt cgcgtattcagagagcggaaagtttcgaaaggcagacgcgtgatcagaagaagccact ataaactgcgtcctggcgatactgtgttattggcgagagaagcataggcgaagggc gttcataacaaaggcacatatgtgtgacagacgccaagaagtcagtgctgttaagaaa gtagagaagattatcatgaggtgggtatagctgttaaaatag	gtcaactaccgcgactaaagtcgagcgctt gtctcgcgcaagctaaagactgcgcgtgta catatgtactagctgattagcctaattgtgttc gggcactacgttatccgcacaaaatcagggc acctggggattcttccacgtccctcgtctg cgggcatgcattaaacatctctgacgtgcag gagaaggtgtcagtcatacacaactcgg ataaccttggcgatgggaactaccttgaag ggagaactgcataatactgcctctggcaa gcgtatgcaggcgtgcgtaagcgacactt attcatcaatcttagctcagctgattaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
72-73	Ga0075125_10000791 GENOME_ACESSION: AntAceMinimDraft_18_\$ F_3300005935 GENOME_ID: 22974 CONTIG_ID: 790 SOURCE: JGI	atgttagtatatactgaataaacaaggtaaaccattaatgcctgtaaaccacaaaagc aaaacaccttctcaagaaaaaaagctagagttatccaagtagaacattatcatccaa ctattgtttggatcaagtaattatagacaaccaataaccattaggtgtgatccaggataccaa aatgttggtttatctgtatataccaataaaaaagaatataccgaagtgtatgcaattacga atagatataccacaaacttttaattggaacgcaaaatgtttagaaaaaatccaggaacagaa agactcggatagaaaaccacgttttgataaccgtagcaaaaaaacggtcggctgtac caagtatcaagcacaacttgatagtcataatcgggttagtaaatgcagtcgaaaaaatatta cctattactaaaattattattgaggttgcatcttttgattcgcagaaaatgcagaaacctgaga tatccggagttgaatatcaacaaggaaactctgcatggttatgcagtaagagaattattattg gaaaagtgagtagagagtggtctattgtgaaaaaggaatgtacccttgagatagaa cacattgtgccaagtcgaagggtgtacagatagatcaaatcttaccattatcatgtca cgaatgcaatcaagaaaagagcaatgatgactgcagaaaggtttggacatccagagattca gaaacaggcagaaaaatcgttgaaggctacagctttcatgaacattgtcgttgaaattg gtaaattattttgaactgtaagcatacataggttaataactaaacgcaacagaataaaaaat gatatgtcaaatctcatagtaattgatctttgttatagctgggggtactaaagatgtggatc gatcagatgttctcatcaacaaaagcaggtacgcagaaacaatcgcaaacgtgtcaag ggtcagagaggtgagatatacaaatgtccaagagaagtattcggattcagattgtttg ataaagtagatataataacaagaataattttgtttggggacgtaggaaaggaggtatctttt gcttaaaccttattctgtgtgataaaattgagagaacataaaaaagtacaaaaggatgtg gtcagggttctttctgtgtgagatccaatttcatccattcaaacgaagggtttttaagggtat aatcaatgaataa	ggfcaacagccctattctaataaacaatctt atgaagttgactaggaggcataatgcaga cgftaagaagaagaacacatacacacctta ccgtaccactctagcggcaagctctgtgacc atgcaattaaacagacctgttagggtaggggt tagtgttgcaaaagctctcaaacctcttaact ctccgaagagaccattactcggaaacagga ggtatattttattgttag
74-75	Ga0180109_1294567 GENOME_ACESSION: EasRivAnnotation_4_\$F_ 3300020067 GENOME_ID: 25739 CONTIG_ID: 523 SOURCE: JGI	atggcagtttctgattggacaagaggagcaaacgcctcctcaatcgaggaggccgc gggctctcttgcaccctccctgcgccaccgagtgatgacgacgatggcatgggtgcagc gtatctcccgatgggaccccgtagacgcgacgtgacggagctgggtgcgttttgataccc aggcgttgcagaaacccggacatcgaaggcagtgtaataccagcaagcgcaactcgccg gctatgagctacgcgagatctgctggaaaagtgggctgctgctgcctactcgggc gcgacggatgtgcccgtgcagatcgagcatatccaccaaaggccaaaggcggaagc aacgcgcggccaatctcacgctcgtgtgtgcccgtcaatacggcgaaaggggcga gggacgtgcacgagtttgcagggacgacacgcgcctgaaacgcctcctcgacaggc caaagcgccctgaaggatgcggcggtgtgaatgaacccgcgggatgctttgtcgtg cctgaaggccaccgcactaccggtagaatcgcgctcaggcgccggacgaagtggga atcgtctgcgtcggcattccaagaccatgccctgatcggtctgcgtcggcgcg ttgaagcggtgaaggatgagcagatcgccgcttggcaactcaaggccaccggacgcg gaaagtaccagcgacccgggtcacgaataacggtttccgcgcggctacatgatgcg gaaaagcggtgaagggtttccagacagcggaattctgtggtcgccgagtagcagcg gcaagaagactggaatctatactggccgcgttgccgtgcaaaaatccggcagcttcaac attcagaccgcatacggcgtcgtcgaaggcatcgccacaagcattgccggcggaattca gcggtctgatggctatgggtatttttacaacatgggcccgtacaaaaggagagcggga gggactttcg	gtcaatcaccggcctaaaggctggggct tgagggggcaaccccaaaagcccgagatt gaccagggaagcggtgccaatccgctg cgttgcacgaagtacagaccacatacag gatgcttctcagtcctgacactggaagcg gcagatgcagacaagggtccggtagctac gaaacggctcgtgtaaggcgaaacccaa gctgcgttgcaacatcccaggggagcggt tgctgcaaggcaccgtaacaaggcccgta aggcgattttagtgacaagaacgatggca
76-77	Ga0073583_1167572 GENOME_ACESSION: IMG_3300005253_\$F_3 300005253 GENOME_ID: 27780 CONTIG_ID: 10098 SOURCE: JGI	ttgaacgcgaagttaaagaccgacggtaggggtgcttcctcagccctgcccctcgaagct catgctgcagacaacccggatcaggaggaacgggtgtggggcgatgctgcccgttc aacatgtgtaggggagcccaacggaaccgtcaagccgttggcgtcacggcggaag cccagtgccgctgaagccgcgtatttgtgctgagcgaaggggttacccttgatgccttg tcatccagcccagcccgaagctactcaggaaacgttagagcgccgggtccatcggtg attccgtttacgatccggatcgtatgacgagacgtcgaagacagccaactgcagccagti cggctcaagttggatcccgttccaaggttaaccggcatcgatcgttctgtgaggacgggt gctgatcagcatgtattgcacctggccgagcttaccaccaggagcaggtcgttcgaag aaaatgcaacagcgtgcgatgtacgttagacgtaggcgttctgctaacctccggtataga cagcagcggtttgtaatagaagcatagacttaaaatcggaggaaaatggctaccgcc gtctatacatagttgctgaagtaataatacttaattgctgttaattacagagggtgtgtcc ggtttccggacttactgcagagaccgccaatcgcacaccagttgatgcagaccgccg agatctcaggcatagataatcagcagggcggaactccaaggctacgagatccggagttat ctgctggagaagtggggcaggatgtgtgctattgtgctggttaagacatcccgttgcat gtggaacacgtggtgtcgaagctaaggcgccgtccagccgtgtcagcaatttgacgtt ggcctgtgtcccgtcaaccaagctaaagggtgctcgggacgtccgtgagttcttgacagg ggcgccctgccagattgccatctactgtctaaagcgaagggctgctgcgaagtgcga gcagtaataaactctacacgtaatgcttgcctggtgcaattcgcctcggcggtttttgtga gacacagctacaggaagtcgtaccaataacatcagaaacgggttacgcgtcgcaaaaga cccatgcccttgacgtgctgtgtgtctagttagccgcttttgggtggggacagcg tgtattgacaataaaagctcaggccgaggaaaatacaaaagaccaccacgaccaagt atggcttccctcgaagctattcgtatgaggaagaaagatcatgggttcaatctgtgtgat ctgtacgcgcagtaatacacaaggaataatgtaggtatatacacaaggaagtgtatgtg gtccgctcgcggggcttttgcagtaagacaccagaaaaatgtgttattggcgct attgctcactactcagagatcgaatggctactctgtagctacagctaa	gctgcagtcagacatcttcttctggagaa agccccatgggctcatgctgaccagacaa gttcacagaaggctgacagccgcaaaagg aactacgtgaacgcgaagttaagaccga cggtagggtgcttccctcagccctgccttc gaagctcatgctgcagacaacccggatcag ggacgaaacgggtgtggcgccgatgctgc cgttcaacatggtcgaggggagc

SEQ ID NO:	Name	IscB sequence	ωRNAs
78-79	Ga0070741_10072506 GENOME_ACESSION: CoaSoiStandDraft_11_\$F_3300005529 GENOME_ID: 24556 CONTIG_ID: 72505 SOURCE: JGI	atgtcaaaagttttctgtggagataccaccgaacccgcttgatccggttcatccgggc cgcgctcggcttctgtcaagcaaggcaaggcagcggtttcggcgctcctccgttcg ctcctctgaaggtgcagggaacctatccgaagccacagccgctgcgagtcgaagatcga tcctgggagcgacgtccccggtctggccgtctggatgacgccacgggcaagggaatc tgtcggtggagttggggcaccgtgcaggcacgagtagtcacactggagagctgcg gagcaggccggcgctcgcggcggaacgctacaccgttaccgccaaggcaagatggc gcaatcggcgggccaggtatcgactcgcgaatgccagggtcagccatctcttgactg tctctggagagcagactcgcaatgtgctgacccgggtgcaacggctcgtcgggttcg ccgattctgccatcagtcaggaaagggtgacgtttgatatgcaactggtgagccatctg aaatcagcgggcgaagaggaccaccaggcgatgtacggggccttgaagtgtatgata tctgctcgagaagtggggcaaggcagggtcctctgcgcgacaaaacacgttccgtcc agattgagcatatgctgtccgcgccaacggcggaaccgaacgggtgagcaatcgtt cctggcctgtgagccctgcaatcaccctggcgaagaaggcggaggtgcttgagcgtg tctctggcagggccacgctccgttgaaggtatgcggcgccgctggatgccatgcagg ggcaccgtatcgccgctggtggcgttgggggtgcccgttgaggcagagactggcggg aggaccagatggaatcgaccacgcgccaggttcccgggatgcacgggtggaggca gcctgctcgggtgccagt	tcatgaacccctgtattcaatgcaggagcttg cagtagggcatcagctcccctgtaagccctc atcatgaccttgaaagagagcctgcgttttca ggaaagcgtttacgttccctcctgggtgcg cggccatagtgaacaggtcagggaggtg aaggcagtagtgagaagacagttaccgct gaaaaccttgcaggccaacgttaccgtg caagcggagtcctttcaggggcagaagt caagcatcatgtcaaaa
80-81	JGI12273J12029_100009 67 GENOME_ACESSION: SoiCnureplicaC0C_FD_\$ F_3300000833 GENOME_ID: 310246 CONTIG_ID: 106 SOURCE: JGI	gtgccgtagtaagtcaagacgcaaaactaatgcccaaaagcaaggcgcaa gaaaggggtgaaagtgtgtaagcaatcgcaaatgttcagacttggaggtctactatg tgcactgtctcaaccatcgaggagaagaacgcagcctattgtgttgagtagacccc ggtaaatcttattgtggaatcgaggttcaatccgtcaataacacctactaaggctgcact aatccttcccttggctgtgttcgagaaggaatgataaacggcgagagttgaggcggtt cgccggggaagaagaattaacgcgaagctaccacgcaagctacgcaatcaccacacgt gcagggttcgacaatcgtaaacagctctaatgttcccccttaacaggcatctcgcaact ggaactgagagtagttaaggaaatatacaaaatattcccaataatcacgataggttacgag cgagtcagagcgaatgtagacagaaccgaaggagaagaagccaagtcaggtaaaaggt tcttccgtgtatgtcaggtcaaaactgggctatcgaacaaatgggagctattgcacctg cttattcgggaaggatggcaaaagatggcaacgggtacatctcagattagaagcatct gggattggaaggacaagaaaaacaaaggcgagggttaagccagaaactcatgcagt agatggtgttgccttagcagcaggatattgttcaatacagccgtcacatccgtgtaaatat tcaggatatactcgaaaggcaat	ttcacgaacccctggctagtagttgagcca ggggattgcaggagataacctgtcatgaaa tctgtaaatcggctctcgttggcgagagaca aacacttccgaatgttccctaatcggatga aatgtacagacctggaatcttaggggtg gggtaaagcccagaaacaatcgcccaat ggcgtaaaggactataaactttagctctaaa ggattatctccatcata
82-83	MIS_10001517 GENOME_ACESSION: IMG_3300002026_\$F_3 300002026 GENOME_ID: 305123 CONTIG_ID: 1516 SOURCE: JGI	atgaataatgtgtttgtttgatacggggaagaagcccttaagtatgttagacctgctag ggctagacagttattgaaagaaggcaaggctgctgtatataggaaatatccatttacgata attttgaaggtggcaatgcctgaagctgtgtgaaagaggtattgttaattagaccccg aagtaaaacaacggggtgcttttagtatctgataatcgggtgtgttctgctgaattg agcatagggggcaatataataaaatagaatagtagtagacgtgctcattaaaggcgtaatc cagaagtcgtaagacaggttatcgtgctgcgaggttataataataggtgccgtaaaagctgga tggttaccccttagtttaaacacacgtgtgtaataactacgatgacgtgggttaataaattaga aagttgcatcagtagtgaattggcgatagagaggggttaattgtatgcagaagatggt aaatcctgaaataaggtgtattgtagtagtagcagagccacattgcaagggtatgaggttcga gaatattattagagaagtttaataagaacgtgtgtatattgtatgctaaggatgttcattgc aatgtgagcatatacatgtaaatcaaaagggtgctcaataaaatttctaatttaactttagc ctgtgaagcctgtaataagaagaagataattgataatattgtatttttaagaataagcct gagttattaaaaaataattgaaaaagggttaaaacaccttgaaggatgctgctgtgtaaat gctactagggaatgccttattaaagcacttttagatactggattacctgttgaacagggaact ggatcacagaccaagtaacgtacacatctaaattacaaaagagcattgtagatg gctgctgtgtgaggtattctgtgttcgaggttaaggtggatgctacttttaaaccttaaaa ataaagagtagtggtcacggtaataagacaatgtgtattactaataaacatgggtttcctatt aagcacagaactaggaaaaagtagattttgatttcaaacaggggatagtagtaggtg tagtttttaagggggttaatacgggtatccatgtgggtagggtatctgtaattaaagggt acttttctgttagaactactgtaggatafatgaactatctgtataagaagttcatcttactt caaaagaatgatggataggtataaataa	ataaagaataactcgttatatacaggggcg catcagtagatccgtgtcgaaggttcaattaa aaccttaattagtttaattacgaccagacta agtttaagtaactacgtacatttaatac acctatgaatgcttccagttttagctgtgtg ttaattatcaagatattatggtgaacataat gaaggtgattagcttagtaagttttatgacat tgtcaggaaaaacataaccaagaaagttg cgtaggagatttaaaacatgaataat
84-85	MIS_1002273 GENOME_ACESSION: IMG_3300001765_\$F_3 300001765 GENOME_ID: 305120 CONTIG_ID: 2272 SOURCE: JGI	gctgtctggaacggttaccggttaccatcatctgaaagatgtttaccaaattccccaaat tgcacccttgcgattaaaaatagatccaggcgctaaatttactggcttagcactgtagatg atgcaacaggagaagtagcttggcgccgcaactagaaacatagaggctttcaaatcgc gatgttttaaccaaacgcccgcaatgaagaagggaagacgccatcgaaaaactcggta tcgtcaaccacgcttgaataccggaacgcccacaggtggtcgtccaccacagctgca aagccgagttgagaatcgtacttgggtctgaggttacaaaagctagctttgatttctga tatttctcaggaaatagtcagtttgatactcaactcatgagagaatccagacatcagtgga cgataccaacaagggtgagtagctgttgcgaagtgccgagaatttctgctttcaattca atcatacctgtatgtactgtggggcctaagaatactgttggagattgagcatcttttacc agtctaaaggtggttcaatagaattagtaatttaggcatgtctgtgaggaattgcaacaaa aaaaaggccgtcaagctctcagagaatttttagcccaaaaacctgactgttacagcgc tctgagacaagtaaaacaacagaagccgacactgctgccgttaatttaccggttgggc attatttaacgactcaaaagactgga	gctatccgcgaaccactgacaaaactgga caatcccgaataatcctgttagagatagcc cggtaagtctaacacagactacgttaaga aggcacgacacctgtgaggtgtccgac ctcctgtctgtcgtaccaggttaaacatctt atttcgtaaggaggtgctgtagcatgaca agccctcttaacattaccaaggcaaaagnn nnnnnnnnnnnnnnnnnnnnnnnnnnnn nnnnnnnnnnnnnnnnnnnnnnnnnnnn

SEQ ID NO:	Name	IscB sequence	ωRNAs
86-87	Ga0377217_000700 GENOME_ACESSION: AD121II_FD_\$F_330003 6959 GENOME_ID: 312260 CONTIG_ID: 699 SOURCE: JGI	atgaccactaaacaattcagtttagttattgattcagagggaacacagtttaagccaactaat gccaataaggcatggtatttaattaggaacacaaagcctaaatattacaagatttccaat gggtattcaattgttcaagaagattattcctaaagcaattgataatcaaaattatctgtggca ttgatatcgcccaataaacacacgggcatcgcgattgttcaagaatgccaaactaaatctaa agtaattataaagggaacaattgagcatcgaaatgatgttataaaatattgaagagcga gcgagttatcgagataatcgtagaaaaataagagataatcgcccgaaagatttaacaata gggggttcaagcaaaaagaaggcgcatcgcccatctattagacaaaagaagagctct attctaaagagtaattaataggttaaaagagacatgatctatccataaatatatagtagaagat gttttaagataataaaaaattcaagagccaaatattttccaacggaataccaaaaatct aataaattagattctaatattcgaatgctgttatgatgcgtgataaatttaaatgccaaagt gtaaaagagggtatgctaaattagaagtcacatataacgccgaagagattagtgaggaa caatacaattgataatcctaataacgctttgtgtcgtattgtcataaacaacacatggaaaag aagagggtattataaaaaatattataatataaaaggaaaaatattattcaggggacg cttcgtcgtaattgcaaggcaaaacataattttagagtgtaattaaataaactagggatagtc gagttacaactggatcgcaaaccttttataaaagacatgttggaatattataatctcatt ctaacgacgccatagttattgtcaataaagagggttgcagaacaatgcaacataatcgat tggattatcaagccattaagaaggcaaaacagaataaaagaaatattgtggattg acacatcgaggattttgtatgttactactattagaacgaaatagtgcactggatatataacc gcaatgcctattggcaaaaaccaagtaaatattcaatctaagaataaaaaatggactaggg taagagccgaaagatgcttttaattaatcgtcctaataagaataatgtggagaattatgatta a	tcgaaatcttattaatattgtatttattcgggttc aagctcaagtgtgataatctgaaaggaa gttgacgatgaactacaattttacagaagg ctaaaaatacaccacgaatgaattcagttt gtagctctgtaagcacaaggatgaccgac acctacgacatgtgttcaaacaccgcaactg catagtccctgtaatgttagcagaagaagta aaacgctcgaagagaggatggttagaagatg accactaaacaattca
88-89	Ga0180007_10009944 GENOME_ACESSION: MM_PC_MetaG_2_\$F_3 300014656 GENOME_ID: 32273 CONTIG_ID: 9943 SOURCE: JGI	atggtattcgttatctcagagatcggaagcctttggatatgtgttccgagcgctgccca ggatcctgttagacaaggaggagagcggtgtcttttagacacttccctttgtaattgtctca aaaaatcctgtgctcttgcagccgaaactcgaatacaggctgaagatcgatcctggcg cgaaattgtctgtcttccatcctgagagagataacggacaggtgttatggcgccggg agatacagcatcgccgctggaagataaaaaagacatggacaggagacggaattgca ggagaggaggagaaggagcagaaaaacccgacataagagagccgaggtttgataatagaa caagaccgaaggcgctggctgctccgtcgtgatgcacagggcgttgaccgtttacacct ggctcaagaggctcatgaggtattgcccattagagcacaataatcagatcagctagcaagtt tgatacgcagaagtttgagaacctgacattgaagggctgtaatatcagcagggagagct tcagggtatgagctcgagagaaattttgtggagaagtacaacaggctgtgtatctactgc aaaaacccggcgctattccaacgttgagcataatgtacctatgagccggggagggtac ggacaggctaaaaaacctggttctgtcatgtgagaatgcaacatagagaagggaagca ggaccggccgaagaatactttaagtcttacgcaagcgcggaagagaggaaattagagcgg gagctgactgagctggagcatggcggttttcagctaaaggcgcgatccgtcctgagttt atgaaggcaatagccgtgaccaacagcacaagaataaactggctgaactcctgagatc aaccggaatacctctggaagtttagctatggatacgtcacgaaaaagaaccgccaggagc tgcaacttgagaaggcccatggatcgacgccgctgtgtgggaacacagaggcagcc agataaggaaatctattgacaagatcaaacgttgcataatgtcgaaggacgtggaac gcgccggagggtccagggtttacggccttataaactggatgataatggaagcctattata ccgaaggctaaaaaggatttccctgccgtagctcctgggatgccatgctcaaggcccaaa agcggaaatgagtttttgcgtcctaagctgtcgtactacgttaaggccaggactaagcaag ggggagaaaacaggagactcttctgtgaggaacccggaagctttactacagaa agggataagcaaaaatacagcgtaaattgtaagaattgcagcattatacagcggaatgat ggttacgaaattttaactaa	aagaattgtgaatattttcgcaaatcgattt ctgcggaaaaacccgctgtattgcgcaaat ccgtctgcaattgcataattgatggctacaga gttgtaaccagcctaaagacgctttataacggt ctacgttatcgccagagctaaagacgtacc ctggagctgtccagctccagcctctataat tcattgagttaaacaggcgtatgagtggaag ccagtgctcgtgagaagagctgaccgataa cattggcgaggcgaacgttaccgcaagaa cggagttcccttacgggaaaagggaacgca gatg
90-91	Ga0233412_10000478 GENOME_ACESSION: Na_anoxic_4_MG_2_\$F _3300023210 GENOME_ID: 33123 CONTIG_ID: 477 SOURCE: JGI	atgcagaagtttaattaaagttaaagaacacccaagagatgctccacaagtctttgctc tgtaaattcttcattaacaagatcaaaagcttagtgtgaagaataaagtactggaattaaa cagtcgccgatgtggatctacctcaacgtlaaagaggagaagaacttgagatctctctatg gtttacgtattaaataaacaggagaaacctttgagcctgttctccacagaaagcaaagaa actgattaaaggagaatttagcatgtatagttagaagaagtcatttactattcagttaaat cgactggagagaataagcaagaagtaacattaggaatagttctggagctaagaatata gggttttagacagtctcaaaaagaaagaattaataagtggaactgtcatttttagatgaca aatgaaatctagattagataatagaagaatgtatagaaaacataaaagaagtaaaacttgg atagaaaacctgattcaataatagatttctactaagaaaaagggttgctaccaccttct attttgagaaaatcagactcacttaactcttactataagattaaagattattacatttaa acaaagtaatagtagaagtaggttaattttgacatacaaaaaattaaagaatccagataattaa ggaaagagtatcaagaagggtatttattagatataacaacataaagcttctacatttttgcaa gagaaggtataaatgtcaactttgtgacaagctgttaattggtagaagactaatttacatc atatattccctaaacctgaaggaaactgacaagcttgataatttagcctatttgacaaaaagt gccataaaaacttcatacaaggttgcagggttaagttaaaaaagagtagacaataca aagaagcaacatttatgaattattgtaagaatgatttcaaaaagattttagattgtaagtagt tttggatatgaaactttactaaaaagaatgagttgatctgcctaaaaatcatgttaatgat gcttttgaatttctgttgagagaacagattaggactatcccttttgaggttcacaaaaa agaaaaacaatagattttacaaaagaatgaaaagggttttaaccttctatttaggaaca aaggctactctatcagaccctaaagattttagtaaaaatagaaggtaaacctttttaggttaagg	gtcaattaccctctttttaggaaaggcgcttg aacctgaggtttaacgtaagagttgattag ggagctgtattgaagaagtttaattatgcaga gttttaattaaagttaaagaacacaccaagg atgctccacaagctcttctgtgtaaaattctc attaaacaagatcaaaagctttagtgaaga ataaagttactgaaattaaacagctccgatgt ggatctacctaacgtaagaggagagaagaa cttgagagctctctatg

81

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ttaagttagccttctcaaggagcattacaacgatgcttattgcatagctgtaactttaatac aagtaggctttgcaatcatcatttaataagggttcatacctaggcatagtaggattgcatat gcagaaattcagtaaaaggcgtgtaagacgaagtgctagtgcttcttattggcttaactgtg gtaaaccttcaaaaaggcggagcaatgtttccatgtttgacaagggttaagttcaatggattg tttfttcattagtggaagtagtaattgttatgctgcattagagataaaattggaataagggt cacgggtgaagacaactgtactgttaataaattagcattgtttctcgaaaggcgtggcag catgttgggggaattatcgccgatag	
98-99	Ga0180007_10001273 GENOME_ACESSION: MM_PC_MetaG_2_\$F_3 300014656 GENOME_ID: 32273 CONTIG_ID: 1272 SOURCE: JGI DATE: 2019-01-27.	atgcaaaagtattatgctcgataataacatgaatccattgatgcctgtcatccagcacg agctaggcaattactcaaaacaggtaaaagccgctgtctatcgtcattatcctttcattattat ctgaagaatagaaaaggcggcgtgatgcacgccagtcgaattgcggatagaccccggca gcaaaagtaacgggctggcgggtgttagctcaatttggtagaggacgaaccgttctctggg cagctaatttacaacataaggcgtggagagtcagaaggctttagataaacgtcgtattctt agacggagtaggcgatcacgaaagttcgtatcggccggcaggttggaggaacagaa aaacggaaaagggtggcgtgtcaccctctctgtatgctgcggcgtgaataatatacggcati gggcccagaaaataacaacactaataccaataaaaactattgctgggtgagactatacgtttt gatatacaaatgatggaaaacccgaaatagtggtgtagaataatcaacaagggtgaattgc aggcgtatgagggtgcgcgagctattagaaaagggtggcgcgaagtgtttattgtag tgcagacaataaacagcttgagattgacatgtttggcccaagctcgtgagcgagta agggtgagcaacctgttatttctgtgagcctgtatagagcgaaagggtagcagctctgt acaggaattcttggcacatgatccaaaacgacttgagctcatattagcgagaagcgaaa gcccctacgagacgctgctattatcaatgctatttgtatcgtattgtgatgagcttaaaagag ttgggaatgctcgggtacgttttggagcgggggcttactaaataaatcgttgaatgctgg ttatccaaaagatcattggatagatgccgcttgtgtaggaaactcatgcagcgagatacta gaagggaatgctccgctgaataatcaaaagcaatgggttagagggaaacgacgggtgtgtca acctgataagtagcgttttctctaaagcgaaacccagaacgggtaaaacgggttcacggctt cagacagggcactttgtgaaagcggttgtgcccactaggacggaaagctgaggggacac atgttgcgtgagtgacaatacggcgagtgagataattttgtattagcaagatcgacggcatt aactggaaatattgtaattgttcagcaaaagtgatggctatgagtagcactcagatatga	ggaagtcgaaatgctcttaagttagatagct gaagtaacgttatcctcccacgtatgtattg gataaccagccttagtaattagttttcaaaaa ctaattactactgtacaagcgaatacatagcg acttcggaatgcttccagttccgaactctg cgggtgaaagttaataaactgtacgtgtg aacgggttagtctttcaacattaaaccgctg ctaactattggcgaggagacataaccgccc aaagcgagactgaggttaactcagatgcaa aaa
100-101	Ga0256407_10000103 GENOME_ACESSION: RumenRJG_04DNA_2_\$ F_3300028886 GENOME_ID: 280890 CONTIG_ID: 102 SOURCE: JGI DATE: 2019-06-13.	atggtttacgtattagacacagcgggtacacctttaatgcctacacagaggcattacagagt aagattcctcctcaaaaagaagggaagcaaccgttgtgtccgtatccgttcaccattaaac ttaccactgacaaagccaaagcattaccacaaacgtatcacttggattgactgtggaaccag acacatcgggtgttctgtctacaacggaaaccaagggtactttatcagcagagggcaactag cgtggggacactactaaactgattgccacaagaaggagctacgtaaaacacgcccga accgtttacgttacagaccgaaagattcaataaccgtataacatcagaaaaaaggatg gctagcgccttccattcttaaaaagtggtcattccatgccaaaagtcgttaagtacgtaagg aaacttaccgataaacctgataaatttggaaagttgcacgttcgatatccaacaaacttgcata tcctgatataaaaggggacggactatcagcagggggaacagggggctatataatacac gcgagtatgtactgtacagagaccatcacgaatgtcagcactgtcacggttaagtcgggtg ataaaatctgaatgtacatcattgaaagcagaagaacggcggttaatgcaccgaata atctgtttacgttatgtcatatcgtatgcataatcagggcgagatagagccttaac attaaacgggtgaatgcattgcgtgatgcagccgggtatgaatctaataaagaccgtctgt acgaggtgttaaaaggagaaaacgggtgacatcactgttcggtacacgtatgggtatataa cgaagtacaaccgtataaagtagcgttatagagaaaagctacacgaacgatgcacaagt ataagtggttaacttaatgcgaagtgtatgtgtatgtgggaggttaaggcaggtacgtag acataacagacagatacacaagttcaatatctgtaaagcgccgcaacttaacgtaatca gagtgatactgtgtaaaagggttatcgtttatggagcgtggttaagtaaaaaagtaacat atctgataaaggctaaacgaagcagtggttattcctgtgatgtttagatggagagga acgtgacgggtgtgggttaaaaacatttgaagaattgtactatattgaacagactgattaaat aaaaagtgaaaagagagggtgcaatttaccacgggttaa	gtcaataaccacccctaaaaggagtgggc ttgataggcgaggcaagcgtgccttactgaa aacaagcccttatttagccttgggtatag ggatttatcccttgaactccgttatccgtgaa tgcataaggcaccggtggatatactccaagt ctgccgctcgtcgggtcaggtgtaaacagtc ctgagaggtaggagcagtgctgtcagcaaa taaaaggaaccacgggaataacattggcgat gggagccttaccacacttaggttgagactt accctacaggggttattttaaaaacaaaa acaatg
102-103	Ga0373630_0001637 GENOME_ACESSION: YL18G_13C_MG_FD_\$ F_3300035179 GENOME_ID: 312754 CONTIG_ID: 1696 SOURCE: JGI DATE: 2020-04-07.	atggtggtgacaatcgaaacttgactagacggcccgctcgtcagtgagggttagcgg cgtggaaagtacaaagcgtacaggtgcaagacctagcctgtacctctcgttagaacc actaaatggcgtgctgccaagaattgtagcctaagccttccatgcaacgtcgaagggtca ctcaactcaaaactggggagcccgcaatcgggagagttccagctcgtcctccgagacagta cgtccttgatgccatgtaaacatcgaaagcacgtaagctcctagaacatggcctagcgg agaagcgatggagcaagcttggacaactcagccttattcgggctagaaaccgctgtg ctacgtgaaaggttcagcctaccaagaatccagtcaaaaggagcttacctgactacg catgctgtagacggcatagccatcggtccgctgagatcggaatgtgtaatccgtatccg ccagagttctgggttggaaacgcttcgagtagctcgcagtcgaacttcacaggcttgaac ctgacaaggcgggaattcgcagaaggtacgggtgcataatggagttaccaccattcaga aaaggagacgtcgtgttatggcgtggcaagcttcgagaggttgaggcttcatgtatg ctacgggafctcactacactcttccgggctgaagaacaaacgttcacacaagccccaat ccaaacgaatcgctcaggctattcaatcagcacgtattcagtagacgggagcagccccg attccttccaccatgaatggtgtgggtttccttgggggaccttga	gtcgattcgacactcttaaggaagtgtgaaa cgacatcgaggaatatgtccagctcgcacac aacgaacctgtccaatagggtgagggagt cgggttcacgtgatgaactccgaccggttaa actgtgttaacaagtcctagaggacacagt gttgacgccgaatccccgagcctccggtag cttcgatgtgtaattcggtaggaacccatg cctaaggcatgggctgtgtatgtgtgtgac aatcgaaacttgactagacggcccgctct gcgagtcgggttagcggcgtgaaagctaa caa

SEQ ID NO:	Name	IscB sequence	ωRNAs
104-105	human gut metagenome genome assembly, contig: NODE_298_length_7096_9_cov_7.505500, whole genome shotgun sequence GENOME_ACESSION: UPDC01.1 GENOME_ID: 16413 CONTIG_ID: 297 SOURCE: WGS DATE: 2019-01-27.	atgtacaggcttaccaccgaagcgggatcccggtgatccctgagagactacatgacac caaaaaaagaagagacatcgaggatgctgataatattatgtcatagccatggatgg agatccacagatgccgatcaaacgaaagagacatgctgaaaaatgttacagacaggaa gagcacgcgtctccggcaagtcccttatacaatacaattaacttataagaatcaaccagt ttacagccagttacgacgcggaagatcccgacgtcttaataatcggaatggcagtgctc ctctgagtgagagttactatccgcagcggtgtgacagaccaggaaacaggaaatgfta aattaatggagaacgtaagttgcatcgacagctccaggaaatggagagcgaaggca agacaacgactgctaaagagtgctatcagatgctcaaaagctggtttctgatcgaaaa cttccgatgatagaaaaacgaaacgagtgaaatgtctgtgatccgcaatacggaaagca agattctgtaatcgaaagcgagaagatggatggctgacaccaagtcgggaacatctggt gcggacacacatcaatctgacataagatgcagaaatctgctccgatcaccgatgttgc gatcgagatcaaccggttgccttctatcactggaagatccgagatcagtgagtcgattt ccagaatggcccgttaaaaggctatgatctgaatgcagcggttagaagatctacagga cggacactgtttaatgtgtaacagcctgatcgaacaccggcatcatatctaccacgaagt atgcagggtccaatcagatcggaacatcgacgattatgctgcaaatgcatgagaga gtccataaagatgcaagattgagggaccgtctgaagaaaaagaaagccggactggaca gcgttatgacggcgctgagtccttaaacaggcgatccggtttatctgcaaacgactggaa caggaatcgggaaagaacatgctcattactgttctggccgagatactccatggtcga aggctcttggatacataagacaaaaagagcagctacatgtggtcgtatgctggtgt atcgacgttcttctttacggcagatccagagaaagctccagatgctgatctgcatga gatcctgcaatttcgacagagatcgacggcgatcaagctcagacctctcgtgcatga ctatcatgaaggaaagaaatgtcaggaatcgcaaaaaagcggaaaggcaaggagaa gatcccttcaggaaatggagaaacgacaggtgatagatatggagaagagcagaccag aaagatgacatcagtgtaaaagtcagaaaaatgctccggcgttacaatcgtctggaccgt ctcatgccagagctgtcttttattacaagcgttcgatgtcatcgtggacaacattgc gaaggacatatctcaggcgggtggaatgggaagcaaggttttctgcaagaatgc aagatcgttctaacaatgaggacgtggtcttctgtgctataa	gtcaattaactcatgactaaagtacagagctt gcagaaaagtattatgattagtcgcgtta atgagcagaaacaagatgacagcgtctct ttgaaaagtgggacactctgggattctctagt ccagaccctgtcaagctgacaccaagtftaa gaagaccaaaatctgtcatgtacagccttac caccgaagcggggatcccggtgatccctga gagactacatgacacaaaaaagaaagg agacatcgaggatgtctgaatat
106-107	human gut metagenome genome assembly, contig: scaffold2386_25, whole genome shotgun sequence GENOME_ACESSION: UMBD01.1 GENOME_ID: 15693 CONTIG_ID: 2562 SOURCE: WGS DATE: 2019-01-27.	atggtgtacgtacaagacataaatgtaaacctatgatgcctacaacaaggcatgtaag gtaagacgactgcttaagaaaaaacaggcagttgttgaacctatgtccgtttaccatcaa attaacgtacgtacatctgattacaacaagaatgtgttagcggtgatgctgggacca aacacgtgtgtctatcagctacaacgaaagcgaagaaactttacagcagtgaaatgattct tagaaatgatatcgtatgattctttgtctaccagaagagagctacgaaaaacaagcggaa agattgagacatagaaaacatcgttttaacaatagaataaaaagtaggcgtcccggtatgg gtagcaccctcggtgaagtacaaaatagacgccatattcgtgtattgataatgtttttctat actgtctgtatccgtattgttattgaagtacgtcaattcgatactcaaaagattacaactct aatatatacaggtaagaataatcagggaggtgtatcaacttggatttggaaacgttagggaa atgttttagcaagatggacataaatgccagcattgcaagggaagtcgaaagaccaca gtattgaattgtcatcatatgtaacacgaaagacaggtggagattccccatctaatcttatta ccttatgtgaaactgtcataaagaataaccataaaggtaatatagatttaaaatc	ataaccaattgtattgtattatgcataatagcc aaaagctattccgattattagcctaagtgttga aacaacactacgttatttaagaatagatggt acctacggatatttaccgaagtctgtagctct aagtgaggtgattaaacagttctggatttcag gaacagtggtgcttacgaaaaaccttaataaa cattggcgatgggtactaacagagttttactc tgacttatgttgaataaacattaaaacgtttg tagatatg
108-109	Ga0315279_10002905 GENOME_ACESSION: YL17G06_20_MG_2_\$F_3300032070 GENOME_ID: 281630 CONTIG_ID: 2904 SOURCE: JGI DATE: 2019-06-13.	atggttccagtatttgaattcaaaaaacaacgcttaagccctgacagagaagcgtgcac ggaaattactggagaaaaagcaagcctctctgttctggaagaagggaatctttgtattaag ttgttcgagaaacctctgatagacagatcagagatgttgtgtaggtattgatgttgatcg aaaagagaggttataccgttactactgtcaaacatgatgtattgaatttccacgaataca cctgatgggttaaagacgcagtaaaaagtcgtagagaaaatgcgtagaggtcgcagattt cgtaagacaccatgtaggcaaaatagaactaacgctgctattatggagcttctgctgtgat gttagcaccttcgacaaggcagcagtggaacgaaatgaaggattctaaattgctatca acattgtatccgataactgatgtcgtttagaagatgtgaaagccaaaactgtaaaactg gagcttacgctgggtcaaaagaatggaatatgtcatcttaccattggaagtgggaaaaact tgggtgtactcggagataaagaagttctggaatttgaatttgggtcagggtatcgaaacaaa agaatggcgtgatattgcaaaagtataaaaagtagcaataaacataaggactgttggga atctcacaatgtatgtctatgttcttggatcggctctggaggtgtatcaagccattc aagaagatacttagactagatttctacacctatagaaggcaggttgcattgttctacaattt gccaaagggtgtatctgtaaatgtatgggtgaacgagaagcattggatttaagcgtggc agttgggttaaacatatcaagcattggttgcctatcgttggcacatcaaaagaaagag ttagctccacgagatgtcaacagggcaaacgtttgacgcaaacgcaaaagccagaggatt gtagattctgactatgcaacgtggagaatgtcatttttga	gtcaataacccacgcataaatcggggggc ttgagtaactgggaagccaaccagaccaatc gagacgaaactgttagtagactatgttttag aaaccaattgagctatctgatcgtggcaggt caaaagatgtcatagacgccacctatactta taaccaatctgatagccaatgtcgaagggtc gtttaacaaggaggtgtaaaccta
110--111	Ga0194137_10007470 GENOME_ACESSION: TAR4v2_2_\$F_3300018 411 GENOME_ID: 35909 CONTIG_ID: 7469 SOURCE: JGI DATE: 2019-01-27.	atgaaagtatatgtattaaataacatagtaagccattgatgccttcttccaagaaaggca cgattattattacgtgataaagaagctaaagtatgcaaaaaatccatttactatacaatta ctatatggatcaagtgatcggcaacaacatcattgggtatcgatcgggcagcaaaa actatcggttatcagcaaacccgaaacatgaactatattcggcagaagtcatttgc gcaatgatatcagtgataacatcactaccaaaagcaactaagaagaacacgtagaaatc gtttcgctatagaaaacagggttttaaatcagtgagtaataaaaaagaaaggttggcta cctcatcaacacacacaaatgatacgcataaaattatggttgaacagtcataaatt gttgcgaattagtaagattatagtaaacagcagcttttgatttcaaaagattgataatcc aaatattatgggaagccaatatcaacaaggaaacacatggttctggaacacacggg	gtcaactaccactgggctaataacccagtggt cttgataatatcaagccacagttgattaga ctaaggattatttaccacgtgtacccgtc atgatacctatgatgtacaatatagcctagt ccatagctctatctatatttataaacgatttca aagtggttaatggaatcagtgtaatatatgta aaaaagccatttcaacattgtcgaaggcctaat tacttacttcggtgaagaggccttattttag cccaatttaaggagaaattatgaaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
		agtatgttctttatcgtgatgtcatatcagatgtgtaaggggaaatcacaagataag atactgaatgtcatcatatgttatatcgtgcatggcggtactgataaacataaattga ttacattatgtaacacctgtcatagtcctaaaaaccataagaagggtgcaattatttgcaat ggtgggaaaaatgaaaaaaggtaattactcatataaagatgctacgtttatgtctataatgc gtgtgtcattttataattgggtgaaagataagtatatagatgtatcaatgacctatgggtatacc actaaaaataaagaattgaatcaattaccgaaagagcattatattgacgtattgtatta caggcaatattaatgctaagagattagatggctattggatataaagaaaatgcgtaagcat aatagatcattgtatatgttaaacactgaaagggtgtatatgaaaaagagacaagcta attattatgtaagggtctgtttatttgataaagacgttatgataatgttattgtttattatg ggacgtagggtactggatatttcaattaaaggactattatggaaatgaagtacataatag tgcgagtataagaagattgagtatgtaggacgtatgtgtgggtatatgtatgattatgtgtg acattag	
112-113	Ga0310696_10080563 GENOME_ACESSION: AusTroGonzalMGv2_2_ \$F_3300031993 GENOME_ID: 279930 CONTIG_ID: 80562 SOURCE: JGI DATE: 2019-06-13.	atgctggtatatgtccaagacaaggaaaggcgataagccccaccagagggtgcg gacgcgtcggtatctccgcctgcacggacttgcctgggtcatcatgctcgagccgttca caataaaaacgccgaggagaatgccgaattttagtctgcataggaatcgggcgggg gtgacacatcctctggaaaaatggtatataatgacccgaaagaacccgtccagaagt gcaaaaggcactctgcgaagagaaaacggctcaaatgtcaacgggacatgacacccgc catcgtctgaaaggactttggcgggcaggctgtgtcaacggcccaatcgtcggggcg aagcctacggcaagcattcgcgtcttcaggcgcggaagcaggttgacgatacggctcgtga cgatgtccacgggacgagcatgtccaggcgcgacgctgtgtgacacggggggc gaagcacatcgggctcgcgcgtcgaaccaagaagcgaactctatctccacaggtg aattgagggacgatgtgtcagactcctgacagcccgaggagaaatcgagaggac ggcgggggcgacaggcacaactggatccgcccgcgaggtggcagaacccgcgtgat gcgcgggaagagcacttggccacaagcatagtgcacggcgaggctccatgttcgcg caatcaagaacgcggcgaagatactccactgcggaagatagtggttgagattgggaag ttcatgtgcagaagataaaagaccccgacatcgaggcggtcgtgagtagcgcgcgtc ctgagatgggtggaagaacctaaaggcatagcccgctgcgcgtgagtgccgagaagt gccgcatctgcgggaagtcgtcgttcaaggacaaggcggaagctggatgccaccacatc atccagcgtgcatatggcgggacggatgtcccgcgaagtgtcgtgacattgtgcgagaa ctgccaccacgccaccacccaatctgcgcaggaataag	gtcagcggcccaaccctgaagggttgggct tgtgggtgaaagagaccgcgagcccgagc tgactaggcgactgaacgaaaggagaaa aactgagtgacagtagaaggtgatagagaat gcataggcaccgtgggaggtgttccaaat cccacgctctgcgtgtgcggttaaacagtc ctgcgaggtagggtgagtcggtacacgg gaaacctcttcaacgcgccgatgggcacc acaccgcctcggcgaggactaaccc gcaagggtgataataggaacacacatgctg
114-115	human gut metagenome genome assembly, contig: NODE_66_length_12806 8_cov_9.837282, whole genome shotgun sequence GENOME_ACESSION: OWGP01.1 GENOME_ID: 11987 CONTIG_ID: 65 SOURCE: WGS DATE: 2019-01-27.	atgtatgtattcgttatagggtggatggaacaagacttgcctgtaaaccaagaaaagc ccgaaaattaatagaagctcacaaggctgaatctataagaacagccgtttacaataag attactgtataagacaggctgtgtcttcaacacccataacattaggtatagatacaggtccc agcatatcggaattgcggtaacatcagaacaaagtccttataccaggcagagatagaac tgcgcagcactatgataagcgttactcactggaacacagatagctaccgcagaagca gaagataccgcaagaccaggtacagaagtcctaaattcagatttcataaaaacggacat attctgaaacacttgaagcgttaagaccacaggtattatgaccactgggttaacatatta actcaatgagtacaacagaccggacggatggcttgcacatccatgcagtcgaaagta gatcatcatatccgatgattaacagattttatagatgtgttccaccagatacgaactccg gctggaattgccaggtttgacatggcgaagatgaaaaatccggaagtgcataatgaatt ataccagcatggtccacagatgattacgaaattttaaagcttatgtttgacagagatca ttataagtggtgtgtctgtataaaggaaattaggagtaaacagccggatggacattcttaa aaggaaatgatcatcatataacattccgggtcaaaaggagctacggacaatcctgatcagc ttgttacagctcgtgaatactgtcacttccagcgcatataagaagggtgagatctgt ggatctgaaagagaacacaaatcagtcagcgggggactaagggtatgccacgagat gaacataatccgtacacgactgtgtaaaagcatttccagatgcggagctgacttatggaat attacggctgctgatagaagaagatgcatctccaaagtcacatgctgggtgatgcagttg caatagcaatgaaagggtgaggatgttaatatctgtgaaccgactgtatataaaacagatt cgggaagaaaaagcgttactccatgaagcaacgcctagaagggcgctaaaaagcccta ataccaggcagtaagaatcctaagaatacaaaaagcataggtatttatcatatttatgac actgtaaagtccagaatgaacaggattatctcaggatttcacaggcaagctgcatatgt ccaggattttgacgcaattatatacaaatccgtggaataatgtataaacagataaattatc gatttgaagctgttaaagcataacaacaactgtgattcaaaagaccagaaattaa	ttgaataaccgagcatgttataaattgtcg cccgtgaaggacaggtgactagcctaagtc caagagaactacatcagaagtagatccca gtccagacggtagccctaatctgagcgtcg gcggcgcaacatagatccaggagaaagg atgatgtaacccggtgttacaagtagctgtat attggcgaaagggtaatgctgactttggcca gcatagac
116-117	Ga0307373_10037003 GENOME_ACESSION: OX2MG_2_\$F_3300031 672 GENOME_ID: 280771 CONTIG_ID: 37002 SOURCE: JGI DATE: 2019-06-13.	gtggtctctgtctgtgatggcgcaagaagccgtgatgcttgcagtgcgagcggcg gcgcacatcctgtgcagcgcggcgggcggtgtccaccgcctgcggcggttaccatc cggctccgggacgcacggcggggagtcgggttcaacccctgcgcctgaagggtg gcgcccgccacgaagggtgacggggggcgctctgtccgcgacgacgttccggtatgct ggccgctgtcttgcggcgagatcaagcacaagccggatccacgccaagatgc tgcgccggcggtacccggcgccgcccgttcagcaacccctgcctaccggggcc ccgcgtttgacaacggcgccctgagccgtgcatcgttttggcgggcaacgcccgcc acggcgcaaacctgtcccgatgcgtgcaggcggggacgaacccgcaggccct gggggcgccggcccgccgtgcggccgacctgtgcgaagccggcgcaaacaccc gctcctgggtgcccggctgcggcgtggggccgggtgacggcgatcggtgtgtgt gggggcgctacgacacgaggtctgtcaggaccccgaggtccggggggtcagttacca gcaggggaccctgcagggtgacgaagtcgcgagttacctgctggagaaattcgggca	gcccgttaccctccgggatgcttaccaggc cgagcgcccgaggttgcagccggttggag gcgctacgttgcggcagcccaagacc accccgggatgcttctcagtcgccggctct ggaacccggcgagcagacaagcccggt gggcgggggcagcaaacggcggtcgga cgcggcgggcggtgacatggcgagggg agcggtatcgggcaagccggccggctcc gtcacacagggccgtaaggggcaccggc caaggggcctgacgcgagggaagggggc gcggccgtg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		<p>ccgctgcgtttactgcccggcctttccggcgaccctgtgctcaacatcgaccatgtggtg ccgcgacgcccgcaggggcaccgatcgctccaacctggcgctggcgtgcccggcg tgcacgaggcggaaggacgaccggccaccgaggagtgggcgggcgcgctggcac gacctgattccccgctggcgctgcgcggatggcgggatgcaaggccgtcggggcc ggctgcaggcgccgctgcgcggccggcgccgtcaacgccactgctggcgccgtc ggcggtgtgctggcgagacggcgctgcccctggaagcgggaccgctgcccgcacc aagtggatcgcgacgacctggggagcgcgaggccggccggcgacgcgccctg cgtggcgccctgcagccagccgactgctggggcgggagcgtgggtacacgccat cacggcgctcggcgggggcgctatcagcggacgaacggatgccacgggttcc gcgcgctacctgctgcgcggcaagagcgtgcatggcctggagcgggcgacctgtg gacggcggggggtgcgccatcgggcgggggcgctgcccggcggttggcgag agatcgaagccacaggccccgtcgggtgcgcggcaggcggtccttccgctcggcg cgctgaccccgctggtggcagggccctccgctgctgcagcgcggggatgggtacgcc taccagctggccccgctcggcgacaccggcgccggcgggcgcggggtggaaact ggaggaggcgccctcgga</p>	
118-119	<p>Ga0373633_0030533 GENOME_ACESSION: RO12tremer_MG_2_FD _SF_3300035661 GENOME_ID: 312627 CONTIG_ID: 30532 SOURCE: JGI</p>	<p>atggttttgagcgtgtttgtatggacaaccgaggtgtggcgtgatccgtgtacacaga agcgtgcaaggctgtgtctgctgcgtggacgtgcccgcgtgcatcggtgtgctttgt catccgatcggtcgatgctcaagacagggatggcagctgagccctgtctttgaagatc gaccagggcagcaaacatgtgccaagaccatgcgtggatgcggcggtgtgtcgg tgagatggactgggtatcgcccagcctgcacatcaatccacagggcggtggcagctac cagcgacacgcctggtgctgttgggttccgcgcggctacctgatcgcaataaact ggtcaagggttttccacaggtgatctgtcaaggccacgggtgacacacggcaaaaagc aaggcggtaccgagggcggtgtggcaatcaggggccacagcgagtttcaacatccaaac ccgagagggtgtgttcaaggcgctctcagaggcattgcacgtgtgtgcaacgcgggtga tggatgtgttaccacagcagagcaaatcagacgcaggaagagatgctgttcgcacg atcgctctgctccccggcatgaatggcggtctccccgcgaattcaagggtga</p>	<p>gtcaactaccccgactaaaggccgaggct tgatcagagatgttaagcctggttaccag accaagaaggagtcattttttctacgttgg ttaaggcgaagagtcgacccacgaatgctt cctcagttcgtgctctcgaaggcggtgtg cagacaagcaaatgggtacgcacgaacag gccccgtcattcccacggcgacagcaaca tggacgagggggagatgcaccgaagggtgc gctcacaaggcccttacgggtggtgct ggaaagacagcactattttggagagatgtt ttgagc</p>
120-121	<p>Ktedonobacter sp. 13_2_20CM_2_56_8 13_2_20cm_2_scaffold_1 557, whole genome shotgun sequence GENOME_ACESSION: GCA_001915055.1_AS M191505v1_genomic GENOME_ID: 113898 CONTIG_ID: 45 SOURCE: NCBI_Prokaryotes</p>	<p>atgtcaaaagtgtgtctctggatatgaccaaaacggcccttagaccagtcacccggcg cgagcacgcctgtgtctcaagaggcggaagccgctgtgtatcgaggtaccctttac cctatcctgaagacgcaagtggtgactccccgtgggtgtgtgctgcccgtcgaacttgat ccaggggcggaagacgagcgggcgctgcccgtgtgtatgacgcgagcggggaagtgtg gtggcgggcagaactcggccatcgggagcgacatcaagaagcgcatagatgcccgc gcgtggcggtgctggcgcaagagacgctcccgtttcacgcggtaccgcaagccgcgttc cacaaccgaggagttccagagaaaaaggcggtcttccgctcgtggaagccgg gttccaacatgctgacctgggtagggcgccgtgagggcggttggccccattgaggtgatc agcatggaactgttgaagttcgtatgcagggcattgcaaaaccggagatcactggcg ccagatcaacaagggggaacgcatgggctatgagacgcgggaatactgctcgcgaag tggggcgggcgatgcgctactcggggctgaggtgttccctggagattgagcacat cctctgccgcgcacgagcggcacgcatcggtgagcaacctaccctggcctgag ccgtgtaacgtcaagaaggcgacgaactatcgaggatttctgaacaagaaccaga cgtgttggcagcatcctggcccaggcgaaaacggcgctcaaaagctgctgcccgctc aatgcgaccgctggtcacctcttgaacgggtgaaggcgaccgggttaccctggagac cagcagtggtggttgaccataaaccggcggaagcgcatctgcccgaagacgat tggctcgtatgcagcgtgtgtcggccagagcaccacaagccttggaaacgagccagg tggctccttctgctatcgaagccaccggctatgcaaccggcaaatgtcgcgctgatg agcatggctttccatccgcatagcgagcgggaagaggtgcatcttggctaccagacg ggcgacctgtgtcggcggtgtgttcccacggggcgaggcggggacgcatgttgg gcgcgtgctggcagagccagcggtctttcgaattgaggacaaagctggagccag gccgggatcagttatcggtatgctgcccacatcgcacgatgctatgttaccagc aaggaggcgggcatgcagtcgccgccccaatccacctga</p>	<p>gtcatgaaccctgcatcgaatgcggggcg ttgcagcaaggctcatgctcgtacaagc ccggaaatgaccagactcaggttgaac agagcctcgttaggagcgaatccataggg acgtccgggtgagcgccagcccgagccg ctgcggcaatcgattaaacagggtagcgg gttaccacagtgctgttgcgtcaaacgctc cataacctgtcaggcggaacatcaccgg cgcttccggagcgctcgaagagcaaga ggaccccatcatgtcaaaa</p>
122-123	<p>Ga0209048_10010134 GENOME_ACESSION: FrelakCRCRSPAdes_SF_3300027902 GENOME_ID: 26208 CONTIG_ID: 10133 SOURCE: JGI</p>	<p>atgtcacaggtcttctctatagtcacaacgaacaggcattgaaccgatccatccggga cgccccgctgtgctcaagcaaggcagcgctccatcgacgctaccgttcac gatcattctgaagggtgtgtagagcagccacactcctcattgctgcgtcaaaatcga tccggggagccacaccatggcatcgtctcgtcaatgaacacacgggagaagtgtgt ggcgggctgaactgacctatgggtgagcagatcaagcgcgatttgacaagcgctc agcagtcagaagaagcagaagacagcgcaagaccagataggaagccacgcttgc aatcgccgcaaacgcacaggcggtgctcctcttttagagagccgctatgcaacgt gctcacgtgggtacggcgctgatgcagatcgccctgtcaccggcatcagccagggaac tggcgcgcttggataccaggccttggagcaccggatagcaggcggtgactatcag cgcgagacgtagcaggtatgaggtcagagatgtctcgtcgaatggaaacatca gtgcgttactgtatgcgcgaggtccgctggaactgatcatgtgcagcctgcag caagcgctgctgtatagagtgagcaactgacgtgtcctgtcgtatcgtgaaccagc caagggaatcgggatgtgctcgagttctcgtatgacgatcgtcgtcgttgcggcat cctggcacacctaaaagtccgctcgggatgcagcagctgtcaacactaccgctggg cactcaacgcggcgctcaagtatttgggttccgctcagagtgagcagcgcggttg accaagtataaccgctgatcgggggctcgacaagaccactggtcgtcgtcggcca</p>	<p>gtcaggaaccgcccgtgtagaaggggca ggctgtgtgaggaagcagcctatcttga ccagtcacgaccaggtcttcttcttattg aagagggcgagcggtcttctgctgagc gaatgataggcagctccagtgcttacc agcccgacctctgagagtagcattatca ggttgacgaggtaaagcggtgtgtgtc aaatgaaccgctcaagaacactgacgagg tgaacatcaccgcaagcgagggtcgt aagagcacaaccaaggaatccagatgtc acag</p>

SEQ ID NO:	Name	IscB sequence	ωRNAs
		acgtaggcaggagcacgccagcctcgtcatcaaaaggcatcgttccttgcacatca ttgcaaccgggacggcagccgccaaatgtgcatggacaagtacgggtccctcgt accggaccaagcagcgcaagcgtgtccaggggtccagacgggtgatctcgtgcgcg ctgtggtgacaagtggcacgaagcagggaacctacgtgggcaaaagtgcgtgctact cgtggcgtctcaacatcacctcgtcaggggtgtggtcaccgacattcatcgtcgtact gtacgctcatcgtagaccgatggtatctatcggcgcccaaggaggtggcactc tctcccatcgctga	
124-125	human gut metagenome genome assembly, contig: NODE_2129_length_960 8_cov_1.985450, whole genome shotgun sequence GENOME_ACESSION: OOZA01.1 GENOME_ID: 8927 CONTIG_ID: 2128 SOURCE: WGS	atgttcgttcagctcattagcaagactggcgtgagactcatgccaaagtgaataccgg gctcgtgaagctgctgaactccggaaaagcaatcatctatggataccgtccgtttacgatcc agttaacagaaaagggaactggcgtcttcagccagtagaactctgtgtgataccggat ataatcatattggagtatctgtaaagtcagagaagcacgaatctggaattacaggtcgtat acattaaccaacgagaaagaaagcatgatgaacgccggatgtaccgtgaagcagagac gaaaccggaagcgataccggaagccgggtttgataaccggaacgggaatctggctg ccttgcctcgtcactagacataagaagagggtcattcgcagtcataacaagaatgatg cgtgtatatacaattgcagacatcactctggaatgggaaacttgatacacaggttctga aagctcaggaaaaaggtgaagccgataccggaaggaaccgattaccagcatggggaac ggatgggattgcaacactccgggaagccgtcttcaagagatgagataaatgccagt gctgcgtcaggaatcaagatggcgcaatccctgcagtcgccacataccagatcgtacg agccatggcggaaccaaccggatgtccatctgtacacgggtatgtgagaatgccatc accagccaatcataaccgggtggaactgtacgggtggaaccaaaggctgcctcgt tcaaggcgcaactatcatgacgattgctccgtggcagttatatacaaaaagtgaagaag cactggcagtgattgggtcaagattacatgtggcgagaaactaaggaaacgcccgagg agcatggatgtcaaaaagtccatgtgaacgatgcattcgtaataggacgggtccatcaa agcaccgttcacacccgggtctttataaaaagaaacgccggaataaccgggtgtctggaga acttctatgatgcgaataatcagcagcaggaagcggaaagaaacgatccggccagga actgtttagtggcaggtatcagccggaaccataagaagattccgagaacctgcatcgata ccggaagaaaaaggtgtcaggggtgaagcgactattcgtacacagcgttacaagatcc agcctcatgatactgctctatttgacggaaagaaatagagaccacggatgccatacaa ggggaccagggaatcctgtactacggagaagaaatcgaatccgtggataaacttaca ttataaatatgcaggtggatattatcatcgaagtgttgccttag	gtcaacaaccccgctgaaataaattcagg cggagcttataaagatctgattttattagat cttagccccgtcttagatttattggtggac gaccgaaaggccgtaatgattagcctcagt gcaacgaaactggtcagacagacacgggt gaattattgactacgttaccgtgaaatta ggcactccgggatacacccttagtcccgga ccctgggtatcttataacatctctgaggg taagagaagtgtgagatacgcgaaactacc ggataactttggcgaagaggaccacggatt ctgagttaggtactgtaccgaaagatagag aacctgtatgggtatccataaggatctgta tacagaagcgtgaagcagattttaggaag gaggacagctgatgttc
126-127	Proteobacteria bacterium isolate F1-120-MAGs186 F1-120__c560309, whole genome shotgun sequence GENOME_ACESSION: GCA_005777235.1_AS M577723v1_genomic GENOME_ID: 284306 CONTIG_ID: 130 SOURCE: NCBI_Prokaryotes	atgagccatgtcttcttggaccgtgcccgtacccctctggaccctgccaccggcg cgggcacgggaggtgttcagcatggccgtgcccgggtgtccggcgttcccttcacc atcatcctgcgcgaccgcaagcgagccgagtcgggtggtacacgaccaccgttaagat tgaccagcgatgaagaccagcgggtgtggccctcgtgcaagaccagcgctcgtctgg gctgctgagctgaccaccgagggcagcggatcagagacgccctggagagtcgcccgg ggccttgcgctagtcgcccgaacgccagacgcgctaccggcaaccaggttttctcaa ccgcacacgtccatccgggtggctcccgcgtcgtctggggagccgtagccatcag atgacctgggtagagcgtctggcgcggtcgtgtccatcaccgccctgagccagagct gggtgcgttcgataccagatgatgaagaccctgagatggctgggggtggcgtatcagc agggcgagctggctgggtatgaggtcccgagttatctgctgagagaagtggaagcgac ctgtgctactgccacaagacggcggtgcccttgcagtcgaacatctatccccaaagc gcgaggtggctcggaccgctgtcgaatttgacgctgcctgtggccatgcaatcaac agaaaggcacacagagggcagcaggttcgggttccagcgatccaggctcaggcca aacagccgctgaagacgcgcggccggcaacgcgacacgttgggggctctatcgtg cattagcaacgacggggctccccgtcgaacaggcaccgggggacgcaccaagtaca accgaccccgctgaacatcccaaggtccactgggtgacgcggcgtcgtcgtggtgc cagcacgccagagcggttgcgtgctcggggatccagccgctggcatccgtgctatg gggcatgggacgcggcagatgtgtcgcaccgagcccatggcgttcccaaggcgaccc gcacacgccagaagaaatattggcgcatgcagaccggcgatctcgtccagcggtggt ccccaaaggcaagtatccggcacctgggtcagccgtgtggtcgtcagggccagcgg gtggtttgacctgtcatccacgcaagaaggccagcgtgacccacaacattgcacgt ggctgtgggctggcgtatgctacgtacaccctgcctgctgtgcaggcaccgctgtgt ctccccccactcaagcgagggtgtccacagcggaattctttag	atcagagtcgacgtactccccgactgaagt cgggggactcttaggaaactgagagtcgggt gtcagcagactccgtgccgaaggcagc ccgtcttttggctatgacaccctggaatgcg tggctagtccagccctgtcgtcgttga aatgtccgtggcgctaccacgctgtgcgg cagacaggacaagcagatgagaactgtcgt aggccaacgtgaccggcgcaagccgagc ggaaaggtaactctatgagccat
128-129	Ga0256831_1000378 GENOME_ACESSION: 132544metaG_2_SF_330 0028030 GENOME_ID: 20588 CONTIG_ID: 377 SOURCE: JGI	atgtctaactgtgtattgtattagatacaaatcgtaaaccattagctccgtgccatccggcta gagctagggaactattaaggaaaggtaaagctaatgtatttagaaaataccctttaccatt attttaaataaagctatcaaatgctaaactaacatactaccataaagtagatccgggt agtaagcagacaggaatggcattatgtcgaagacagaaataaaacccgtgtattttg gtttaaatagttcatagaggacaacagattgtgtgctttactagtagaagacaatcta gaaggactagaagatccagaaaaacagatatagaaaacctaggtttgatacagagct attcctaagggtggttacccttcaatctatctagactatcgaatattaaaacctggatc tgcgttttagtaaaacttactactatagctaacgtgtgtaggaattagtaaaattgatatga aagaataaagaaactctgataaaaaggtgaagaaatcagaatgggtttattaaaagatttag agttaaaggaaatctcttatacgcataataaccatacttgccaatattgccatggagtcca atgattctattcttgagaaggaaatataacccctaagtctaaaggagagctgattctgttg acaatctaccatagccgtgaggactgttaatttgcgaactaatctgctattgaatgattg	gtcaacaacctgcctaaagaacgaggtt gtattaattataagctctattattagggtga gccctaagaaggtatgttgaccagactctg gtttatgtaaataacactctgttaccaaagaat atataggtagccgtaggtggttccagccc acggctctacgtacgtgtttaaacaatgct gaggaaaggcatagtcagctatagataaa aaccttggataaacattgtcaggaacaacgtt atggtttgggtcggatgtgcccagactagatt ttaaatggaacatttataaaggagactgta atgtctaactgt

SEQ ID NO:	Name	IscB sequence	ωRNAs
		<p>gttaaggattgtctgataagaagataagctatctgtagctagaactaaatataccctaa agttataaaaggaattaaaccgagtttaaaagatctacggttagcaatattcttcaatcgt accatagctttctgtgaatctatgggtttaaatacagaggctgcgccttctatctcactaaat ataafcgtagaatcatgtttacaggaaagaccattggatagatgcgactatgtgtgtaat ctagacagaaaaatagagataataagctatgcgtgtgataacacccattgctcagaaaa ctaataatagacaaatgtgtctaatggataaatatgttttctaggactaaggctaaagg caatctactgtatatgggttaaaacaggagataattgttagggctgtgtgcttagtgtaa gtataagggaacatataaggaaagtatgttaggaataatgttaggtttgatgttaaaa ccattaaaggtaaaattatcactaattattcttattgttaaaaccacgcataagaagaagggtt tactatcaacacgggtactaaaactattctaaaaataatcctaagaagggtattgtattaa agatatttttaactgtataaactcaaa</p>	
130-131	<p>Ga0315550_1018591 GENOME_ACESSION: SalMarSW160190MG_4 _\$F_3300031653 GENOME_ID: 280938 CONTIG_ID: 18590 SOURCE: JGI</p>	<p>atgcagcagattctgtgctgagcagcgaccgcgagccgctagaccctgccaccgg cacggggcgcaagctgctcaaacaggggcgggcagcggtcttcggaagtggccgt ttacgataatcttgaaggatcgccaggttgccgagtcggtcacacaccgcaccgggtc aagatcgtacggggcagcaatacgcgggcttggagttgtgcaggagcaaacaggc gaggtcgtctgggcagcggaactggagcaccggggccagcagatcaagcaccgaatg acgagccggccgagttgcggcgggcgccggctggcaggaagtccgttaccgcaa gccgcgttcgacaatcgggcgctgcgacggcgaggcttcggcggcggcggcggcgtt cagagccgtgtcagaaatcgagcaccctggctgaacggcgtagggcgctactgccc gtcgagggcgctgctactgagctggctaagttcgacacgcagagaatggagaacccga aatcagcggtgcgaataccagcagggtgaactggcaggtacgaggtgcgcgagtac ctgctggagaatggggccgcaagtgcgcctactgcggctcggagaacgtgccattgc aaatcgagcatatctacccaaggcccgaggagggtctgctgagggtgagcaacctg gatctgtgcggccgtgcgaaccaggagaagggaagctgcacggcgaggagttgg acaccccgaggttcagggcgaggtcgcgaagccgtgaaggatgcggcggtactgaa cgcgacacgctgggtgctgttgcgacactcaggcgatcgccctgccgtagagatcg gcacggcgggcggaacgaagtaacacgcagcagctgggacttccaagcgcaatt ggactgacgggtgtgtgttggggagttgggtgagcaagtgcgcgttccggccgctg gtaccgtgcagatccggcgacggccacgggaagacgacagcggtgcgggacgga caagtacggcttccgacccggcagcgccctggcgcaaaaagtgttgggggttggcag acggcgacatcgccaaggcggtgatcccaagggaagtatgcaggcggtgcatgtg ggctgtgcgccattcgcatagggctatgttcagggtgaatggatcgacgtcatcca agtactgaggctcgtgcagagggcagacgggtatgcatactctag</p>	<p>gtcaactacccacggatgaatccggggg ttgtcgggaagcattgcgagccggagttg accagttccagccaccagccgagaggctg atggggctacgttatcagaagtccaagg ttacacacctgggggtgcgagcagccccc aagctctgaaccggaagctaaacagccat acggggtagaaggcagctgtttcgggagat ggccggccgataacatcgacgaggtcact ttaccggcaagtcctcgggacgaaccgag aatccgtaaggaggaaatggaatgcagcg a</p>
132-133	<p>Ga0137384_10001405 GENOME_ACESSION: SagtaG_31_\$F_3300012 357 GENOME_ID: 35418 CONTIG_ID: 1404 SOURCE: JGI</p>	<p>atgtcacggattttctgttgatgccacggcgaatccctgatgccttctactcctgcgcg tccccgcctgctgctcaaacaggcgaagcgccgatcttgcggcgccctctgtgtg tgattctcaaaagcgctgcggcagagcggtgtgctcagccattgagcaactgac ctggagcaagcaccagtggtgacgctgctgtcaacgatcgtcgggggaagtgtgtg ggctgcagaagtgaaccatcgccggcgagagatcccggaagctctcaccagaagacg agcggtacgcccgacgctcggcgctgcagagagagatacgcacaaacgcttgc aatcgccgtcggcgctgaggtgtgctgctcctcttctcagctgtgctgaatctgct gacctgggtggcgctcgtcgaagctttgtccgatagaagcactctgcggaactggc cagattcgacacccagggcatgcaagatccgacattgctggcatcagatcagcagg ggagcctgtgctgatagataagatcatacttctgataaaagtggcaaaaggcgtgtg cctactgccagcggcctgcaggaagtgcgaagtcgaacatctgattccaagagccg ggaggggagcagagaattccaatgtgtgctggtcgtcgaatctgcaacatcgccaa ggggtgacagggcgtgaagagtttggtttgcagctgattggctcaggcaaaagtacc ccttgaagtgcggcagtaataatgaccccggtggaggtgtaccaggagttgagggc gataggcttccagtagaagtggacacgggggggagaacaggctacaaccgtgccatt cgccagcttccaagcagcactggatagacgcggccctgtgggggacctcaactcctg aacagctccaacttcagcagcttgcggcctggcagatcacagccacggctggcagcg gctgcaaatgtgcctgtgatgagcaggatttccacgacccgcgccaagcagcgat cctgtgtaaggattcagaccggcgacctggtagttggcgtggtcaagacgggaac gaaacagggcatctacaaggcgcggtggcggtcagagcttctgttcttttaacatcac gaccgacaagtcagatccagggcacatcatcgtgtgtccgggtgctccagcg cgtgatggatgcatatcagcaacgagaggagcgcgcttctccaccccttga</p>	<p>gtcagatgcccaagcctctgtaaggggcg ggcttgcctctatctggcagttccctcggt gttcgactgaggagcccttgggagagc gccaaggtcgcacgtcgggtgaccgga tcagcccgacccctgtaagtgcgtgtgtaa acagcttgggtggagaagcagttcagtg gacagatggcgctccaaagcgagcag gaccatttactgagcaggcaactgctcag cgacggccattatcggtgtccgagtgcg ggaaaggaaagcgtgcggtgacgg</p>
134-135	<p>human gut metagenome genome assembly, contig: NODE_13340_length_31 49_cov_39.749515, whole genome shotgun sequence GENOME_ACESSION: OIWW01.1 GENOME_ID: 6110</p>	<p>atgtcaaccggccttggggcgagcaagcggagcctgcgggtatccgaaaggagac actttgaaagtattgtttgaacatgcggcgaagccgctgatgcctgttgcctgcaaa ggcgagacacatgctgaaagcgggcaagcgctgcttgcgacgcacgcgttcacg atcagctgacgtgcggcggcgagacgaagcaggaagtacgcttggcgctgatg caggcgccgaacacgttggcatttccgccacgacggaaaaggaggaggtcttcgctc cgaagtgcagcttcgacagacatcaagggaacttggctgctgtgctgcaattccgacg tgaagggcgaatcgcaagacgcgttaccgcgacacgcttcaacaaccgcgttcga tccaagcacaagggtggttgcggcttccgtggagaaccgattcagggcgacatc gcgcatggaagcggttgcggagtgcttccgacacaaaactgtagtgaaccgcatcc ttgcacattcagaagatcaaaacccggagattgaaggcgaagctatcagcaggcg</p>	<p>gtcaataacccccgctgaaggcggaggct tgaaaaagcctttgttactagtctcagaaa cctcttcggggaggaggagctagttgttgg gaattgacaggcaccgtggatgttcatcct agctccacgcttgcggccaggtgttaaaa gctctgagaggttagagcgggtgctgctggc aagaacaccttccaacattgacgaaggatg tcaaccggccttggccgagcaagcgga gcctcggttatccgcaaggagacattt gaaa</p>

SEQ ID NO:	Name	IscB sequence	ωRNAs
	CONTIG_ID: 13339 SOURCE: WGS	agcagcttgcttttgaacgtgcgcgaatatgtctgttcgcgcaggtcacgtttgcca agcctgcaaaaggcagatcgaagatctgattcacaagtgcatcattgagatcggaa aacgggagggcgacgcgccgggaacctcattacgtctgcaggcggtgccacaaggc gtatcacgcaggcaagtgaagcagttcagctcccgccggtgttcttcagggcag agactttcatgggcatcatcgtgttgacggtgctcaaccgctgcgcgagcgccatccc gaattgcctgtcacgaatacctacgggtatctgaccaaacacaaggcgtatcgcgcg	
136-137	Ga0137365_10006127 GENOME_ACESSION: SagtaG_29_SF_3300012 201 GENOME_ID: 35416 CONTIG_ID: 6126 SOURCE: JGI	ctacaagagagcgtgtcctaaatgacctggcagcagaaccagcggcatgacctggt gcgggcagaaagagagccagcgggagaggtgcatcatgcgtcttttctccgaagt cagcagcgaggagagctgtccatcaggcgaagcagaccacaagcaatgccgacg gcgtagacgcagtgccactctgctgacccgcgcgcacgcttcgacaaccgggtgttg ccaaagggtggctccgccctgatgtctccgtgtgggtaacctctacctgggcc aggcgatacagccgatgggtcccgtcggacactggacgggagagcaacgcggtg gaaccgggcgcgttgggcatcgcgaaaacgatgccaacgacgcgttgtgcgtggg agagatc	gtcaacgacccagcgtgaagcgcgggg cttgaggatcttcccagggcccgtgtg tccagcctgagttctgatcccaggtagg gagcaggggaactccgttcgtccgtcag gacaccaggaacgcctcctaatctgtg ccctgtgtccagcattaaaagcccgtcgg ggtgtcggcggtgtgctgacaagagca agccggagggaacattggcgagaggagctg cctgggaaccccgtccaggcgctcaga cagccccgtaaaggggtcccgaagggaac agaaaggctcaacagcatg
138-139	human gut metagenome genome assembly, contig: NODE_3863_length_375 3_cov_1.622173, whole genome shotgun sequence GENOME_ACESSION: ORBQ01.1 GENOME_ID: 10332 CONTIG_ID: 3862 SOURCE: WGS	ttgagagtattgtcctaaacacgcggaaaaccgtgatgccctgttaccggcaaaa gcgcgccttctgtctaaagagaagaagtattgtgaagaggcgcaacgcctttactattc agctgacgattgcaacgggtgagtcgaacacgcggtaagcctgggtgttgatgccgga tacaacatctgcccctttccgcatcaacggaaaagggtgagctttatgcatcagaagtcg aactccgtcaggacgtctgtatctgtctctgtcgtcgtgcttaccgacgtcgcgcg taaccgaaaacgcgctaccgtgctccgaggttcgacaaccgcacccgcaccaagcgc aaagggtggcttgaccatcagttgaaaaccgaatcaacgcgcactgtcgcgcataga agcgtgtcttgcactgctgccgtgacgaagatcaccgtggaaacggcgtccttcgacac gcagctgctgaagaattcagacattgcagggaaaaagtaccaagaggggcaacagctc ggcttctggaaacgtccgcagtagctgttctttcagagacgggcagctttgtcagcattgtca cggcagatcgaagatccgggtcctaatgttcatcacttggagagcagacgtacgggcg gtgattcggccggcaacctgattacgctttgtgagacgtgc	gtcaactacctggcctaaaggccgagcgt gaaaaagcctctagtgtgactagcctcagcc gtcgtttggcgactacgtgtgtcgggaacc tataggcaccgcgggatgcagatcctagtc ccgcgtctgcggccgatgtttaaagctct gagagggaggagcgtgtgtgtcgtcgtga aacccttccaacattggcgaggatcacaa ccgttcgaaagaccgaggagataaaacttg aga
140-141	human gut metagenome genome assembly, contig: NODE_801_length_3381 9_cov_6.114886, whole genome shotgun sequence GENOME_ACESSION: OGQE01.1 GENOME_ID: 4614 CONTIG_ID: 800 SOURCE: WGS	atgagcgtacttgaatcggcgcaaatggaaggcggtgatgccaccaccacaagaaa agcccggtatctgtgaaagaaaacaaagcttctgtttgatgcagacatcgttgcagattc atttgcggtatataaacccggtatgcgacacaggaaggagcgtggtatgcatactgggt cacaacatcgcgaatcggagtcacctgtggaataaagtgtatcctcaagatgaacatg aattacgtcttttattggagaagctgtccctgatggaacgagggcaacatgcggcggtg gacgcagatcgaaggggtgcggtaccggaagcgaagtgtgaaacatcacaaaaacg tatgtatttcgagaaggcgaaacgggaaggccaacactgggagaaagtataaaacaacc acacagtcaccaagaccggaaagctgggtgccaccgtcactacagtcgaagtgcgacc accattccgtatcattgaccggtattgaaatgcctgccgattcaatcacgaagaatttag tgatcaggtaggacgttttgatagggcgcatgaatgatccgacgatccatgtggaat gtaccagcgtgtccgatgtatgatacaggaacctgagagcgtacatttgcgaaggga tggatataaatgtcctgtcgaagcgaaggcaggtatctccgaaagaaagatggga catccgtgaaactgattgccaccatcaggttccgcagtcgcggagctacggacaatc caaaatataatgatcagcgtatgcgatcactgccatcagcaaaaggcacaccagcctgggtg ggatcctgtattcctggatggaagaaataaaaaagttaccgcggattaagggtatgcga catttatgaacatccttcggagacgattatttgacagatttccaggcagcatttcatatg ggatataacagcgggtgaccggaacgtttactgtctcgaagagccatgcaaatgatc ctgttgcaatcctcctgtttgggaagatgtttctgtcatcaatgccacttgcagacgttac gttataagcaggtgcgaaatccaacgttctctgcatgaagcaacaccaagggaagga cggaaggaaacaaatacaaaagcgtatcgaataaagaagaacacaacccatgtaattgg atttaagctgtgggatagtgtctgtaaatggacagaactatttattgctgtttcacagg ctctagtgcatactggtgatcagatggacattatgtatcgcctcctgggaaacataca aacagtgacattatcgaagtggacgggtgcatccaatggtaactggctgatggcgta a	gtcaagttaccaccgcttagatctctgatga aatctgaagcggggcgtgaaagaccatt tgactagcctcagtgatttggtaagtctgtc tgtccgccggatagatactgtcagataa actacgttatgagtgcatgatacctcgt gaaccttagcaggaagcggtatcctggga accgatcgtgtgtgtgggaacccacagcga cccgaaaactcaagcctcatgacattggcg aaggacattacgcggcacgcggcagg gacggttcgttccctgtgaatgcaataaca taaggaggcgcttatgagc
142-143	human metagenome genome assembly, contig: NODE_23_length_11264 7_cov_10.9368, whole genome shotgun sequence GENOME_ACESSION: ODKX01.1 GENOME_ID: 3080 CONTIG_ID: 22 SOURCE: WGS	atgacaaaaactaaccacaactacgcaagcttgcgtgttagaccaacacggtaa accgttaatgcaacacacgtcttggcaagtttatcgtcttttaagacacaaaaggctc atatgtgtcttatgagccattcacaatcgaattgattataaacctgacacacatgttatca accaatgacacttgggtgtgatagtgggtctattcattcaggtattctgtagctaatgaaca cgtgaatttatagtagcgaagtattgcacgtgataataattctgtctgtatcagatagac gtatgtaccgtcagactcgggttcacgaaaacgcgttatcgtaaaccacgtttcaataat cgtaaaaataagaaaaaggtgtgttaccacatctcttgaacaaaaaggtgtgttcaatt aaatgaatcgaacatcttcatcgttatttccaattgaacaaattattgtcgaagtgtgag tttgatattcaaaaaataaaatcctgatatcttgggaaaagattatcaacaaaggaaacttac aaggctataatctcgaattatcttgaagacatagtcgtaaatgtttttattgtgataaa gaagtgtcaactttgaagttgaacacatgattccaaaagctagggtgttctaatcgtatt gacaatttaaccttatcatgataaaatgtaatacagaaaaaggaaactaacggcagatg	acataaataataaccaccccgctgaagtgg gcggttgcgtgactccttaccgtgagttat caaactttattgattagcctcagtgtaacta cgttactgttaatatataggtaccttgatg tctgcctagtcgaagctctacgcgttatcatt aaacagttctaagggttaggaacagtgtaaf gacgatataaaactacaagataacattggcg aaggcaacataggggtgtttatcccgctta ccgcataaaataaaatftaaacgaaagg actttgtcaaacgtatgacaaaaactaac caaaacaaactacgcaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aatttatcaaacgaactttaccggttgaatgccgctaaaaaattaaaacaattgtctaaag aaaaacgattgtttaatatatggcgcatatgaatgctacaagatggacgttatataatgca attgatgacaataatccaaatgtcaaatgacttatgttatattacgaaatacaatcgatttc aagcaggtcttccaaaagcccatcatatcgatgctaaatgtattacgggctttgctcaagta ccatcatttgatataatggttgaataaacgaagatgcgttaggcataatcgtaattgcatcga gcaacatttagtaaaaggtcatgtacgtaaagggcaagtgtccgacagttatgtttgcttt agattatagaccgtgtattatataataatcaccattattatataaaaggctgcgaagtagcg gtttattgcaacttgcttctgttgaaggtttgaagacgaaagcagaacttatagaaaattga catttttggcgcatacgaacgcttatttgactaacgatacattaacaatgatacagtaacaa tttaattaaaaaatag	
144-145	Ga0070706_100018127 GENOME_ACESSION: KBSK5StandDraft_5_5F_3300005467 GENOME_ID: 30504 CONTIG_ID: 18126 SOURCE: JGI	atgtcacaagtctttgtcttagacacgaccaagcgaccgctcaacccgggtcatgcgggg cgtgcccgttgcctctgaagcaagcgagcgccggtgtatcgccgctatccgttcacg atcatcctgaagcgccggcgtagcaaccttctctgaaccgcttcgagtcgaagtcgac cggggagccaaaccaccgggctggcggtgtggtcaacgatgcagcgccgaagtcgtgt gggcccgtctcctgacgcaccggggtaagcaaatcacgcgcgatctggcgagccggc gcacggtcagaagagtcgcagggcaacgcagaccaggtaccgcaagccacgcttcg acaaccggcgcaaaaagaaagcagcgtgcccgtcgttagagagccgaatctgca acatggtcacgtgggtgcggcgttctgcggcgtgtccggtacggccatcagtcgaag aacctgtgaagttcgacctgaacagatggagacgcccgcacatcagcgccgtggagta ccagcagggcacactctcgggtatgaggtgcgcgagtcacctctcgaagtggcagc accagtgctcctactgtgaagcccgtgaggtgccattagaactcgtacacgtgcaccta gagccaagcatggctctaacgggtgagtaactcgtggcgccctgcacgactgtgaac cagcgcaagagcaaccaggatattcgacgtctcctccgatgactcgtgagcggtggc gcgcacctcgtcctcaggtcaagacgccgttgcgggagtcagcgccggtaaatgcgacc cgctggcgattacacgaccgactcagcgggtagggtgcccgtggaatgcggcagtg gcgggaggacgaagtacaaccgcgtgagggcgggggctgcccaagtcacattggctc atgcggcgtgctgtggggccagcagccggagcacctggacgtgcggggcgtggcg cccctgcacattaggcgacaggcgac	gtcagagaccgctccggtcaacgggacg ggcttgcgtgtaccacgcagggcccgtc tgaccagctcagccaccagctccaagg actgacggggctccgttgggaagcgaatgca taggaacgtccgggtgctcaccagcccgg accgttcggggtagcattagcaggtgga ggggtaagccaggtgtgcttccaccggaa accgttcacacgtgtgacgattgagcatt acctgggaaccagagggccatgcgggca catcacaaggagtcattgtgtcacaa
146-147	Ga0137383_10047051 GENOME_ACESSION: SagtaG_26_5F_3300012 199 GENOME_ID: 35413 CONTIG_ID: 47050 SOURCE: JGI	atgagcaagtgtttgtcttataccacacaacaacattaaacccgggtcatcctggca gagcgctcctcgtctcctctggcaagcgaggtgtgaagcggtacccattacgat catcctcaaaactgtggtcgagcagtgctgcttgcgggtgcaagattgatcct ggcagcaagacgacggggctgagcagtggtcaatgatgccacagcgaggtcgttttcg cagcggaactggagcatcgaggtgagcagatcacaagcccttgcagggcgagcctt ccccaaagacgattggctggtgctgctgtgctgcaagaatagcctgagcgtcttcg cctcaagggggtgtgcccattgctcatcaaaagccaacgggcatggctgtcggcaactgtg tttgatggatgagcatggcttccacgcaccaagccgaacagagaagaatttcggcatct gtttcggacgggcatattgtccagggcagagttccggctcacctcaaacaccagggcgt ccatgtgggtagagtgctgctaaggcaaacggagccttcacgattgcgaccagatcag ggaaagtcaccgattattggcaagaagtattgccgtgtctgcaacggcgatggctatg gctatctacagaaggagaggaggaatttctcccgcccctcaa	gtcaggaaccatcccgtagaacgggatag gcttgcagaagcaagttgaacctgaccga ctcagctcgacaacgggctacgttaggagc gaatttaggtacgtttgggtgcgagggcagc cccaacctctacggatcaacattaaacagct gtaacgggtgaaggccgtgtgtgtacgtca aaacctccataacctgtcaaggccaccat taccttcgagaggaggctcgcattgagc aaa
148-149	Ga0373956_0000940 GENOME_ACESSION: WV9ome_35_5F_330003 5119 GENOME_ID: 307019 CONTIG_ID: 939 SOURCE: JGI	atgaagggtttcgtgtatccagggaagaacaccaggatgcgcagacttccaggcg ggcccgcctctgctcaaaagcgagcgccccgagtggtgcatatgagccatgtccc atccagatagcgggtgacactacaacgtctactacgctgtcacgttgatgtagagacg ggatccagacaggtagatcgccagtggtggaggtgtttccctgctgaggtcgtcctc cgcaccaacatcagccacacagtggggcaacgacgccaggactgcagaagaggcg aggccggacaactcgggatagacagggccgatgtgcgaacctgcggtgaagccagg gtcgttgcactgtcagtgccagccatcgccaagacaactgtcaagaccgtgcgcgttgt ggccactggagtgccagtggggggggtcgatgtggacgtcggacgcgttgcagcga gacggtgcagactcccacatctccgggtggcaaggccagcatagtgcgttcaggact ctcatcttctggaagacctgcgtgccagatggcacacgccatcgccctctggcacctga gtgcagtgccgctccaggagagatctgagacccgtcagccgaggagacaacgattg cgctcgcacatggcaatcgctgtgcggcctgggaatcaggcaaaagggaagcacacc gccgccgagtggtgtttccacaggttcaggctcaggcatgggtggccctcacagatgca gcgcagtgagcgagtgaaaagacagtgtgtgtacaaatgcagggtatcttgcggatc ggattgggtgacgacacagtggtcgtgcaacgaagaagacacgtatgcaaatccccg gccttccagctgcacaccaacgatccatcgcatggcctgtgcaggaggcgaggt ggtaaaaccacatcggtgccttccactgcacaacgggaacggagtgagcacgcg gtctgggcatcttgcagtgcaaggctggaaggtggacgaagtgggtgaacgtcaagg gcggtatcggttcattgtagatgtgtctcagagcgatgtgtggtcaaggatctgac gagtggtgaaagcatgtgtgaggtgacatcacgcatactgcgaacggttggccgtccg tacaaggctggatgacatcgtcttcttctcacaacagaggaaaggagggtggcg ttctccctgtctga	gtcaatgacctgtccctccagaggagattt gtgagagcaaaagctcttgcgcgtcgattg accagactgagcctcagtcgacggcgagc cgagcacgtgagcgtcccatctgggaag aagaggcgacgagatggctatgtgtgtct tctgactccagcgtgtgccccggcgctgg accttccatcgctgtatcaccagcggtga ggtcagcagcagtggtcaggagaacgaac ggggctcttctgtgtcaggggatcacatg tcttcaatcggaagagaggagccttctctg at
150-151	contig_6736066 GENOME_ACESSION: mgm4547164.3	atgcagtatgtattgtattatccagaacgggttcgccgctgatgccgacaacatcggtca tgccagaaaactcctcactccggcagagccgataatcgaaaagcacgagccgttcacaa tccgtctagatcatcagccgccgggtattcagcctgtcaggtataagacagataccgggg	gtcagccacccatgactaaagtcaggggc ttgcagggaagaaagaaaaggcgttgcctt ggtatttctttaggaaagccatagctgac

90

[illegible]

92

SEQ ID NO:	Name	IscB sequence	ωRNAs
170-171	Ga0075125_10001675 GENOME_ACESSION: AntAceMinimDraft_18_\$ F_3300005935 GENOME_ID: 22974 CONTIG_ID: 1674 SOURCE: JGI	atggaaacaaaaataaactacgagaaaggaaactgaaaaaacaccaacagacgcttact aatctgttctctgaagagaacccattaaactctacgagtttagaaaacaagagaacactc ttgataacttaagttctttttcttggggaagtgaagataacttccgaataggaaagcggttt gtagccgagtggttattaggggtgatgggaaaccattaccccttgaacataaaaga gcaagggtgtaataagaacaaaaaagcaaacctgtatggaataagtttgagaaatttg gaattaaatgttagaagaacagcgagtgaaacagcagaataatttttagggatagataa tggacaagaattggaaggtattctataattgtgataaaacaaataatttaattgtagtgga aacttccagataaacaaaaactgtaaaaaaacttgaaagaaagtagattaagaagag cgagaagacaaaagaattgcaggagaagagaagcaagatttaataatcgttcaaaagat ggttttattgcaccaagtcattacaggtgtttcaagtaggttgaaagcaataactgagtt tttaattcttaccataaagcaaggtgcatgtgaagatgtgaagtttaaccaggagataat agggtgggggaagaatttttctacaattgaagtcggaaaaaatatgattaaagattttat aataaaattggaaggaaaaatttaataaattttgaaggataaacaataaacttgagaaa taataataattgaaaaagcttcagataaatcaaacacagagttttatagccattgtgtag agttttatatacaatggaactattgaaaaacctattattcttaagaaagtataacttatac gatgataactacagacctgttagaagaaggttacacgacactcaatttcaaaaggtggaa ttagaagtaaatcttactggaaattcaaggataatcaaaaggggtgattataggagatg aaaatggttggttggaacattggttggtggaactaaagataattgctggtattgtattttg aatgagaggtaataagaaaattatcaaaaaggaaaaaccataataaaattagttggatt tcacatcacttcaatataataaactaaatataaacaccaattatccccccactgaagga aggggacttcttggtcagacaataatgttaa	taaaaagaatacgtgaaaaagaaaaagaa ttaatgtaaaaaaagaagtctcaaaa aagaagttagaaattgttgatgaggat ggaactattcttgattgggtatgctggat cttaaacctaaagaaaaagagctctgatgc tgttaaaattggaataatgttctctgaaat ggggaagaggtgaactacccccccataaa tgaaggggcttctatgatgtaagaacgca agagttactacactaagttgaggaacaa agaaatggaacaaaaataaactacgagaa aggaactgaaa
172-173	Ga0209347_1000563 GENOME_ACESSION: AutmicBR23SPAdes_\$F 3300027640 GENOME_ID: 23478 CONTIG_ID: 562 SOURCE: JGI	atgacgcgatgtctacttcttccaggaacggccgctctgatgccaccatgccggcc agagccagacacctgctcagagcggcggaagcgagggttggaagatcaagccgttc accatcagtaaccattaacacgccagaagtggcgagccgggtatagcggcgccagg acccccgattaaactcagggaagtgccgcagtaaggaacgacggcgaagttttgtttaa gccgaagtcaaatgccgacatgcatttcagagagactcagggaacgcccggagttacc gcaaaagcagccgggaaccggaaaacaggggtaccggcagcccgaggttttaaacgccaa aaagccggaaggctgggtggccccgtcaatcaagcagttaaaacacgagcatgataaa ctccggcagttggtggaagcatcctgccgtgacggcgtggttcatagaactcaacaaa ttcgacttcagaagatgaggaccggaacatccagggtgttgcagacagaacggcccc gcagaagagctgcttcgacgttaagggaatagctcctcgaacgcgacggctacgcctga ttctgtgcggcggtgtaaaaaaccgaaactctatcacttccgcgcaaatccgagcgtc cgaagaacctggtgacccctgctcggggaatgccacagaagggcgtcaacaaagaaat cccttttgcgtattgctggagagctaccgttggcgagccagggtcaacgtgatgctgtgc gctgtggggccggtccgacaaatcaatttgcacagcagagcaggcagccgctgcc gggaactgctccgcatggctaaaacgcacagcaacgacgcttggccgagttcatgc ggcttatggaacagtaccgccggcggaatgcacccctgcacggccgtacgtccgg cagaagaaccggcaactgcacggcggaatcccggttaagggcggtgaaggcagttg gctaaccgaacccgtacgtgtaagcaaggcgaggagtgagggtgcaaaatagac ctggtatataccgtacccggagtgccggaaaatcaccgggtacgtcaacaccctgttc agccgcccgcggcgccgtagctgattacgccggcagggaactgtacagcggagc gagtgtaacaaaactcaagaagctgcagaatgctgataattggtatggaggtgtataaa	gtcaaccacccccccactgaagtgggag gcttgcagaaaaactgcgagcctcgggt gactagccccagcctgtcagttacatgccag actacgttggaaaggccatcacacctgga gcgtagcccgaactcaggctctgtgtgc gggattaaaagccctgcggcgaggggcg gtgttcgggacatacaaaccttcaacat ggcgacggcgcaataactcaacgtgg aggaggtcatctaaatgacgcat
174-175	Ga0209941_1000055 GENOME_ACESSION: ThiBioThioSPAdes_5_\$F 3300024978 GENOME_ID: 36135 CONTIG_ID: 54 SOURCE: JGI	atgacgggtattgtcttggaacaaactaaacggccgctgatccgtgtagtaacaaacga gctcgtttactgttagagcgtggcgccgctgtgttcaccgcttlaagcccttaccattcgg ctgaaagaccgtattcaggggcagctgctgtttacacatcatgctgggtatagaccgg gttcaaaagaccacgggtctagccctcacacgcccgtgacgggtgaggacgccgtactggtg tttgggtgtgagttacagaccgggggttggtatagggccaagcttttaaggcgagtg cttaccgctgaatcgcgctcccgaagaccgctatcgccctgctttaaaccg tacgaaacaaaagggtggttaccctccctctctcgaatcagtagaatacaactgtgaca tggcggggtcgttctgacgcttagcaccagtgaccgcttggcttatgaagctgtggaat tcgacacacagcgtttacgtatcctgaggtctctgcatacagtagcagcaaggcaggt gcaagggtataccgtccgtcctacgtacttcagaagtgaggattacgcctgtgctgactgt ggatccaagaccgtctgacactgtatcacgttatccctcgtctcgaacgggagtgac gcggtgacaaacctgtctgtgctgttacggctgtaatacaactgaagggaacccgtcgg gtgcaagagtttttagcaaaagaaccacggctactaaaactgttccctgaacaaactgaaa aagcccttacgagacgttagtgcggtgattagtagcgcggccggttacataaggcggtta ggggaggtaggaattccactcactgttggtactggggcggaaccagctacattagaca ccgattgaaactccctaagatcatgtcgtggacgctgctgtgtggcgctaacagggaac ccctaaaggagaatgtttaaaccattactgtctgttggctgctgtagggagcttacaac gtgtcagaacagatcgttctggatttctaaagcgacagagttcgcgtcaaacggccatt cggatttcaaacaggtgacctgtgctgtatgggaagttattgtaggacagctgtaaga atgacgggtttttcagtttcaacacaaacatcaaaacttaacgttaaaatggctgaaacta acactgggtcagagaagtgacggctatctctattgtga	gtcagcggctgcgatcaagttcaggtacgt gactttcccgaacctagtttgcgagggaa cagactcaatctgacgtaaaagcagactacg ttactcaagaaataggtatcttcatatccct tcagttcgaagctctacgttattacgttaac agctttatgggttgaaggcagtgctgaatgc tcaaaccttgggataactttgtcagaggact gtgggtgaacggccactgtcactgaaggc ccgtaaggcggttgaaggattttaccat gacggt

SEQ ID NO:	Name	IscB sequence	ωRNAs
176-177	Ga0208186_100002 GENOME_ACESSION: Uti3Day2SPAdes_SF_33 00026735 GENOME_ID: 372111 CONTIG_ID: 1 SOURCE: JGI	atgctgtgtctacgtattagataaagatggaacccgttgatgccaacacgcccgtttaggac gtgttcgtcattggcttaagatggtgaacaaatcgtgttcgcatcaaccattaccattca atttacaaaaaacactgaacgtcactacagacttaaatcttggtattgacgctggtataa acacattgggtttctgtgctaacaacgtctaaaaatgaagagatctactcagaatctgtcg atcaacgtgcaaacggtgtaaaaagaattgaccgtcgatgtatctgtagaactagacg taataaactacgtcaccgtaaacccgttttgataatcgtaaacaatttaaaaagtctgaag gctttacgccctcaatttatcacaagttccaagaacattgaaagcaattggccgtgttaaga aattcttacctgtgtcaatccatattgaagttgctcatttgataccaaatgattaaagct aaatttaataacgaagctgttaaacgtcaaaatggcgatagcaaaagacattctgacattaa gtcttatatttgaacgcgacaactacacatgtgcagtcgcaatftaataaaagcaaggtgc acaattacattgtcatcagctcaaatftaataacaaagcggaacgaaccaacctgataatt tacttacagatgtactaactgtcatcacacctgacaatcatcaaaagggtcacgtactctatc aaatgatgcaagcaaaacaaacccattttatcaaggggcggttctcatgtcagcactcaa gattttattagaagaacaccttacatttcaacaagccttgggtatgaacctctgctaaacgt attagcttggctggaataaagaccatcaggtgacgattagtcattgcccgcgcaata atgacacaaacacatattctgttcatattaacctgacaaaactacaacgaataatcgaagt ttgtctaagttctatgatcggaatggttagactctgtgaccacaatgtcaaatccggcaa agaattgtctaattggacgctgaaccgtatcacaccaataattctgaaaatgagcacgttt accgttgcagaaaattaaaacaggacgtatctctacacgaaaacaacattatcaaatcgc accacatgatattcctaatttggattgttaaggtgtacaaaatcttggtaaatacctcaag ctaacacgtggaaaagtgggttttactaaaacccgttagagtagtctgtcatgtaaacgggta tctagtagaactaactaa	gtcaatgactcccactaaagtcggaagctt gccaaagcacttttaattagaaaagtcttta caggcttaattgattagactcggtagagaga atcttacgattctgtgaacccgttagctaaaga atacataggcacggttggatactactcagctc tgaaccactgcgtgttaataaataacacctta agagtaagggggaagtgttaataacaaac cttagcataacattgtcaagagtagcaaca ggagcaagcttctgaattatctttaattaga tacaacaacactattagaagagacatgtca catgctt
178-179	Ga0315284_10000153 GENOME_ACESSION: YL17G09_16_MG_2_SF _3300032053 GENOME_ID: 281635 CONTIG_ID: 152 SOURCE: JGI DATE: 2019-06-13.	atgcagaagattataaaggtcaagaattcaccagagaattctctagtctctgtctct gagagttctgtctaaacaaagaagaattcttagtgataagagcatacagccagtattaat caatcccgaagagaatcaagttcaacaagtaggacgctcaatacaagcactgtactgtg cctttcaaaagggaaggttaaaccttggatgcttctgtctatgctaagacaaatgtatgtaaa atcgggtgaaggctacagtaattaaaagatttccgtttgtatccagcttaattttagctgtgaa gaaaagacacaagagatttttttagatagacaggtctacgaatgtcaggtgtatcag aagaacgggaaagagagaataatgaagaattgaagtaaaftaagaaccgatgttaagt acaagattgaagaagagctatgtatcgaagaataagaagaacagattatggtacaga gaaccaagattgcacaacagaagaaggttcatatttgcctccagtcagtcacataagattg attctcatataggattctgataagatttcaaatatgttctgtatcaagagttgtgtggaat cgggaaagtttgataatcagaaaattttaaattccggaaatatcgggaaaagaataatcaaca aggaattctttatggttatgaaatgtaaaggcttatgtgacaacaagaagaatcggaaaat gccagctatgcggcaaggaatcaagtaaggtatggttttagattgcatcatataatcca aaacctaaagggaactgataaacctgataatttagctctttatagagaatgtcatgaaag actacataaaaagaacttgcacatcttttgacaaaaataagcaatacaagaatgcaca atgatgaataataatcaggaaagagtagtccgaagaacacaggaattgtatccaacagca gtaacttatggatatgaactaaagttaaaagaaacgaactaaaccttgaataatctcata ctaattgatgcttttgaataggtaacggaacagttcaagaagatgtaaagaagtaaggtg gacacaaaaaagaagaataatagagccttacagttaaacagagtaggttttaagccttcc ataagaaaaacaaaatacaaggttcaaaataaggtttgatttatatagatggtcaaccatt cggttctaaaggatgtcagaatttaggaacaagagttgcttataatgattatcagggaaga acagaactgtcagaatagaaaacatagataggttttaataatcaggaggtttttatgtatga	gtcaactaccccttggctaaagaccaaggg gctgttacggtgacgtacacagatgaaaag agttgattaggagagcttaggaaactatgcag aagattaataaaggctaaagaattcaccag agaatacttctagtcttctgtctgagagtt ctgtctaaacaaagaagaattcttagtgat aagagcatacagccagtattatcaatcccg aagagaatcaagttcaacaagtaggacgct caatacaagcactgtgact
180-181	Nocardiosis sp. JB363 genome assembly, contig: Scaffold25, whole genome shotgun sequence GENOME_ACESSION: GCA_900143625.1_AS M90014362v1_genomic GENOME_ID: 187582 CONTIG_ID: 87 SOURCE: NCBI_Prokaryotes	atggctacgttccgacagagacaagaccaccaggccgtgtctctcagcggcctgc tctggaaccggagtcggtggacacgccccggatcgggcacgaacgggactcgcaca ccaccaccagggtggtatccggtgacacatgtgcgaggggagaccaccacatccca cctgatgtcgttggcgtcaccaccagccttgcgtggggagaggccgcgtgagcggc accatccgtcttctgtcgtggacaagaaccaggtcccccttcagccctgtcaccggcca ccgatacaccgctgaacctggcccacgtccatccccgtcccgtagcgggctggaaca acgaccgcccactcatgtcgttgcgtgcggaccaggtggaaccggaccgcgac cacctaccgaaacccacacccctggacgccctggccgtgggcaaggtcgacaccacc accacggcaccgtccaaggcatcggacacaagtagatgcgtctgtccaaagagcgg acggctacggctacacctggaaggagaggggcgtttctcccggctctga	gtgaaggatccggccctgttggaccgg gctttcagtccttaggtgtagaccgtcttta ccagcaccagccatcgcctatgaggagtg acctgatggctacgttccgacaggacag aagaccaccaggccgtgttctcagcgg cctgtcttgaaccggagtcggtgacacg ccccggatcgggcacgaacgggactcgg acaccaccagggtgtatccggtgcgg accatgtgcgaggggagaccaccacatc ccacctgatgtcgttggcgtcaccacca
182-183	Wastewater metagenome scaffold_5, whole genome shotgun sequence GENOME_ACESSION: RCMZ01.1 GENOME_ID: 14570	atgtcaaatagattttgggtgtagacactaacccggtaccacttatgcccgtgccatccg cagagcgcgggagttactaaacggtgtgaaggctgcgtataccgagatcatcattc acgattatcttgaagaacggtgaaggagcgatgtacagcacatcgatatcaaaatagat ccaggtacgaagtacaccggcatggcggtgtgacacttctgtagaggacaccagcgtg tgtataggttctgcacatcaaacacggtgtgacgtgttcgacagtcgatgactcagcg agcagcgtctagacgcacagtagatctgttaattacgtttacagggcaacctagattaat aatctgactctgccaccagggtgggtaccaccctcgtatcgaatcgcgggtgtcaacgtg acgacgtgggttagacgactattcgtgttaacaccgctgacaaggcttacttcaagtg	gatgttaacgttcaagttactagtgcccttta ttagcttataactaaatagtaaccagatctagc taccgtgaggtagctactgttaggaatac ataggtaccgtgggagtcggcgtgacgctc cacgctctacgtctgtcagtttaacagacat gtgggtaacgtgtcgtgtgtacagatata aaaccttaggataacatgatcagggcgac

SEQ ID NO:	Name	IscB sequence	ωRNAs
	CONTIG_ID: 5 SOURCE: WGS	gtcaaatfcgacacgaacgccatggcgaacgtaaacacgcgatcagtagacacagacg ctcgtatccgtactcagatgcgtcactatttactgacaactcgtggtaagtgtatgttcgtactg taaggaggtctccgagataaccgggttagaacgcgaacatgtaataacctgacttagagg tggtagacagatgcactcgcgaatgctgaactcgcgtgtagacgggtgtaacttagacaaagg aacgatgttattgagcagtggtggaatgcaactaaaggacgtatagatcctctctcgtatg gcacgggttaaacatatacctaaatgtattcgtcgtatcagacattcgttagcgtacgccag tatcatgaactggacacgctacatcctcgtgaacgcatacgtgacttagggatcgaggt cgtggagtgtccagcgtgggagaccgcgtaccatcgtcatgtcggacagtagctgaaga cccactgggtcgcacgcagcgtcgtatagggtacgtaccgtacgtggtatgactgactacg atctacaccgcagtcgcatcaggatgtgtaatcgcagatgataaatccgacaagtat gggtcccttaggggacgtcctaaggaccctctcctaataatgggtttcgtagtggagata tgtttaaactaaaccaaccgggtgtaaatataaaggacatattgtgggaaaagtaacga tacgcacaacgggtttttgatataaaggctcctaactctaaatcgttctaaatatacaa actatgtaaagttaacatcctgggtggatattgtatatacaggtacaggtatcctgtcta taacaatcaaatgaattgatacagatctatctag	gttactgtacgatctcgtcgtacaaaggag ataatgt
184-185	Ga0371488_0019773 GENOME_ACESSION: MB2SIPfraction_8_\$F_3 300033983 GENOME_ID: 280663 CONTIG_ID: 19772 SOURCE: JGI	atgaacaacgtccccgaccgcgtcttctcctctacgacaacaaatcaccggctatgcc tgcctccgcggccgcgcgtatggcaccctgcctaaagccggacgcgccgactacccggc tcgtgccgtttaccatcatctcctaaagatgcacaccctgggcgacgcacagcctgtcgaac ttcgcctcgcaccgggctcgaacacctccgaatcgcgtggtcgtacgacacccgtc gtcttcgccccaactcgaacaccgcggccagccgtcaaaaaggcgtcgtcagcag cggcgccctccgccgcggccgcggagcagcgaagctcgtaccgagcaccgcg cttcgataaccgcaccgcggccggaacctggctcccgcaagtctcgtcctccgcgtc acaacgtcgtcgtggcgcagcagctcgcacacccctcgcgccactgacctatcgc gtcgaacccgtccggttcgatacgcagcagatgcagaacccggaaatcctccggcgtc cgtaccagcaaggcaccctcggcgctacgaaatgcgcgaatacctcctcgaataatg gaaccgcactcgcctactcgggagccaagaacgtcccgctcgaatcgaacacatc caagcagcgcgaacgcggcgtccgaccgcgtcgcgaatcgcacgtcgcctcgag ccctgcaacgtacgaagggcacgcagcagctcgcgtgttcttagcccggaagccga acgtcctgaagaagtgtcgtcgcacaagcgaaggccgcctcgaagacgcagcggcgt gaacagcagcgcgaagaatcgggcgacgactcgtcgcactcggcttgcgcagcag cttctggccgcggcgctaccaagatgaaccgcgtcgcacaggcgtatcacaagaac cactggatcgcagccgcctcgtggcggaacggcgtcactcgtcggatcgcacag acgatcgcgcctcttgaatcaaggcactcggccgcggatcgcgactcctatgccgacc ggacagatagcgaattcccgactcgcgcgaagagcgtcgaacgcatagccggcttc caaacaggcgaatcgtcgcgactcgtcagccgagcgggaagtacgccgaacgcac atcggcaccgtcgcgtcgcgcgcgagcgcgactcgaacctagaaccgcggcgca atcatcaccagcgcggccggaactttacgtcctccagcgaaccaacggatacgcctat ggccgcagcctga	tccaatcagggtccgtagctcagttgtgatt cggcgcttcacggcttcggtccgtgaagc aaagtgtcagggccacacgcctagcaacg tcgacaaccagcctcagaaggataatc cgatctacgatcttcgagaatgatataggca ccgtcgaatgtcttccagttcgcagccctg cgggtcagcagcagacagcagatggaccga aacaacagggtcagctcgtacgtaaaacc tcgacgactcgttgcgaggagaccgtcccc cgcaagggggtagcccggtgaagccg gatccgcgaaaagaccacaaatgaacaac gtccccgaccgc
186-187	Ga0302192_10002069 GENOME_ACESSION: II_Bog_E3_2_2_\$F_330 0030507 GENOME_ID: 280464 CONTIG_ID: 2068 SOURCE: JGI DATE: 2019-06-13.	gtgctttgtcagaggagcagacgggtcaagccgtctcgttaccgaaggacaaatcctt gagactaatccgtaacggatcggaaaccaaaggcgaaggagaaccccaagtgttctg gatcggtaaaagaaatctcctcatgcccgtcattccagcagagcaagagaacttcta accaaggtcaagggaatcgaatcctcgtatccctcgttctacggttaaaaaaccgca ccgctggatccacacaaacatagaataaactcgaaccagggcgttaagcttctggg ttagcactggctcaccaataaagcaataatctcctagtagagatcgtcagagactcagg aaatcaagaagccctcctcgaagaagggttatcgccgtaggcgtagaacctccaac cttcggagcagggcagcagcgtgctcaatcaacggagaagaaggcgtggttacctc catcccttctcctcatcaacaatcattaaactgggtaaaacgttcgcacatggggt cccttaacgggaatcaccattgaacgtatcaaatgtatccagaaactggagaaccag aatcagcggagctgaatacaaaaggcactcctcgtcgtgggaatcgtggaatc ttactgaaaaattgatcataatgtgtctattgcaacggagcaagtaacgacccaagct gacaaaagatcacgtcattgctaccacaaatggcgtagcaatagatcagcaacctgt agttgctcgtacacctgcaaccaagaaaaggagacactccattgaaagctatttggc aggaaatccccagctcctcggcggtatcagatctcctcaaaaaaccccttacaaggag cagccaaaatgaactccatcagaatctctagtctgtgaaatgaaaacttctggactaccg cttacttttagctcagagcagaacaaataatagggagaacatagatcctcctaaa tcccacgctcgtgatgcagcttcaccgggaacgggtgcagaccgcaagaactggagac aacctacttttaccatcacagctcaaggccgtggaacacatcaacgcaccaaacctgaca gggttggtatcccgctcctcctcctcccgtaaaaaactttttaggattcaaacaggc aacatcgtccaacccccattcgggtaggaagaatcggcgtccgatcaactggttactttg ccttaaacggcgaagtaaccatcaaacacacaaatgtcgtcgtcctcaagagcagac ggtacacatatacgtctcctcctccgcttaa	gtcaaccacctctccctaaaggagagcgtt gaccagactaagccactgaaacgtgactac gcacactaacaacacacaaaggataacc taagggacaacgcgaaggcgttgcacctg gcctcagtcctctcgtcgtcatttccccg aaaggggcaaggtgtcttctgttcttgcga gaggagcagcaggttcaagccgtctcgtta cccaaggagcaatccttgagactaatccg gtaacgattcggaaaccaagggcaaggga gaaccccaaa
188-189	Ga0376455_0025877 GENOME_ACESSION: Beaver2Cecum1_2_\$F_3 300035501	atgccttgcacacaaagaaaagcgagaatccttttaagggtatggaaaagcaaaatata aagtatcaccatttactatcagctgacatgaacaggagaacaaacagaggtgta gcatagggttagatattggtatgacatattcgtttagctatcacatcagaataaaat attctcaagggtgaagtagaactcagaagaatgtaaaagtcaatctgacacgagaaga	gtcaacaacccaccacttgaagaagtgggg gcttgcacaaagccatgttactagcctaa ttttttagaactacgtgttagatgttatc cctagaatgatttcttagtctagctctat

[illegible]

SEQ ID NO:	Name	IscB sequence	ωRNAs
	CONTIG_ID: 7893 SOURCE: JGI	gttttcgaaagactcctgtgcgaaagaataggagtaatcgcagtagcctaagtaagaagg gaagactaccgcttctaccgagccagatgggaaactaagttacgtttaatccatagctg gcgaaaggcttcccataagtgactatgtatgcgaagatcgtgcggctaccacaacc aagaaaacgacgttggaaftcatcttttagccgttagaagttggaagaactggttctatg atgaacttcggaaactcggtaatttagttaccaagctggctgggaaactaagaatctgcg ggatgctttgggattagtaaaatcgaataatccagatgcgtttgaagctcactgtctgc attcttgggttttagtaattctgttgcggagacaccaagctcccagataaaaaaatct tattcctcagccctcttcgattccatctgcggcaattacatcgtctccagccgactaaggaa gtcggctgattagcctacgggtgaacgatgagcttaggattgaaaaaggcagtttagta gaacatcctacttacggctgatgctacattgggtgtaataccaaaggacgtttaagctgca ttctctggaacccgggtgcagcttaaccagacagctaaaggttgaagagtgtcaattcctt agctataatagctggcgttggcgggttctcctccccctaa	aattgaaccatctaaccgtggcaggttggaa tatgtctggcagctctcccaagttgctagcct ctatgatgttcaatggcgaagggaataa acctcagtttcaatgaaagctgaggacttacc cgcaagggtgcgttgaag
196-197	Ga0394872_0157437 GENOME_ACESSION: CP_S2_MG_FD_SF_330 0037308 GENOME_ID: 312363 CONTIG_ID: 157436 SOURCE: JGI	atgtccaaaatctttgtaattgatacaaaaacaccattaaacccatccaccagcaca agcgaggcaactattaagaaacaaaaagcagccgtcttttagacgttttcttttactttgat tcttaagaatcaacccagattcatctatctctctgagattgaaaattgacctgtgtgc aaagttcactggaatcgcttagtcacgattctactggcgagggtgtcttttctgggaaaa aagttggacatatattggcggtgtagctgtaagatctcaggaagcttfaatgttcaacta agaatggactggttcaggaatcagtcacaaatattgtactcattaccacaaaggatgg ttattcctatgtgtattag	atcaacaacccacgataaatcgggggctt gaaagagcctaagttgaccagataagacc tcaaggtctactgttaaggtaaggttaag acctaccagggaatgcgtagctagttccttg ctctagaaccaaagattaaacagccttaa gggttaaacagctgtcttttggatagttaccg accttaaacattgtcgaagctaacattacca agcaattggaggacttatgtccaaa
198-199	Ga0376456_0000023 GENOME_ACESSION: Beaon1_8_SF_33000355 02 GENOME_ID: 304479 CONTIG_ID: 22 SOURCE: JGI	atgcgagttattgttagtaagccaaagaaataaacctctgatccatgtacacagagaaag gcaagggtattattaaagggaaggcaagcgaatataataataatcatttaccatcaa attaaaatagcaaccggtgaatcactgcagccgttcataattggaattgataccgggttaa acatatggacttgcagtaaccagtaaatgataaagttctttcaagggaagtcgaattaa gagaatatggagaacctcagaagatgcgaacggctcataatgcttttatgacatgtgttg aaaaagaaaaatgataggcggaagtacgtaataagaaaacacgttacaggggcacct cgtttctgaatagaaaaagccggaggtgttggctgcccacgacaacgcagcgcaagctt aattctaattcaaatggatagattattagctgagctgtacctaactctattctcacatagaa atagctaaagttgatgtacagaagatgaggtccagatattgaaggtgtcggttatcagaa cggacagactaaaggcttttggatgtacgataatttgttggcaagagataattatcatg ccagggttggaaaggaaagctaaaggacatttttaaggactcaccatctgtgtataaaa gcatgggtggaaacggacagagctgataactaattactatttgaactctgtcataccggt aagaattataagccaggtggaatcctttacgattggtgtcaaaatgaattcaaaataaac atataaggaaaccaccttttatgaataataatcagcagcagaataaagaaacagataccggct gcttatatgacttatggatcagtaaacaggcttaaacgtacagagttaaaacttgaaaaaac gcattacaatgatgctatagcaataagtggttaattgaagataftaaagaaaactcaatgatct gttttatgtaaaagcagattagaagaaaagccgacaggttcattatatgcagccatataaag gacataatccaaatcagacaagacgtatgtcgaataataatgtaaaaggtatataaaa ggtgataaaagttcaatatagaataaataatgatatgtaaccggatttaccgactcaagtc atatgtaaatgataaaaatggaaggctgcctattccggaaaataaaaactcaaggtgtaaat caataagtaagcttaattagtagtgcataacgcaactggatgtacttcacaaccctgtat aa	acctcaaatgtgaacacctacaaaaatgtgt gtacctcattctagcctaagtcttaactgacta cgttaattatgttatcacacctacggatgggtc cctagctccgtagcatctgtcaggctctgttaa aagcttcacagactcttagattggaagcagtc aacctggaggtgccgatacggcaagcatttt taacattggcgaaagggaatgctttatggc atgacaactaacatagatgattctgtagtt ggatggcacttgagatagccataaaaacat taagggtatgtttagtacccttttacttttta gaaaggagcttttagtagtcga
200-201	Ga0370511_0001040 GENOME_ACESSION: Frozen_pond_0310_2_SF _3300035210 GENOME_ID: 304715 CONTIG_ID: 1039 SOURCE: JGI	atgcaagagttattagtagttagtaataacaggaatcaacttatgccgtgctctcagcgag agcgcgatgctgttgcggaacaaagagccgcgtcttaccggaagtatcccttcacga tcatctcaagaccgagaggaagcgcgtatcacgtctcagttcaagcagatccc ggcagcaaggtcacaggcatcgtctcgttgcggaattcgaacgagaggtgaagacagt agtcttggccactgagcttcacagagggcagatcgatcaaaaggtctctgattccaga tgtcgggtccgagaagccgcagaacaggaagacaaggtatcgtgtccacggtttga taaccgaacaagaccctccggtggtgcccgcattctaatgtcccagcctacaacgt tcagacccttagcactgagacttcagcgttttctcccccttcttctattgcggtcgagaccgt ccgtttcgacatgcagaagatgaccaaccccagatattccggcagtgatcagcagg gcacacccaaggtctacgaagtcagagatgtctgttgagactggaacagacaatgt gcttactgtggcaagaggataattcctttgcagattgagcactgtttccgcgagccgag gtggaactgaccgggtgtcaaacctcactatctgtgtgagacatgaacacgaagaag ggaacgaagactccgcccaggttcgggtccccgataccagaagcaggttcgac cgcttaaggtgctgcagcagtaacgcgactcgttatgccattggggatacgtcaaga ctcttggactgccagtatgttctgtcgggtggagcagaagttcaatcgacccagc aaggctatccgaagtcattgtagcagcggcctgctcgggaatctggctccaatg tccatctcgtatcctaataatttactgcttaccagttaaagcttgcggtcatggatcaggcaaa tgtcccgaatggaaggtcgggtccctcgtacttcagccaaagcgtctgtgtgtcca aggtttccgaaccggagataattgtaagacgattattccgtctggcaagaaggtgggcaa ccatttccgtaaaagtcgctattcgcacttcggggagcttfaatatttcaactagcggcg ttgttcaggggcatagtcacaaactcgtcagtagtgcagctgtcgtgacggatattcgtat ccccggccaataaaggaggtctcctcagcaataaattaccgctatccgcgccca ggaggtgttatga	gtcaacgactcgggattaagatcccagcat gcgaagtcagtcagctcaagtttgccagg ctcagtcctaattgttaggactacgttaccgg cgaatcacataggcacttgggatgcacgcc agtcacagcctcgcggcagacagttaaac agggtcaagagttaagccagattgtttgcat ataaaaccgtcgataacattccgaggca cacattaccgcgcaagcgggaacgatg gagtaatcgtcaatgcaaaaga

98

SEQ ID NO:	Name	IscB sequence	ωRNAs
	SOURCE: WGS DATE: 2019-01-27.	<p>tcaagaacccgggcattgaatgcacggagtagacacaggcgaccagctcgcttctg gaacgtccgagaatacgtctctccgcgacggccataatgtccactgccacggca aatccaaagatccgattctcaacgtccaccacctgaaagccgtaaaacgggtggaac gcgccgaacaactgtatcacgtttgcgagacgtgccataaggcgtaccacgagaaa gatcgttccaagcagaagcgtggacagtccttccgagacgcccttcatgggaattat gcggtggacgttctcaatcgctgaaggcacaggttccggaactgaagtcggaaca cctacggctacgtgacgaagaacacgcgcatcggcacgattggaaaagtcctcat acggacgcttctgattgccggcaacttcgaagcaaacgattgggtgagtacttctcc agaagcagaccgcgaagcaaatcggcagattcacaagatgtcccttctaaaggcgccg gtccggaaacggcagcaggtcctacagagtgaaagggttccggtgttcgacaaggt gcgtttcaacgggtcaagagcgttctgttccggtcgcgcaggtctgttcgttgatgtccg gacactgctgggacagaaactctctgccgaggttaactgtaaaaaactcggctgttga aaagcgacggactttctaacgcaacttttaaggagaacgcgattcctatgaagaactt cgatttgacgtcaccggcatgattgcgccgctgttccgcccgcgtggaaaagcgccg ccgcagcacggcgtgtgacggacgccgcggtgaatctttgaaaaatcgttggctg ccgtttgctggttccggtgacggcgaagcgtgacggcgccggttccgaagccgtt gaaaaagcgggctacggggctcgtccggcggaagaccgaggacgccccaaagc aacgtcgcataaaatgaggcgcagaaagcggcgacggcggcggcggcgttaa aaaaagcggctttgtcttccgtcgttctgtctgattcttccggctcgcgtgggcccga tgatcgggggtgacgggtccgggactgtatccatgaagaaccccgccggggtggggct cgcgcaattatctcgcgtccccgtggcgttttaaacccgcaatttttgaacggcgc aaaggccctttaaaccgtcacccaacatgatacgttgggtgcgacgggttccgggg cgtcgttctgttcgggacttttgcgttcttccggtatgtcggaaagtgacggcggaat cttgcggcggcgacgactacgcgatgaatctttttgattcgtcggcgatgattctac cttgatcacgtggggaattttttagggcccgcgcaaaaggcaagaccacgagccca tttcaagcctcatgaaattgtgcccgaaccgagccgtgcgttgaattcggacggcgcg aagaattgtgtggcaaccgacctgcgggttggcgacaaactcgtctcaaacccgg gagcgcatcgtcgtcgcggcgtgattcgaaggcgccgggaacggcggaacgaatcg gcgatgacgggggaaagcctgcccgtcacgaaaaagtcggtgaccgggtgtcggc gcgacgctcgtgacgtccggcgcttgtgatgcggcggaataagtcggcgaaaca cggccttctgcagatcatccgttgggtgatgaagcgacgtcgggaaaagccctgtgt cgagggttagcggacaaagtgaagtcggtcgtccg</p>	
210-211	Chloroflexi bacterium isolate CF_154 14_0903_05_20cm_scaff old_224, whole genome shotgun sequence GENOME_ACESSION: GCA_005879655.1_AS M587965v1_genomic GENOME_ID: 285751 CONTIG_ID: 235 SOURCE: NCBI_Prokaryotes DATE: 2019-07-22.	<p>atgtcacatgtcttctcatcgataccgacaaacagccgttctcccgatctggggcg agcacgcctgtactacaagagggaaggtactgtttacacgtcttccgttttgagggtg cgtgaatatgtattgcgaagtggaaccgcgtgtgctactgtggagccaggacgtg ccctagagctggagcatactgtccacgcgcgggtggcacggatcgcacatgataa tttgtccctagcctgcgagtcctgcaatagacgaagggaacgcaggatattagcactt cttggccgatcagccccgcacggctcgtccgtgtactagctcagacgaaagcgccctca aaaatgccacagcggtaacgcaaccgttgggaactatcgcggcgcttgcagccact gggctacccctggaacgggttcagggtggacagacaaactataaccgagtgctgctg cttccaaaagcccactggacagatcgccgctcgtgtggagcatcaaccctacgcctc tttccactgagggggtaatccattgtcatcacggccactgttcacagtagacgaaaat gtgcaatacaacgatttaggtctccgaccagtcacgcaagggtgcaagcgctacttt agctaccagactccgatctgtgagcgagcgtgtgtccagatcgaactcaagtcgcgg ggacgcatgtaggcagggtgacagttaaagctcgagaacctttaccatcaaacccgg catggcaaatcactgatgtccacacccgtttgtcagcccgtccatcgtcgtgatgat attcttattcctaagtgtgagggttccccaccccccaacacccgaaaggagcgctg tttctcctccgctag</p>	<p>gtcaagtcccccactgtgacgtgtcacagt ggaggcttggaaagcaagactttcaagcccg aaacttgagcagacagaggtttgaaagaca accacgttatacagaaggtgtaaaagaacct acctacagctgttcaccagcttgaacctta gaaccgttcagttaaacccgctagaggga gaaacagtgctgagcggaagtagcgac tgataaccgtgtcgaggtgagtagtaccctgg gaaacagaggccccacggggcacaaaa caaaaagaagggaacgaacagcgtcatgt cacat</p>
212-213	Ga0376455_0000343 GENOME_ACESSION: Beaver2Cecum1_2_\$F_3 300035501 GENOME_ID: 304484 CONTIG_ID: 342 SOURCE: JGI DATE: 2019-11-15.	<p>atggtcattacattagacaaaacaaaagcctgttgattttgtacggaacccgtgcta gaatcctgatgggtaagcggagagcgtgtttgtatcgtcatttctctatcatcatcctca aagacgttgatgtcgggacttagaatactgcacagctaccgcatcaaaatgaccccc gctctaagtacaccggcattgcgattgtagacaacacagacaatagttgtgtttaccatg cagattgagcatcgtgctactacaatcgtaaagtcattaaaaacacgcaatgccgtacgta gaaaccgcagaaatcgggaaacccgttatcggcgttgcgaagtgattaaactactacg aaaaagggttagccgtacaaagcagattcggcccgccctgatggttgggtccctcgtct gtcaagtcattgtgacaatatcatcaactgggtcaagcgggtgtgcaaatggattaaat tacagagtgacgttgcaggcagtgctgttcgacactcagttgatggacaacccgataatt gagggtgtggaatatcagcaggaactctgtacgggttacgaaatccgtgagtatctgtg gacaagtacaagcatacttgcagattgtcaacggcgatccaaagatgatgttctggaat gggagcataaactcccaaatcagaggcggaagtgattcgggtcaagaacgctacccctt gcctgtcacaagtgcacacaggaacagggaagtgatgaccccgctgagtggttgaag ttgtcaagaacacgcttctccaaactcgtgatgctgtattcagggaatcagcgtgtc attgataaaaaaactgttagcaatcgttattgtcgttgggtgtcgaactcgtagata catcgagcgatttctgttgcgttggcagatgcgaatgtccagcggcgtagaacca</p>	<p>gtcaacaactcggctattgaataaccgagca tgaggacggggaatccctttaaactcat gtgaggtgtttgatagattgttgagcagact aaggtctctcgaatgagagaactcgttattt gagtgccgactaggtgtcttctccagca ccagtatccgaggctggcagctaaacagg caatccgagatgaatagccagtgctgtcag cgtgaaacctcagataactgtcaggggg acgtgctttttccacacaataagcccgtaa ttgggtatttaaggagcaacttatg</p>

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aatacaaccgcacgaagctgggacttccaaagaccatcactacgatgctctgtgtctggtacagtg	
214-215	Leptolyngbyaceae cyanobacterium CCMR0082 Scaffold_1b, whole genome shotgun sequence GENOME_ACESSION: GCA_011009535.1_AS M1100953v1_genomic GENOME_ID: 325887 CONTIG_ID: 11 SOURCE: NCBI_Prokaryotes DATE: 2020-04-22.	atgtctaactatgttttgaattgaccaagacaacaacccctcaaccctgtgccccagg tctgggccgaagatgcagaccgaccagaaggcggaacattcaaaagtatccgtac acgctgatacttaatactctgtgtgctgcgccaaggctaaaggcgtaaccctcaagctc atcctgttccaagttttggggctctctatcttggatggcaacaagtgattggcgcc gaactagaacatcggggatggggatccatgaaggacttgagaacgttccggctatc tcggaatagacgtcactgaagactggtaccgaaagaaccgggacagtttgatttaa gcctgatggatggctcccccatcactgttcacgggttcaaacacaatgacttgggtg aatcgaactgcgaagcttgcctccattgattgtcaagatggagctagtcaagttgaca ctcagttgatgacaatccgacattgaaggacccagctaccagcaaggacattagcc ggatataccgccagggaatacctacttgaataatggggcgagctgcgttattgcag caagtgcagcgttctctgcaggttggacatgttaccaccaagtcgaagggggatcaa ctcaatccgttaactgtgcttgcctgtgagaaatgcaacaccgcaagggtagtaagct gttggagggttctaaagaacaacacctgaggctctaaaggattcaggccacactcaag ccccctctgaagatgccgctgttgaatgcgaccggttggcgcttaccatgccctg caagacacagggttccagtgcagccgcacgggagcgggacacaaatgaacgg ggctaaacaggggctacctaaggagcattggatcgtgcgcgtgtgtgtgtgtgtgtg gagaaaacctcgatataagaccgtcaaccgtgagaattcaagcaaggggcacgg tagccgccaaatgtgaagggtacggcacaaaaaagacgggtgaaccgatccgagga ctgccccaccatgctccccggcatgcttaagcgcagaaggacggcaccggagag gcccccaagtttgcctgcagtggttctgttgatcaggcgacattattgatgccaaatcc aacgggcaagtacaaggcgatttacaaggagtcagagtgctgtcagggtgacggtc ggatagccataagaccgaagggtttagctctaagttcgtatcaccaccgcattctatca caagttgttcaaaaggaaagatggctgacttactcgttcaatga	tctgactgctcgaccctatcgggcgagatga gaaagccgttagcttaccagctcaccgggtt attggttttcccgtaaaaggataggacacc gcaagggttgaaaaacggcaacacctctct aatcgagagcgttaaccagctcgagagcca gttgtgtagagggttggccaaagtaagcg aggtaacgggggtgactacgtgcgaggg cagagcggaatgctcaaaagctgcttatccga gcgagggtcacattaccgattatcgaggg actcttatgtctaact
216-217	Scytonema sp. HK-05 NIES-2130 Scaffold_41, whole genome shotgun sequence GENOME_ACESSION: GCA_001904675.1_AS M190467v1_genomic GENOME_ID: 113419 CONTIG_ID: 115 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	cgctcaagaaaaactagatatagcacacgcttttctaatctagccgtaggcaagggt ggctagctctcttacttaccgctttctagtatagaacttgggtcaaaagactttgcaa atattcaccatcacagaaatagtgatggaattagtgaggtcgtatcccaaaaaatgcaa gcagaaacgatagaagggtgtacagatcagcaaggaaactctctgggttatcaagttcgt gagttatctgttagaaaaatggggcgttgcgtgttactgtatcttctgtgttccacttc aaattgatcacatacaacaaaaagtaaggaggcagcagatagaatttcaacctaacct tggcggtgaacgttgcattctcgcaagggaataaaccagttgaagactttctaaaaaa agattctgcgcgacta	cggttaccggcaatccccaccatgctcgca atctgggattggtactgggagcgttggcaca gggtgcaacaatttgggtctatttcttct ctgtccattgtgattccacatttccggcg cgatttagtgaactgctaggaaatagcgg gtggcgattggcttcttcccgcaatttgc caaaactttttagccggcattttacctgtg acagaaccatttcaaatgatgaccagatag tctttaaaacttttaggggaactgtagaagt attcagacacgagctacaacaatcagaacct acgatg
218-219	Ga0310150_005743 GENOME_ACESSION: K76_2_\$F_3300033987 GENOME_ID: 280576 CONTIG_ID: 5742 SOURCE: JGI DATE: 2019-06-13.	atggtgtacgttatttcaaaagatggtaaaccttatgccacaaaaagacatggtaaagt tagaagacttttaaaacagggtctgtcaaaattgttagaagaacatttacaattcaact attgtatgatactacaacctatacacaaactgttactgcggaatagacattgggtcaaaagt agttggtattcagcagtaacaacaacaagaattgtgcagtgctgaagttgaactaga caagatatcagaagttgctactgaagagaagagaacacagaagggttagaagatatcg caagagaagatatagaaaacaaagattttgaacaggcgtcgatacaaaagatggctgc accaagtttcaatggagggttgatgcacacatcagactggttaattatagccaagata ctaccagttactaaagtgtgtgtaaatgacacatttgatacacacaaaaatgtaattcca gaagttagtggcaaaagaatacagaagaggaccacaaaaaagtttagtgatgtaagaga atactgctgtggagagcagggttataaatcagaactgtcgggtaaaaaaggcatactgaa gtccatcatatttccaagaagtaagggtgtactgataatccctctaatttaatagtgttaa ccgttgaggaaacaaagcaatcatgaggggaaaattaaatccacgtagtaaacfta aaaaagttaaaattctcaagatgca	gtcaactaccaccacttaagcctaacggc tttgaagtgggggttgcgaagccctagtgtg actaccctcagccagggggaagttaattctct atcgactactgttagctgtcatgacacc tgggtatgctcgaattccaggctctgtctg atgtacctaaacagtcctgagggttaggga cagtggtacatacctaaccagccagcttaac attggggatgagcaccttaactctcaaaagg aggcttaccacatag
220-221	Leptospirillum rubarum LeptoII_Scaffold_8524 genomic scaffold, whole genome shotgun sequence GENOME_ACESSION: GCA_000205145.2_AS M20514v2_genomic GENOME_ID: 242318 CONTIG_ID: 3 SOURCE: NCBI_Prokaryotes DATE: 2019-05-30.	atggcccgtgaaggcatttgcaaggagagatcgtggctgtgttgttctggacaagaaa aagaaacccctgatgcgctgctcgaaaaacggggcagactcctcctggaacggaaaa aagccgtggttaccggatggctcccttaccatccggctgaaagaccgggtcggggg cgtgacccaacccgtccgggtcaagctcgtaccggatcgaagagcaccgggctcggc gtgttccgggagggagggggacgggaagaccaccgccatgtcctgtttcaggcg gagatccatcaccggggagccgccattaaaagaaactggaccagcggcggtcctcc ggagacggcgaggggggaacctccgggtaccgaaaccccggttcgacaaccggaca cgtccggagggctggttcccccgagcctgcgcaccgggtcgatacaccctctctg ggtcgaccggctccggagacgggttccggtcacggggtacggagcgtgtgtccg gttcgacatcgagaagatcgaaaacccggaaatctccggagtcgaataccagcaggga accttggccggatagagggtccgggaattctctggagaagtgggggccgacctgtgc ctactgtgggtccgagaacgtgccccctgagatgaccacatccacccccggagcctg gcggatcgaccgggtctccaacctgacgtgcttgcgttccgttctgcaacctgaagaag ggaaacgtccgggtcgggagtttctgcggaagacggcgagcgtctgtgatgacct cgccccggccaaggccccgtgaagacgcggcgccgtcaacacgacccgggtggg	gtcaaccaccccgccctgaaggcgaggagc ttgaaggaggttcgacagcgtcgggttga ccagggaaagcgggtaccaccccgctccg ttggcaacagggtacaagaccaccccggg atgcttctcagtcgccgggtctgtggaagccg ccgatgcagacaaccgcgaggggccaggga cgaaacgggtcgccgcaaggagcgtatc ccgaagccgggttccaacattccgaggg gagacgcgtcgtgaagacgcggttcatgg cccgtaaggcgatttgcaaaaggagatcgt tggct

[illegible]

102

[illegible]

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aatctctctccttgagagagcaagacattgctaataagaaaggagataggctctctcc ccatgtctgaaggcaggggtctccgagacctga	
238-239	Ga0257068_1000081 GENOME_ACESSION: GoafecGen5Rep2v2_SF_3300023705 GENOME_ID: 280338 CONTIG_ID: 80 SOURCE: JGI DATE: 2019-06-13.	atggcagttttgtactggaccgtcacaaagaccgtggaccgtgcagcgagaaacg cgccagacagcttctggagcgaggacgcgcacgtgtccacaaactgaacccttcacg attcgcatcgtggaccggctttttgaaacagctgcgtcaacggcgctcggtcaaaattg accccgagcgccggagacgggcatcgtctgtccgcgaggacgggtgacggagcc cacgccctagcgttcatcaatctcagacaccgcggtctgtcatccgaagaagctggag cagcgggccgcatatcgcgctcggcgccggtctctgaaatctcgggtaccgcgcccgc gttttaacaatcgtcagggcccgaaaggctggcttgcgcaagcctgcggaccgtgtc gactccacggctcgctggatgcgcagactgtccgtattgcgcccgtgcgctatctcg atggaactcgtgaagttcgacatgcagggccatgcagaatccggaatctccggcgtcga gtaccagcaggagagctcgcggctatgaggtgcgcgaggtacgtgctgaaaaatgg ggccggaagtgccctactgcggcaaggaaaacgtgccgtggaatcgagcacatc acggcgaatcgtcggcgcgagcaaccggtttcaacctgacctcgcctgtcatga ctgcaaccaggccaagggaacatgccggctgaggcggttctgaagaacggccgga agccctggacaggatagcgcgagggcaaacagccgctcaagacggcgctgcgggt caatgccacacgtcgggagctctacaggagctgcaggtcttcggactgccgtggag actggcagtgccggcaggaatggaaaccggacgcgtctccatgtgccaaaggcg cactggctgacgcccctctgcgtcggcagtggtgacgccgtgtccgggattgggaagc cagtgcttgagattgcctgcacagggccggtcgcacagcgaccccgctgcacag gaacggcttcccgcgtgtgtctgcggcagaacgtgtccacggctttgcagcggg cgacctggcgccgctgtcgtgccaggggcaaacatgcagggaacatgtcggaaag actcgccgtgcgagaaacgttctctgtgtgctgcggcagacgtaagcacgacg gcatatcgtggcgacactgcaggtcgtcgcagcgtgccgacggctatggctacgggcat ctgctcggtaacgtcaaccaggaaaattttggaggcagcggtttcccccggcttga	gtcaatcaccggccctgaaggacgaggct tgtaaagctgaatcccagagccgattgac cagcccaagtgagggtgaaccggactccgt tgtatcagggctaaagaccaacgccagaat gcttctcagttctgaccttgaattctgtt gcagacatgcccgggtaggcggaac gggcaggagtgacacgggacagcaaca tgggcgagaggagccggcgggcaaccgc cgcgtaacaggcccttgcggggcagga aaacgacaatggca
240-241	Desulfobacter sp. isolate S02.Bin034 NODE_29_length_48907 _cov_12.536371, whole genome shotgun sequence GENOME_ACESSION: GCA_011391705.1_AS M1139170v1_genomic GENOME_ID: 329340 CONTIG_ID: 28 SOURCE: NCBI Prokaryotes DATE: 2020-04-22.	ttgaacgtattgttttagatacaaaaaaaccaaaatccagtgcatccggcaaaag ccagattgctttatcggagagaagcgccgggttttcaggcagatctcttcacaattattt aaaagaagagattgggtaaatccaaagcacttcgggttaaaatgatccaggcagca aaacctccggcattgccgtcactgatgatgccacaggtgaaatcgttttgatgcaatta gaacatcagaggccaacaaatcaaaacgatctggaatccagacgggcaatcagaagat ctcggagaacacagaaaaccgatacagaaaaccacagattgaaaacagaaatcagacc gggaaggatgggtggcccatcattgaaagccgggttcacaaatagagacgtgggtcaa tcgattatgccgattcagtaattatcaggcaatcctcaatggagctgttctgtttgatagcag aaaatactcaatcccgaatctccgggttgagatcagcaaggggcactttcgggata gaggtacgggagatctcttgagaagtgggcagaaacgtgtgcctattgtggcaaaacg aatctgcactggagattgagcacattctcccgaatcaaaaggcggttcaaatcgcgtc agcaatttgacctggcatgtaccgaatgcaatcagaagaaagggaacaggccgattga catctctcttcaaaaggccggaattactgaacggattcaagccagggcaaaagcacc gctcaaggatgcggcagccgtcaatagatccgatgggatctattccgcactgaaag agacagggtcctctgtggagattgttccggcggttgacaaaattcaaccggacaatta gaggggtttccaaaacacattggctgatgcggctgtgtgggaaaaagcacactgaaa aattattcaaaactgacaaggctgttttgattgtaaaagcaaatgtgatgggagcagcaa atgtgccgggtcaatacgtttgattccccagacaaaagcgaaatcccggaaagaaaa agttaacgggttttcagacaggtgatattgccaaagcaattgtactccgaaaaaagggtt gaacatattggctgtgtcgggttagaaaaagtgatttttaacattaagacaagagaa acgaccatcaagggtattattggaagtattgtcatatgcttcatatgtgatattcat ataacatttga	gtcaactacccctcctgaatcacagattcag aaggggctgtgtaaaaagccctattgttacc gtctaagtgcttgcgactacgttagatcg aaataggtaccctggggtgttccagctcc aggcgtactcggcaagtgttaacagggtg aagaggttaaacgggtgtgtgtgcttaa cccggtcataacattgacaaggcaaacatta ccctggaacaggagaatttagattgaac
242-243	Anaerobic digester metagenome 6175, whole genome shotgun sequence GENOME_ACESSION: MTBK01.1 GENOME_ID: 1684 CONTIG_ID: 79072 SOURCE: WGS DATE: 2019-01-27.	atgggcattatgcgctggacagctacggtaaaccaagaattataccctaattgcaaat nnnnnnnnnnncatctcggtctatcagccacacagataagaagcttattgaatctga agtcaattaaagaaacgacatagtaaaactcctatcaaacagagaagaaattaggagag gcaggagataatcgaacacagatagcaaaagccaaagttcaataaccgcaaaaacct aaagggttggttagctccctcaatcaggaacaaaattgatactatctgaaggttatcgctta cgttcgaatatactcgcgatacaaaaactgaccattgaagtagcacagtttgataatcaaaa gattagaagcccaacatcaaggcaaatcaccagatagctgtttcatggcattatg cgctggacagctacggtaaaccaagaaattataccctaatgtcaaatgacctacgggt atataactaaacataccagaattaacgcagggttgaacaagtttcatcgtacagatgcacg ctgcacagcggcaaccatcagctgaaccactaaacatctgtgactacttcaagcaagt aagaggccagaaccggcagctgcacaagccaaccgaaaaaaggcattcgcaagg ccaacaagccccgcgatacgttcacgggtttcagttatttgacaaggctgtatcaggg acaagaatgtttatcttggcagacgatcatccggctattttgattaaaggaaattgaggtg agtgaagtacacgcacatcgccagccataagaacttaattgctagaagcgcaaatata ttattatcgaaagagagggaagtgcttccctctacgactaaagtcgggggtatagcc gcgtaa	gtcaatcaccggccctaaaggcggggg cttgcaaaagcttgattgactgacctcagct taattgactcgttatatagaatcatagttac cctgggatcgtgcccaggttccaggcact aaggctgggttaacatagttctgagggtga ggaacagtgctgtggtacaaaactaatat aacattggcgatgggcaaccaactccaaa aaaggaggtaggcttaagtcta

SEQ ID NO:	Name	IscB sequence	ωRNAs
244-245	Fischerella sp. PCC 9605 FIS9605DRAFT_scaffold 12.12_C, whole genome shotgun sequence GENOME_ACESSION: GCA_000517105.1_AS M51710v1_genomic GENOME_ID: 55988 CONTIG_ID: 3 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgaactcaaatgctcgaattccagttttgtcaccagatggttaagccattgatgccaaactt gtatcgacgggcacaggtctgggtagaacaaggcaaaaggatggcaatgac ctgaatattaagcaagttcatctgttacaagaacctccggtaacgcaactcagccggtg cgatcggcattgaccctggtaaaaagtttgcaggaaattgctgtcaatctagtcagttcattt gttcgacgacacttggtattgccgttccctaatgtcactaaaagatgacggggagacga attttaagacgtgagacgctcaagacgtatcaaccgcaaaattccattccatttaggg cgatcgtcaaaaacgcttgacaatcgacggcacaacaaactgttccatctatccgtgc taaccgtgaatttgagttgcagttgtcaaaagattgatgcgactatccagtcacaactat tgtctatgaatacatagaagcaaaaggggataaaagcgttttagtctgtaattggtgggtcaa aaagtcattgttgagttcttaacagaactagctccagtggaactgttttgggtggcaaac cgtaacctcagaactcactgaatttgattaaagcacaacaaagccctcacaatctct caatctcatggagtagacggaattgctgttggcttctagccaattcgtaaactacgaagcctt ccaaaccaagcgagaacatggcgctgctgggttggaaagcgtcgtctgacgcttctc cgtttcgggtgattactcgcctcaatctatccgtccagttgcaacttgagaaattcaggc aaggcgctgtacgtaagcgaaggggggaactgttacgccattcggattcgtctgtgga tttcgttcaaggtgaaaagcgtggaagattcatagaggttgattggtggttgaagtga	gtcaagaacccccacctgcactcttgaagt gggggcgtgaacagccaggtttaccggt aactcctgactagagccaatgagccactatc tcgcacggactccgctatttccctagagcg gattatctccaaactatggctgtaggtgct tgagggtcgtcaaatgattcggaggcaac ctccgaatcccgacctccagaaaggacat cgtgatagtgtggtgctgaagg
246-247	Ga0187846_10005139 GENOME_ACESSION: TC_taG_2_\$F_33000214 76 GENOME_ID: 35912 CONTIG_ID: 5138 SOURCE: JGI DATE: 2019-01-27.	gtgttcgtttactgattaaattgtcagggtaaccgttaccgttagaccacgaaaagc acgcctgctctgcaagagggcgaaggcgaaggtgtgagaattgttcggttcacatcca actgctgtatgacagcaggtgctacaacaagcaatctcttctgtgctgacgagccac ccaacggatcgggtttctgacacactgagcggcaaggtctttcgaagcagaggtgc agccccgaactgacatccagcattgctggcgaccccgccatccgccatgccagg cgagtcgcaagacgcgtatcgccagtgtcgttctcctcaatcgaagaagcgatcagg ctggctcagccttctgtcgccgacaaagttagcagcccatctgaaaacaatccgtttggt caccagttggtcgtgtgagtagaccaccatcgaggtggcacagttcgacatccagaa aatacgaatcccagatcgagggcagagagtagaacacaggtcccccagctcgtgttct ggaatgtcagggcatalgttctgcccagagaccgtcatgtgtcaatggtgtcaggggca gtcgcaagacccttctcaggtccatcacatcgagtcggcaagacagcgcgggata ggccccgagaattgtgactttgtgtagacgtgccaatgatctcattcaccgaatgcacca ggaacacacgatcgaacagaagtcagggggtttcagatcgccggcgaatgggcatc atcgctgtggcgcatctatgacagcggaagcactctcccgatgtccatctgacatag ggtacatcacaaagcacactcgaattgccaaccatttgaataatcgatgtgatcagta ctcgctgtatcagtggtcatcgctggcatgctggatgagacgtggtatctgatcaagta cgtgaggagaaataagcgtcagttgcacaaagcgaccatcagaagggcgcgcaagag gcagtgctacacggctcccaagtagtccacggcttgcgtttgattgctgtaactatc aagggaagtgctgtttgtttgttcagagcgcagttctggtacttgcgtctgctgtg acgggacgaag	gtcaagaacccaatcgccctcagcgatgg gcttgtgtgaacaagccatgctgactagcct gagcttagactacgtcgacaggtcatgcta cctacgggtgcctactctagcctgtagctcta gcgtctgtgattaaagcctgtgaggtagg aacgtgtgtgcagacagcaagcgtgtg aacctggcggaagtagtaccatcatccgaa aggatggtaaacgtgtgtctc
248-249	Moorea sp. SIO3G5 3G5_NODE_179, whole genome shotgun sequence GENOME_ACESSION: GCA_010692495.1_AS M1069249v1_genomic GENOME_ID: 325541 CONTIG_ID: 878 SOURCE: NCBI_Prokaryotes DATE: 2020-04-22.	gtgaagccagtgctcaggagcatgcacaagctttgatagccgagcgaggcacactttacc ctcaacgggagtaaacacatcgcccatgtgtaattacgtcttcttttagatgcaaatcataa accctcaacccctgccgtccgttaccgccaggaaagctgttaacggctggttaaggcgg ctgtatcgcgtttacccttccactatcttgaataaagcaggtagagcgcaacccaa accgatgtctcacaagatcgatcctggctcacaagatgactggtctagctattgtgtacggga atcaggtgtgtgtgacgtgaaatcgaacatcggtgctcccaaatcaaatcagctcttga ctctcgtcgggcagtgctgctgttccgtagaacacaggaaatccgctaccgaaagccac gttcaacaaccggaaacgccagaaagttggtgtgctccgagcctacagcacagagtg gcaaccactatgacctgggtgctgagattaattaactacccccattggtcgataccca ggagctagtctgggttgatacccaaaagctgcaaaatccggaaatctctgcatagagtac caacagggggagctgatgggtatgaagtgcgggagtagtctgtatcagaatggggctc tcaatgtgttactgtggggcacactcggtcaagctagaggtggaacacatagttcccaaa tcaaaaggaggaaacacgtgtcagtaacttaacctagcctgtcatcgggtataatcagg ccaaggaatcttaatgctcaggaatttttattgggg	cagaagcagacaattactaaaccttagcaa acctgctcaaatcgggtttacctgagcttcc agattggtaagctcacaatggttaaacgc tatcaaggtcaagatcaccctgggtgagag ccagctccaggtctatcgtctgtattaaac aggtgatgatttaattttagtgaagcaggtg caggagcatgcacaagctttgatgccgag cgaggcacactttacctcaacgggagtaa acacatgcgccatgtgttaattac
250-251	KEY_1011702 GENOME_ACESSION: IMG_3300000938_\$F_3 300000938 GENOME_ID: 308319 CONTIG_ID: 3733 SOURCE: JGI DATE: 2020-04-07.	aaaactcgtatcgtaaccacgctttgataaccggaacgcctacaggggtgctccca cccagctcgtcaaaagccgagttgagaatctgtaacttgggtctgagttacaaaagctag ctttgatttctgataatttctcaggaatttagtcaagtttgatactcaactatggagaatccga catcagttgattcgcataccaaaggtgagttagctggttacgaagtcgagaaattctg ctttcaaaattcaatcatacctgtatgactgtggggctaaagatactgtttgagattgagc atctttaccacagctaaaggtgttcaaatagaattagtaatttaggcattgctgttaggaa ttgcaaccaaaaaaaggccgtcaagctctcagagaatttttagcccaaaaacctgactg ttacagcgacttctgagacaagtaaaacacagaagccgacactgctgccgttaattcta cccgttggcattatttaacgactcaaaagagactgga	gctatccgcgaaccactgacaaaacctgga caatcccgaataactcgttgtagatagcc cggttaagctcaaccacagactacgttaaga aggacacacacctgtgagtgcttggcag ctcctgtctgtcgtaccagttaaacatctt atttcgtaaggaaagtgctgtagcatgaca agccctcttaacattacaaaggcaaaagnn nnnnnnnnnnnnnnnnnnnnnnnnnnnn nnnnnnnnnnnnnnnnnnnnnnnnnnnn
252-253	Ga0172380_10006798 GENOME_ACESSION: Leawe162metaG_2_\$F_3 300014205 GENOME_ID: 31497 CONTIG_ID: 6797	atgcagaagttaaagcaaaagtaaaagacttaccataaatgcttcactagttgtgactct acaaacacacaattaacagagacgacagcttagttgtgtgagtggtcaaaaattgac aaactgacttttaaacatcccgaagtgaatcaaccggaggaacactccgggacagaggt gagactacctgtcaaaagttttgtattgaatctttaggggcaagccccgatgccttgcctatg ccgacagggcaaggtgtctttaaagagaaaagggcaaaagtagttaaagggagctctt cacattacagttgttagttccaacaggtgaacaaaacaaagatattgttttaggaatggattc	gtcaactacccccatacaagttaaagggctt gtccggttaacggacagacaaaagttgatt aggagcgtttttaaataatgcagaagttta aagcaaaagttaaagacgtacctacaatgc ttcactagttgtgactctacaacacacaatt aaacagagacgacagcttagttgtgagtt

SEQ ID NO:	Name	IscB sequence	ωRNAs
	SOURCE: JGI DATE: 2019-01-27.	aggttatcaaatgttggaattctgtgtactaaaaaacaggaaattatgagattgattgtgttttagaaaatggaatgagtaaaaggcttgagaaaaaacatgtacagacgcgcgagaagaaataaattgtgtatcgtaaaccgcggttatgaatagggtttcaactaagaaaaaaggatgggtaccgccttcaactaagacggtttgaacacatataaggcttattgaaagtattaaaaacttttaccataacaactgttagaattgaaagggttaatttgatattcagaaaataaacaatcctgaaatttcaggaagattaccagcaagggttcgatgtatgaatatacaaaacaaacgaaattattatgagccgtgaaaaatggaagtgtcaattttgtgaaaagattttaagggtcagtcgagtcataacacataaaccacgaagtaagagggtactgacaaaactaacatctggcaatctacataaaaaatgtcatgaagaacttcacgccaaacatttagaaaaaacactgaaaaaaaataaacagttcaagatgcgacatttatgaatattattcaacataaatttcaagaagtgttagattgtgaaatcacttttgggtatgagactttcataaaagaaaggagtttagaattttgaaatcaattcaaatgatcttttgtaattgcaaacgataactaataaaaaagggttaaagaattcaagtaatacagaagaagaaaaataatagattttacaattaaatagaaaggttttaagccaagcataaggaagaaaaatctaaataagtcacatgatttattttggattggtaaaaaacagtatactgtaaaggatgcacagttatgtagatatgtctttggggggaataagaaaaaagaatattgttagattttcagatgtaactaaattttcgttagtgggttagtatggatataa	ggctcaagaaattgacaaactgactttaacaaatccccgaagtgaataaccggaggaaactcgggacaggaggttgag
254-255	Ga0114359_1005163 GENOME_ACESSION: IMG_3300008122_SF_3300008122 GENOME_ID: 305744 CONTIG_ID: 5162 SOURCE: JGI DATE: 2019-11-15.	atgcgaagaggacgtagaggaagacggaftaacgccaaacttcttttaactaagagcgcatcgacaaaaacgattttcaaatagaagacaggaataatgctccctcaatcagagcta atcgtcaacttgaaactcgaagtcgtttctgaactaacaaaatctatccaattactgacattta ctttgatgtatgcaaaacgataattgatctacttccagttagaaaaggagctaaagctggaa aaggtttctgcgcagttatggtcagacagaatggggcattgagcaactatcctaattggc aaaaagtcatactcgttttgggtggcaaacctctaattctcagaaaattttgcgactagaaa aatccaaaaataaagcaaaacaatcaccagaagatcagtaacgatggcattgtctaaga gggaggggacatctcactttgttaacccctaa	gtcactaaccccgccctaaaaggagggg gcttgcctagcttaatgcgaagagacgtag aggaagacggattaaccgccaaacttcttta atctaagagccatcgacaa
256-257	Ga0394881_0010240 GENOME_ACESSION: WSA_S1_MG_FD_SF_3300037311 GENOME_ID: 312717 CONTIG_ID: 10239 SOURCE: JGI DATE: 2020-04-07.	atgtctaattttgtcttgggtgcataactaattgacaacgccttaaccagttaccacgggga caagcaagcgattgttaaaacttcaaaaagctgctattttatcgcgtttacccttgacgct cgtactcaatacacagatccgctcccaatacaacccatcaatataaaatcgatccg ggtgccaaagtttagtgattagcaattgtccgagacgacaaagtaattctgggggtgctgaa tttaaccacaggggacaacaatttaagcacgataacgaagccaaagggaatcaggaaac tttaagatttctagcacagcaactgagttgttcgagcgcaatcgcaaaacaagccaaca accattgaaagatgcagcagcagctcaatagcacccgctggcaactctttaacagactgaa agaacaatcattaccagttgaaatcggaaacggcggtcagaccaatacaatcgtactcg attggaaatcaaaaaacattggttagatgcagctgtgtgggactcgtgtcccaatfaaa aattttgactgccgggctttgttaattcaagcttcaggatggggtagccgtcagatctgcc agccaataaataatggtctcccaatcagacaaagacagcggtgcaaggaatttttcggctt taaaactggcgacatggtacaagcaactctgcctacaggaataatttcgggtactcatag aggcagcctaattgttcgagcgagtgagtttttgagatgattcaccaacgggaaaggtt agcccagtagtcacaaatactgtagggaattcaccataatgatggtatatacgtatacatt gtccacttgcgtccagtaa	gctatccaccaaacagcagacaagagtgga caaccacttccaacgctcgagagatagcca gcctcagttacaactacgtttcagggtcat gacacctacaggtgctttccagctgtagcc ctgtcgtcaatgtaaaatccctatttgggtt aagggaatgctgattggcattcacaagccctc aaaaactggcgagggaacattaccgccg aaggaggacatccatgtctaatttt
258-259	Erysipelotrichaceae bacterium SG0102 DNA, complete genome GENOME_ACESSION: GCA_003925875.1_AS M392587v1_genomic GENOME_ID: 208690 CONTIG_ID: 0 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgacaagctatgctttttagtttaggtgtgataataaacattagcgccaaccaagaaca gaaggcttggtttctattcgtagaagacgagcaacattgctcagcagatatacattggttaa tacaactcaaaaagaaatttcagatcaagaatctgcaaaagatgaattctgttgggaata gatgacgaggtctctcatgttgggtgccttagtagcagaatgtcagacgcgaacaaca gtcattttaaggaactattgaacagcgtatgatgtaaacatcttatggacgttagacgt ggatttaggcgttatcaccgtgatcatalaaagatataagaccagtgagatttgacaacagaa aatcctctaaacgaaaaggagaaatgcacaagatttttcaaaaacgtcaatcaacaat aagagttatcaatcaacttaacaaatgggttaataataacgaattattggttagaagatgttc tattgatataagagcattgacagatggctataaacatatacgggtggcaatacaaaaatcaa atagactggagcagaaatccgtaaaagctgtattttacgagacgggttccaatgtatgga atgttgaaaatctaattgttagattagaggttcacacattaaagccaagaagactgaaagggtt caaatacgccttggtaatttattacgttatgtacaggtgtcaccagaaaacagaagggtgta gaagaattatacatgaacagatacttgcgtttgttaatttcttgacaataagaacctgaatt atgcacagcatgtaatgataggttaaaaatggctgagaaaacagttatcaatttaggaat gttacatttaaccaacggaggtgatacagccaataagcgtattgactgggtattgcaaaa tcacattctaattgacatctgtatcacagacttgcggccagacacatgtgaaatcaaa gaaatgggtaataaaacattatcggaagacaagaaggctaaagacagataatgttcttggaa taaacatagggtttgttgagtacacttttatgaacgggtgaacacatagagggtatgtaa cagctttatatacagaacaaacgttctaattttcaaaagccaacaaacattgcaagaaa gttaacgcaagaaatgcaaaagtcgttggaaatattctaagattattggttagataattgta gttag	tacgtaaacatagacaggagtggaacattta agtctcttctgtatggcttaacaaataggttctg agcctaagtgtactgtctatcgaagatata gttgcagatagtaactacattaaagcagtaag gtaaaacacacctttagatgtaattctcagt ctgaagctctgtgagtgcaaaccaagaac aatgctaattgtctgattgataacaggga acacatgtccttactcgaattggcacgaag agaaatattccgaaaggaaggtgcagaaa tgacaagc
260-261	Ga0247609_10038159 GENOME_ACESSION: 172DNAHGlowp2_2_SF_3300028888	atgacaagttactatttgggttgatgctgattgtaaacctgtaagttcaacgaaagaaac aaaagcgtggtatatgatacgttaaaggtaaaagccaactcgtacaaaatcccaatgta gtacaactgaatcggataatccctactgatgaattctgtaagatgaagttcgtggcggaa ttgatgacgggtgttgcatacgggtattgctgtcgtacagaaatgtcaacaagaaataaaa	tagtcaaacatactcaaaagtggaacagatt aagccttttgatgtgagcttagcagtaggtt tcgagcctcagtgagtgctgacagtgaaaa ctgttctgcacgatgaactacgatagagaag

SEQ ID NO:	Name	IscB sequence	ωRNAs
	GENOME_ID: 279758 CONTIG_ID: 38151 SOURCE: JGI DATE: 2019-06-13.	gttttattaaagggtactattgaacagcgtaatgatgtgaaaaatcttatggatgttcgtagag gttatcgagataccacagacagcataaaaggtatcgtgaagcaagattgataaccgaa attcaactaaacgacaaggcagaataatctccaagtatcttcaaaaagcgacaagccaaa tgcgagtgatatacatattgaataaatggataaataaacaactatttggcttgaagatgtgc aatagatataagagcattaacggatgattacaagcccttatagctggcaaatcaaaaatcta atcgttttagatgaaaataatacgtaaagcggtaatactctgtgatggatgtaaatgtatgaat gcggaaaaaaaatatcaaacctcgaagtacatcataaataccacgcagaaaagaatggct ctaatacgttaggaaatctgataactcttgtgaaaagtgtcatcaaaaacgggaaggtaaa gaagaacaatat	taaggtaaagatacacctgcggatgtacttcc agtctgctgctgtgagtgccaaccaagaa acaccgctaagtctctgtgtgataacaggga aaacacacgacctcttctgtcattggcagc aaggaaaaatctccgaaaggaaaggtttcaga aatgacaagt
262-263	Pseudomonas sp. OV081 Ga0215794_101, whole genome shotgun sequence GENOME_ACESSION: GCA_003391585.1_AS M339158v1_genomic GENOME_ID: 169706 CONTIG_ID: 0 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	atggcgggtctacgtgttgacaagcaaggcaccatgatgccgtgcagcgaaaaacg ggccaggctgcttctggagcgtggctgtgcacgtgatcagccagatgccgttggatc cggctgactgaccgcctgcagaccgaatcgagatccagccgctgtcgtatcaagatcga tcttggcagcaaatcactggatgcgcgttggcgcagcgggacaagcaggtcttggc ctgtcgtgatcgaattggatcatcgtggcgccatgatccaaaagacattgctgcaacgc gccggctatcgcgtggagcgtgcagcaggaacctgcgtaccggcgccgcgcttca ataaccgtacccgcaaggccagttggctggcgccgagcttgcagcatcgggtggacac cacgatgagctgggtgcagcgctgcgcgcgtggcgccgggtgactgactggtgtgtg gagcgagtaaaattgataatgcagctgtgacaaaacctgaagtcgcccgggtgtgagatc agcgggggtgcgcttcaaggctatgaggttcgggaatactgttagagaagtgggtcgc ctttgcatgtattgcaacactcctaactgccactacagatcgagcacattcttgcagagc caacggcgagcagcgaccgcttcaaaccttggcgcttgcgtgcagccatgcaatcaa cgaaaaggaaagctgctgatcgaagtgttttgaagaaaaatcagaattgctgaagcata tttggccagagctaagactacattacgcgatcgcgccggcggtgaactccacacgaac gctatattcgcactctctgcagagaccggcctgccggcaggcaggtacaggaggtca gacaaaattcaaccgctgcacctatggcttgcacaaaacacatgcgtcgtatgctgtgt gttggggatattcagggggttaaaaactggaaaattccgacacttgcgtaaaagcgatg ggacgtggaaattatcaacgaacgcgggttttggcagtgcttcttagaggatattctaa cacggcagaaacggcatttgggttccaaacgggtgacatggtgaaagcaaacgtaacc agaggtaaaaaaattgggcatttaccagggcgcatgtgagagcaactggaaaatt caacatccaactcatgatcgtgtttcaaggaaatcaaccagaattgtgcgcaatcgcc aaaggcggtatggctacggttaccagcaataataacagaagcttattaa	gtcaactaccccgccctgaaggacggggct tgctgaggaaatcggtgagtcgggttgacca ggggggcggagatttttgcgagatgtttgc aacaggtcgttgagaccactccggaatgct tcttcagttccggacactggaaattctagatc acgctggcgaagtgaaagcggcgaaggt tttggatgccgcgaagcgggagccgggtg caaacccccaggggagcgagccctaaa aacctccgtcactaggcccgtaagggcata aattaggaggatcgatggcg
264-265	Ga0307373_10033165 GENOME_ACESSION: OX2MG_2_SF_3300031 672 GENOME_ID: 280771 CONTIG_ID: 33164 SOURCE: JGI DATE: 2019-06-13.	atgcaacatgttttggcgtggatgccagcagcgtccactgcgcactgcccccggc ccgcgcggcgtgttctcaccagcgcgaaggccgcgtctgctggcgttacccttca ccatccgcctgaagcagcggttccagcggcgcttccgcgtctcggcgtcaaacgtg gatccaggagcaagacgaccggctcgcgttgggtcaacagcgtcaccggcgaggtg gtgtttgcgcctcgaatcggtgagcaatctcagcgtcgcctgccatccctgttaacacg gccaaaaggcgcccagacagcggcgagtttggctatccgaaggtccagggcgagggc agcggcccgctcaaaagcgtcgcgtctgcaatccgcgcgtggcgctctatcaccg cctcaaggcgggcggtgtatgtggagcgctagcgggtgcgagcaaccggatcgtgcaac ctcaagacggcgacaggagcaatcaaggcaatccatgcccgtatgttcagccgctcca gcggggagatggctatgcctatcgaaaaggggcgccggcgttgcctcccccacgcctga	atcaacgacttcacgcagcagcggcaatag ctgaccagtccttgcctctcggcgggca gggagccgttcacgagagcgcccaagct cctaccctggatagcatctcagtcaccagc cctagaatgctcaggttaaacagagacgtg gggtatgaaactcagtccttgagcaagatg ggcgtcgtgagccgggagcagggaacc cttacctgcgcaagcagcgtccgcggtac ggcgcccccctacgggggatagcagacag aagaaagcgaccgacacgatgcaacat
266-267	Ga0118725_1008328 GENOME_ACESSION: IMG_3300009377_SF_3 300009377 GENOME_ID: 311715 CONTIG_ID: 8327 SOURCE: JGI DATE: 2020-04-07.	atgcaaaaagtatttgggtatcaaaagagcgtcccttgcactactcatcctgcgcg ggcaagacaattacttaaaaagggaagctggtattatattggaaccgtggaattaga acttctggcagtttattctcaagcgaacacgacgaatacaaggatcatgattcggta ctgtcatcttcaactgctgtatgggtacaaactacaaaaaggaggaacggcacttct ccccccattcaatgggggtgggttccgtgcctaa	gtcaactacccctacccttgaagggttagg gcttggcttggcgggttacttctgtaacc gcttagcaactcaagcccggagttgaccg ctttagctcccagaatcgggactccgttagg gggtaatataggtaccgtcgcgatcttc accagtcggacgctctacggcaagtcgttaa acgagcatacgggtagaagccgggtcgggc ttgcacttactgaacgatatttactttaggtt caaaaccaccactaacattggcaaggtga ctgttacttgaacttaggttctacgactgaaa ggagactttaccaatgcaaaaa
268-269	Ga0247841_10001800 GENOME_ACESSION: LaCruzMarch2015_16_ SF_3300029286 GENOME_ID: 280592 CONTIG_ID: 1799 SOURCE: JGI DATE: 2019-06-13.	gtggctctgtgtctgtcattaaaagtctgagggtaggaaacgggtgtgagcagactaaaaa cctttccaacaacctcgatgctcttaactctgaaaggagaacgtaacttgagagtatac gttgttaatttaagaacgaacctttaatgccactactccagaaaagcaaaaatcctgctt aaaagtggcgaagcattcggtattcaaaagaactcgttaccatacagttgctacatgcta gtggcgaaacaaaaaacctataacattaggtgttgatagtggaattcagaaatgttgatta tctgttataacagaaaaagaagaagctttagtgacagaagttgaagacagatatagt aaaactaaattcgagaaaaaggcagatcgcaggcgcaaggcgtaatcgcaagacttgg catcgtgagccaagattcctaatcgcaaaaaggatagtggttgggttagcgcctctatata acacaaattagattctatatacaggttagatagattgtaaaagaatattgccgataacca agataaattaggagttgcaattttgacatacagaagataaagaatcaagacatagaagg aacagattatcaaaatggagaacagtggtgcttctggaacgttctgtagtatgtattgcatc gtgatggctatattctcagcattcgagaggtaaatcgggtgataaaatctagaagttcat catataaacaccagacaaacaggtggagatagaccagataacttaataaccttattgtgtt	ccatgatttgggtgaagattagcataaatg cggaggcttgcaaaaagccttgattgactaga gggtgttaattacacagaagttggttaagag aaatacatacacactcaaatattacgcaa gtttgtgctctgtgtgtgtcattaaaagtct gagggtaggaacgggtgtgagcagactaaaa acctttccaacaacctcgatgctctctaact ctgaaaggagaacgtaacttgagagtatac

SEQ ID NO:	Name	IscB sequence	ωRNAs
		tatgccatgagaaagtatcacaaaataagttacaactcaaaatcaagctagtaaaggata taaagccgaatcgtttatgtcaatggtagtgccgtgtgttaattataagagacattgg agatattgtgtctcactatgtgggtatataaccaagggaatcgaatagcactaggtataa gcaaatctcatgctaatgatgcttttggattcgggaggtactaatcagatacgtcttaattgg ttattaatccaagaagtcgaaagtcgaaccgcaactctataaggcatagcaagtcac atcaagaatacagccactagattcattaggggattccaacgttttgataaggtattgtggaa aggtgtattccctactcggcggagcagaggtgttgaggtttcataatttgctgtagaacat cgggctattttgacataagaagactgtgatgcacgaaattaaagtaacagtttaataatcag gaattacgactattagacatttagaacattttcactagaaaggagagtgagcagactttcc cgaataa	
270-271	Ga0194121_10011106 GENOME_ACESSION: TA2MahCa50mmetaG_2 _F_3300020200 GENOME_ID: 35901 CONTIG_ID: 11105 SOURCE: JGI DATE: 2019-01-27.	atgttagtgtacgttttgaacaaagaaggagagcccttaagccctgtagccctgtaaagc aagaaaactcctggcaccggaaagcaaaagtgtcaaccgagagcctttaccattaa acttatgtttggctacccggatataagcaaaaggtcactgtcgggtgtgatagcgggtcga aagtagctgtcttttcagcaacgggtgagtaacaaacccctttacgtatcggaggttaatta agacaagatattagatccaacatggaccaaaagcgtatcttttagaaggatgagaaggctc agaaaaactcgttatagaaagcctcgtttcaacaatagaaaagggtgatgtgtgacg cctacagtgcacaaagaaagtaacctcacaacaaagggaactggcttacattaaaaactta ttcccgtaacaaaatcattattgaacccgcaagtttcgatttcacaaaatcacaatcca gaggtctcctctgtgttatcaagaaggagacttaaggactctataatgtaaaaactat gtattacatagggatcatcacacctgccacaatgcacaaaaaacaaagcttgcgctcat gtgcatcacattttttgatctaatggaggagcagctctccggacaaccttattacattt gtaaaaactgtcatgagactctacacggttcctcctcaagctgctcctcaaaaaactc ttcacaacaaacttaaaagcaagcctacccttgatgccacgcaagtgctaccattgctcttt tttgaacaaagaggtatgtatcgaggaaacttttggttatgaacaaaatcaaaaagagag tccttaggcctgcacaaaactcactaccatgatgcaatttgcattgcaatcaagcaagggc aacctatgcagattgggggtcctcttttaaaaaagtcacacatagctcaaggagattacaa gccttgctcaggagatagatcagaaaaatcctctacaggaaggtatggggttaag aaatttgacaaagtaattctaattgggtaactcgtttgtcaaggtagaatgtcgacagg ttatgctatccttatgtatcaggggtcataagctaaatttaaacccatcccaagcttaa agaattaaacgcgttgcgcagaaaatcatgcccttaacagcttgttccatcgaataa tcctttaggtaccacatccttggcacaacaaacagaaaccactttttgcaatgagtc cataa	gtgaagctcccttagactaaagatctaggg gcttcttagaaactaggaagcctacttcacaa gaacacaaacaaaaggagtaaaaacatgt ttgtaaacgatagaacagaaaatataggcac ttcaggatgccgactcagctcgaactctgc gaattgcagttaaacagctcgtagggaagg gacagtgctgaattgcacaaactgtttatc cgctcagagtagaagtcagaaaaagcatggc catcctgggtctgtcttcatgcataacct agttcaagggggtgcactttatgta
272-273	ElkS_mat_CD6ADRAFT _1002000 GENOME_ACESSION: ElkSloMetagenome_4_5 F_3300000354 GENOME_ID: 25895 CONTIG_ID: 1999 SOURCE: JGI DATE: 2019-01-27.	atgaaagtctactgtataaacaacatggcgctcccttgatgccacaactcctagaaaag ctagattgtactcaaaaatggtaagccaaatcgttggacgtgacccgttactatccag ttaatttatggctcctcgtgggtacactcaactgttggaacaaagggaagcgtcgggaattag agatagaaaaatctaccataatgatgtctatgtctgctgtgggtgagactcagccctgt gtactcaagccttggatgcaacatgacttagttctgtttgaggagcaaatatctcgtttgtc gagtcacaaatttgggaactcgcctcagtttaagccagagcaggatgcacaaacaaagt ataaaactcggcaatgaacaagggttaaaccttttaaaacaaaggagggaatctgtgagc aacaagcataa	ccgtcacgaacccaacgctgaccgatacga gctgtacagcgtgggcttgagaaaactt aagccctcgtgattagcctcagattccgac gaaagcaaggaaactactcgttgatcggaat acataggcaccgtgaagtacaggggttgca acccgttccctagcttcacgcctcgtgtttc attaaacattcctattggtaggggaagtgtg aaaacatcaacatctcagctccgtcgtatgt gaaactccgatgaacattggcgcaggaaac cacaccgcaagggtgaaaacactctgttagg tctatgaaa
274-275	Museum specimen metagenome DNA.contig: SAMD00035458.scaffold _600, whole genome shotgun sequence GENOME_ACESSION: BCQK01.1 GENOME_ID: 159 CONTIG_ID: 600 SOURCE: WGS DATE: 2019-01-27.	gtgtgccaaggtgacaacttatttcaacattggcgaagggaatctacgggagttatccga attacatttttaaaaatacaaaaatggtttatgtgttagacaagctcggaaaacctatgatgcc gacaaagcggatggctgggtacgtcgtgcttgcgtgatggacgcgctaaagtagtcag acgtgatcctttacgatccaattaacctacgagccggacacacaggtgtgcaagacact acgcttggattgacattggctatgagaacgtaggtacaaagttgctaacaaaacaaaag gaattgtttagcggcaattttaagctacggacagacatccaaaagcgtttgcatactgtag gatgtacagacgaactagacgaggacgaaagacgcgttatagaaaagctcgtttttgaa tcgcaaaaagtattatgccgacgctccgagcgtgaagacacaaaggtacatagccacgtaa gaatcattatctgataacatcaattttgccgtcaatcaattgattgtcaggttgcacatttt gatgcacaaaagattaagaatccagacattcagggtgaagaatatcaacaagggtgtcaa ttaggctatcagaatgtcaaggagtatgaaagagtcgagattgtaccgttgcactttaa cacaggcaaggttagcaaaaagctagaagtacatcacattgtattcaaaagtcaggcgg atcggatgctcgaagtaatttgattactgtttgtataagtgccataagtcattacgatgg aaaaagccaaacccaaaagctagtaaatcaaatcactaagcctattagctttatgaaaaa cagtatctaagcgtctacaagaatgctaccagacattgagtaacctatggtaccagac aaagcaaaaacgggtataatcctaatctcgaaagtcacacgtaaatgatcttttgaattg caaatgggtacaagtcaggttcgatgtaagacaacgaattatgtttcaaaagaaaaataa cagacaattgggtaagctaaagaaaggcgttgcctacgcagtagaagaaacacgttatcc gatacaacctaaagattttagtgatgcgaaggaaaaagattatgctattggcacacatt gtaaaggtaaaagagtgattgtgatgcacggcaaaaaaaatcatagcaattaaga taataaaattatttttaaccaacgaagttttagccctcacggcgtga	atccacaccaagcgtcaatgcgtatgggaat gggttgaacaaacgtgttgattagcctaa gtcttaactgactacgttgaaataattatcata ccctaggatgcttcccaagcttaggctctat gggtgagtattaaaagttctgatggtaggaa cgggtgtccaagggtgacaacttattcaacat tggcgaagggaatctacgggagttatcccg aattacatttttaaaaatacaaaaatg

SEQ ID NO:	Name	IscB sequence	ωRNAs
276-277	Ga0170573_10661034 GENOME_ACESSION: IMG_3300013232_SF_3300013232 GENOME_ID: 28678 CONTIG_ID: 11314 SOURCE: JGI DATE: 2019-01-27.	atgttagtctttgttctaacaataatgcacacccgttaatgccctgcgtccgcgcaaggc gagactgtttgttcgaattgagaagcaaaagtgttaacagaactccgttcgtgattaag ttaaatttcggctcctcttgatatacccaaccagtaggttttaagtttgacacaggggctaa aaatgcgggatttgcgcgacagcggaatggaaaggtttgtacacaaagttaggtaaaagt gcgtacagacattccagaaaaatgaagcaaaagtattcgttcaggcgaaactgtcgcg gacgcaatcacgctacagagcgctcgtgattgataatcgtctcagactggagggtgg ctctctcaacgatgactcaaaagttaattctatgttcgagaaataattttattaagaaaat aatgccagtgaatcaatgctcgtgagattgcgcaatttgatattcacaaaacacaaatc cagaagtgttaaatgacttaactgggaaaagtattcaaaactggccgacaaaaggattctat aacacgaaaggctttttttatcgcagagataaattacgagtgtaaaaaatgctcggcaaga aaaataacccaaaactccacgttcacatattttttgttcaaatggcgggacaaatcgc cagacaatttaattacgctttgcgaaccgtgcataacgggattcatgcacacaagacccc agaaaaagaatcattaaaattttcaacagaatacagaagcaaaagacacaccgcaagt ccgttcaagtttcaacggctcagttcgtatcgcacaaaataatccagaaatgaattaaact acggatagcagacaaaatttaagagagattttggcgctaaaaaggacattataatga cgcaattttgtcgtgactgtttgaagggtgaatttgcgaaatgccaaatcatatttataaaa gttcatacgcgaattggcgattacaataaagaatggcaaccactcgaacaagtaattcc aactggaaaaattatgggttaagaagttcgataagattgagtatcaggagaaaagtatt ttattaagggtcgaatgcgacagggtatgcaattttaaagaaacacgaaactgta aaattaaaaccgataccaaagttttcaaaaatgaaagaatttctgcgaggagtcgtgcc ttacaagtcagattgtatcggaaatcactctcaaatatcacattcttctgtcgaataac agaaaaaatctccttgcagagaaaagaattgtgtgaatttatga	gtgaaggctccgctaagctaagaacttagc ggcttctataggaggcggaatttttcgcct cagaccaagagtagcttcacagaacaaca acagaagggttgagagaatgttttaaacga tagaacagaaataggcactttgggatgtcga cttagtcccgaacaatgcgaatgcaagttaa acaacgataagagtaatcgcagtcgttcgat tgttaaacctgttcatactagtcgaaagtga ccgtgcacaaaacaaaattcctgtgttttag tgcgcactactcagtttagcgaggaagact ttatgtta
278-279	Beggiatoa sp. 4572_84 ex4572_84_scaffold_492 2, whole genome shotgun sequence GENOME_ACESSION: GCA_002085445.1_AS M208544v1_genomic GENOME_ID: 120561 CONTIG_ID: 281 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	atgttagtctatgtcattataaaaaatgtaaccatttaagccttgcacaccgcaaaagcg aggaaaacttttgcgcgctggaagggcaaaatcgttaactgtgtcccttcaccattcaatt gcaatgggactgtggaagaaatgtgcaaccctgtacgttggggattgataaagggaagtc ttacaccgggttgcgagtggtgttggcgaattttatcttctggcattattaatcgcac cgatattaaagacaaaatgaccgcacgacgtgtgtaatcgtcgtcaaaagacgtatcgtaa atggtatcgtcccaacgatttttaaacgtgcaactagtaagcacagcgacgattacag ccttcattaaagctaatgcagaagaggtaataaggggtgtgcgcaaatcttctccccc agtcaaatgtgattgaagatgtgcaagttgatattgtcgtcacaacatcctgattgtcgtg ggattgagatcaacgactaatcgtttagattgaatgaatcttgcgtagccacattgatgcga gataaatatcaatgtatcttgcggaaaaaaaagggttcaacttcaagcacatcacatagt gcctcaaatcaaggcggtgaaggtatcgttaaaaacttaataacgctttgtcaatcctgtc ataataagtagatcaaggcccaatcactcttcagctgtgagcagataagtggttttaagac cagatagcccaagaacatgcaaggaatactttttatccaaattttagaacacgttggcc ccgggtttcaagtggttgggtaccacacgcttcgttgaataatatttaagcttacctaaag aacatgatgttgatccttctgtgtggcaaccttagataaaggaaacaaaagtgcctccttgg tgtggcaaccttag	atcataatagttcttaactatagaataatgtt ctaaatagaaattttcttagacatggaaac taacagtgggctactgtgaatgagtcgagaa gttagatagggcatcacacctccgatgtgat tccagtcggtcgtcgtgtgtctagtattaga ggtagcgggaacgtgaaggtgtgctaaacg taaaaaacatttaacaaccacaaggagta cctcactcactaacaggaggtcctaaga tattatgtta
280-281	Groundwater metagenome DNA, contig: GPS12011_contig_3985, whole genome shotgun sequence GENOME_ACESSION: BBPF01.1 GENOME_ID: 157 CONTIG_ID: 3224 SOURCE: WGS DATE: 2019-01-27.	atggctgtattcgtactggacaagcacaagaacactaatgccctgcactgaaaaacgg gctaggctgttactgaacgcagacgtgcgggtgtgcacaaagtgtcgccttttactatcc ggctcaaaagaccggacggtgtagaacaagccaactcaaccgttactgtcgaagttagatc caggtagcaaaactaccggtatcagatcttacaagaagcagtgctccgagaaagcagat gtcgtcttttggcgagctgcgccacaagccgggcatcaagaagaagctgtgccgaca gaaagacacagcgcggaatcgtcgaaccgcaaaacacgctaccgcgagaacgggt ttgataatcgtaccaggccaagcggtggtgcggcgtcgtcgaacccgggttaatc aggctcagaataaccgtaaacaaattgcaaaagctgttgcacataaccgcaatcagcacc gaacatgccaaattcgacacacaacttatgcaaaacctgccattgcccgtatcagatc cagcaaggcggaacttttaggttacgaggttcgggaataatcattagaaaaatggaaacaca aatgcgcctactgccgtaccgacacacccgctggaataagagcacattacccccaaa tcaaggggcgagtgacagaatcagcaacctaacattagcctgtcggcgtgcaacca agccaaggcgcaacaggacggcagaaggttggccaccctgaaatccaaaagcaggc caaattgcccgttaaaagacgcgcccatgatgaacgccacgcggtggcgattgtttaacag gctaaaaagaaactgcctcccggtcagtgccgttactggccgaaccaaagaacag cgtatagaacataaactccgaaaaacccattactacgatcgtgtgtggtggtgcccgc acccccgcaaatcttgcgatcaggcaaaagtatttccatttggaaagcgatcggcaga ggcaccaggagatgtgaacacagacaaatatgttttccg	gtcatagaccccccaggctaaagccggggg cttagcatttgcctcactatgaccagcctaag tcccacgaggactactgttacttagccatgac accctgggggtgcttccagctccaggcact gtcgtccggactaaacagcgagtaggggtt gtcagcagctgccccggagcgacaaagc taagctaactgtggcaggagagactaaaa atagcgttaccggcccttaaggggcctcg aaaggagagaaaaagatggctg
282-283	Ga0315276_10000018 GENOME_ACESSION: YL17G05_0_MG_2_SF_3300032177 GENOME_ID: 281627 CONTIG_ID: 17 SOURCE: JGI DATE: 2019-06-13.	ttgactagacagttgttaactgaactcactggacgtgfcagtcacaaaagacgtagagat gcttccctagtcttctctctctgtgagtcagtgctgaaggacgtacaaacaggttttca aggttctcgtctactgcacaaaaaccttgagttgaggagttcaataaagattgtttcct gttactgacaagaatacaaaccttgatgtcgtactactcctcaagagcgaggaaagtga taagacagggcaaaagcgacacgtatttggagcaaccaaattttctgtgtcgttgaatca agaaccaagtgcagacaagaagcaaaagattgtatgggtatcgcacgggcagcaaaa agagaggcattcagatcaagtcacaggttcatacatatctgaacattctgtctgatcggg tgacttgggttaaaagatgctattaaataagaagagagcagggcaggtgaagacgtcaaa gaaaaacgccaatcctcaatgtatgatctaacagatcaaaagagaaaattgcacctcaa	gtcactcatgagccccgggcagcgcatgga gggtaatgacgtcaaggacctggccaagc agctcgtgcggacactcgggtacaaagacc agatcctacgcaacgcgaatcacgaag gacaacatcgacagctcgtcaaggcggtg aaggaggagaagcgttcgcgtgtgtcaa caaaaagcgcgtcttgttaccgagctgac gcggtactgaacgccttcgcgagatgtac ccggaaactgttatgatgatcatatataaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
		taatatcaagatggcaactgaaatfaaatacatcaacagaatgatgcgtatcttccaatc acagattttgttgaagacatcaaaagctgaacaaagaaaggaaagagacgttgaac aaatcttttcgcctcttgaatfaggaaagagtggtctataatgaacttaggaaatttgta atcttgagacacgacagggatagaacaaaggaactacgtgatacatggttattgaataa aacagagtcaagatgtcattgtcatttgattctataatgttattctgggtctgtctaatta cttgttggaggtcatattaaccagataacaaagaaatcaaaaggatgataccaattcagt ttcatcgaagacaattacatcgaatttaagattgaaaaaggaggtgaaaggaaacgatatgg cgttacgatgatttggatttaagaaaggaggttggtaagaataagaagcatggtattt gttatatcggtgaaattcatatgtaaaftgagcttcattctcttaaacaggaaagagac tatgtcaaggagctaaattgaagaaatfaaattttaacatattgtatttttagatga	cttattttaaagaaagatgtaccctgaggtc aattacctagggtatttttaggcacaaagaa a
284-285	Ga0117909_1048905 GENOME_ACESSION: IMG_3300009320_SF_3 300009320 GENOME_ID: 311703 CONTIG_ID: 48904 SOURCE: JGI DATE: 2020-04-07.	atgcgtgtgtatgtaataacaaaaacggtaaccacttatgccctgcaaacaggccaaa gcaaggcatcttctacgggacggcaaggcaaaatttgcatacgaacacattcacaattc agttactttgggattgctcagaaaattgcaagacgttcgttggattgacaagggtatgc atggtaacgggtgttcttgcgtaggtaattggagaatcctctctgccaatatcaacaa cagaaatccggtagcattacaacaaaagacgggtcaacgaaacatttatacagtgag agcagaaggcgcacaaagtcgtagacatcgccacaaatggtatagaacaaagattca ataatagagcatcatcaaaaagatcgttgtaggttccaccgaccattaaagatgaatgtcat ggaagtgtcaggggtataaagaaaattccttggccaatctccatattaccggtgaagatgt ccagggtgataattaggagaattacccctgacgtggaaggtagtaagtagcagaatc taatcggctagatgagaatttaagacttgcctgttattacagggaataatttactgtcagaat gtgaaaaaagaaatcccgattggaagctcatcatgtaatttggactgcaaaaggcggaa aagatagttttacaatttgattactcttgtgaggttgcattgagaagtagacacaaacag gggaagtagcaaaagtcagctcaaaagaaacaaagtagtaactggcatggatgggattc aaggataagatagctgccgaacatgcaagggaacattttagtatcaagaactggga aaagattgcccgtctgtgtgtatagcgttatcagactctgaatttagaaaagcattaga tttacataagagcattggatagatgctatttgttgcctatttagaaacgggtgagattgtt ccatagatagtaataatcattattctatattgttgcgggtaacaaacagacgacgtttgt acacgcaaccatcaaaagggtgaatgataaagcagtgccaaaagtagcaaaaggattagc aagtaattggcaagattgtatttttagtagacaacggactaaaacgttgtttaccagaag gctataccctttacaaaaaggcgaatttattgacattttaggtttacagactgaaatagcttc tatcaatggaaggcgaagggtttattattggatatacagcctgatggaactagaaagt atgcttctgtatccataagaagtcaggttggtagaatatgctaaagacactacatactatc tcaagtaa	gtcaacaacccctctgattcacagaatcag aaggggcttgcgaataactcggttaagccct gggttagaccagctccccttcatcttgacttt gatgaaggagcaggttactgtgtatcacac ctttgggtgttctccagcctgaagctctgtg gtttattcgtaaacagttctacgggttaggaa cagtcggataagcgtaaaaacacgggtgaa gacataggcgttctgtatgtcttaacaagga cgaggaggacatactctttacacgtagga ggttataaaaaacataatgcgt
286-287	human gut metagenome genome assembly, contig: NODE_25_length_91722 _cov_8.591391, whole genome shotgun sequence GENOME_ACESSION: OIYR01.1 GENOME_ID: 6152 CONTIG_ID: 24 SOURCE: WGS DATE: 2019-01-27.	atgttgcacaaaacaagaatfagcgttagattggatagtgatgggaataaacttgcctcaac taagtagcagaaggcttggcttctaataagaagaacgagcaaaactaattcaaggat cctatgttattcaatfaacacaaaagtgaaatagataaaagatggcacaactctagaatg tggaaattgatgatggttctaaatattgttggcattgcattatgcaaaagtgtaaaactaaaa caaaatttctcacaaggactattgaattaaagacaagattgtaaaaagaaaatggatattcc gtagagaatcgtatagatagaaggatcataaactattatagaccagcaagatttaataat cgtgttctatgataaagaaggactgccccacaaagtatttaacacaaaagaaagactca ataaattagattgtaaatgtttaacaaaattgtcaattatgataatcatttagaagatgt aaaaaattatattagaaaatgttgaaggcaagaacatttattggtgatgaatataatattcc aataaagaagaaatgccaatcagaatagcgcattttagtagataattatactgtcaag agtgtagcaaaaaacaaatttgaaggttaccatataattcctagaagtaaaaatggttct aatagtattataacataaactattatgtcggattgtcatcaaaagacagaaggtaaga attttattcgtgtgtaaatattaaaggcaaatagggaatttcttaagggtcttaattatgc acaacacgtaattgcaaggcaagaagtatttaagggtgattgagagaattgttaaacttg gacacaaccgatggttttaacaacttcagaaagaagggaattgtgggtatagacaaatct catagtaattgatgcagtttgcacacaggatataaacctaatgacatagaattactgaatat actataattccacaaagaagaagagtaaaagctaaagtaaaagggttaaatgggttaaac atcgtgatatagttatcatatcatcacaataacatagattatattggtacatcaccgcaa tttatactgatggaacaaatcattgaattgtaaaaacaaagaaaacattttaaagagttt cttacaagaatgtaaatattttataggtttgataaaaatttgggtgttacaagataaattat aatttatataatttag	tagaaaaatataatgataaagtagatgcttggc ctttattacatagatttttaacagggttcaagc cttagtgattgtcaatattgaaaatatttggc gatataactgcattgagagaaaagttaa gacatacctttagatgttattcctcagctgaag ctctatgaactccaacaaagaacatactaa catcttgatgataacagggaagaaagtaaga actcctctgtatttggcgaaggagaaaact ccgtaaaagggaaggtagacagatgttgca aaaacaaga
288-289	human gut metagenome genome assembly, contig: NODE_113_length_8314 5_cov_5.672704, whole genome shotgun sequence GENOME_ACESSION: OLRJ01.1 GENOME_ID: 7925 CONTIG_ID: 112	atgtccgttgcgtgattagcaaaactggcatagtactcatgccaacaagtgaataccggg cacgttaagctgctgactccgggaagcaatctttatggataccgtccattcagatcca gctaacagaaagggaacccgggtgatgttcaaccgggtgaactctgtgtgatacaggata tatccatattggagatctgttaagtcagaaaagcagaatctctgggttacaggtcgata cattactgacgagaaggcaaaacacgatgcccggtatgaccgagcgaagac gaaaccggaagcgataccggaagtcgcggttgataaccggaacggaaatctggctg cgctccccgtcactggaacataagaaagagatcaacttcagatcataactgctatctgt ggagtataatcattgcagacatcacgttgagatgggaacatttgatacagggttctga aagcaaaaggaaaggttaagcagataccgcaaggaaactgattaccaatcaggagaac gggtatggaaattgcaacgctccgggaagccgtcttacaagggaagggtataaatgccagt gctgtgacgaggcatcaaaagatggcgaatctgcacgctcatcatatccagatcggga	gtcaacaaccccgctgaaataaattcagg cggggcttataaagatctgatttttttgg tctagtcccgtcttaggtttattgattggcgac gaccgcaaggccgtaaatgattagcctcagt gcaacgaacccggtcagaccagacacggt gaatcattagcactacgttaccggttaaafta ggcactccgggagtcacactcttagtccaga ccctgcgggtatcttataacatctctgaggg taagagaagtgtgagatacgcgaactacc ggataacattggcgaaggaccaccgatt ctgagttagtgcctgtaccgaaaggataga

SEQ ID NO:	Name	IscB sequence	ωRNAs
	SOURCE: WGS DATE: 2019-01-27.	gtcagggtggaaccaaccggatgtcaatctgatcacggatgtgaaaagtgccatacac cagccaatcataaccgggtggaaaactgtatgttggaagccgaaaattgcttcgttcaa ggacgcaacttacatgacagcagtcgggtggcaactgtatcgtaaagtgaagagatgtt ccgggggatccatgtcaaaattacatgtgagcagaaccaagggaacccgtagaatcc tgatatcaaaaatcccatgtga	gaacctgtatggatcacctaggatatctgata cagaaggcgtgaagccaaatttcagggaagga ggacagcttatgtcc
290-291	Ga0074469_10883752 GENOME_ACESSION: IMG_3300005832_SF_3 300005832 GENOME_ID: 311650 CONTIG_ID: 72826 SOURCE: JGI DATE: 2020-04-07.	atggtttatgtactatcgaaacaggccaaccgttaattgtcactgagcgacacggaaaa gtcaggcgtctgttaagtcgggcagagcaaaaggttgcagtcggacgcccgttcactatc cagttgctttacgagactacccaatactcaacgtgtactctgggcattgatcccgatat aagaatgtgggttatcccgagtgacagacaagagcagggtcttcagtgacagagactaa agttaggacggatattccgaaattgatggaggttaggaacagatcgcagagctaggcg caatagggaagaccaggtatcgcaagcccagattcaacaacagggtacgtactaagcata aggggtggctagctccatcagtagaaaacaagatggatgcacatatcaagctgttaaaatt gggtgtgttctatactgccataagcagattgtgattgaagccgccaaattcgatatcgaga aactaaagaaacccgaaaccaaattggactgactaccagcaggagagcaagctggcctt ctggaatgtgagagatgtattatggcgtgatgcaaatatcatccagcattgtgaaggc aagtcgaaagacgttatcttggaaagtcacaccttgagagtagacggacagcggggga tcgacctgcaaacctgttaacactgtgcaaacctgccataataggatcataaggacgggt ttgaaacttcaaaagccgggtgtggtatcaaaagcagcggcgacgcaaacataatgcg ctggaagtatacgaacgcgcaaaaggcgttaggcttccctgtaagataacgtatggctat caaaccaagtgcagtcgcatatcacaaagctggagaaaagtcatacaacgatcggttt gtcatgtcaggcggcagtggtgcagtcggaagcaaggagattccagtttcggcaggt acggagcagaatagaagctatacaggggtatcagaagtcacgttaggagcaaatgtt ctagagttctatcggattccgtcaattgggacaggtgagatacaagggtcaagagtatttt attaaggacgcagatcatctggctatttcagcttgcgatatcatggcgaaggttttca ctggacgcaagaaactagatgtgttaagtattcagaactctgtttagtagaaagagcat cgactttattatcaaggaggagcgaactctctccccgatgaatcggggagtatccgtcgc tgataattca	gtcaactacccccgcataaatgcgggagctt gtagaaagcttctggtgactagcttaagtgc ttcgagcactacgttgggtgggtcagatcac cctggaatactccttagttccaggcactatc gttaggcattaaaagcctctgaggggatcgg agcagtggttttagcgtgaagccctacca cattggcgaagaggacacacagataacctc ggatatctgattatccgaaaggattctctatg
292-293	Ga0209253_10060444 GENOME_ACESSION: FrelakBRBRSPAdes_SF_ 3300027900 GENOME_ID: 26207 CONTIG_ID: 60443 SOURCE: JGI DATE: 2019-01-27.	atggcagtcctttgtttggaccgcagcggcaagccgctgatccgtgttcagaaaagcgg gcgcggctgctgtggaacgcggccggcgccgggtgcacggcgctgcccgttcgcc atccgcctgaccgaccggcgctgtgcccgaactgcagagaggcagccgctggcggtgaag ctggaccggggcgcaagacgacgggcataagccgttgtgtgcacagcgggcaaaagcg gtggccgtgcttggcttattcgaattgcagatcgccggcgacgcgacgctcgggacgcgt cacgtcacgcagccagaacgcccccgtcggcgggtgcgaacactgcgctaccgcgc accgcgtttcgacaatcgccgtcgtccgacgactggctcgcaccgtcgtgcagcacc gcgtcgataccgtgctgacgtgggtgcgcaggttccaacgatggcgccgacgacagt cttgcgtcgagcgcgtgaagtcgacatgcaattgattgcagcaccggacatcgccgc cattgagtatcagcagggcacactctccgggtattcgggtcgcgagctacgtctggagaa gttcggccgccagtgctgtactgcgatgcagagtggtgcccgtggaaatcgagcacg ttgtgccggggcagccggcggtatgaaccggccgagcaatctgacgtggcctgcc ggccgtgcaacgcgaacaaggatcgctccggtcgagcagttcctgaacgggcagc ctgaccgtctggcgccgctcaaacgccagttgacagcgccgctcgggtcgcggccgc ggtg	gtcaatcactccggcctgaagtgaagggccg gagcttgggggggtcactccagcgtcgg gattgaccagagaagcggttaaccaacc gctacgttatgtagaggttcaagaccaccg gcaatgcttccctagttcgcgctctggaa cctgcgaccgcagacacgcttcgggttaagc acgaacgggtcgcaaggatgtgcccgtca cacaactctcaggggagacagccgaa aggctgcgtacaacaggccgtgaaggcag caacacaggagagattgccatggca
294-295	Ga0172378_10009394 GENOME_ACESSION: GroweLOW332metaG_2_ \$F_3300014203 GENOME_ID: 27072 CONTIG_ID: 9393 SOURCE: JGI DATE: 2019-01-27.	atggtttatgtaattaacaagacgggagtcggttaattgccttgcaaacctgtaatcgcaag attacttttaagcagggtgaagcgaatgtatcaagagaacgccttttacaatcaagttatt gtatcaggcaacatcgtacactcaacattgacattgggaattgatactggtagcgggtaca atcggaaactgcgtggttgataacgataaccgaattggtttacgtgtcgaagttgaagtca gaaatgatattaacgacaaatgacacagaggcggaagtacagggtgaacgacgaag cagaaaaaccagataccgaaaagcagttgggttaaacagggcaaacagatcatcagaaaa gaccgtttttcgcctacaatgacaagcaagattaattcacactgaagggaagtcaagttgt aaattcgatattaccgattagcaaaataattgttgaaacggcgaacattgatcctcatgcctta aaaaatcccgcctctacaaaacgaattggtttatcaaaaaggcgtgaattacggttttg caaacactaaggcgtatgtattggacagagataggcacaaggtgcaatattgcaaaaggca aaacaaaggattcaaggttgcatgttcaatattacattcagacgaacaaagggtccgat aaacctgaaaacctgttaactttatgaagactgtcatgacaagttgcagccggggaaa ttatattaaagacatcggcaaaactaagaataattttaaactatgtacgcagatgaatagt atacggattcaattattaaagtgttgcggatcgccggaagcaacattcgatacataactaa ggaacatcggaattgatggacttgccgaagaacattgtttgatccgtgctatcgctt gtttgaacaatatctccaatgacgggtgttgagtgtagattttaaagcgacaaaaaatttt aaagaagtgcataccagatggagattatcagcagacaaaggcgtgagaagcgaaaca aggattccagcgggtaaaaacaaggatttaggaagttgacaaggttatgtatttaggaa aagagtatttcattaaaggtagatgtcaacagggtatgcaattgagatatttaacgggt aataagggtgattcgaacatttccaaatttgacaagatgaacgaataagatcgagaa aatcatggatggtgagcgacgaattacccctactattag	gtgaactaccaccacttaacctgaagtttga agtgggtgcttctacggaaagcctaagttcac cggactaagttgtagaataatcagctacgtt gaaacggttatgacacctacggttgacgcac cagaccgttgctctgtcgatattgttaagta gagccgaggttaaggttcggtatataatgcat gtaagccgtttcaacattgtcgaagtgaagtc ggattccttgcatggttacgtgcaagata cgcacgacttaccgaaagcgaagattttct gaaaggattcacaactg

SEQ ID NO:	Name	IscB sequence	ωRNAs
296-297	Ga0137388_10019024 GENOME_ACESSION: 15con2h14AmetaG_2_\$F_3300012189 GENOME_ID: 20760 CONTIG_ID: 19023 SOURCE: JGI DATE: 2019-01-27.	ttgagcaaagttttctgctgcacacgaataagcagccattgatccagtcacaggac gagctcgtattttgctaacactgtaaaagcagcagctacagaagatttccattcacaatta ttctacaagaagaatacatgatccagagggtaaaagagctgaggatcaagattgatctag gaagccgggtcacgggaattgcaattatcaatgatcagtcaggtgaagtatttttcagc cgagctttcccatcgcggacaagcgattagaacaagaacacggcctgcatacggac gagfattaacacgtgcaactggatcattgatafacaacgcgggctggacgtctggtaa cgtcaattcccaattattgccacatccatcaaaaggacggctatagtattacagaaaggt ggcgggctctccgttcgctag	atcatgaaccaccactgaaagcgggagc ttgtagtaccgaaaggggaactacaac ccaacatgaccagactcagtgctaggccttg tctccagcactccgttaggagcgaactagg tacgttggatgcgcggccagtcaccaact ctacgggtacagcttttgcgaataatctacagt gttaaggaaagtcgtgctacgtaaacctg ccataacgtgtgcgagccaacatcactgg gaaaccagaggctcaatgaggagcaata caggagtatacattgagcaaa
298-299	Ga0126377_10033159 GENOME_ACESSION: PanS0iMetPlot_22_2_\$F_3300010362 GENOME_ID: 34141 CONTIG_ID: 33158 SOURCE: JGI DATE: 2019-01-27.	atgagaatagcgttttagacacagccaaaaaactctggcggcagcagccgcggc gcgcccgtattgctgttgaatcgggcaaggccgcgggtgtcaagcgtatccgttcacga tcattctcaagcgagagatcgaaaacccccctccgacctaaggttaagattgac ccggctcgaaaacgaccgcgtcgcgtatcgaacaaagcggcgaggtgtctt cggggcgagatcgagcatcgcgggcaagcgatcaaatccagactgatgcggcg ctcgcttcgccacggacggcgccgcaagaagaccgctatcggaagcgtcgttcgac aatcgaaaggcgccggaaggtggttccgcttcgctcgaagccgctcgagaatg tcatactggagcgcgaggtcattcgtcgtatccgatcaggtggcatcgcatggaac tcgtcaattcgacacccagttgatacagaacccccgagatcgaaagattgaataac aaggcgagctccagggtttgagctgcggaatacgtgttgatcaaatcaatcacaat gcgtctacggcgcgacgacgtccctgcgtacacggcctgaacgttgatcacatc ccaggctcgcggcgccgcgaaccgcgtcagtaattcgtttgcgctgcgcaagc aatgaagaaaaaacatctatcgtcgaagaatcggcgcatgcgaggcaaaagactt tgccatgtgaaggcgtgctgcaagagcgccgttgaaggtgcccgcggcgtaacgc gacgcgatggcgctcttcaaccggctcaaatcgaggaggttgcgcatcggaacaggt cggcgaggactcgaagttcaatcgagggtgagagattaccgaaggcgcatggat cgacgcggctgtgtcggcaaggagacgctgaaaagctcgacatacaaacgttcacc cgctccggatcaaggcgatggggcatggggcgctcgatgtgcccgaaccgataagta cggcttccgaaagcgacccggacacacaagacgatgtttatgggatttcaaccggcg acgtcgtgaaggctgacattccaggtgga	atctgtacgcccgcgagcatgagaatgctta cggcgggactcaaccgactaagcgtctct cgcaagagagcgtcgtgagcgtcacacac ctcaggtacagcttttgcgaataatctacagt ctctggatgcgacggcgtccagactct gaggcttgcgattaaacaggcgacaacgc gaagccgctgtcgcgagcgacaacacgc aaaagatcatttgcggaataatcagctg acagcgatgtcagccgaaatcgaaagattt ccctttgaatatgagaatagccgt
300-301	AAAGATAC- AGAAGACT_L001_R1_001 (paired) trimmed (paired) contig_11 [cov=8 23346] GENOME_ACESSION: mgm4581772.3 GENOME_ID: 237567 CONTIG_ID: 9 SOURCE: MG-RAST DATE: 2019-01-27.	atggcgtactagtgttggacaagcgcaagtgtccgtgatgcttgcaccgagaagcgt gccccgtgtactgtggtcaggtcggggcggtgtgtgtcgaattcaccggttcacgac cgttgaaggatcggtatcggtggggccaccagccgggtgcactcaagatcgaccccg ggtcgcagacaaccggcatgctgcgtgtcgtgagtgatagcgtggccaagcatgtt ctctgtcgcagaactccgccaccgagctatctatcagcaagcaactgaccagcgc cggtaccgtcgaagcgctgcgagaagcagctacggcgtcggccagcagattcaaaa accgggtccgcccacagggtgtgttccccaagcacaacacccctgtcgacaacgtc atgactgtgggtcagcgattccgacgtcggcgccgacacacgacatcattcgagaat ggcacgggtcagacggcactcggggagaccctggcgatcgatcacatccggcagca aaacgccccccagctcagaaccacaggtgtcgcgagcgcctactgcatctgcacgat ggcacctgtgctactgtgaagggtgtcaggtgaaactgcctggaagtagagcatgt gcaaccacgaagccgaggtggcagtcagcgcctagcaaacctgtgtatcattgtccgg agatgcaacgaagacaagggtggccggaatgcggcgaatggggcgaggcattggcc aggtctcgaagcaggctggggcagacgcgccatcgcaatgccatgtgtcaatggc ggcagcgacctagcgtcgcgagaccctcgggcagtcacatgcaccagctcggcactatc cggctacgtgaaagcaacgggcctgcccctgtcttctgcagcggcgtggttaactcagg agaatcgtcagcccttgggataccgaagaccatgcatgtgagcagcctgctcgtggt ctagtgcagacgctgtgtgggtggaggcgccgacccctgtatcacggccgcggggc gaggaggatcggcggaacaacgtggaccgtcaggggttcccgaggagctaccggc cccggcgaagatgtcatgtgggtacaaaccgggtgaccagctagagccaccgtgcc gaccgaaagaagcgggcactcatgtcggcggggtagcgatagtgccggcagcg aggtggacatgtgacagccaccggacgtgtacagagcatcagttatcgccactgcgc ctgatccagcggggcgacgggttatgggtacgcaactcttccctaccacggatggagga agccgggtagcgaagtgaacgaaccgcccctctaccacccctattcgttcgaccga agtaa	gcctactgcccccggtgatcggcgagc ctgtgaacaggtcgggtggccaggagga gcggtaccacaaccgctacgtcggcaaca ggtcgttcgcatccactccgggtgcttcc cagccccggaccctggaaggtgaggtcat gctggtgaaaggcaaaagcaccgaaggctg acaccgctgccgaaggcaggagccaggt gtcgacattcccgaggggagagcggagcac ttatgctccggtcacaagcgtccgtgaagg cggagatcaggaggtatggcatggcc
302-303	Streptomyces sp. SAT1 chromosome, complete genome GENOME_ACESSION: GCA_001654495.1_AS M165449v1_genomic GENOME_ID: 102820 CONTIG_ID: 0 SOURCE:	atgactacgtttcatacaggtgagcagaccaccctgcgtacttctcagcggcaggt ctggaaccacggctgcagacaacccccggagctgggacgaaacggcgctgagcg catcaccgccagggtgatgcagccggatggaacatggcgaggggagaccgtacg cacgtcacctcccggctacggcgacccccgacgctccgctccggtgtgag ggaggagaggtgttctgtgagagcagcttccgaggcaggagcagtcacccagg gcaggggcacaatgtcgcgtcctcgcagaggaacgggtgtccgctgatgcctg atccccgtcagccccgagactcatgcccgttggcgccgggtgtgtgtgtaaaaccgc accgttctcatccgtcgaagaccgcgttgcggagagctccgaggtatccggcgttgc ggtagcatcgacccccgttgcgaagggtacgggcatcgccgtaccgccgatgtcgag tgtccgaccggggcagcggtgaggtacgacaagtcgccgggggtcgcagcaacg	gtcagccgctccccggcgaagaggccgg ggccccggcccccctagtggcgagtc tggttgaccagcccgatcaggtgaacg gaggtgttcgatgactacgtttcatacaggtg agcagaccaccctgcgtacttctcagc ggcaggtctgtgaaacccacggctcgagac aaccggggagctgggacgaacggggc gtgagcgatcaccgccagggtgatgca ggcggtatggaacatggcgaggggagac cgtacgcagctacccctccggcgtac

SEQ ID NO:	Name	IscB sequence	ωRNAs
	NCBI_Prokaryotes D ATE: 2019-01-27.	gaactccagcaccgtggcgacagatccacaagagcatgcgacagcggcgatcac cgccgagggcgacggcggaatctacgctaccggcgccccgatttcaaccggg cacgtcccaaggagtggttccccgtccctgcagcaccgggtggacagaccctgag taccgtcgtcgggtcacaagatgggtcccggtcactgaactccatgtcagcgtgctgct ttcgacaccggcggttggccttgatcgtggcgccccgaacagtggtgagtagcagca ggggcgctctcgggatacagaggttcgcaagtacctacgagaaagtggtgacgctcct gcgcctactgcggagcgaagaacgtcccgctccagatcgaccacatccacctgcgc taccggggggctccgaccgcatcagtaacctgacctcgcctgtgcttctgcaatcagga caaaagccgcgcgtccgggtggaggagttctggccggcagggccgtcagttggcgcg gctctcgcgggagcagacagccactgcgggacggcgagccatgaacggcacc gctggaagctttggcagggcctcgagagcctggggctgctctgtcgcgtggtccggc ggacggaccaagtacaccggctgatgcaggggcctcgcgaagtcgcacacactcgatg ccctcggcggtcggcagggccggccccagaccgggttgcgataccccggcacag tactgtcaccctacgctgtggggcgggctcctacgcacgcacgctcggacaagca cgggttccccaggtctacctgccacgtcagaagcagcaccacgggtttcgactggcg atctcgtcggggccacattccacggcgcaagtagcaggtacacacaccgggtcgtg gctgtgctgtcctcggcaccaccgcatcctccatccctggcgggtacggcagaccag ccacagcaacctgcgctcttcaacgaggcgacggatagcctacacatgaggaag gaggacgcgcggccttag	
304-305	NODE_1262_length_193 46_cov_6.071588, whole genome shotgun sequence GENOME_ACESSION: GCA_009843725.1_AS M984372v1_genomic GENOME_ID: 319939 CONTIG_ID: 76 SOURCE: NCBI_Prokaryotes DATE: 2020-04-22.	atgccgggtcaggctcgaacggggcggaaccgcagcggcggtcaggtgtgcga ggggagtcggcgcaagccgctgtctgtcctgcagccgggtggggcgaccgctgatgc cgtgccatccggcgcgagcaggacgcttctgaggaaagggcgcgccgtcgtggtgc gtctccatcccttaccatgcgctcaggggacagggaccggagggaagaccaggaggt cgccctcgggacgaccagggtcgaagaggacgggcatgtgccttgcgtgacgac ggcgaggtcctgtcccttgcgagatcgagcacaggggaaatagggtccggaaactga tgcttcggcgcgcgatgtcgaagggcgggcgatcagccagcctgagataccgga agaagcggttctgaaccggcgctcggggagacgggtcccgccctctctcagtcgcg ggcggaacagctgtcgtcgtggcgacgggttcagcaggttcggtcgcagtgaccggg atccgggtcggagacggtgcgttcgacatcagggccatggacagccccggcatcgagg gcgtggagtgccggcaggggacgctcggcgatagcagcgaagagtgcttctg agaggtggggcgcaaatgcgctgtcgtcgcagcgtcagggcaaacgctccagatac accacgctggcggaagccctgggggggtcccgccgggttgcgaacctgacgctgc ctgtgcgtcctgcaaccggcgagggatcgcggtcctggtgagggcagggcgccgttcacgacccg ccggcgatcgtgttcggaggtatcctggctgagggcagggcgccgcttcacgacccg ccggcgtaatgcagcccgtcgcgtccagttcagggtcgtggagcgaacggggcctcc tgtatcggcttctcaggggacgaacgaaggtcaaccggcgcggttgggagatcccc aagccgcatgcgtgacgcgccgtgtgtggcgagacgtcttcgcttcgggtggga ccagcccggttgcgaatcagggccatggggcgggggagccatgcacggaccgctg catgcgttcggcttccgtcgggtcgtgcggcgcgaaagtccgagatgggttcc ggacggggagctgtcgtcgcggcgccgcgcccggcgacggcgcaagggccttcat gccggggcgctgcgggtgcgcgctcgggtccttcaacgtccagacgggtcagggc accgtccaggggatctcccgccggcaccgcccgcagtcgagcgatgacgggttcc ggataccacatcgacgctcagcgaagaaggggactgcgacgatgcaggaacacgg aatcccgatctcccccggcgaggtgcagagtttctcggggctggaccgatga	gtcagcgaccctccctgaaggaggggc ttgagggcggaagccccgagagcccatg ctgaccagcgcgagaagagaccatcat ccaatctaccctgaaccatactgtgaagacc gaccttgggtgtgcctcaggtccaggt ctcgaaggttcgctcgagacatcgccgggt caggtcgaacggggcggaaccgacg cggcggtcaggtgtgaggggaggtcc ggcgcaagccgc
306-307	Ga0376441_00010 GENOME_ACESSION: stlouis1_FD_\$F_3300035 388 GENOME_ID: 312032 CONTIG_ID: 9 SOURCE: JGI DATE: 2020-04-07.	atgcaaaaagtgttttttaattcagagaaagccccctgacccctgccatccagcaa ggggcaggcaattgcttaggactaacgagcaagggttttcgaggtatcctttgtcatt attttaatcgcaaaaacaacatgcaattcaaccgattgagattagttcgtatccggga gtagaacaacgggagtgatgttggggaacaggtgagaacgttttaattccagaaaacc acgttgcgcatctaatcaaggccatgggaagaggatccaggcaaatgtcgggtgat caatttggatttccccggacagggcgaaagtctgcaaaaaccgtcaaaagattccaaact ggggaccttgaaggaattgttaccagggaaggaataatggcatatatactgtcga gtagctgtcggacatctgtgattttcaacatcaagactggtacagaaacgggtccaagg attagttggaatactgtatctaatcaaaaacggatggttatacttatgggttaa	gtcaactacccccccctgaaggacggagct tgtagaagctctcatgttgaccagcctacg aagcttgaacaaggctacgttactgaag aatatatactgtactttggaatgccttccagtt ccaacactacggataagtgtaaacaggtt aagggttaagccgggtccttatacaacaa accttcgataacattggcgagggaattta ccggcttaaacgccgagcgacaggtactg tcaatttaaggatgttgatgcaaaaa
308-309	human metagenome genome assembly, contig: NODE_158_length_5482 9_cov_5.54544, whole genome shotgun sequence GENOME_ACESSION: ODWR01.1 GENOME_ID: 3365 CONTIG_ID: 157 SOURCE: WGS DATE: 2019-01-27.	atgttagtgtatgtattaaaacaaaacggacaacctttatccaacagaaacgctttggcaa gggttcgtagattattgaagaaggaaaagcgaaggtgtccgtagagagccattcaccat caagttgctttatgaatctgaacagatgtgttcagggaattgtactgtgtgtatagacag gctcaaaatattgtgtgtagcgggttgggaatgacaaggtgtgtaccatctcaaca gaattaagaagcgacatcaaaaagaagatggggcgtagacgtgttttagaaaagtaag acgctcaagaaaaacattacagaaaaccaggttttttaataagaagaattctatcaa aaagatagacttccaccttcagtaaaaacataaagttcagggtcacatagatgagatagat tctgtaaaaagtatttacctgttccagacttaatttagaggtgtgtacgtatcaacaagtttga tgaataatcttagtttaataatgagagagtaaaacactggagctatcaacaagtttcaa ctatggctattcgttagaagaagtgcatcttcataagagtggttatactgtcagtgctgt ggaaagaagaactgtagattagaggttcaccacataaatttagaagtgatgtgtgtagacag atgatgaagaaatttgattactgtgtgaagattgtcataaagggtattcatgcaggtactat	gtaaatatgactagggttaaaaacctagtacc attaggtacgattgtttacagactaagggt agagaatttaacctacgatagttagatgaaa cacatacacaccctcggttgaatgctcaaga ctgaggtcgtgtgattactgattaggtgaggt gagagtgctaaaagaactcgtgtcagtaatt tcaaaactctaatatcttgcgtgagagaaga ccgacaccaacttttgtaacagaagttggt aggcgctcaggttcgataagttactgtcttaa gagttaaactgtcatcaacgaagggaattg aagtgtagttaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
		agtcttgaacaagaacctaagaagaataagaatttgaatacgcactcacatgagcata attaggagttggttgaagaaataccggacgctattgaaccttggttttattacagca gaaaataggaatcacttaaaattagaaaaagaccattatatagacgctgtgtatcgc gtggcggttagagttcaaatgttagatgtgattatagaaaaagacgtgtccggttcaa gataggattttgacaaaagggtgttcgagatgaagaaaaatacctaagcaaaatcatg gatttaagagatacgacaaaagttaataatcttggagaaatcatttcattaaagtagaaaaa gtattggagcgtttttgatggatattgacaacaactctattgatttagagatagaggagg aaaacagaatccaccatatagtttatcaagagattaaacacgagaagaagtgtttatgta ttaagaaaaatagaagagaggtaaggcttattcctgtattaaagctaa	
310-311	Ga0307374_10008841 GENOME_ACESSION: OX3MG_2_SF_3300031 670 GENOME_ID: 280772 CONTIG_ID: 8840 SOURCE: JGI DATE: 2019-06-13.	gtggctacgttgcgaagcaggcgagaaagaccaccaagggtgcttctcagtccttgc ctggaatccggccagcagacactcccgggtgggcacgaacggggcgccggagg caggccgagacggcctgcaaccggcttgcaacttcgtcaagaggagactgccttgggc cacaccgcccaagacggcgtcaccccccaacagagaccgctcctggcaaggcgct ggacggagcaatccgtcaaacatgtgtcgtgctcgataagcatgggcatccacttatg ccttgcacacggccaggccaggcaactgtcggcgacggccgggcaagggtcca ccggatcgccccgttcgtatccggctcgtggaccgcgaacagggcgccgcgacaact cagccgatcccttgggtagcagccgggtccaaaggcaccggattggccctctaca ggatcgagagcccccagccagtgtgaggtgcgccatgtgtcttccggctggagctccag caccgctcaggatcatccggaagcgagataggccagcggggcgccgacccgcgcg gcggcgtcgtccaacctccgttatcggcgcgggttcgacaaccgcaccacaagc agggctgtgctggcggccacccctactctcgtgtagcagcatcgacagctggactgcg cggcttttgcggtggggcccaatcaccgggacggatctggaacttgcgcttcgacatg cagcagatgcagaaccggagatcagcggcgtggagtaccagcaggcgaccctgcc ggctacgaggccgggagtagctgtggagaagtggggccggaagtgccgctactgc gacggcgagaacgtccccctcaacctcgaccacgtggcgcccaagtctcgagcggct ccgaccgggtctcgaacctgactcttgcctgcatccccgcaaccaggcaaggacaac cgccacgtcgaggagttcttgcacatgacgcaagcggctgggtcgggtgaaggcc caactcaggtcggcgtccgcgacggcggcgtgtcaatgccaccgctggggcgtca agggcgacttgaggcccttggcacgctgtggcctgtgtgttgaggggcgaaccaa gtggaatcgccaccggcttctcgtcgcgaagtctcagctcggcggcggcgtctgtg gggagcttcagggtgtgcgcatgctggagtcgtttatctgtgagcggccactggc cgggggaagtcaccagcggacacgacccgacgcttgcgctcgcgcgctctacttgc ctgaagtcgaagcgggtccacggcttcgccaccggcgacgtcgtgtcggcagcttcca ctggcgcaaggccctcaggacatcacaccgggtcgggtcgagatgctgtccttggct cggttcgctcggaaacctggacgggatcaattggcaacactgtcgtgtcctcagcgg ccgacggctaccgctacacacaaggagcgctcctcctccggccctgaacgacggg tctccggaccggcgggatttga	gtcaaccaccccgccctgaaggacgggctt gtgaggcgccgagccgcttccgggctct gctcgaagggaacgggtgaccagaccaa gccgctgacaggaggtgtgacacaggtg ctacgttgcgaagcaggcgagaagaccac caagggatgtctcctcagtccttgccttga atccgcgccagcagacactcccagggtg ggcacgaacggggcgccggagcaggc cgagacggcctgcaaccggcttgcacctc gtcaagaggagactgccttggccacacc gcccgaagcggcgaccc
312-313	Ga0111052_100017 GENOME_ACESSION: IMG_3300007996_SF_3 300007996 GENOME_ID: 28344 CONTIG_ID: 16 SOURCE: JGI DATE: 2019-01-27.	ttggatctgtagctgtaataagtgcatcaggcaaaaagctgatgccaccaattcttaca gagccagagggtctttaaatacaagagagctgtacacacaaatagacactgtatttac cattaaactgtagataggactaacgggtatcacacggcgatagaatacaaatgtatatac ggatatcaaaatcgtgtattccattgttcagatcaaaaagagctgtcaacgagcaaag agatctgcttaagacgaagtaaaaaagcatagcgacagccgtaaataccgcagaacca gacaaaatcgcttgcgttacagaagaacgattcaataaccgcagggggtatgattagca agatgggtttgccccatctatccgcaacaaaagagcgtacataatctgtctgtatgagga gtattgcaaaatattccaatcgcaaaagctatctttaaattggggcagtttgatcgcagg tcttaaagccatagaagctggacttccagtccttaagagcagattatcagaggggtg aacagtatgttatgccacgttagagaagctgtatttgcgaagatacatataaatgtatct gctgcgggaagtcagcaatcaagatgggttaaaactaagaatccatcatatcgatatac tcatgtggggacagaagcaacagaatggcaaaccttgggtctgtatgtgaaaactgtcata cgccaaagaaccataaacagggtggaaagtgtataacctaaaccaaagctaaaagaa ctcaaggagcgaactttatgacgaggtccgctacagcatgtgtaaaaatcttaaagctg caacaccaaacgtagatatcaaatgtacatgtgagcagcaaccaagctcgcagaaaaa gatctcaatgttaggaaaaacgactctaacgatgatactgtatgggagaatttcatccaaa acatcgaagcgaacttttagcattacaagaagtgtcagcgtataatcgtgtgtaagtaggt tttatgatgcaaaatattgtatctagagatgggttactaaacaggctcagttatctctg cggacgaacaaaccgaaagatctcaagacatataaagcttgatgaacgtattatcagact cacaagatctccaaggggcggtgttctacaagaagacaactactcattaagacctgga gacaagggtttatagatggaaaagtatactctgtaaaaggagtacagagtaattgggacc agcatcaaatagcgaatggaaaagtgtgccatataagaaggttaacattatgcagcact gcaacgcgatgggctttatttgcataa	gtcaacaaccaaccacttaggctaaagcctt gaagtgggagcttttgaaaaaagccctgtt gattagcctaagcactcagtgctactgttaa agcagaatagataggtagtcagaatgcttc acaagtctgacctctacggatataattaaac atctctgacgggttaggagaagtgtgtatc gaaaaacctgttataacattggcgtatgtaa ccactcccgtaaagggaagaaccgacatct tcatgtcggaggcgtgaagccaaatattta gaagaaggaggtatcagtttggatact
314-315	Ga0210003_1012872 GENOME_ACESSION: LanmetSPAdes_SF_3300 024262 GENOME_ID: 31358 CONTIG_ID:	atgcagaagttagggaaaaagctaaagacgtacctatgatactcacagtcctatagc tctataagttcttcttaaacagagttcaagactcagtgagaagaacatgtgtctgtccat aacaatcccgaagtgaataacacccaatgtaaaggggcgagaacttgagagtatctacc aaagtttacattatgaatgtcaggtatcccttaatgccatgtatccgagaaaagcaaaa aaattattaaaagaagggaagtagtattgtcaaaagatttcattacaaatccaattgtta	gtcaaataccacgcactaaagtacgtggctt gactcgtgagactcaacgtgaagagttgatta gggggcctaaaggaaattatgcagaagttag ggaaaaagctaaagaacgtacatgtgata ctccacacgtccatagctctataagttcttctt

SEQ ID NO:	Name	IscB sequence	ωRNAs
	12871 SOURCE: JGI DATE: 2019-01-27.	attccaacaggggaaacaaaaaaataacattaggaattgattctgggtatatgaata ttgggttagtgtagtatacagaaaaagaggaattatcttggtacagttaaatagatgaa agaacatcagaagaattaacagaacgtaggatgtatagaaaaatfagaaggagcaactt tgggtacaggaaatcaagatttttaaatcgtccaaacaaaagactggcttccatctat acaagaataatgatgcacatttgaaatctttaaacagaattaagaagattatctgtttct gaaacaattattgaaagtagcaaatctgatattcaaaaattatgaacctgatatttcagga acaggttatcaacaagggtctctatatgaatatcagaatcagaagttattatgaccaga gaaaaaggaaaatgtcaactttgttgtaaggattttaaggacagtcacatatacatca ctgtaacaaaaggaaatgagaatggaagtaataagacctaataatctgtctatttacataagg ctgtcataaaaaactgcaa	aaacagaggttcaagactcagtgaagaa catgtgctgttccataacaatcccgaagtga atcaacctcaatgtaaggaggggcagaactt gagagtatctacaaa
316-317	Ga0310695_10070301 GENOME_ACESSION: RumRJGDNav2_\$F_330 0032007 GENOME_ID: 280889 CONTIG_ID: 70297 SOURCE: JGI DATE: 2019-06-13.	atggctctatgactagcgcatcagggaagccccctgatgcctacaaaacgttcggacat gtaaggcgcatcctccgcgcggcaggggcgaagtcgtaagacgcacgcccctttacga tacagctgacgtatgatggcagcggttacccagcccatactctggagtcgacgccg gggtcaagcacattggctctcggcgacaacgggaacgtctgtctctacgaggcagac ctggagctgaggaaacacatcacagggtctctcagcaggaggaggagtcgagacgtt caaggcgacggcgcaagacgggtacccgaagccgagggtttcacaccgcacaaaat caaaagcataagggtggctggcgccctctgttgagcagaaggtagacacacgtgaca cggtgaaggaaaggcctgaaatgctccctgttttcggagacacggcagagacggcg cggtcgacacagcttttaaggcacaggaaaaggatttccgtcccggaaaggagaa ggctaccagcagggggaccagctcggtctgtgaaacgtaaggaggtagctactgttca gggacggtctcactcgtccgtctgtaggggggaagtcacaaagaccccgtactggaagt acaccatatacagagccgtaagaccggaggcgactaccggacaacctgtgtgacgctg tgccgtacctgcccataag	gtcaactaccaccgcttaggcgtgagcgc ccttgaagcggggcgttaggagaagga aaagcgttcttattcttctcctcaggccc gggtgaaccagctgagtcctcggcactac gttaagagggtcaggtaccggtggcgta cagcctaactcgggtatgcccgggtgg ttaacagtcctgaggcgaaggacagtcg tgcggcatgacagcccttaacattggc gcaggtgcactacagcgaagctggttta ctgactattgtcagttaaaggaggacgatg
318-319	Ga0137371_10000096 GENOME_ACESSION: SagtaG_43_\$F_3300012 356 GENOME_ID: 35430 CONTIG_ID: 95 SOURCE: JGI DATE: 2019-01-27.	atggttttctgctgtatcgtcacaagaagcgtgatgccaccacacccaacgagcgc cgctctgctggctcagagtcggcggtcgtgcatcgcgtcagcccttctgtcattcgat tacgcgaccgacgctcgaggagagtcggttgcgaagggtcggcctcaaaatcgtatcct ggcagcaagactaccggcatcgccctggctcgtgtggaagaggagaggttcatcacg cgctcttttggcgaagtggcacatcaggggaccaggttcagagaacaaagtcagg caagctcaagcccgcagagctcgcaggagtagcaacttgcgtaccgcgaccccgttt cgaccaccggggtatccccgctggttgctggcaccctgcctgctcctcagagttggca acacgctcgcctggagcgcacagactcacgctggcgcccccacacgcttggatgac gagcaggtgcgttcgacacgagttgcttcagaaccccgaatcggggcggtgcagt atcaacggggcgaactggcaggttgggaacccgagcgtacctgtcctcaagtacgc gtatcagtgctctattgtgtaagacggattgccccctggagatcgaccatcctcgtcta gaagcagagggggctccaaccgctggcgaaatctgtgctctgtgtcagcagctgcaac caggccaaagggaacaagacggccaccgagttggacaccccagggtcaggcaca agccaacgccccctcagcgtatgcagcagccgtcaatgccacgcgctcaagctgttg gaggccttgcgggtctgtgacttccatcgccacgttgacaggtggcggaaccagat ggaatcgggcgcgtttggcgtggagaagacgcacgcttggatgcttgggtggga gaactggcggggtcagcatagcagggtcacaacactggtgtcaaggcgaccggcg agaggcgagcactgccgaaccaactgggacaggtacggcttccgcgtggctacaag atcgcgagaaaggtgtgtagggcttcaaacgggggaccgagtcgagcggctgtg ccagctcccccaagacggcggtacgcacattggacgcgtgaggtgcgcaaaagc ggctcgttctgtgcaaacccgcgacaaggacctgatggatcgtgtgcactacata catctgatccagaagcagatgggtatgtagtctgtgaggagacacaaacacac aggaaggccacccctccccctcccccaagggaacggccattcctccccgtcctga	gtcacgaacccccggattatccgggggct tgcagcggggcgcgctgttgcgtcgc aagccctatgtgtccagcctcagcgtcaa tggcgtaggcagccacgtcggcgctcc gttcgacaggtcaggttacgccggagct tcttcaatctgctcactaccgtccagggt aaaaggtccgagggggcagaccggtgcc ctgggcaagacaagcctgttgaaccttggc gagaagagatgcctgcgacggcggtgca ggcgctcagacggcccccttaccgggact cttcttctgggaagagaaaggtgtccag atg
320-321	human gut metagenome genome assembly, contig: NODE_633_length_3411 8_cov_2.922379, whole genome shotgun sequence GENOME_ACESSION: OPBE01.1 GENOME_ID: 8983 CONTIG_ID: 632 SOURCE: WGS DATE: 2019-01-27.	atggtgtacgtacaagataaattgtaaacctatgatgcctacaacagcatgtaagg ttaggagattgcttaagacaaaaggcagtcgttgaacctaattgctttacagattaat taacctacgaacaacaattacaacaagaattgtgttaggcgttgatactggcactaa gcatgtaggaaattcagcaacaacaaaaggcaaaagactttatagtagtgaagtatcctt agaaatgatafctgttgccttttggcaacaaggagagagcgaagcggacaaggcgtca agattaagatatagaaaatctgtttgataatagggtaaaatcaaaagcgtgaaggatgggt agcaccttgcgtgaagtacaaaatagacgctcatattcgtgttattataatgtttgtctata ctaccaatatctgtattatcgaagttagctcaatttgatactcaaaagattacaatcctg aaatatcagtaagaatcagggaggaatcaacttggttttggaacgtaaggagata tgttttggcaagagatgggcataatgctgtcattgtaagggaatacgaaggataagatc cttaattgtcatatattgaatctgtaaaataggagggtatctccgtcaaatcttgaacct tgtgtgaacctgtcacaaggaatatcaaaaggtaatatagatttaaaatcaaaagg taaatctttcagtgatgcagctaatggaattatgaatgagactttatgactttaaa atcaaaatttctaataattcaatgattttgttatattacgaatacaatcgtattcgaataac attgaaaaatccatctctgatgccttggatttcaataactttaatgcaaaaagggttagg attcttataagataaaattggtagaagacataatcgtcaatacataaaatgaaaatcca aaaagggtggagtaaaagagacttaataatctctttttaggttttggattcgtttgttgata gagtgaaaggttgacaataaattctttatctatggaagacgtaggactggcagatttaac	attaagttttatactatcttgtgttgaatga aagttcaattcggttattagcctaagccttga acagaggctacgttatttaagaatatatagtta cctacggatgtttgccaaagctgtactcta aggtaagtgattaaacagttttgtatttgaga aacagttgtgttcaaaaaactttaataacat tggcgatgggtactaacagattttactctga tttatgttgaataaacattaaaactgtttgtag atatg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		attcgtgataaaatggagaaaatcaaaaggatgttacgcataaaaagttaaattgtcaag gtgtaagcgtttatgggaaaattgaaatgaaataa	
322-323	Ga0129284_10010797 GENOME_ACESSION: IMG_3300009514_SF_3 300009514 GENOME_ID: 311720 CONTIG_ID: 10796 SOURCE: JGI DATE: 2020-04-07.	atgggtgatgtaattaacaagaacggaaatcccttgatgccctgtaaggcagcaaaaggcc agaaagcttttgcgaagtgtaaaagcgaaggtcttctcagaattccctttgtaattgctctgt tgtgggattgtgaagaaaagggtcaatccgtgatcgggtgaatcagatggcagcaagg taattgggtcggcagcagttggaaatggaaaagtgctgtatcaggcagagactattcttcg aggcgaggagatcaggaagaagatggacaagagaagaatgatcggcgaacagaag aggaagaaagacccgctaccgagaggtcaaatcttcaatcgtatccgtccacaagaa aggcgagactggcaccacgcgtcaatcataagtgtagaagccatctgagagagaaggg atttatggaaaaattctcccgatagcaaaagggttgctgaggtggccagtttgataatc acaggattaccaatccagaagtaaggggcgctggtatcaggtggcaacaagaagg ctttacaacgccaaggcctatgtgcttcacaggagcattataaatgccagaagtgcagg gcaagaattgcgcccctccatgttcataataattttcggagcaggggaggaaccaact ctacagagaatctgatccctatgcgagagttgccacaaccgtttgcatgaggagagaatt cgagatcaaaagcgaagatcgaaccaga	gtcaactaccccgccgctgaaggccgagct tgtaaaagctcaagttgaccagagggtcaac agaggatctgttaattgttagcagccgttaga accgaaaatcacatataccatcggtatgcc acccagctccggcgctctatggcctcagatt aaacagctctgtggtagggacagtgctgtg ggcgacaaaaccgggtctaactcttcgagg ggcaggtcggattccagtcgagactcctacc gactgatacgcatacctcaacatacagg aggctctaagaacatg
324-325	Human gut metagenome isolate S6163 scaffold_1873, whole genome shotgun sequence GENOME_ACESSION: QZLU01.1 GENOME_ID: 14549 CONTIG_ID: 1871 SOURCE: WGS DATE: 2019-01-27.	ctgactaggcgccggcttgcggagccgaagagatagagaacctgtatggcgctgttta aatactcttggatgttggaaagtgtgtacagaaggcgtgaagccatcccgcttgcggg gaaaataataacacaaaggagggcgacttatgtcagttgcagtcattagtaagacagtcg aaagattgatcccaacagtgtaacacagacacgttaagttgctcaaatcaggcaagca ataaaacatagttatcctcgtttaccatacagcttactgaagggaactggagatacca gctaataagactgtgtatgataccgatatatacatattggaatctgtgaagtcaaaagaa gcacgaatatctggcagaacagattgacacgttcaacagacgaagaagcaagcatgatg catgccgtatgtatcgaagacagagacggaacagaagagataccgtcagccacgattc aataatagaagaagataaaggctggttgccttcttggaaacacaaaagaaaaata catattcaggcaatttcacgcatcagcagagattatgcctgttactgatatcacaatggaaat gggaaatttgcacacagattctgaaatcctaaggaggaaggagaccattaccatgg cgagattatcagcagcggaacgttacggcatcgcgacgttcgcgaagcagtttttct cgtgatagtataaatgccaatctgcggaagaacaataaaagacggagcaatactccat attcatcataaataacagagccagggtgggaaccaacagatgtcaaatctggtctacgg tatgtataaatgtcataccgaaaaatcacagccaggaggaactatataattggaa accaaagctttcatcctttaagagagcaaccttcagctcgatagcagatggcagctttaca atgaagttaaagctttattccagatattgataatcacatt	gtcaataaccgcccagaggtaaacctcg gacggggcttgcacaaagaataattttcttat ttccactttgtttatgccgtttacagtttacgg aaagctgttggctgtaaaagccttattgattag ccttagcagtaaacgaaccggccaaagccag tcaggttaaatcttaactgctacgttaccgg taaaataggcaccgttgggtgctcctcaagt ctcacgctctcggtatgtgttaaacatctct tagggtaggaagtgcatatgcgctaaa ctatcggaatacattggcagtaggacaacc gattctgactaggcggcggcttgcggagc cgaaagtagagcaaacctgtatgggctgctt ttaaatatctcggatgtttggaagtgttgg acagaaggcgttaagccatcccggcttggc ggaaaataataacacaaaggagggcagctt atgtca
326-327	Ga0315288_10027781 GENOME_ACESSION: YL17G11_20_MG_2_SF _3300031772 GENOME_ID: 281639 CONTIG_ID: 27780 SOURCE: JGI DATE: 2019-06-13.	atgtttagagtgccctgtataactgggtgaccgcaaaagcgttgatccctacgcatcaagcga atcgagagtgcttataaagagcgtctgtactacacgcttttttcataaaggattttttgcatt cgtctcaataaatggttaagtgaataattcaacctatatacaataggcattgatcctgtgtcc aaaaaggaagcatttactgtgatgtgcgaacatcacactctgctaataattcaattggtatgc ggtgactcatgttaagatgcagtggaagcgaacgaatgctctagagcaagacgct atcgcaacactcgtgcagaaaaccaagatacaaaatgcagaagaaaaggatgggtt cctccatcaaccaaggcgaagatgggataataaataatccaattattgcttgaataaaag tttatcctatttccgtgtgtgttgcgaagatgtgtccgcaaggacaatgaagaatggcaaa agatggacaagctcttctagctctgtacagtggtggaagcactatttctgtaattatataa tcaacttaacttaatttaactttgaaagaaggttgggaaacggctgctctacgagcaattat tctctacctaatacaagaataagatggataatacatgggaagcacattgctgtgactctgt gggttctgcgaagcgttccgcttgcacacccaattgtaataaagcaatgatagttgcaa agccattgcgggttaccgcagacaattgcatgtcttcaaccggccaaaggcggtattag aaagtcatatgatacaacaggagtttggacttctagaggttctattgttactatccaaa atttggaaaattgatttagggaggttcgcaaaaggacgaatatctttacatagcatttccga caataaacgattaacacagaattctaaaatagaagatattaaattttgtgtataacaattgg aattttaatttaatccaccggctaa	gtcaactaccactgattaaatattatgggc tttagagtgaaacaggtgtctacgaggtag cactacgagattcaaaagcagttgactagcct atgtcagttgttcacagcaaacgtgaatgg atatgacgaactgaatcaactgactatagca ggtcgaattcgttcgaatgctgctatgtttg acacgagcgtatgtgtcagtgccgaagatgt gacagggcgttaagcctatataccaccg taagggtgagattgctcgtagagcctaaag gagatgaattttaga
328-329	Ga0123349_10013493 GENOME_ACESSION: ElkDunMetagenome_6_\$ F_3300012983 GENOME_ID: 25860 CONTIG_ID: 13492 SOURCE: JGI DATE: 2019-01-27.	atgtcaaacatattgtatgtcgttactatagtggtacacattaagtccaaccaagaagtt aaagcatgtgatcagctaaaggtaaaagcaatactcgtatcaataatcctatggttatt aaacttctaaagttatacctaagaagatataatgaaagatgaatacgaatgggtattgat gatgggccattacacaggtatagcagtagtacaaaaaggtaaaaagtataataaagttc tatttaaaaggacatagaacatcgtaaaagatgttaaaagaaaaatggaattaaagaaaatt tatagaaggaatagacggcttaataaaagacatagagaagaagatttaataatagaactt ctataaaagatctaataatagctccttctataaacaagaagaacagctaccattagagtta ttgataataataaataatttaataatagattcttataatcagaagatgtcaaaatagatattag atgtctatctgatgactatacaccatataaatgggaataccaaaagcttaataaattagataa taacctaaagatagcttataataagataattatatacgtcaaaatgtgtggtgaaaaaga aagggtgacttgaggtacatcatattacctaagaagctatctgtgtaacaataatagata atcttattacactatgccataaattgtaagaagatgaacaataaagaactaaataatata cttattttcataagatacaataatgaagataaagatataataaaaaacttaaatatgcttct catgtaatggtaggaaggttatttacagaatagattaaagatagaggtatgtcttttttaa ctactggaggtgatactgctaataagagaaggtattggaaaatagaaaaactactctaa	cctttaaataattataaagtgagcgttagtc tttatattattattattgacagcttttagccta agtgaagtgttctgtgtgaaatcacgctgca tgatgaactacgttgaatagtaaggtaaga cacaccctcggtatgaattcttagcgttgc tctgtgagtgccaaccaagaacaagtgcta acgtctgcacggattatcatagctaaatact atgagtaacagggaaacacatgctcctattt gacattggcgaagaagaaaaactaactttg attgtaataataaataattatataaactttaga aagaaggtgatagaatgtcaaac

SEQ ID NO:	Name	IscB sequence	ωRNAs
		tgatgctatgtataacagatttaaaaccaagaagagacactatagataaaaagattgga atattaagccattaagaagaagtagataataaaaagaaaacgataacttagttatattc agcatagagattatgtttacatacaaaactaaagataataaattacatgaaggatataaacc gcgttatccaaataaaaaattacttagttttaaaactaaaaataaagttataataaataa gtagtaaaaaatgcaaatcttctggcattttgataaaaaatgtatttataa	
330-331	Ga0117908_1041818 GENOME_ACESSION: IMG_3300009314_\$F_3 300009314 GENOME_ID: 311702 CONTIG_ID: 41817 SOURCE: JGI DATE: 2020-04-07.	atgcctactaagccagctaaggcaagaagctggtagagggtggagtgctaaagaat gctggctaaaggtagtgctgtttctacataacgatgttaaatctgtggcgaaaggtaca ggatgtggcactggctattgacccaggcagcaaatacgacggctacgccgtatcaggct caaaagatggcgactcaagcgaatggcagtgatgcctcaaaaggtacacaagaaggt aacagaacggcgagtttaaggcgacgagacgctatcggaataaacgacatggcaa gtgcaatttaacaatagaagcgtgaagcagggtgatgatgacacaaagccaattagctaa gggtcaattccgcatcaagataatcgcgacctgttaagggtattccctatcaactcattgc agtagaagatgtcgttaaccattacaagaacgggtgggcaagcacttttctactgttg agataggtaagagtagtctatagcaggttagagaacatggcaaggtaataaagtatg caggctggcagacagcagaagcgaggcagttggggtattaaagaagtcgaagtgtctaa ggatgctttgacacctgagtcctatgctaatgatgctcttgcgatgctcaatgaggtattcgg ggataatgtggataattcatgcatattcatggtatggcgctgacttgagttttcaagacgctc attgcacggcaaaattatcagaagggtgtactgtccccgcttggcggaactccaac ggacattatctccgtaaggagatattgtctatggagagattggagacaggcaattgtgg gatgggtatgtgattaccgacagacaggacaaaagccgttgctatagctgatgcttcag gtaaacgttttagcacagctctcgtgagcaaaagtgagactgatacgagaccactgtga tcacatgggaagtcagtataaacagggtccctatggccaatcatagtacaaaagccga ttcaattagatatacctctggcaattctcccccttag	agtcaataaccctcctgaatcagagattca gaaggggcttggggacacaggactcccg acgcaaggtgtgactagagtgcttaggaact cctgagcagccgactcgtgatgtacatac gttaggggttttcgctagctcttacctataa actgtctcttggggcagtgaggataaagcc ccgacatactcgcagtcgaatctcaaaagc gacctttacttagaggaagggtgccccctt ctccagcctgtatcaggctccgagaggga gacaatcatgaga
332-333	Ga0395718_000631 GENOME_ACESSION: 713C1X0metaG_FD_\$F_ 3300037107 GENOME_ID: 312125 CONTIG_ID: 630 SOURCE: JGI DATE: 2020-04-07.	atgataccagtaattgataataaccagaacacattaatgctcctgcagaaaaagagcaa ggaaatgattgaatccaagaagcaacaccattttggaagaagagattttgtattag attgaatcaggaaacctcagcaaggaatctcaaccaatagttgtgggattgacctggat ctaagaagggaaggtatgacagtaaaagcggaggtctatcattattaaatcatcaatgcag atgccgtaacatgggtcaattctcagttgaacacgaagaatgataagaaaaacagaa ggcgtagaaaaaacacatgtcgcaaaagagattaaataaattatcataagcctgggt ggctggtccccatcgaccagggccagatgggagtggaagattgatatgtaattgctt gttaaaatgtatccactactgattttgtaattgaagatattaaaggcaaacgcaatgtgcata aagggttttctacttgaactagttaaaagtgttttataaaaaatcttcaaaattaggagt gggtatattagacaagggtgggaacaaagagcttcgtgatagcttggattatttaagt taaggaaaaattagccgaagtatttgaagctcattgtgtgatgttgggtacttccaactg tattgttgggtgacattctaaaccagagaatactagattgtatgtataacaccaattagatta cacagaagacaactcattatctacgccagaagaggtgttttagaaagacgttgggtg gcactagtagtgtgtttgaaaagaggtatcatattttataaaccaaaagtgtgggcttttt atcttgggtgtgaaatgaatgttaaaataacatgctttcaattattgatgtgaagaaaaat ctcaaaaaacaatcattgaagtaaaatttaactgtataatcatgtgaaaagacagta atagcataa	gttaattatccagaaatggattggaaggagta aaaaagccaactttgaaacaacaattaaaa tagaaactagagtggtcaagattcgcgtaaa ccgggtcgtttggaaggcaactgaattaaacta accatggcaggttgatgatattgaggatgc ctccctagctctcattctcagttattcagctc cgaagggaagtaatacaagggtgcgtgaagc acatttctagaaataatgatacca
334-335	Ga0187860_1009162 GENOME_ACESSION: June2016WEW_10_4_2_ \$F_3300018014 GENOME_ID: 30387 CONTIG_ID: 9161 SOURCE: JGI DATE: 2019-01-27.	atgcagaagttgaagacagagaactaaaaagaacacacacatcggtactccacaagtcgg tagcaactgtgacagttattaaacagagaccaaaagctcagtgtaactgatttaaaacct gttttaacaatcccgaagtggatctacccaacaacatagttggcgacttaagctaaggt ctatgtattaaataatgatgtacaccattatccgtctatgctccatgaaagctagaaaatt gttaaaaagtaaaaagctgcagtaattaaattatcccttcacaatcaaatgacttatga aagtgaattcatgttcaagtcattaaattaggcattgatagtggttataaaaacattgtttt ctgccattactgaaatgaagaattatcggaaggtgaattatgatagataaaaactagt aaagattaaagtacaaaagtatgtatcgtgacttagaaggagaaaactttgtatcgtaa gcaagatttctaactcgcaaaagaatgttggatttttaaatcccgatatacaagacgtta taatacacacataaaataatagataaaataaaaaacttttaccatttccgaaataatttg aagtgtctaatttatattgctaaaatgaaaatcctgataatgaagggaattgaattacaaga gggaaatgtatgtttatcaaaagcataagaagtatttattggcaagagagaaggaaat tgtcaattatgtgtgaagacattaaaaataaatacaagccatattcatataattcctagaa gtagggtgggaacaaatcgtcctaaaaatcttgcataactcatgaagattgcatgaaaa attacacaaacaaaattttattttataaaaaaatcaaaacataataaactcagaattttat gtcaattataaataaaaggttcaacaagatattccaaattaaaaattacttatgttatatta cttggattataagaattaaagtttagagaatctcatcataatgatgcttttataattgcg ggaggagggtactaatcaattagaattaaacaaattattataatacaaaaacatagaataa tagaaaacttcaaacacaaagaaaaggattaaaacgaggtattagaagaaaaatataa aatcaaccattagattttttggaattataataaaaaatttttctaaagggtatgtgccataa tggagaacgagtgatgataataaaaaatgaagtttttacttaaaaaaggtgaaaaaataft tcaatttggcacttttgttttaataa	atcaattgtcaacgaattgttattcaaatcgtg aggtttgagtgtaagtgttattaggagactta aaaattaaatttatgcagaagttgaagacaga gaactaaaaagaacacacactacggatactc cacaagtcctagcaactgtgactagttatta aacagagaccaaaagtctcagtgtaactgattt aaaaacctgttttaacaatcccgaagtggatc tacccaacaacatagttggcgagcttaagc ctaaggt
336-337	Ga0074046_10018987 GENOME_ACESSION: BogForStandDraft_3_\$F	atgcaatctacgttgaattgcaagtcgaagaccacacccggagtgcttctcagctccgg accttggaagcttgggtgacgttaccgggaaggtgaagcccgacgacgagacgg gttcgctacgcgcggcagggcgtgcttcaacatgtctgaggggagagtcggcaagg	gtcaactaccacgactaaaagccgtggcg tttagaaatcagaccgacgggtgaccag accagaagacgaaatgcaatctacgttg

SEQ ID NO:	Name	IscB sequence	ωRNAs
	_3300010339 GENOME_ID: 23904 CONTIG_ID: 18986 SOURCE: JGI DATE: 2019-01-27.	accgtcaccttctccgcaagagagaagagagcggggtaacccgcgttttctgtgtc gaaggacggcaggccgctaataccatgccacggccgagagccgagaactgctaag gaaggaggaggcggtcattgtgcgccgtaccggtttgttattgccttaaaacaaccc gaaccaaccacaacgcagcctattactatcaaaatgatccggagcagaacgaccg ggatgcacattgtgagactaaccttcatctgctacatcgttttgcacctgtccgaactgacg catcggggcgcaaggattagagagaacctgatcagcggcgagggttcagggtctaac gacgaaggcgcaaaactgggtatcgcgatcgagattcaataaccggaccaggcgta agggtggctggcgcaagctgtcaatcagggtagacaacgtgactcatgggtgcgc gttaccgatctggcgccgattacggcgatcgtcattgagacgggtccgtttgatacgc agaagcttatcaatcctgaataatcgggtgtcagtagcaaacagggcacattgtttctac gagcttagagagtatctgtgaaaaatcagcgatcgtgcgcctactcggaaggac gaactggccgttgaaatcgaccacgtgatccgcgttccagaggtggaccatgagcc ctaccaaccttgcctggcggtccacgggtgcaaccaggctaaagggaatcattggtg gaagacttctgctgagggagccagaacgctgaagcgtatcaaaagtcagctcaaaag cccactaaagcaaccgcagctgtaaacggcgcgcgcaaaatcctgagcggactat tcaagacgaagctaccgtagaataagcaaccggggcgcaagaccaagttcaaccggg ctcgtctccatcccgaatgcaacgcccgtgacggcgatgacaggtgatacgcga gaattgtgggtggaatgacccgtgctgctttaaagccggcggaagggtcgtat cagcgacgcagctggacaagtacggttcccaagggtacatgatccggcagaaaa aagccaaaggcttcagaccggcgacatcgttcgagcatcgattgccaaggcaaaaag gcaggcgatgatgacggcgatcgatccggcgagcggttcattcaatccaatc cctcaccaccacattcaaggcatcggtacaaaaatgtgctgatccagcgagcgga cggctacaactctgcaataaacaaggacagcgattcatccacggctga	aaatgcaagtcgaagaccactccggagtg cttcctcagctccggaccctggaagcttggt tgacgttacgggaaggtaagcccgga cgatcgagacgggtgctacgcgggac gaggctggctttcaacatggtcgaggggag agtccgcaaggaccgtcaccttctccgcaa agaga
338-339	Ga0307380_10077270 GENOME_ACESSION: UN3MG_2_\$F_3300031 539 GENOME_ID: 281081 CONTIG_ID: 77269 SOURCE: JGI DATE: 2019-06-13.	atgatttacgtacagaacgcagacgggaacgccgttaattgccgacacacgggcaggat atgtccgccggttatgaagaagaagaagctattgttatgtcagaacacattgtgtga ggttagcgcagcagacacatacaagctgtcaaccccttattttgggtagacccctggga tgacgattgggttgcgtgatacagacgaacggggatcccttttactcgagagctaac cacgcgggtcagcggaaatccctgactaatggaagaacgacgcatgcatagaatggca cgccaccgctatcgacgaatgagaacccgtagcggggcaaaaaaggcgggaacgata tatgacggagagcggaattccaactccaggggacaaacccgtgacgggtgacccgctc gttcgttcatgccaatgcaataaagccgaggttagcggtgttttcaaccggacacgacc ggacggatggctaacaccaacggcgctgcatttgcgggcaacgcatacgcactgttg attatcttctgattttccaatatctggattgttgaatagctgccttggaccaacag aaactgatacccccggacatacaggaagggttatcaacaggtcaactcttgggattt gaaaacgtcaagcaatcgttctggagcgtgatggacacctgtgccaattatgtaaaaag aaatcacaggtatttctccacgttcacatgtaatttggcgagtcgaaggggcggaata cgcaaaaaaccttctgacgctctgtgacggatgccacgacaaggttcatacaca aaacaacacaatgcttcaagaaaaatgccggggatagcaaacgaaagatgaagac aacgctcatcaatcagattatgccacattttatcgggtggtagaatgaaacacgatgctg ggcgaacgtatggattcctgacaaaacacgtcagaaaagagcttgggttgcaaaaaca catgcccttgatgcctatattatcgcacttctggggcaccacgacggggcgatcaata tggatgttaaaacatttccgtgatgtatgggaaacaatctgctgccataatcggaaga gattacacgacaacagatcggaatattataatgggaaaacgtgtgttcgaaaaaccg ggcaaaacgcgagtcacagcatgacgacgtctaactgaatacgtcaacaattcgggg agtcatatgttctgcaactgagggctgcaagggaacccatgctaaaaaagtgggttact ctggtgcagagaggagataaagtcgatattgggtgacatgtaaaagctgcagaggggtt gggagctatgggacaagattactgattgagggcgaaagggaagcaccaactccaaccc gaaacgccttcttcttttaaaaaacacaggcattgtgtgggggaataacatgga	attgacacattacttcatctcgcgcgctgca cgggatgccataacaattgagtaaaaag gagtaaacactatgctgttcacatcgcgcg cgtgcacgggatgccatcagcaacattgac accaggcgagagctgccaagcttcatctc gcgcgcgtgcacgggatgccataaataat tatcatgaagatctgagtaaatccaagttca tctcgcgcggtgcacgggatgccatagcg tgcggggatgaagacagatagacagat accagtaacacccctgtgtgagccgagctg caggccctgtgcaaatccggtatctaatacca cggcaacctactcgcctggcggaaccccta taggagatt
340-341	Ga0209647_1007495 GENOME_ACESSION: GraSoiAngeSPAdes_43_ \$F_3300026319 GENOME_ID: 26992 CONTIG_ID: 7494 SOURCE: JGI DATE: 2019-01-27.	atgtctgtctatgtaaaaaatcatgatggagcagcactgatgccctgcacagaggcaaaa gcacgcaagttgttagaggcagggaagcaagattgtggactatcgtcctttcacgac caactcagctggcagtgtagggacaggttcagggaagtcacgtgtgggtgataaagg cagcagatcacaggccttgcctgcgtgggcaatgtgtgttcttgccttgcggcagat ccagcatcgtagagatgtgaagacaaaatggagatcgacgggacgcagcaaaagc cgagagcatgtcgtgtatcggcgtgcacgctttcctaactgtggaagtcacctacgg agcggagcgttgccttgcgtatcagaaacaaatgtggaggaagtcattcgtgtgtaag cttcttcttccgatcagttcttctgtatcaggagcgtccaggtggatattgtctgttga acaacccgaaactcaagggaagtcagtaaccaagatccgacgcgattggatgagaattg cgcatggcctgtttgatgcgtgatgttatcagtgtagcactgcggcagagtgccagaa aactggaagcccatcatctatcttccgtgtgtgagggcgaagacacttgaacaaatct gttgacgctttgtgagcgggtccatccaactgcacaaagggaacaaacacgctcaaggt aacggagtaagtgtgtatcgtgaccagattgtcacaaggactatgcaggggaaacac atctatatacaactttgaggaggttgccttggaccaccgtgttggctatgaaacatccac gtaccgtgaagtaccggagtttgcgaaacgcatattattgaccattgtgtatcgaaccc tgggactggcgaagtag	gatcatagttgctaaaaagtatgtgttttcgga catcctattcaaggaaactaccagcgggcctc ttcgaggagcagcagttgtcaggtcatcaca ccctgggatgttcttcaggtccaggtctc gtgttccggtcgaagggtgagcgaaggc gaaagtgtcggagcatgacaagcgtgatg aacaatcgcgaggagcaattcacttgcaca gaggaggccttatagctagatgtct

119

SEQ ID NO:	Name	IscB sequence	ωRNAs
	60_cov_3.164799, whole genome shotgun sequence GENOME_ACESSION: UYDE01.1 GENOME_ID: 279396 CONTIG_ID: 1019 SOURCE: WGS DATE: 2019-06-06.	gtgggctacaactacatcgcatcagcgtgaaagcgaatcccacgaatatgtgtctgcg cagtatgatacattacagatgagaagccgtccacgacagtgtgtcgaatgtatgccgtga cccgcagaacagactgcgttaccgttaaaccgcgttcgataaccgcaagcgacgcga agggtggctgtctctcttggaggcataaagaactcaatgtcaacgtgttcaagatgt attgtgcagtaatcccattacgcatgcaacgggtgaagttggtctttgacacaaatgttgc gcaagccatccagaaggaatcaaaacgggaaggtgtgactaccagaaggttccc cgctacaacttgcaacctgtcgtgaggcagtggttaccgtgataaattacacctgccaag tttgtggacgcaaaatcgcgatgtgtccattttacatgaccacatgtttattggaag gaagacacgggtaccagcttgatgagttggttacagcgtgtgaaaaatgccacacgcca gcaaatcatcaaaaagggtggcaagctctacggatttggcgaagataaagaatttgccaat ctttcaggtgcagcatttatgaacgctgttcggtgcgagatagtaaatgcatgtacgccac atacggtaaaagaattgtaaccatcattatggcgctatgacaaaagaaaagcgtatcgcg cttcacttgaagaagtcataaacgatgcgtatgcaatggcagctttatccagttaa ccgctgcgcgtttgaacattatgaaaagatgaacgcaataaccgattctcgaaaagttt atgactcgcagtcattgacactgcactggtgaactgactaacggcaaaagcttattcaa cggtagaatcagccgcagccataaaaaggattccgagaacctgcacaagtaccgtgga aagaggattgtaaaaggcaccgcgtctacgccaaaaaagttggccctcaatcccgg agattttagtttctcaacggagaattctgtgtccatggcactataccaataaaaagg tgctgtaaatgtagaattcaaaagcaccggcaaaaaacggcaaaaatccgcgagccttaa aaaactcaaaattgtaaaagcagcagattccatcatcccgatgggaaaaagtatctaa	cgccggcaggaaactacgttgttactaata atataggcaccttactcatgtccacaagtgg tgagctctgcggatgtttgtaaaaatctctga gggtaggagacgtgcaaacatcataccgaa aggtaaaacagtacaacaacattggcgaag tggaccacaggcgcaagccctgacttata gttttattactgttttacgaaaggagtgccgtg catgagcacttgc
350-351	Ga0190367_1011244 GENOME_ACESSION: 4872181011_MG_2_SF_3300022470 GENOME_ID: 20998 CONTIG_ID: 11243 SOURCE: JGI DATE: 2019-01-27.	atgtcacccgggctgagaccggcggggtaatcccccaaggtgattgatatgggaaagg ggaggagtgaggaaaggcactgtgaagagagtagccggtagtgtatgcagagggtttcc attgatccctacgacgccagtaagagcagcgccgaatgttgaaagaaggggagcgagta cgcgagaaggaaacagttggggcattttctatattcagttgaagcgcgtgtcgtatccggtgc cgaagaaggtgcgaagaagcagccgagccgatagcagcgtccgtggtatctgggtttctcc tttgcgtgttctcagtgtagggactaaggacacgattctcaacataatgacggagccag tgaattgggtggaagatgctctaaggaaagcggcgagagatgagaagattgagaagatac cgcaagtgccagcgtcggaaagaagcgtttcgacaatcgtatagagaccagaggatgg gttccaccatcaacgaaggcgagatgggatactacctaaggataatagaccatcttagg aaaatagtagcatcagcatgtcggcgtagaggagggcaaaagcgaggacaaaagaag ggacagaaaagatggacaataactttccctttacagaacggcagaaaactacttca aagagctgcaaaagagaggcttgatgtgactttgcttccagcgcgagagggttgcgagg ctgagaaaagagcatggtttgacaagggtaaaagataaaagccgagaagagttttactca cactgcgtgagatcggtgggttatatcagcttgataaacagggtgctgcgacgccgaactgt cttgatactgtgtacgtgttctttaaagggtttcacaggaggcagttgcacgggttaaaagcc gtcaaaagggggagtgagaagaggatgtgtgtacaagatctctcggttcaagagg gggacgctggtgaaatctgggtacgggctgtgtacatcggcggttcgacgagaaaa ggaaaaggctgagttgtgcacgacgitaagcggggagagaacaacaaagctgccaa accgcagacatcagagttcttaactctgtcagcttcaggagtttttacattcccgcaattcc cccgcaagcttgcggcaccccttgcgggggtga	gtcaacgacctgggctaaagcttggcttgg gttggcgggcccaaggccgaaggagctt cgttgactagcctcagtcgcccgccgaccca ggcggggctacgccccgggctacaag gacggtgggtgctccccagcctcacc gctccggcagaaccgtgaatgcgtgtgtg ggcgagcgtcacctaagcctcgggggca ttggcgagggggaatgtcaccgggctga gaccggcggggtaatcccaaggtgattg atatgggaaaggaggaggtgggaaggg cactgtgaagaga
352-353	Azohydromonas sp. 13393 Scaffold36, whole genome shotgun sequence GENOME_ACESSION: GCA_009760915.1_AS M976091v1_genomic GENOME_ID: 317215 CONTIG_ID: 145 SOURCE: NCBI Prokaryotes DATE: 2020-04-22.	atgtattccgagaactctgtgttcgtctggatcgggcaaggcgccctttcgcccttggc gcccagcgcaggtctgcgcgactcctgcgcgacgggaagccgcagtgctgcgccgt accggttaccatcatcctaccgaagaaagccgcaggtgacccgaagccgctggc gttcaagatcgaccgggctcaaggcgacccggcttggcgctgtgacaaaggcggg ccgctgtgtgttcgcagccaggttgagcatcgcggcgagtcgatcaagaaggggctg gacgaccggcgcatgtaccggcgcaaccggcgagccgcaagaccgctaccgcgc cccgcgtttgacaaccggcgccgcgcgaacggttggctgcccatccctgcaaac cgcgtggacacgaccatgacctgggtgcggcgatttggagcagttccaacgtcgcgc agctgtgctgtgagcgctcaagttcgacacgagggcgaatgcagaaccctgaagtcagc ggcgccgaataaccagagggtgagttggccggcaacgaagtgcgggagtagcttgc agaagtggcagcgccgtgtgcctactgcgatgccagccgcgtgcccgtccagatcga gcatgtctgtggcgctcgcgcggcgccaccaaccgggtgagcaacctaccctgtctt gcgggcttgaaccgggccaaggcgccgatccggtggagcagttcctgtgcagga agccggacgtgtggcacacatcaaggcgaagctcaagcagccgtcaaggacacca ccgcccgtgaacgcgacacgttggcggtgttgcagcgcgttggcgaacgagctggc cgtcgaggcaggtatggcgcgccgcgaagttcaaccgcacgcgacagggtaccc caaggcacactggatcgacgcggcatcggtgggtgagagcgcgccctggtcgcgt cgatcccgccctgcgcccgttgcgcatcaaggcatcgggccatggcctgcgccagcg tgccgccccgacaaataggggttcttagacggctgcggcaggggccaagttcttctg ggattccagacggcgatctgtgtaacgccgcgtgcctacaggcaagtacgccggcc ggcacacaggacgcatgccaatccgggtccgcccctatttgcctgacctcaaggaca ccacgttcgacgtgcatccaagtaacctaccggcggaagggtgacggctatgcc tacttctga	cccgattcgatgaattggatggagcttgcga aagcaggtctccccggttaccaggtcagt gattcaggagcgtatctgatgaactagttca gacgtgatgacaccctaggatgctgcca gtccaggctctgtcggcaggaatcatgtg gcccagcagggaaggtggcgaaggttc ccgctcaacaagccgttggacattgcg aggcaaacattaccggccgaaggccgtg aaggagtttgaacctgtattccgagaact ct

SEQ ID NO:	Name	IscB sequence	ωRNAs
354-355	Draft_c051504 GENOME_ACESSION: IMG_3300000053_SF_3 300000053 GENOME_ID: 27425 CONTIG_ID: 4381 SOURCE: JGI DATE: 2019-01-27.	gtgagtgtcttggtagtgacaacagaaacacccgttaagccatgccgagtcacgg gcccgtcagctcttccgcatggcaagcccgaggtttctgttattcgttcacgattat cttaaaaganagagtgggcggtgatacncagccgctgtctctgaatgttgacctggcag naaaacaacccggtgtgntctgtgtagcagatcagcgccgagtgccgagtgattgg ctattcacattggccaccgtggtcagcagatnaaatccgattggangggcaggcgcggt atncgncggtcacgncgaaacccgcaaaacccgctaccggccacgggttttaaatc gtacacggcctaaagggtgttagcgccatcattatgtacgggtgcacaaatgttgatac gtgggcaaaagcggttaalcggcttgacacatcattagcgcgaatgtgaaaccgttg cttcgatatgcagttgatgaaacccgtcgatggccggcgactatcagcaagggt cattgtttgtggagcttcgggaatctgtgtatgcacaaagcatacctgtgcctatt gcgatggattgacggcgatcagtgctggagaaagacacattattcgaaggcattag gcccgcagcaaccgctggcaaatcatgttatcagttgccgacctgcaacgaggataaa gggtctctgcaccccaatgcctggcgagctttgcatgcagcgcggtggaaatataa acgacacggcgcaaaaacatgcagcgattctggccggtaccggccatcttaaggga tgacggcgagtgattgctacgcggtatccgtgggtggcggtgttaagcgcttaatacca gatagcaggtttgtgtgtggccgcacaaagaaaacccgctcagaccaggggtatcac aaagaccactggattgatgcgacctgtgtcggtgagaaagtggtgcggtgtcactgta tgcgagctgtgtctatggcaatgccaaaggacatgttgcacgagatgtgtgtgtgg acaagtacgggttccgcgaccagcgcaaaaacacccgtgtacacggatttaa accggcgacatgggtggcgccagtggtccaacggcgcaagaacaggtgtgtatgtgg gacgagtagcggtacgtacgagcggtttttcaacattcaaaacaaatggcggtgtgc agggtgtatcccaagcattgtcgatactccagcgcaacgatgggtataactttaccta tggagcgcaatctccgacgactgattcaaaagattcagtcggtatcctgtccgaat agttgtgtga	gtcaacgaccacgactatttcgctgtgcga attaagtcggggctgttaaagccctaagttg accagaccccatcaggcaacggatggaga cgttatcgaagaatacatagaccgtggg atgacttgcagtcctccgctctgcgcaaaa cgattaacacagggtatgtgggttaagccagt gtcgtttgaaaaaaccttcgaataacacgggt cgaggcacacataaccacgttcggtggaga acgacgggttaaccgtcataatgttgagatt ggataacgtgagt
356-357	Moorea sp. SIO318 318_NODE_1, whole genome shotgun sequence GENOME_ACESSION: GCA_010692325.1_AS M1069232v1_genomic GENOME_ID: 325532 CONTIG_ID: 0 SOURCE: NCBI Prokaryotes DATE: 2020-04-22.	atgcgagttttcgttttagatcaaaataaaaaacccctggatccatgccacccggcacgag ctagggaactactaagaaaagagggtctaaagtatttaagcgtatccattcactattatc atgcaggacagaagtgtaataattctgtaactcattccacacagaatcaaaatagatcctg gttcaaaaacaacccgaatagctgttattcaagaacagactggacgagtgacaagtgcacat tagaaattaaccatcgaggacagagatcaagaattctttagaatccagaagagcccttaag gcgagtgtagggcaacagaaaaactcgtaccgcaagccctcgaattcttaacggaagc gccagaaagggtgttaccctcgtcactaatgagtcgaatatttaacgttgagacttggtg aagacgactgaggaaagtgtgtccagtgactgcaataatcgaagctagtgcggtttgat actcagaagatgcataatccagaagtatcaggtgttgaataatcaacatggcgagctatac gggtttgaggttaagagtacctcttgcgaatggggacacagttgtgtttattgtgtgtgt atgaatagcccttagaagttgaacacattgtcccaaatcaaaaggtggcgcaatcgg gtcagcaacctaaccctgtgtgtcgggtgtgtatcagaaaaaggttaattgacccattga aaaatttttaagaaaaaacagcaattctcaaaaaagtatttagctaaagcaaaaattcttt aaaagatgccgcagctgtaacatcaacccgttgggaattatgagaagactacaatcaac tggattaccgtcgaacacagctcgggtggtgattgactaagtttaacgcaagactagggg tctcgaaaaactcactggatcgtctcttgtgtgttaaaaccccttgagcaaatat tactgaacggaacaaagccactaacgtaacagctaaaggctatgtattcgtcaaaagg gcccacataataatcgggtttccaaaggctcatgctcctctgtcgaatgttttaaaaggc tttcagacaggagatatactgcaaaagctgacataaagaaagaaatatacggggcagtat acaggacggatagctattcgttatcgtccaaagttttgactcagactcccgagaagaagt cgatgttcaccaaagtatctgagaataatttaaaagctgacggatacgaatatgcgtcta actag	gtcaataactaccgctaagtcctatggagct atagcgggagctgtgaagctcatgttga ccagactgagacttagattaggtctacgttat ttgattataacacctgttagtgcgtgccag ccctcagctctgtgttaacaaatcaacatcgt agcgagttaaagaggtgtgttaacctctcaa gctcttataacattgtcgaggcaaacctttacc cacatcgtggaggtgtcaagtaagtcga
358-359	Methanosarcina sp. 2.H.T.1A.3 scaffold7_1_size162551- refined, whole genome shotgun sequence GENOME_ACESSION: GCA_000979425.1_gtIE nvA5udCFS_genomic GENOME_ID: 234720 CONTIG_ID: 76 SOURCE: NCBI_Archaea DATE: 2019-01-27.	atgttagttttcgtatcaatcaaaacaaaaaacccctgatgccctgcaaacctcctaaaagc cagaaagctactcgaagcagcgaagcaaaagtggtccgaaatagccattcacaatca agttacttttcagaagcagtggtatatactcaacctgtactgcaggatggatccggctct aaggtagtgggtgtgacgccattgctaacggaaaagtgtgtatcagtcggaatctacc tgagagaaaacgtttcgaaaaagatggaacaacggaagatgtaccggagaaacccggag aagtcggaagacaaggtatagaccgcaagatttgataacgggggaattcaagaagaa aagggaagattggctctctatccgaagcaaaccttgaaagctatftccgggaaaagaggtt tgtggaatccctgctcctgtaaccgagtggaaggtagaactgtctcctttgataatcaca aataacaaatccggaagtttccggatcggaatcaggaagggaccttaaaagggttcta caatatcaaaagcttacgttctggacaggggacggctacacctgccagcactgcaggggaa agtcaaaaggattccaggtcattgccatcatactgttttcagggtcacaaggggaacaga tgcaccggaaaacctgatacgtctgtgaaacctgtcacaagccctgcac	gtcaactacccctgagctaaagactcaggg gcttctcctaagccctggttgaccagatc accgattaggagcaacggaaaatcggtaaa cgataggaaaagatacatagttacccttgaat gtcgcctcagtttaaggctctaaggatgccg gttaaacagctcctgagaggtagggacagtg cttgcacgttaaacctttccatcagatcga gaggaggacggattcgaattgactccac aattcggatagcataactcttcggaggaaa actatatgtta
360-361	KEY_1008596 GENOME_ACESSION: IMG_3300000938_SF_3 300000938 GENOME_ID: 308319 CONTIG_ID: 2543	ttggaatctcgccgttcattgctgtagaagtaggagaaatcgtacatcgttatcgtcaagg gaggttctcaatagactcgcctaaagggtgtgtgtccagctgcgaacatcgcgt cgaaacaacccataacctgggtactagattgatgaagctcgcctcattctattactc aagaactgtacgggtcactcaacaaatagaagaatccagagattctggaatcgaatat cagcaaggtgtgtatgtgctacgaggttcgggaatctgtctcaacaagtggaataga aagtgtgcttactgtgtgttacaataacacattacaagtcgaacatattatcccaagct	gtcaacaacccggcaccgagtcggattaccg ataccgttgcgggcaagaattcgcaagtt gaccagctcaaatgttcgccactactgtta aggtaagcgttaagactaccagaaaaatg cgtgtagttttctcctagaaatfaaagtt aaacaggtgaagggttaagccagtcgtttt

122

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gacacagtgcaaggtattggctggaataattgtaagaaatacattgcatagatggctata actacaacaatagaatggaggcggaattccccctgagctaa	
370-371	Ga0211577_10055383 GENOME_ACESSION: IMG_3300020469_SF_3 300020469 GENOME_ID: 278952 CONTIG_ID: 55382 SOURCE: JGI DATE: 2019-06-01.	atgtcccaagttttgtgtcgcacaaagacgtcgtccactagccccctgcacccgagggc gagctcgtgtgtcctcctcgagtgcaagcttccgtatttcgacgatacccttcaccatc atcctcaaggagtgccacgccacagccactccacgacctcagggctcaaatctaccct gcgagtaaaacgacagggtggctgtcataaatgagctacagctgaagtggtcgtggct gccgaacttaagcatcgtagccaaactgatcaagaagacgtggagagtcgtgtctttac gtagtggagcagcagctgcaaaactcgtaccgaccagctcgtgctaaaccgctga cgtataatcctgtgtttaccaatcacaggggggtgtcatcacaggaaagtggtgctc cctctctcaacatcggttggaagtagtaatgacctgggtagaacgctccagcgctatctt ccgatcacagcgctatctcaggagattatgcgtttgatacagcagaagctacagaatccg gagatcagcggtgtgtgagatcagcagggtactctgcacgggtatgaagtgctgagatc ttctagaaaagtggtcagaaaatgtgctactgtggagctagagatactgtctggagat aaatcatatcgttctcggagtcgtggggtagtgatcggttcagtaatttgacctgactt gtagatcctgccgagagcagagaggagcttccaafttgaagaatttctggcaacaagg cccgcgtgtgtgatgaaactccaaagtcaggtcaggtctccttgagagatgtagcagcg atcaactctacacgtttgtgtttgtagcgttggaagctagaggctgcccgtgaagtt tcagcgaggagagacaaaagtttaactgtaatcagcaacaataccaggtcccattgg cttaatcggtgtgtattggggccaataccccagaaaatttaaaatgggatcaggtacagc cgttagcaatcaagcgaatggccatggcaagcgccagatggtcaatgttgatgccttg gattcccagggtgtaaacggaaggggactccggttcacccatttctacaggtgatgtcg tacgggcagcaataccaagggtaagtacgttggaataacgaggaacggatctcatca atcaagacttcagaaacaagggtggggttccaaataaaaaaggccaaggaaccatcta tctacagactaaatatactacataaaaaattttagttcagatggctttgactatgaatttcta ccagtgaaatcataa	atgttcattaaaaagagtgaactctgggctga ctctcgcaaatgaagtgcgttttccgcagc tacaaccagactaagtcttcgagaaactagct tattttctgtgatacctacgggtgcgtgcc agcccgtagctctatggccattgcgttaaac agagacttttgagggtgctcagtcgctgtg gcgtgacaagcgatgataacattgtcgagg caaaccttacgtgagggtcagcttccctcgga agcaaacattcgtacgtcagcttcaaggcta accaaattgtccca
372-373	Burkholderia turbans type strain LMG 29316 genome assembly, contig: contig000014, whole genome shotgun sequence GENOME_ACESSION: GCA_001544655.2_AS M154465v2_genomic GENOME_ID: 98378 CONTIG_ID: 67 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	atgccgtgcagtgaaaagcgcgcgagattgctgctcgcgcgggaccgcgcacgcgtg catcaggtgtgacgttctgttctggctcagcgcacgtgaacccaattgcagcaaatcca accgttgcgcctcaagttcgaccgccgtgtagcaagaccaccggctagcgcacgtgcgc gacatcgtaacgccgagcaaggagtcgctggaactatccgcggcgagccgtacttaa tctattcgaactggcccaacgcggccggcaaatcagtgaggcgcttacggcgctgcgg gcaatgcgttgcggcgacgcgtgaagctgcgcctaccgtgcaccacgcttctcgaacc cacgcgctctatgggttgcgttgcctcgtcctcagcacagggttgatactacaaggc ctgggttcacgtatcatgcgtggccacccatcagcgcattctcagcgagctcgtgcg catcgcacgcacgtcgtcgaaaacccataagtcactgggttgggtatcagcgggga acgctggcaggtacgaaatcgcgaatactctttggagaactggggccggcgttgc atactgcgacgcaagtggtgcgcgctgcaagtatgatacatcgtcgtcggggccgcg gtggcaccgatcagtgctgaacctgacgtcggcctgcgaaccttgcaaccgggcaaa ggggcgctgtcgtgtaagcatttctagcgaagcgccacgcgattggcgaagatcc tgtccaaagcgaagcgtccactgaaagatgctgcagcagtgaaacgacacgcggggc gctgaccaacgcactggacacgactcggcctactcttggaaacttgcctcagcggaagta ccaaattcaatcgttggcagtcgggtgtaccgaaaacacatgcgtcgcagccgctgtg tcggggcagtatctgtcgtacaaaactggattagaccgatcctgtctatcagctcaccgg tcggcgagttaccagagagcagggctcgaaccgtgtggttccgcgcggcgat ctgaagggcgaagcgtgtgcacggatccagaccggcgacctcgtgcgcggatg taccctttgaaagaaggctggcacatacgtcggcagagtgccgatgctgtaccgggt acttcaacattcaagagctgggaaggtgttacaggggcgtcgcgcatcggttccgat tggttcaacgtggcgacggctacgcatactcccggataagccccggccacgcagggc ccgacagcgcgtgaggggccgtcgcctatccctgtga	gtcaatcaccgggcaagcccgaggtttt tgaagcactcgaagctcagattgaccag acctagcactgtaagcgctacgttgcag agacagcaaacccaccgctgaatgcttcc agttgcacgctctggaagtcgacgacgag acaagccccgggtaggtacgaacgggtc gcgacgggaagcgaaggcttccctgtgc acaacatgtgcgaggggagccataccga aggatgctcactagccccgtaagccaa ccgcgccgaacggactcggcggtgatc aataggggctgcttggct
374-375	human gut metagenome genome assembly, contig: NODE_575_length_2940 2_cov_7.466760, whole genome shotgun sequence GENOME_ACESSION: OLPO01.1 GENOME_ID: 7878 CONTIG_ID: 574 SOURCE: WGS DATE: 2019-01-27.	ttgcaaccccgtgcatcaagttctgtgaccttggcgtatgggcaacacacaccgattgct gtgaatcaggagtgaggacgcagagatcgaggacagtggtgttctctgtcatatctg aagaaaggaggcgtgaagcctatgaacagatctgacaagtcagttctgttattggaatgaa tggtatcgggcttatgctcctaccacaccagctaaagcaagacttcaatcagacaaaaaa agctgaagtgtgtacgaaagatgccattcacaatcagactcgggtacaaaaccggcagtg ccacgcaacatgtagaacttggcatagatacgggttctcagcataattgagtcgctgttgc ggcagacggcaagttatgcaaaaatctgagcatgtctccgttcaacaatggataaaag agtgtgaacgaaaaacgcaaggtcagagaagatccagacgatacgcgaaaacagaa taccgtcatccgaagtttagccagcatagcaaatatgtgtacgtagagaaaacagtcactc gggaagacgacaagcgcactgggaagaaatcagcaacaggtacggtaacaaccgta aagaaggtggttccgcatctatcagcagaaatcagcatcatatcatgatcatcag gagatacaggaagctcgtccactgtatcgtcaacaataatgcaaafttgaagttgata tcagaaaatcaataagccagatcgaaggtcattgatacagctgggcagaaatgtatc agttgaaaatgtaaaagcatatcgtatggaacacagctaccgctgtcccgtctgtgg tacggaattcgggacaaaaaagaaatctgtggaatcatgtcttaccagaacttcacat aagcatttcagaagcagaggagccacagacaatccggatgagtatatgctgtctgtaca gtagatcatcgtcgtcagagcatgtgacagtggaatacttggaaaagctgcgaaaagc	acaagtaaccgagcatgtcagaccgacgt gcctgtgtagcttatgtgggtgtcctgcata ggatattccgtcagactagcctgagatcgtt atacagattgcgactaagtttgcagaatc aatactcgaaccagtgatgccaaagttgat agagttggggcaatgtatttcagagagaaca gaatgattgcaaccccgatcgaatgttctgt agaccttggcgatgggcaacacaccgat ttgctgtgaatcaggagtgaggacgcagag atcgcgacaggacagtggttctgtcatat ctgaagaaaggagcgtaagcctatgaaca gatctgacaagta

SEQ ID NO:	Name	IscB sequence	ωRNAs
		agaagaaaagacaatccgtggacagagagatgatcattcatgaatcctcagaaaac gaatgtggaggcttttccaacagctgtgtttacatgtgcaatgtgacgaatgcggacag aaaactatcggcttatcaaaagactcatgcaaatgatgcgggtgcaattgccatgcacag acagatactgtctggcgtgcgaagcgttcagatgctgcagatactacatacacaagca ggtacgcaaaaaaagaacgttcactgcatgaagcaactgcaagaaaaggtagaccaga ccaatactgctgcaaaacgcaatgcaaaaaacacaagctttgactgtaaaaggaaag aagtactcctgcgtgcaaaagtagaatacaacggacagattgcgtggaattgctggctttt caggaaaaacaggtgtgataatcagcttattgatgacaaatattatcccaaccaggaaa atcatcacatcaatcaatctttcagacgttaattgtactcaatcacacaataactggatcgt aggatccatcacagaataatcaatataa	
376-377	Ga0061017_10387965 GENOME_ACESSION: CoaofCornStover_3_SF_3300031119 GENOME_ID: 280051 CONTIG_ID: 7826 SOURCE: JGI DATE: 2019-06-13.	atgcctacggaiaaacatgcaaaagtcctgtattctgtaaaagaacaaaggcaaatgta attaaaagggtgccgtttacatacagttagcgtatgatagcacgaattatagcgaagacgt tactttagggtgtagatgcggcagtaaacatctgcctttcagctaccacaaaagacaaa gtattatttgaatctgatgtggaaactcggaaacgatatggtgattgctttaccctgtagac aaaaccgtagaaccagaagaacgcgaagaccgttaccgtaagccacgtttcgataat cgtaaacgcaaaaacggatggttagcaccctcgggtcagaataaggtagattcgcacctta cggtaatccgtaaaatacacgaaatctgcctatcggtaaggatttttagaagtcgcttctt tcgacatacagaagattaaaaatcctggaataagcggtagcaggatcatcagcgggtgac agttaggcttttgaattgacgggaatacgtactttcagagcggacacgcttgcagtg ctgtaaaaggcaggtcaaaagacaaaatccttaacgtacaccacattgaaagcagacatac cggaggagatgctcaaaacaccttattacattgtgcaagacctgtacacagataccat aaaggtagcagtaaaactgccaaaaactataaggcggaggatgccttttaagatgctgcc ttatgggaattttacgatggctatgtatacaaaattaaagcagatatacttaagttaagtct gacttatggatatactactaaatccactcgtatagagaatggctcctccaaaagaccattaca tagatgccctgtgtataagcgggt	actaatcccacggggcaagcccgtgggggt gcttgacaaccagagtgtagcctaagt gaagccttttggctgaactcgttaagagag aatataatagtcacctacgagcgtaatgcctaa ctgttagctctgagtcagtgattaaacaatc ctgtgtataaggatagttgtctgatataaa acctctcattaaacttgcgaaggcatacaa ccgctctcggagcagtaattaaatccttat gggatttggaaaggagacagacttatt
378-379	Ga0310140_0030813 GENOME_ACESSION: K1A_FD_SF_330003570 3 GENOME_ID: 312523 CONTIG_ID: 30812 SOURCE: JGI DATE: 2020-04-07.	atgggtaccagttatagatataataaagttccaatgatgcctgtcagaaaaaagggcaa gaaagtattggataaaaaagaagcaaaacatattggcaacatgggatcctttgtataatg cttttgaaagaaccaagcaatagaaattatcaaaaggtagtagcaggtatagatccggga agcaaacgagaagggtatagtagtaacggaaaagaaagtaattctaaatgtactaca gataccacattatgggttaagataaggtagagacaagagagaacattaaagaagaccgg aaggcaaaagaaaaactcctatagagcgtgtgtagaataatcgtcatctttaagaagatt aataggctccaccatcaacgaagcccgtgggatgccaaagcttagaattcttaaatat tatctaagcttattcagaatcacagatacaatgtggaagatataccaaagctatgaccagaaa ggaaaaagcaaatggaacattcttttccctcttcaactggaaaaaattcctttataata aaataaaagaatttccaaaagtttaatttagtttaacagaagggtataatcgaattaca tcgggataaaagaagccttttcaaatcaaaagccaaattagattatttgggatcggcata atgttgatagtcattcattatgtgagatggccttgaataaagagataaaaccatataaaggga atgtataaaatagatttatgcaatttcatagaagcgattacatagaagcagccattaaaa ggaaatgcccgcaagcaatagtcgggaacgtttcattagggtatcgcgaggtacatt taagataaaagagagatacaaaatataatttttgggtggtactggagttactcgtaaaaata aaatagccatacatagtggtgaacaggaagagaataaagcaacatacaaatattatcag atatagagattatgtataataatacaattagaacgcaatttctgccattttaagaagattggg catcttttgcacttttagataa	gtcaataacccttaataaaattaaaggccttg tgcgaagaaaaatagcaatttctttgatcaa gactaaaattattgactagacaaggcagtat gcaaatagctggatcagctttcgttatgacaa gaaggcaaaagtcattggtgatttgcgta gcttttaacctctatgattcaggtatgtcgaag cgagttggacaagggtcgtgaagcatcggtt ttgtagtaaaagaaaaataagtaaatggtac
380-381	human gut metagenome genome assembly, contig: NODE_1950_length_253 70_cov_3.785187, whole genome shotgun sequence GENOME_ACESSION: ORVF01.1 GENOME_ID: 10838 CONTIG_ID: 1949 SOURCE: WGS DATE: 2019-01-27.	atggggagggaacagtgagctagctatataacatatttaattaaaaccttatataacattggt gatggtaaccaccacaataatgtggagaacattagtagctgtactatttattagtagcattgta gtatacagaaagagacgtaaatcattttattagaaaaggaggaacaaatattggtgctgta ttaagtagtattgcaactaaatgatgcaactcttaattacagagcacggaattactttaa aaggacgtgctaaaattataagtagaccatttacttactgatttggatagggaagag ggtgatactcaaccaattgagtaaaattgtgatactggatatacaatattggtattagttct gttcacaaaagcagaatattggaatttacaagtagacattgtaaaagacgagactgaaa ggcacaacgatcaacgtaaatatcgtagaactagaagaaatcgtctcaggcatagagca cctagatttaaaaatagatttcatctaaaagaaggttggtagctccatctgttaggcatt aaaaaagaatccatattcaatgggtcaaaaaatattatgaggttaattgcctataacagatgc tacttttgaagttggtgaatttgatactcaacttttaaaagcattacaactggatcctctatac caactggttaattgtatcagcaggacacgttagtataaaataagtagattgagaaatgcaat atttacaagagataattacattgctgtatattggttaattggaatcctcaaaaatcattata tttgtacatcatattggtatttgggtggagatagaacagacagactagataatttacttaca gttgtactaaatgccatgtgccgcaaatcatcaaccaggtggcatattatattggtcttaaa ccaaaattaaaaaatttcaaggtgccacttttatgacaattatcagatggcaattgttagat gagttaaaaactaattcctgatataagttccatgcaacttatggtagtgaacaaaagaa agaagacgcataattaaagggttaagaataatcagtagtaattgatcttatgttaattgggtatttt catccaaaacatagaactgatttttactactaaccaaaaagagatgtaataatcgcataatta gaaaagtctatgatgctaaatatatagatagtagagtgatctaaaaaatcaggcaagg atttattcttctgtagaacaagaagggtatatttctcaaaccttagattttatagacaac aaaaagttctaaaggagacgttctcatcagaagaacacattatttattacaaccacatgat	ataagagtttaattattcttttgaataattttcta gaaatcagttggggcctgttaattcttttgcaa gccttattgaaatagcctaagtatttccggatact acgttatataagaataataggtaccaggt gatgcttaccagtccttgcctctacggatta gcttttaacagttctatgggaggggaacagt gagctagtcatattaacatatttaaaacc ttatataacattggtgatggttaaccaccaca aatgtggagaacttagtagctgtactatttta tttagtactatttagta

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gttgtaatatggaacaatcaaaatatacagtaaaaggggtacataataatggaacacgtgtatattaaagaaaaataataaactgttaagatatctgataftaaattattagacattgtaatgatactattaa	
382-383	Ga0120377_1000528 GENOME_ACESSION: IMG_3300013830_SF_3300013830 GENOME_ID: 278117 CONTIG_ID: 527 SOURCE: JGI DATE: 2019-06-01.	gtggcgaaagcaccaccctcgcaagaggagaaatccggcggttctatgaccggggaa gtaattcacgaaaggagacacaaagtgatggcatagcttgcagtcctgtcagcatctgga aaatcgctgatgctcacaactgcatacaaaagcagcggaaactctaaagagtagacgtgca aagatctacagctacagaccgctgttccacattcagctgcaggaccgtgaagaagggtgct acacagcctgttggctgaagatggataccgggtgcacagcatatcgggatcagtgctgc tctgaaaaacacgaatattggaacaggcgctacgatatgctgcctgcacaaaaagagat gcatgatgaggcaagaaaaaccggcgttaacagaagaaacagccttcgtatcgcgctc cacgcttggataaccgactcacgggtcataatcgcaaggagataagtggttcccccttct ctgaaacacaaagaaacatccacatcagcttgcacaaaaaactgtgcagttgtcccg gtaacagacgctcacttcgagatggggcagttcgacatccagacttgaagccctatgaa gcaggaagcccatccctgtaggaaacgactaccagaaggtgagcggatggttatgac cactcttcgcgaggcagcttcttcgctgacaaattacacctgtcaggtatgcggcgccaa actggatgacaagcaccatccgattttgagaatgcacatatacggctattggaagcacgat cattccaatcgcatgagcaacctgatgaccgtttgcacaaatgtcataccgcagcgaatc ataagcccggaggcagctgtatgctgtggagccagagcttccaaactcaagggtgccc gcattcatgaatacgggtgcgggtgatatgtttgcacagctgaaaaaagcattgccgaatg tcaaatgccacatgacatatgggtccatgaccaagctgaacgctcagaactcaatgtca agaaaacgcacagcaatgatgctatgtatgggaaagctgcattccgaaatggaggacc gactttcagcattatcagaaagtacggcgcaataaccgagtgctgtccaaattctacgatg cacagtatatcgtacacgcgacggagagaaaaatgagtggtctcagcttctcgtcgggc gaaccaaacgcagtgagtcacagacagcataaaagatcttcgtggatagaggggcat cgggcaacatcaaaaaagggaataatccaaaggccatatttcaacaagaagacataa atatgtattcaacaccggagacattgtaatcgtcgaaggcaagaactgacgggttccgg gacacagcaattacgggtgaatatgtagtgttcgggacaaaacacatgcactcgcacaaagc aaaagaggtaacaatgtatcgacacgggtgatggctgggtacgtgtgtatataa	gtcaaccaccccgcccattactcggtaaat ggacggggcgttcgggggaccagaatacc ccgtaagcccggttgaagccaggttacatg ttcgaggagaaatatacaataggcacttcgg aacgctgctcaaatctgaacactgcggcg tgcattaacatcctatagggggaagggaat gtgtacgtctcggcgactgacagcaacat cgccagaatgccgtaaacctctgataaca gtggcggaagagcacc
384-385	Ga0209066_10011920 GENOME_ACESSION: AllZonSPAdes_SF_3300027851 GENOME_ID: 22535 CONTIG_ID: 11919 SOURCE: JGI DATE: 2019-01-27.	atggcagtttttcttctgaccgaagcggcaagccattgatccgtgcagcgaaagcgc gcaagaagctgcttgcggcggtgctgcgcgctgcatcggtgatccgttctgctgat ccggatcgtcgtatcgccctgcaggatagcgcgttgcacacgcttcgcgtcaagatcg accgggacgacgaagacgaccggactcgcgctcgtgcgcaagtcgattcgtatgctg tcgagtggtgaagtcgcccgcgaagtcttctgctcaacctcttcgagctgtgctcatcga ggctgcgcagatctcggagacgtcaccagcgcagtgcccttccgctgcgtcgcgtgag cgcgaatctgcgtatcgcgcaccgcgcttctgaatcgcaagaaggccagggcgtg ctcgcctctcgtcgcgccaccgcatcgagaccacgctcgcctgggtgcgtcgggttca gcactgggcgccggtcaccgcgtctcgcaggagctcgtcgccttcgacacgcaaaag atggagaacgccgagatctccggcgtcgggtatcagcaaggcagctcgcgggtacg agctgcgcgagtatctcgcgagaattcaaccgcactgttctacttcgacgcccgaag atgtccgctcaaacatcgagcaactcaccggaaggcgcaaggcgccacgaaccgcgt ttccaatctggcgtatcgcctgcattcccgtcaacacaaaaagggcgcgcgtcaatcg aggtcttctcgcgaaggatcctgcgcgtctgcacgcatcgggcgcagttgaagcgtc cgctcaaggacgcggccgctgtcaatgcgacccgctgggacactcttcgaggcgctcaa atcaacggggctcgaagtcgaagtcggcagc	gtcaattaccggccacctggcgtaaccgcc gaggtgggggcttgtgaacaacaagcctga ttgaccaggcccgatgcagtcgag ctacgttgcacgaagtacaaagacgaccg ccggatgcttccctcagtcggcgctctgcaa gcgcgacgacgacacgcttccgggttgc gcacgaacgggctgtcgaagatcgga tcatcgatcgaaagctgttgaacatggcc gaggggagcggcagaccggaagggaat gcccgtacaaggccgtaaggcgatgc taaggagtcgacatggca
386-387	Ga0373621_017898 GENOME_ACESSION: YL18G_03A_MG_2_SF_3300035099 GENOME_ID: 307034 CONTIG_ID: 17897 SOURCE: JGI DATE: 2019-11-15.	atgaacaatatctaccaggacaacccgctacaagcaccggcggtatgcttccagtcgc ccgctctgctgctgcgcccagccgatgaccataagaggtctaagcggagttctgctg gtcgaggagagaatacttcccgaagggaaggccgattatcggcaagcatggcgtgtg gggtggtgggggtgacgggaaaccgttgacccccaccctcctccgcccgtgccggaa gttgattgaggatgtcaggcagagaaggcgtggaaatggttggagtattcgtatcag gatgtttagccgtcgggaataactccccgacagcaccctgacggtggacaacggc acgaagtacgaaggctatagctgtgtggacggagaaacatcagcgggtgaacc ttaatctccctgacaagaagaagattgtcgaaaaactggaagccgtcggcagatgcgg cgggcaagacggcaacgcaacaccgacggagagcggcgagattgacaatcgaagt cgcaatggttctatccccgagccagatggtgatggtaattcccgctaaagatcatca atgaactgtctcgtgctatccgatttccgatgcgggattgaagatttcttttaacctg cccagcatcgttggggagtgaactctcaacgggtgagattgcaagaatttgattcggg agttctttaagtctcggggcattaaagataattgatttgcgggttacgagacgaagaactg cgatccaaatcggctataaaaagattaaggataagtcggcggtatcgggttgaagcgtat tgcctctacgcccgtggcgtgctgcgaagtcgggcctgggggaacgaattgagcttg accattggttatcgtgacgacacatccgcccgggttcgacggcagttatgacaccca accggcaaaaggcgagattcggcgagttattctcgggtacggtgttcgggttgagaa aggccctgctgattggccgccccgatggacagatggggcggttctgctggcgagatcg gggggataccgttatttgaattcaagggaacatcgccaatcgactaaaaagctgtatttta tttcggatcaattataaccagaaggaggtggcgcttctcccgcccttaa	gtttgatcgttactccttgaagctgcgggag ttcacgaatgccgaatgcgggagatgtatg ttgcgcttgaatcgttattatgattattcag caggaaactaatgggtatcttgatgacataaa ctatttcaactaagttcaatgcgatttgaagaa tcggtcgaagtttgaacctcagagaagatt gtataaggagaaaaagagtaactaccgc cccataaaggggacggcgtgtgttaaggac tatcacaacgtaagattgactaggccaaga gaggttcag

SEQ ID NO:	Name	IscB sequence	ωRNAs
388-389	Ga0209726_10008685 GENOME_ACESSION: Subgroc54mSPAdes_\$F_3300027815 GENOME_ID: 35774 CONTIG_ID: 8684 SOURCE: JGI DATE: 2019-01-27.	atgttgaccagggttaagcaggattaggcaggcaaccgcttgagtctgtctcgttgcg gatagtgccaaagaccgacggcagagtgcttccagccctgccctcgaactcacgc cagcagacacgcccgggtcaggcacgaaacggggcgcgagagacgaccggctccga acatcccaggagagcccttcgagacggatgccaatctcgacgctcagggcccaagg gtacgaccgaaggccgtctcgtcgcgaggggttaacccccgctgttctcgtcaccgg catggacgccactcatgtcgtccgattctcggggaaggaattgctcgataagggt caggccgtgtctacagatccgccctttacgattcgtctcgttgaccgtgtgggtgcgc gggtccaaaccggtccgcgccaagatcgatcccgctcgaagatgaccgcatcgtgttg ttcgcgatgttgagcacgttgatgcaacgacagggtgaagtcacccgcgaagccgtgtg ctgcacctgcttgagttagagcatcgcggcgacgattcgaaggcgcttcagcagcg tcgggattcgtcggcgtcggcggcggcgaatctccgatcgcgccccggttcg acaaccgaacaagaccgatggctgctcggcccgagcctacaacaccgggtcgtcac aacaatgacgtgggtcgtcgcctgagacgagcgtgtccatttccgctctcgtgtgag agggtacgcttcgacactcacgctctcctcaaccggaaatcgatggcgttcaaccag cgggggacgtgttcggcacggaaattcgtgagtatctcgcgaagtggagccacgc ctgcgcgtattgtgacaaggcaggcgtgcccctgaacacagatcactcgttgcacggc gcccgggtggctcggatgggtgagcaatctcgtgatgtcctgcatcactgcaacacac gaaagccgcatcgctgatcgaggagtttctcgcgcatgacctgagagactcgcgag cattctcgcgacggcaacggctccactgaggacgcccgcgagtcacgcaacgcg gaatgccctgttcttcgctcgtcgggacacggggcttcaatcgaagccggaactgggtg tcgaaccaagtggatcggctcggctcgacatcccgaagaccacgctctcgcgacgag cctgtgtcggcattgtgattcgggttccaactggcagatccgggtcctcgtatcaaggc cactggccgtggaaagccgaaagcgacgcggtcgtatggtacggcttcccgagg ctacctcatcgggggaagaccgttcgaggattccgaacaggcagatcgtccgggcc gtgttccatccggcacgaaagcgggacggtgacggggccggtagccgtcggggcc aacgggttcgtcaacgtacagacccggccggcagatcaaggcgtcctcacacaggc attccgctcctcatcgaggagacggctattcgtacacgacgaagtgcgctcctcc ccgctggaacggcagggttccgcccgggagaaattga	ttggcaatcaagacaagcacaatctttgaagt ccaagaccgcaaggttagccgtatcctc gaacgcaatcatatggcagcacattagcggt gggtatcccgaatgagatcgccctgat gatatctccaggcaatcaagcccccaattt catccaaaaggaccctctgtatcgtatagc cgacgtattattgttgaaccaacttcattt gggaaacactccaggtgtatgtactgtgtg gtgttagggatagacaagagtcggttaaac gggaaagtgcgcactgctcatctggttga aacgaagaagggaaca
390-391	Ga0376502_004060 GENOME_ACESSION: B26_SRodWell665_2_\$F_3300035540 GENOME_ID: 304280 CONTIG_ID: 4059 SOURCE: JGI DATE: 2019-11-15.	gtggatcacttgcgaatgcaggcgagagcttgagagctctgccaaataatcaaagga atctcaatgagcgttgtgtgtaaaaaatagcggggcgccacattgatgctgacgacccg caaaaggcaaaaattctattgaaacaaggaaaggcagtcattgccggtatcacccgttta ccatccagctgaattatgctaccgggtgaggccaggcaggtctgtgttcttggcgtatg ccgggttttaaaacaatgggagtttaacacggcccgacaaaagagttctattcttcgaa atcagccgtcttgagggacaggttggaacgaataaggaaacgaaggatgacggattca aagaagatcccgctcgtcgcacagaaagcagcgttggataatcgaggagaaagtac gggttgctgcccccaagcatccagcataagcttgatagccatatcaacatcattgagcgg ctgaagtcctgattcccgattgcaataaccatcgtggaagtgccggccttggatattcaaaa aatcaagatgcctgaaatcaacggcagcagatcaggaaagggtcctcaatcaggattctg gaatctaagggaatatatgctgacccgtgataaccatactgcaaaatccggactgcaa gaacaaatccaggcatcccggttggagtagacacataggggtattggaattggacag gtcggacagaccgggaaatctgatcacactatgcaacaagtgctacacaccggccgaac ataagaaaaatgttttctatgtgttgggaacccaaaaccaagtcgtcaagccggaac attcatgagtacagcttccgttggaaactgggaatacacttaagtgcgaccacacataggt tataagacaaaacaa	gtcaacaacccctggctgaagaccgagg ggcttgggggaagcccgcaagcccggtt gattagcctcagccctgcgttattggtgaagc ggggctacgttactgcagaatacataggtac ttcgggatacagatctgaccaatcgtccaca cgtcccgaatgttacggcagtggttttaaacat ctgtgagggtagcagaagtgatgctgatatt aaaaacctcgggataacattggcaatgtgga tacttgcgaatgcaagggcagagcttgag agtccttccaaaataatcaaggaaatctca atgagc
392-393	Oribacterium sp. NK2B42 G625DRAFT_scaffold00 048.48_C, whole genome shotgun sequence GENOME_ACESSION: GCA_00042445.1_AS M42444v1_genomic GENOME_ID: 51832 CONTIG_ID: 47 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atggtttatgtgataagtaagaccgacatcctcttatgctcaaatcggtatggcaaaagt gagacggatgcttaaatcaggtcttctgtaaaagtaataacagatatccgttcacaatacag ctttatataatgctacagattatacaggatgtattatagggtgtagatgcaggcagtaag aacataggattatcggctacgaccattaaaaagggtattgtcgaagccaaagtcacactaa gaaacgacatagttgaacttataggcactgacgagcagcgccgtacaaggcgtaac cgtaaagacaaggtatcgtaaagccgctgattcgataaccgtaaacgttcaataagtggttag ctccatctgtaagacagaagatggagctcatatcaacatggtacagaaggtgcataagat tctccctgtatcaggtataaaggtagaagacgcttcttctgatatcaagaagataaaggccc ccaatatacagggtactgactacaaaacgggtgaacaaacgggggttctggaatacaagg gaatacgtactgtgagcagcggctatgtatgccagtctgttaaaggcaagtcgaaggat aagatacttaattgccaccacatagaagtcgtaaaacagcgcggtgattcacctgacaac ctcgtaacactgtgcgagtgatgctataaacagtatcatatgggaaaaataaaacccctaa gacgatacgcagggaatgagcttcagggtatcatcattatgggtatcatgagatgggc gggtgtatgacagactgaagggtattatctgatgtgaagcctcatatggatacaca aagaatacagcagcaggaataacctgcctaaagaccacatagatgcacgggtatc agtggcaacccgacagctgaaagtgtggcacagtatatctataaagaaagtcgggtg catacagggcagatacataatgtaagatactaaagggaacacatagcgtaaagaaacca ggcaccataaaagtacatgattcaggctgtatgataagggttctgtataataatgagctgt ataccatatatggacgaaggccagtggtttttcgtataaaggaaacctgatgtacaaa agtcaaatggaagtacagctgcaaaaaactgaaactcgtggctatgcagcatggctat atagtagacatgactgcacccgatga	gtcaatgataccaccacctaaggtagtgg gcttgcgtaaaaacaagctatgttgactagc ctaagtgctcggcactacgttaagagagaa tacatagtcacctgtggtatgtatgacaaagtc tgacgtctgagggtcagtgattaaacagccc tgtgtataggggcagtggtgctgacaaaaa cctcttttaaccttggcagtggtacacaaac ggtcagatatgtaccggcttacagcataaaa catatcaatacagaaaggagtaccgcatat g

SEQ ID NO:	Name	IscB sequence	ωRNAs
394-395	Ga0099364_10003290 GENOME_ACESSION: TergutTh196P3_SF_330006226 GENOME_ID: 36060 CONTIG_ID: 3289 SOURCE: JGI DATE: 2019-01-27.	atgcctaccgctagaacagggtgggtacgtaggcgcttacgtagtgacgtgcggtcgtt acttcccgcattcccgttactatccagttgtgctatgattctactgaacatgtgcaagaatgc acttgaagtagatagcgggaagctacgtttgttggaatttcagtaactacaggagaaaagat tataaaagaggtattcgcagcagaagcacacctgagaacggatattgttgattgctatct gctagaagggtgcttcgtcgaacccgtcgcgttaggaaaactcgttatcgaaggctaga tttaagaatcgtaaaaagcctgatgggtgctccgcttctgtcgttggaactgtatgca cataagcggtgattgctatgtgcataagtttttgcctataagtgacattttagtggaactg cacctttgataatcagaaaaataatcctactatttctgtattgttatcaatgtggagat cagttaggctccagaatgtcaaggaaatgtttgtatcgaataaacacaaggtgcagatt tgtgttaaatcgaagtgaaacttcattgacaccatttgatcaaggcaaacaggtgga gatgcgcctaataatttgggtctttatgttagaatgtcatgatattgtcataaaggcgaag taaaatfaagaaaaacgagggcaatcttttcgcatgcaacacagatga	agagctccgctcataaagtaaacacccac acgtttattgctgtagacatattgctttgttt actagattaaagtttaagcccggttcgttc gggaacataactcgttgattaggaataca taggcacttcgggatactgtcctagtcccgga atgctcgggtttatgattaaattctgatg gtaggagcggtgtcataagcttaaacccct tttaacattatcgaaggcaccctacggga ggaatcccgcttatagaatttt
396-397	human gut metagenome genome assembly, contig: NODE_6243_length_689 5_cov_3.894444, whole genome shotgun sequence GENOME_ACESSION: OGOW01.1 GENOME_ID: 4581 CONTIG_ID: 6242 SOURCE: WGS DATE: 2019-01-27.	atgagtgtgttcgtgtagggctgaacggatgccggtcgtatgcctacatctgaaagagag gccgtttattactgaacatgggaaagccttctgtttatcgaaagtccttaccataaaa ctaaattataagacaggcagtagaccacagcctggttatttgggaatcgataccggatcgc agcatatcggaggttccgtgttccgcgaagatgtgacgtgttatacagaagagatcg gttcagggttccatgagtaaaagaaactgtatggatcaaggcttcttaagaagag gaagatgcatcgaaagaccagatattcgcacccaaatggagataaaagccaaactgt gtttattacgaactccagaccgaaaagggaagacactggaagaagcaaaagatcaggt cgcgtccaaacgaccgagggctgtgcttccatcgtcgtcagtcgaaagcaggtatc atatccagtgatcagaagttacaggtatcttccagatgatatcgttttctgacgaac ttggctgttgcgtccgcaagaatgaagatagacagatccatggagatctataccagaa aggaccacagtatgactatgaatgttgcgttctgttctcgtcgtcgtcagatgacagatacttg tcaggtatgtggaagaaagggtggaataatcatatataccatctctgtacagaagtc ggtgcgaccgataatccacagatattggtacgatatgcagcgattgtcatagcacagag aacatcaaccgggaggtatcttattcagtggtgcaggagcagaagaatttaccag aggactaagagatgctacattcatgaacatcttaaggaaacgtctgatgaagcatttcca aaagcaacttttacctatgaaatatacaaaagcagacagagagaactgaagattcct aaaagccatggaatgatgccactgcgacgcaatcgtaaaaactgggatcatgacggt aaaggataacgaactgtgatctatataccacaggctccgaagaaaaaacgttctctgcat gaagaaactccgagaaaaggacgcaaa	tagaataaccgagcatgcagaatgtatgga tcagttatctgtatgacctgaagggcagct gactagactaagcacaacggaacttacagc ttccatgctgtgaagatatacattagtctac gtccagatataatcccgaccgggtggtta cctaagccggcacagggttcaggaacg atgtcttccggagaagatataccgacctg gtcagggaatataccggaactgtcgaag ggtgaacactccctgtgggagggcagga cttttgcgtacctccaataataaaggag ggaccagccatga
398-399	Ga0209948_1000490 GENOME_ACESSION: SalPonR2_CSPAdes_3_SF_3300026120 GENOME_ID: 35496 CONTIG_ID: 489 SOURCE: JGI DATE: 2019-01-27.	atggtattgtattgaataaaacaaagaatcagttatgccttgcagtgaaaagaaagcaaa aaagtatttagaaaaagaaaggcggttattaccggcctgtgcccgtatggtataagacta aaggaaaacaaagattatgaataaaagggttgagattaaactgtatcgtgfcacaaaga caacaggatttgcagttctgatgcaaaaagcgaagcgaagcgtatgtttagg ggaaattatacaaaaacagttatgaaggcaggttgatgacagcggttcaatgagaag aggaggagaacagggaacacgagatagatagactcgaagatggttaaacagagcagct tcaaggagaagagtttacccttctctgaatcgagattgaatcaaacgttccacgctgt acaaaatgaatgaagtggttccgataggtgcaatcggtagaacacgttaagtttgata ctcagaataatggaataacatgcaatgaagggaacggatccaacaaggcacaattagc aggttatgagatcagagaatttactcgaataatccagagaaaatgcgctattgtgggg ctgaaaacataccgcttgaagtagagcatatataccaaaagagtaaaaggcgacacaa aggattgacaaccttgcgaatgcttgcacgattgcacaaagacaaaggaaacaata cctaaggatgcttgaagaattgcggacaagcaaaaggaaatcagacaatgacagagc gaagaagtttgaataactcagagatgcaagaaacactgaagatgctgcggctg ttaattcaaccagatggcgtttatgaataaagaaagaaactacaccccttgcgtgaatgtg gttcaggagcatttacaagatgaatcgattcggcataatttgcctaaagaacattattt gacgttgttgcattggaaaagcagccgggaataatcaacttaaaacaattatattg aagaatggaagcagttgtagagggaagacaaaataattatgcctaaataatcggat tcccaaggagacacaggtatttttgcggaagataaaaacgggaacaaagtaggacac aggagaagaaagaaagttaaaaacggattcatgaccggtgacattgtgaaaggcagc aacccaaaaagggttgcgtgaacggatgagtgacagcgtgaaagccacaggctct attatggttctcacaacggtaaaagaatagcgttttagctctaaaaacacccgttttgattcag cggggagatggatggcaattcaaacgttaaaatataa	agcgtttatagaatttgcgtcaagggtccagtt ttaaacgtttatagatcctcagctcggctgat tgaaccagactaagttccagagaactacgt tattcaggttatgacacccatgagcttctcc atgtcggagccatgtcgtatgcttttaaca ggcctaaaggatttcagggaatgccagtga gtatagtgtaaaagcctttataacattgtcga ggagagagaacgaagttccgtaacaagc cgctgaagcggaaacccctcgggggttttt gaataccagaggaggaataattag
400-401	Ga0123355_10000410 GENOME_ACESSION: TergutEmb28metaG_6_SF_3300009826 GENOME_ID: 36034 CONTIG_ID: 409 SOURCE: JGI DATE: 2019-01-27.	atggctgattttgtttgtattgacgctgaaagtaaaccttttagcacctacgaacgttaacg cggttggatccttattcgtaaaatcgcgcgcttttagtagaacgttttccaatggtataaaa actacacaaaattgtagaaatccaattgtcgaatacaactcttgaattgtatgacggaagt aagtataccggtattgcagtaattcaagaattgtccactaaaacaaatgtattcaagg cacattggaacaccgacaagatgtaaaaaattattggaacgcgcgctagatattcgtaa attacgacgtttcaacaacgacatgtaaacgcggttcagtaacgttaaaatgcaag cgactcgaagagtagcccccagtttttgcagaacggcgaacagtagtaacccgattacta aacaactctgcaaatggatacttataacatcatccttgaggatgtatcaattgatatta gggcacttaccgatggatagggcttacaagtggcaaatatcagaaaagcaatcgtttaga cgagaatctcgaagaaacaaatcatcagagatgacaataagttgtgcaatgcggaa aaagaaatcaagaattacaagtagcatattgtaccaagcgtatgaagggtcgtat	taaacaaatgtgacacaaatggtcatatgtt tataactattgaggggtttcaagccataagt agtgcgttaaacggaagcattatgcacga cgaactacgattagcgggtaagttaaagac gcaccaatggatgtatgcatcagtcattgct ctgcgagtgaggcaggtggccgactacta atgtcctgtatgtaaaacacggaatcacatga acttccgtctgtattggcaagaagcaaatg ccatgaagtgagggtgtcagaatggc

SEQ ID NO:	Name	IscB sequence	ωRNAs
		attagctaacttgattacattgtcgttaaatgccacgaaagtataaatggttggaatatga atttatacaactattccaaaaagatagagggtcaaaatatacagatttgattacgctcaaca tgaatgcaaggcaaacattgggttaagaaccaattatcaattcttgcactctggaattaac agacggtggcacaacagcaaatcataggattggtgggcatagaaaaatcacattgtaa cgatgctgtgttattacaggattatataccaacattctgtcacaagaagattggatataag gcctaaacgtaaaaacgcaaaatgaacataaggttgatgttgcggtttcgacatggt gattacgctaataatactgatacaaaagggttacttggcagggtatataactgctatatac ccgataaaaagcagtttaattacagtgcaaaacaaaacattgaaacgggttaagtctaca aaaagtacattgttgcccaagtcaccacaattgtccatgttttag	
402-403	Ga0134588_000160 GENOME_ACESSION: IMG_3300038309_SF_3 300038309 GENOME_ID: 311742 CONTIG_ID: 159 SOURCE: JGI DATE: 2020-04-07.	atgctagtgttatgtttaaaaaaattgggtaccattgatgcttgcacaaaccagcaaaagc aagaaaactgttaaaagctgtaaggctacagtagttagaagaacaccgttactattcaa ttaattgggattatgaaatcacattcaatctgctactttgggtgtagatgctgggtataag gtatgtgtatatacagcagtttaagaacaaaagaattattgtctacagaagttaaagtaaaa acagatgttagtaaaaaactaacgtgaagaagaattgtatcgtagaatgctgtagaaacaag ctatgtgtatagaaaaccaagatttctaacagaaaaaagaatgatagctgggttaacaccaa gtgttcaacacaggttagatagtcatttaaaggcttaattgaattgtatgctcaatcttaccat aagtaaaaaatattgaacacagctaaatttgatattcaaaagattagaatcctggtaattca agtactgaatatcaaaatgggtgaacaaaaggatttttgaaatgtagaattatgtttattata gggataatcatcaatgtcaattgttaaaaaatccaatataacctcttaattgtcatcatataaa accaagaaaagatggtggtactgataaacagataatttaataacgttatgtgaacctgtc atcaattatatacaggggtaaaaataattaggttaagtaataatcaaaagggttcaag cagagagtttatgagtataattagatggagaatatataatctcaaaagctatttattcaaat tttaatttaccatggatataatactaaaagtaaaagaatagaactgggttatcctaaatccc atgtaaatgacgctttgttatagctgggggcacagagcaaaatagaatagaattgtga tagttatttcaacagaagaataatgatcactcaattgaatagaatgggtttaaaccatct gttaggaacagaataatcaatacaaccagggtgatatagttgcttaataatattatctatt ttgttaaaagggtattttaaacaaggtaaatatataaaattaatagataagtcacaaaatatag ttaattgttaattatacaaaagtgagggttaataacttatggttaaggattacaattcatcga	gtcactaccaaccacttatagaagtggtggg ctgtgagaacaagttagttgattagggag ctatgtatgagtaagcagcagttatcaagaa gagatacatcacacctacagatacttctta gctgtgtagtaactgtgatctatcattaaacatg catgagagggttaagtgcaagtgtagatgatt taaaaactctgtatacaaatcccgaagagaa ccaaccatatctgatagtgggctttataagca aaagt
404-405	uncultured Clostridium sp. isolate RUG13060 genome assembly, contig: RUG13060_asm_28, whole genome shotgun sequence GENOME_ACESSION: GCA_902785295.1_Rum en_uncultured_genome_ RUG13060_genomic GENOME_ID: 346965 CONTIG_ID: 27 SOURCE: NCBI Prokaryotes DATE: 2020-04-22.	atgacaaattatgttttgaatagatgtaattgtaaaacattagcaccacaaagaaca aaaagcatggtatcttattcgaagaagggtgctacattgtaataataatccaatggtaat acaattaaataaagtgattaaagataataatattgttaaaagataaaattcgttggaattgat gatgggtgactctcatgtaggaattgcatttaattcaaaagatgtcaacaaagaacaaaggttct ttttaaaaggaaactattgaacaacgtaatgatgttaaacatttaagcaagttagaataatcagtagtcttcta aaagacagagcagaatagctccaaatttttcaaaaacgtcaagctataaattagatgaata aatcaattaaataaagtgataataataaaaaattattgcttgaaagattgttctatagatataa gagcattaacagatggatacaaaccttatagttggcaatatacaaaaatctaatagattagac gaaaatattcgtaaagctattattttaagagatggctgtaaatgatggaattgtgtaaaagca aattgtagattagaatcatatataaaaccagaagattaaatggttcaaatacaataga taatttatacatattatgtaaaaaatgtcatcaaaagacagagggttaagaagaattgtttat gaacaaatattttctttattgggttaaatcaaaagataaaaaatttaattatgcaagtcag taatgataggaaagaattgttaagagaacaattgtcattctgtggaattattttaacaaa tgggtgagatagctcaataaacgaattgattggaatagaaaaaacacattctaagac caaatgtataactgaaatlaaacctgacacacagatataaaagattggattataaaagc tatcgtagacaaagtaaaagctaaaacagacaatgtattaggaataaaagcatagagattta gtggaatataataaaaaatggagaatacatagagggtatgtaacacgattgtatccag aataaaaggcattaaatttcaaaagtcacgaagcattgtaaaaaagtaaatgctcga atgtaaaactctttggaaatatacaaaatatttgggttagataatgtgaattaa	tgttgaacatagacaaaaggagacagatta aatctttgtctatgcttagcaataggctcaa ggcttagtgactgctactatcgaagatagtg tgcagatgtaactacgtgtgacagtaaggta aagacacaccaatggatgtaattctcagtc attgctctgtgagtgccaatcaagaacaat gctaatgtcctgattgataacagagaacaa catatccctgtatgacattgccaagaggaaa aatactccgaaaggaggtgtcagaatga caaat
406-407	Ga0066903_100182758 GENOME_ACESSION: PanSoiDrversion2_2_SF_ 3300005764 GENOME_ID: 34135 CONTIG_ID: 182757 SOURCE: JGI DATE: 2019-01-27.	atgagctcgatcttttgaatatacaatagcgccccggatgctggcgccagtcaccag gctctgatgcttgcgattaaacaggcgacgaacggaagccagtgctgcaagcaaacaa accgcaacggctcattgtccggcgcaacattacgtaacagcgctgctgctgcaaatcga aaggtttcttctgataatgagatgctgttttagacgcgacaaaaaacctttggcgccg acgacccccgcccgcggcgccgattgcttttgaactcaggcaaaagcagccgtcttctgctgc taccattacgctcattttgaagaggagggtgctggcggttcaaacgcccgatctccgat tgaagtcgatcccgctcaaaaacgacagcgctgcggtgttcacgagaagccggc gaggtcgtctcgcgctgaaatcggaatcgcggcgaggcgatcaaaaaagagcctcg atcgcgacaaagcgagcttgggggctttgaactgcgcgaaggcaaaagcgccgctg agaaagtcgagatcgcggaatcgcgacgcttggatcaaaagccactggcgaggga cgcgacaaaagtgctgacgaacaaatattgctttccgacatcgacatcgacagggcg aaaacattaattgggtttgaacggcgacatcgtaaggcggaattccgagaggcaa gtagtcgggtagatattgctggacgggtgacgattcgtcaaaaggccgtcttcaatttgaac ggcttcgatacgatccgaagtat	atctgctatgccgagcggtgacaacgcttt cggcccccagctcaactctctccggcgtctc tcgaaagatagcgctgctgagcgtcaaacac gatgagctgcgactctttcgaatataacaata ggccgccccgggatgctgctggccgactcca ggctctgatgcttcgattaaacaggcgacga acgcaagccagtgctgcaagcaaaacaa ccgcaacgctcattgctggcggaacattac gtaacagcgctgctgctgcaaatcgaag gtttcttctgcatatgag
408-409	Microcoleus chthonoplastes PCC 7420 scf_1103659003824	atgctacgagtagcaggttctatcaagaagaattcaataaacgtaaacagggtaaactac ctccctctatcgagccaacaggcagttggagttacgagtagtcaaacactgtgccaact atttccgattagtcgattcactacgagtagtcatgctgacgtagataggactagggga	ttcaacaaccccaactacttctgttgggttg gggattgcccgaacaaacggcaatttaaac gttgaataccgattgagctcgttggc

SEQ ID NO:	Name	IscB sequence	ωRNAs
	genomic scaffold, whole genome shotgun sequence GENOME_ACESSION: GCA_000155555.1_AS M15555v1_genomic GENOME_ID: 40746 CONTIG_ID: 45 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	agaaaattagcacgatccggcgttggtcttccccggctcatggtgggacaaggacaaat gctgacttgggtatctgaattagctccagttacaacccacaaggatggcagcgagacgg caatggaaccagtcacactcagaagtggtctgggactaccaaggataagaagaataaa gcttgctctacacctgcgacccatgctgttgacgggtgaattctagccgtttcaggttcatt caatggcgcgaatggcattctgatcatactaaacacgggtgactggagaggtgatgttcag attacacccgcgccatttacgatcatccgcagacccccaatcagtcgtaggcaggttcatt atgtgtccctctaaagggtgaaagcggcgtaagtagcgtggcactgttactgtcacgt agagacgcgcacatggcgcgtctttacgcaagggtgatcaggttatcgtgaaaaagccg ggaagtcctatgttggttggtctgttgatactgatcaagcgatttcggttagtgatgcta actggaagcgacttggcagtttacggctaaaaaagtcggttggcagcgaaagcagc ggattaattgtcgcgccttcagttggattgtcaaacctaaccgcacgcagcggcgcggtt ga	agacacccgggactacttccctagtcggatca aatctaaagccttttgcgggctgtgtaaga caagacatcttaaccgagatggcgggaag ggactacattcacttatctcacttcgctcg atacaagtcgctcaatgttaaacatttcactc gtgaggattatctcatgctacga
410-411	Ga0376082_0035695 GENOME_ACESSION: CLC_S1_MG_FICUS_ FD_\$F_3300035674 GENOME_ID: 312359 CONTIG_ID: 35694 SOURCE: JGI DATE: 2020-04-07.	atgtcaaaactacgtcttctaattgaccgcaaccaaacaccgctaataccagatcatccaaa acaagctagaaaagttgttgatgcaggcaccgcagcaggtgttcggcgatattccgttcaat ttaataactcaaacgagttattgacaatcccaacgtttatcctctgacactaaaattgacca ggctcaaaagtttactgtattgcgttagtaaccaacaaagggaatgtcatttgggggatgg aattgcaacatagaggtaacaacaaatcaaaacgcgccttagcacgtaaagcactgcgt gaggacggggaatcgcaatctgttatcgtaagcgtcgaactcgaacccgaagcgtc caaatgttggttagcaccatctttgaagcgcgcgtctgacacccaaacccgggtcaa acgaattcaaaagtttgcttcaattgctcaattgttcaagaactgtttaagttcgacacca ggcagtcaccaaatccgaaatctctgagattgaataccggaccggaaacttgaagggtta cgaatgtcgcgaatacttattggagaatggaatcgtcaatgcgcttactgtgtgtcaag gatgttcgctcgaaattgaacatacgaaccaagtcgaaggtgttcagaccgtatata ccaatctgtgttgctgtcacaagtgaatcaatgcaagggaataaagatafcaaaaga ctttctaaaaggtaagtcagagctgttgaaccgcgttttgaacaagccaaactcccttaa aggatcgccgcatcagtttaactcaaccggcggtggcattgttcaatactttgaagcttttgg attgcctgttag	cggtaaaacctaccagagttggacattact gacaactctaccgagcgtgcagactcagct atttcgtagctacgattttgagtcataacac caacaaatgaaccgcaagtttggctctgttg attgtcattaagccacggattaacggtttataa atccgctggccacggcgaacaaagctctta aatcattgtcagggaacacttactagaaat caggctgacagaatatgtcaaacatc
412-413	Actinoplanes derwentensis strain DSM 43941 genome assembly, chromosome: I GENOME_ACESSION: GCA_900104725.1_IMG - taxon_2634166338_annot ated_assembly_genomic GENOME_ID: 183929 CONTIG_ID: 0 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	ttgaacacaggtgagtcggccacggcggtgtcttccagcccgctcctctggaatc gcggtcagcagacaaccgctcgggtcgggacgaacgggagcgcgacgtccatcgtt gtgtgtgggcaaacgggtgttcaacatggcgaggggagaccgctgttccggttccgg cacagcggcgctcaccggcgacgtgtggaaacgtgtcgggagcggcggttaacgcc cagcaacagaggttgaagagtcggggctttgtactggacagcgccggaaggcac tcacgcatcgctcggctcgcgcggccgacgtgtgtgtcgtcggcgggcagccgtg caccgcggcacccttctgtcatcgggtgcgcgacgtgtgcccgcgacgaatcggctgt gccggcgctagaggccggagtcgacccggatcgcggtacacgggacagcgtcttc accagccgacccgacgagaccggcgccgagccggtgtgtgtcgcacccgggtcctac agcatcggggtccagcatcgaggcggtcagatagcgcgacagactaccgcacggggc gcgctgcggcaggtgcgggacgcggaatcaggtacccggcaccgcggttcgac aacggccgacgaccggcaggttgctgtcctcgtgttgcggcaccggatcgagacga ccatgtcctggaatcaaccggctcggaggtgggtcctccggtcaccgcggtgcacgtgga acacgtcgcattcgacactcaactgctcggatccgaggtgcaagggtcgcggctatc agcaggttgaacacgcggcgccgctcgcggtgcgctcatccgcaggttcaacatccg tacggcaatgggactcgtccagggaatccaccacggccatgctcggctgtgcaacgc ggcgacgctgtctaccgctaccacgaacaggcgtttcgaatccccgcggcttga	gtcggggtcggccggagagcccgggtcgg tcatagcgtgtgtgaccagcccagtgagc cttggaaatgagggaactcgttgaacacag gtgagtcggccacggcgccgggtgtcttc agcccggctcctctggaatcggttcagcag acaacccgtcgggtcgggacgaacggga tcgcgacgtccatcgtgtgtgtgggcaa accggtgttcaacatggcgaggggagac cctgttccgggttcggcagagcggcgta ccccggcaggt
414-415	Ga0310375_1000145 GENOME_ACESSION: IMG_3300028923_\$F_3 300028923 GENOME_ID: 278994 CONTIG_ID: 144 SOURCE: JGI DATE: 2019-06-01.	atggggaaaggcggcgtttaccggctgaaccccttaccataatcctgaaggggaagt gaacaaccccgatgcgcggagcggaccgaggtgaagtgcaccgggcagcaagat tagcggcatagccgtggtcggccaggggagcgtgatcttcgcatggaaactcatcacc gggggcagcgggataaaggagccctttacaggagggtcgtctaaagcgcggcagac gcaaccgcaagaccgctaccggcgccaaaggttcaacaaccgcgcacgcccgaag ggatggctcccggctcccttcagtcgggggtgacaactgcgtcagctggatgaggaa gctcatgaggtttgtcccggtcacggagtgccatgtcgagaccgtgcggttcgacactca gagactcgagaacccggagatacgggtatcgatgatcagcaggcagcgtcatgggc tatgagatcaggggatacctgtcgagaagtggggaagggaagtgcgttattcgatga gaaaggagtcggctcgggttgagcacgtagtgcctcggccaggggatccaaccg ggtgaagcaacctcactctcgctcaggtcctgcaatgagaagaaaggcaataagagca tcgaggaaatttctgaaaaggaaaccgataggctccagagatcaagtcacgctcaga aagccctgaaggacgttgcggcggtcaacgcgacccgcaacgccatataatggggccc tcaagtgttcgggtcccaacgtccatgtggagttccggccgcaggaagctgaacagg gtgagcgaggatagagaagtcaccactggatagacgctcctcgtcggcgagtcgc gtgcccaggttccatcgcgggggtgaacccctcgagataaggccatggcgaggg gctgccggcaagtacgcatgacggacaatacgtttcccgagagggaaggcgggac gcgtgaagcgcgtcttcggcttcagcagaggaccgcgcagggctagcactgccaa aggcaagtacgcgggcacatgggaaggggcattgccgggatcagggaagggggct atcacgacatcaggtgtggcgccgtgaagatcaaggcaaggcactgcaactttaactttt acaacggcgagtgatgcctacgcatga	accttcgacgtcatgtctcaggtcaaccgg gataactcatgcttcggcatgagctcaacca gcatcagccgcaaggctacgttatggcgga atacataggcacttcgggtgtcttcacg cccggaccctcgcggccggtttaaagacg acggcagccgtgaaacggcgagtgaccg cggcatacaaaaccggcgataacatttgc gaggagaccattacccccgaaaggggaga aagtgaactttattatgcagaga

SEQ ID NO:	Name	IscB sequence	ωRNAs
416-417	Ga0265297_10088569 GENOME_ACESSION: Munlanlwell13791_2_\$F _3300029288 GENOME_ID: 280701 CONTIG_ID: 88568 SOURCE: JGI DATE: 2019-06-13.	atgaataatctagctttcgttaaagcagaactggcaagcccttaatgccatgcagccgag ccaaagcaaggcatttgcacaagctggcaaaagcgaagtattagggtgagccctaca caattaaacttctgttgattgtagtgaacagacacaagagggtttagctgtatggattgtg gaagtaagaatttggagtgtcggtgttagtaagaatgataatgaattaaaggattgttt aaaggaggaaagtcattcttaattggtgatgtaagaagaaaatgactcagagaagatgt atcgtgttaataaggagatataaaaaactcgttatagacctactagatttttaaatcagcaa gtcaaaatagaattggaagattgacccaagcattaaacataaaggattgccatctaaa ggaaaagaaaatagtagaattcttattaccagtactaaatggaatttgaacaacccaat ttgatccataaaatctcaaatcctgaagttgtagattaccaagtgagatcaaaaggga ttttataatactaaagaatttatcttatacagagataatcacatgtcaaatctgtgtgttaa aataaagaacttcaagttcatcataaattgaacgatcaaaagggtgactgatgaccta aaaacttacaacttctgtgtagaatgccatgataaagttcatagtggaagattgaaaattt aaaagttagaagaagtattactaaaaatgctgacatgttaattataaagttcacaaataaa aaagcattttggagattatattccaccttggatgaaactaaataaacgagaattaatg ggacttctaaaacacattataatgatgcttttagctatatgtctaatgatgaagaagctaaat ctaataatatacaaatgtttaaattattattgtaaaaaagattgtagctaaaggagattaca gcaacatagagggaagtcgttcgaattaaagatacctacaagaagttattgttttagaa aatttgataaagtgaaaacacctaaggaatagggtttgttaaagctaaacgtgcaaggag ctatttcatatctgtgatgtgttgtaatactgtagtagatggtattaatgtaaaaaattac tagagtgctgctagaaaaaattatgatgtagactataaaaaatacaagaagaagattga	atcaaacacccctcccgaaggaggagggttc agaccgtgaggactgaaagtaattgtttgat cagaagacaacaggaggatgttttaacaga atgttttagccgttaaaatgaagaaaacat acacaccagcgtgtgccgctcagcacgtt gctctgtgactgtatattaaacagattcaac gactcagtgatagtttaaaaaactatttta actctcagaggatgccgaattcacacaaa cacttcaaatgtgtgatacgcatacgtgtt actgtacagggaaagtatatgaataa
418-419	Ga0209608_1000352 GENOME_ACESSION: AD_des_41_\$F_3300025 609 GENOME_ID: 21341 CONTIG_ID: 351 SOURCE: JGI DATE: 2019-01-27.	ttgagagtggttcgttaataatcgtgtggcaaacccattatgccatgtaaacacgaaaaagc aagattgttattgaaacaaggtaaaagccaaagtgtgcaacgagatccatttactatccaatt gacaattgccactggaggagcaaaacagaatatttctctgtgtggttgcaggttagtaaat atatgtgtatctgcaactaccaaaaaacatgtttttgttgagcagaagttgaactcagaa atgatatgttaaaacttttagctgtagagacaatctcgtcgaagtaggcgttatcgcacaa acgcggtatcgtcaagcaagggttttgtaaacgttaaaaaacccgtgaaggttggttagctccga gtatacaacacaaaataaatagtcatatcaaaaataatagatatgctaaccaagttattacca taaccgatatacaaatgaggttgcaagtttgatatacaaaaaatcaagaaccagatatt caagacaaaagaatatcaacaagggtgacgttaggcttttacaatgtaagagaatattgtct gtggcgggataattaccaatgtcaaggtagaaaaggatgtaaaaaacaaaatactgaacgt tcaccacatcgaagtcgcaagacaggtggcaatttcacaaacacccattataacactttgtg aaacgtgccataatgattatcatgcaggtaaataaaagttgaactaaaacgtggccaaatg tttaaaatgcaacatttatggcggttatgcgatgggcaacttataatgcactcaagagttat tatccaaatgttaaatgacatatgatatataacaaacacagagaatcaaatatggtctt gaaaaatctcatagaatggatgctcgtgatgtattagtgttaatcctttggcgaaaccattagat acctattatttcaaatgtgttcgcaacaaaatcgtcagctacacaaagccactccatct aaaggcgtgttcgtaaaaaataaaagctccaagatactcatatgggtttcaactatttgac aaagtatttaccaaaacagaatgttttatattgtgtagacgaactagtgtttttgatftta cggaaatagatgggacaaaagttactgcatcagctagtgtcaagaaattgaaattgattg aaagacaaaatacattattatgtgaaggagggaaggaattcccccctgttacactat cggttag	gtcaaccaccaccactaaagaagtgagg gcttgtaaaagctctagtgtactgcctaagt cttaactgactacgttggaggtgtcatgatac cttagtactcctcagcactagtctctatcgt tcggtgtgtaaaagttctaagggttaggaac gggtgctgtagatgacaagccatttcaaca ttggcggaaggagcactaaccccttttataaaa ggaggtacgaacctgtaga
420-421	Ga0208824_1000448 GENOME_ACESSION: AD_des_30_\$F_3300025 629 GENOME_ID: 21329 CONTIG_ID: 447 SOURCE: JGI DATE: 2019-01-27.	atggtattctgacttaacaacacaaaagaacctttgatgccttgcctcgaagagagacaa gaaagttactcgaaaaacaaaagcagttatatacaatacactcgtttacgattcgcct aaaaaacgaagtggaggattgtaagttgagccattacaataaagattgatcccggaagt aaagaacaggaatcgtgtattcaagagaagaagacaaattactttacgttatgccg gaatagtgaagcataaataaacggtcgcggataaccttaaacacaggtcgcgaatgcgc cgtggccgcagaaaacagaacacaggtatcgaccagtcgttgggccaatagaaaaa acagcacaaagaaaggcggttcgcacctctatgctttcaagagtatactcaactttaac ttagttagaaaaatgaagtcattagcgccttcaaaattgtatcagttgaacacgtgaagttc gatgtcaaaagtttagcaaccccgaggtgtcgtggtgaatatcaaatggaagttat tcgggtatgaagtgaagaatactgtctgaaaaatcggtaggaatgcgttattgttg tgctgaaaaagttcctctgaaattgaacacatgattccgaaaagcaaaagcggcgacgga ccgcattgataatctgcaatcgtctggttaaatgcaatcaagagaagtcgaatgatgc cgaaagatatatagaatacttatcaaaacaaaaggcgataaagccaaaaccatgatc caaatctgaaaaggcgataaaggacgcaaacagacccttaaaagatgctgcaagcgta aacacaaccgatgggtttatgcaacaagctcaaaagggaatttgagacgttgagctc gctctgcgggagaaacgaaattcaacaggcacatacaagggttgcacaaagaacatta ttttgatgctcgttgtgtttgtgctgataagcacattgatgttaaacacaatacgaaca ataacaagataatggccgttggaacagacaaacgattataccgacaaacatggcttt gcaagaggctatcgagtcgcaacaagacaaaagaaggtatcatggtggcattttgta aaagtcaaaagcataaccggagcgcgaatcgcgggtgaaaagcgcgggaacagttcata tcagggataaaaatgggaaggaataagttgtagcactaaaaaggcactaatgcttaac acggcgatggttggcagcgttctattttgaaaataaacttttaacagaaaatcagaaaaata caaacagaagaaaacaaattga	atcacggattaattttaccttatgttagcggtt tgttaacttatctatttacaagcgggtccaga ctaaggtcttagagacctacgttattttgtca tgacaccagcagggtgttctccagccggct gctctgtcgaaccagtctaacaatgcgattgt aaacacggaaggtacgaagtgattggctc gtaaaaagcattataacattgtcgaaggaga gacgaggtcatgtccctgcgttacaaaggcg cgttaagcgaagccttcgggctgtaataaaa aggagtgatacctgtatg

SEQ ID NO:	Name	IscB sequence	ωRNAs
422-423	Ga0223845_10169920 GENOME_ACESSION: 3300021387 GENOME_ID: 241586 CONTIG_ID: 33254 SOURCE: JGI_Dark_Harvest DATE: 2019-01-27.	atggtatatgtattaacaagaatgtgaacctgtatgccactaagcgtcacggaagg tgcggcgtatgctgcgcgatggcaagcacacgtggcaaggagcttccattcaccatcc aacttgactattccaccagtggtctatctacagcccgtgctgctgttgacgcaggagac acagcacatcggcatgctggcaagcagcgatcaaaaagaacttttcagtcagaagtgg agttgcgcagtgatgtgaagaagaaatcagagcgaaggatgtaccgccgtaacgt cgctataggaaagacgcgctaccgcaagccccgtggcagcatcgtggacgcactgaaa attggctcacaccaaccgtgaaaaacaagattgaaagacacttgacgttatcggtatgg cacattccgtcttccataacgaagacggcgtcattgaaactgccagttgacatagcaaa aataaagaacccaacattagcgggtgtgtaatacagaacggcccgaagaagattacg gcggcgctacgcgagtatgtttgtggcgtgacggacacaagtgctgtcattgcaaggga actcgggtgacaaaattctcgaagtgcacacattgagaccgtcagactggcagtaatg ctccagataacctgtcacgctgtgcaagacttgcacaaaggcataccatgaccacaaaa tacatttgacgtgaagagcggcatcggcatcccccttccgcatcttcacagatgaacat tatgctgaaggctttactgaacaaagctagaacgatgttcttaattgccataacactatg gctacatcactaaagacacgcacatcagtaattggcatagaaaaacacactgcggcagc gcttctgcatagctgtaacttgaagccgagagactagatacttctgctcgtccactgt ctaccaagacacaaaggctcattgcatgtagtaacttttaggaaggaggcgtaagacg ccctaccgtcgtcctcactggttaaatgaaatctacgactacagagatagacaggtg gaatggaacggcatccgtgttctattccggaagtacaaatggaaggcgtgttctaagag acataaactggaagctgttaacaccaactacttccgtaaacgcaaaactgttaggttcta tgcggcctacacggaagactattatcatcacaacaaccgcaaggtgcaataata	gtttactcacaanaatgctgtcattaacagtc gatagtgagtaagccgattagcctcagtcatt taaacatcatatcgcagaatgtgaatataaat gaactacgttagaagtgaattataggcacc ttgaaatagtcgcccaagttcaagctctcg gtcagtgattaaaaggagcgaaagcaacag tgtgtgcacagaaccatttcataacattg gcgatggcgcataacgggcattagccccg acttacagatttttaaattttagatgatag
424-425	sediment metagenome genome assembly, contig: NODE_20_length_14530 _cov_3.617064, whole genome shotgun sequence GENOME_ACESSION: OFOG01.1 GENOME_ID: 4133 CONTIG_ID: 19 SOURCE: WGS DATE: 2019-01-27.	atgccaatccgaaacaaacagggttaaacggggagactagatgggtcaagatcagcagaa aagctgctgtttaagacacactgaactgtctgtgtggcaaggcaaggatgttgaaa tgcctccctagtttccaacctctccgacagtcagcctcgaagggaagacattgaggga ttctgaaaagaacccccgtgtgaccgggaaggtgaactccccgttctatggatcccgttg tttctaagaccgggaaacccctgatcccactcgtccaaaggggcgagagagttgatg aagaaggaagagcgatccgaagtggaaaactggtatcttctatctcactaacgaa aggaaagacggagatgttcagaaagtgtagttggaattgatccgggcagcaaaagag aagcattcacgatcaagagtaacagcacacttattgaacataaattgatgcggttact gggtgaaagaggcagttgagggcgaggggacaagccagaagaggtaggaggtacaga aaaactccttctcgaacaaacagagaaaatgccctaagaggaaagttggctccatcaacc aaggcaagatggcaggcaaaactgagaattgcaacattctgaaaaaattttccgggtga ccagtttattgttgaggacataaaggcaacaaactacacggttaaacgatggaacaaatc gttttcacctcttgagggttggcaagcaatggtttattctgagctaaaagaatttgaattctg aggtttaagcgggatatgaaccaaagaattcagggtatgtttgggttgataaagcgca aagtaaaatggaagaaaaatttctgctcacacattgactcgtgggttttagcaaatcagt ggtcggtggtcacactcaacagacaatactcaattttccgtttgattcctttaagattccat cgtcgccagcttcacgcttccagcagacaaggaatttcaggcgatcttaccggcgga acactctcgaaggcctaaaaagaggttcattgattaagcataaaaaatgggggttgcct atttaggtgataattctgaaaaaacaggattaagctgcataagctgggggacggcaaga gattgtgccagaccgccaaggaaaaagacttaataatttgaaggtttaataattggcgatgg tacaagagaaacataa	gtagtttacttcttccattccagttgaaggcact gtggttcgcccaccgggtcagggtcctctt cctgtgtgcataaaactactgttcggcgatt tactgtgtgcgccgtgatgacagccgttttga cgatcatcgggtcgggtgtgtcttcttgg caacaaccccccgatattgtttaccacgt gacagaggatgacattgacaaaataattac gagattgatacggggattgttgatgggtggg aagataatcagttttgaccagaaccagaaa tttttagcaatacaaacgggaatccctaagcc tagagattgttaaccaccccatagatgaaa atctaattgggcttgaaggagtaggaatgc caatccgaaacaaacagggttaacgggaga ctaga
426-427	Ga0310139_000809 GENOME_ACESSION: M274_2_SF_3300034759 GENOME_ID: 292993 CONTIG_ID: 808 SOURCE: JGI DATE: 2019-08-09.	atgccactccagctccggacgatgaagtccgtatctgcgagggagaccaaccggc aaaagccgtgttttaccttttttttgaaggtgggtgattacgaatgttatgttaatacca gggaaggcaggcccaatgcccgaacagagatatgcaaaagtcagggttctgctgaag caaaagagagcgaaagtgtatgtcaaaagccatttaccgatacagctgtgtatgaaacc acagcatatcagcagccgatagtgccgggatgatcctgcgcaactcaccatcgat aacggctgttaagaagaacgggtgaagtgcagatcattcagagttaataagcaggga ataaagaggtaccgaaactgatggcgaaaagggaagatgcacagaatgataaggcgtc caatcgtcggatgaggaaaataaggcatcgcaaaaggtgtggcgccacattcgaagtga cgaaatataatatacagccgggtgcagacaagccaattccagtcataataataaagtgc agaagcgaggttctgcaataggaacggccggagggtgcttaccgcaacggcga ccatctgcttcagacgcataatcaattatcaagaagatccggaaaactgcccgaatacc agggtggtactcagatagcgaagtttgacatacaaaaactgaaaatcccgatattaaag gtgaggatatacggggcagggttatgtgataatgccaatacgaagggaatataagaa gaaaggcaaaaccacaaatgtttgtatgctggaataaaatcgatagaacatttacaccaca taaaagcaagcgatgaaagcggaagtgcaggttgcaggaataaagctggtttgcaata aatgtcatataaaagtgcacaaaatgaaaaagtcagggaagcgttccgagaaattgg aaggcttaagcaagcaatgatgtaacaaacttttaactcaattatgccgtattgtataa ggagatacagaaaataatgggtgcaaaagatgtgcaaatgtctatggatacgaacga aataatgagaaagagtcgtgggtcgggaagactcattacaacgacagttatgcaatgg cattaatggcgcaaaacaacagcaagatcagaatataacccgtatattgttcaaca atacagaaggcataaccgtaattctgtgatcggaaggacaggctgtataagaaag gcagggaatagcagccaggaaacagaagaagaaatggagcagaaaaatccatcgc tggaagattatcggaagaactatacagaataggcaaaaaggaagcgcgagggt	gtcaactactcctaatgaattaaggagcttg cagctttgtagctgtcgtgtgtatctatagg gaacctgctggtacagatgtgaacaggcg gttgatcgccagcgacatccgaggtgccca ctccagctcccgacgatgaagtgcggtatc tgcggaggagaccaaccggcaaaagcc gtgttttacctttttttgaaaggtggtggatt acgaatg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aatatccgggctaaccgggtataaagcgataaagagaataagaacgtctatcaaggaaat accgatgccgcaaggcagcatagtgatgtacaaaggtaagcgatgtagtgggc gtattaaacaaaggtaatttattggtgttgaaacacatgaaggttatgtaccggtaaagg atgcaaaattactggcagaacacagcggtatagtagtctgtag	
428-429	Activated sludge metagenome, whole genome shotgun sequence GENOME_ACESSION: LAHR01.1 GENOME_ID: 1476 CONTIG_ID: 1779 SOURCE: WGS DATE: 2019-01-27.	atgcaacatgttttfgttagaccagaacaaaaaccacttgacccttgcacccgcac gggcaagaaagcttttgcgttctggtaagcggcggtgtccgtagggttccatttcgatt attctcatggagcgaacactgaaagaatctgtgacgcacaagtactgtgtcaagattgacc ccggaagcaagcaaacgggaatggcctttgctgagaaggtgaccacaaggtcgtttg ggcggcgataattcaacatcgtggacaggtcattcgagacaacctgcttgcacgtagagc tatcagacgcgggaggaggcgacatcattgtcgtaccgtcctgcacgattgataatcgt catcgccaaaaagggtgttggcggcgaagtctgaaagtcgtttggcgaacatcgagact tgggtgcgaaggtgctggtgcacacgccattgactgccatttcaatggaactggcaaat tcgatactcagaagatagaaatcctgaatatcagggggtgagtagcagaaggcaagc ttgttggctacgaagtcgtgaatatctactcgaaaaatggggacgaaggtgtcctactg cgggaaaactggcgttccgttgaatgaacacatcattccaaagtcgcggggtgggt caaaccaagtttcaaacctaacactggcgtgtgagacatgcaaccgcgccaaaaaac ctatccgccgttagaattcggccatccg	gtcaacgactgctaggctaaagacgtacg gcttgggaaaaatccaagccattgtgaccag ccacagccaccagcggtaacgtgacggg gctacgttgcaggagagtcagaagaacaca cgttgggggtgctgccagcctcaaccactg caactggatgattaaaagaacagggtcttg ttcggtgtcatctggaagtcgtgactgataac attggcgaggcaaaccttaccctcgatagg ggagtacaaccgtaagggaaggacata agaaaaatgcaacat)
430-431	Ga0373625_0000621 GENOME_ACESSION: YL18G_12A_MG_FD_5 F_3300035701 GENOME_ID: 312749 CONTIG_ID: 620 SOURCE: JGI DATE: 2020-04-07.	atgaaatgcagcagttacaattaaagttaaacacatacctacaatgcttactagttgt agctctataagttttgatttaacaaagataagctcttagtcaagacaagactggtgatt gtacaatctcgaagtgaatcaacgcttacatacaagcggctttaaagacattgtgtatgta ttaaatgtaaatgcgcacacattaatgcctgtgtagttgctcaaaagctagaaaatttataaaa aaaaggagagctaatgtcgttaaaaattatcttttacaattagattaaattttaaattgtaaaa tgtgtgtcaagatgtcatgttagcatagatagtgattgaattatattggttattcagcaatta ctaaaacaaaagaattatcttgaacattgacattagattgtcaaacgattagtagattag caaatcgcaatgtatcgaagaataagaagagatagattgtgtatagaaggtctagatt taataatagaaaaactagtaaaaaagggttgggtgctgccgtcaattaaaagaagtataata ctcatgttaattgattaataaattaaaaaaactcttaccattaaagaagttattatagaagttg caaaattacagtaataatgaatcgtcaaaattatgataatagtaaatgtcaattgtgcca taaaaatttagtaataactaatcaggttaatgaacatcacattactagactgtatggaacta atagaaaagctaataatagcttatgtacatgaatctgtcacaaaaaaattcataaaaaaaa atgtatgatatgttcgagaagataaaattatagacagctgtcgtttatgaacattattagtaa aagatttttagaagatggttataagactattttggattgtacacttttaataaaagacaagttta taatttagaaaagactcactatgatgatgctttattatagcaggaggagataatcaagaaa gagcaattataaaaaatgtagttaaaaaacatagaaataatagaattacagtttaataaga aaagggttttaaacgttcaatcagcgtcaaaagggtcaaaatcaacccaaagatttgatag gataaacacaagaataatatttcaaaagggtgctttgaaaacgcaaacgggtgacata taatgatgaaaatgtgtaaaagggtcaacaatacatagatagaattgacaataattataatttt ggaagtttaattttctatga	gtcagctgctaatttttagtttaaatggtagat ttaatgaaaaagtcgattagagagcatgaaa atgcagcagttacaattaaagttaaacaacat acctacaatgcttccatgattgtagctctat aagttttgatttaacaaagatgaagcttag tgatcaagacaaaagtactggattgtaacaatc tcgaagtgaatcaacgcttacatacaagcgg tcttaagacattgtgtatgtattaaatgtaaat gcgcacacattaatgcctgtgattgtgcaa aagctagaaaaattataaaaaaaggagag ctaattgtgtaaaattatcttttacaattag attaac
432-433	Ga0307928_10005039 GENOME_ACESSION: AceLakmetagen232_2_5 F_3300031227 GENOME_ID: 279923 CONTIG_ID: 5038 SOURCE: JGI DATE: 2019-06-13.	atggtagttagtataataataatggaagacctttaatgccatgtatgctcgtctaaagcg agacatctattaagagatgggaagctaaggttaaaaaacgtacaccttttactatccaact taaatatggaagttcaggttacacacaaagatttaattagttatgctcgtgactaagaaa gttggaaacagcagttcgcagaagtggaacaaaagaatttttactcttcgaaattacttta agaacagatatacaaaaaagttgaacagagatctcttatgaaaaaacctgtagaataa gaagacagaatagatagaccaccaagattcttaaacagaacagaagtgaaagttggttac ctcttcagttcaaaagtaaaatttactactaagaagagatagattatattttagttacttc ctaaactagaataatctttgagttatttcaatttgatattccatagactcacaaataaatttga gaggtttctggatcaattagagatagtaggttatgaatctaccaaaagctatgtcttaga aagagataattataaatgtcaatcgttaaaaggaaatctaaagataaaaaagttagaagtac atcatataatctatagaagaatggtggaactaacaaaccttcaatctacttactctgtgtc tactgtcatgacttgtacataaagatcaactatttttaactaaatatcaattaaaagctgtgt taatacaatagatgccactaagttatctataattagtaagagatttggagtagtctttatct cttaaacagattatatttagtcaagacatagtgatagtagtcaaaagttaaaagaaggta cttaaaattaaagaattctatctttggacgcagtagctatctcttatggttaggaataagaag tacagaaaaggacttaggaacacagaataatagataactttataacaagatttgcacttc taaaaggagactaccaacagactaaaggtagacattcagaaaagcaaatccaacagata aatacaagggttcaggaattgtctgttaagtagcttggtaagttatagattatataaagg aagaagatccacaggatagtagaattatggtatcaaacagttgaagaatctgaatttaaga ccaatacctaagtttgaatttaaccagggttacaagctagaattcatggattatatttccaa atacaattcaaaattattga	gtaaactacctacgaactaaagtttaggct tcaaggtgacttgaggtagtaagtattttac cagactcagaactaagaattagttctccgtt atttaagttatcagaccaagaatgccgcctc agttcttctcctctggaggcactgtaaacag agattaatatgtctcggtcaacctcagtagca cctacgaagtagtaagcttagataacattgt cgagaggaaagttgaaaagtttagtgtaa cagctaaatttagtacacacacttttaactta ggttaggaggtagactcaatggtga
434-435	Ga0315291_10038423 GENOME_ACESSION: YL17G12_20_MG_2_5 F_3300031707 GENOME_ID: 281642 CONTIG_ID: 38422	atgcagaagttaacagcaaaagtcaagaacacacaaaggatgcttcttagtcccttgc ctgtaaaattcttataaacaagagaaactcttaggtggaagataaaaactgactgtgta acaatcccgaagagaatctacaccaatagaggtgtggtcagcttgcgagtttcgaatat tgtctatgtacttaacatagaggacaacctctcatgccactactccacgaaaagcaagg ataattactacaagcaaaaaagcaaaagtcgttaaaagaatcctttcaccattcaattac atatcaaacaggtgaatctaaacaattatagatctttgtgttgattctgggtacaacacatt	gtcaactacccctcacaaagggtggggcctt gtgaaggtattcatggtgaacgttgatttag ggagcttaattcaaaagattatgcagaagtt caagcaaaagttcaagaacacacaaaggat gcttcttagtccctgtctgttaattcttcatt aaacaaaggagaaactcttagtgaaggat

133

SEQ ID NO:	Name	IscB sequence	ωRNAs
		acctgacacagtagcttatgatattgaagactatcctgctaatagacaactttggcaacgaac ttacaatctttacaagaagacatttatgtaggatcgttatttgaacatttgacacctgaa aaagttcaatttgaatttgggttcggacttcatatacaacatttgatcgaacagtatcagc cgaatgcagatgatgaaaaaacacattagaagtttaagtaactaacacaggcgacaagtt ctgagaaaaaaggtgttcaggatattacgaagcaccacagggaactttgggacagcc tgcagacagccttctgcatagcagaaaaacagaaatctgcaccgggtcagagccaga cattaaagattgttcgatattaacggaattgcatatgatgactcaggcggttacaggca acaagtcagctacgtattagaagcaggagattacaacttctacgtaggtaacagcgtaaa aaataataaactagcttatcatataaaattgaagaacttaaaagtaacagaacagctttcag aagctgtctgctaacgatgaaaatcttaattaatgaagccggcgagagaagagaag acggaacatacgaattacatatgtaccatcacagaagcctacagttgataggctaagag aatagaagacaatcttccaaaagacatgaaaattacaggcgatgtaggaaftacattacag gatgtaaaagcaggtaaaaaactattgaaggttcgttcacagcttacagttgcagaat tggcacagattgtaagaggtgaaggaatgagtaaccctaggttacaacagagagcct cagcttttgggtgattaagcgatacattgtttccttgggaatccctgcagcatgttgcgtg acggccctagtgacttgaatggaaggaagctacacagcttctatttgaacagca ctttcagcatatggaatcctaaacttgaagagaactttacacaatggaaggtcaggaaft atacggaaatcaggtagatatacttctggaccggagtaaacattcacagacatccttta aatggacgttaactttgaatactattcagaagatccatacctttcaggaaatgtcagttgca tcaacaggcgggtattaaagacgggtgtgctggggaacaattagcaatttcgcattaaatg gtcaggaaftacatagattgaagattgacgcgtttgctcagaagagcaatcagacagat ttatctcaaatcatttgaatggcagtaaaagccggcacagttaaacattatgacagctt acaatcctattaacggacactggcgacttcaactacagcctttgtacaacaatccttaga aatgaatgggatacaggggtatcgtaatgactgactggtggccaaagatgaatgacgtt gtgaagagtgccgaagaatcaaatcagatacaagagataggttgcgtcacagaacga cgtatatagtgttgaacaataacggcgacgaaggttaactcaacaacgacaacacaga ggaatcaattaaagagggaagacttacaatcggagaacttcagcgagctgcaatcaacat ctgcaacttcatttctcagcacctgttattgaaagagaattgtgacacagacgttgcaaa acattacgattcagttccaaatgatcaggccaaagtatgaagtatttaacattgaaaaagata ataaggtaattgttcaatagcggagcagaagcaacattagaagttgaagacgaaggggaa tacacaattattgttaacatctcatttgacaagtcacacttatacagtcacaagtaaacgtta atgccaacggcacaacaatggtgtaataccagactaacggaacagacggcaactgatt acacagaagctttgcaaggttaacttgacaaggggtatatacaacttaaaacttgaagaag tattagcaggaaftcaaggttaaatatattcagtttaagaagattcctaagaaaaataa	
440-441	Ga0172382_10042448 GENOME_ACESSION: Leawell138RmetaG_2_SF 3300015214 GENOME_ID: 31498 CONTIG_ID: 42447 SOURCE: JGI DATE: 2019-01-27.	atgttagtttatgtcatcaaaaaacggcaatccgttaatgccgtgaagccgtcgaagc tcgtaaactttacgcgacagaagaagcggaagattgtgaactatgccaccattcacattcag ctcaatgggactgtgaagaaattgttcaaaaagtgtctgttaggaatagatagaggttcgtc ttacacgggttactgtctatttctaagacaaagtattgattcaggacgaattgaccacag attagatattaagacaaatgactgtcctggctcggtaatagaaaaagtagacgaagccg tatgtgtatcgtgaagccagttttctcaatcagcatctagcagacgagcaggcgctta ccaccatcaattaaagcaagtgtagaagaagtgtttcgtgtataacgaaattaccaattcc tatttttgaattactgtgaaagtgtgctgattgataacgaaattaaatgatcgaagtctaa aaggtagtgcgtatcagaagtcgaataagctcaacgagaactgcgtctagcttttacttc cggtataattcacatgttattatgtgggaacaacgcaagcatgaaaagctagaagcgc atcacatcgtgcccgtttcacagaatggtgaagaatgatttataattttagtaacgttatgta taagtccatgatgatgttcatagcgagaagttgaagttgaatttaaaaggtatgggtgta ttcaagatgtatgacacaacgaactatgattgtaagacatatttataaatttgtgaaaaa atataaacaccaaaattatatttactatttgggtatgaaacgtcacattatcggaggaattag gtctagtgaagatcatgacacagatcctttttgtatagcaaatcaccacgcacgttacgat ttaacatatgaacgtgataatgtatataatgtgactttcagagcctaacagacgagacgtag atatcatgataagcctcagaagaggtgaagggtcgagttgaatatcaagtgaaatgaagctta gaagggttccggaagggtgatttgggttagtaaatgattgtgaaacagattaatcgtata tattctgatggcggttta	gatatggcttttttagcatagttccttagctag tagtatttcatagtagtgaactaacaggaggc tggctgttctagcagaagttattaggtgaac agtttacttaccctgggggttctgctccagct ccaggctctaaggacaggtttaagggttag cggaaacgtgaaggtgactgtctgaggaa accctaataacaactcgaaggagcaagaa ctctgtcaaggagaataataactgtta
442-443	human gut metagenome genome assembly, contig: NODE_63_length_10954 3_cov_6.260805, whole genome shotgun sequence GENOME_ACESSION: ORUX01.1 GENOME_ID: 10830 CONTIG_ID: 62 SOURCE: WGS DATE: 2019-01-27.	atgccctgttcaccggcaaaagcgccctctctgttaagagaagaagctattgtgaag aggcgaacgcctttcactattcagctgacgattgcaacgggtgagtcacaacagccgggtg actctgggtgttgatgccgggtacaaacatgtcggcctttccgcatcaacggaaaaggct gagctttatgcatcagaagtcgaactccgtcaggagctctctgactgctctctgctcgtc tgcgttacggcagtcgccgtgtagccgaacacgcgtactgtgcaccgaggttcgaca accgcacccgcaccaacgcaaaaggctgggttcacacgtcgggtgaaacccgaatcaac gcgcactgttcgcgcatagaagcggttcttcgactgtgcgggtcacaaagatcacccgtg gaaacggcgtcttcgactgtcagctgtgctgaagaatcccgacatttcagggaagaataac caggagggaagacgctgcgttctggaacatccgcgagtatgttctttgcagagacgg gcatgtttccagcattgttacggcagatcgaagaccgggtgcttaattgtcatcatctgg aaagcagacgtacggcggaattacccggcaacccgtattacagcctttgtgagacgtgc cataaagcccttcatcgcgggtgaatcacgctggaaggcaagcgcgacaatcgttccg	gtcaactacctcggcctaaggccgcgactt gaaaaagcccttagttgactagcctcaggcc gtcatttggcggactacgttggcgggatttt ataggcaccgcgggatgcagatcctagtc cgcgctctcggccgatgttataaagctctg agaggtaggagcgggtgctgctgcttga ccccttcaacattggcgaagatcacaacc ggtcgaagaccgaggagataaaacttga ga

[illegible]

SEQ ID NO:	Name	IscB sequence	ωRNAs
448-449	Ga0116188_1006392 GENOME_ACESSION: AD_JPNTR2_MetaG_2_-\$F_3300009658 GENOME_ID: 21268 CONTIG_ID: 6391 SOURCE: JGI DATE: 2019-01-27.	gtgaaacttgaggtaattggtgattaggggcctaatttgaaggaattatgcagacgtt acaaagagagtcaagaacgtacctacggatgcttactagtctgtagctctacaagctgt gcgttaaacagagaggaaactctcagtgctcatggcaagactgatttgaacaacccc gaagtgaatcaaccccgcaaggagggggattcgaacttgagagatcgaatattgtctttg tactttctatgagtggaagacgttaattgccatgtaaacacaaaagctaaagaaattgta aaaggaggtaaaagctaatgtttaaagatttccatttaccgattcgaattactatgggtact ggtagacagtagcagatattaatttaggcattgatagtggtataagcatgttgattttca agtataacagaaaaagagaggtgttttcagggaacattaaattctgattcgaagactaaaga tagacttaattgagaaaaagatgatcgcaggaaatcgagaaataagttaaagatatagaaa gtcaagatttaataataggaaacgaaaagataattggctaccaccatcaattgagagaaaa tatcaaaactcatttgactttgatagaaaaaattaaaaatctattacctataaaaattgtatgtt gaggtagcaaaatttgatatacaaaaaataatgaatcccaaaattaatggtaagaatatca gcaaggcaatcttttaattatcagaatatggtagttattacaggttaaggcaaaataatc tgctcttattgttaaaaaagatttaagggtgagccaaaagctacacatcattttatagacat ggagattcaagaagatctaataagacctgatgtttattgctttacataaaagttgcatgta gattacacgaaaaacatagagaaaaagagtttcaaaaacctgttaaaaggtatgagcctt caactttatgctataacataagagattttatgaagattatcaggtttgcaggtaacctat ggttattatcacaaaatgaaacataacgaatataatagaaaaaacctttcaatgatgct tttattattgcagggtgaactcaacaagtaagatgtaaacctataattatagagcaagacat agaaataatagggttttgcagttaaataggaagggttttaaacatcaattaaagagaaag aagtaaaattgcccgaagattgttttggctcaaaataattataatattcttcaaaaggagtg tttaataaaggtaaatattgtttattggtgactttaaagaaagaaatataattaaagttacatta atagataaagatttacaattttggaagttttctatggaattcttaa	gtcaatcaccctgaataaattcaggggcctt gagttgtgaaacttgagggttaattggttgaatt agggggccttaatttgaaggaattatgcaga cgttacaagagaggttcaagacgtacctac ggatgcttccactagtctgagctctacaagtc gtgcgttaaacagagaggaactctcagtg ctcatggcaagactgatttgaacaacccc gaagtgaatcaacccca
450-451	Ga0307376_10003019 GENOME_ACESSION: TR2MG_2_-\$F_33000315 78 GENOME_ID: 281014 CONTIG_ID: 3018 SOURCE: JGI DATE: 2019-06-13.	ttgaatcaaacggaacacagaccggaacaggaacaaatctttgtcatatcaaatgcgggg aaatctctggcgcacacccgtgctcccgcgaagtgctgcatcttcaaggacggaaa agcccgtatctattgctacgaaccattcactatcaactaacctacgagagcatcgagttgt tctgtagaataaacttgggtattgacctgggttccagcgacactccgataggcgtgaa gagcatgtgctgttggtcagcatatgctccatctacgcgaaagagatcctgctgcgg acagacatctcggcgcaattgaagcgcaggtctggtgttcgacgtagacgcagaggtga taagatacggcaccgttaaaccaagattgataatcgggtgaagtgcgtgttgggtcttfg gtaaaaaacgcacccctaaacactggaagaaaggctcaatcgtaaaaaaggcggtaaaag cctaaagaaagtagaaaacggaagggtcgtgcatctgtcgttaaatgtcaatgaacggg ttggcgagaaagggaacatgatgccgacaagatcctcaatcccacgcttcagaaacaaa gttaacgccatcgtatccgaggtcaagaattggttgagataatcccgtgacgaagatc cgagtagagttgactgtttcgcacccagaaagatggccaatccaaaattcaggggaga ggaataccagcaagggtacattgttcggatagcaggtgaaggaatattcttccataagtat ggccataaatgcgtttactgtaagggtgaagtaggaatcccgtcttgaagtcgagcat gtcatcccgaagaaacgtggcggaacgaatatagtatctaatctggttaattgcatgtgaaa cgtgcaatcgtgaaaaagggttcacgaaccgtgcatgaatacagtttctaatatccaaa acaagcggttaagtttgcggccttccgttacgcgcattgaccagagttacaattgggc actctggcgggaactgaagaaacttggcataccgggtgaagcaacatttggatatcagac aaagtattaccggttgaaagatgcgtttgcctaaagcccaggttagtgacgcgatggtgata gcatccggtggtcgcagtttgcgtgacgcaatgctgacgaaaggcgcctcaag gcgcggaagcctttccaccgacttgcgaacgaaacagaaaggcaagacatgtgaga aaactcccgcgatgagacaaatcaatgttttgccttatacgaaggtgtcttttggatg ggaatggcataagggtctatgttatgtcaccgggttgcgtactcggggaacctttgaggt gtcccacttgggaaggaaatgatctcgcgaaggttgaagaaacttaagttggaaga ccatatgtatcgaacaaattgatcgaagaacgtcaattatcgttacgatttcaaaact tgaagggttaagggtattccgaaatggattcaggaaagggggtgcgcggcgctcctccc catggataa	gtcaattacccacgcctaaagccgggggct ttgaggagaaatcctgatagtcttaggtttta gacggggagcgaagctccataagagacag cctaagcttaactgactacgttcttccgaatat cagacagctgagtaaaactgatcttctccgg gtcagccagctcgtgcataactttctggaggc tctgtaaaagctctttgggagtgaggacgtc aacctcgggacgacctgcccattgcaggc ttcgaagaacattggctgggaatgaacg gacgaagctccgattattccccgaagg gaaccaaggagaggtatttccattgaatcaa agcgaacaagaccggaacaggaaacca
452-453	Ga0247608_10100524 GENOME_ACESSION: 176DNAGHGlowlp2_2_-\$F_3300028805 GENOME_ID: 279762 CONTIG_ID: 100517 SOURCE: JGI DATE: 2019-06-13.	atgaaagcgggaacaggtgaaacccgttaaggagatgaaccggatcctgtcaaggag accggcaaggaggtccccgtctatgtgcaggacaaggacgggaatcccctcatgccga cgttccggaacggcaaggtacggaggatgctcaaggatggtctgccgttgcgtgagg ggatgcccccttactatccgcctacattatgagccgaagactcaggtggtccagagggtc accctcgggatcaccgggatacggcactgtcggcttctcggcggtacgatcgtgcg tgagctcatttgcgggtgaagtgctcttcggaacgacgtcgtggagaaggtgtcgaaca gcgcgaattgcgaagaccagacgtccaggaaagcttcgctaccgtgccccaaaggttca acaacaggaaggaagaaggcagcttgagccgtcgtgcgtcgcgtgaggtgcgacg cccactgtcgggtcataaggaccgtgtgctcgttccggttaagaggatcttctcga gatgacgagtttcgacgtccggaagctcaaggaccgggaaggtcgggggaagggtat cagcacggagagaggacggcttctcaacactagggaaatgtctccaccgggacg gacaccgttgcgcaactgcggcggaagtcggcgacaggtatcctcaggtgcatca cctcgagagcaggaaaaccggaggcgacagccccgggaacctgtcaccctgtcag gacctccacacgggataccatcggggacgttcgaacttaagataaagcgtgcggcc cctcgaagctcgcgaccgtgatgaacatgatgaagggaagggttttctcctcactgagaa	ctcaacttctgcaagtgaatcccaagtcactg aaagggtccattggccgcccctgcgaagg ggttagccggaggtccccgcaataatagg aactacgttccggaagaactactgcaaggata ggcactccgggaggtgcgcctagtcccgg acactcggggtcgggtaattgagcgaaga gcgacggtgctccgtccacgaacacattacc gaacatggcgaagggtcgtttgaccttcc ggtcctcgggaaggactatccctgaacag gggagtcgggaacggttctcgttctcggg tcagctccggacaagaatgaaccgaaaaa ctaagagtcgcatgtggacaaaaaac aaaagacgcaatgatgaagcgggaacagg tgaacgccgttaaggagatg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aggcataatccgcacaaggaggtgctcggtccttcgcctaccagacgaagtcgcacag gatagacgagggactcgggaagatcacgcgaacgacgcctactgcataccggcaac ttggggccgaccgctgcccgtctcctacccgggggaagcagatcccgacacacga ggagcctccacgtgcagaagacttcgaaggagggaagaggagaagcacggtagcc ccgcaccggatcgggaagtcgacgtccagagatacgcgtcgtgaagtaccgctggg gaaaaggcggtcatcgccggcagcacgaacggaggccggtcctccgcacatggac tggaaaccacgaaggagcgtcggfgaacgcgaagaaagtgaagttcctgtaccgga gacaagatcatcatcatcatgacgtcccccgcaggagacggaacttaagcgggg cggccacctgtgggacacctggatggatgtcatcgaccgtaacaggagtagtgggatg atcttgcggggaggttcggcggaatgaaaaatcggagcaatga	
454-455	Ga0194060_10004581 GENOME_ACESSION: SepL225m_MetaG_2_\$F _3300021602 GENOME_ID: 35568 CONTIG_ID: 4580 SOURCE: JGI DATE: 2019-01-27.	atggaaatcaatcactgtgagcagaagatgcaaaaatcgaaggggggtctaaacgggt actgaataagatgttgcgtttactatccgcacgtgtgatctcaaggccgaaaactgcgcgt ttccaggcactgcgcgtcaaaactgaccttgaagcaaggaaacaggcattgcgtgtgtg cgtgagacagaaagcagtgccgtctgtactcaactgtttgagttggctcaccgtggcc agcagatcagcgaagccttgacatcaaggcgcgggcatcgagggtacgcaggggcaa aacttcgctaccgttcgccccgtttgataaccgtgccaatcagaagcggctggttgc gccagcctgaacatcgggtcgatcgcacgtgacctgggttaaaccgattgaaagt gacccccatttcagccatcagcacagcaactgttcgatttaatatgcaggcgtgggaat cctgaaattgaaggcgccagatatacaagggaacactgtcagggtacgagatgcgcg agtatctgttgagaaattggggccgactgtgcctattgcgatacaaaaatgtccctta caaattgagcattatccaaaatcacaagggtgctcaaacgcattagcaacctcaca ctgcttgcagtcgtcgaacacaaaaaatcagctttaccaattgaagtgttttagccaaa caaccagagcgtctaaagcatatcaaggcagcccaagcggccactcaaggtcgtg cggcgctcaattcaaccgatgggctgtgtcaatgcacttaaaacgcacgaactggcg attgaaaccgcacaggtggcagaacaaaattcaatcgcacatcattttgattccaaaaa ctcatgctgtgatccgctgtgtgttggaagttggatcaatcacagattggcaaaagcc gacctgtga	gtcaatcctccctgcctaaaggcaggggcgtt ggagtcgaaagacaacgagtttaggtttaa cagggaagcggttaaccaaccgctacgtt aataacatgtcgtcaagacgcaccagcaaat gcttccccagcgtcgcagacaggtcag ttgctgctgcaaaaactgaatcatgaaa ccaaggtaaatgtcgaattgttttagtcgc gactcgaagggtgggagcgggttactgac attccgatggagatgagctgcaaggctta cgtcactaggccgtaaggcgcaaaaggatt gtttaggca
456-457	Ga0209056_10004009 GENOME_ACESSION: GraSoiAngeSPAdes_47_ \$F_3300026538 GENOME_ID: 26996 CONTIG_ID: 4008 SOURCE: JGI DATE: 2019-01-27.	gtgaacgtgtttgtattagcgcctgacaaaaagccactcatccgactacgcccgcgagg gctagggtgtggctcaagcagaagcgcgcgaattgtcaatcgcacaccatttactatc caactgcgtttcgaaccagcagtggtatcgccagcatgtgaagggtgggagtggtatc aggttcaagactgtcggaggtgcagctctgtcactctcccaagttctctatcagcgagag attagcttgcgtacggacattaaacgaaggttggaccaacgaagacaatatcggcgcaat cgtcgaagccgcaagacacgctaccgccaggttcgatttaacaatcgaagaaagccga gaggttgggtcccccacgtcgtgctctaaagccaaggcgacgttaacggcggtcgtg caggtggctcgcacatctaccagtaagcagtcaggggtgagggtgctatgttgcatac cagaagatgcatttctgaatttcggggagttgagttaccacaacgggagagctgatgggc taccatgtgcgcgaatacctgttggcgaagtgggggagggaagtgtgcctattgcagtagg gagggcatccccctcaggtcgcagcattatccttagaattcagaggcgccaccaatgc cgtgtccaaccttacttggctgtgagacctgttaaccaggcggaaggggaatgcactgct gaagagtcggctaccagatatacaagcacatgctctgaacctttgagagatgctgtac aggtttccatcataaagccatggataatgggacctgtcaagactattgggtcaagaaaa tgtttcaccacctatgctatgagaccaagtacaagcgaatgaaatgttgaagctgcct aagacgcactattttgatcggttgcattgcctcgcagcttggggaatagttatgccag gcacggtatggtatcaattcaagtgcgtccacgcggtagtatcagctctataacgggag tcggagcgagcacagagtgtaggtccgaagaaagtatttggctggaagctgtttgagct tgtcaaggctaatggtcaagtagctacatcagttggagtagggttaagtgccgcttctc agtgaagacgcgattatggcaagctattgtggatggcatagggccacaagaagatca tccgattagcaagagctacgcacgcgttattgtacacgttaattttagttagtga	gtcaacgaccatcgtgaaatgtgtgtgcct ggagagcaacctccaggccgtgttgacca gactacgcctttgagggcactcctgtgtgc aggtttacatcaggtgttcggaaaaggct acgattgcgctaagaggtttagttcgtacct ggagtagtgcctcagctccaggctctaac tcttgattaaactgtgtgaggtcagcgac agtgtccagagaaagtagccgggtgcaatc attgtcagggggaagcctgtactccgaaa ggagctttacagct
458-459	Streptomyces sp. NL15- 2K DNA, Scaffold95, whole genome shotgun sequence GENOME_ACESSION: GCA_003851625.1_AS M385162v1_genomic GENOME_ID: 249446 CONTIG_ID: 94 SOURCE: NCBI Prokaryotes DATE: 2019-05-30.	atggcgaggggagaccgacggcagcgaccgagcccgctcgcgtcacgccaa gcacgttgcacggcagcgaggcgccagagaacggggggggtgacgctccacgc acaatccggctacgcggcggtgtcggcgcaagcagggtctctcctctcaagga cgggaagccactctggaaccggcctcgcctcaccgacgataagaaggaagtcgacg ggcaaggagttgtatcaccgtcagggcggggtgtgtcgttgaactccagcaccgc ggtgaacagatccgctgtgcatcgccgacgcgcggggtaccggcacagcgccgc tccgcaaccgcccgtaccgagcaccgcccggaacacggcccccgtccagcagga tggctaccgcccctccacgccaccgtgtcgtacacgcctactcttggcgtctcgcctct gccgtacgcaccgggtcactgaattcatgtggaacaggttgccttcgacgtccactcat gagtacggcagggccctcgcgggggtggagtagcagcagggaactctggccgggat cgaatcccgcgctactccacccaaggtggaacagcgctcgttactcgcacgcc acggcgctgcccctgaacgtcgcagcaccctcagaccccgagccgagggggctcaaatc gcatttccaaactcgtcgtacgtcgttccctgcaacaggccaaggaacaatgcccc gtcgaggtcttctcgcgacggcccccgcgcctcgcgcctcgcgcctcgcga agatcctccgcaagccagacaccgctcaacgacgctaccgcgatgaacgcgacct ctgacagctgttaaaggcgctggggaccctcgcgacacagtgatcctctgctgggtg	gtttcggtaccggacgtcggaggtcagtc gctccccgaccgcaaggttgcggctgacc agccccagtcacgagatcgaagatcaag gagggtccctagatgactactgttcccgacg gcgagcagaccacagaggctgtgtctctc agcagcctgcttgaatcgttgggagcag acaccctgggagaaggacgaacggct cagggacaccccccacggcgggggca ccggcagggaatggcgaggggagac cgacgggcacggcaccgagggcgtcgg cgtcacgcaagcaggttccacggcagc ggaggcgcgagagaacggggcggtgac gcttccagcgaactccggctacggcg ggtgtcggcgcaagcagg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		cgcgcacaatgtgggaaccgcgaggccatggggctcgacaagacgcacacgatggccttccgggtccgaatga	
460-461	Ga0376465_0015542 GENOME_ACESSION: Beaver4Cecum1_FD_\$F_3300035510 GENOME_ID: 312344 CONTIG_ID: 15541 SOURCE: JGI DATE: 2020-04-07.	atgagagtattctgactgaacatgcgtggcggaacattgatgccatgcacacagagaaaa gcccgtatccttctaaaagaagggaagcgtgaatacaccgatccgttcacgata cagcttacttatgctacaggtgaacaaaaacaggactgtcatatcgggtctggacaccgga agcaagcatatcggaacttgcggtgaacatccggaaacaagtttgcacaaaggagaagtc gaactcgcacagagatgtaaaaccgaagattgatgcacgcagatgggtaccgcagggaaca ggcgggaaccgcaagaccagatacagaaaagcaagatttctaaccgcaagaaatctaag aaatggctgcccgagatccagaacagggtcaaccatacaccattggatcgccgt gctgcagagtcttgcgggaagcagaactccacgtagaagtcggcaagtttgatactgc taagatgataaatccggacatcaatggatggattaccagcacgggcagacctacggtt ctatgatgaacggctattgtgtcgaagggtgattacacctgtcaggtatcggttaaat ccagtggcaagatctgcacagccaccatactgttaccggagcaatggcggcaccgga cagggtggataacctataaccgtatgtaccgactgccacacctcagaaaccacaaga aggcgggaatcctttataagtgccaggaggaacacagaaggtaaagcagtataagga accgccgttatgaacacactccgc	gtcaaccaccccgacatgctaatgcctt agaggctcggagctgtaaaagctcagattga ctagcctaagtccttcgggactcgtgttt atgttatcacacctacgaatgattacctagtgc gtagcaactgtgcaggcactgtaaaagtct gtgagcgaggaaacggtaacctgttattgc cgattacggaaagcatttacaacattggcga aggtaaaacaaacttcatagaaaggggcg gcacctcgggtgaccgccaaaggttaaca agtaagtga
462-463	Ga0373622_0034557 GENOME_ACESSION: YL18G_05S_MG_FD_\$F_3300036766 GENOME_ID: 312747 CONTIG_ID: 34556 SOURCE: JGI DATE: 2020-04-07.	atgaaaaagcagacgttaaaacagagagcgattaaatgaacacacctacggatgcttc acaagtcgtagcaactgtgtttgtcattaaacagagatgaaagtctcagttggcgagc tcaaaaacctgttttaacaatcccgaagtgaatcaactcaatcacggcgaggaggactta aagtaacacgtgcctgtactttcctaagggggaacccgcttatgccctgctctccagcgaa agccccgaaaactattaaaacagggaaggcggaagtggtcaagcgaatacatttattat caaaactcaatttgcgtgtgtaaaacaaaacgcagagaattacttccggacttgattcggtat acgccaatatggattctctgcaatcacttctaagaagagctcatctctggaagcgtatgt attgacaacaaaacatccgaacgaattaccgacaagagatgtatgcaggaatcgag gaggaggttgcgtatctgcagcgtaggtacttgaacagaaagattctaagaagcaaat tccgccatcagtacaacgaaggattgatacgcacattcaattagtagaagatggcgaa atggattcctataacaaaagtaaatattgagattgctaatttcgatattcaaaaatcatcaat caaaatataaaggcaagcagttatgcagggaatttatggatagcaaaacttaaaag cgtatcatagcgcgagaatatagcaagtgctcagttatgtaaaatgggatggataagca aggatggcatctacacatattatcgaagagcaaatggcggaacgaataggtcagaca acatcgacttacttcaaaaagatgtcacaagaagtacatacaaaaggcctaaactaa gcctaatagtcaatttaaggctgagacatttatgtcaatagcaaaatggcgattgtaata aattaaaattaaaattaccacaaaaacaatttggatatgaaactaaatcagaagaat gaattagaattagacaagtcgcataattatgatgcgtttgtaattgcaaatggaggcaacca aaacagatgcctaccaataatcatcactcaaaagcacataaataatgcctgcgttgattac agaaaaacggatttcccttgcacccagaagatgtagatataaaattcagccaaatgattta atttgggtggacggaaaaaagaaccgcaagggtatgttttaatacggagaatgggtt gtattaaagaatgacttattcgaagaaaaggaaatctgtaagcataaataagaataaagg tgtataatttgggaagtattattatagtaattaa	gtcaatctccagcaatgggctgagccgtg aggcgcaaggcgcaacaggttgattagggtg gcttcaaaatcgaaagatgaaaagcagac gttaaaacagagacggaattaaatgaacac acctacggatgcttcacaagtccttagcaac tgtgtttgtcattaaacagagatgaaagtct cagttgtgcgagctcaaaacactgttttaac aatcccgaagtgaatcaactcaatcacgg gcaggaggacttaaaagtaaac
464-465	human gut metagenome genome assembly, contig: NODE_399_length_6336 5_cov_4.825295, whole genome shotgun sequence GENOME_ACESSION: OQMP01.1 GENOME_ID: 9944 CONTIG_ID: 398 SOURCE: WGS DATE: 2019-01-27.	atggtttatgttttctgactggaccgggagcgggaaccgcctcatgccaccaccgctgcg ggaaggtccgcaggatgctaaaaaacggcgaggctgaggtgttttccggatcccgctc acaatccgcctctgctatgagccggcttctaaaggaaacccagcgctgtatagcgctgc gaccgggggctgacaaatattcggagcgccgtgttcagggaacacggctgctgtgtcta cctggcaaatgtgccacaagggaacgggaatcccaacagctgatggcaagcgccg gcagcaccggcaggttcccgccgcccggcggaacggctgtcaaggaagggctggcaaa acggctcgggaccaccacaagaactgtcggacaggtcctcctccgggatagagaag ccggtccgggtgaaggatattcatcaatcaggaagccgggttaacaaccgtttccggccg aagggtcgtgacgctacggcaaggcagcttttaaggaccaccctgaacattctgaa gaagggtccggaagatcctgccgataacggatgtgtgctggaagccaacaggtttgcatt catgcaactggacaaccgcatatcttccgggtggcagttccagttacggccccctgcatgg gaaaggaggtgtggaaaatccgtaaaggagcagcagggcggaacctgtatcttctgta aacatgagatcagcattaccacacatcatccccggagccggggaggaagcgatac ccttcaaatatggcgggttatgcaatgcatgccacgataaagtcataagtcggaaga atgggtccggaagctgaagaaaaaagaaagccgggctgaataaagaatacggggccctg tccgtgctgaaccagatcatccgtacctgggtggacagttacacggaactgttccggac catacctatgtgacagccgggtacagcacgaagcagttccgggaagaccacggcatag aaaaagaccatgacagcgaccccgctgcatcgttcgacattctggagaaggtgga caggatcataatccgggaatgacataccagatggaacagttccggaggcatgaccgtg ccaagataaaatcttccggaacaggtattattacctggggaagaaaggtggcggtca accggaaaaaggcaattatggcagacccttccgggaagagaaaaacgggaaggttca gaccaggactcgtggaggactgtttaccaagaactgagagaacggcgtgcag gaagcggagagaaagcgttccggcctgaggcggtgaaagcatccggatccggaaac aacatgcagaggccctgctcctgggagatctttatctttgagggaagaggtatctcctga caggcaatcatgggaattattatcagacaaaagtgccaggaaacagtagaattcctg	gtcaataactcatgactgaagtcacgagctt gttgatgatgcgcgtcatccatcgga cagcttctggctgatccgggttattgagca gagcagtgatacggcatcactccggggc gtactcaaacctcggacactgtctataggcg tacctatgctatggaacacttactgccccacg gggcaggcttaccgcaaggagtaacttat gggttat

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aagtcaaaatgtaggatagcctcggggaaccaggggctgggtatataccggacaggtaaa aagcggcgcttctcccatgtctaa	
466-467	Ga0373620_0052976 GENOME_ACESSION: YL18G_03S_MG_FD_\$ F_3300036760 GENOME_ID: 312746 CONTIG_ID: 52975 SOURCE: JGI DATE: 2020-04-07.	atgaagactcgaaggctcaactgtctggggcgtggtagccgaaagcacgcagggg gatgctcgccctagtcctccctctcgacagtcagtgccgaaggtgaatacaaccaag gcccgtaaaggcatttctgctagcgttctgtctgtagatcgagatcagagccctctgatgc cttgcctcgtagctcgtgcttcaagatgattcggagtcggaaggcaacgcccctctcaaa gaagggaactcttctcgtcagcgtgaacgtcagccttcagcccgccgactcagccga tcgctgtggggatcgatccagggtcgaagaaggaaagccttctcagtcgtctcagcgaag tacaccttctgaacctccaggcagatgccaggacagggatcaaggagaagatcgaga accggcgaaagcctacgaagatcgctagaaagtaggaagacgcccgtcgggcacgccc gcagcaaccgatcttactcgaaataagggtgttccccgagtagcttggcgcggtgg gattggaagctgcgctcctggagggtctagctaaagctctatccggtcaccgacgcgatc gtagaagacatcgctgcggccgacgaagaagtagcagagaaggtgaacagggtccttct cgccctcgtgaagttggcaaggctgtgttctacgcagagggttcgaagattgcccctca gacgcgcaagggttacgagacgaaggctctgcgagagcgctacgggctcccgaagct caagaacaagatgtccagcgacttctatcccactgcgtcgacgcttgggtgtagccgc agaagctgtaggcgcagcggctccgacggagaagcacgctctcgtacacgcccgtc gacttgcagggcggtcacttactaccaagtagctccgaggcaggtgaagcggtgggtc gcatggagcagcggaagcttaggaagccgtaggggaaagcctctacgatctccaaag tatggcctgtatatttgggtgatccaaagccggatttcttgcagacctagagac tgggaagagggtgaccaagagcaggatcgacaaagtgcaggatcctgagtcataact cagtcagggttttctccccgctgagcacgggtcgtcgtcgggcagcagagggtttgtct ga	atgtccgaggcgaccagcacgtccatctgg cggatgtgttgagacgtgctcctggttgc cggcataggcagctccatgtccccaggcg aggcatgggaccggaagtagctataacagc cagatcggtgcgcgacagacagccact gctggcgagacagttcgggttagcgaag ctcagcgagagctgacacgctggggcg ctcgaagagctgcccgtgcgccccatccc cttgggttggatccgacatagctaaccgtg tacgcttatagca
468-469	Ga0114925_10000117 GENOME_ACESSION: SumtaG_3_\$F_33000094 88 GENOME_ID: 35784 CONTIG_ID: 116 SOURCE: JGI DATE: 2019-01-27.	atgcctttagtftcaagaaaagctagactgttacttaacaaggtagagctaagggtgttag taagtgttcttactattcaactactatttggaaagttcaggtatataaacaagaagtaaaagctt ccttaattccaagtagtftcaagtaggagtagtctgttcttctttaggtaaatgtttatattctt tgaagtagaattaaagacaagatatttcaagaaaatgaaggagagctgtctatagaag aacccgtagaataagaagactcgttatagaaaatcaagatttcttaataagaataatctgata gaaagtttactccaactatgaggtctaaacttgcagagtcagcgagagaggttaaaagAAC aactaaactacttctatcagtagtgggtatttgaagaactctattagaagaattataga gggtctaaagaattagaatgttgaatctacaaagcaaaccttttggagagagatagattttaa tgtagtattgtagaggtaaatctaaatgctatgaacttcatgctcatcatttgatactaaaga gtgaagtaggtgaagatactttagaaaatctgttaactcttgaagactgtctatagcttat cataaagggtgaaattgagttaaagaataaaagtaagggaagaaagctaaaattataactg aacttaatatattagaagaattttagaacttccaagtagtatttagtaaaattacggcttga gtcaaaagctaaaagaaaaggttagatctggaacctacacctattacaatgctttagtgt cttagagatactaccaataatagttattattataaaatgttctaaaggagattatcaaaaga acaaaaggaggttaggtgaaaaattattacataaaggtaaaatattagaattttctaagttt gataaaggattatattaaaaaatactattttttaaaggtagaatgagcactgggttattcat aggaaatgaatatttttagtaaaagcttgaaggttaagactttaaagctaaagaatgcgaa ctaatttag	gtgaattcccatagtggtttcacaactgtg ggcttcaatcgtgagattggagagtaattgt tcaccagactcagttggagaataattctaacta cggtattcttgaagtagaaacctactaatgtc cctcagttagtagcttcttgcgtctgttaa cagagattaaagctcagtcacacacaatgc gaagcaagataaacattgctgagaggaggt agaattccttaactgtgaacaggttgaagata ctcaactacttaggaataggagataatactaat gtta
470-471	Ga0370498_000007 GENOME_ACESSION: Frozen_pond_05D_6_\$ F_3300034155 GENOME_ID: 280199 CONTIG_ID: 6 SOURCE: JGI DATE: 2019-06-13.	atgccaaatttagtacataaacaaggttaaatcaatgactagatccgggtcttctcgtcgt aagagtaggaagattaaaggaactgaatcatctcctgtggcaaggaaaataatatttctaa tctttctagttggaaatccctatgatgacgtctcgaagagagcaattatccaaattatgc aaaggtaatttggactggaaggttaacttcaaaattgttacgaaaatatagttccagtaattt ctgtactggagttccattaatcccttgcctccaaggagagcaaaagaaatgaatgagaaaa tataaagcgaagaaagcagtgagatagttatttctatattaaattactagaagagatata ggaaatattcaagagatcatgtggaatagattctggaatcaagagagaagcgataacg gtaaaatccagaataaaacttttattaatgttttagcagatgctgaacaggagtaaaaga atctcttgaaggtagaaaaaatagagaaggcgcaagaagatttaaaaaaactcctgtaga aagaataaattaaatagaaaaaggaataaaaactttattcccatcaacaaaggcgagat ggaattccaaggttaaggttgataaataatataaagaaaatataccaatcagatatagttgt agaagatataaaggcaaaaactttaaagggaagaaaattggaataagaatttttccaca ttagaagtaggaaagaataatttctataatattaaaaacatatggaaatcttatttgaag aaggatgggaacatttacaagaagacagaattagggttagtaaaacaaactaagaatt agataaaatatttctgcgcataatgttgatgttgggtttggctaatttccctttaaattacaa acatactcgtaaaatattgatatgtatttcttcaacaattgaattgcatagaagacaattaca catgttacagtttgcgaaggagggaaaaggaaaagatatgggaacagtaiccttgg aataccaagggaacagtagtaatagcataataaaaaagaaagaaatatagtttatatt gggtggaatatgaatgttaattatcaatccataatgcagaatggagaagaataagta aatgtataataaaagaagatataaaatataatgggatatagcgaagtggaagtagaaaa aataacagaaaaaacgagcattcgtggaagacatgcataattaa	ctgtcataacccccatcaagtttcatcttgat cgaggcttggtagtcattaatgccaaattt agtacataaacaaggttaaatcaatgactaga gtatccggcttctcgtcgtgaagagtagga agattaaaggaaactgaatcatcttctgtggc aaggaaaaatatttctaatgtttctcagttg gaaatccctatgatcatcgtcgaagaga gcaattatccaaattatcgaaaggtaattt
472-473	Ga0247608_10002058 GENOME_ACESSION: 176DNAGHGlowgp2_2	atgggtacgttttagacatagaggcgcaagccgtgtatgccactgagaggcacggaaa gggtcagaagacttctcgtgacagcaaggtcgtatgctgaagactgcagccattcacaatt cagttggattatgagagtaccctacaagcaagaagtagtttaggcattgacgcaggc	gggtataagattcaagtttgcgaagtagtgg ttagccagtcagtcaggcgagcgagtagg cctaagtccttgattggcaaaagggaactac

SEQ ID NO:	Name	IscB sequence	ωRNAs
	\$F_3300028805 GENOME_ID: 279762 CONTIG_ID: 2057 SOURCE: JGI DATE: 2019-06-13.	agcgtacataattggagtgctgcacaacagagaagaagagttgttcgctcggaggtgttcttcggacggacatagtaaaacttgcgagccgtctgaaatgcctgcacaaaggcgtaacgcaagactcgttatcgttaagccgaggttcgacaacagacgaagaagggaaggttgcttcgaccgagtagataaagaacaaagtgtagccacataaagggttatccgtttggtcattattgtcttcgttacaaagaccactatcgaagttgtcagtttgatgcgcagaagataaagaacgatgccattcaaggtgtggagtagcagcagggaacagatgggtttctggaatgtgagggagtagctctcgaagagaccatcatacttgcagcattgcaagggggaagagcggtagataattatcgaatgttcatttggaaagccgaaagactgggtgtaatgcacctaacaacctataacactatgcgagacctgtcataaggcgtaccatcgtggagagtttgagttaaaagtgaaagcgtggcacttcctgctgacgcagccgttatgaatacgcgggtggcgggtgtatgaacaggcaaaagcagagtttgggaacgtccatctcacatacggctatgtgactaagcatacagctattaaagaacggaattgaaagacgcattgcgcggacgcttctgtatcagcaagaatgtaaaagcagtgctgttaggttcgtatctaaatgccgttcccttgcgaagacatacaagaacatgcatgtatgcagtcgaaagggcgggtattaggcgtagcgcagtagcaagccactggataggcaagtcacgtctacaaggtacgatagtgttgagtggaacggaagtgaggtgctttatcttcggcagtagacacgcgagaccagttatgctgacattgacggaaatctatcactccaaagcgtcaattaatgccaaagaatgattttaacatagaacaat aagattatgatgcgaagaattactctgcgaacttga	gttagagtggaatgatataggcacttttcgatggctgtccaaagctgaaacctcgggcaagtgattaaagagcgaaagcagcgggtgtgctgttaaaaccacctataacattggcgatggcgcataacagcagaatgctgacttacaacaaagttgtttaacaattaaaaaataaaagaatg
474-475	Ga0401359_0000081 GENOME_ACESSION: Site1_salt_MV_1_FD_\$F_3300037341 GENOME_ID: 312663 CONTIG_ID: 80 SOURCE: JGI DATE: 2020-04-07.	atgctacgtcttccaggacatcttactgagatggcgggaaggactaagaacttaactcgaaggattatccatgttaccagtgccagttttatcaaaatcaggtaaaccttgcagccaacaaagctagtcgagctagacgttgggttggaagaaggaaagccaaagttgacacaa cgtatctgaatgcttcgcggtgcagttgaccgtctctctagagaagacttgcagccgttgcctgttaggattagatccaggaacactctattcgggaattggagtgcaatcgatcgtctac cttgtggatggcacatttgattctgccgttcaaaacgggttaagaccgaatggagctacggcggatcatgcgtagagcaagacgagggcgacgaattatcgcataaattccatacaataat agaaaaatggctgggattagccaaagacagaagaaaaaaggcgtgatcaaacctcagcgacacacgctgttgatgggttacccttagcccggtttgagttcattcgggtgagagtgca ctcgggttcaacaacatgggagctggaaggggtgacgttcaaatcacactgcaccctt tgcgacgattcgtagaccacctgttagccgttaggcaactgcactgtgtgttccgtcaaaa ggtggaaagcggcgcaagtagcgccggaacattaccgcccgaatcaggaagggt gataaggtaattgccgaaaaagccggaaggtttatagaggttggtgtctggagaccc aaaacaaactgtctgttccgattgttactggaaacgaattggacagtttactgcaaaaaa agttcagttgttgacgcaagcacaggattaatcgttgccttcaactggaccgggtcaaa ccttaccgcatcgaacgggtcgggttagacctggagttccctccaagttcctgcgtgaact ggagttctacgaggtttacgata	gttcacgacctgacttactcgttgaagtcagggttgaaggcaaaattacaatttaaaccttgaataccgcattgagctcagtagtgcatacgaacttccgaatactccctcgttcggagcattctaaagacttttgcgagtcgctagtagaga cgtagcatgctacgtcttccaggacatcttactagatgctggcggaaggagactaagaactta actcgaaggattatccatgttaccga
476-477	Ga0181589_10009757 GENOME_ACESSION: 071ome_50_\$F_3300017 964 GENOME_ID: 20511 CONTIG_ID: 9756 SOURCE: JGI DATE: 2019-01-27.	atgcacaaagctttgtagtagataaaaacggcagacctctgatgccttgcagtcacacac gctgcaagcagttcaagaagaaggctgtagtagctctatcgggtgaaaccattcacga ttcaactgatagatgttgaaggtggcgagactaaagatgtcagtagacaaggttgatcgtgg atcaaaagttactggactgtctatgttggaaactttgagaagcaaggcaacgtgttgcctt gggtcgcacaaacttcagcacagaggtcaagaaattgaagtgcccttgagtagcaagcgtt ccttgagggcgagtagacgcaactgtaaaaaacgctatcgggtgcacgatggagattt cgcaatggaatcaaaagcggactagcaccatcactgatgagcagagtcacaacatcac caattgctaaagaagttgatgaatgtgtcctgtgaccaagatcgtcagtgagttagta cgattcgtcgcagaactatgcaatccagagatcagtagcattgagtagatcaacaaggc actcttcaaggctatagggtcaaaagtagtcttttgagaagtggtgtagacagttgtgtgt actgcaacaaaaagatgtgccactccaagttgagcatattcatccaagcctaaggcg gaacaaatcgaataggcaatctaacaatcctatgtgccaaagttatgagaagaagagtg ccagatctatcggagacttcttctggcaaaccaactcctgaagaagatcaaaagctca aatgggtgaagccactgaaggtgtacggcagttataccacagcgtggagattaaagg aagaactagagtcattcggactgccactagagttgagttctggaggaaacaaagttta atcgattacaacagggtttacaagaagaccactggctagatgcagcatgtactgttcca gtggtgagttctgttagcttaaacagaatcaaacactacagatcaaggcgatgggttagag gttccagacaagtgatgaagacagacaagtagtattccctagaggatctgctggcacag tcaagcgtgtcaatggcttccagacagggggtctgtgaaactggatcaacctaaggga aagtagtcaggagcttgattggtctgttagtaggagttcgaagtcgtggaattctgacatt atgactcctctggaaagcagggagcagatggaagaactcaccattattacagcatagtg atggatcagagtagtcttga	caccccccttaggggttggtttataagggaaggttcgttgatccctgacttagcaccact caacgcattccatgctgctgaattgtgtttttt ggctcaaccagactagggtctgtcagggcc cagtagtcacagaatgatataggcaactt tgaatgcttaccagttcaagatagcggctt caagttaagttaggggggtcaggataacca ctctgtgctgttagcaaaaacctgttcatcat tctcaggtgactgatttttaacaacattt cg
478-479	Ga0080699_1005680 GENOME_ACESSION: IMG_3300006137_\$F_3 300006137 GENOME_ID: 305391 CONTIG_ID: 5679 SOURCE: JGI DATE: 2019-11-15.	atgttgaccacggaaacggaatgccatggggctctgccacagcgcgatgaaggg acagatgcccgtagagccacctgtttatgcaggagtagatcacgggagggcgtaa gccagaaaggaaacttatgacagcagtagcagtaataggcaacacaggcaaacgct gatgccgaccagccgtacagggccagggcgttttaaaagtcaggcggtccaagatct atggatacaggccgttcacgatcatgacttgcagggaggatggcgctgtgcaggaa attgagtagacaaagcgacaggggttactcgtatgttgcgtatgttccaaaaaac atgagttcctgcgggaacagaggatctgctccctgatgagcgggagaacataacgac cggaggaataccgccggacgagaaggaaccgcaaacgttataggaaaccgcggtt	gtcaattacccccctgatccagtggtcag atggggcttgcacaaaaaaagagaagatg aagtttttttttttgaagcccggttgatta cccctagctcgggtacgttaccagtttaaat ataggcaccaggggatactccacgtccc ccgcaactgcggcatgtgtttaaactccctg agggaaaggggaagtgattatgcatgaca gcttatcgtataagccgtctaaaacatctggg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		<p>gacaaccgcatcgggaagagccgtaaggcggagaaagagggcgctatgctgcct ccgtcgtggaacacaaggtggatgcacagctccgtctttcaccagcatgcaggatc atgccgataagcagcggcgattgagatgggaaagtggaccggctgtttaaaagcg gaagagagcgggtggccagttcccaagcggtgattaccagcatggcgggcggtacc aggccggcacaataaggcagctgttttccgggagacacacatgcctgttctgtg gcagggggataaaggatggtgttcttccacgtacacatcgggtattggaagaaag accgttcgaaccgtctggggaacctggcgacatgctcgggatgtgccacacatggaa aaccacaagcctggcggaatcctttacgggaagcagccggaagtatccggactggccc ctgccacatacatgaatacggtagctttgaaactttgagaaggctgaagggaagcgccc cggtctgtgatacctatctctacgggcaaggaccagcatggtccgcaagggaacgg gcatcataaagagccatacgaatgacgttactgctgggaaggcttctccgaagcac aggcatcagaggaagtgtccagaagaccggcgcaacgggatactccagaagt tctatgatcggtatattgacggccgcacaggagaggaagctaaggccaggaactt acgaacggcaggattagcagaatacataaagaccacaaaatcgcacatccg gagcaggaaagtcagaaaggcggtttacgacgtccggagatccaggacagcggtaaa gcctggcagctgtgtggaattcaatgacgaagtgttactgtgcatggcacacatacagc cggcataaatacaaaagacagcgaacacggctgtcagatcatatagagtttaaacag ccagccaggaatggcaaaagagcgcggaattgagcaaatccggatcattaaacag catataatactggctgaaaaagattactgctcggctttaaataag</p>	<p>aacattgggtatgtgaccacggaaaggaa tgcgcatgggctctgccccagacggc atgaaggagacagatgccgtgaggccacc ctgctttatgcaggatggatcacggggag ggcgtaagcccagaaggaa</p>
480-481	Fischerella sp. PCC 9431 genomic scaffold Fis9431DRAFT_Scaffold 1.2, whole genome shotgun sequence GENOME_ACESSION: GCA_000447295.1_AS M44729v1_genomic GENOME_ID: 52896 CONTIG_ID: 0 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	<p>atgtccaaagtattgttttagtgcagaaaaagaccactcttacaattatccagcacag gcaaggcaactattacgaacaaaaagcagcagatattgacaaatccattcacaatta tttgaaggagctagtcagatgcacctaacaagattacgaattaagattgaccccg tgcaaaatatacaggaattgcgtagtcacagatattcggagaagtgttatttgcagcag cagtcacgctactagattgtcttacttaaaagtgttgaatcaacagggttaccagtgatg tggttcaggagggctaacaagtcaatcgaatacaacagaatttagcgaagcctattgg ctggatgctgctgtgttgggaagtcacacggatcttaatacaaggtattaaacattg ttgattacagccaatggcgatggtactgcagctatgccgaacggataaatttggatttcc gagcagatatgtccctagatttaattgtcaaaagtgttcaaacaggagatattgttaaagct attgtcaccaacggtaagaagattgtgtgtatgtagggcgtgtagctgtcgttctacgg gtagctttaacattcagctaaacaaggattgattcaaggaaatcagttcaaatattgtttacc cattcataaaaaggacggttacgcatatgcgttggatttgacggcgaatga</p>	<p>gtcataaagcccagccctaaggcggggg cttgaagaagccttatatgaccagcttaagt cttaattgactacgtttaaggcaaggttaaa gacctaccagggaatgcgtagctagttccct gctctagaaccaactgtttaacagatgtac aagggttaagtcagtgcaattgtgatgtac cgaccttaaacattagcgaagctaacattac ccgcaaggaggagacttatgtccaaa</p>
482-483	Ga0310136_005546 GENOME_ACESSION: M271_2_SF_3300034520 GENOME_ID: 292992 CONTIG_ID: 5545 SOURCE: JGI DATE: 2019-08-09.	<p>atgttgtgttcactgttgacaaatcgggcaaacctggcgatccaactagaagggttgacat gataaagaagcttaaaaagcaggcgagggtgaagatcgtcggtgtgtggcgcttccggc aaaccgccgggtgtgtgtcttctggatagggaattcgattattcaaaaactgcagaagaa agctgtcatagcgtctgacccgggtgtatagatatactggcttggggctcgcgaaccaa aagcgggaactaagctgtactgcaaaagtgttcttgaacaggagataccggaatca aagggttaattgacagaagaaggtgacatcgaggagatttagacgttactgttccgccata agaaaaggcgtttatccaagcggcgaaggcagaagtcaacaaaattcaagcgccgag aaatgtaaggcgaagaacagcaggaatgctacactcaagcatggcgtagagacacat atcaatcttggcgaggcttttgaagtgttccatttcaaacatcaggtcgtttttgtcatg gaggacaacgtttttgacgtcagggaatgacgtggggcaagacatagcgtgacaggata tcagaatcgccgagaacagaagtgaagaagcagatgtgtaatctgcgggttcgacggaga acttgcacaaagcaccacataacacagcgtaaagacggaggaacggacattgacgaat ctggtttacctctgcaggagctgtcatgaagtgtgcacggcgggcggtatataattccca taaaaggcatgaagcagtgccgctattgggaacgatgaacgccataatagcggaact gcggaagataccctgtttagaattatccctgcccggatcgcgcaaaaggcgagaaaa acagctgggcttgaagaaggacatgggaacgacgattggcaacagcgcggttact gcaatcccgtgaaattgacacacacaatcaatggaactgcatctgtttaagtaagaag gcacagcaggggcggtatcatgctgtgagagcagggtatacaaaatgaatgggaaa attgtagccaggaatagacagaaaagaaccgatcaaatggagccgttcttggctgatgtg ctgcttttaccacgcgcagcagaggatctgaaggtgtatccaggagtcgaagtgtc aaaccgttttaggagagacatgccgtccgtggaaggcagtggtgggtcattttagcgaca ggaaagcgttttatagtgaccagcgtaatatcgaagaattacctgtatttcccgatgag agagattgtgggaaaccc</p>	<p>gtcaactaccctgattaaaatcaggggctt gccagcaggggcgctgtgtggcggtc tgttaacggcaggttagtgaaacgggtgtg cgcgggcggggaactgacgttccgggacga cactcccagttccgggagacagcggtgaaa tccgcagccttacacagggtgtgtccgca cgcttcggggagaccaccgcttttggcg cccggtacacaggctcctgtaacgaggag ggaaagc</p>
484-485	Ga0063591_100011 GENOME_ACESSION: IMG_3300003966_SF_3 300003966 GENOME_ID: 305235 CONTIG_ID: 10 SOURCE: JGI DATE: 2019-11-15.	<p>gtgttcgtgttgatcgagcggcaagccgctgatccgtgcagcgagaagcgcgcca ggaagctcctggccgcccgtgtgtgcgcgctgcatcggtgtaccggttgcggttcgg ctcgttgaccggcacgtcgaagactgcggttcagccccctgcgcctatcaactgacct ggcagcaaggtcacaggatgctgtgcgcgcatcgagccaaagctgtcggcgag actgatgaatttcgagccgggtgatgcacatcagcgtgttgatggagatcagcaccga ggcgcgagatcaaaaaagacgtgcaaaagtcgcggcgctgcgtgcgcggcgccgc agccgcaactcgggtaccgcgcgccacgtttgacaaacgcgcgttctgacggct ggctggcgccaagcctgcagaccgcggttgacaccacgctgtcgtgggtgcgcggct tctccgcttccgcccgtcagcaccatgcccgaagagctgtgcgttcgacatcagg ccatgcaagctgaagagggaagcaccggcaagcgaatcattcgaagccttgagta</p>	<p>gtcaatcaccacgactgaagtcgggggg tgaagaagcaaaaattcgcagccctggt gaccagaccaagaaggattcgaagaac accaatctacgttcacacaggtcgaaaga ccgacggcaggatgcttccagctcgtccc tgcgaagtcacggaagcagacaagccaa gggggtggcacgaaacggctcgtggcctg gtggcgcatgtcaccaaagccggtgtgcaa catggtcaggagagcgccaagccgcaa</p>

142

143

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aatgtttcagttaccgattactttctcaagtggtggcggaaccaaatttaaccgtgtcaagc taggctacttaaaagaccattggattgacgcagcctgtgtcggcgaaagtgtgtcaagc tgttattcctaaagcgataacacccctaattatcacatctaaggagcgtgggtcaagacag aagtgtcgtatgaatcaatcaggtttccgagaactggccctaaaaacagaagcaggttc aaggggtttcaaccggcgataattgtaaaagccatcgtgactaaagggaagaaagataggc acgtatattggcggtgtgtgtgctgatccgggggaagcttgataattgtgtgtggcgaaa agaaagtgggcagcatttcataataatgtccaattgtctcaaggcgatgtgttacga gtacacataa	
496-497	Ga0223825_11034434 GENOME_ACESSION: 3300021255 GENOME_ID: 241584 CONTIG_ID: 33 SOURCE: JGI_Dark_Harvest DATE: 2019-01-27.	atgctagtatactgacaggataaagagggcacaacctctaattgccaccaacgggtcgg gtgggtgcatattcttaaaagcataagggcggaaggtagtacggcgagagccctttac cattcgaactctttgcgatgtttccatcgcgaagaagtactctgtgcgtgatgttga agcaagcacatcgggatgactgcatcaacagaaaagaaggaactatactctgcgcaagt tgatagacgggatgatgtgtctaatctattgactgccgtgaggaaatgcggaggggggag acgaggaagaagcacaaactggtatagccctgcgaggtggcggaacagggcggaacg aagaaagaaatgcttctcctccgctgtgtgaagcataaggtgattctcacatagcagc aatagagttcgtaaagaagattctgccaagtaagcagattacgtgtggagattggaaagtt gatactcaaaagatacagaatccatcaatcaaaaggagaagatagcaaaaggcgcaat ggaaggatgggagaatttgaatcctatgccaaagtggcgggcggaacaaatgccga gtgtgcggtgcaagtcataaaggacaagtcctgaaggttagaggtgcatcataccgc cgtaggcgggatggcggtagcaattctcagacaatgtgtgtgactttgtgccacgaatgc cacgaatcccatcaccagaagaagaaggtgttgaagtttaagagaccgcctatacacaa aaatgaggcacataatgaattccatgcgaaataatctgattgataaactgtgtcaataattc gaagactccagtagaggttaccctatgggtatgagactcggatggcacggagggaacaca atgtcgagaaaagccacagaaacgattcttctgcatagcagggaatttgaatgcacttctc aatcatataatgtctatcgtctgtatcagagaagaaggcatagggaattctgcatgaca acacgattttgtctccaaagtcaatcaagacaaaagtaagatgaccgctaaagaactaa agtgcggatagtagagacctcaaaagatgagcggaaggtatcgtggaattctctttggg atactgtcaagttcaacggagtatatatagatcgtcagtaaaagggtcagataacaga gtgtgtcttagccgtagcgatgagggggaagacaattgtgaaagccatgagcaagtgcga gttgcctcgcataacgggaacttgtaacagaagagttggagcttga	catttacaacctagaacctgcaagaggcttta tttcaactaatttttgggaatccgcaaggat agttcaactaggagactagaaggaacga aagaaacagaaaacctagtagaagtacag ggaaatgcatagggcaccggcgaaatagttg ccaagtttgccgcactgctggatgtacattaa acagttctgtgtgtagaggaacagtgtgtgca tcgaacaaaggaaactctctgcaacctcccg atgggcaccaaacaccagaaggagtaatct cagctgtgactaccgagaaatgggttaaca aaaaaagtaagacacagaaaaaagaatgct a
498-499	Ga0116592_1000084 GENOME_ACESSION: IMG_3300009305_SF_3 300009305 GENOME_ID: 305874 CONTIG_ID: 83 SOURCE: JGI DATE: 2019-11-15.	atgctcgtattcgttaattatcagcatggtgagcgctcatgcttgcctcaacgcgaaaag caagactattgctgaatgaaggaaaagcggaagattattgtatcagccattcagatcca gttattatattggatccagtggtgctacaacagctgggttaagtaggcacgaaatcgggga gaaacatacggcattgcaattatccgagggccgtgtactgatgaaggtgaaatcga attacggactgatgtgaaaagtctaactgaaacacgtgaaatatacagacgcattcgtaga caacgaaaaacaagataccgaaaagcacgcttcttgaatcggaaaagccaaaagggt gggtgccaccgagtttgaaaagtcgaaacgaaacacttttcgttggaattaatcgttttagta gtttgctaccacccctacacatcgtgaaatcgcgaagttcgaagtcataagctaact attccaagcgtaggcaacgttgcctgcaaatgggtcatagcgacggctattacaatcgc gtattatgtattcgcgagagacaactatatttgcgaagtatgtaaaagaaaaataaata gtaaatgcgcatcatattctacaggacacaaggtggaagcgatcgtgcggataacttg atcaccgtttgcaccgattgccatagtttgacaatcgcggcccgagatctttggcg gtggatggaaggtgggaaaaaactggggcgctacaagaaccacctttacgaacggt tgcgaaaccgcgtttacgagcaatcccgatgcaatctctcatatggagcggaacg acgacctatcgaataatgattgggttagaaaagtcactgtcaacgatgccatcgcgataa cggggataccagctattatgaagataactgacaccacctcatggtgaaacagttccgca agaaaaagcgaagcctgcatgaagcgaccgcagtaaaaggaggaagcagccaaata ctgaagcagacgcaacgcaaaagataccaaggcgctccaaaggattcttttgaacgatg aagtgcgaacacctgtgtgaatcggctacattagtgcttactggaaaaaacgcctg ctacgtcaagtaattaatggccagtatatcgtttccagagaaaacacataaacacaat cgctcaaggagactagtaagagaagccatcaaggcaattggcgctaccaactctcatat aatgaagaagcgcaacagttctacataaaggattaa	gtcaactaccaccactccaactactagct gtcgaagtggggcgttgaagagctaaagt gtctagccatagctttttgactacgttggttc ggatcatcacccctgggataatcctatgct ccctgcaactgtgtcggcgctgtaaaagctct gttgggtcggagcggtcaaccgaattggc gacatttaattgctggaagccttaccacattg gcgaagggtgacaactccaaaaggagg gaatacgtatgctc
500-501	PROU1_101214 GENOME_ACESSION: IMG_3300000530_SF_3 300000530 GENOME_ID: 305015 CONTIG_ID: 147 SOURCE: JGI DATE: 2019-11-15.	atgtcagtttatgttttaagtaaaacggaaaaccgctaattccgactaattcgggtaacgc aagattaatgcttagacaaggcaaggcaaaagtaattaaaagaacgccgtttactattcag ctgacttatagcagtgataactaaagcaggtgtgttttaggcataagaccgatataa gaatgtaggcttatccgcagatcggacaccaagaatattatgctccgaaaccaagttg agaacagacatcgttaagctgttatccgaaaaagcgaatccgcagaggaagacgga acaaactttggtatcgcaaggcaagattcttaaatcgtggtattaagaaggctggcttgcg ccgtcagttcagcatagcgttgatagtcataattagtagtaatttcgttaagtcattttgcc tatatcaagtattatgtgaatccgccctttgatattcaaaagattaataatcctgaataat caggcgttgaataataaaacggagttcgaagaagactctggaacgtcaggggaatagtt atacagggataatcataaagccaaaattcaaaatacagacatacggacaagataatgaat gtcatcacatcgtatcaaggcagacagggggatagacctgataatctgataaaccttat gtgaaactgtccataaagcgtatcacaaggcaagattaggttagagattaagaacataa aggttttaagccgaaactataatgtcaattattacgtcggaataatgttaggttaagg	gtcaaccaccgccgcttgaagaagcggg ggcttgaagagctggctgataaggagat taaattatgtttatcagcagttatttaagaga accgttaaggaacacaccttaagatttgcg caagtcttaagctctgtggattatataaaca gtcttgagagggttaaaagacagtgtgcta cttaaaaaccttaataacaatcccgatgct atctaaccctgtaaaaggaggttaaacatta tgtcagttat

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aactcggttaataccgttaataattacatgcttatcttaccagtcgcaagaatagattag gattagacaagacacatagtaattgacgctttctgtatagcggggcgggtagggcaggaa aaagttatcaatcctattttatcaagttgttcgttaaatgtaaccgttcgtgtttaaagctaa ccttttaaaaggcggaagcgtaaggtaacactatgaagcaggcttccgctccacagg tttgataatgtttatagcggcaaatggagtgctttatcacgggttgaggctatctgctattt tgataattagacttttaacgggagagaaaatcaatgattccgttaaatatagcaaatgaaggt aatagaaacttttaaacacggaggatagctctctccccgccttatag	
502-503	Ga0256407_10017433 GENOME_ACESSION: RumenRJG_04DNA_2_\$ F_3300028886 GENOME_ID: 280890 CONTIG_ID: 17432 SOURCE: JGI DATE: 2019-06-13.	atggcaaatatatacgtttctacaaagacgggcagccttaattgctgtccacactttcgg cagagtcagaaggctcttgaagaaagcggcaagctcgcattgtcgcacatattccgtttact atacagctgacctatgatatcaagaacctgtgattgacgactccttttagggatagacc tggcaggacaataatcggcttatgctcatagacagcaagccaacgtcttttcgcatca gatgtagaacaaggaacaaagcgatagcaaaactgatgctgaaaggaagacagcac gacaggcgtcacgaggggcgaaaggtccgcagacagagagtgccataacggca gacaaaacggcatgcaaaagcatagagagttctggaggatgctcccggtgcaagg agccggtgtgctgcaagatcataaggaataccgaagcccgttcaataacgaaaacgt aaaaaggggtgaactaaccactgcccacggcaaaacacctattacaggcacatctaacct ataaagaagacacaaaactgtgcccctttcaggcatagcatagatcaataaagttc actttgcccgcagtgagaaacctcaataagggaactgggaataccgaaaggcaggctc tttgattcagggtgtttatgacgggtacaccaccggcaggaaggcaagtgccctttt gcgggaagcagatatagaacattatcatcacttaacccaaggcatcaggacggaagc aatactctgtataacatagccggattatgaatggtatgcacagcaagtgcatacagacc aaaaacaaaagggaattactctcaagcatgaaggcatcaagaaaaataccatgca ctgtcagtgataaacagataatgccggaaacttttgggaaatgtgcagaaatccttcca catagctactaccggggaagaaacatgctgacaaggggtgtgcttgccttgaaaaa gaccactatgtagacgcatgtgcatagcagcctctcgataacggaattttcggaaagatt acgatatccccgattcaaaaggcagttatcataatcagcgatctccgctcagcacag agccattatcaaatcacagacagaaagaacatataagctggatggcaagatagtcgcaa agaaccgcaagccccgttctgagcagcaggagaccggcgtgtcagacttaacctgac aaagcaggagatatcaagacttaaaagtaacaaaagcacgcccaggtacataattaaag acagactgatccccggagcggcattctacataacgggagctaccatataatgagcgggtc aattaaactaacggcgactacctgagagcataggagacggcaaaactaattaccaaca aggaaatgtaagatactcaaacacatacaggattagtttatagcataa	gtcaactaccaccgctacgtaacgctaa gaggcgggggctgtgaaaaacaagcaata cacaagtattgtattgagcagagaccaaag ttgagtaagcttcacgcttacaagtccgacat gccacgttccatgggggtgccactccagcc ccatgtattcaagggtatgccgagctcgg gagacactacggaccagcccgggtttgc agccattaaagggttaattcaagaaagga caggattatgcaaat
504-505	Ga0190346_1003230 GENOME_ACESSION: 48721345_MG_2_\$F_33 00021486 GENOME_ID: 20994 CONTIG_ID: 3229 SOURCE: JGI DATE: 2019-01-27.	atgcaagagattatttctggaatatacaaaaagccgttgatgcccgtccatcccgcaa gagcacgcaaaccttaaaaaacggaaaggctgctgtttatcgtcttattcctttcataatta ttatgaagtgcgtgtgattggtgatttgcacctatcgagatcaaatgtatcccggtagc cgcaaacctgggagtcgtttgttgggcaatttgcgtggccagaggtgatctggcg ggcaacctaaaccatcgtggacatcagataaaaagaatcttgattcaagacggcggtga cgtcgttctcgaagaaatagaaaactcgttatcggccagcagcttttaataaccgacagc ggcccaaaagctggttccccctcattgaggtcacgagttgataacgtaaagaattggtat gattaaattgtcgcgattagcaccactcactcaaatgagcttgaacggttaggttcgata ctcagaaaaatacaaaatcccgaatatcgggggtgaaatcaacaaggcgagcttgcgt gctatgagttgcgggaattttctagaaaaatgggagtcgtgaatgtgcctattgtggtgct aaaaatgtaccgttgaattgagcacattcaagctctgtagcaagggaggtcagatcga gtatccaaacctaaccttagcttgcactcattgaacaaaaaaaggcaacaaggacatta aggaaatttcaaaacgcaaacctaaagcgttgaagaaattcaagcacaagcgaaggtac ccttaaaagacgctcggcagtcattgctactcgtacgctactggcaattgttcaaaaga tttcggtttaccatcatttcaactgttggggaacaaatttaaccgtacaacgctgcg gctatccaaaagccattgtagacgcccgttctgtaggggaaagtgtgtgctaaagtgcg ttatccccccagcgataacgctttgattatcactgccaaggacgtggctcaaggcaaaa gtgcagtatgaatcaatgatttccgagaaccagccccaaaaacacaagcgagtaa aagggtttcaaacgggtgacattgtaaaagggtcgtg	aactacagactcttataagtaaccgtccataa ataacttaggcaactaagttatttaaggttacc cgttttagttcttcggaactacgttagggagc aaaatatagtactctgtgagcttccagct ccgagattctacggggagtcattaaacaagt ctgatggaaaggctagtgtggtccccattca aaaccgtctataacattagctaggagacca ttaccagtgcacaactgagttagtgacaattt atgcaaga
506-507	Moorea sp. SIO4A5 4A5_NODE_11, whole genome shotgun sequence GENOME_ACESSION: GCA_010672005.1_AS M1067200v1_genomic GENOME_ID: 325451 CONTIG_ID: 12 SOURCE: NCBI_Prokaryotes DATE: 2020-04-22.	atgcagaattacgtattcgttattgacacaaacgaaccattaaacctattccacaaa gaaagctcggcggttattaaacaaggtgaaggctgcccgttttaggtatgccgttcacac atcatcttaagactgcgatataataccaacctctacattgtcaataaagattgacct ggtagtaaggtactgatttgccttagtccaaaacaaccaagttatttgggaatggaatt agagcacagaggaggttaattagaaaaactagagctagaagcgtctgaaggcgtta ggagacgtaacgcgaacaccgcctacaggaacccagattccttaaccgtgaagcgtcca gagggtatggctccacctagtctagaacacaggattttgactattcagacttgggtaaaac gattgattaaatctgccagtcattgagatttggctgaagggttaagtttgacacccaa aaaatgcaaaatcctgaatcagtggtcattcagtagcgaaggagattgacgggat gagggttagagagtacttacttgaaaaatggggaaggaatgcacttactgtgtgaagcaat ccgttccattgcaaatcgaacacattcaccgaaggtcacttgggtgaagcgtatcgctaa gtaactttttgttcttgaaaaagttaatacaacgcaaggtaacaagcctatagaagac ttcttaaaaaaaggaagcctactgcaaaaatcaaatctaaagctaaagcagccattaa aagatgcacacagcagtaaatgcaactcggaaacaggttagtgaaggcttcaattaatcaa	cgttaagcctggcaattccggtaaacctct gggattgcctagcgaagatgcagactaagg attagtttctacgttgcgtggcaggtatcc cgagggtgacctccagctcccggctctatc gctaactattagaaaaaggcaaatgtgtg ttagcatgacaagcattccaactgtcgtg gaacacattaccgaattgagttgcactaatca tcgagaattac

SEQ ID NO:	Name	IscB sequence	ωRNAs
		<p>ggtftgtgtcaccggaacgggagcgcaaaccaatacaaccggactagattagaactac ctaagcagcactggattgatccgctgtgttggggatattgagacctagtgttgagaacc tctcagccgctgttagtacttgaaggaccggaggagacagaaagcagcacttaa caaatcagggtaccccatcagacacatccattaaaacaaatcaaggctgggttactgg ggacatagcccagcatcctctactaggaataggcaaaatcaccctagaaacaaaggaa gcttggatttaccgggttaggaaccaagggtacaaaagtgcacacccaagataatc ggcaataaccgaaaagatgatacacttatgtcttctag</p>	
508-509	<p>Ga0207747_1000200 GENOME_ACESSION: HigsoISP1DSPAdes_5F_3300025534 GENOME_ID: 27183 CONTIG_ID: 199 SOURCE: JGI DATE: 2019-01-27.</p>	<p>atgcctacaacccaagtaaacgcgaaatattataaacaggtaaggcaaaagtagta aaaagagaaccatttacaattcaattattatcgaactggtgaaccaagcaggatataa cattaggtatagatgcaggaaatatttggatgatctgctacaacagaaaagcaaga gctatttccagcagaagtagaataaagaacgatatgtagacttactatcactagtag aattcagacgtacaagaaggaatagattaaaggtatagaaaaccagatttttaaatcgtgt tatagtaaaaataaagggtgttagcacttctgtgaaaataaatacaaacatttataa ataacaagtaagatacagaaatfttgcctatatcaaaaataataagtaagttgcttctt atatacaaaagattaaaatcccagatatgaagggtgaacaataccaacaggagaacaat tgggatttggaaatgaagaatattgcttggagagatggatataatgtcaaggtaaa aagggttgcacaaaataaattataatgtacatcatatagaaatcaaggaaactggtggaa attcccaagcaaccttattacttgtgaaagttgccataaagattatcatgctgtaaaat aaaattaaattttaaaccgcgtcaatcgtttaaagatgcagcttttatgggattatgagg gacatttataatagactcaaaataatattctaatgtataaataacttatggatatattacta aaaacactagataacagaaatcaccacaaagacatagattgatgcattatgtataagt ggtaacctaattgttaaagattagattattggtattatattaagcaagtaagaaggcacaat agacaaatcataaagccaatatattaaggcgacaaaaagaaatataaataagctcctt atttagtaaaaggattcagattgttgacaaggtaaaatacaaggacaaagatttttatt cggtagacgcagttctgtgtattttgatttaagaaactgtatgtacagttattcataggagt gccaatgttaaagattttaaacttataagcaaggcaaaaacattattatgggaaaggagga tagatgtgctgttccctcatgactaa</p>	<p>gtcaataaccaccactaaagtggcgggctt gcaaaagcctttattgactagactaagcgaa agctacgttacctatatacaacacctcgaa tgttaccctagtctgtagctctgttctaact ttaacagttctgatgggttaggaacagttag ttaagtgtaaaaagtagatgaacattgtctaa gggtactaactccagaaaggagacataac ttgaga</p>
510-511	<p>Ga0315268_10018380 GENOME_ACESSION: C1_top_MG_2_5F_3300 032173 GENOME_ID: 279979 CONTIG_ID: 18379 SOURCE: JGI DATE: 2019-06-13.</p>	<p>atgccgtgctcgcggcccaaacgggggtcctgctcggcggggaaagccgaggtc gtgaagagaacccctttgtgatccggctgaagtacggagcgggggaccggcaac cggtagaccgcggggctggacagcgggtacatgaatgctgggtgctcgtagtgtcggg tgggaaagagctccacgcggaaagggtgtgtgagatccgacatcgtgtactgaaca gcgaagaagggcagcaccgcgcaaccgaagggaacaggaaaacctggtaccgcaa ccgcggtcctgaaccgcaagaacccggaagggtgtggtgcgccttccctccagcaca gctggacagccaggtgaagctgtgtgtcgggttgcgaaggtagtaccggtcacgaag gtagtgtggaagtcgccagcttcgacatccagaagatcaaaaccccgagatcgcg gagtcgggtaccagcaggcgagcagcgccgggtttcccaacgtgaggagtagcttct ctaccgggacgggcacaggtgaggtgctgcaacggcaggtcaaaagacgagcggtt ggaagttaccaccgggagagcagaaaacccggaggaatcaccggagaaatctgtg gacgctcgcgagacctgtcacgaccgggctacggccggagagatctgtgttggaa aaactcctcttgggttcaaggcgagacctcatgacaacggtccgatggaagctgatg cccggtccgggagtggtgtgtgaggtgtctcacaccttggctacatcacaagatgc gaagggaagaagcgcgatcgagaaactcacgcaacgatgcttctgctatcgccg gagggacgggggaatacagagggcagaggggtgatcctgtgaaaatgcaggtccgga agtgcacacgggaagctgttcaaggcgacggagccacctcaggaacacggcgctc ggctgtgcaagggttgcgggttcgacaaggtccgatacaggggagtcgagtttt gtgtccgagcaggagaagcaccggatatttcgacctgcggaagctggatggaccgag tccatgcaagcggaagtggcagatctgagactgctggaacggggaggaacgatgc cgtcgggatacaagcagccaggagaggggccgcgcatcctccccaggggcaagcc ccggggtctcctgcgcgatgctatga</p>	<p>gtcaattaccctcgggacgagccggggg cttgcgggaagcaacgcaacaccgcag acctcgattgactagggggccgggtggaaa caccggcagcaggtgtcttcgtagcaca ccccgggataatccgacgtccgggacct gtgtccacgggttaaaagggttcagggtta aaaccgggtccgtagacatgacaagccg aggcaacaatcccgatggcgaccaaccc cctgaacggggagagagactgtctgtt agtcttc</p>
512-513	<p>Ga0209317_1000713 GENOME_ACESSION: RifGroB2SPAdes_5F_33 00025013 GENOME_ID: 34862 CONTIG_ID: 712 SOURCE: JGI DATE: 2019-01-27.</p>	<p>atggcagtttctgtatcgacaaaagaaagatccactatgcttctgtagaaaaagg cgagactctgtgaaaagggttaaggcagttattaccgtcgaattctttactatcagg ttgaaggaccgattgtgtgagtagacggcgttgcgtcaagggtcagtcaggcagc aagactacgggattgctggtgactaccgagtagtgcaagggtgttttgcagcagagata cagcatcgagggaacaagccagagaagcactgaaggacggaaaggattcaggaa aggccgtcgagccgtaattcttgcctaccggcgaccggttcgacaaccgaaccaga ctaaagggttggctaccgccatgttgagagcagaatgcgaacatcgtacatgggc ggagagattacgcaaacctcgccagtgacatctgtcacaggagttggttcgcttcat ttacagcagatggagaatccggagatttccggatcgagtagcagcaaggcactttggcc ggataggaagtacgcaatattcttgaaggtggggccgtaagtgtgctatttgcgac gcggaaaacacaccttgcagttcgaccacatagcgaaggccagcggtgttagca accgaatcagcaacctcacttggcctgcgggctgtgcaatcagaaaaaggactcgtt gatattgaggtttctgaagtcaaaacgaaaatctcaaacgtatatttcccaggcaa aacgaccgctcaaggatgctgccgtgtaaacgcgacacgatgggaattgtataacagg ctaccgcaactggcctccagtcgaaatcgggtctggggggcgacgaaattcaacag aaccagcagcatctccgaaaagctggtacgcacggcgtgtgttaggtctcttggc</p>	<p>gtcaactacccggcctgaaggccgaggt tgtgagagattacaagcttgagggtgaccag ggggagccgtaaccaaccggctacgttgtta tggaggttcaagaccgacggcagaatgctt cctcagttctgccaatcgaagtccgggtt cagacacgcccagggtagcccgaaacg ggctggacatagccgccatacaacatc cccaggggagatttccgcaagggaagc gtcacaaggcccgtaaggcggtttgaaa gaagcttcaaaaggatcaaatggca</p>

SEQ ID NO:	Name	IscB sequence	ωRNAs
		<p>gaacagggtgtatgccagcaggagtagctgttttgaattaaggccatgggacgcggat cgtatcagaggacccgggtgaatgctccgggtttctcgcgggtatctaccaggcaaa agcagtagtttaggtttccaaactggtagatgttaatcgcggacgtaccaaaaggga gtctgggtgtccatgccggcagagttgctatcgtatgtctggatcttcaacattcagactgc tgatggggtcgtccaaaggaatcagtcacgccactcccgtaatccagcgtggcgtatg gttataactactcgacaacgctgatagcaaacataaaggagagcgggcaaaagggtcat gctttgcacgacgcgtatcctccccggcctgaacgccggggtttccgcgcagcttga</p>	
514-515	Ga0373927_0000546 GENOME_ACESSION: GW9ome_3_FD_SF_330 0035695 GENOME_ID: 312474 CONTIG_ID: 545 SOURCE: JGI DATE: 2020-04-07.	<p>gtgcgagacgcgcagcgcaccgctcaagctcctgactaccaggccgatctgcggcag aagcgctataggacaacaagacgttcgcgtgccaccctgttcggacgacatctcgcca cgcgcagcagacgagaggaggcacacctctcgcacccctcagtgccatcgggaaga gttctatgacccgggcatgggctggagaaaggagatccctaatcagcgcataccagt cggagacgcacggggagtagctctacacccgtacacccgtccaaagctcgcacccctt gatgtgtggcaacgccgggtggcctgtgtactctcgttcgaagtgtacaacgaacagg agcccaactaaccacggttgggtggcaggtgtggatgcaggcaggacgtctgaggcgga ccgtgtgtggcagcatagacacggcgctcaacctctgggagaggcccgcatcaca tccgggaggccgtggaatatgcagaacgctgcgacgagctatgtccgactcacacgt cggcggaaccagagcccgcacccggcctgtagggaagggtgtgccgctgtgaca cgaccgagcaaaagggtggcaggtccatggcgcccttgcggcggttcagggtggcgaag agcacttctccggccttctgcgtgcgaggggacgcatgtgtatctcaaggaggtgtg cagaccagagaactgcatgaaaggtagctgtgcagaggccgaaacgggagagcctg caatttttagtagtatcgtcgatagacgcccgggtcctggcagccagcagcatccgtggg acatgtgtccgactagcgcgcaccccccaacagcatcgtttgcacgcagccgggg tggtctgcgcaaacctctgcaggaggcatgtgcctggactcaacgtggcagcgtgtt ttgtcattccgggtatgggtctgtgcgtgggtgtgtttggccagagagcgtccattatca gcgtgcagctgtgttgacgaacgcacgacggacgcgaggagcagaggcaacgatt gtcggcccttgacccgggtggcgtgtcgtcccgactgtcgggtgtatccaaagaag cgaggcaaggaggccacccgactcctgcccatctcagaacgcctgtttttccact cggggaacgctggtggcccgccacagcgggagggttga</p>	<p>acgaacacaaccgctgtgttctgtcattcag atgacgggtcactcggggccagaacgtgtg atctctgcaaacgtctctgtgtcgtcgtggc ccttcgacgaatactatgtatcaacgtaaggc agggtcaattttccagggtattatctcagtg tctgaaactggaagcgccgaagtaccgc gtccctctcctctagccctgccacagactgg ctgtgtggcaggtataccggcgccgacg agaagtgggggtgtgtgtgtcctcggcga gatagctatataata</p>
516-517	Ga0136449_100279210 GENOME_ACESSION: Sb_comassemblies_SF_33 00010379 GENOME_ID: 35517 CONTIG_ID: 279209 SOURCE: JGI DATE: 2019-01-27.	<p>gtgcttctcagccccggacccttgaagcgccgtcagcagacaacccagggtgggc acgaacaggggcgccacggcgtaaggcaccggcttccgggaaggcgccgcgagc gcaacatggccgagaggagaccagccgggaaggctggcgtcaccgccgaaggggag accgccgaggtgtgtgtactgacaagcggtgaacccgctcatgccctgtcatccgtccc gcgcccgcgcgtctgtgttcggggcgggcggggttcacaggcacacgccgttcgt catccggctgaaggaccgtgacggcgcggtgtgttcagcaggtccctcggcggtggat ccgggctcaaaagaccaccgggatggcgttaccggcaggacggcgacgcgacccg ggaagtcctgtactactggaggtcacccatcagtcggtcgtatccgcaaaaggatcca acaacgcgcgcgagccgccgacgtcgcgatcagccaacctccgctaccgcgcgcc gaggttttagtaatgcaccaagccgagcggtgtgctggccccgtcgttgcgatcccg tgacgatgtttgagacatggcgcgacgactccagcgctgtgtccctatcgtcgcgac gacctggaactagtcgggttgataccagggccatggagaagcccgagatcgcgga cggaaatcagcagggcacccctgccggatataggccgtgagtagctgtatagaa gtgggagcgcaggtgcttattgcatgacgcaaatgttctcacaacatcgaccacat cgtgccccgctcgcgcggcgagtgatcgcgtctcaacctcacgttctgttatccc gtgcaaccaagccaagggtctcatgccgggtgcgggttcgtcgcggtatccggaccgg ctggcgcgatcctcgacaggcaagggttctgctacgggacgccgagcggtgaact cgaccagattcgctgtctcgcgcgtacgcggcctagggtccccgtcgaatgctgg tccgggtggccgaccaagtggaaccgccctggaccgggacaccgaaactatgctc ttgacggcgctgtcgtcgcgaggtgagccagctgaaaggaaacctctgcgcctgtctac gtggtcaaaagcgccgggtcgaggttgcacgacggacgtgtccagatgcctacggatt ccctcggctgcgcatgccgcgagtcgaagcaggtgtacggattccagaccggagacct gtgcgggcacacgttcttaacgggagatagccgggatgcatctcgtcgcgtcagcgt ccgaagcagagggcagttccgcatcggcgacgggacatcaactacaggaattgcatg ctaattccagcgagcggagggatgaccatgccatttag</p>	<p>gcctggtcctccgctccctgagggtgaggag taaccgatcccaatgttctacatcggttacc aggccaagtgcctaaggaggtgactacgtt gcaggcaggacagaagaccgactccggg gtgcttctcagccccggacccttgaagcgc cgtcagcagacaacccaggggtgggac gaaacggggcgccacggcgtaaggcgc ccggcttccgggaaggcgcgagcgca acatggccgagaggagaccgccaag gctggcgtcaccgccgaaggagaccg ccga</p>
518-519	Ga0213873_10000897 GENOME_ACESSION: VelepcomBERX_R3_2_ _SF_3300021358 GENOME_ID: 37230 CONTIG_ID: 896 SOURCE: JGI DATE: 2019-01-27.	<p>atgcaagttttcgtcctggacaccggccgcaagcgctcgtatcgtcgtccggcgcg ggccccggtcctgtgccaaggggcgggccgctgtcgtccgctaccccttcacg atcatcctgcacgacgggaaggttgaggattcggctgtgcagcagcagcgggtcaagatc gacccgggttccaagacgaccggcatcgcgtcgtgcagaggccacggcgcggtc gtcgtcgtcgtcgaagtcgagcaccggggccaggcgatcaaggcgtgtcgtcgcgac gcaagggccctccggcgaatcagcgccggccgcaagacgcgtatcgcggcccccgg agcaagggcgaggccgcgaaccaagaggcgagaaaggccaagggttggctccc tccgagtttgcagcggatcgccaacgtcgtcgtggctcgtcgtcgtcgtcgtcgtcgtc gggtcccgtcggcgatcgtcgaaggttgggtcgggtcgtcgtcgtcgtcgtcgtcgtcgtc gatcccgagatcctccggatcgtcgtaccagcaagcgacgctggccggctacgagttgc gggagtagctcgtgagaagttcgaccgcacgtcgcgtatttggcagggaccgacgta ccgctccagggtcgaacacatcgtaccggcgagcaaggcggtatccgaccgggtcagc</p>	<p>gtgaagttccccaccggtcaacggacccca gccctaaggcgggcggtgtgttgcgaag acgccaagcaggtgaacagggtcagccc tagcaataggctacgttggagggtgacga cagacgggatcggaagcagctccgatcc tgtcggcccgattaaagttccgaaggtcta ggacgggtgtcggggcgtaggaggtcag aagcgacacaacattgcccaggcgacgtt actcgaaagcaagggacctcgtgaagggc aatctacatgcaa</p>

SEQ ID NO:	Name	IscB sequence	ωRNAs
		<p>aacctcacgtggcctgcgagccgtgcaaccgccgaaggaaaccgaccgtc-gag-gactttctgaagagcaagcttgaggcttgcccaagatctgaagcgagccaagcccc-gctcaaggacgccagccgtcaacgcgacacgctgggagctgtaccgcccgtccg-gggcaggggcttgcctgagtgccgtcggcgccggagcgaagtccaaccgcgc-gacgcggcctgcccagacacactggctcgacgccgatgcgtcggggccagcac-gcccaggtgctcgacgtcgaaggttgcggccgtgtgtccagggctcggggcac-ggcacgcgacagcggcagaagttggataagttcgggtcgccaaggggtaaccggcc-cagaggaaacagcagcttacgggtccggaccggggatactgtcgtgcccacatcc-cccatggaagcagcgggatacacacggggcgaatccagatccggaatcgaccgt-cattcctgtagaaaggccgacgtgcatcccaagtatctcacgatccacgacggc-gacggatacgcgtattcgtcggcgagaccatcgttatccggggcgagccggaccggat-ttgcctccccccacggctaaagccggggggccccgcaaatcctga</p>	
520-521	<p>Ga0070717_10012599 GENOME_ACESSION: LARL11StandDraft_3_\$ F_3300006028 GENOME_ID: 31022 CONTIG_ID: 12598 SOURCE: JGI DATE: 2019-01-27.</p>	<p>atgtcacatgtgtttctcctggagcgtgagcggcgctccgtgtcctcgggtgcaccggga-cgggcacggcctcttactcaagcaggcaaggctcgccttcaagcgttcccccttacc-ctgattctcagacagctcgacgcgaagccgcccgtgagccgtcgtttgaagatga-cccgggcagtcacaccactgtgtcgtcgtgtagcagagagagtgccgaggtgctg-tggcagggggagctcaccatcaaggggaggcgatcgagcgcctgcggaaagcg-gcgagcgggtgcgagaggcgccgacacgcccatacagcgtaccgagaggcgcgct-ttgccaaccggcgaggaagacggctggctgcccctcgtcgtcgagccgggtgca-aaacgtgctgacctgggtccagcgttgcggcgccctgtccgacgcggcgctcgtc-ggagttgtgtcggttcgacacgcagggccatagacggccgacatcgccggagtcgt-taccaacagggcacgctggccggcttgaggtaaggatgtgttggaaagtgggg-caggcgctgcgctactgcgagtcgagcggcggtcggctggagatcgagcatatcatg-ccgcgagccgtggggggagcagccgcgagagtaatctgacgctcctctgtgtccct-gtaaacaggccaaggggacacaggacattggcgtgttccgtgcccacgaccccgagc-gactcgcgcgcatctggcgagcggaaggcgccgctgcgtgacgtggcgccgggtca-atgcgacgcgggtggcgctctatgaacggctcgtagccttagggctgtccgtcaggtg-ggcagcggcgggcggtaccaatacaaccgcaacagcgaacagatgcccaagacgcat-tggacggatgccgctcgtgggggagagcagcccagcaggttcggaactggcag-accgtccgaccttctgattactgccaccgggagggcaagccgtcagatgtgccacgt-ggataagcggcggttccgctggtaagcccaaggggccgagccggtcacaggggttc-cggaccggggataggtgcgagccgtgtcaccaaaggcgtaacatgggcacctacg-tggggcggtggccatcaagtcgatgtctactcaagctcacgaccaggccagggt-ggtggaggcgatccatgccctgtattgcacccactgcatcgtagcgtgagctatgggt-a-cggcattgggagctagcagcgttctccccagggtctcaaggaccggggctccgct-gcttggagggaaggggcaccggaatcgtcaagcatccatcccttcgaaggatggg-ctcgtga</p>	<p>gtcagcgacccacggggaaaccggggg-gcatgtgcttgagaaagagcatgccctgatg-ctgaccagctccccgtgggtcctctgtgag-aggcgagccaggagccgttctgtgggag-cggcaccagtcggaccctgggatggccctt-caggtcccaggcgctgcaatcctgtgttaa-acaggtgcagggggaagcagagccgggtg-cagcgggaagaggccgccaacgaacag-gagcgaggaaaactcaccgcgaagcg-gaggcccaagcgggacagaccacgaag-ggaagcagtagatgtcacat</p>
522-523	<p>Ga0163147_10008687 GENOME_ACESSION: OliLakLV19MP6G1_2_\$ F_3300020192 GENOME_ID: 33721 CONTIG_ID: 8686 SOURCE: JGI DATE: 2019-01-27.</p>	<p>atgcaacagattggtgctagataagagcaaaacccgctgatgccgtgcatcccgca-cgcgcacatcaactgctcaagcaagggaagcgccggcttctgcaatatcccttacc-atcattttgaaagaacgggatggtggggatgttcaacccgtcactatcaagtagaccg-ggcagtaagaccactgggataccgatgtgcgtgatttcagagcgggtaagcgtgtgatt-tgggctgggaagctcaccatcgccggcagcaaatccgagataggtgttatcgaccg-tcaataaaggcggggcgccgtgtcctgaagacacgctatcgcttggctgttcttaaat-cggcgtcgtgtgggaagcaacttgcgccatcgttcaaaagtcgagtcgagaacacact-aacttgggtcggcgctgagtcgtgtgccttattaactgatctcaatggagcttgtgc-gcttggatagcagttgatggaaatgtgaaatctcgtgcgtggaataccaacaggcgga-acttgcgtggtatgaggttcgagaatctactgaaaagtggggacgcgagtcgcctat-tgtcatcccgccgataggccattagaatcgaacatatccccaaaatcacgcgtggc-tcaaatcgtgtgagtaacttaccgtgttgcctgactgtatcaacgtaagggtagcca-aactgcgacagaatttgggcacgccaatccaacagcaagccaagcaaccgctcaaa-gatcgccggcggttaacgcgactcgtggcggttggtagacgcatggacacattgg-gtttccattggaagtcgggacagggggacgcagcaaatftaatcgggtaaaacaggac-tatccaaaagcgcatggattgacgcggcttgttaggagaaaagtgagaacaaattac-accagcctaataatcgccactcagatgaagacgacagggcatcaatcgccgcaaat-gtgcggcggtgataagtagcgttcccaagaaccagcgccaaacaggcggtgtccact-ttggtttccaaacaggggatagtcaggcgcaatgggtcacaaaaggagtaaaacgcgg-atgtatgttgggagagtttagtgcagggtagtgtgttttataatcactaccctaccgga-acgattcagggtatcagctatcgtacttcccccaattcacaatccgatggctatcttac-gagaagagagggcgcttcccccgctaccacaagggtagcggtacccccgcgcaaa-gttcaatga</p>	<p>gtcaacgaccacccctatgaagggggtgg-cttgcgataagcaagcccatgttaccagcc-ttagtctcattcgtgagacaccgttattcggt-cacgataaccggcggtgagtaaccagcc-cgtcgtctatcgttaccgttaaacatgccg-agggtctaggaagtcgggtgggcatgac-aagccgaataccattggcgaggttcacttt-accgcgcaaggcgagaataatgaggtaa-ctcatgcaacga</p>
524-525	<p>TPA_asm: Euryarchaeota archaeon isolate Hy Vt- 292 Hy Vt- 292_k295_120690, whole genome shotgun</p>	<p>ttgtgttacttaactgggaaatatacacaagatataagcgttggcatagatgcagggtattc-agccataggggtcagtgcaaacagacaaaaagggaatttaattcaggtgagtttaatttg-agaagagaatttcaaaactcattgacgaaaaagaaattacagaaaaaggtagggcaaa-taagtatggcacagaaaataagatttaataatctagcaaacctgaaggttgggttcac-ccagtattcagcataaacttgagacacatctcagattgatagaaaaaactgaagagatact</p>	<p>gtcagtcaccacggcgtaaagcacgtgggc-ttgtctgtgaggagcaagagcaattgggtga-ttaggaggcattgaatatgcagaagtattgg-tagagtttcagaaacaccagaggatgtccc-tcaagtcctctgctgtgaagtgaagtattaa</p>

149

SEQ ID NO:	Name	IscB sequence	ωRNAs
		taaaatataatggatataatagcgaagtggaaagtagaagaataaaagggaagaataaa gagataacgagcaattctccatactgcgtag	gaataaagcgaagaagcagtttacaatagt attttctatattaaattactgtaagagatatag gaaattattcaagacatatcatgtggaatagat tccggatcaaaagagagaagcaataa
532-533	Ga0197142_1009778 GENOME_ACESSION: GBS60_MetaG_2_5F_33 00020153 GENOME_ID: 26410 CONTIG_ID: 9777 SOURCE: JGI DATE: 2019-01-27.	atgtatgttcctgtagtgtatcgaatcaaaaccccttatgccaccacaccaagtcgcgc caggcgttggattaagctctgcaaggcgactcctttctggaagaagggtatctgcgtc agactcaatgttgaacctctgaccgtcgaactcaaccgattgctggatgcacgacctg gcagtaagaaggaagcgtgacagctcaagctcaagctcctatctcaaccttcaaa ccaacgccgtcacttgggtgaaggagcacgttgagctacgacggcagatgcgtcgtact cgacggtatgcgaacacccctgtcgtcagcccgagctaatcagcgagaggtggtat tcttccctctacaaaagctagatggcagtggaactacgcttagcaaaatggttgatga attatttccaatcactaccttgggtgaagacatcaaggcaaaagacgaaaggtgaagcgg cgggtgggatgtctgttctctcttggaaagtagggaacagtggtttacgagcaactgg ctaaactatccaggtgttaacccgttcaggttggagaccaacaataatgcgggatactc taggtctcaggagacgggtataaaaactgctgaagctctcagggtcattgctgcgtac ttgggtgttagccaactctgttgggtgacatcttcagcctgacaacaaaagactggtgt cgtgttacccttacaatttcatcgtcgaacccaccgttggcaacctggcaacaaggaa aacgcctaccctatgggggaactttaaagctaggggtgaaaagggtgattttagtcaaca ccccgaatacggactcacttgggtggtgctccaagggtaggtgagttgcatgg cctagcagatggaaaaaggttaacgcaaaacgccccaaaagaagatgtcaagttttaac ttttttctgtggaggtctatgttagaagcgcaataa	gtcaagtacccccagcttttagctgggggctt gtgggattagagcctagccagctggaaacc acaggggcgaagaaagactagcctgccca agcaacctaccttcgagtagggggcgttga agggaactgaaccagctggacgtggcag ccgaaagacgtagcgaatgcctccctagt gtctacctctctgtcgtcagtggttagcg aagcgagccgcaaggcttaggggaagtata cactccactaggagacttactctcat
534-535	human gut metagenome genome assembly, contig: NODE_725_length_3696 1_cov_16.589144, whole genome shotgun sequence GENOME_ACESSION: UEPH01.1 GENOME_ID: 15077 CONTIG_ID: 724 SOURCE: WGS DATE: 2019-01-27.	atgccaaaacaaagtgtatgtcatcaacaacacggcgctcgcgtgatccctgtcctccg gcaaaagcccgatctcgtggtgatgaagggaagccaaaataaagaagaggacgccct ttaccatccagctggctacgggaagcagcggtacaccaagaagtcattttggcggtgg acgctggaagtaagaccatcggaatgtccacatccacaagaaggaagattcggctc catgtccatcatctggagagcagggaaggttagggcgacgccccggacaaccaagtta ccctttgcgagagctgtcacgagaactccacaagggtgtgattacggcggaaggacttca aaaagcgtgaagcgtcttcacacgggatgccatcattgggcatcatgcggaagactt taatgcagaggctctctctgagttgcctatccccgttcattgagacaggggctacatcac aaaggcgacccgagagaagctcgtggtcttcccaagagccacacacatgacgctctg gcaatcgacaggggaagcagggtctcaatgtggctaccgtccgggaatcgtaacaat taataagattatagcatttgcctcgtgaaggcatcacaccgccagctgcacaaggcgact atcctcaaggaggatccgtaaatccaatcaagcagaaaaagtc	gtcaataacccccacaaatctaacgatttg gagggggtctgagaaaatcgtgaagccctg attgactaccctaagtgttcgagcactacgtt actcaagaatgaagaataggcaccggcg gcgtgaatccgaacctccgcactgcgtgtg tgtgattaaaagctctgagggtgaagagcg gtgtgtcacacgaaaacctgagataacatt gggtacggatcctgacggccgaaaggcc gcgtggtctttctgttagccacaatcaagaa gggtgagaggagggaaggcatcatgccaa aaca
536-537	Ga0307988_1006404 GENOME_ACESSION: EllFjometagen183_2_5F_3300031741 GENOME_ID: 280157 CONTIG_ID: 6403 SOURCE: JGI DATE: 2019-06-13.	gtgaaagtctatgttttaacaaaaacaaaaatcattgatgccacaacttgcctaaagc aagaatattattaaaaaaggcgaagctaggggtcataagagacaaccttttacaattcagc ttttaaataagtagctggagaggttgaacatgttaattcattgacagcgagcaaat aatattgggtttccgcgcatcagggaagagaaggttaattcaggaaacagttatactgat gtgatgatgaatcaagacttgatgacaggcgaatgtaccgaagaacaggagggaacc gcctctggtagaaaaaccgagattgataacagaaaacgcgctgacgactggttgcgg cccagttataaaaggcgataccagctacatcgtactattatcaataaaatgaagtcactatta cctgttaagtattttagatagaagtaggttaattttagatacaagctattaaaaaaccgggaa taagtagtagcaggttatcaacaggcgtaggtgattgattgcaaatctaaagcttatataa tagccagagaaaaagtcattgtcagttgtcggttaagtggttaggaacaagaattaa tctacatcacataatcaaggtgaaggcggaacgaataaagctgataatttggcactgt tactatgtaaaatgtcaaaaagaatataaaaaaggtcttggaaaaaccttaagagaaa taaacagtaccgggaatctacatttgaatataattaaatggaaatttaagcaggatttagtt tgcattttaacttttgatttaaaactcttgcgaagaagacagactaataatcccaaaacg cataacaatgacgactttgtattgcccggctacagaacaagcgagattgcttcatctga agttatcacagaacgaaagaataaccgtcgtctacagaaaacaggaaagcgttgcctc cggcaatcgcaggcaagggtattcaatagcccgaagatctggtgaagatcaaaagc aaatggcttaattaccaacggttgcactgtaaaggaaacaagaatattagtaataaaagatc aattaacattaaacagtggaagcgtttcaacgttaggaacattaggaggggtgcaattc ctcccctgttaaaaagtataactag	gtcagttaccccatagctaaagctaggggg cttatccgcaagatataattagctgattag gaggtagaaaatcacagaagttaaaaagggtg aaatacatcacaccaacaattacctccta gattgttgcactgtgactgttaataaaca agttcaaaagacttagttgacgttttaaaac ttttttaacagctccgaagggaattaaaccac gaaagttaggtgactgtgaaa
538-539	bioreactor metagenome genome assembly, contig: NODE_150_length_1057 57_cov_135.353352, whole genome shotgun sequence GENOME_ACESSION: UPAD01.1 GENOME_ID: 16334 CONTIG_ID: 149 SOURCE: WGS DATE: 2019-01-27.	atggattttgtattagatgcaataaaaagtcattatgccttgcctgaagcagttgcaaga aaatgcttaagcaaggtaaggctgcaattattagaagatacccaatttacaataatccttaaa aagtagtagaagacactaaaaaagcaagaatatagattaaaaattgattttagta agcatacaggattagctatactacaaaataaattgaatattgctaggcgcaaatagacca tagaacagatattaaaaagaagcgttgatgaagacgtatttttagacgtatagaagaac agaaaaataagatatagaaaaccaagtttttaataagaaaaagaaaagaggatggatg tcaccaagtttagaaagtaggatttaataataaaaaacattgggttaagttacaaaaatfa attccattactacatactctatgaaatgttaatttgaactcaattatgcaaaatcctga aataagtggtattgagtatcaacaagggtacactcaagggtgataaattagagaataattact tgaaaagtttgtagaaaattgttattgtgaaaagaaaacattccattagaataagaac atataataccaaaatcaagggttggaacaacagaattgataatctctgttagctgtcat	gtcaactaccaacgactaaagtctgggact ttagagccctatgttgaccaggttaaggttg aacagaaacctgattatctatgtcatgacac cctaggatgccctcctgttcttaggctctgtc gtacaacattaaacaggtgtatgggttaag ccagttgtgtgcatgtatagcatagataaca ttgccgaggagatgttacctgcgtgaagca gagaagagaaaaatcctatg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gagtgtaaccagaagaaggtaatatgacagcagaagaatttggtatccagaggtaaa agacaagttaaagaaaca	
540-541	Ga0070738_10015183 GENOME_ACESSION: CoaSoiStandDraft_9_\$F_3300005531 GENOME_ID: 24565 CONTIG_ID: 15182 SOURCE: JGI DATE: 2019-01-27.	ttgagcaacgtcttcttctgataccacaagcagccgctggccccggttcacctgggc aggccccgctgctcttaaaacagggaagccgcggtcttaagccttaccatfcaccc tgattttgaagcgagccatccccaccgaagagggtccacccccccgctcggtacaaa atcgatccccgcagcaaaactaccggtctgctgctgtagtaccacacgcgctggt cgtcttttggccgagctgtccaccgggggcagcagatcaaaaaaagcctcgacaccc gaagaggagccggcggggacggccaccgcacacactcgctaccgtgctccccgt ggcgcaatcgagcgagagccgggctggttccccctcgctcgtagccgcacccg caatgtggtcaccgtgggtcgagcgtttgtgcccgtggctcccatccgagcgatcagcct ggagctgtccggttcgacgtgcacgcatgaggaacccggccatcagcggtctggag taccagcaaggcactctgagcaggtacgagatccgcaatacttctggagaagtggca gcggacgtgtccattgttggaaagaggcggttccgttccagggtggaacatctgtgccc ccgtgcaaaagggggcagcaaccgcgtgagcaacttgggtctcgctgctgcaacccgtgc aacacgaagaaggcacgtgtgataccagggttctcaagaagaagcccgaggtgtt gaaacgcatccaggagcagggcaagggccccctcaaggatgcagcgccgtcaatgc taccgggttcgcttctggagcaggtcaagcaactggattccgggtgaggtgtgga gcggggggcgccacgaagtcaaccggaccacggggcttgcacaaacgcactgg ggagatcgccgctgtgtggagccagtagccagagcggttgcaggtcagagggttc cgccccgttgcagcagcggtgtggccatgggtgtgcacgggtgcgaatgtcactgcc attggctgtccccgctcagcgcccaagggcccaaacacgtgcaggcgctccagaccg gcgatacgcccgccattgtcaccacgggtgccaagcaggggacatcatgagggtcg gggtgtggtcggggccagcagatcattgtatctgtgacgaagcatggacggattgctgg gctcaatgcgcgttttttgcacatgcacggctcggatggctatcgtacaccaagga gacgtctatgcaacccctgcccattccaccgataa	gtcatggaccccggttgaacacggggcctt ggagcaatccgagcgattgaagcaagccc cgccctattccgaccgaacatgagcagac ttggccccgaagggtccggttggggcgga atacacaatagttaccgggtgggtcgccgg ccagcccaccgctctacggccagggttaa acagatcgagggtacggtcagtcctctgg caacaacacccgttccaacattgtcgagggc accattacccttgaagaggaggtcagttt gagcaac
542-543	human gut metagenome genome assembly, contig: NODE_9436_length_504 2_cov_3.339082, whole genome shotgun sequence GENOME_ACESSION: OVZI01.1 GENOME_ID: 11798 CONTIG_ID: 9435 SOURCE: WGS DATE: 2019-01-27.	atgccaaaggtctatgttctacccggcagggcgctccgctgatccatgctccccagca aaagccccgtcatcttctggtatcggggcaaggcaaaagtgaagcagcaaacaccgtttac cattcaactgctctacggcagtagccggtatcgccaggggtatcctcgggtggacgc cgccagcaaaacatcgacgttctccgctccacgaacggaggaaactgttctccgag aggtaagccgcgcaacgatgtgtcgagctgatgcttccgacgagcaggttccgcccgt gcgcggcgcaatcgacacccgggtatcgtaagccagcttcgacaaccgcgtgcgga gcaagcataaaaggctggctgcaccttccgtggaagtaaaatccaagatcatacaccg ccatccggcgctgtgtggtatcctcccgctcagcaaaagtgtgtcgtagagaccgccgag ttcgacttgcagctcctcaagccggttgcggaagggaagcctgttccgagggcgagga ctaccagcatggggagatgtatgttactactaacgtgcgccagtagtattttgtggcgcat ggctatacctgccaatgctgcggagcgcacgcgaccagaagaagagggtcgggctc catgtgcatcatttggaaagccggaaggtaggcgggcgacggccccggacaaccaagtta ccctttcgagagctgtcacaagaagctccacaggggattgattgaggcgaagaacttca agaagcgcaagcgtcgtccccacagggatgccacattcatgggcatatgcgggcaac gttgcgtcgaaggcttcttccgagttgcctatccccgtcattgaaccagaggctacatc acaaaagcaaccagagaaaaagctgtgttctgcctaaagccacacaaatgatgctct ggcaatcacgaggggaagcatggttcaatgtgggtacacctccaggaatgtacaaa gcgataagacttacagattcggccggtgcacaccataacggcaactgcacaaagcg actatcctcacgggagcgtccgaaagccaaccaagcgaaaagtacgtctgtggtct ccgtctatagacaaggtgctctaacgggtatcagtggtttgttggggcagcgggacc agcggtctcttttctacttctgagcgttagacgggaacaaaggtcaaatggtgctggacac agaatccctacttctgctcgaacgaagcagaattatctcattgcttaa	gtcaagaaccaccacttaactctgacgag ttttgaagtggggccttgggtgtaaacccg aagccagattgactaccctaagtgtttcgag cactacgttactcaagaatgtaagaataggc accggcgggcgctgaatccgaacctgcccgc actcggtgtgtgattaaaagctgtagggt aaggagcggtgtgtgcacacgaaaccttga gataacattgggtacggatacctgcacagcc gaaagcggtgagcggttatttttagccgcac ccaaagaagaggagcatcaacgtatgccaa ag
544-545	Ga0071116_1002195 GENOME_ACESSION: IMG_3300005077_\$F_3300005077 GENOME_ID: 311639 CONTIG_ID: 2194 SOURCE: JGI DATE: 2020-04-07.	atgccatgttcacccgctaagggaataaactttaagagaaggcaaaagtagta aggagaactccctttacaattcaacttactattgcaacaggagagacaaaacaggatgtta ctctgggaatagatagcgggtatctcattattgattatcagcaatcacagagaaaaaga attattttcaggttctgttataaaggaagatagttaaattgaattcagaagaagagatt atagaagatgtagaagaacagaaaaactgtgtacagaccggttaaggtttctaaacaga aagaagaagaagggtgttagctctcttattacaacacaaagagaatagccataaaag aattatagaaaaactgaaaacataattacctgttaagatagatagatagaaagtagcaaatftt gatatacagaaaaattaagaacctgaaatagaagggttagaataatcaaatggagaaca gaaaggtttctggaatgaacgagatgtgttatatagagacggccatacctgccagcat tgcaggggaaaatctaagataagatacttgaactcatcatttagaaagcagacaaaca ggaggtaatagacctgataatttaataactctgttaaaacctgtcacagtaagggtatcagc cgggaaaattactctcaggttaaacctctaaagggttacagagaggcaggtattatgacg accataagatggaattacttctctctctaaagaaagaaaggaaacttcatatacttat gtttatattacaaaaagaagaatagaactggcacttccaaatcacatattatgagtc tttgttatagcaggaggaaaagaacagagtagaagtaagcttattctataatacacaca ggcgacataacaggtcattacaacaacaggtcaggttttaaacctctattagaagaa aaaatatcttaccagcctggagacacagtaagactgacaggtaaaatctgactattaaa	gtcaatgaccccgccacaaaggggacag gcttgcgtgacacccgtaagttatcacattg cttagaggcttaaggaataattgttttatgca gcagttattcaggagagatacacacacct ggcgatactctcagctcagtcagctctgtg gttttaagttaaacagtcctgtgggttaggga cagtgcttaaaagcttaaaacctgaataac acctcgatgggacacccggttttaaaagg ttccggaatagaacttgagatgattattggc aaaacttttagtaaggagtaactcaagttgtat agcttt

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gggtgttttaattatggaactgggttagattaaaaatagaaccggacgattagaaatgt ggcaactaagaaaatagaactgtattaatatgtggaggattttctttgaattataa	
546-547	human gut metagenome genome assembly, contig: NODE_8540_length_2630_cov_2.261748, whole genome shotgun sequence GENOME_ACESSION: UZOQ01.1 GENOME_ID: 279611 CONTIG_ID: 8539 SOURCE: WGS DATE: 2019-06-06.	atgtcagtttctgttttgttctggacaggaagcaccggcctctgatccgtgccgaccgc tcgggcggaaggtcctgaaatcggggcgcgcccgcgtcgtcaagcgtctcccgtica cgatccgctgtgtggatcggctgacgagaactctgacgttcagcccgtcctcgtcaagtt cgatccgggtcaccgcgagaccggcctcgggtgttcgagccgacgggaagaggcg ccaccatgcgtgtttcttcaatgacctgttcacgcggatccgggaatgcctga gcgcgtgtgtgccttggccgtcggcgccgagcaatctgcgtatcgagccccacggt ttctcaatgcaccaagccgcaggggtggccggcaccgtcgtctgacaccgcgtggat gcccgtagccgatggcgcccaagctatccgtctcgtccccgaacgggattgatgga agagctcgtcaagttcgacgcgcagaaagctcgagaatccggaaatttcggcgccgagt atcagcaaggagcgtcttcgagtatgaggtgcgtgaatactactggaaaaatcggca ggaaatgcgtctactgctcagccgagaacgtccgctcaacatcgagcacgtgtggcg aaggcacgagggcgatccaaccgctctcgaaactcgcgtcgcctcgtgtggcctgca accaaaagaagggggcgagcccggtgaagttttctcaaaagaccgacctgaggtgctt gagcgcccaagcagcaatgcagcggttctctcggacgcggcgccgtcaacgcaa cccgtgtgtcgttctcaatgcgtcaagacgttcgggtccctgtcgagacgggttccg gctgattgacgaatcaatcgaagttccttggcgattgccaagaacattgctgatgc gctctgcgggtcgaaatcaacgcagcgcactatccaagagcatggcgctctgaatt ccgatgcaccgggtcggggagctatcaacgtacgaggtcaccagcacgggttccc cgcggtcttctgatgcgcaaaaacgcgttcacggctcgcacggggcgacatgtgcaa ggccatcgtccatcagggaagaagccggcgctatcggggccgctggccgttcgt gcgtcggcgagcttcaacattcagacggcgaggcgctattcaaggcattagttggcg cactgtcaactgcttctacaacgacggataggggtatgcgtggctgcgccgcacct cattctccccgtctga	gtcaaccacccctcttaaaaggagagcgtt gagccatcaagccggtgaccagcctcagt gatccgaaaggagaactacgtgtcgtcag gtgacaagaccgccccggggggttctc cagctccgagctccggaacggcagaagc agacaagctttgggtatgcacgaaacggtct gcccgttcaatggcgagggcaacattggc gaggggaattcaactttgtgatgtcacggg ggcaaccccgattcgaaccgtaagggttcaa caccatgtcagttct
548-549	Ga0116164_10013910 GENOME_ACESSION: AD_UK109_MetaG_2_\$F_3300009775 GENOME_ID: 21298 CONTIG_ID: 13909 SOURCE: JGI DATE: 2019-01-27.	atggatttgggttgaataaagacaaagacccgggtatccgtgtctgaaaaaaggcaa ggaagtgttagaaaaaagaagctgtgatacatcgtcttacccgatgttattcgtttg aaagaacaaaaagactacgaagtgggggattgaggtgaaactgacccggagcaa agagaactggctcgcggtgttgcataaaaaatcagagcatgaagcggatgctatactttt aggtgagattatataaaatttaccgtcccgcaaaaggtaaaaaagattgtaaacgca tatatatcattggcgagctcgtgtgtatataaagaagaattgtcaggatagcgcgagcg aagcagaaaaacagatcacagaccggcgagatgtgttaaacagaaaagcatcaagaaa gaaaggggtaccgccccttttgaatcaagaataaacagactgttcatgcggtagagaag ttgatgaaatggttgcggttgcgtatcagttgaacacgtcaagtttgatagcagaa aatgagagatgcggtatgtagggtgttgggttaccacaaggtacactgtcagggatga aatcagagagtatttactgaaaaatttgggagaaaatgcgttattgcccgtcggaagaa atcccgttgaattgagcatattgtgccgaagtagaggtggcacaacagattgat aatctgtctatcgttgcgttaagtgttaagataaaggaaacaagctgccggacgag tgggttagctgagttaaaaaagcagtagaaaaatcagacaagctgccgcaaaaaacttt gagaagcggaagagatgttataaagactttgaagacacggcgaatttgaactcaact	actgttgaagaaaaagaagtaagggtcag actcgcgcgtgcggttaccttagtatttca ataactaactgaaccagactaagttctacga gaactacgttatttgggttaccatcctcag gtgtcttccagccgggaagctctatcgtga cattaaacagagtttaagggttaaacggaatc tcagtgtgtgcgaagcaatcaagcctttataa cattgtcagagagagaggtttaatccgtt acaagccgcgtgagcggaaacccatcggg gttattttaaataccaaaggaggaaaaaaatg
550-551	Ga0074478_1419092 GENOME_ACESSION: IMG_3300005827_\$F_3300005827 GENOME_ID: 311646 CONTIG_ID: 9600 SOURCE: JGI DATE: 2020-04-07.	atgttagtgtttagtttaacaataatggcaaaatctaattgccatgcagtcacgaaaaagc aagacttttataaaataggttaaagcaaaagtttaattcagaaaacttttactatccagctt attttaggttcagcaggtatataaacaacaaatttatttaggaatagataaaggtttaatttac aggaattagctatattataaagacaaaataattttacagctcaataaaaccatcggttga tgtatctgataaaatgattacaagaagtcaaaatagaaaacagagacgtagtaggctttgct atagaaaacaaagatttaataacagagcatcaagtaaaagaaaaatagattatgccatc agtataaaacaaatatagaagaactcttaaggataattatagaatcaaatctctacaacfc atattgctatagaagatgtcaggtgtgattgcaagattataaatcctgacttaacaggaa aagattatcaaaaatcaaacagattgaatgaaaatcttagaattgcatgtctaatgagagat aattatcaatgtaattgtctgtaaaaaaagaatttaagattagaagtcaccacattattcca agaaaagaagagggaagactcaataaccaatttaataagcattatgttcatctgtcattc aaaagttcacgacaataagattaaattagatattggtgtgaagcgttttaagataaagat ttcacagcaggcaatgcaaggcaaatcttatttatatgcagaatttaagtaaaattggataa	ctcaatctgatagttgctagtctataaaat atttatttttatagatatgttaactaacagagg cttagaaaataagcaaaagttgtaagctatc tacctttagattgttattcagctgaagctcta aggttcataattaaagggttaagcgaagccg aaagtgttatgaacgtataaaacttttacac agcctcaaggataatcactcttattatagg agaaacactttatgtattgta
552-553	Ga0400266_0006374 GENOME_ACESSION: UnnLak12425_FD_\$F_3300037330 GENOME_ID: 312715 CONTIG_ID: 6373 SOURCE: JGI DATE: 2020-04-07.	atgggtgcgtgcgtatccgtttacatccggctgaaaggccgtgccgatgcatcacacag ccggtccgtatcaagatcgtccggcagcagaccacgggatcgagtggttcgag aaaacggccagaagcaacacgttctggcctgtggaattagccatcgcggccgcacg aaattcaatcggcagagactgggcatccaaagactcagcgctgacgcggcttgcgt cgggagaagtcgagatcgtagaggttgggacgtccaacctagcgatcaaaagctacc atccagggaatatccatcgccactgtgttttaacccagcgagcagatagctacggctatc acatccaacccaacacgaaggaggaggagacagggaacgagtcgcgtga	gtcaactaccccgccctaaaggacggagct tgtagagaacacccataagccaggttgac cagggaagcggacacccacccgctacgt ttgccacaggtcgttaagaccactgctggtg tgcttctcagcaccgacactggaaaactg gaatcacgctggcaaaaggcaagccgcg aaggtttcagttgcgcgcaagcgggagcc ggtgttagacattccgaggggagagagg cttcggctccgtcacaaggccccgtaagg cattcattgaaggaagtaattgtcg

SEQ ID NO:	Name	IscB sequence	ωRNAs
554-555	Methanocalculaceae archaeon isolate CSSed10_120 CS-Sed10-C174733, whole genome shotgun sequence GENOME_ACESSION: GCA_003560275.1_AS M356027v1_genomic CONTIG_ID: 125 SOURCE: NCBI Prokaryotes DATE: 2019-05-30.	atgcagaagttatcagtaaaagttaaagacgcaccaggggatgctccacaagttccctgctctgaagtgatgtttaaacagagaggaaactctcagtgatcgtcgaagttactgtctga taacagctccgaagtgatcaactctctggcaagagagaacaggacttgagagtcctgt tataaacatgcgtggagaagcatatagcctactaccctgggatggcacgaagaatgtt gggttctggcaaggcaaaaggtgttaagagagaccattcgttatccaaatgacaaaagct actggcgaaaacaacagctgttacttggggattgatatgggttatgggaagattggta ttctgtgttacagataaagtctgaattgttggcggtaggtaggtggacaacagaaacag ctaaaaggatggaaaagagaagaagtatagggaatactgtagaagtaggatgtggcat cgtgaacctcgtttaacaatcgtaaagaaagacaagaggtgtgccaccatcaacccaa agaaggggtgatacacacattaggttagttgagaaactgtcgcagtggtgccgataacga aggtgagagtgaagttgccaaattgatattcagaagattaagaaccccaacattgaagg agtgaatatcaacagggaagtagtatgtgttaccaaaacaccaaagaatacattctattca gagaanaaggtaaatgtcaactgtctctaatgttgggaanaacagataggtgggaa ctccaccacattatttcgagaatgagggggaacagacaagaccgacaacttagcact attacataaaaaatgccacgaaaagttacataaaaaaggattaaaactaaaaccgaacag acaatataaagcagagacgtttatgtcaatagcaggtggaaaatagttgatgttaagg gaaaaattttacaccgaacacactatgtttataaacaaaagttaagaaaacaggttaa atttggaaaaatctcacgaaaacgatgccttgttatagcagcggaatgggcaagtaa gatgtaagaattacatctgttcaaaagcacaggaataatagatcttgggatcatcagaga aagggtgtgcaccatcatcaagaaaacaaagatacaagatacggccaaagatttagt gaagattaatggagaatggaaagacaaaagggtacattgtaaggagaacgtgttga ttgtagaaggaaaatcggttaattgtaagaatagagggaataatacaatttggaaagttt caatctctccacatcaagaagggggttacttggaggttaacagatga	gtcaatcccccatgactaaagtacaggggct tgagccgtgagcgttggggcaattggttga ttaggagcgcaataaatgcagaagttatcagt aaagttaaagacgcaccaggggatgctcc acaagttccctgctctgaagtgtgtttaa acagagaggaactctcagtgatcgttcga aagtactgtctgataacagctccgaagtggat caactcttggcaagagagaacaggacttg agag
556-557	Ga0265294_10038476 GENOME_ACESSION: MunlanPumphouse3_2_ \$ F_3300028602 GENOME_ID: 280699 CONTIG_ID: 38475 SOURCE: JGI DATE: 2019-06-13.	atggtttatgtaattaataaagatgggagtccttaagtccctgttaaacagcaatagcaaga ttattgttaaagatgtaaaagcaaatgcataagaagaactccattfacaattaaatactat atcatgctacagactatatacaagatttaactttaggaaatgatactggtagtagtaaaatag gaagtgcggtagcaaatgacgaaaatggaagtgtattatatacagaagtgaagaattgaa tgaatatactgacaaaatggacaaaagagctaaatataggagaactagaaggaataggaa aactagataggggaacatgatttaataataggaaaacagcattaaaaagatagatttt ctctacaatgacaagtaaaatattcacacttaaaagaattgatttgcataaataat tccattacaaatataataagaacagcaacattgacccacatgcttlaaaaaatccag cagtataagtaataaaggctttatcaaaaaggaaactaattatgatttgcgaataactaaag cttatgattaaagtagagataaataatatttgcagattgcaagggaagtaaaagacagc agattagaggtacatcatagatttttgaaaaaatgggtgctcagatgatgcagaaaattt aattacactttgtaaaactgtcatgatagtttgcacaagagagaattaaactaaagaaga atggtaaaatgaaggacaattaaaacacgcaacacagatgaatagataagacaacaatt attaaagacagctaccaaatgtaaaagaacacttttgattt	gtgaactccccgactaaagtcggaggctt cttaaagttctaaagttaccagacttagtaga gaaatttatactacaatagaatgttagatgac ctttgggtgacgcatcagacattgtctctatc gtatatactaaagtagagttggagggtaaag ctccgtgtgatacatgtgaagctattttatg tgcgagatgaggacggattctatatgtgta atagtatatagatcgcataacctgctttaaag cagagtggttatccgaaaggatttgatttag g
558-559	wastewater metagenome genome assembly, contig: NODE_7314_length_445 1_cov_2.613512, whole genome shotgun sequence GENOME_ACESSION: OVQC01.1 GENOME_ID: 11612 CONTIG_ID: 7313 SOURCE: WGS DATE: 2019-01-27.	atggctataacttacgttttaaaccaagatggacaaccatttgcacaccactaggtgtgg caaagtaagacgacttaataactcaggtcaagcgctgtgtgacgcaaaagacctttcact attcaactaaattatgagcccaaaactaattgttttaagacttaacattgggtgtgagca ggctcaaaaactattgtcttcagtcactcagaaaatagagaatactatgcttccgaagtc aaactgagagatgatgtttcctaactatttagagctagaagtaactctagatatacacgtag aaaccgtaaaatccgctatcgcgaaccacggtttaaacacagaactaaatcaagccaaa aggttggattgcaccgacattaaacaaaagttcagacacatgtagatgctgtaaaaga catcatgagcttttaccataaaacaggttaattgttgagaccgctcagtttgatactcaaaa aatggtaaatcctgagatttcaggagtagagtagatcaacaagggtgaactcatgggtatcac atccgagatgactgttgtagaaaattcaagagacaatgtgtctattgtaaaaagaagaatat tccacttgagattgaacataaattccaaatcacgaggtgtgtcaaatagatgaataaattt aactatgtctgtcgcgaatgtaactagcgtaaagacaactgacagccgccgaatttggc ttcccaatgttcaaaaacaagcgtcaagggtcttaagcatgcggcacatagacacttatt cgttaagtccttttgacgcatgacaagagcttggaaattgatgtaaaagaacactttggatat atcacaagaanaacacggatagacgcccggattagagaaaacacatgcgattgatgtag atgtataacaggcaatgtcaatgtacaaccaattgaagacctatggctcaagaaaaagtg agacgccataaccgacgatccacaagatgacgctatataaagggtggagtcacaacgag caaaacaggcaccatagaaacgcaaggttatagacttttgacctgtgtctgttgacggg gccttatgtagtattcatgctcgtcgtgtaaaaggggttactctaaaacgcagtgctgat ggaaacagcctgaataaagcccatcaaaaataacctttgttgcgcatcaacactcttatat acaagaaaagggtgccaaatactggagtaataacattgatgaataa	tccatcatcaataatcaccacctaagtaga gcttctgtattatcaggtgggttttagtga acaactaaacctgattgattagccttagctt acatagactacgttagacttgaataggtacc ttgggtgtttatcactagctccagcgtctacg gttgaagcattaaacattttagtggtagaaa tagtgtgcttagcgtlaaaccagtcataaca ttggcgaagggtatttttaggtcctgtgtccc agcttaccgcataaaacacgaaaatcacag aaaggcacatctagatggctat
560-561	Ga0206225_1000096 GENOME_ACESSION: MancosB2_2_ \$F_330002 1064 GENOME_ID: 32426 CONTIG_ID: 95 SOURCE: JGI DATE: 2019-01-27.	atgagtggtttgttttagacaaaaacaagccctcatgccgtgccatccggcacgca caagggtattgttgcgaagaagaaagccgccgtatttggcgtgctcttttacgattattt taatcaaacgagcaaaatgggatttgcgaagatcttattccttaatttgcctcggcagtaaaa cgacgggagtggtctcgtcgcgaattttagtgtagcgtatgttcatgttggcggtcat ttaaacaccggggagccgctataaaaaaactctggatcaagaagggcgcttcgcag gggagaagatcgcccatcacgctatcagctctctgtttgtaaaatcgaacgcggcc gtccggatgctcggccatctatcaatcagggtagatcaagtagctcgcctgggtt	gtcaaccacccctccctaaaggaggagct tgaagaaataaacaagctcgaagttgacc agccttaacctttagagaaatcaaaagctactg taggagcgaatatataggtaccgtgggatg ctctccagtcaccgctctacgtaagtggt taaacaggcgtacgggttaaacagtgctg cttacacacaacccgtccataacattggcg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		<p>gcgactctctcatagccccgtaacatctgttgcgttgaacggttcgttgcgatatgca aaaacttaagaatccggaatttcaggagcagcctatcaacaaggcactcttttggctatg agggtcgagaatatctctcgaaaaatggggaagaaatgcctattgcgaaaaacg gatatccgattggagattgaccacatcgttcccaaaagtccggaggacaacacgagtc ggaaatctcactatttctcgagaaattgtaacgaaaaaaagggaataaggcacttcaa gatttcttaaacacgctgtaaagggttgcagattctttctagcagcaaaaggactctaaa gacgccgcagccgtaaatgcctcaaggcttgcataaggagaagctctgtccgttctggc aaaccgatttcttactggagcggagggaacgaagcacaatagacagaaacaaggatt tcctaaggacattggctggatgccgatgcgtagcggcattttgtctcgtgacaattccg caagaaatttccgtactggaaatgactgcaagggaagaggctctcgccaaaaatgctta gtcgtatcgttttgggttccaagaagtgtctccaaagcgcacaaacgagcttttggcttca aacaggtgatctcgtttctgcaaccgtcccttctggaaaaagcaaggacattaccgagg atgtgttctgttaaggagcagaggtaatttaattcaaacctcatgtggagtattcaagg aattcacgccaacactgcgtcttaccgacgctatggatggctattcatacatacatttaa aagaggagcggcacttctccccggcctaaaggccggagttccgtgccgtcaagagg atga</p>	<p>aggagaccattaccgggaaacggagcg cgggtaaccgaaaaaggatttttaaatgag t</p>
562-563	<p>Ga0335049_0000303 GENOME_ACESSION: TYMME1r0156_2_\$F_3 300034272 GENOME_ID: 282112 CONTIG_ID: 302 SOURCE: JGI DATE: 2019-07-18.</p>	<p>atgtcaaaagtgtttgtttatgatacacaacaaaaacctctagatcctgttcatcctggacaa gctagaaggctattgaatcaaggaaaaagcagcaatcttctgtcgtatccttttactatcatt ctgaaaaatggaggttaattgatgttaattccactccggttfaaaatcgaccctgggtcaa aaacaactggattgtttagttaacgaccaaacaggagaagtaatttgggtctgaatt aaccatagagggttttctattagagaattcttaagcagtcgtcgtcaatgaagcgaagcc gaagaaaccgcaaaactcggatcgcgaacccgtttttcataatagaagcgaagcagacg aaatgggtgccacctgacttaattgacgaatttacaatctctgacttgggtgaaaaagctg aaggagattgttctattatggtatctctcaagagttagtaaggtttgataccaaagctattc agaacctgaaataagtggaactgaattatcaacaaagggaaggttagcgggtacgaggtt agagattatcttttagaaaagttaacgacaatgtatctattgttgggttaagatactagatt agagattgaacatttaacccctcgttcaaggccggtagcaatcgttcttaatttaggaa ttgctgttcaaatgcaatcagaaaaaagggaacaaagattgtatcaggttcttaaaagg gaaacaagatttagctaaaaagatttagcaagagcaaaaaaacgttcttctgactgctgt gctgttaattgacaagatgggaattgttaataaataagcacttgggtttacctgtagag gtaggaaagggtgttgaactaagtttaactgttgcgtcagaaatcgtgataagactcattgg ctgtatgcagcttgttaggaaagtcaacacccaataaattatcaaaaggaattaaagccctta ttgataattgctacaggtcatgttgcagacaatctttagaacggacaatacgttttctct tctagatatgtgcctagaacaaattgtttaaagggttttcaaacaggagatttgcgaagcc ttttagacatcaggcaagaaattggagtatatacagggcggttagctgtacgttctagtg tagttttaatatctcaacagctaaaggcttaattcagggaatttagcataaataattgtactcata ttcacaaaaagatggctattcgtatgcaacatag</p>	<p>gtcaacaacccaccgctaaagcgggggctt gaaaaagctctagttgaccagactaaggtag caatcctcatgttagtggcaagcgttaaggt cttaccttcagatcgtagctagctgaagct ctaaaactcaaaaggtaaacaggtctaaagg gttaagccagtgcttttgggatagaccgac cactaacattgtcgaagctcacatcacctta gaaataggagactcccaaggttatgtcaaa aa</p>
564-565	<p>Scytonema sp. HK-05 NIES-2130_Scaffold_20, whole genome shotgun sequence GENOME_ACESSION: GCA_001904675.1_AS M190467v1_genomic GENOME_ID: 113419 CONTIG_ID: 92 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.</p>	<p>atgcataaattatgttttctattgacaaaacaaacacactcaatccagtttccagca cgagcaagagagttactacaaaacaaaaagctgctgtatcgggtttaccggttgaat cattctgaaacacgtgttgataacccccagagcaagcatttaactcaagctcgacc cggtagtcgattcacgggtattgcaatttggatcaagacaaggttgggtggcgagaa ctggaacatagaggtgcaaaatgaagatgcttagaatacaagcgtctttacgtcgca gocgtgtaatcgcaaaactcggatcgcgaacacgtttcaataaccgcaagcgcaaa gaggatggctgtccatcgttgcattgtgctgtcctgacgattgagacatgggtaaaac ggcttctgttattctctatcactcaaatagcgatggaggttgaattgatactcaaaa aatgcaaaaccagaaattgatgtgttgatgatcaacaagggaactcgtggttatga ggtgcgagaatactgttgaaaagtggggcagggaagtgcgttactgtgatcatgctgg tgtaccctccaagtgagcatattcaccttagagcaaaaaggaggttagcaacagatcag taatctaactttgtcatcggaacgatgaacacaaaaaaggaaactaagctataggcgag tttctcaaaaagatgttctaggttagaaaaaattcaacgacaagccaagcaaccattaa aagatgcagccgcagtcgaatgcaactcgttgggaattttgcaccttaaaaaacatcct accaccacaacgggaacgggtggacaacgaatataaccgaacagcgttgaacta cccaagcaacattggattgatgcagcttgttggagaagttataaatttaaatctgttaacc caacagccactgaagattaaattgactgttggggaactcgtcaaatgtcggcactgat aagtatgggtttccgacctgacagagaacgcaacaacattcatttgggttcaaaactg cgcacattgcaaaagctgttgcacttttgggaagaaggttgcataatattgtcgcgtg ttatgccgaagactggcagctttagacattgcaaccattagtggtcgttggcaggtattag ccacagattttgtatataaattcacaanaaggacggctattcgtatggatttttag</p>	<p>tcgttaaatcccaacccatgctaaaggca aggggcttgcgttgactacctgaccaga ctactcgtttagggaaggttaaaagacctta cttcaggatgctgtgactgctgaacctaga accgaacgattaaacaggttacagagaat aaaccagtgctgcgctgatagaccgacc ctaaacattgtcgaagcaaaccttaccgcga aggagtatgcagtaaatgcaaaattat</p>
566-567	<p>Ga0120161_1001269 GENOME_ACESSION: 3300012005 GENOME_ID: 239592 CONTIG_ID: 1268 SOURCE:</p>	<p>gtgccggagcgtttccggtgttttcaggaaaggtgcattttgtcgtgttcttggatcag aaactgcgccccttgatgcggtgcagggaagcggcgcgcaagctgctcgtcgtg gccgcgcgcgcaattcggctgtatccgttctgtattcggatcgttgaccgcgcgcttga agattccgtgttgcagccctcgttcaatgaattgatccaggcagtaaaagtcactgcttgg cagtttctggatctgaagcaggtgacgatgatgtgtcgtcgtccctcatgcatatc cagttctgatggaactgttgcacgcggagacagaatcaacaatccctacatgcacgca gcagcatgcgtcggcgccgctggcaacctgcgttatcgtcaaccgcgcttggataatc</p>	<p>gtcaactaccccgccctaaaaagacggagc ttgaaaggcgaaagcctcgtatgttcagggtt gaccagggttagtgataatcggcgcaactc gatgttactccgttgaatgaggtacaaaac caacgccggagcgttccaatcccgat agaagtgcgaaagcatttcggttgaattgc gatcgcagacaagcgacagggcaaggca</p>

SEQ ID NO:	Name	IscB sequence	ωRNAs
	JGI_Dark_Harvest DATE: 2019-01-27.	ggacacgtcaaacagggtgttgcgcctcgtctccatcgctgctgacgatagaa acgtgggtcaatcgatcagaacatggcaccgatcacagatcgcgcaaggttagt gcgggtcgatagcaaaagatgcaaaacctgaaattgcccgtatcgataccagcaag gcacgctcgcaggctacgagctgcgcgaatacctgctgaaaagtggaaacctgagtg tcttattgctggtcaaaaggacctcttctgagatggaaacatcgttctaaagcgaatgg cggtagcaacctgctcgaattgtgctgggtgtgaatctgcaaccagaagaagg cgcaaaaggataccgcctctcctcaaaaatgacccgattcgattggcgcaaaatttga catgcaaaagatgctctacgcatgaggctgagttaacgccaccggcaacgctctact gaaagtattgaaggcaactgactgcccgttgaaccggcactggtggccgcacgaaat ggaatcgacgtcggttggcatgcaaaacctgctctgacgagctgtgtgggtgt tgttccgatgttagcatcaatgcgcctgcattacaattacgtcatggccgcgga accacagcagaactcttttagaccagatgtgttcccgccactacgctccccgaaca aaaactttcttggctcaagactggcgatgacactgctgaagtcacaaaggcaaaa gaattggctccataaaggccgcgtcgcggtacggctcttgatctcaacatccagtc tggtatcaaaaggcatgaaagcagttcaaggcattctcataaggattgcccgtatcgcaa cgcccgatggatataactattgtggaagtcggttctgacaccaaaggcaaaaggca tcgactgctgtgacgtggctccttgcctaccgcccgaaggacggtgtatctcggagc atatctga	cgaaacggatcgcaaccgggtagcaaaagc accgctgcttaaaaccgcttacaacattcc cgagggtgagcgggtattgtcaatctatcgccg gtacccgtcactagcgctgaagcgcattg cgccgtgaggcgtcttccgtgcccggacg gttccggtgttttcagaaagggtcattttgt cg
568-569	Cyanothece sp. PCC 7822 plasmid Cy782202, complete sequence GENOME_ACESSION: GCA_000147335.1_AS M14733v1_genomic GENOME_ID: 40504 CONTIG_ID: 2 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgtcaaaatttgtatttcttagatgcgaagaaaacctcttagtccctgcatccatcagt agctagaaaacttctcaagcaggggcaagctcggttttaaggcggtatcctttactataat tttgaaaaaaagatgccaaaaccacagaaactatcaaacataagctagaccctggaa gcaagacgactgggattgcatgtaacaagcaagcttatttggggagcagaattaat tcacagaggacacagatataaatttgttaactggaagacaaatctgctgtagccgt agaaacggaaaaaccgatacggcaagcgcgatttcaaatcgtactgacttaaaaggtt ggcttccctccagctcctgacatcgtagaaacgacctgacctgggtaaaacgtattt caaaattgttctatgttactaatattcggtagaactgttaaaattgatactcaagctttagataa ccccgaaatatacaggtataaagataccaacaaaggagaacttttggatagagattcgaga atatctactgaaaaatggggagacgatgcgttactgtgttattaaagatgtaccgttaga agttgaacatattctagcaaaatcaaaaggagatcagacagatgttcaaatttaactattc ttgtgaatttgaatcaattgaaagggaatcaagatataaagattcttatcaaccaacct agtttactagaaaaagattctaaacaatctaaagcaatcccttaagaattgctgctgtttaa acaactcgttggcattatttaacaagctgaagaacacaggattatcaattgaaacaggaa caggaggtagaactaaatataacagatgtagattaaatctagaaaaagacattttatagat gcgggatgtgtcgaaatcttgaagcftaaactatataacaggcaaccattgttaattaa agcgacagacacggaaatagacagatgtgcggaatcaataatacgggttcttattc acatcggtctagaaacaaattttacaaggatttcaaaactggagctcaagttaaagcagta gttacaaaaggaaaaaagggtgttattttaggtcgagttttagtcgagcttcgggccc cttcgacatagcatcccatcaaggcagaacaacaggaaattacctaataattttgcacagct attcaaaaaaagatggatataactatgattttaa	gtcagcaccccgcatgaaattgcccggct aggctacgctcgtcgccgtccatgcccc cagtttcaagtttagtgaccagccctaaagcttaa atgactacgttttccaggctataatctacata atagctgtctagttttagctctattgttaacaa ttaacagttttacgaggggtagacaggtgtt gttaacccaacaagccttaaaaaattggcg aagcgaaactttaccgaaaggagagacag aaatgtcaaaattt
570-571	Ga0193910_10667 GENOME_ACESSION: 4E_4869_18_06_2_\$F_3 300019455 GENOME_ID: 21018 CONTIG_ID: 666 SOURCE: JGI DATE: 2019-01-27.	atgaaaagtctatgtcattaacaagaggctcgtaaaattgaccttgcctcccgaagg caagaaagctactcgctgatgaaaaggccaagtcagtgaaacgactaccgtttacgattc aactgaaatgggactcggaagagaattatcagaaaataaccgttggcattgataaaggta gccatacgacgggttattgtccatagccaattggtgaaatcttaattgtagcagctatatcaac catagaacagatgttaaaaagaagatgcaaggacgtgctgtaactcgacacaaagaag atcaagggttgggtatagaaggacacgggttgataatcgcaagccgcaaaagcgtcttt agggttcccgtctcaatcaaaaccaatgtagaagaggttataaagacgataaacaattac ctttaccgattaaagtgttattatagaagatgtcaaaattgataattcgcaaacgtaagatg ccaaactaaatggttaattatcaacaataaccgactgcatgaaatctcagactgggt tcgagggttaagtttttgaagaaaaacttaaacctcgaaggttgattcgagatcatttacct gccagatttgcaggcaagcagacaccattggaagcacaccacatcattatcgaaaa aatggcggttccgagacgattaaaaatctgataacgggttggctgcaattgtcatgacgggat tcattgatggaacaatcacactcaaaagaagggttgatggctttaaagatcag	attataatagttacctaaccaattgattttaa gtattgtatggaactaacaggcgctaataat tagcagaaggttagataggttataacacctta aatgttcaatccagtttgaagccctgtgtttat ctttaagggttagtggaacacgaaagttag ataagcgcaaaaagcctttttaaacagcca aggattacttcaacttaaccgagggccta gccactattatgaaa
572-573	Bacillus sp. AFS014408 AFS014408_98_D11_Co ntig119_140514, whole genome shotgun sequence GENOME_ACESSION: GCA_002557915.1_AS M255791v1_genomic GENOME_ID: 141756 CONTIG_ID: 20 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgcgtgtatttgcagaagtgtgaaggagaacccgtcatgcttgcagtaatcgcaagg cacggctcttctcaaacaggaaaaagcaaaattgtgaagtacacaccatttaccgattca actcctatagccaccgggtgaagcgggtgcaaccggttacaattggcggtgtagcggagc aaagcatatcggtattgcatctactgcagataaagtgtagcaaaaggaaaccacca gctcgctcaagacgtcaaaagaaatcttatacgaagctacattacgcagaggtagaag acaacgaaaaacaagatcgcagaagtacgttttctcaaccgaaaaaagaagaaggat gggtaccgccatcgattcaagcagagtggtataacaaattcattggattgaaacatttgc tcgttattaccatctcaaaagtgattgtgaagtaggggaattgatgcacaaaagctaaa aaaccccaatatacaggacagaattatcaacaaggagatgatacttttggcttttggaaatc agatactacgtatttgcgagagacaactacacctgtcaaaattgtaagaaaaagggtgat tttgcatagcatcatattgaaacgatgcaatggcggttcgaatattggcagataatctgt aaccgtgcatgatgaatgtcatcaaaactcatcaagggaacatcaaacacagtttcaag aaagtcaacaataataaagaactgtttatgaatatatttgcggctgcaaatcatgaatcgt	gtcaactaccaccgcttaaacgctaacacg ttttgaagtgggggctgttaaaagctctggt tgtctagccctcagttcttgcgtgactccgttcg taggttgcatacccaagatgattccctagtt cttggctctatgggtgctctgtaacagttctga ttgggaaggaaaggtaaccacatgccttct tgcatagaaagttgcctacacctacaacatt ggcgaagggaacaaactcttagggaggga caaaatcagcgt

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ttagattgtgacattacgtatgtagctacaccacccaagagaaaaaacttggattag ccaaaacacatgcaaatgatgccattgcgattaccaatccatacaactacaagaatacg atcaaaagggtgaattttgcatcaagcaatttagaaagaaaaaacgctccct	
574-575	Ga0315282_10014819 GENOME_ACESSION: YL17G07_20_MG_2_\$F _3300032069 GENOME_ID: 281633 CONTIG_ID: 14818 SOURCE: JGI DATE: 2019-06-13.	atgcaaaaggtagactagaactccttcgaggcctacgtgtatcgggctacaagaacgg caagatgcctccctagcttgcctcttcggaagaaccactaaattccacgctcccaaaa gggaactaagcctgatgcacattgtcgaagggaacacgacttggaaaagaccgtaatc ggtagagttccagtagtgcagcagatggaacacctttgatccctgtaaacatccaaag caagaaactacttgaaaacggaagcaacaaacaaatggaacaaactaagcataattct acctgcaactacactcaacctaagaacgcaactcaaccattagcattggcggttga ctgtggaagcaagtctgaaggcttttctgctgcggcacaaaagacactgtgtaaacatc atgtccaaagcaacacatgggtcaaaaaagcagttgagcaagaagacaaatcgcaa agacaagacgcaacagaaaaccagacgacgtgaatgcagggtcaacatcggttagc ccacaaaaatatattccccctcaacaaagcaagggtggacacgaaattgcgtgaatt cgccagttggagaaaattttcccaattcaacagtagtgcgttgaagacgtcaaaacgtaa cccgaagaaacggaacgatggaacaacagtttttcccaatagaagttggcaaacaa tacttctatgccagataaacaactcgtctgtaaatcagggtgtagaacgaaaatgctgc gagagcaatctatgttgaagaacttgaggataagagcaaacctgttttgaacccactg cggtgatgcttgggtgcttgcgctcagagacgggtgccaagcagcctacaacaagaa gcctttattatcttcttccgctgcgttggcacagacgacaattgcacaggttcagccagag aaagcggaagaaagaaaccttacggcggaacatgctcgttggactaaaacgcggaa ccttagttaaacatcgaaaacacgggttctgtacgctggcggaacctaaccggttaact cagcctgcacagcgtcaaaacgggtgaacgctgaccaaatgtgcaaaaaagaagag tgcaaaatcttgactaaaatttgcgttcgaactcaaaccaattcccccctcaagcagg agggtacccttggaggcttttga	gtgtattccaccggaaactcgaaggcgaag agaaatcatgacgagccccattgttgcatt gaccaaaccgctagcgtgaagggaagcaac acaataatgttccaccgaaaaataaagaa actaatcgttggtagcaaggaactagtt ggcataatcacacaccgacatcattagca aaatgccaaacatgatgtccacgctggaag cactgttagcaccgcaaaagcgtattgagc tccactgcaaaacgttgaagcatttaagt gctacgcaataacatattaggagcaactgtg aacttgcactaccactcaacaagttagtg g
576-577	TPA_asm: Chlamydiae bacterium isolate HyVt- 353 HyVt- 353_k145_730505, whole genome shotgun sequence GENOME_ACESSION: GCA_011056925.1_AS M1105692v1_genomic GENOME_ID: 327094 CONTIG_ID: 88 SOURCE: NCBI_Prokaryotes DATE: 2020-04-22.	atgcagaagtttagtcggaagagatacacacactacggatactccactagtcctgac aactgtgatctgtcattaaacagagaggaaactctcagttgtgtagattaaaaactcctt aacaaccccgaagtggataactctagcatagcagaggccatagcggcatatgtgtatgtg atttcaagaatggaagcttggatgcttctcaaggccaaagcctagaaaattatataa ggaggaaaaggctctgattgttgcacaaacacttactattaaactgttttcaaatcgga aatcaagttcaaaaaactcttaggtattgacccaagatatgaaaatttggttatcagct atttcagaaaaagggaattattctcagctgaggcaaaaattgaagaacaaatttctaaacta tttagcaaaagaaaaagctatctagaaatagaagaacaaactctgtataggaacca agatttttaataagaaaaaaactaaaaattacctccagtttagagcataagctagattca catcttagtagtgtaaaaagggttctctctttttatctatttcaaaaatcaatagaaagtagc taattttgacattcaaaaattaaaaatcctcaaatagaagccttcaatatcaaaaggagaga ttttagtgataatcaaaaattaaaaagcatattctgtagagcgcgagatggcaaaatgccagc tttgcctaaaaaatcaactaaaggcaattcttttaggtatcatcacattttccaagaaatga gggaggaacgaacaagactaataacatcgtgtctacatgaaaaatgccatgataagt acactccaaaatctactatatttataaaaagaatagggcattttaaactgagacttttat tcttaacttgcgttggaaactagtttctgaattgaaaaaacttgcagcgatactgtcctatcttt ggatatatacaaaaattaaaaaattctcctcaagctcgaaaaagaccatcacggatg catttggattgcgaatggaagctcagaaaaagagctcaaccttcttttttgcaaaaaa gaaaaaataatcgctctacaaatagaagaagggtttaaattatcgattagacgacaa agataaagattcagccaaaagatgaagtgcagtgaaaaacaaaaatagcagttata ggcatttttaacaaaggtagttggtaagagttagagacaaagctaaaacttaattttcag atttcgagagttgaaaaacattttataataacggctgcgaattcattcctctaaagaaa gggtttcttgcctataa	gtcaatcacctctccctaaaggaaagtggtt gaaccgtgagtagcaggggttaactgttga ttagggggcaaaaaggaactttatgcaga agtttagtcggaagagatacacacactac ggatactccactagtccttagcagcactgtgat ctgtcattaaacagagaggaaactcagtg tgtcagatttataaaactccttcaaacacccc gaagtggataactctagcatagaggggcc atagcggcatat
578-579	Ga0223845_11964739 GENOME_ACESSION: 3300021387 GENOME_ID: 241586 CONTIG_ID: 316 SOURCE: JGI_Dark_Harvest DATE: 2019-01-27.	atggaatacaagaatgtatgtatgaattgataaaaatgatagccgtgcaatccgatcact catggaaaagccagatattttataaacaataatagctgtgtataaaaaaccatgatccattt gttatcaaacgaactgatgattttaagggttttgagggtgataatcattatgtcttaaaaat tgacatgggatataagcatattggaatttcaataactctgtaatatgatgaggaatttggg caagttgaattgttgaagggaatgctctaaaagattagcggaaggggcaagatatagaaga ggggcgagaaatagactcaggcatagcgttaataagaacatagataattaaacaattaaag aatccaaattacaaaaatggcaatgaagacggctgttgcaccgtcaataagacataaaa atggatactcataccagactgttgataaattgaggcggtgattcctattgaccgtattgaa ttagaagtttcaaaacttggatattcagcagatgaagcagatctgaaagattatgagatgcat ggtagctgattatcaaaatgggtgaatgaaggatagataaattgaaaattatattaaagag cgtgacaaatcacatgccagttgttgaagaaaaaacaacatctggggaagtacacat attattccaagaagtgggtggaagttaacgcccaggaaacttatatacctattgtgtaga atgtcactccaaatgtcataggaataataatgataatgactatttagggatatccaaagaaa gagagtcgatgggtgattttaaagaagcgactttcatgaatgctgttcgctgggcaattttag atgcttttaggagagcaatttcgatgtatgcatttttgatacgaacaaatagaatgc aatgcagctaaattgccaaaattcatcaacgatgcagtggttatttaattcatttaataata cttctcttcaaaatcattatatacattaaacaatcaagatgataataatggtcagatgaagat ttttttagtcaaaatattatgataagactggcaagtagcaagcgggaacgattttaa	ctaaaaactcaattccaaaacaattttagaac attataaaattatgtggcaaacaaaaataga ttaagagaagttctatttgaacaatactacg ttattatgagataatattattgttaactctag caataaacatatacgaaaagttaaacatcca aagccatacaataattatgggaagtactgag cagtaatgaattgaatttattacaagctcgt ataacattatcgaaagatgaccgctattta gcgagattttattataaattataaaagga taaataatggaatacaagaataat

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aaaaatacacaagagggtaggcttaagagatctactagaaaagaagataataaattta aggggttttagacaagaaaaagctcaagctggaaatgaacgcaattctgcatagttattgt ttaaagcctggcgatttaattaggatcattaaagataaagataatcgaagtaatactatg caaaaaaggaataatggtttataatgtatgcgataatccagatgaggatgcaactagcg aacagttaacattctctataaaatcagacgaatacgaataaataaaactacagggaaagt taatcgcataagattgttagaacaagcggtgttaattgttacaggtatgctgtgtga attgaagaaaaacatgtggatcagtatcatattaaagaagtagatgcagcagaaaaagc agctagatcaaaaagaataaagcaagaaaagacagaaaaagaaaaagaaatgc attgagagaatccaatgaaaccgctctttcgttaa	
580-581	Ga0376681_0131425 GENOME_ACESSION: CalMonMetagenome_5_ FD_\$F_3300036805 GENOME_ID: 312378 CONTIG_ID: 131424 SOURCE: JGI DATE: 2020-04-07.	atgtttgtaccagtcataatcaggaccagcatccactgatgccaacgacccagcacgt gcacagcgttgatcaagagtggaagccacggcttctggaaaggtggatctctgt gtacggctcaacgtgatccatcgcccacgtctcgaaccatcgccataggaattgat cctggaagcaagcgtgagggtatgtgtcattcagccctgcatacctacctgaatatcc aagccgaagccagggatgctgcaaggatgcagaaaaggaactgcaccaggtatgctg ggacgcgccggaatcgaaaacgcccgtccgcccagccagccagaatcgcaagcag agcaagcagaagcttccccctcgaccagagtagatggcagtggaacttcgggtgct cgtctctgaaccagattttccggtgagtgcttgggtggaaagatgtgctccgtgacg aagaagggcaaacgacgtggaacagcgttgcctccctggaaagtcggcaagcact ggttctatgagaacttaggaacttgacactgtagcaatcaaacggggtatgaaacca aagcttgcgtgagcagttaggcttcaaaaagaccggcaagaaattggccgaagtgtgg gaagcccatgctgtgatgctgatactggcttatagccgactggcggaagcagacc cctgataatcagcagctggtctgcatggtgcccgtcaactggcatcaccggcaactccat cgttcaaacgggagaaaaggggcaagcgaagccctacgggtgtacactctcgcaag gcatcaagcggggaacgctggtcaccatccaaggtggggaagggcagcagttgggg gcactatggagcggcaagctcagtttgacgatcctcatacgaacagcgactgacgaa acggcgaaaagtcagtgatgctgcccataagttgtacgctggaggacgcggctgtg cctctgctgctgacccccgctccctgccccaaaaagggaagactgcttctccccgg ctgaa	gtcagacccccgcacatgaatcggggg ctgtgaagcgaaccgactgccatgggg aaccaacgagggcaactgcgagactagct gtccaagctccggcaccggggcgttgaa ggaagctgaatcacctgggcatggcagcct atagaacgagcagacgctccctaatccc gctcctctcagtggtcagcagcgaggga acatattccccgcaaggggcttatcgaca t
582-583	Ga0315902_10061306 GENOME_ACESSION: LakEriMetagenome_16_ \$F_3300032093 GENOME_ID: 280599 CONTIG_ID: 61305 SOURCE: JGI DATE: 2019-06-13.	atgcctacgcacaataagtcacaagacaccagcagatgcttctcagtcgtgctc tagaatcgttccgatttccagaaagtggaataaggcaacgaagcggctggtgctgctc attgtcaggggagagcccttgcctaaaccgcaagggcggtaccacctggcaacaggt gagactgagtcggtaacggctccagcatctctcgttggccaaatcgaagacacccct catgccctgccatccagcacgggctagagaactaataggaaacggcaaacggctgatct atcgccaccaacccctcgtcattcggcttaccgttagaacggaagggcaatgtgcaaccga tccagctcaagctgacccggcgcaaaaacgacgggcacagcgtatgtgttaaccgct accgctccagcaaatggacaagtagtgcacacattgagctgaatcaccgcaaaaga aaactggaagaagcggatggcgcaacgcaaaaactgttcggccgctgcggcaccgc aaactgcgttgcgcaaaagctcgttctcaatcgcgtaaaattgggaagatcgacc atcgatcaaatcaaccctcgaccaactcggggatggatcaatcggtacgcccgtgg caccatcactcgtatgtaaggaaacgctcgaattcgtatgctcagaaaattcagaaccc gaaatttctggcgtcgaataccaacaaggaacgttggctgcttcgaagtcaagaatac ctctcgataagtggtgcagaaaatgcgttactgtgggataaaaacccccactcgaa atcgagcatatcgcccgaatccaaggtggatcagatcgagtcagcaacctaccctt gcctgcaactcgtgcaatttggcaaaaggaatagatgttgcgatttacttgcggga aacccggccgctcaagccattctgctaccgcaaaaagccgttggcgtatgtgccc accatgaacatccttaaacctagggtgatgcaaatgctcacgaaactggctgcccgtga caactgccacagcgagcatgacaaatcaaccgcaaaatgttgaattccaaaactc atgcgttgcgtccgcttctcgggccaatgaaaaaatttgaagggtggaaccaat ccatctcaaatcaccgcgacaggacgtggttcgtatcaacgcactcgcaacgacaat tcggatttccaaaggctacgctgccacgaaccaatctgtccgggcttcaaacggcg acctcgtcttaactccaaaggaaatgtgctgcatcgctacccgctcctgggatacttgc gctcgcaacggttgatgcagcaaaagccacatcaaccactcaaaactgcgcccgtctcc aaagagccgacggctaccaataccaacaacaacaaaaatctcgcagcaagctgc ggggttctaccctaccgttaa	gtaaaatccccctagcgaagcctagggcg ttttccagactaagacatagcaatatgcctac gcacaaacataagtcgaagacatccagca gatgcttctcagtcgtgctctagaatggt gccgatttccagaaagggtgataaggcaac gaagacggctggtgctgctgattgtcaggg ggagagcccttgcctaaaccgaagggcg gttaccacc
584-585	contig_9137_[cov=16] GENOME_ACESSION: mgm4527699.3 GENOME_ID: 237371 CONTIG_ID: 9136 SOURCE: MG-RAST DATE: 2019-01-27.	atgcagaagttatcagaagagttaaagaacacacaaaggatacttccaaagtccttgc ctgtaaacatcaattaaacaaaggagaaactcttagtgtgatagtatagctgtctgata acaatcccgatgggaatcaaccgagtcacaatcggaacagaaacagagtgatctgttt atgtcttaaaccaagagtgtaaccgttaatgccttgcaccagaagacaaaagtgttgc ttgaacaaagaaaagcaaaagtagttaaaagaagtcggttacaattcaattactattgc tacaggagaacaaaacaaagaatatttaggagtagatcaggttattcaaatgtaggg gtaacagctattacagaacaaagaattattatcagcaacacattcaagttagaacgaat ttgagacttattaaagaacgctctatgcacaggaggttagaagaataggcttggat agagaaccgagatggaaaatagagccaacgcaagaaaaggaagaaatgaatgcctt caattctacataaagttatctatatttctatttgaagattaaaaagttattacatttc aaaaagtagtattagaacaggaattattgatatgcaaaaatggagaatgacaagattaaa	gtaactaccacgccttaagcagctggtt gtaagtaacttacaacgtaaggttgattag ggagcttaaaatttaaaagggtttatgca gaagtatcagaagagttaaagaacacacc aaagatacttcccaagcttctgtctgtaa accatcaattaaacaaaggaactcttag tgttagatgtagtactgtctgataacaatc cgtatgggaatcaaccgagtcacatcgaa acagaacacgagtgatct

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aattatcagtagcaaaaaggcgagatgtttgggttgaagatgaaagcttatgtttcttcaa gagacaatcataagtgctatatttaaatgcaaggttcttcaaaatgaagtagaccataatata attcagaagtcaggcggtacagataatccgaataattgattactttatgtgaaaaatgtca taagaaagttcatctatgtgaattggagttaaacattaaaagcataaagaattaaaatctac cacagcaatgaatgttataaggaaaagattgtgaattttatagagaagcagaagaaca tttggttatgagacaaaagttaaagaagagaattgattagaaaaatctcattctaatgat gcttttgaatcgcaaatggaactaatcaattagaagtaagagtttgaaataattcagaaa agaagaacaataggtgtcttcagttaaaccgtaaagatacaaaccttctataaagaaag aaagggtcaaaaatacaacccttcgatttgggtttaaaggaacaatacgtatgttaag ggaggttcaactatggtataatattgttatgttagtactaaatataaagagatttttaaat agagctattagagaagcattataatcaaggtgttttagtatggaattaa	
586-587	Ga0307377_10069547 GENOME_ACESSION: TR3MG_2_SF_33000316 73 GENOME_ID: 281015 CONTIG_ID: 69546 SOURCE: JGI DATE: 2019-06-13.	atgttagtatatgttaaaaactgtaaaggtgaaccctctatgccttgcctcctacaaaggcta agagactgttgaaaagtggtagaagctaaagttagtgcgaatgcctttactattaaactat tggttgtagtagtcttcaacaagaagtagtagtagtggatactgggttcaagtattat ggatgtgcagctaaaagtaacggtaggtttgtaccagtcagaggtacagataagacaa gacgtttctatgaaaatgaacaagaactcatgtatagaagaacaagaagaagcagaaa acttagatctgtaaaccttagatggcaaatagatcttctcaagaagaaaaggagatta ccacctagcatcagaagtaaaattgattctcactaaagagagaaaaggttattgaatctatt ttacctgtactagatggatagtagaaactgtcaagttgatatctcatgagataactatcctg atgttaaaggaaattgataatcaagaagtagtcaaaagagttttacaatgtaaaagcttatg ttagacatagagacagttatattgtcaactgttaaaggaaaatcaaaagataaagctctaa tagtcatcatattactctagacttataggcactgattctccagataatttaataactttatgtg tactgtcatgacgattttcatgctgtaagattaaagttataataaaaggaaaacgactca aactaaacatgctactgagataggaatgttaagtcacaattgaaaaagcaatggtaggtt ttgaggagacttttgggtatgaacaaaat	gtcaactaccctcccctaaagtttctttaagg agaactattattaggtaagtgaccagac cactaacaggagaataaagttgttagtaaa cgatacaaaaagaattatgcacctgggag ccgtcccaggttccaagctctgaggttacgaa ctaacagttcttaggttacagaaacgtgtt gtaacgataaaccttttgatctaggtcggag gaaagttgaactctataattgctcctaaattat agatacactactctgttttcagtcaggggt atactttatgta
588-589	Ga0223826_10002115 GENOME_ACESSION: 3300021256 GENOME_ID: 241585 CONTIG_ID: 2154 SOURCE: JGI_Dark_Harvest DATE: 2019-01-27.	atgacaatttatgtcttagacatcacaggaaaaccgftaatgccgacacacaagcccggta aagtcagacatagctgaacgacgcgagacgaatcatcgtcaactattatccgtttacgatt aagttaacgtatgtaacaacaactacgttcaaccagtaactctcgggtgtagatgctggaa gcgttcacgtcgggtgagcgttcgacagagaagaagaactgtattccgctgagattg atttgaggagcaagcaaatgccgaacttatagaaaaacggcagcagcaagaggaat gcgacgttatcacttgcgtacagaaaaccaaagattcaataatcgaactctctcgaaaaaa gaaggatggttggcgccatcattgaatcatcgtgtcaattctcatacagaatcgtagagaa tgtaatgaagattctccaatcagcaaatcatcgttgaagttggtattgttgatactcagaa gatttccaatccggaatctcaggagaagatgatacagaagacagatgctggttctga caatcagaagcgttctgacattcagagacaagaacactgtcagcagtggtgttcgaa gaacacattgaagttcatcatalccagcatcagaagatgattgtccagacaggccagat aacttaatttgtttatgtcacaatgtcattatgaacatcacacaatggttggttctaaagaa gttcaaaaataataacaagaagaatgccgttctcattgcgtgacgagcagcgaatctt atcaaaagacaaggtattcttaaaacttaaggaaacgtcattcagatattgttttgagaacg tatgtttatgttacgaacataacagaagaataataacattgataaactcatgctgaatgat gcgtttgtgattcgaagaatttcaacgctgaaccttggattatatgttcaggggatttcaa ttagaaggcataacagaagaatatacaaaagacactattctaaaggtaacaaattgaaaaa gaaccaatctgcacatttgggtgttcggtttacacgatttgatagatgaatgatgggaa agaatgtttcatatcgaagacgaacttctggaatttgcacttgagagatagatggtta gcaaaatacacgctactgctcgttaaggaaaatcaactaatcaacatgagaatagtat aataattaaagaaatcaaaaaggatagtgctgattcttctgctgacgctaaggcgtcgca gtttcctcagcagatttaatatga	gtcacaaccggcgacactaaaatgtcgc ggcttggcttcttaaatgagcggagtcgatgg tgattagcctaagtgaagagtttcaactctt gaactacgttactgacgaaaataggtacc aagggtatgtatgcttgccttgcctctacgt ctgaatgttaaacagctcctgtccgtagggg cagtcattcagaatacaaaacgctgggata acattggcgagggcgtcattacagcgtcgt ctttagcgacgtcgtgtatcccttaattggg aataaaaaataatttcaaaaaacactatga ca
590-591	fermentation metagenome genome assembly, contig: NODE_1217_length_248 63_cov_7.295993, whole genome shotgun sequence GENOME_ACESSION: U00U01.1 GENOME_ID: 16052 CONTIG_ID: 1216 SOURCE: WGS DATE: 2019-01-27.	ttgacaaggaggacacatagcaatacatgtattgctaatttttaataaggagaattctatg aaatgcatttatgttaattaatcaagcggtaagccactgatgcctactaaaagattaggcag ggtgagtcgttggtaaaaaaggcgccagcgcattgtgtataaaaccgtcgtgataccatt caatttaactgtaaaactactaattatgtccaagaagttatcaaggtatgtatcttggatc acttaggaatgtccgttatcactaacaatcaagaagtttatgccagtgaagttactgtaag gtaagcaaacgcacgactaatgcaaaaacgaaaagaactgcgtcgaacaagacgtaat cgtttacgtatcgcaagccacgttttgataatgcaaaaagaaaggatacgtcctctgat tcaaagaagctagattttcaaaftaaagaattaaacgcttgatagatgttttaccatcact aaacgtgttttgaaggctcgaacttcgataatcaaatcactcatcagctgcaaa aaaggttataaaactacgtttgattttgtatgacgggaatcgtgctgtgacgcgctagac ggtaagcattatcctaagaaaaatgtgtatccaccatttagtacagcgacataacggcg gtacaacaatcctgataacctgtttattagcacgtaaaatcatacgcaggtaatcata aaaacggggttttagacaaattagtaaacagcgtcagaaaaataaagaatgcggata ctcgtggagcataatttcatgaatgtgctaggttaaggaaattcctaagttttgattttgccc aacatatggctatcactgctaaaaagcgttaagaagtagcgtattgtcaaaactcatcac gatgatcccttgaattgctggggaactaatcaacaacacgttttgatacatgctttatc gtgaaaaacagcgtctcaataatcgtctttagaaaaattttagatgcgcaatcgtgatt acgtactgttgaaaaagaaaaagggtgctgttttatcatcaggtagaacacgtcgttcttaa	gtcaaccactcctgactaaagtcaggagctt gtaattattacaagctcagttgattgtctaag cattaaattgtctacgttactgcgaataggtta ttaatcatgctaacacgagtcaaaacgttgc tagttgataactttaataggtcattgaaacgc tactccaatttcaagctctcaggtatcggttta aacagagctgaagagtaaggctcagtgatc gtattgaaaaactgtcggataacattgacaa ggaggacacatagcaatcatgtattgctaa tttatttaaggagaattctatgaaatgc

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aagatccacgcaacaatcaacgtgtttccgtgctaacaacttaaaaaggcagaagaa cgattcgtaaacagcattatcaattgcgtcctaagatttagttagtacaatgataaatttat cgagtaaaaggcatgcagaacaatggcacgcgtgtttattaattactctgctaagataa atcagtagccattaaaaacgtaaaatctgtttcataatcaacgggttttatcaacggaaag gagaactacggcattcatctcctga	
592-593	Ga0117909_1085896 GENOME_ACESSION: IMG_3300009320_SF_3 300009320 GENOME_ID: 305879 CONTIG_ID: 85895 SOURCE: JGI DATE: 2019-11-15.	atgttagtccacattttccaaggatgtaaggctttgatgccttgccatcctgctaaggca agaaaattactcaagaagtaaggctaaacctgtcaaggaaagactggatattttacta ttcaattgctctatgttagcagtggtctataacaagaagttgtgtgtgtagatactggc gcaaaacgagtgccaatcgtcgtgtgtgtaacggcaaaagtgtattatgcaaaagagaag atactaaggacagatgtaagaagcagtgctgtataggtcaagatacagacgcacaag gcgaagtcgtaagacagcttatcgtaaaccagattcttaaatcgggtaagacaaaatg ctccaggtgtgtgttaataacgtgcccaagcgttggaagagtgcaagcgcaaaacag gtaagagcaaaaagaagtgatgtatgttagggcgcaactttgtcggcagtgctgtgggta aaaaggggtacacagagaaccgcagtgactgtcccatcagtaftaaatcgtgctttgag tattctcaatgacattcacaagttatctctacacccctataatccagggtgattatgagatg catcattgtatcgcagaagatgacaacgcctttcatagacggcgctcgaatatcagcatgg gacgctctttgctacgaagtcgaacataatctctgtacagttcacaagcataagtcgctt attgcgggtggtgctgtgcaataacatcttacaggttgagcatatatttccacaatcctaag gcggcaccgataaggtaataatctaactatctctgtgtgggtgtgtaacgaagctaaagg aagtatgacgttgaccagtgaggaggaggtgtgtcgtcttcccaagcgatataaacga gaagcgattgaagaacatacctgtgattaatcgaagagcaagctaaagaagggtatcc agtacagcgcaattacacaagctacaagaattttgctaagtaactacataaagatttc atagtggaaagtaccttggcgcaaaaacaagtaaatcgaactcagttgggctgtgtctaa atcgcaaatcaatgatgctctctgttatagcatcagaaggtgaagccagtgaaatgccagg gtgtgacatccttgaaaaacaagtaagaagcgttactatatactacatttctccccc agaaggagcagccaattgttaagtgtagagagacttagagatgttccggttccgctatg ggacaaggtagagtgtaatacctcaagaacgggaatgtagtggatgttcaagggc gtagaagtagcggtagcttggccatcgtctctgtatggggaattactatagaggtatat cctacaagaagccttctgctaaagaaggcggggagtaattacgttagagaaaaggaga cggcaattctccccgcctga	gtcaataacccccctgattcaagaatcag aaggggctgtagtgttatccctgaagcc ctatattgacagcctaagttagttagtcta cgttattctgttatcagacccccggggtgctc caccagctccgactatcattttgtatcgcat ggctctctgatggcgtctgatcttgggtatc agacattgtaaacagtcctaagaggctctg gacagtcaacataggacgacttagcatct aagcaagctggaaataacattggcggggtg aatctaaccctatcttatagaggtgtgaga cttcattgta
594-595	Ga0175859_1302585 GENOME_ACESSION: IMG_3300013314_SF_3 300013314 GENOME_ID: 28680 CONTIG_ID: 20373 SOURCE: JGI DATE: 2019-01-27.	atgaaagttacgtttatcaacaacatggctgccttaagtcccacaaccccaagaacg caagattgctgttaaaaggagggaagcaaaaattatgtgctgacccctttactatccaa ttaatttacggttctagtgtggtacacacagccgggaaggctaggaattgatgctggtatga gcattatgggtatagtggttaacgaaaagaagaattgattgggtgggaattgatattgt taaaaggcatgtacgcccgttaacgggaacgggtcaaaataccgtcgtcaagaagaagaat aggaaacgtcacccgcaccagaattgtgataatcgcaggcgtaaaagaaggtggttgc tcctagtattcagcacaagcttgataccaccacaaaataatcctaaagattgtatgattgt gccagtcagaanaagataatagaggttgccagttttgatactcaaaaatcaaaagacact ggtattgacggcaattgatacaacaaggcggaacagtcaggtttgacaatgttcgtgaat atatactcaccgtgacaagcacgaatgtcaaaatcccagatgtaagaacaagtcacactg agcctatcttacaggtgcatcatataggttttgaaggaggacaggacggatagaccag caaatcaattacttttgcgataagtgccatacacaagaatcacaagaagaatggtttt tgtttgggtgggaacaaaacttaaatcctttaagggtgaaccttcacaccacagtcagg tgccgcttaagcaatgaagggaatgataatcaacttacggttacattaccaaggatga gaagagattttcagatagagaatcacaccacaacgatgcctttgtgattgctggcggtac taccacaaaaggattgagccattatctagaacagattagacgaaataagcggttctcta gaacagttttatgacgcgaatatatgtatagatgatttcaaaaaacaggctcaga tttgtcatctggtcgtataaccgttaataaaaacaagagtggtgaaaacttaagacatttc gaggtcagaaaataaccggcgacaacgacgaattaaagaagtcgtgatalcgttaccaa ccgcgagatttggtgaagtttgagaacaagccttatgaggtgattggaatgcaaaacttag gtactggtgttaagttaaaagactatccaggagtcacaaaacagggtgtacaagtaagttt gtccagccactaaagaagagatcgggtatctgtactaggcttggatag	ctgtgacgactccacaccgacgccctttg gggtacgggtgtggctcttaagaattagag gttcgtcattagcctcagtcctctgacgc aagagtcgctccgtttccagaatataagcc actccgaggtgacaggggtgtagccctgt attcctagtgcgggactcgggtcactgatt aaacattcctactgagaggggtgtggttg gtgacaaaaacctggacaaacattggcaa aggaaaccacgtcgaagaagcattctctct aattaggatatacaataatgaaa
596-597	Ga0394874_0000250 GENOME_ACESSION: KP_S3_MG_FD_SF_330 0037309 GENOME_ID: 312525 CONTIG_ID: 249 SOURCE: JGI DATE: 2020-04-07.	atgtctaattttgtcttagtcataatgccccaaaacagcccccttaacccggttcatcccggt caggcgccgctgtgtgtaaaagcaggaattgctgctgtatttcgagatttccctttgtatc attctcaacaatcctgttgcgttttgagcgcgctcaagacttagagttaaaaatcgtccg ggatctacaaccactggttagcttgcgtcaagacaaaaaagttgtattgtctgcacagtt gacgcaccgtggacaagcgattagagctaaatfagaacgcggcgcaaccaacgcagtc tcgaggagaagtcgccatactgctaccgacagccgcgatttctaaccgcacccgtcgt caaggttggtagctccagcttgacgacccgcgtcgaacaacgattacttgggtgaac aaacttatccgatttgcacccattggctcaattgticagagctgtgcaaatcagctgcaa aaattagaanaaccctgagatttcagggaattgaatatcagcaaggcgaaattgcagggtac gaagtccgcgaatatatgtgtgcaaaatgggagcgaaaatgctgctgaattgcgggtattga aaatgtccggttacaggttgaacacatccatcccaaatcaaaaggaggctctaatcgaatc tcgaatctttgcctggcttgcgaaggtgcaatatcaaaaaggaaactcaagttatgagc aatttcttgcacaaaacccggatgtactgaacagagttttgtctcaagctaaatgtccctta	agcaccatttggagtaaatcgcaagaatct ggtttatttcggctactccagaccgctgaattt tgattcagcagccgttatcttcaagcgttaa gttcgcaccttgggattcgctagctctgagct tctgcaagtcactgttaaacatctcctgattg ggtaaaaaaatgcagtggaacaagaccg ggagataacagggtcgaagcaaacattacc ccgtgaaggaggacttagatgtcaattt

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aggatgcagcggcggttaattcaacacgatggactttgttcaatcgattaaaagaacgctgttaattaccctacttgacgggtacgcggcagggtcggatagacaagtttgggttccttgttctaaccgccagaattacgatattgctggcagaccggcgatcgcgatga ctgtcaaaagatgggtaagatattgttgaagttgtcgtgcaatccgaaaaacgggttga agtaagaactggaaaftaagaatcgccggcacgcttgataaattcgtaaaattgcactcg caagacgggtatcaatagccaagattgacctactaa	
598-599	Ga0007854_10018219 GENOME_ACESSION: APCry1669191629_5F_3 300004806 GENOME_ID: 21780 CONTIG_ID: 18218 SOURCE: JGI DATE: 2019-01-27.	atgcaaaagattttgtattagataaaagtagacggcctttgatgccctgcacggcactag agcaagaagactctttctcaaggaaaagctacatgctcaactccatcctttacgattct tattcaagatagagaaggaggcagggtgcaaaattgaagtaaaatcgatcctggaag taaaatttcaggaatgtactttatggacactttaaaaaaggatgacagtcataatggcgag ccaatctgaacacagaggagctactattagatcctcgtgaatctagaagagctattag acgaagtcgaagacatagaaaaaccgttatcgagcagctcgatttaataatcgaaaacg taaagaagtatgtagacaccttcattacaagctagagtggaatacattgtcattggatca aaagactggaaaagttagctcctatttctccattgcaaaacgggtacgggttgatagc aaaaaatccaaaatcctgagattacaggagagctttacaacaaagggtgagcttatgggtta cgagatccgagagtagcttcttgagaaatggggcgcgaatgtccattgtagcgggaga acatcaagggttagaaatcgaatcatattgtcccaaaagcaaaagggtggaactaacagagtc tctaactcaccattgtttgtagaacctgtaattgaaaaaaagcaaatcatcctcgaagaa tttctccataaaaaatcagctcttgcgctaagattttatccaaagcagaaaagccttggac aac	gtcaactaccacccctaaaggaggtggcctt gaagaggtgactcttaagcttgggtgacca gactaagcatcatgttgatgctacgattac aataggtcggttaagacacctcttgggatgcta ctccagctccaaagcctagaagggaagatc acgatgggtaaaaggtaagaccggaagggt ttacttggcgaaggagccgggtgtaatac ttgccctggaggagcttatccacaaagcgct tttgcgtaagtgagaggtgactacgtgcta aatttttttttttgaagatatgcaaga
600-601	Ga0209720_1000657 GENOME_ACESSION: AD_des_11_5F_3300025 605 GENOME_ID: 21309 CONTIG_ID: 656 SOURCE: JGI DATE: 2019-01-27.	atgatacatcgttggattaaacaaggaaaagctaagtttaggaacgtaactagtccaagt atttaacaattcgatccagcaaaaactataccagctaaattgtattggtatcgatccgg atacaaaaatattgtgtacgcagtttataagattataataaaaattacagagctagtcca ggtgaagtagtcaccagaactagttagaataaagagctattagatagacgaatgtttc gtagactcagacgctattatcgaagaagaatgtcttaagaaaatcggtaaaagttaagttt agagctcccgatggaagaatcgaagaagaaccatcgcctctactcataatcatttga ttcagagtcattgtaactcttaactaggttatttaattgtaactttacagaatttcactag aatactctagcttgacactcagaattacagaatcctaataaagaactggcaatcaga aaggctcctcagtttgattgaaaatgttaagcatatgtgagagctagagataattatcagt gccagaattgtgttctggagataacttaagagtacatcatatagtagagcgaagtaaaagg tggcagtgacaggtctgacaatttaacacagtttggagagctgtcacaatcgtattcatca aatggattatcttccagcgatttcaactgataataaagagagatagcggagttctgaa ttcatgtctgaagaagctatagaagttctagctgacagatataactactgttaaacatcgggt tacatcactagtagcttctgaaagatctaccagtttagagaaaagtcacgaactgatgctaa attaatagctttatctgatgaaaatggattagctgtagatttagaaaactgcgattattctagta gtgactcaactataattttatcagttcagaaggcatcagcgtagctgggttaagagatatg tagacagaaaatatatagaacagattttatgcaacagttgcttggaaatagaagacgtcgt tctgctcaagatgaggagaagccaagcttcaagaactaaaagctgaatacctgacgctc gcttaatagctaaaccggaaaagtcgtctatagaagaagtcataaactaaatcagacga cccgtgatatcttttagatcggaaacagattgatgtagcgaaggatggtcgtctactat gaatagagtaggtgtctagacatcggtacggtccgataagtaaaactactaaaatctgt aacaattcaggattagtatactaccgctaaagcgggataa	gtcaactaccacccgctaagcaggtggcctt gtaagagtcactcgttagtgggtgggttga acaggcacgtggaaggttaagcttctgcag gatctgattaaagcgttcacccctgggttgcga ctccagctccaggtctgaaggtaatagct aaaagtttagctagaagcttagttagtagca agtaggtaacgaagcttagttacgcttaa ccaagccgaggagactcacactgatacc cgaatctgggtatcgttttaaggttagagg taactaaaattgaaacagata
602-603	Ga0177923_1152199 GENOME_ACESSION: IMG_3300013414_5F_3 300013414 GENOME_ID: 28682 CONTIG_ID: 6545 SOURCE: JGI DATE: 2019-01-27.	atgacagattttgtatcagtaaaaacgggtgaacgcctgatgcctacttcccgggttgggaa agttcggcacatgatcaaggatggcagagccgtcatttattgccaaatccattaccatcc agctgacatagacacaaccggctacacgcagccattgaaattgagaggataccggg tatcagcacataggaatcagcgtaaaaagcgagagctcgagaatattgttcggcgcaatac gatctccttgagcacgaaaaggaaaacatgatgactgccgaaaatcacagaagggtcga aagaaaccggaaagggtaccgcgcccaagggttaataacccggagcgtccaagaaa gaaggctggctgctccatctctgagaataaagcaaatcggcatatcgatcaatccag aaatatgtcaagggtgctccatcacttccatcacgggttgagcttgggcaattcgatacgca ggtcctgaaagcagtgacagaaggaagccgggttcagaaggcttattaccaatcag ggccctcagtaggcattgatacttgcgggaagcagcttccaagagataatcacacctg cattttttgcggaaagaggttaaaaaacggagccatcctgcatgttcatcatgtctatttctg gcgtgggcagcatggtaattcttctgacgaactggctacgtgctgtgaaaagtccatact cccaaaaaccataaggagggtgcaagctctgggttacaacaagaagcttccagata caacgggtgcggcattcatgaatctgctgttggatatctacaaccaggtcaaggacctt aatcctggcattgataattcagttacttatggcgtgctacaaaacgcagccggatggattt aggacttgaaggcccatgtcaacgatgcctattcaatggggtctatttcatatacaaaa aggacagccaaggagtttttctaaaacgccgccgtaataatcgtgctcctggaaaattct atgatgcaaaatcatagatgccagagacgggaaaatcaagtcggctccagcttggga tgcgaaggataaacaggagagagcctagaatactgataaaaacctcgcatitttctg gaggaaaaaagtcaaaagggtcaccgatccattagacgacagcgatatggaatcgtccc ggagatatcatgctttagcaacaagaaaattcgtactgggttcattgtaattgggac aagagttttatgctgggtgaaaatcttaacaattggaccagatggagttatcgacatcgc gcggttgggaaaaactgactaa	gtcaacaacccgccactaaatcaaaagtatta gtggggcgctgagaaacagctccattcgact gtccctttcaagcttagttgattagcctcagtt ggctgggtatgcttgcagcgaaacaaacta cgttatcttggaaatccttatttggaaatag gcaccaggttgggtactccacaagcccaat gctctgcggtatgatcaaatcctctaacgg taggagaagtgctgaatacaaaaacccaag ataacaattggcgatgtagccacgggtactgc gctgtcatcttgcagacagttgcgaaggc cgcatatttccccacggggagtcaggagt aaatcacatgaca

SEQ ID NO:	Name	IscB sequence	ωRNAs
604-605	Ga0256405_10019152 GENOME_ACESSION: RumenRJG_02DNA_2_ \$ F_3300028048 GENOME_ID: 34937 CONTIG_ID: 19151 SOURCE: JGI DATE: 2019-01-27.	atggtttacgtattaaatcaaacggagacgttttaagctacaaaaaacatgccaaagt gcgtgttcttctaaacaggcgcaagcaaggtgataaacaaggtgccgtttacaattcaa ctgtgtatagtagcacgaactatgcacaaaaggtaactttagggtgtagatagcggaagta agcatattgtctcttcagcaaccacaaaggataaagtattttgagctgatgtagagctta gaaacgatattatggatttgccttctctctgtagagaattaaaggcggtcccgtagaaatcgt agcttcgttaccgtaaaccaagatttaataatcgacagcgtggtaagggtggttagcacc ttctgtaaagcaaaagtagattctcacttaacaatggtatcaaaaagtatgcaaaatactcc aatacaaatatagtagtagaggtcgtcttctgcatatcaaaaagattaaagatcctacaat aagcgggtgctgattatcaacatggcgaacaattagacttttggaaatgacgggaatatgtgt tattccgtgacgggcatactgtcagtgctgtaagggtgaagcttaagacaaaatccttaaf gtacatcatatagaagccgtgaagacaggcgttaatgcacctaacaacctgatacacta tgtgaaacctgtcatactggatataaggggaactgtaaaactgcctaagacaatacata gaggaaatgtctttaagtagctacgtttatgggtatcatgcgttgggctttgtatgggaaac ttaagctatataccagacgtaaaacttacttacggatataattacaacccatttgcggcg aaagccacgggttttaa	ttcaactaccaccgcctaagggcagtgagg attgtgcagtagcagtgctacgaagaatcc taagttgaatagcctaagttctttgagaactac gttaccttgaatatataggtacttgggacg taatacctaattccgaacactacggcaatga ttaacagttctgtgaggtaggaacagtgtg tagtcatacaaaacccaggataacattggc gaaggtatgcaaacccctttcggaggagat aattaaaaccttattgttttggaaaggaga atgcgtaag
606-607	mouse gut metagenome genome assembly, contig: NODE_1060_length_375 73_cov_4.643504, whole genome shotgun sequence GENOME_ACESSION: UWSV01.1 GENOME_ID: 279227 CONTIG_ID: 1059 SOURCE: WGS DATE: 2019-06-06.	atggaaactaaaagacataacattcttttttagtagatagaaggtaaaccccttagctcct actacaataatagaggatttgaactattaaagaaaagaaagctactcttattagtaaatc cctctagtaattaaftaaataaagaataagaaatcctaaatgtaatttgaattggtatag atgatggctcaggtcatgtaggtcttattactcaaaatgtaaaactaaaaataaagtag tctttaaagccaaaatagaacaacgtggagatgttaagactttaatgactaaaagacgtga acatcgtcgtcatagaagataccataaacgtttataggaaaccgaggttataatagggtct cttctaaaagaaaggaaatcccccatcttataacaaaaaaagatgctattcttaga gttattagacaactttctaaatggatttaattacttcaaacacacttattttagaagatgtg gctattgatataagagctcttacaagatgcctttaaactttagatgggaatatcaacaatct aatagattagatgaaaatttaagaaaagcggtcttttttaagagataaaaatacatgtcaaat gtgtcaatgtaaaaggacaaaagaagtcacataaaggcgaaggcgtttacatg gagcagatactataagtaatttattgtctcaagtgccatgaatcctataaaagataa agaacctctttatgaaaactcttttatgaaaaaattaaatctattggaacattcgtttgacta tgccacacatgttatgcaagaaaaacttattcttagagaagaactttctaaaatagggtgtc acaacttactaatggaggagatacagaataaaagaatagtggaatatagaaaagtc tcacgcaaatgatgctatctgtataataggctttaaaccgtataccactgatatttttgaatgg actataaagcctattagaaaaagaacaacttataaaaaaacagattttgtagaagggtct aatggcatttataacataaagatttattttctatcctataaagatggtaatacttatgtagg ttatgtaactgctttatctctagtgataagaaaacaaatcagatttataatttcagagta aagaaaaacattgtaagaaatgcaaaagaagacaaagctcttatggagcttcagtc atctttattggctatag	tataaagatgtattatataatacatctgtattgt ataattatattattatgtatttttaaaagatcca agcctaagtgatagcaccatacgaaggtta tgcaataaaagaaattttatttaagtgctga tatgaactacgttgataagtagctaaaaata taccttttagtacaccttttagcctaagactcta taactgccacaagaacaaagttcttactat tactttgataacagggaacagatgtgctctt attgacattggcaaaaaggacaataactcatta aatagtagtttctaacctagtagaaaggaga ggcgaagggccaaaacttatggaactaaaa gacataacattc
608-609	Ga0114843_102905 GENOME_ACESSION: IMG_3300008138_\$F_3 300008138 GENOME_ID: 28389 CONTIG_ID: 2904 SOURCE: JGI DATE: 2019-01-27.	atgggtattgtattagataggtctaaaaagccttagacatgatttcacatgctaaggctaga ataattgcttaaaaacagattagcagtagtccataaagtatactctttactattagactaaagg acaatagttgtgaagtaataagaacttactgtttaaactgatccagggttcaagaacta ctggagttgccactgaaatagatcacatcacaagaagtaattggcggaactaaagtt catacaatctaactccagcttcgagaccatgcaatgaaaagaatacaacttatcactaaa agaatttgggaagcttatgaacaaagactattctcagtagaactaagaacactgcctaaa gatgcagcaatagttcaatctgctagaactatagattaaagagataactaaactggtttct aatacaacttcgtatgacgttggtaactaaatataatagagatagactagacttatctaa caacattactatgatgctttatcagtaggaatatagagaacttttagatttctacagataagg tattacaatatcagctaaaggtagagggtcttaggcaaatgtgcctttatggataagtttggt ttccaagaacaaacaaaaggtagtaagtttagttaaaggcttcagactggagatatgggt taagccacagttccaaaagggtttaaanaaggagaatatcttggtaagggttcagttaga tctagtgatattttaaattacagactaaaactcaagttatccaagatataggatataagattg ccgtcttattcaaaagcagtggtactcatcaattataaggagtgcgacttctctcgg ctattcataaccgagctccatcgagattttagatga	atcaattactcagtagttaactactgagcttga ttgaccagactaagttgctaagaaatagctt actcagtagataagttatcacaccgcatg atgcttccagtttgcggctctgtgtaggctc tgtaagaatgcttaaggcaggtcaacctag gaacggctcagcaagcttatcatattgtc gaggagagactacttcaaaaaggtgg gcgataccgtagcaatcagagattattaaaa gaaaggaaattactgatg
610-611	uncultured Erysipelotrichaceae bacterium isolate RUG13468 genome assembly, contig: RUG13468_asm_35, whole genome shotgun sequence GENOME_ACESSION: GCA_902789375.1_Rum en_uncultured_genome_ RUG13468_genomic GENOME_ID: 347331 CONTIG_ID: 34	atgtctgtgagcagtagtactccagcaccggcgagaagctgatccgacttgaactacaga gcacgggaagctcttaaaaagcaaacgtgctgtcattggaatgctatcgtccgattttacaat ccgggttaaccgaccgggaggagggaacacacagccgattgagtagctcgtcgatacag ggataccagcatgtcggcgtatcaatcaagtcggaacaaatcaggttcgtccatgcgcag tatgacatgctgagcgtatgagacagagcggcacaatgactgcggcaagtgccgcagaa ccaggcgaaacagactgcggtaaccgtaaaccgagattcgacaaccgttcgaagaagaa caaggagatggcgcttccctgcggcaccgcaaggagaaccagatccgctgttcgaat cggttcgcaagtcattgccgatcacatcgccggcattcgagatggggaagttcgacactc agcttctgcaggcaattgcagacaagaagcctctccgaaagggaaggactaccagcg gggatcggaagtactgtatcagaccgaacgggaagccgttttcggcagagatcactatc ctgccagggtatgcgttaaatccgtcaagacggcggtattctgcatacgcaccatcagg gttctggaaaggataccgtccaaccgaatcagcaatcttctgacagtatgtgagcactgc catagggcaaaagatcaccagccggagggaagcgtctgggagctggagccgaaagc acaaatctggctccggcgacgtatagtagtaccgtgcgtcggaatgtatcgcgatctg	gtcaacaaccccgctaagcttaacggact atagacggggcgttgagaggcttattacagc tcttcaagcccgttgataagcctcagctcgc gtaagagccttcggactctgaagtgccgga ctcgttgccgtgagctacgtacgaatgctc cactagttcgtaacctcttgggtcgatgcta aaaatccccctagggttaaggagcgtgcatct gtagtattcaaccacgggataacattggcga agtggaaccaccggtgcgtggttcttcgga accgttaagtcgagatgctgcatgtcttcag gaatgcagaaagcgttaagcccttaacgg ggcaggcagaaaggagcatcagaagat gtct

SEQ ID NO:	Name	IscB sequence	ωRNAs
	SOURCE: NCBI_Prokaryotes DATE: 2020-04-22.	gtgctgacgcatccggagctcgatattatftccagfatggtgcaaaagactgctgtaaccag aaagtcgctgcatcttgaagagacacatgcaaacgatgcttactgattgggtctctgcatc cgaagcatcgaaacagcagagctgtctatcagaaaaagcggcgaataaccggattcttg caaagttctacgatgcaaaatgacatgacatccgggatggcgaaagaagtcgggttcc gcactctcatcggggagaaacacgcagggaagcaggcgctctgataagaatcagc gcatctatcgcggaagaagtggtctcaggcgacttcgggtcggaagaagcgttattc atatcagcctggggatcgattctgttccggctcagcgtttcactgtcaatgggtctcattg caatggaacccgctgaattctaataaccggaaaatcagcaagatggcagatttgacctg cattaaaatggaaggagatggtgcttctccccgccaaagcctga	
612-613	Ga0105046_10007489 GENOME_ACESSION: MAT03_2_\$F_33000090 84 GENOME_ID: 31818 CONTIG_ID: 7488 SOURCE: JGI DATE: 2019-01-27.	atgcaacgagtattgtctagacaagaacaaaacccgctgatgcctgtcactcggcg cgtgcgcggcaattgtctacacaagggaagcggcggtcttctgcaataccctttacc atcagcccaaatcatcgccgcttcaattaggcgaccgggcatcaatctcgccaaatg tgccgtgtcgataagtagcgtttccacgaacagccccaacaaggcgcaattcactat ggtttccaaacgggagatattgtcaaggcgtgctgc	gtcaacgaccacccataaagggggtgac ttgcgataagcaagccatgttgaccagcct cagctcattcgtgagacaccgttatccgggt catgataccgacgggtgactgaaccagccc gtcgtctatcgctcaccgttaaacatgccga gggtctaggcaagtcgggtgggcatgaca agcccgaataccattggcgagggtcacttta ctgccgcaaggcgagcataatgaggttaact catgcaacga
614-615	human gut metagenome genome assembly, config: NODE_671_length_2545 7_cov_5.408117, whole genome shotgun sequence GENOME_ACESSION: OIYO01.1 GENOME_ID: 6149 CONTIG_ID: 670 SOURCE: WGS DATE: 2019-01-27.	atgcaagtagtatgtattaaacaaagattaccactttgccaacgcataaactggg taaaagtaagacatcttttaaagatgtaaaagcaaaaattgttaagcgtatccattactatt caactaaattatgagtggttaattacattcaacaaattacattagggtgacgtggttcta aacacataggactaagcgctcaacagaaaaagaactatattctctgatgtagaact tcgtaaagatattgttgatttattcaacacgtagacaaaacagaagaacagacgaatc atctaagataccgtccagcaagatttgataatgtaaaaaagaagatgttgcttagcacc atctattagacaaaagattgattctcatttaaaagtaattgaagatgtacataagattctgct atcacaaacattattgtagaagttgcatcgtttgatacctaacttttaaaagctgaggctgaa ggcaaaaccattagtggtactgtattacaaaagcgaatgttaggtataacactcgtg agtatgtattattagataatcacacttgccaacattgcatgttaaatccaaagataagg tattagaggttcatcatcttgaagctgtaaaacaggtgtaatgcaccaataacttaatca ctttatgtaaaacatgtcacgaagcatatcataaaggaaaaatagaattaaacaaaaacg agggtcctaataataatgacgctgcctttatagggtattatgagatgggcttctacaacaaat taagaaatctatccaaacgtttcttaacttatggatataaccaaataataatgcattaatt agggttagaaaaagacattataacgatgcttattgtattgctggaaatttaacgctaaacc acttaataacctttatctacaaaagaaggttagatgtcataatgcaccaatacataagctcaa ccgttttaaaaggcgggtgttaaaagcgttaaccagacccttacttagttaaggatttagatt attgacaaaagttaaatacaaaatactgaattgtttatctttggaagaagatcgtctgctac ttcgatattagaaaattagatggaactaaaatacatagtttttaattataaaaaattggaatt gcttaatgcaagggaataatttattaaagaagaaggagcagcaattctctcagccct cttaaaagggtgaatttcttgcctacttgtga	gaagcggaggagctgaatgtcaactaccc acggtaaatccgtagactgttaataagccta gttgattagactaaggagaaaactttaagtta atatctctgattatagaataatataatgttacct acaaatgtatagccaaagttttagtcttaagg tatgtgtaaaacagttttaggtgtaaaac agtggtgcgttaattaaacattatataacatt gtcgtggtctacattacaaatgtaagattg aattatctattaaagtagatagaaggaaaagt tatgcaagta
616-617	Ga0209647_1002274 GENOME_ACESSION: GraSoiAngeSPAdes_43_ _\$F_3300026319 GENOME_ID: 26992 CONTIG_ID: 2273 SOURCE: JGI DATE: 2019-01-27.	atgtcgtgtgtgtgtgtgtgacaccgagcaccgcccactggaccgggtgcatcggg cgccgcccgcggcgtgctctcagaggcgccggcggtgtggcggtgctccccgt cacgtgatcctgaagcggcgccgtgcccggcgacgcgcaccactcgggctcaa gctcgatccgggcagccgaccaccggcctggcgctcgtcaccgatccgacggc accgctgctgacgaggtatcaatcgggcggtgtgtggtggggagctgaccca ccgggggagggccgtccatgagaaactgtgaacgcgcgtgccattcgtgctggccg ccggcagcgccacatgcgtaccgcccggcgcttcgccaatcgtcggcgccgga gggggtggtgcccgcgtgctgagagccgctggccaaccagagacctgggtgcg gcgcctcgtccgctcgcaacgtcgtcgccatctgcaggagttgtcaagtttgacac ccaggcggttgagaatccggagatcagcggggcggaataccagcaggggacgctgac gggtgatgagctcgggaataatctgctggagaagtgggggcgcgctgcgcctattgc cacgcgacagcgccccgctccaggtggagcacatcgtacccaaaacgcgccttggc ggctctgaccgtgcccgaacatcagctggcctgcgcgcgtgcaaccagcgcaagg gcacgcggaccgcccaggagttcgggcacccagaggtgcaggcgagggccagcg gccccctgcgcatgccgccgctcaacgcgagccgctggcggtgtttcagcggctg cgggcaaccggcttgcgggtggaacaggtaccggcgccgacccaagtggaaccg gacgcaacgaaacctgcccgaagacgactggtgacgcccgtgtgtagggcgctc cacggcgcaacacctgctgtgtggcaggcatagcccaattgacgatccggcacgggc cggtatgcgcgcatgctgcgcgatggaccgtttggcttcccgcacggggccca aggccaccagcaggtagggggctgcgcagggcgatctggtgcggcggtgtgtg ccggcgccgagcgtgaaggtggcagctacgtaggcggtggtggcggtgctgcccagc ggcatgtgcaacatctgacggccggcagggtgtgtgagggcatccatgtgcggc actggcgccgctgcatcgagcgtggctacggctacggctacggctacggcgatgg ggcatccatcacactgaacagaggcggtga	gtcagggacccacggcgtgaagccggg gctgtgtgtgagcagcccaccctgaccag cctcagctcgaagagctccgataccgg cgaatgcataggcaccggcggtggcacc tcagcccgcgcgtcggggcagcgtta aacagggtccacgggtgacgaccagtgcc gtgtccgaagaaggaaaccggcgataatc ctggcgaggggaccattaccggcgaag cggaggcccaagggcagcaagatcgg agtgcgtgtgtgtgtgtgt

SEQ ID NO:	Name	IscB sequence	ωRNAs
618-619	Ga0401364_0036777 GENOME_ACESSION: Site2_nosalt_MIX_4_FD _SF_3300037317 GENOME_ID: 312665 CONTIG_ID: 36776 SOURCE: JGI DATE: 2020-04-07.	atgcaacgtgtactgtactagataaaaaacagagagccgctgatgccctgccaccggcg gccccggcggggtgtttctgctaccgttaccatcatcatcacagatcgggaaggcg gcgacgtgcaacctgttgcctcaacgttgaccaggcgagccgcacgtcgggcctggcg ctgttcaacgggttcaaatcatccggcatccgctggaacagggaacaggcgacgca cgaatacaatcgcgtcgccagaactaccgaaagccattggattgacggcgctgc gtggggcgaagtggtgcgtatgtgtgtacatccgatccgatcggccactgctatcaaa gctaattgggctcaatcgcttgcgtgacgtccagataagtacgcttccgcgcacg aaggcc	gtaagcatccccacgctgaaggcgggg cttttaacagcccaatgcttaccagcctcag cccttcgggggctacgttaccggcgaatac ataggcacctccgggtgacgccaccagcc ccggactctgcgtgagtggttaacaggga cgggtggggacaacggcccgctgctctca catcaaacggccgataacattggcgaggt ggacgttaccgcgtgaagcggagattagag gggtaactctcaatgcaacgt
620-621	Proteobacteria bacterium isolate F1-120-MAGs186 F1-120_c247377, whole genome shotgun sequence GENOME_ACESSION: GCA_005777235.1_AS M577723v1_genomic GENOME_ID: 284306 CONTIG_ID: 42 SOURCE: NCBI_Prokaryotes DATE: 2019-07-22.	ttgtggcgagttaccgtgttggggcagacacgacaagcctgttgaacattgtcgaggca aacatgacccgagtaacggagagacatgtccacgtgtgttctgctgacgactgac atgtcccacaaacccccgtgcatcccgctgtggcgcggaagatgttgcgctaggggcc ggcgagtggttcaagcgcttcccttaccatcatcttgaagacggcagaggccggccac cttccgtgcacacgcacgcctgaagattgaccttggcagcaagaccacgggactggc gttgcctgatggccccagggtcgtcgtggcgagagatcacgcataggggccaacgg atcacagatgcgtcgtcgtcgcacgcgacctacggcgtatgcgcccaacggccaga cacgctacggccaagcgcttccctcaatcgacccggcgagggcggtgctgctcctc atcgctgcacagtcgtgtggcgaacatgatgacctgggtgcaacgtctcgaacgactgt ccccatcgtggcgctgagtaagagctgggtgctgttgcgcagctcctcgcagcacc cgcaatcagggcagtcgagttaccagcagggcacgttggcgggctatgagtgctgag tatctgttggagaagtggtggcgctgttgcctactgtggcgagccagctaccgctt gaagtcgagcatatcgtgccgaagtgctgggggctcacatcgctcagcaacctcac cctggcctgctgtcctgtaacagcagaaggcagccagacggcgggagagtttgggt ttgcgaagatacagggcgagccagggtgctgctcaaaagatgccgcgtcgtcaacgc cacacgtggcgctgtatgtgcgttgaacacactgagctgccagtcgaacaggga cagggggggcgacgaagtaacgggactgcccagggcattcccaagtctcactgga gcgatgctgctgctgtggcgccagcagccagagggcattgcatgttccggtgtaca cccttgggtattcggcgatggggcagggaccagggcagatgtgtcgtgtaggtggcc agggtctccgaaggcgacccgcacgcgcaggaagcagtttggatgcagattggc gatalcgtcaagccagtgctccccagggcaagtatgccgacatggaccagccgcg tcgtgtcaagaacagtggtgtgttaccgtgttatccatggcaaaaaggccagctgca tcacaagcattgcacacgcctctggtcgtggatggctatcgtacacccgtgctggt gcaggcaccggcgttctccccccactga	gagtgagtcgacgtactccccgactgaag tcgggggactcttagagtcgggtgcgagc agacatagtgccgaatgcgtccgttcagc aggtcacgataccctggaatcggtccaggt agggtctctatcgtctgcttaaatgttgc gcgagttaccagttgtggggcagacacga caagcctgttgaacattgtcaggcaacat gaccggagtaacggagagacatgtccac ctgt
622-623	Ga0137390_10046836 GENOME_ACESSION: 15con2h24AmetG_2_SF _3300012363 GENOME_ID: 20762 CONTIG_ID: 46835 SOURCE: JGI DATE: 2019-01-27.	atgagtgctgctgcgctcaccggaagagccgctgagcacgacaatggtatcgtctgc acgcttcgccaaccgagctgcaagctcgcgagggggagaatcccgccttgcataaa accaatgtcgaaagggtgacgggtatcaagcgcgcttgccttgcgacgacgacatc atcatcgaaagatgtccaggttgatcgcgcggctgaataatcctgagctcaaaaggag ccaatatcaagaccgacacggctgagtagaacctcgccctcgtctgctgacgcgcg acggctatgcctgtcagcactcggggaacaccaggttcgctcgaagcgacacatctg gtgtatcgagaacacggcggaagagatagcgtgaccaatctgctgacgctgtgtgaggc ttgccaccagaaggctcatcaggacaagctaccttacagtgactgtgtgtgagtgggc atctcgaccaggagcctacttctgctctccctaa	aatagctgctaaaaatctagcgttgttttacgc cgggttcagtgagctaccagcgggcttctgc ggaaagcagcagttgggatggtcatgacacc accagatgtgttccagtcgggtgcaactgt cggatccattaagggtgagcgaacagtgga agggtgtgatactgcaaaaagctctcacaac agcccgaggagcatttcaactctgaaagga gggcttagagccaccatgtat
624-625	JGI11876J14442_100221 72 GENOME_ACESSION: ComassMDCD2ACD6A _SF_3300001357 GENOME_ID: 24872 CONTIG_ID: 22171 SOURCE: JGI DATE: 2019-01-27.	atgaatcgtgttccagtaacatcagaaggaagccctctgatgcctaccaagccgtcta gggctagacgtgggtgcaaaaggaagggcggtgggaagtggtcagacctggggcg tctattatgtccaactaactacaccaccatcggtgaggaagcacaaccaatcgccgttgg agtcgacccgggcaagcttattccgtgtgctggtatccagtcgctaaatgcaacccgtcta caactgcatcttattctcgttggcggggtcagaagaacggatggaactcgcgcctatgt acgacggggggcgacgagggcgacgggtaacgggatgccccattcaagcagcgttaa ccatcgggcaatgcaggttcgataatcgaaacaatgtaagcttccccatccattaaagcc tcacggcaactgaggttaagggtgtgtacggattgtgtaaacatcttccgggtggcagcg attggttatgagcaggttaagccgacatagaccagactaagcgcaaacgcgccaagtc gggcaagggtcttccccggtgatgttggcagaattggcgaatcccatgagtgga aaatcgccccgtctacgtccgacatgttggcaaaaggatgggaacgggtacatcccaa ctccgaaccaactagggttagagaaggataaaacgaataagtcatagtcaaacggga gaccacgcagttgatggagtggtttagcttgggtatttcatcaaatatgtccgttcat tggctctaactccatggctacaccacaagggtggcggttaccgtgaccccttccccgttc aaaattatccccgtctggtggtggaagcgggggaagagtagtgggttttccgtcgc caattgcacttgaagtaccggataaatctggagtcagggaagcgcaagggggaacaat cacaccatttggagcaaggattgtgattgtcagagcagagaaaagctgggaagtctta catcggctacgttgggtgggttaccgatacaaaaagctcagaaggttctgttgtgacta tacctggaacggatggcgagtttgcctcagaaagtcgagttatcaggagaataaa cggctctatgcgtagcgttaa	ttcatgaaccgcgcttacttctgtaagacg cggattggcagagataacctgctccgaagt catgaatagggcatgagttggtaggggata atcacttccgaacgcttctatgtcggactc catgtaggtctgaatcttaggactggg tcattccaggacacccccaccaatgcctta agagacgaaaaattacagagtgga
626-627	Ga0209616_1000698 GENOME_ACESSION: FloCalmetaSPAdes_SF_3	atggtatttgttttagacaacataaagaaccgctaattccgtgttcggagaagcgcgcg gacaactattgcaacgcaaacgtgcagtagttcacaaaatgcagccattcacgattcggt aaaagaccgcatagtagagaaaagtaattgcaacccttagattaaaactagattcagg	gtcatagacccacgcctaaaggcgagg cttgaaaaagcccaactatgaccagcctaag cccgccgggttgagcgggactacgttggaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
	300025091 GENOME_ID: 26000 CONTIG_ID: 697 SOURCE: JGI DATE: 2019-01-27.	cagtaaaftaccggctfttgcggtgctccgggaatcggcacagagaaatctgtggcaat tctgatgggagaactgcatcacagcctggaataaagaccggtttagataacagaagag ctcttcggcgagccgtcgaaaccgcaagaccgggtaccggaagccaagatttatcaat cgcaccgcgtcacaaggttggctaccgccatcgttgcgggcaagagttaataaacccct aagtgttgaaccaagttaagaacgtgctaccgataatcacaaatagcactgaacatgctc aaatttgacactcagctcatacaagcccaaatatccagatcgaatfatcagcaggcg aattgttcggctacgaagtcgaagagtacgtcgtgaaanaatggagacataggctgcct actgtcacgaaacaaatgtacctttacatgtagaacatgttataccagaaatcctaaacgt gggtgtaaggcacaacccgcatatccaatctggccttggcatgtaagccatgtaataatg caaaaaaacctacagcctgtagaatggtagacaatgaatgtgactaaaaaagctat cgacaaaaagcgggcaaaaaacttgcctgaagtactaaaaacagctaaaaacccctga aagacgctgccataatgaacataacccgatgggagctattaaatgccttaaaaaacttgg cctatccgtagagagcggcaccggcgcaaggaacaaaaagcagcgtattgacataaa cttccaaaaaccactattatgacgcttgtgttggccctagtagtccgcaaaacgtcga actttacaaaaatattactaatttggaaagcgtatggcagagcagacagacagatgtga acagcgacaataacgcttccaaaaggccacagacaaaaaagaatcatttgcgtt ttcagacaaacgatattgtcgaagcggtatctcgcggttaataaccagggttaacacg ccggcgctgtgagttagaactagtggatatttgaattaaaaatattactggcaaccgt attgtcaaggcatcaaccacaaatatttcaattatcgcaagagctgatgggtggcagta cgaaaaataaaaaataacatag	gggtgcatacccaagaatgcttctccagttctt ggcactatggaacggagctttaaaccgagc aggggcacagcaaggtggtgcttctgtga gtcaacccctcccaacttggcaagagag acgggataaccgcttactagccccctacgg ggctccgaaagagtagttagtttatg
628-629	Okeania hirsuta strain PAB-10Feb10-1 PAB_NODE_716, whole genome shotgun sequence GENOME_ACESSION: GCA_003838195.1_AS M383819v1_genomic GENOME_ID: 204924 CONTIG_ID: 685 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgtccaactcaactaattatgtatctgtgctgacgcgagtaaaaaaccattaacacatg caaacctggcatggcaatgctgacgaatcaacccctaattgtaacagcaatgggacatg gttgcaggcagatgtacagatggacaataatgtgtttcccccgaagggttatcaggctaa aaaacctgtaccaggttggaaaactggagatatgatcaatgtcgtcaaaaggcaaaacat tggattaaaggagtcaggattaaaactgtcagaagtaaaagttaattttgacatccgacaf ggggataaaaatctgtctgtatctgaaatcatatccaacccattcacagacgagatggata caattactcgtttga	tcgaagactaaccacaaatctaatcaagcg ctcaaatcgttaaatggttagtcagccctac ttagcaataagtaaacgttatttggtcacaat acgtcagaatgcgtggccagttctgacctct attgtttggcattaaacaggttaagagatttg aaaaaccagtggtgtcgaacttaacagccaa aataaccaggcgaggccaactttacattttt gtaggaggacgcataaatgtccaactcaa ctaattat
630-631	Okeania sp. SIO2H7 2H7_NODE_770, whole genome shotgun sequence GENOME_ACESSION: GCA_010692585.1_AS M1069258v1_genomic GENOME_ID: 325546 CONTIG_ID: 3798 SOURCE: NCBI_Prokaryotes DATE: 2020-04-22.	atgtccaactcaactaattatgtatctgtgctgacgcgagtaaaaaaccattaacacatg caaacctggcatggcaatgctgacgaatcaacccctaattgtaacagcaatgggacatg gttgcaggcagatgtacagatggacaataatgtgtttcccccgaagggttatcaggctaa aaaacctgtaccaggttggaaaactggagatatgatcaatgtcgtcaaaaggcaaaacat tggattaaagtga	tcgaagactaaccacaaatctaatcaagcg ctcaaatcgttaaatggttagtcagccctac ttagcaataagtaaacgttatttggtcacaat acgtcgggatgcgtggccagttccgacctct attgtttggcattaaacaggttaagagatttg aaaaaccagtggtgtcgaacttaacagccaa aataaccaggcgaggccaactttacattttt gtaggaggacgcataaatgtccaactcaa ctaattat
632-633	Ga0256832_1033259 GENOME_ACESSION: 134614metaG_2_SF_330 0028400 GENOME_ID: 20589 CONTIG_ID: 33258 SOURCE: JGI DATE: 2019-01-27.	atggaaagttttgttaattaatagcatggcgaaacattgatgccatgtagctcaagaaaagc aagattattgcttgaagtggaagcgaagctaaaggttatccgccgtatgcttattatccaact aatcatgtttcaacaggttataagcaggatttaaccttaggtgttgataggggtcattcaga ggttgattatcagttgtatcaaaagacgaaggggtttttccagcagtagcgaagatgcgt aatgatatttcacgaagatggatacaagacgcgtgtatagaaggcagaaaaagaacaag cttagataccgtaaaccaagatttttgaaccgctctgccagtacaaaaaaggacgttttag ccccatctgtacaatggaaagtgagcgtcatatcaatttaattaatcaacttaaatcattact accgataaccaaggtgttttagaaaacgggtacatttgatatgctgaagataaagaaccca aacataacaatagcaatatcaaaaaggcgttcaatatgttttgagaatgttaaggctta tgttttaagtcgggatgctatcaatgccagagtaaaaagaaggatgtagcgacagatta caagttcatcatattaatatcgtctaatgtgttcaaatgcacctaacaacttgattacttt atgtgaaaaacatcataaagcattacacgttgcgaagtttagctagatattaataatcata aaagccttaaaatcagcaacgactatgaattatccgcagtcgattattacgtctatttccag aagcgattgagacatttggctacatcacaaagcaaacgcgtacgcataatattgaaaa aacgcatactaatgatcggtttgtattgctgtgggtcaaaagcagaagagagcagaaga aagaaccatccactttaagcgtaaaaataaccgttcaatacaaaaagacccaacggccta cgccccgtctattcggaggcaacgatatccgattcaacaaaagactgtgtcacatttgag ggtagacaatatcaggcgggtgggatgcaaaaaggggcttacctgaaaatgacaga tggattgagaaccatgttaagtcgtgaaaaaattgaaatagtgtttcatcaaaaagggtg tgattacgtatga	gtcaactacccctcagctaaagacagagg gcttgagtcgtgagatttaagataggcttctg gattgaatagtgattaggctaagttcctcga gaactacgtgttttagttatgacacccctgga tgccctccaaagtctgaggtctgtctcgaagt attaaaagacctcaacgggtcgtgtgcttg gattgacaagcctttacaacttgcggtggg aatcaacctcgcaagaggaatcgttcaatttt aacgattaaattttaaggttttaaatggaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
634-635	Ga0105041_100006 GENOME_ACESSION: LS_189metaG_2_\$F_330 0009983 GENOME_ID: 31076 CONTIG_ID: 5 SOURCE: JGI DATE: 2019-01-27.	atggcagttttgtactaaatcggataatcccccttagctcctgttcagagagaagggc cagaaagtactagaacagggaaggctccatgtgctacatagacctttcacgttgatc ctaaagtctgtggtgactgacacgtgagctgcccattgctgcaattgagctacgtatcga cccaggtagcaagttcacaggattcgctctagtcaacttggcacagggtgacgtgtgca ctgcgcagaaacttaacaccgaggttacaccataagttgcgacctgactcgaagatcgca acttcgccggagccggcggtccaggaacacacgtcaccgaccggtagatttaataatc gaactaaaccgaggggtgctcgcacatccctcgaacaccgtgtcttgacaacgttgt cctgggtaaaaaagtatttgagactatgcccaatacaaaatgacgccgtagaatcggtaaa atftgatacacagaagttgcaaaatccagaatatccggaagtgaataaccagcgagggcg aactgtttgaatcacgaagttcgtgagtgtattcaacaattcaatcacacctgctatact gtaaacgctctgtgtaccgtgacattgaacacgtgaccccaagatctttagggggac taatcagtcagcaacttagcggctgctgtagaacgtgtaatcagaaaaaggtaataag gccccgtgaacaattttacggacgccgaacaatcagcagaatcaaaatctgacccgggt ttccactgaaagatgcggcgccgtaaactaactcgaatgctctagttaagcagctatc gtctataggcaaacccgtgtcgtgtgacccgggggaagaacaaaatacaatcggactc aagctgaactgacaaagaccattgctacgatgccgcgtgtgttgggacaggcctgcc aaaaaagtacaaaagtgttgtaacgatcaagctatggcaggggtaatcgccagatg gtgctaacgacaaaatttggcttccctaaagcaagcctcctgacgaagatcaagagc gtccatggcatgtgtacaggggacatcgcaaaacccctgtgtgtacagggcgataac cgggtctagaacaaaagggaagctttccctgaaagtaagggaattgtatctatgaca ccccgtaaactaaaagtgttacaagaggcaacgggtacgagttcggtcaatcctgcgc agaagctgaatga	accaagccataatcccttctcttctaccc tttgaacaacctagcccgtggcaggggta cacacggactcctacagccataatccctct ctttctaccccttgaactcgtcccaaaaa ctctttgactttacgcgataattacagtactct gcgagaggttaattgtgaccgactcagttt gaggttaactcaagctacgtactcgaaaac caagactcacctaaagggtgccctaccagc ctttagctctgtgattcggaaatcaatgaccgc ttgcaaaaggggcaggcgtaatgatccggaa tcggtaaaatcgagataattgtctaggtaga caattacaagccgttattggcattttatttgg ca
636-637	Ga0370539_00069 GENOME_ACESSION: 20181212_24i_2_\$F_330 0034631 GENOME_ID: 304144 CONTIG_ID: 68 SOURCE: JGI DATE: 2019-11-15.	atgtctaactttgtctttgtatggatatacaatcccaaccgctcactctgtacagccggg atggcaaggtcattatataaaagcgggaaaaagcagcgggtatttcgccgtatccctttacga ttatfttgaacaaggcagtggtgactgagcactcgcagaagctcagtggtccaccttgcgga agtgtggctactgaccttcccttcgacacgtcagggcaatgcgactgcccttggcga gtcaggaaccacgtcagggcgaagtgtcgaattcgttgaagtgtgactcacaaccc gttgcgtatgtcgccgaagggtcatggaacccgccagatgtcggcactgataagtatg ggttccctcctggcatcgtcagcgtgacgattcataaaggcttccagactggagacct tgttgcagcaacagttacaacaggtgaagaaatcggctctatgtggggcgggttcttgc cgcgcttctggcaatttgaatgtccacttctcgggaagagtggtgctgtattagccaca atatgtccaagcaattcacaaaaaggatgttactcctatggattctga	gtagactgccgattctagcgggatacaag ccacaacatgaccagcctaagaccttcag gtctactgttttcacgcgtcatgatactaca atgcgtcgtatgtttgtactctatcgttggtt gtttaaacatctgtatgtgggttaaggaagtgc agccaacacaacaacacctgaaaaccttgg cgaagcgaacgttacgagcaattagaggga gacgcaacaatgtctaacttt
638-639	Okeania sp. SIO4D6 4D6_NODE_771, whole genome shotgun sequence GENOME_ACESSION: GCA_010672015.1_AS M1067201v1_genomic GENOME_ID: 325452 CONTIG_ID: 2197 SOURCE: NCBI_Prokaryotes DATE: 2020-04-22.	ttggtagctgaagaaatcaaggcttgcattaaaaattgacctgttctaaacaacag gctttgcaatgggtcacccaaagtgaagaagtattcttgaatgttttaaccatcgtggtc agcaaatcaagaatgctttggaacgacggcgaaactcttctgcaggtcgtcgtatcga aaactcgtaccgtgaagtgcagatttttcaatcgcaaacgaacaaaggatggctgccac caagtttgaggcatcgggtttgatcatgtccgttggggcgcttctgtgtaaaagttagc gtgatactacaggaatttaagtttttcttgcctaattgcttccctgttgttgcctgttgc gttgcctgttgccttctgttgttgcctgttgcctacttttctatacaataaccgatgctaccg gacatgatatgactgtagaacttgggttaataagactttgttaactaagtcgataagtagtt tgactatggaattgtcaagtttgacaccgagaataacagaaacccagaaatttcaggagt tgaatatcaaaaagggaactatttgggtatcatgtccggagatctattagaaaatgg ggctgattatctgctgattgtggggcaatagatacggccttagaagttgaacatatagttcc acgctcaaaagggtgtgtagtaacgggtgtctatatttagccatagcttgcatacaataacc aaaataagagtgctatggatattcgggaatttctgaaaaataaccgtctatactagctcat gttctaaagggttccaaaacgccattaaaagatgctgctgcagtcaacacaaccgcact caaatcttgaactttaaaagcaaaagggttgcgtgtgattaccggtagttggctgggtact aaatataaccgttgcgttgaattgccaaaagacattggtatagatgctgctggcgttg gtga	tttaggactaaccaaatctaatcaaacgc aaattttgttttaalcgttagtcagccac tttttcaagtaaacatttttggtaacatc gtccggatgctagggcagttccaaacctatt gttcggcattaaacagccaaagagatttgaa aagctagtgtgccgaacgttaaaagttaaaa taaccggggcaggccaactttacatctgaag aaattgttagctgaagaaatcaaggcttgc aattaa
640-641	Acidithiobacillus ferrivorans strain PQ33 Aferri_contig000098, whole genome shotgun sequence GENOME_ACESSION: GCA_001857665.2_AS M185766v2_genomic GENOME_ID: 111397 CONTIG_ID: 99 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgctgaataaagtgttttttggacacgcatgaacaccgctcatgccctgccatccgg cccgtgcgcgggagttgttacgaaggcagcgctggtcttccgtcgttccattca cgatcatcctgcgggaacgggttgcggagacaggcagaataaccgccaaatcgga tcccagacgaagacgaccgctggtgctgttagcggcggttcaagcgaggactcacc gtcatatggcggcgtgaactggcgaccggcgtcagactatcgtgccgccctggaga aacgcagtaatacagcggaattcccgaggaaccggaacacggtaaccgcccacgc gtttcgatcatcgcactcggcggagatgttggccaccatcattgagcaccgctctt aacgggtcatcactgtttcaaacgctcatgctggtggcgcccaactgattcagcatg ggcggggtgcgcttgacatcgaggctatggagaactcggagatatccggcgccgag accagcaggggactttgttcgttacgaagtgcggagtagctgctggagaagtggatg ctgacctgtgctattgctgcatgccagaatgtaccgtcgaatagaccattgtaccgcc gcagtatggcgcgagtgaccgggtgagtaacctggcgatagccgttcacgattgcaac caggcaaaaggacaatgctcgactatcggaattttgacagcggacaaggcagacagac aagacagcaggtttctgctgcagtatatgctggcaacgaccccaaaagcgccggag cgggaacgtcacgagtcgaactgctggagcgggtgctgaagcaggtcaaggcacca ttgcggatgccgcagcgggtcaatgtcacccgcaacatctgttcgagcgtttgtgaa	agcgtcatgaacacatggataaatcagg ggctgtagagccctgattcatgaccaccc cgagaaagctgaaggcgatcatgcttcaa gggtgtcatggcactccgggatgcttgcagt tccggactctccgttccaatcatgctgtc tacctggggatgtgcgaaggcgggcaac atgacaagcaccatggacatggcgaggc acaggttaacggcgcaagccgatatttaca gagtaatccttatgtctgaat

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ctgggggtgccggtggagacggggtccggcggcgccaccaggttaaccggagtcag cagcattatccgaaagcgcactggatagacggcgttgcgtgggtgaatccggagcttc agtttgcctgaatcttgaattaaaacccctacagatcaccgccaccgggtcatgctgcaga cagatgcaaaacatgacgaagaaaggggtcccagaggaaggaagcaaaagtcggcag aagacgtattttggccttcagacgggagacatgggtgcgggctattgtccaaaaggagg ttcgctgggaaacatgttggccgggtggcgtgcaagaatcaggaaacttcaactcaa ggcggcgggaaagagttggatgtgtttcatggcgtcactgcactccgggtcatagagg tgacggctatgcctacacattga	
642-643	Ga0209318_1005124 GENOME_ACESSION: RifGroE1SPAdes_SF_33 00025000 GENOME_ID: 34868 CONTIG_ID: 5123 SOURCE: JGI DATE: 2019-01-27.	atgacaaataataaaaaaagaatatattgttattggataaaaaataaaaaattatctc ctacactgttaataatcttgggtatcttataagaaaacaaagcaatattatttctaagtat ccaatgggtattaaaaataaaaaaagaaataaaatgatgaaaaataaatgataaatcaga gtttgttgggtatagatgatgggtctattcaccgttggaaattgctataattcaaaatgtaaaa ataaaaaataaggtgtattaaaggtatttgaacaaaggcaggatgtaagaaattatg gaattaaagaaagatatagacaatccgtagaaaacataaaagatatagacctgtaga ttcaataatcgttcatcttctaaaagaaaggtagattatgtaacactattaacaaaagaaa caagcaatattaaagtagtgaatcaactaaataaattgtagatatacataaaaaataattta gaagatgttaaaattgatattagacattgcaagaagggtcataaattatataaattggcaatat caaaaagtaataagtagatgagaatttgaagattgctactttaaagagagataaattatacc tgtcaagaatgtgtgaagaaagattgtagattagaagcacatcataaattgctagaagatt aaaaggttctgattcaataggtatttattactcttggatatttgcataaaaactgaag gtaattgaagaattatttattgaaaaatatacaaaaataaaaggtaaaaattatgattcg attatgttcaacacgttatgcaagggaataattttaaaacaggaaattaaagtaaaatagca gaattagaattgactataggaagtgaactgcaataaaaagaattgattggaataattaataa atctcatagtaaatgatgctattgtgattgtggtttaaacctgatacatgtaataaaaagaat ggattatfaaacctatgagaagacaaagtaagcaaaaacagataatgtattaggtataaa acatagagatttagttcttatacatttaaaaatggagaacacatattggatatgttacgctt tatatccgaattgaatgcattaaatttcaatacaaaaacaaacattgtaaaaagtaaatgt gaagaaattgtttacttggaaatacaacaaatttattgtgtgtgcataa	tagacaaatgtatagtagtggacatatatat atctatataacagactattgagtggtttcaa gcctaagttagtaccactaacgaaagtattt gtacgatgaactacgattatataaatgctaaa aacacacctttagatgtaattcagctattg ctctgtgagtacaaggtgaccgacttcta atgtctgaagtcaaacaccgaaatcatgt gcattatattgtcttggcaagagaaaattc tccgaaaggatggttagtcagaaatgacaaa taataaaataaaaaga
644-645	NODE_3724_length_286 10_cov_9.117721 GENOME_ACESSION: mgm4743560.3 GENOME_ID: 238062 CONTIG_ID: 1440 SOURCE: MG-RAST DATE: 2019-01-27.	atgcagaaaaacaaacagtagtgaatcgggaagaaacggccgtggccttatgccatg cacaccaagaagggccagaattcctcttaaaagaagaaggccaaagtagtaagaaga acgcccgttcacgatccgattgctgtacaagacagggctgtgcgatacagcccaacacgct gggctgtagactcgttccagcacatcggcattggcattgtttcagacgacattgtcttta gcaaaagcgagtagcaactccgtccacaaatggaagaaacgggaagctgatggaacaag aaaacagtagtccgctggcaggcggtatcgcaagacacgataccgccacccaaagttc ctgtccataccaaacggacatatgtggaagaaacaaacggaagcgaacatcctgcaca cactggaagaagaaactgttctgtatgtcaagccgcagacagggctggttccgctg tccgtgcagagcaagtgatcatcacattcggattatcaaccgctatatcgaagccctgc cgccaagcacaagagtcgttaattgagattggccggttgacatggcgacatgctgaatc cagatalccatggtgaggagtaccagcacggctgctgtatgatttgagaacaagaagg cttatctactcaagctgtacgactacaagtgccgatttgcagtaagaatttggctcaag aggctgcaggtacaactgtcaaaagcatgtatgcaccatgtcctgtaccgctcaaaaggc agtagcagacaatgccagtggttctgattcctgtatgtgacattgccacacggcagaagctc atcaggaaggcggaacactggacaaattgcagaaggctgctgccaagcagaacaacg gtatcagagcgctgctgtagtcaactatgatgaacgttggccaagcgtcttagtagc gtttcctgatgcctgtacaccttggaacatcacgcatccgacagagaatgatgcgt ttagagaagagccatgctaacgatcggttagcaattgccaagcatctgataaagggtat tcggcgactacacaatcaatgattgcattatcagctctgtataaacagggttcgaagaa gaagcggtcactgcatgaggctaatccccgaagcgagaagacgccaatcgtgag gcaaaacgcaatgcaagaatacaaaacagcggtttggcatctgtctgttgacaattgtc gctatgctggaatacaggcgttagtaacaggcttcaactctctgtagcagagcaacgtt gccggacggcacatatctgtgtgactgcagcagaaatcaccatcttgcggccat cgcaagttgtattctgcatcatacaacaactggcttgaaggaggctagtcataatga	acaagtaaccgagcaagtcaggacttccag tccgtacttgcgaagtgcaagtagagcacg cgtaaagcgctgcagatgactagccctagc agagaacacacggctctgatttccgca atgtgatgatgcaaaactacgatgccataa gcgcagatccaggcaacgtataccgcaca aacgagaatgtgcatcgacctgatttccaa gcattgcgaacttggcgaaggcgcaactgt aagtaactgg
646-647	2206908067 GENOME_ACESSION: IMG_2199352033_SF_2 199352033 GENOME_ID: 308158 CONTIG_ID: 638 SOURCE: JGI DATE: 2020-04-07.	atggctactgaggcgagagttatctgctagacacaaacggcgaacacttgcgcctacc atgcccgaagggttaggcgttctcaatgcaggcaaggctgccgttttagggcgctg gccattcacatcatctgaagcgcgcggttgaagctgtatccacagcccgcgattgaattc aaggctgatcctgtagtgcacatcggggcttccctcattgggaatttcccaagcaa ggcggggtgtcctgtggcgcccaatctgcatcaccggcgaggctatcaaggaca ggctcaaaccccgccgtcacttgcctgtagccgtcgtctgcgaagacgggtatcgt gcaccgcgatttgataaccgaacggcgccgaagggttggtaaccgctctttacaatcg cggtgcttaacgtcgttctgtgttgagaagctactgtataggcgagactactgagt gccatcgaacggcgtggttggacacaaagcattgcatctctgagatattccggag ttgaatatcagcaaggcgagctcggggtcacgaggttcggagtagctgtagaaaaa tggcagcgcaagtgcgcctattgcgggaaggagaattgtccattggagatcgaacatatt aaccggcttccaatgcggaatcaaacggagtgagcaatttaacgcttgccttgcctg gtaattgagaagaagcaatcgcgagtgccgagagtttctggagaaaaagcccgatgc	ctgcacatccatgtgttctgtcgtatcatt ttcgtcaatgtgtccgcatattgactgctc aacagactcagcggtatggccctacgt tatcgagagctcaagaacacactccggg gtgcttaaccagctccggacctgtaacgga atcgttaaacagtgcggaccgtgggcaacg gacagtgcggttgggtgcgctgtgataa cattgtcgaggttaccattacttagttgtcg agagcttattgctactgagggcag

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gcttaagcgcatttgacgcacgctaaagcaccacttaaggatccgctgcagtaaatgc aacgcgctacgccatagggaatgcaatgcgcgctattgggctaccaacttcg	
648-649	Ga0272428_1021842 GENOME_ACESSION: FingerMnordMG_2_\$F_3300031520 GENOME_ID: 280175 CONTIG_ID: 21841 SOURCE: JGI DATE: 2019-06-13.	atgagcaacgtgtttgttttagataccatgcataagaccacttgacctgatacaccaggact agcacgtaagtattgaccagtggaagcgcggtgtttgtcgtatccgttcacgac atactcaaaaaaagtgagggaaccaataacacccgttacgcatcaagcttgacc aggcagtaagacaacaggcatagcgatagtaacgatgcacggcgaggctgtatgg gccgcagaactggcgcatcgagggaagcaataaaagacgcatagagagtagacgt gccatagccgttcacgtcgagcgtcacacacgctatcgcaagccacgctttgacaat cgacgccgcaaaagggtgtggttcacccgtcgtagagagtcgtatccatccatcgt acatgggtcaaccgctatcaggtattgcccataagcccaatctcacaagagttatca agtttgatgcaacttatgcacaacccccgaaataagcgcgctagaataatcagcaagg aattacaaggctacgaactcggaataacctgcttgaagggtggggcgagtggtgtgcct attgcggcggtgaaggatgtgtcattgcaaatgagcatatccaggcaagagccaacggc ggttccaaccgtgtgagcaacctcacacttgcgtgtgacacatgtaattgcaagaaagg acacaagacatacgggacttctgaagggaagcccgatagtggaacggcgacttgc tcacgcgaagcccccttaaggacgcggcgccggtgaacgcgacacgagtggtg tacgaacgcttgaaagcaagggttaccgttgagacaggaaaggagacgacca agttcaaccgcatgacgcaaggcttagacaagcagcattgggtagacgcgcgtgtgc ggggcgctcaactcctgtcctctgattgatcggttacgtccctgtctcatcaaggcaacag ggcatggcaaccgtcagatgtgtgtgatgataatagcgttccctgcactggcccca aatccgcgaagtggtacacgggttttcaaacgggagataggtcaaaagccgtgttacc cgggcaagaaagggtggacgtacatcgccgcgtcgagtagggcaacgggcagc ttcaatgtgacgacgaaacacatgggacgtacaggcattggtataaagctgtgtga atactacataaaaccgatggctatgttactaa	gtcaagaacccacgcagagaatgcgggg gcttgtgaaagcaagcctgtcttgaccagtct cagggaacctacgcagagggaagaatgtat aggtagctcaggtgacgacacgtcttgac cactacggttgggaattaaacgtctgtcgg ggtgataggcaggtgttccaacaagaacc ttccatagcattgacgaggtactcattacct ccgaaaggagaggctcatgatgagcaac
650-651	Ga0310134_003633 GENOME_ACESSION: M172_2_\$F_3300033999 GENOME_ID: 280646 CONTIG_ID: 3632 SOURCE: JGI DATE: 2019-06-13.	atggtatagcttatataaaaaacgggaacccctcatccgacaaaaggcatggcctg gtgaggatactgctaaagagaaaaagcgctgtgtatgacgcgcaagccgttactat acagctgtgtgacgacgacgacatacacacaagacataacacgggtttgacacag gccgagcgtaccagtcaatcacagccgtaaatgccggacagcgaggtttgtactcgt ctgtgctggaacgcggacaaggaagtaccgaagcttatgaaatcgccaaaatgtat cggggcgataagcggcacaacaggcggtgaaaaaagtaagacgcgtgtgaaggaa aaaacatactcagagcgccgaagaaagtgtgtgtggccggcgcaaaagagccgataa ccgcaagtatatacaacccaagagcgcggttcaataacagggaagagccggaag gctggtaacgccgacagcagtgcatcttttgcggaccacctcaactattttaagaaggt cagggaagatactgccgattaaagcgtgtgtgtggaatacggaaaatttgacattcagaa gctagagaacctgatacaaaagtgccgaataaacgcggcaaatgtacggctaca ataatctggggaatatgttatagcggtacagcagggcaaatgcctgttatgtgaaaagaa gccgatagaacacctgcaccacgtaatgccgggtcaaaagaaggcagcgacacatac aaaaacatagccgggctgtgcaagtgccacgcaaaaggtgcacaaagcccaaaa gcaaaagaaaagctggccgagaagagtcggcgactgcaaaagaatacggcgacacg agtatactcaacataatcatgccatactgtacagcgagctaaaatcgtgctggcgacgg aaaacatagcattatgctacggctacgaaaccgagggcagcaggaatcttccgggtg gctaaaagccacagcaacgacagctatgcaatggcgcttatggcaatcggacagcgctc caggatagagaagatagagccgtaccgggtacaacaatacagcgctcacaacagggc gttttgcgatgccagcgacaggctgtacaagaaggacgggaagatagcgcgag gaacagacgccgagacggaacagggaaggcatctcttgcgggaataccgcagcga gcttataactgcgttggcaagaagattgcgacgagagagatataaagctcaagggtga tagagcagtcagaaaaagacacgtatcaatcaaaagatgtgctatttccaccagatcg agtcactacaaggaagacagcaggtgtcaaaagtttttcaacaaggcagctcgtc gatattggaagcatttctgggtatgtgccggaaggactgtcagcttattactaagaaa gcggggatagtatgtctgtga	Gtcaactaccccgcttaagacggaggc atgaggtttcgttaaccaagtgtgttacctcgcg ggtacagtagttgaacagcgccgcatgctg gctgggactccaggatgccactccacgtc ccggacactgccgacgagccgcatcaagc cgggggagaccaaccggctaacgatagcc gagcaattacctaagagaggtgcaaaagga aatg

SEQ ID NO:	Name	IscB sequence	ωRNAs
652-653	Ga0247610_10094209 GENOME_ACESSION: 174DNAGHGhighp2_2_SF_3300028833 GENOME_ID: 279759 CONTIG_ID: 94202 SOURCE: JGI DATE: 2019-06-13.	atgatatagtattgaataaaggaggtaaacctcttatgccgacttctcggcatggcaaggt aagacgtttattgctgataaagccgagtaaatgaattcaaccctttactattcaatta actacggagacaagaaatgaagtagatgaagtttcattaggtgtgataccggttatagata tattggttatcagccacaactaaagataaagcttatttgagtgcaagcagaattaagaat tgatttagtgacaatttaactgctagaagagaattaagacgtaccagacgaatagaagat taagatacagaaaacctagattcttaaatgtaaacgtgaaaaaggatggttcctccgct attaaaaacagattggaatgccataaaacatttgcataaaagacacaaagttttgcctgta ggcaaaactattgtagaagtgcataatcagataacaaagaaatcctgatataaa aggctgctgagatcagcagggtgatcaattaggttttgaatgtacgtgctttgtttat agagataattatactgtcaatgctgttttggtgaatccggtaatgataaattacaagctcat atttaaaacaaagaaagatggaggaagtaatgctcctgacaaacttaattctttatgtgac ccatgtcataatgataatcgtctaaaaggaaaacattaaatgttgacctaaagggaatc ttataacgcagaaccccttatgaatacattgcgtaaatctgttagagaccttaagaaat atacaaaaatgaagtgttactatgtgttatacaaaaacagtaagaattgaaaataaatta gaaaaagaccataatgtatgctagatgtattcaggaaatccttagcaaacctaatgg tatttactatttaaaaggtaagatgtcataatagacaactcataaattgtaaacattaca gggtggaataagaaaacttaatacatcattatagtagacatgggttcagactgtttgataaa ttaagataaataatcagatagatttatatatggcagaagacaacaggatatttgcatt aaaaatagatgggaaaactattcaaaaagtatttcatataaaaagttaaatctcgttgaa aaacgaaaagggtggtgtgttttaaacatctgaattgtaa	attgacatagagtacaagtgaacgcataca agttagcaaaagtaactgtatataaataaactt tacaattgtttatattctttacatagctcaagt atttctgaaaaagtaagtaactacgttatg agagaatgactgtaaaaagctcaaatgttac cttcggatgttcatctagctgaagcttaag gtgtatgattaaacagtcgggtattgctgacag tgtgtacacaacaacacctcaataacattg acgaaggtgcttatcgttttgcagtatctgg ctttgtgaaagcataaactgaacctattttaa ggaattttaaatg
654-655	Ga0265292_1000080 GENOME_ACESSION: Munlanlewell296A_2_SF_3300028028 GENOME_ID: 32898 CONTIG_ID: 79 SOURCE: JGI DATE: 2019-01-27.	atgaaaaccacaacagagttttcgtttgacacaaatcataaacattgacattatgccg accctgttagggcaagcgaattgttatggacaagaaagcggcagttactcgcgcgatac cattcacaaattatttgaagaatcaaaactaaatgctgtgttgaatccgtaaaaggtaaggt tgatcccggtgcaaaatactactggaattgcgttagtagacataaaaaagaattgttatatt cgcatcggaaacttgaaacaccgtgggttgaattaaatctcaatagattcccgccgctcac agcggcggtctcgtcgaatcacaatcagatatacgcagggaagattgatcaccgga ccagaacagacaatggttgcgcgtcattacggcatcgtgtagaacaacaatgacct ggattaatagattatgcatgttgcacctattgaatcgtgtctatagacacgcgaatttg atattcagaaaatgcagaaccccgaataatcgggtgtagggatcaacaaggcgaactat ttggctatgaagtgcagagatctacttgaaggtggggcggaaggtgtgcatattgcg ggaaagagaatattccgcttgaagtgaacatagatttctaaagcgaaggcggaacac acagagtgtctaatcacaattctcgtccgagaatgtaacgagtcgaaggccaatcacc ttagcagaatttttagcagataagtcgatgtccttaaacgcattcttagcccaagccaaat ctccgttaaaaagtgcggcaataaccaagccaccctgaatgtttacattagcgtatgt cgataccaatcttatagtgaiaaccggaactggcgacaaaacgaagatgaaccgggttaa agttgggataaaaaagcccaattggattgatgtcgtgtgtcggcagttccggcgtggc agttaatctgaatgtcggatgatgtccgatgtcattaaatcgaaggacatggcaatcggc agatgtgcggaacggataaattcggctttcaataagacatcgtcagggacaaaaagagt attttggtttgagactggcgatatagttaagcggatgtcctcgcggcgaaaaaaacgg aagataattggacgagtgctgtgccgtaaaagtggctcattgatattcaacgctatcg gtcgaattgaaggaaattcttggaggtgttgcgaagtagtccatgccaatgacggctatag ttatagacaaggggagggaacattcctctcatgtctaa	gtaatgaacccacgacttcagtcgagggct tgtagaagctccaattcattaccagcccaagt cccgtatttccgggctacgtttaattggaga tggtacctacgaatcgtgcaggttcgtgac aataccgctgttttcatgctgcgtagatga ggatacggcggaaggaggtggcgtaacaa ctcttttagacattggcgaggcaaacattact cccgaagggaatggacggctgggattaa cccagcaacaactgtgtgtggcggggag caatactctcccgacttcagataaacgag tcggagtcgtaaacatcggggag
656-657	Ga0257070_1000485 GENOME_ACESSION: GoafecGen0Rep3v2_2_SF_3300023711 GENOME_ID: 280332 CONTIG_ID: 484 SOURCE: JGI DATE: 2019-06-13.	atggtttatgttcaaatcaaacggacagcctataatgccgacatcgaacctgcaaaagg ttcgtattctcttaaaaacaggtgaaggcaaaagttaaccacaggtgtccgtttaccatacagt tacagtatatagtagcacgaactatacacagggaagtgacttggaaattgatgcagggaagca agcatacggagtatccgctacaacagaaagtagagttttatgaagccgatgttgagctt agaacgatatagtgatttactgtccacacgtgacagaatcgcagagcaagaagaag ccgtaagaccctgtaccgttaaccccgggtcaacaatagatttccgctaaaaaagaagg atggttagcacctctgtaaaactgaaagttaaatcacttaactgtcatagcgaaagtac ataagatacttccatatacgaataatcgttgggaacggcatctttcgaattcagaagatta aaaaccctgcgataagcgtacagaataatcagcagggcgaacaattaaacttctggaatg tcagggaatatgtgctttcagataggttcacacctgtcagtgctgcaaaaggcaaatcaaa agacaaaatccttaacgtgcatcatatagagtcgagacatacaggtggagacgaccca ataatctgattactctatgctcagctgtcataaggacaccacaacggtagtgcagttta ccaaaaccatcagacgtggaatgtcttcaaggatgcaacgtttatgggtatcatgcggt ggtctgtctatacaaaactgaaagcgatctacccaatgttaaatctacatacgggttatatt acaaagataactcgaattgagaatgtttaccaaaagagcattatagatgcccgtgttat aagtgtcattcgttagcggtgaagggtggcgaagtgttctatcagaagaaagacgtgttc acaatcgtcagatacacaagcaccataaacaaggtggcaagcgcaaacctaatcaa tgtccatacgaagtacaaggatacggctttttgacaaagtctgtataaggacaggaatg cttcataatcggtagcgtagggatggaagattcgcctcaggtgtgtagacggaaacaa actaatgaacagatcacatacagaaaactgagattttagaatcgcacaaactttatcac agaaaaggagacgccgtctctatcggggcaagccccgacgttcccgcgctgtttaa	actaatccacgggcaagcccgtgggggt gtgcaaacacccggagtgtagtgcctaag tgaagctttttgactgaactgttaagagag aatatatagtcacctacgagcataatgcctaa ctttagtctgtgaggtcagtaattaaacaatc ctgcggtataggataggtgtctgatataaa accttcaattaaactggcgaaggcgtgcaa ccgctctcagagcagtaattaaatcttat gggatttggaaaggagcgaaccttatg

SEQ ID NO:	Name	IscB sequence	ωRNAs
658-659	uncultured Erysipelotrichaceae bacterium isolate RUG10685 genome assembly, contig: RUG10685_asm_16, whole genome shotgun sequence GENOME_ACESSION: GCA_902761625.1_Rum en_uncultured_genome_ RUG10685_genomic GENOME_ID: 344834 CONTIG_ID: 15 SOURCE: NCBI_Prokaryotes DATE: 2020-04-22.	atgtggaccgccataaggacgagaacccgtaacccgttggggataaggacgtgcg taagcacacaactttacagtagcaggaggagcctctgcatgaaacagtattgtattg cgaatgacggcacacggctgatgcccgaacataaataaacgtgcaagcggttgatgaaa cgagagaagccgtcatttgcagacatgatccattcactattaagctcaccagagatttgc aacataatgttcaggatatcgagtttaacaggatagcggcgataagcatattgcatatcc gtctgctccgaaaaacatgaatatacagtgctcagtgatgctcctgaaagatgaacaa agaagcacacgatcagcgcaagctcgcgaacccgcagaacccgcaggcgctatc gcaagccgaggtttgacaacccgaaagaagacaggggaggttgcgccaggtattga gcacaagaaggaaactgcattgacagctctcgcagaatacaatgatgtctgtccaatcaaa aaagcgggtatttgaagtcggaagctatgatccatgcgaatgcaggaatatgaacagaac gggtcaggtctgacaggcactgattacagaaaggacaaagatagcgcatgaccacatt cgtgaagcagtgctgatacaccagaattatctgctcctctgcagaagagcttctatcgg gcaagaactgctattcatcagaggattccgaacaggcgacagatcaaacagactgaa taatctgatggcagttcagcagtggaacacacatcagccaatcacaacccggcgatt attatgggacatcaagcctgatcacaggccattcaagggtgcggcggtcatgaacatcgt agaaaagcaattgcggatgaaatcgagaacggcagctccgaatgttgcctgatacgaac ttatggcgagagacaaagctccgcaggcagagatctgcacatccgaaatcacatcgga atgatgcctatgcaatggcgagtatcatcctaacacagatcgagacaatgcatttca gaagcagcgagaacaacccgatactgagcaaatttatgattcgaagtacatcgtatgt cgggacggcaggaaagaatcagggtgccgagcttagctgcggacgaacaacagatgc gtgccccgcaacaatccggagaacaacagagattatcagagggcagaagctcgttaag gcagagatcatgtaagaagacggagataccaattacacgcaggagatgctgttattgca acagaaaaaagctgattgccagcgccactgcacattcggtaatacgtgacacttgcgc aaaggtcataaagatgtaaaacatcacaagtgaggatccgctgccatgcagcggatg ggtagacagataaa	gtcaataaccccgctaaagttaatgaacta taggtgagcgttcgggaaaacgcaggcctg attgattagcctgaggtcttcgagcactgcgt tatacaggaaatgfataggcacttcagaatgct ccacaagttctgaacactgcgagcccgagt aaacatctctgaggcgaggagaaatcgga ggcagctctgatacttcggatagcaaaaa ccctgtataactttggcgtatggacccgctc cataaggacgagaacccgtaacccgtttgg ggataaggacgtgcgtgaagcacacaacttta cagtagcaggaggagcctgcatgaaa acagta
660-661	Ga0223845_11721960 GENOME_ACESSION: 3300021387 GENOME_ID: 241586 CONTIG_ID: 22774 SOURCE: JGI_Dark_Harvest DATE: 2019-01-27.	atgcctaccctccgcaggcaccacgtggaaaagctcttaaggcgccgggaaggcaagg gtcgtggaaatagctccctctgcatcacagctcatgtatgactccctgcgtgacacagc ccctctacggcgagacagaccccggaaggaccaacatcggcaacgccgtcatggagc gcaggggcagacgtggtctacaaggaccacgtgacgacaggaacaaggagatttcaaa gctcatggctgccaggaggcagtagcaggtggcctcccgaggggagagcgccgtgc caggaaagcgctcgcgaagagctcggcactacgatgaaagctgactggaaagggtgta ctgcccggctgtgacggcgctgcaagggtgaaggacatcatcaacactgaggcacgggt caacaaccgcagccgccgaagggtcgggtaacgccttctgtaaaacagctcatccgc accatgtgaatatgtccggcgatcagggaagtacgtccgggtggaacattggacact ggaacttaacaggtttgccttcatgcagctggatgacgggttccgttatggcaccgactcc agaacgggaagctcaggggctaccgtgacgttaaaggactatctggcacctgcagga agggaagtgcctgtgctgcaaaaaggcaggcatagagcattaccaccacatcctgccaa ggcacaagaacgggagcgaccgtggtacaacctggcagggtgtgcacatcatgcc acgacaaggtccaccggggggagatatacaaaaagggaaggacccgccggcgct ataccgggacatccgtcctgaaccaggccatccggggatcctgaaggaaatggagga aatgttccctctgtgacacatgcaccggcagggaacacgcccgttaaggagctg ctctctgtgagaaagaccacacggcagatgctgtgcatcggcgctacggggcacc agtatctggcgtgacagacaatgcgcatatcttaaggtaaggcagttccggcgccatga ccgagcccggtgcaacagccagctgagcgacacctaaggtagacagcgggcttcaat aaaaagggggaaccggagtagaccattatcgcgaagaaccggaagccccgctatgaa cagaaggggcctccctgagccagctggggctgtccaggcagcagatattcggtagctgc acgtcgaaaagagcaaacggcactgtaatacaccggggccgcatgatgcctggagctgt attcatttataagggtgaacgggtacgtgatccggccagctgacaaaaggaaaataattc catgcagcaggtcggggaaaaagattcccgacagcaaatgcagcatctgcagcc acaatagaggattggtacatagcgtaa	gtcaatgaccacgactaaagtacgggctt gcagataaaaagaagaagcactgtttcatc ttttctgtcatcgtggctatcatcatattg agcagaggcggtgacgcgccctcaccgg gggtgcttttagccccggcgatggcaac aggcgccgaagcataaggaacatttttg gcagtcggagccttacaagctgtgaaaca ctgccccttatcaggaaggaggtaccagt ggggtac
662-663	Ga0118727_1075366 GENOME_ACESSION: IMG_3300009128_SF_3 300009128 GENOME_ID: 305857 CONTIG_ID: 75365 SOURCE: JGI DATE: 2019-11-15.	atgttagtgcatgttttaataagcatggtaaacctttgatgccatgcgaacctcgaaagc aagaatccttctcagagaaggaaaagcaaacgtgcaaaagcaagactggctattttac aattcaattattatcgggaagcagcgggttacaacaagatattgtatcggaatagatag cgcgctaaaagagtgccagtagctgctgttgtaaatggtaagtgattacgcaaaaggaaa aaattcttagcagcgactgaagaacaattgtctgacagagtaagttacagacgtacaag gagaagtaggaaaacgagataccgtaaacctcgggttctgaacagaacaagaccaaat gtgccagatgtggcatcaataacgtgccaaagcgttggaaagaagtaaaacgcaaaaat ggcaagagcagaagaagggtgtgcagtgccagagcgcaactttgtcgtaatgtcaag gcaagaaggcacacacagaggccgcatattctgcgccatctgttaaaatcgtgccg aaagcatccttaacgacattcacagggttatctcagtcacttctatcctaagatagctgtg agatagcatcttttgatacgcagaaaatggctgatgccatgataaaagggtgtgagtagc catggaacgtgttgcgctatgaagtcaagcagtagcctcctaacagtagacacagcataagt gcgcatgctgcggtgattatcagaagacaattgtattgcaggtggagcacatataccac agaccaaggcggaaccgataaggttagcaatctgacatctcctgccgagtagtaattg aggccaaagggaagttgacattaggtcaatggagagaatgttgcgtctttcccaagtg	gtcaagaacccctcctgattgttccaatcaga aggagcttgcagtaggcaactcttgcaacccc tatattgacagcctaagttcttttagatacta cgttatccagttatcacaccttgggtgttca ccagcccatagaagtagtttgcctctgtggg ggctctgtaacagtcgtgagaggttagcga cagtaacccccaggacgacctgacattttag gcaagctggataacattggcgaggtgtgtgc gtacaattactactccgaaggagctgcact ttatgta

SEQ ID NO:	Name	IscB sequence	ωRNAs
		agattaatgagaagcgattgaaaaacatacctgctattagggaagcagagcaaacccaagaaggattccagtagacgagcggtgactcaaaagtataagaattac	
664-665	Ga0326512_10006706 GENOME_ACESSION: 1_2617_0518_D_2_\$F_3_300032038 GENOME_ID: 279770 CONTIG_ID: 6705 SOURCE: JGI DATE: 2019-06-13.	atggtttatgtattacataaaaaatggaaaccttttaagcctactaaaaatcattctaaagtaa gatgtttattacaataaacaagcctaaagttgtaaatcagaaccttttacaattcagctttta tatgatacttctgaatatgttcaggacgttaatttaggagtagatagtggtgtgaaacacattg gcatttctgctactactgaaaaagatgtttatttgaagcgtgtgtagaagaatcagataaaac ttctgaaatgttacaaagaagctatgcttagatgtactcgtagaagtcgtaaaacaaggt acagaaaaccacgttttatgaatagaactaagcttaaaaaagaaggatggatagcccttc tataaaacatactgttcaaacacatttaacataaataaagtagacataaaatattacattata aaaatattattgtagaataagcaaaattttgatattcataaaccttaaaatcctaatattaacggt aaagaatatcaggaaggagaaatgaagattattataatgtaaaacaatatgttttagatag ggacaatcatattgtcaatgtttgaaggtaaatccaaatgtaagaagttaaatgtacatca tataaagtttagaagtaacggcggtggaaattcacctttaaaccttataactttatgtgaagac gtcatagtgatttacagcataaataaaataacattaccagaaaaagtataaataaataaaca tttaaatctctacacatatgaatataatgaaaaatactttaattaaagaattaaacaattatat tcaaatgtagatattacatacggctataatacaaaatattatagagaatcataatcttccga aagaacattatagatgtcagatgtatttcaggaaattataccggcagaatcattaggatatt attatattttaaaaaacttcgtgtcataatagacagattcataatgtataaataaatacata atgtaagaccttttagcaaaattatcttatgctatttatggattagaatgtttgataaagtttata tgataataaagaatattttgtttatagcagaagaaataatggcgtgttcagtataagctctttat atgacctaaaaacataatttcagaaaacataaaaaacttaaaccttatagaacctgttaaa catttagtgtgtcaaaaagtatttttaa	ataaaaatttataaagttataatcctaaaaatt aaaaacttttaggttagtaattagctcaagtct gaaataagaactacgttatttgagaatattata tagttaccatgggtgtaataccaagcctgtg gctctaaggtaaatagttaaacattcttattt gtggggaatagtgctatttatttttaaacctc aaagtaacattgacgatgtttttaccacatt tatagtgagcttatttaagtgtttaaacactttt aaagaaggaagggttttaaatg
666-667	Ga0209777_10000441 GENOME_ACESSION: FrelakHBHSPAdes_\$F_3300027896 GENOME_ID: 26210 CONTIG_ID: 440 SOURCE: JGI DATE: 2019-01-27.	atgcagaagtttaagaagagtttcagaacgtacacgaatgctcactagtttgtagctct acaaatttgatgttaacagagaggacactcgtcgtcatcaaatcgcctcgaagaagc acactgactttaaacaatcccgaagtgaatcaggttcaacaacaggacgtcattaaaa gtatttgtttgtactaatatgcaaggcattccttaattgcttcttatgctaaatcaaaa cgatttaataaaaaaggagctgctaaggtaattaaaggatttcttactatccacactcaatt tgagtgtgaaatcacactcaaaacataaatttaggaatagattcgtgttatgaaacatcg gttttctgcttactgaaaaacaagaattatttcaggcactttaattttatagtgaaaaact aaagaaagattagaagaaaaaggatataccgaagggaagaatataaattatggtga cagaaaatctagatttaataatagaagaaaaataaaaaataacattacctccagtataga gaaaaacacaaactcatttaatttaataaaaaataaaaaacaaattcttctatttctgaaatt acatagaattaggaatttttaattgttcaaaaattagaataacacacattaggggagcaca tatacgcaaggaaatctctataaatacaaaatagcgaaattatttttgaatacaaaaatg gaaaatgtctttttgtaaaaaagattgaaaggattttcttctcatattcaccacattaaatca gaaataataacggaaataatagagcagaaaaatttaactttatctcataaaaaatgtcatgta aaatacataaaaaatctttagataaaaaatcaaaatcaaaattctaaagattataaacaatctac ttttatgaatataaattagtaagagatttcaaaaaatcttccagataaattaggatatttgat atgttacttctatgaaaaagaattaggaattagaaaaaacattttaattatgctttcgtga atagctagaggaaataatcaaaattagatgaagctcattgaaattaaacaatacatagaaa taacagagttttacaattaaatagaaaaaggatttaagcctctattaaaaaaaataaatctaaa gtaaatcctggagattttttggattggaaaaaagaataactgttaaggaaatgtttaact atagtagatataattcttttgaaaaatgaataaaaaagaataattttaatttcagaataaca aaaattttcattttggaagttttgcattggaattttaa	gtcaattgtctatccataatagattaaaaatcgt gagattttacgtagaattgattgggagcgt aaaagaaatttttatgcagaagtttaagaag agtttcagaacgtacacgaatgcttacta gtttgtagctctacaatttgattgtaaacaga gaggacactcctagtcacataaatgcctcga aaagaaacacagctacttaacaacatcccga agtgaatcaggttcaacaacaggacgctc attaaaagtatttgtgt
668-669	Ga0268280_1002972 GENOME_ACESSION: saktaG_18_\$F_33000282 98 GENOME_ID: 37736 CONTIG_ID: 2971 SOURCE: JGI DATE: 2019-01-27.	ttgaaagtattttagttgataaaaaacccaacccctgtgaccagtaattctgcgggttgca cgtatcttcttagagagaagaaagcgactgtctataagcattatccatttgaattaaattga aagttagcttcagacatagagccacaggcttacagttgaagattgaccaggttcaagg taactgtgtcagcaatcgtgaaccagggaacaggagaataatatttgggtcgggaattgg aacaccgtggatggataaataaagagacttggatagtaggcgaaagtcagacaatttta gaagggtatcgtacagtttagtatagacctgtcagatatttaataaaaaacccctagttgg gtggatagctcgtctctaatgagttagatatacaaattttaacttgggttaaaaagactgtt aatttatactaatatcaagctcgtcagtggaataaagcacatttgacattcaagaatga atgatcccaataatcttaagagagaataaccagaggggagaattatttggatttgacgctaga cattactttatgcaaaaataaattatacttgtttatttgggaagtcgggggtggctctttcga attagaccatgtttccaaggtcaaaagggtggaacaaacaagataagtaattctgttttag ctgtgaagaatgtaataaaaaagggaaaaatctttctagatgatttttagcaacaaaac ccggattatgccagaagataaaaaatacaatttctcgtctttacaacacgcagcagctgtga aatattactaataatcggctgtgtaaaaaagaattactacaatttaattaccattggacgtgggg ttggtagccagactagtttattcggaaattctcaaaaaatgaagaagaacactgtagatgat gctgcgtgtgttggaacacttaacacgatcttattataatgataacttttgggttttaata aaagcatatggctgtgtaagcgaataatgtgtcagaccaataatacgaataccgaca aagtatagagaagagcaaaaattagatttgggtttcaaacctggagatattgtaaaagcgtt agttcctagaggtaaaaaataaagggttttaaaagggaagggaactgttagaaaggcc cacatttcatattgattgtcagattctataaatcctaataatagacttacttcaaaaa	cccttgaaaactgataccatgatttgaatttt tttgatacaacttttgaatctgtcttaacttt gtgtattcagaagggttaccgggctaagtga tgatagcaatctgaactacgttatttcggtc acgacaccatgggtgctcaccagtcctcat gcaactgtcgcctcgggttaaacagggtaat tccagtgctgcaagcttaacaagccaagata tacattgccaaggtgacgtgtggaacacac accacccctttaatgggagagtatttaattga aagt

SEQ ID NO:	Name	IscB sequence	ωRNAs
670-671	LHMISPF_alex1_c100 GENOME_ACESSION: IMG_2049941002_SF_2049941002 GENOME_ID: 27373 CONTIG_ID: 2 SOURCE: JGI DATE: 2019-01-27.	atgtctaaaaccccaatcaagtattctgacttgacaacagcaagaaacctctaaccccg gcctaccagttagatgtagaaaactattaaatcgaggtaagctaagggttttagactgtac ccatttacgattactctgaaaaaaacaggttgatgaacaaccggaaaccaatagaattcgca tagatccaggtagtaaaactacaggcctagcttggtagcttaacccaagtaatttgggg ggctgaacttacccatcgggggcacgcgatcaaatcccgggttgaacccgcagcgcc atccggcgagccgtcgtatcgtaaaactcgtatcgagccacgatttctaaatcgca cccggccgaaaggttggctacctccctgttttgaacatcgagtgtaaccacaattacatg ggtaaatcgctttgttaagtagcacaatgaagtcataagcttttgaagtgtaaatcttgata ctcaaaaaatgcaaaaccagagggtttctgtgtgaatcaaacagggggaaactgcaag gttatgaggtagagaatactactctgaaaagtgaatcgcaaatgtgcttactgtggcgct gaaaaattaccattgcaagttgagcatattaacccacgcgtctctgtgtgtagcaaccgcgt ttctaacctatgcttggcctgtaaaaaatgcaacacaaaaggggactcaagacatcaa agactttctgaaggtgaagcctgagattctgaaaagaataatgagtaagctaagaccac gttaagggatcgggcgccggctcaatgggactaaatctgcttcttcaagctttaaaggct acgaaactagagcctatttatgtgggacggggtctctactaagttcaatcgactaaat agggttaccaaaaggctcattgtagatcgggcttgcgtgtggagatgtgactcagggtta actttggcaatcaatcaaccttgttaataaagctatgggttgggttagtcgccaagttgta cagactgataaatatgtttgactcgcaagggttatcgcttaagcaaaagtcgaaggtt ggaataccggagatgatctcagtgattgggtggaacatgaggggtgaaatgcaag ggaattaaaactacaaggttaagggttaactttgataatcgagtaaacgatactactgtgatt agcgtttcgagaatacaaatgaagccgtttacagaaatgacggatacaactactcattgt aggaaattgtcgcgaatgtccgagaaaagagtgagagcaataggatag	agttacgggttcgcttaggttctatagactact caaatcttagacatttccgaataaacccga aattgctaaagtaagtcaggctttctattgg caagcgctcaagttcataccttgagggtcgt gccagccccagctctataactcggcagttta aacaggcaaggatttttgaagcaggtgtg ctgagaaagtaccaccaataacaatgccg aggccacctttactaactaggaatagacgca accatgtctaaaaccccaatcaa
672-673	Ga0307249_10042978 GENOME_ACESSION: GoaFecpelsamples_SF_3300029305 GENOME_ID: 280327 CONTIG_ID: 42977 SOURCE: JGI DATE: 2019-06-13.	atggcaacatttggatcgcaaaagacggcaccggcgtgtagccaccaccaatcaa gaagggtgcgaaccttggtaagaagaacgtgcgtctgctactgttatgaccatttacg atccagttgacctatgaaggaacaaacatgtgcagcctatggaactgacacaggacgc cggctaccagcatatcggcgtatccgtgaagtcagagaacacagagtatgtttccgagca gtacgacctgctccagatgaacgggagcggcataacgaccggcgtaaataccgcggct cccggcggaagccggaacggtagccgaaccccggttcgacaaccgagccattccaa aggatggctcccgccatccctggagcataagagcagctcatgtacagatcttcgataa gtacaacggcgtagctcctatctcacaggtatgtgtagaggttagccagttcgatacga gctgctggaagcaatagaagccggtgaagcccttgcgggaagcgaaggactaccagcg agggtgaacagtaggatacgaacccctcggggaagcggtattacgcccgggataagtag acctgctctctgctgtggaaggaatgcgttcaaaagacggtgccaatcctcggatgcacat ataggattccaaaccggcgacagaggaacaccgcatggggaacctggcatctgtgtgtac caaatgccacaccggcagaacccaatgaagccggcggaagctgtacggctggaagcc gaaagctgacagcgctcaaggcgccgcttcatgaacgcgggtgaaatttcagattttga ggcacttcggatgaacacccggatgtagcagttaccttcggcgccgtacgaagcgag cgcgctttaccggcaactagccaagtcctacatacaacgatgctactgcatcgaggtta tagaccagacacagatcgagaccgaggtatatacgcgaagcgccgaggaacaaccg ctgccttggagaaattttacgatgccaagtatgtggacaccgcgatgaaaccgtcaaaag cggcgctcagctgtcttgcggacgaacgaaccgttcgacaccaagaggaatccgctca atgagcgatctaccgggtgaaaagaagtcgaaggacgcagagatttccgggag gcggtactactcgacccctgatcatctgttaagtcacaggcgagatgtacgtttaa ggcgctgcaaaacaaagcgccctatgtggccctcagtgatgtactgtgtcagcatcg caaaaggtcaaggccatccgccacattggcggtatgtccaaggaaaccacaaaaccggc gtag	gtcaataacccccacttaacctgttgggttag aagtgaggcgtgtgggaacgcggaagcg ttcctgtgagcctgattgattgacagggtct aacgacccctagagtgacctgccaagct gcggatactggcaggtgagaacctagctgt acacaaatgtataggcaccggaggtgctc cacaaagtcctccgactcggtgtgtctttaa acatctctgagggttaggagaagtgaggcac acatcgaaactgtgttaacattggcgatgt ggaccaccgccccaaagggtgaggatccc ggacgcttccgtccgggaagcgtaggc ctttgcaacactgaacgaaggagcatcag cttcagtgcaaca
674-675	Ga0163144_10177891 GENOME_ACESSION: OliLak19BULKMAT1_2_SF_3300015360 GENOME_ID: 33717 CONTIG_ID: 177890 SOURCE: JGI DATE: 2019-01-27.	atgctatacgcaattactccgtgtgtgtcggaggagcgcatgatgactaagtcgtctgt tgtgttcgttatcaacaaaatagtagccgccttgcgtctatgagggcgcgag tccttctcggagatggtctagctaaagtgtatcgcgagttccgttaccattaaactttaa cccagcaatcggagtacaacaggatgtaacggctggcatggactggagcaaaaca cattggcgctgctgtatttcaagcggtagcgtctgtataccaggccgaagttaaacttcgc ggggacgaaattaagggcaagatggaccagctagaatgtatcgccgacccgtcgtg gcccgaagacacgctacggaagccaagatttctaaccgggttaactccacaagact ggtcgcataccgcaagcggtcgtcacaaagtacagcccatgaccgcgaagactatt cgtcgaatccatcctgctgtacaaaatgggtctgtagaaaccggcgcttggatgcgt aaactatctttatcagataatgtacgacgcttaaggcgccgattaccagaatggacgcc agaaagggttctacaacgctaaggcattctgctccaacgcgacgcctacaagtgccagt acggcaagggtatgtctcaaaaattctcagctccaccacattatcttctgttgatggc gggaccgataagccagaaaatctcataactctctgcaaaattcatcagacgcccttcatg cagaaaaattaggtacatttcttctgctgaaaaatcttctgcgcaaaagcctctacaaa gtaaaggggagccaccacaaattagcatgattcgcccaacattctcgcaactgggtgaat ataactgagacattcggctacgagactaaftaaacgagaagccctgtgtcttcgaaaa ctattacaacgacccgtcgaatctgccttaacgaaaaagaaaccgtcgaatttcaa acatcaagttattcaagcgccctgtatcaaaaggcgattaccagcagacatcaggttcca gaagcgaaaaaactatccctacagggcaatttggcctcaggaaattcgacctaatatc	gtcaatccccaccctaaacgggtggggctt ctagggtgacctaggagggttaactgttga ccagaggactaacaaggagacgaagaa agtgttagtagactgtaaaggcggaaaaa catacacaccaggatgccatccagtcg ccggctctgtgttctgtctgtaaaagagag gaaactcgtgtgctgctgaactttaaacc gcttaaacctcagagggaaggtcggattcg tgttgctactcctacggatgctatagcatt actcgtgtgttgcggaggagcgatgatg actaagtcgtcg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gacaccaagggaacagggtttgttaagggaagcgcagttcaggatttttccatatac gctactgatggcgaagtcacacccgattccgtaagcgttacgaaaaactgtccgattat ccgcaaggcccttagtctgatttccagaactatccttgcgaatttcccccgttaac gacggaatatccttggaggttcaagatga	
676-677	Ga0265595_1002166 GENOME_ACESSION: saktaG_23_SF_33000281 91 GENOME_ID: 37741 CONTIG_ID: 2165 SOURCE: JGI DATE: 2019-01-27.	atgcagaagttatcaaaagagatacatcacacctacgaatcttcaagtttgaagct ctgtgatctgacattaaacattcctatgggtaaggaaagtgtggaagattaaaaaccttga taacaactccgaagagaatctacagcattacatagagctggttgaacccaaaagtatca aacattgtctatgtacttaccacaaatgtgtgctttaaagccaactaaaccacaaaagc aaggagggttattacaatgtaaaaaagcaaaagtagtaacaaattccattacaattcaaat gttgataccaacaggagaagtaaaacaaagagattacattaggtgttattctgttatgaga atgtaggaaatcagcagttacagcaaaagaaagattattatctataactttaaattagaa caaatatgtcaaaatatttctgacaaatcaatgtatcgtagaggtagaagaaataagttgt gggtatagagaaaaaatttcttaacaaaggatatacaaaagggtgggttagcaccaagat acttcacaaatataatgtcatttaaaattatagagagaatacataagtttctactataacc aagataatatttgatagcaattttgatatacaaaaaaagcattctattattaaaggta aagaatatcaggaaaggtgaacaaaaagggtttgaaagtgaagatgtatgaagaagta gagatagttatcatgtagaattgtagaagaagaattgaattacaagttcatcatataa taagttagaaaaacagggtgtgtatagctcgtgaatttagtaactttatgaaaaattgcata gtgattatcatagtgtattatcatagtgtgtaaaattgaattagataaaaaagagaaaaggct ttaagcagagacttttatgtcaacataagaaaaagaaattatagaagatttaaaaagcaaa tatgatgtatgaagagacttttggattttgacaaatgtatagattagagtttaaaattag aaaaagatcatataaatgatcattttgtatagcaaatgtgtaaatcaagaaagaaagtttg ttcaaaatataattcaaaagagaagaataacagaaaattacaatacaaaaggaaaagat ataaaccttctataagaaggcaaaagatatagatttcaaccatagatttattaaaaatcaatg gtaaagaatattgttaagtaagggtatacattgcaaaaggtgaatcagtaataataaaaaaat ggtaaaaaagagagtatttctgttaagaagtagagaaagatttatttggaaacattaatt atgtaaaggaggatga	gtcagtaaccacgactgaaggctgtgggc ttgcagaagccctgtgactaggaggcata cttagtatgcagaagttatcaaaagagata catacacacctacgaatgcttctcaagtttga agctctgtgatctgacattaaactctatgg gtaaggaaggtgtggaagatttaaaaccttt gatacaactccgaagagaatctacagcatt cacatagagctgtttgaacccaaaagtatc aaacatt
678-679	Ga0111033_1162652 GENOME_ACESSION: IMG_3300008516_SF_3 300008516 GENOME_ID: 311683 CONTIG_ID: 53090 SOURCE: JGI DATE: 2020-04-07.	atgcaaacagtatctaagttatcagcagagttaaagacgcaccaaggagatgcttctag tcccttgcctgcaagtttctcattaaacagagaggaaactctcagttgtgcaggacaaagt actggctgataacctggcgaagagaacctactcgggaacaggagaaaaactctgtatt agtgtttgtactgaacaaggaggcgcaaggcttattgcccctccctcctggaaagcaag acatttatgaagaagaaggcgagggtgtgacctgcaagccgttctactatcaattacttt atggcagttcaggatataagcagccgattacacttggaaatagatcagggtatagtaacat aggattcagcgggtttcagcagagtcagaaatgattgccgggaagttaaattagaaga gaatatgtccaagaactgacagaagacgaatgtatcgcagagagagacgaacaag ttgtggcacagaagccaagattctcaaacagggtttcattccaagaagaagagcgtgtta gcaccatccatacgcacagattagatttctatattaggtgatagaaaaagtcagagac tattaccaatttccaaaacaataatagaagtggtcttttgaaccacaaaatgcagaacc cggaatatcaggaaatcgaatatcagcaggagagaactgcaaggctatgagatcaaaagaa tatctactggagaggtggggaaggaatgtgcttattgtgtaagacaattgttctttgga aatagagcatatcactcctgagtcgcgaggaggaaccaacgggtctcaaaactgacga tctctgtgaaaaatgcaaccggaagaaaggtacaagacagcagaagaatttggatatac cagagatacaaaaacaggcaaaagaaacactgaagcagttgctgtcatgaacagcgtc aggtggaaattagtagactccttagactgtgactggacttatggatactccagaaagcata acagggcgcaaataggttttagaaaaatctcatgttaatgatgctttgtatagccggagg aagtaatacaaaaacgaatttgggattgcagttgtaatacaactcagacgaacaacccgtg tttgagatgaatagaattgattcaagccgtccataagaacaaatagatatgctttccagc cttattctctgtaaaattacgctgtgtctatttgaagtcagggtgattttaaagagggca cctacatcagattaaaagcagaaaaagacatcaatcgagaaacagaaaaagtaacattat atcaacacatgaacggaatggcaatccatccccaccaatcagagattggagggggtctt cttgcgagaaaagataa	gtcaatcacccccacaaattctgacgaattt ggaaggggcctgtgagggcgacctgcaagg gcaactgggtgattagccaaagagaggtttg cgtatgcaaacagatctcaagttatcagcaga gttaagaacgcaccaaggagatgcttctcta gtcccttgcctcgaagttctgcatataaacag agaggaaactctcagttgagcaggacaaagt actggctgataacctggcgaagagaaacctta ctcgggaacaggagaaaaacactctgtattag
680-681	Ga0373927_0000008 GENOME_ACESSION: GW9ome_3_FD_SF_330 0035695 GENOME_ID: 312474 CONTIG_ID: 7 SOURCE: JGI DATE: 2020-04-07.	atgctgggtgtttgtcgtatcacaagaaggaatgccttfaatgccgacgatccccggcgaa cacggatctggctcaaggcgaacaggcccgctgtcggcgctcgagagccgttcaccat ccaactgcgctgcgagacgaaagcctctgcccacactgcgctgtaggcgtggacgca ggatcacacacggtcgggattgctgcatcgcaaacggcgctgctgtttcagacgga ggtagagttgcagataccattccagagaagatgacacacaggtgcaggtcccgcgct cccgttgtgctcgcaaaacacgctatcgagctgcgagatggccaatcgacgtcgcctg gaggaccaaatgcccccctctccgcttccaggggcaacgtccatgggtgaagcggccc gggtatattatctactgctcctcatcggtcaggtgaacatgatgccgtggctattgctgtg agatgggcaaggtgatcaagccctttcagattcagcagaccgatgctgctgcccga ggctgtatcagcgccttcaacggactgtgtagcagcagaaggtttggacgccaagaa gggtcaaaaggcttcaagcgtacgctgggtccaggccagaggggtgatgggttccatg ggagaagggcgagagaaggaatctttgttctcaacgatgttgcagttgggaagagagc gttagaggtgacgtcagcgcagcttactcaggggcaggatatacgtacggtatctcgccct ggagaatacagagaaggagagcggcgttctcccccgattga	atgctgggtgtttgtcgtatcacaagaaggaat gccttfaatccgacgatccccggcgaaac acggatctggctcaaggcgaacaggccccc gtgtcggcgctcgagagccgttcaccatcc aactgcgctgcgagacgaaagcctctgccc aacctgcgctgtaggcgtggacgaggaat cacacacggtcgggattgctgcatcgcaa acggcgcggtcgtgttcagacggaggttac agttgcagataccattccagagaagatgac acacaggtgcaggtcccgccgctccggttg gtcctcgcaaaacagctatcgagctcgag atggggcaatcgacgtcgcctggaggacc aatgccgccccctccgcttccaggcgcaa cgtccatgtgtcgaagcggcccgttatattat

SEQ ID NO:	Name	IscB sequence	ωRNAs
			ctcactgctcctcatcgggtcaggtgaacatgatgccgtggctattgctgtgagatgggcaagggtgatcaagcccccttcgatcattcagcagaccgatgctcgcgcgggtctgtatcagcgttcaacggactgttagcgagcagaaagtggacccaagaaggtaaaaggcttcaagcgctacgagctggtccaggccagagggtgatgggtccatgggagaaaggcgagaaaaggatcttttctcaacgatgttcagtgaggagagagcgttagaggtgacgtcacgcacgttactcagggcagatatcgctacggtatctcgcctggagaaatcagagaaaggagcgccttccccgattga
682-683	Ga0070730_10006957 GENOME_ACESSION: CoaSoiStandDraft_4_\$F_3300005537 GENOME_ID: 24560 CONTIG_ID: 6956 SOURCE: JGI DATE: 2019-01-27.	atggggcgtgaaggccatcctttctgtgaggtcgtgatggtatcgttctggatcagagcaaacagccgctgatccctgtagcccgaagcgggccccggctgtgttggcgcggggcgcgcgcgcgtccaccggctgtgtcccttcacatccgcctgtgaccgccgcgtggccgcgagcaacctgcagccgggtgtgtgtggggtacaccccggcagcaagacgaccggcctggcctgacgcgggaggagccgacgcccgcgggtccgctggcgcacgtcttgaccgtggcgaaactagagcaccggcgggcgtgtgtggcgagcgttgcgcaagcgcgcgccggcccgccgcggcgccggggccaacctgcgtatcgcgccacgggtccacaatcgcgcgttagcgcgggctgtgtcccatcgtgcaaaagtcgggtggacagtgtggcgattggcgcgctgttaccgcccggctggccccgtccgcgggtggcggtcgaaggcgtgcgttcgacaccaactgctggagcaccggatattgtgtgtcgtgagttaccacggggcgagctggcggtgtgagttgtggagttatctgctgctgaagtgggggcaacagtcgcctactgcggcgggagcagctcccttgcagaaggagcacatcacccccaaagcgcggggcgagcaaccgagcgagcaacctctgccggcggtgcggccgtgcaatcgcgaaggccgcgtccgcagcggagtagggccaccaagggtgcagcgcaggccaag	gtcaagtacccccaccgttaaacggggggcttgtgaggagcaagccccggactgaccagccttagcagtgagccgaaggaagcggtgacactgctactgttgggttagagcaaaagacacaccgggggatgcttccagctccccgctctgacgcggcaggtgagataacgtccggggggcgcggaacggcctggcgagatggctgacctacaactggcgaggagacccgttcgagccgtcgagcgcgtcacatggccgtacaggccatcctttctgtgaggtcgtgatg
684-685	Ga0118720_1022568 GENOME_ACESSION: IMG_3300009374_\$F_3300009374 GENOME_ID: 305905 CONTIG_ID: 22567 SOURCE: JGI DATE: 2019-11-15.	atgttagtctttgttaaaacagaacggtaaacactaatgctttagtcccactaaagcagcaaaacttttaaaagcgggacagggcaaaagtggaccagcgacccccctttgtaattaaagctgaattcgggtctagtgtgttactacagaaaaccacggctggaatggacaccggtagtaaaagtattggaatgctgtccacatcccatcaaaaaactctataccaatctgaaactatacttagaggagatgagattaaaagtaaaatggaaacacgcccgaatgtatagcgtagccgtagagagcagaaaaaccagataccgaaaaccacgatttctaaatagaaggcgcaagtactagaagaaggcgactagctccaagcctgtacacaaagttaaagcacatttaaaagaaacgaaaattcattgaatcaattctactatctctcattggaaagttgaactcagccagtttgatattcatgctatctcaaatcctgatgtttctaaaaagatattgtgtgacttaccagaatgtccacagaaaaacttttacaataccaactttacattttgagtcgtgatgaacacactgtcaatcatgtaaatcaaaaaaagaatttaaaacttcatgtacatcatattgaattcgttccaatggaggacagatgcccccaactaaccctataacactatgtaaaaattgtcatgacaagtagacacacacaataacgctcaacaagaatctctaaaaatgtcaaaaaaatcaaaaaaactaataacatgctactgaaaccaacattatagcctcacaatctcaaaactctgactggagtttgaagaaactttgggtttgagactaaatttaaaagagagaattgagactacctaagagcattactttgatgcgtgtagctatctgcttggctgatgatgaatgattaaattgaaccgcaaaacttcaattaaagaattattgctaaaggagattatcaacaacaaaaaggaatcaggctgaaaaagaaatgcgtacagaaaaatactaggtttcaaaaagttgacaaaagttgagtgatgggacataaaagcttcaataaaggcagaatgtcatccggatagccattttaaattgacataaaaggggataagataaatttaaaaccatccctaagctaaaaaatggactgtacgcatacaagcgaggaaatcatcgattatagatcaaaaaatcatagaaaatcctcattctaccacatcatcctgtcaccaa	gtgaagaaccagcgactaaagatcgctggcttctacattaatcatggaagccctacttccagaaggcctaacagaaggagagaaaatgttagcagctgttagaacagagaacacatacagacccccgtgtgccactcagccttggctctctgattatgcaataacagctgtgagggtaacagcagtgtagattttaaaaacgtgtcctcctcagagagatgtcggattctgaatgtatcttactactatcagatcgcattactctgtttctgtcagaggagactttatgta
686-687	Ga0395674_000262 GENOME_ACESSION: 711E1SmetaG_FD_\$F_3300037239 GENOME_ID: 312082 CONTIG_ID: 261 SOURCE: JGI DATE: 2020-04-07.	atgccctgcactccagcaaggcgcaaggtgctactgtcgtggcggaaggccgcggctctaggcgctatccctcagcattatcctgaaggccgcagacatggggatttacagccgatagagctcaaaagccgaccttgcagcaaggctaccggatttactgtggcttgcgaagattgcaaccagagcaagaacaacttaccgcccagaaatcgggcatccgaacattcagaaagcaggcattacgacctctcaaaagacgcttctgtcaatgccactcactatgtatcggcgctgactcaagctttgggattgcccgtatcgttggaccgggtggcaggactaagttcagccccactcaacaggactatcgaaagccactggattgacgcggcctgtgtcgtgagtcgggaatcaggtaacgttgcgtcggatgcaggatcgcaggttaaggctacaggcacggatcacggcagatgtgtcggatggacaggtttgattccccggacttctactaaagacagctcgtgtagtcaaaagggttcaggactggagatattgtcagagcagtttcccttacgggaagagacaggactttagacggcaaggtagccgtcagggcacccgggagtttcaatgtatcagctgtaccggtatagcttgggattagctacaaatctgttctgtgttcatagggcagatggatattcatatcaaaatacagcgctcctcctcggcaatgaattaccgagtatccgcgccaggagttgcatga	gtcaaccactcgggattgaaatcccagcagatgttaagatccgaagggtacgcgcatgctctagttgaccaggctcagtgatttaacactacgttaccggcgaaatcataggcaccttgggatgctggccagctccagactctgcggcagacagttaaacaggtgtgagggttaagcccgtgctgcctcatataaaaccgtcggaataacttggcaggtaaacattccccgcgtaagcgagattaacggagtagtcaattcaacgt

174

SEQ ID NO:	Name	IscB sequence	ωRNAs
		tgaatatactatcgcacgggtgaaacttatcaaggatatataacagcattaaattgaaaa ggaatgcgataagtttctgttcttatacaaggattttacctggaaaaa	
696-697	Ga0209343_10010378 GENOME_ACESSION: RifCSPlow2SPAdes_\$F_3300025311 GENOME_ID: 34854 CONTIG_ID: 10377 SOURCE: JGI DATE: 2019-01-27.	atgcaaggacaaaaacttggatagatttaggcggtaagcatgctggcttctgttgaag aacaccgataaacgaggtggcacattactgcactattgaactcagagaagacattgaagga taagatggatgagaggaggtctctcggaggcgaggagaacaggctctggcatagg gaagcgaggtttgacaataggcaattaaagggtgaaatgcaaatatattgataaagataca ggcgaaatctcggagctaatactccaagaaatccaatgtaaaacatcttctacttgaga ataactctgtaactttaaaatagctgatgaatctaaaggaggaatcagaagaaggaggt ggacagagacacaaaaggtaattacagacaatccttgagaaatttcaataaataacc ttcctgaaaaaacagattaaagacatcattctgaaaagggggaaggagggctgtcttt gcagagagcatacccccttcattatgaacaggttgcacagaggctgagagttctggct gtcaaatcaataagggttaaacaggaccagatactctccgccttaaaagaatagcaaa ggattttaagatagatgaggttggattgaaaggcgcaacttgattgcaaaagctccaga gacctgatgagatagaagcaccctgaagattacatgaagggtcctaactcggcgacaga aacagggttgaggcattgaagcaggaatatggcaaccgatgctgttctgcggaagaag ggfaggatgaagtaagctgaagataggcatctctatccgaaggctaaagatgagat aaacagggtgggaaaaacctataactatattgtaaaatgtatgcgaagcagggttaaaag gacaccagaggagcgaggatggaaattgttaattgaaaggagaagggttttaactcctgc agcagggaagggttaataccataaaaagagaactcaagccgaagccataaatgaatca aagggttaataatataatgaccatactgataattggcataaggagggtcaaaagagaatcc agaatatttttgaaagcataacctataagaaacatacggctatacactatgattttgaa ataaatgggagcttgaaaaagaacattataatgatgctgtatgcatagcctctgacaaa agatttgaataataaaacctgtatttaagatgcagtcctcagacaattaaatcactatcaa ggcggggaaactctttgatacaaatccccctcaggtttgatgagaaagggtttaccagaac ataacccttataggcagaaaggcagggatgcgttcatacaaaacataaaagggtcagag gaatatcagggaactatggctcaatttatatgatgagattgaactataacctcagaatgga agaaaaagggttctctgcgaattaaagagataaacttggtatgtaaaaggagataagaataa gtcttttaagcctgagggaactgatgaatgcaaatctcccttcaggactgtaactattgaca aagggtgttaggagaatctcaacccgcttaatacaatacaatgtatccgtgcctcagct gaagtaatacgcataatggctattcaataatgacggtagaatgaaggcatttgcagat aaaaaaccttaagatatttaagatgccggactccctcatgatttcaaaaaagatattcat tgtaaaaagggggatattgttatggaaaaaagtgaagatggaattgccgtaacagg cagggtgaccaaattgttgacaaaaatggggtaattgatataaaggacatgaataataa atacactcagggaacacccgtgtatattgaaaaagatgatctcctgaaagggtgctga ttttgagagaaaaatcttctgtctcttga	ctcgcataagaaggatactgtctgaaggga gggttgagcctgaacactctccaccttctg atcaattgcgattaatagagggtttatggct tttaattgataattcctgaagggaattgcataa atcttaacctgtgacggagactctatatttct gaaaataggattacccccgacccataat cagttattattcttttaactcatctgacatagt agtccccatgaaatccaccataaacatctgc ccg
698-699	Ga0209737_10031347 GENOME_ACESSION: TergutCu12SPAdes_\$F_3300027904 GENOME_ID: 36027 CONTIG_ID: 31344 SOURCE: JGI DATE: 2019-01-27.	atggatatgttaataagcaaggagggcagccgttgatgccgacggaacgtcacggcaa gggtgaaacacctgttggccaggggcaggcgaaagtgtgacggtctgcccgttcaca atccagttgcagtagcagactggaagccacacgcaggaggttacccttggcatagacag cgggtatagctataggtataagcgcaatgagcgacaataaggaaattgctgtcgtcga gggttagttgttgacggaatgaaagagcggtatataatgacaggcgaatgtacaggagaa accgcaggcagaggctaaagacagggcaccgaaggttcgacaaaccgtgcaggggcg caagggtgttccactagcgtcaggcacaatccgacagccacattaggtgcatat atgctgaaaggactgctccctataactgacccattgcagaggtggcgaacttcgacata cagagataaataaccagacatacagggaaagataaccaacagggtgagcagatg gggtggcagaacgttagggaatacatattccacaggacggacacaaatgccagaagc cagagtgcaaaagcaaggagagaaagtgttgcgagcaccacataatccccgggc gcaggggcgacagatgccccagacaacctctgtgacctttgcaaccaatgccacact ccgaaaaccacaagggtttctcaaaagactggaacccaaagggtgaacgggttcaaggcg gagacctttatgaccaccgtcaggcatactatgtcgcagcagctcaaggatattcccc aatgtcggatcacgtatgggtatatcacgaaacacaggaggtatggcaaaaactggc taagacacacgcgaacgatgggtatataatagcgggaggtaaaggcgaaccacgtcgc gaaacgctgcaattaggcagataaggcggaacaacagaagcctgaacgggtctacga cgtaaatcattgatacaagaacgggaaagcccgaatatgcgagtggttgaacaacg gcaggactacccggaacatgaacctgaacggggagaacctgaaaaaataccgtgggg aaaaagtcggcaagggtcgaaggagcattcgagaaccactatttctatcagcccaac gaccttgtaagtatgaaggcaagatttatgaggtcaagggtgtcgtcaacaggggcaac tatgtcgtctcaaaaggcattaaaaaacgcaaaagggtgaacagttgaaaccatacaga ttcagaaagggtttgtgattgaaaacagaaggcgcaagggaatcaaggggcgcaa ttcaccaccacgcctaa	gtcaaccaccaccgectaacgectaacgg cgtttgaggcggggcgtgtatgggactac gcaacctacagccagttgaatagcccaag ccttgaacaaggctactgttaccagaatg acataggcaccacggatgcttcccaagtc cgtggcactgcggacgggtgattaaacatcc ccgagggttaggggaaggttgcggccaaa aaccttgataacatcggcatgggaacca accggggcaggtcaaacgcaacggcatt tgaccccaacgggttcaaaaggagaaga gcatg
700-701	Ga0075011_10000624 GENOME_ACESSION: TARMetStandDraft_6_\$F_3300006192 GENOME_ID: 35911 CONTIG_ID: 623	atgtttcactccttgaacaaacaaacaccactagaccgggttttccgtactcggctat cctctcatgcgtgtgacgttgtagagctcgtgaagctgttagagcaaggtagagctaaaa agtctgtgaaaaagggtcttctatatacagatgcttcttaccgaagaacagactcaa cctgtagtgtaggtatagacctggatctaagagagaagcgtttactgttaagtgaaca cgctactctgttaaatcacaatcacatcctgtgacggcaatagataaagaagctctag aatttagaaagggtatgcgttagggctgtagatctagaactactcctgtgacaccttagat	ataactatccagctcacaaggctgggtctg ctaaatgctgccaacagctcatagcgaata gctagtcgaaaaggctagtaattgttattag ccattaccaaaactgaacgacctgtggatgg aacacacagtagctacggatgctccctag tctgtagctctagggtgtcagtggtgaagg

SEQ ID NO:	Name	IscB sequence	ωRNAs
	SOURCE: JGI DATE: 2019-01-27.	tcaacaaccgctctagaaaaggctgggtaccaccatctactctagcggatggcagttaa agctcaacatactaatcatctgtcttctctatccataaaaactgtggtatagaggacgt aaaggctatgaccagaaaaggtaagaacaatggaacagcaactttcacctgtgcaag caggtaagaactggctatagcagccgtatatacaagaagggtctgaactgattaaaggtg atggatacaaaaccttggctacgtgaagctgccaatctaggtaagactaaaaataagct tgctgaacatttagtcccactgctagactcatgggtgctgtaactactatagcgg ggcatgtacaaccagataaccactctactactcttaagcaggtaaaagtatactaga caactacatttgcctgtatagatagtaacggaaaaagacagcgttacgtgggaagtatgt ccttagggataaagaagggtactgtatgttcatcctaataatggacgctgtctataagc ggaaactataacggacgaataagttactatcaccatactcacttgagagagttactataca cgtaaaagtctcagatctgaaatagtggcatactaccgtgggtataaaagttaactata atcttctatatactgctaaacgtaacgatagagctactagaaagaacatagctagaactag gtttctcaacaatctatgcgtagtgtgtctcaaaaagaatttaatttacttctccctaa	gtccttactgtctcctctgaacttaccgtta aaaggacacac
702-703	Ga0116227_10001884 GENOME_ACESSION: SITSphgellanicum_5_\$F _3300009709 GENOME_ID: 34953 CONTIG_ID: 1883 SOURCE: JGI DATE: 2019-01-27.	atgcaaacggtttggtagtagatacaaatagacggccttggatgcgtgtctctccgattag cgcaagggaagcttttacaagaaggcagtcgagtttaccagctgcagccttttacaatc gtcttgaaaaaaagagaaggagggtcggatagggtgtccaatttgacaatcgggtgcaga gaatgccatagcaaaaagcttaaccgggtcccttccggggattttacaaggcaaaccttctt acagctctcaaatcgggggaaggctaaagccctcttcccgaatgcgtctccgtcaatag gacggggaaagagattgtatctgctgccacatttgatctcccatatcttgcctcccg gaggaaaggcaaaagttaacaggactctcaaggatataccgaaggtgaaagaagagtc cggcgttttgcaacggcagatttgtaaaagcgggtgtcctacaaggtaaaaaaaggga acatactccggaagagtcgcatccgatcatcgggaagttttgcatcgatacccttaagg gaaaagtagcccgatcagctatagatttgaacatctgcaatagcagacgggtatcaa tattcacacaaccacaacaacaaaggagcgacgcttctaccgagactcaagtcgca ggttccgcgtcgtcagttctatga	gtcaactcccccttactctaagagagaggc ttgaagaaaactaagcctaagtgaccgga ctaagttcttgaagtaacaagagctacgtta gagaagatagataggcaccttggaatcgcg gctccgggttccaagctctgctgggttacggtc aaaaactctcatccgttgcggacgagaac atccccgggaaaagggaattgcaggta actgtaactagtaaaaaagagaacgat gcaaacg
704-705	contig_854547 GENOME_ACESSION: mgm4547164.3 GENOME_ID: 237431 CONTIG_ID: 139596 SOURCE: MG-RAST DATE: 2019-01-27.	atgccttgcgaaccgcagaagcacgaaagccgctcacggcaggcaaggcaaaagtg gtaaaccacgaaccgtttaccatcccaactgctgtttgttctagcgggtacagcgactctgt tacgctgggcacgacgctggctctgtccatagcgggtccctggcatctacgaagaagca agaactctatgcttctccaaaggagacgctgtgtatattcaggggcttcttctaaagca gcttctgtctaaaaaacttgatgggacgaagcttgagattatgccaagcaagattagcttc ggctggcacaataagctgtatttagtaaaaggagcgcggtctctctcgcgggttaaa accgcaagtagcccgccgcaaatgttaa	gtcaactaccgcacctaaggcgtaggctt gcgaaagcccaaacctttagatgcaggcat agcctaagttgactagcctaggctcgtctc atttgcctgggactcctgtttttgtatcacac ctgtggcgctatatacctaccggtgctctgt gctggctctgtaaaagccctgtgagtagg ggcggtcaaccagatgtaggccgattacgg caagcaattacaacattggcgaaggtagat ttactctgattttaggcttcggccttggaaag gaggtagacgtatctgttc
706-707	Pelotomaculum sp. PtaB.Bin104 A4E53_contig000144, whole genome shotgun sequence GENOME_ACESSION: GCA_002067205.1_AS M206720v1_genomic GENOME_ID: 242484 CONTIG_ID: 143 SOURCE: NCBI_Prokaryotes DATE: 2019-05-30.	atggtatagcttttatcaaaagcggcgacccctgatgccaccgaaaggcacggaaa ggtgagcggatgctaaaggacggcaggggcgaaggtagtcgaaaggcgggttcac catcaattaacgtatgaaacaacagctacaccaaccataaccctggcatagacgc cgggtatcaagaagtcggcctgtcagcggtaactccgaaaaggagctttgtccgccga atgcaccctgtttaaaggcgaggtggaacgcaacaagaaggctcatgtaccgagg cagagaagaaccggctgctgtaccgtgcccagtttgacaaccggaggaaccag aaggctgctgctgccgagcattcagcacaagctggacagtcacctgcgcctgatagc ctgggtaaaaagatacttctatcaccgatgtgttatcaaggtagccagttttgacatcc aggcgattaaaaatcccggtattcaagcctagatgacgagaagagagcagtaggg cttctggaaccttcgcaatacttctgaccgggacagtcaccagtgctcaaaaccggga gtgccgtacaacaaaccccggtgttcaaacctaccacataggttttgaaaag cgatattacagacaggcccggaacatgatcaccctctgcgacaatgccaccggcca gagaaccacaagaaggcaagttctgtggggatgggagccgaaggtaaagtcgttca ggcggaacattataacaacagtcgggtggagaatgtcaatagcttggctccgg catactacggccacataaccaccccgaggatagagctaaagctggagaaaagct actccaacgacgctctgtaataggcggggtagcagccaaccgggcaaaaatgct cagcatcgagcagatcggcgcaacaaccgggtcttgcagaagttctacgatccaggt acatcgacattcgaacggaatgcagttgctggccaagaactaaactgcggccgacgt accgtaacagggaataataacgggccaatagcgggtataccggggcgagaaggtgtt taaggcagggtgagcatcagaagcagaggtactgtaccagcccaagatattgtg aatatggcggtggtggtacaggttaaaaggctgcagaagctggcgctacatcaaa ttggcagaattgagcaagccagtgaggacggagttgtgcaaacctgtgcggtacagcaa gggttgtgtgatgtaa	gtcaacaaccgccacttaacggctgcggc ttagaagtggggcttgggtagcgtgagg cgatagcttacgtccagcccggttatt agcttagtggttggtatccggaaggaaacga accgctacgttagcgttggaattgggaaccca aataggcactccgggtactccacagctc cggacactgcggccggtgttaaacatcgc tgagggtaggcggaagtgtcggcgcgaca aacccgtctgaacattggcgaagtggaacc acagccgaagggtggttatttccgaaag gagatcgcatg
708-709	Ga0256404_1034523 GENOME_ACESSION: RumenRJG_01DNA_2_\$ F_3300026549 GENOME_ID: 34936 CONTIG_ID: 34522	atgctgacatagtactggctcagacggctcaccactgatgccacataacataaaa aaggtgagacggatgctaaaagacggcagggtgttatagcaggacacataccgggt ttacataaggcttaattataatctccgggaacaggacgccctcactacaggagactga gttttgcgaggacacggcgaccatcacatcggcgtgtctgttaagctgttaagcatga atacttccacggtagttcgacctgctgacagacgaaaaacagcgtcatgacgactccg ggcgtaccgcaggacaaggcgtaacgtgaagcgctacagaagccagcttttgataacc gccgcaggaaaagacggcgtgttctccatcgggttgaacaagggtacagcgtcacgca	gtcaataaccgcgggttaaacgcaggctt ggtctgaaacacaggactgagcctgtgtgat aagcaatagcagccccgtcaagctgtttttac gatactgcaggggcataagctacgttacc ggcaaatataccataggcactttcggatgct ccacaagctccgtaacactcggcgctgcatt aaacatctctgagggttaaggagaagtgtga

SEQ ID NO:	Name	IscB sequence	ωRNAs
	SOURCE: JGI DATE: 2019-01-27.	gatattcttgcacatgttcgcaaggtctgcctataaagctgctactcttgaacggcgag gtttgacactcagctattagaaagccagcagacagaaataactccacaaggcaag attatcagcatggccgaggtttaagcttaacctaagagaaagcggctttacgagaga tggctatacctgtcagatatgtggcaagctgtatctgacggagcaataactccgtgtgcacc atgcactgtactgaaaggcgaccatactgacagggtatctggctgataacgggtatcgga taagtccacacgcctgagaacctagaaaggcgaggttatggggcataaacca gaagtaaaagccatggcagggcgagccttcataatgtctgacgggtgcacataacaga gatcttcaggtcaggggttcccgatgtggagaccataccactaatgtgtgcgggtacca agcttcaagaagatgcagcgtatggcgaagacacgcctaataacgcttactgtatggg tgaagtcagaccacggcacaaagcaaggagatgcattacaaaagcaccggcgtaac aaccgtgtgtgagcaagttctacgatgctaataatattgatgcaggaccggcaaaaag gctccggcaaggaattagcagtaactgtaccaataggtcgggtcccaaggataatcct aataacctgagaaagttccacggcatgaagctgataaaaggcgatgtatctgtccgaaa aagcattatgacatacagctgggtatgtatgtctctgcaaaaggcgtgaagcgtctgtg cacgcaatacacagagggaacaacgtagagtttgaagctgacggcataagcccgaagt cagcttcggcgacaaggtaaagatcataggaatgataggaggtggcagcccttctg aataa	cgtctcgttcttagaacgtaaaacaccggga taacatttgcgatgtgacagctcttttaaa ggacagaatcctgtgatggctttatcacag gaagccgtaagcgattagtaacttattagcat agcagaaaggagcatcatgttatgtcgaca
710-711	viral metagenome genome assembly, contig: NODE_117_length_3294_cov_0.884433, whole genome shotgun sequence GENOME_ACESSION: OOMW01.1 GENOME_ID: 8612 CONTIG_ID: 116 SOURCE: WGS DATE: 2019-01-27.	atggatfatgtaatttcaaaaagtggaaccgcttatgccgtgtgagaatgtaattgcacg attacttcttaaaaatgtaaaagtaagaggaatgtccattcacaafttaac atatgattcaacagaatattgtcaagattgacattaggacaagatactggatcaaaacat ataggtagacatgtgttggaaacaacaagattctatatcaatcacaagtagaactcagag acgacattaaatcaagatggatggctgttaggcaagctcgtagattcagacgaagtagaa aaactcgctatcgcaaaagcagatttttaaacgtgaagaattctactaaactgatagggtgc ctccatcaattaaaagcaaggttagttcacattttaaagaattagaattttgcacaagatttta ccaatttcaagagaagttttagaagtagcacaggttcgatactcaattactccaaaatccaata ttacgaagtgaagaggttagacattggggatatacaagaggtaggctttatggctttgaaa atgctaagcgtatgttttaagagagataactacacatgtcaatgttgcaaaactaaaaaa ggtacattgcacatcatatattgtatatctgttcaaaaggcggaagtaacacacagataat ttaattaccctgtgtgaaaattgtctatctgtcttgatacacatttttgaanaatactgtgtgc ttgtttaaatgagtttccctga	gtaaatatgtgtcggataaagactgaccgc ctittaggtgtggcacagcacttaaatatttac aagacttaggtttgagaatcttacctacgtta gattggagaacacatacacacccctcggtg atgctcaagaccgtcgtctgtgactactaa taagttaggtgtgtcttaaaaaatagccctgt gtttaggtttcaaaaccagtttaacattgtc gatgagaagaccaatttcaattttgttaacag aaattagataggcactaccaagatagaata tcttgagtttatacttaacaaggagttaaaaa tg
712-713	contig_11853115 GENOME_ACESSION: mgm4547164.3 GENOME_ID: 237431 CONTIG_ID: 486630 SOURCE: MG-RAST DATE: 2019-01-27.	ttggcgaaggtattcaacggtagatgtgtaccgtcttctcgagtaaaacacactatttcaga gaaaaggagtgctgtatgcaggcataatgtatttgaattataaagatggacagccgctca tgcctacacagcggtagcgttaaaatccgcagcgtctgtcgcagacggacaagcaaaagt tgttaaacgctgtccatttaccatccactgtctgtatgaacaacagatgttttcaaccgggt agatttaggtattgacacagggtataagcatattgtgtttctgcctgtactgaaaagaaag aattgtatgtgtccgatatgcaggtgcgtacggacatttccaaagccttgaacagcggga gatttcttgcgcgcgaagagaaacccgcaaaacacggatcgcgaagccgagattca caaccgtgtccgcagtaacataaaaggggtgtgcccacatcagtagaagcaactatcgc gtttgcatacagagttattcatgatgtctgaatgtctccgatcatcaatcacgcttga aaccggcgcatgttacacgcagaaaaatgcagaacgtgaaatcagcggcgtcgatgtac agcagggaaactgatgggtataccatccgcggctacgttgcggagaaatcaatcac cgctgtgcctattgtgcgaagacggaaatagtgagacgaagttcgaagttgagcatgtc gtaccaaaagcacgtgtgtgttctcaagaatcagcaatctgtgtggcgctgccacggc tgcacgaagacaaggcacagctaccgcccgtgaatatggacatccggaagtgcagg cacttgcgtcaaaaggcgatcaatgcgttcagcagctgcaatgtctgtgatgaagtgtg atcttacaaccgtgttaaaaactgaatacagatgatgcagtcgtatgacatattgtgcgga tactgcgacaagcgggtataatgtcaatctgcagaagatcaccatatacgcgcccgtgc caatttgcgcatccaaatgcaatttcttcgcaggaagtataattacatcggaataatcagg gtcacaaccgtcagctgcataggctgaagatcaacaaggcggtactcgcaggaacaat caggtgccatagcaagtatggggcatccggttttgacaaaagttatatatcaaaacagg aatgctttgtgtttggcagaaggaattccggcagcatgatatcagattgctgaatgtgca aaagttaatgccggagatcgtataaaaattgtctgtcatttgcgaagcgggaaatattatt ttgaaagggaaggtgattaccgacattccctgtcctatggatgtctcaaatggtttcctgcg gtga	gtcagtaatccacggatataccgactgtat gtctgaagtaaatccgggtgctgtagga ctataagtcagctgactagcctaaggtgtg aaacaagcactagctgttggctatgacac cctgggacgtcgtacctaattcctgcactgt cggtatgcattaaaagccctgttgggtagg gcggtgtgcataacatacaagccattgcaa cattggcgaaggtattcaacggtatagtggt accgttctatgcagtaaaacacactattcaga gaaaggaggtgctgctatgcaggcatat
714-715	Ga0190303_1007606 GENOME_ACESSION: 48700701_MG_2_SF_33 00021512 GENOME_ID: 20966 CONTIG_ID: 7605 SOURCE: JGI DATE: 2019-01-27.	atggattcacaaggttaatttgacatccaacacgaaattgtagaatggttcgaagattatt gaagaaagggtaaagcaaaagtaatagctgtgggtataaagaggtcagccactattaat acaattgcttgacaaagtatttgataaacttaaaacaactgatgcagaatttcgattgtatt gatccaggatataaacatcgggtatatttttaaaatataaaaaatcatattgatttact attatctggcgaagtagaaacagaaacatctgaagttactgagaatttaagtaacggaag atgtatcgaaatctcagacgtcaatatagacgaagaagtgttaaacgtgaagtttgaagag ctaagtttagacatcctagatggaagaacagagctgataaacatgctttcaacctactataga catttgattaatagtcacattatataataaattgctgttcaaaagagtacaaaagatcaat gtgaagtcatttagaatatataaatttgatgtcaaaagattattaatccaaatatttatgga tggcagatcaacatgtctcctaataatttgaanaattgtaaaatcatatattcgtgatagag ataattatactgtcagatagtagaagcatgttggaatatatacaaaatgaagtacatata	ccaactgttgattataatatttggagatttga aaaaatgaaaaattttttgatgatgattacct aatttcttaaaattatggctacacaggcctg attaaattgtcaccctggaatgccactccag tccaggctctggaggatcagtaattatatt acttagtcttattgtgactagtagttaacga agatcaataattacatttaacagccgagggga gaccaacttaatttaagctgagcttttaata tagcttttttttaagtttaatttgaataaaaa attggagacaaatttaaacatgatt

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ttataccaagatctaagggtggatcagacagaccggataatttgattctttatgtcaaaattg t	
716-717	EMG_10019972 GENOME_ACESSION: IMG_3300001598_SF_3 300001598 GENOME_ID: 27525 CONTIG_ID: 19969 SOURCE: JGI DATE: 2019-01-27.	atgccgacactacgtaacggctgtgtccgttacctctgaagaaggaaaagctaaaatct ttaagtaicaccccttactatccagcttacttactgatactcctgataagaccagcctattga aataggtatggacgcaggttatcagtatatcggagatcggctcaagactaagcagaaga actgttctcgcagaatfcgagcttctgctgatgagaacagcgacatgacgattgcaga aagtaaccgtagaacaagacgcaaccgtcttagataccgcaaacacgttgaacaacag agtccattcaaaagcctgaaggctggcttgcgccgagcatcaataacaagctgagcggc atatagatatcattagacgaatcattagtgtcgtcctgtcaaatctataaccatagaagtcg gagaattcgaccggctcgtcctcaaaagcctgtcgaaggcgaagaacccgctcagggg aaaagaatatcagcagggcccactatcttgcagacagcttgagaaaagcagttattca gcgtgataattatcctgtatataatgcagaaaagctgcattaaaggacaagaacgtaatac tgaaaacatcatcgcactgtcttgaaaggcagggcatggcaacgcctgaatgagatg ataacagtatgcagcaaatgtcatcgtcagctaaccatcagcctacaggaaaactgtac gggttagaacctagattacctaggcttgagggtgcaacatatagaatattgtccgctggt tattatcaacacattgaaagcagcattacctgataccaacatagatttctgttatgtgcgtta acttcacgcaagcgtaaagatttagtatagaaaaactcatgccaatgacgcatattgtat aggcagtttccagcctgaacagcgacgtgttctacaacatactataagaagaacgtagg aatagccgtattcttgaagaattctatgatgcgaatactatgatactgcgcaggaagtac aggaaagggtcagaactaggctgaacagaaccaaccgcagagaactgcgtaatctga aataaaccctgagaagatttcacggctcgtcaaaaatctaaaggcagggcgtgtaaccgtg aactagatacccaattcagccgttgaccaggtttgttaatagcaaaaatacattgtcaaa ggagcgcagaatttaggcataatattgtcttaattgga	gtcaactaccactgcctacgcgagcaagc tcgcttagagcgggggttgaagaagcct agttgattagcctaagctttaaataagcctac gttataccagaatacattgtaggcaccagcg gatgttcacagtcgctgtctcgtcggtga tatgttaaacatctctgaagggttaggagaag tgcataatctttaaactgatataacattggc gatgtgaaccacaggggtttatccctgcattat cgtccctacgggagaaggagaccgttat gactacg
718-719	NODE_9_length_105734 _cov_102.449 GENOME_ACESSION: mgm4784118.3 GENOME_ID: 238329 CONTIG_ID: 8 SOURCE: MG-RAST DATE: 2019-01-27.	atggacacgacaagccctaaaacattggcgaggaaaccattaccccgacaggagga cattcatgtctaattttttgtgtcgtactaaccttcaaccgcttaaccaataccgtcg ggacaagcaagaagactgttagcacaacagaagcagctattatcggttaccattt acgctcgtactcaaaaggcgagtaagagttcccaaacgcacccgcatcaattaaaaatc gatccgggttccaaagttagcggattagcaattgttcgagatgcgaagtaattgggtg ccgaattaacccatcgccgggcaacaaattaaaaatgacttagagtcacgtagcaaccttc gtcgcaatcgtcgcaaccgcaaaactgttaccgcaaacggcgttttcttaaccgcaatcg aaagcctggatgtgtccaccaagcttagaattccagagtgaaaaacatcgtagtgggt cagccgcatccgcccgatgttcccaattacaggcatatcccaagagtagttaaattgata cccgaagccatccaaaaccctgaaatatccgggtgttagtaccacaaggagaacttgcg gggttagaagtcagggaacttctgtcgcgaagtgggggagaaaatgccttattgtggc gtgaaaaacgtaccttttagagtagaacacatctacccaatacaaaagggggagcgat cgcgtttccaatcttaccctgtgtcgcagtgcaaccaagccaaaggcgatcgatatac ttcgggagttctgtcaggtaaaaccgacattctgtccgtattttgtcagcagcgaatacaa cccttgaagaatgcggtgtgtgtaattcaaccggcccttcaattgttcaacaactcaagc aaacaggattgccaatagaagtttcaacaggaggtagaaccaagtacaaccgcactcgtt taggactactcaaaactcattggctggatgcagcttctgttgggaatcaagaagtctgca agttgtgacgaagcagccgttgttaattgcagctaaaggatgggggagccgtcaaatgtg tacgaccaataaatacggatttctgtcaagcacaacacgcgatgcaaaacgtttttgggt ttgggactggcgacatggtaagtgcatttcccatctggaaagtttgacggcactcagc cgcgaggttgacgggttaggcaaaagcgggggttttccatgacaacgccacttggcaaaat agtccgggtgcgtcataaatactgcaagttatccaccgcaacgatgggtatatgtatcggt gtccgctgtgtcattaa	gctattaccaaacagcgaacaaaagtggga caactactcatatgctgagagatagccag cctcagttacaaactacgttttggggctatg acacacacgggtgctttccagctttagctc tgtcgtcaactgttaaacatccctatttggtta aggaaagtagatggacacgacaagccctc aaaacattggcgaggaaacattaccccgga caggggagacattcatgtctaatttt
720-721	Chloroflexi bacterium isolate CF_154 14_0903_05_20cm_scaff old_339, whole genome shotgun sequence GENOME_ACESSION: GCA_005879655.1_AS M587965v1_genomic GENOME_ID: 285751 CONTIG_ID: 335 SOURCE: NCBI_Prokaryotes DATE: 2019-07-22.	atgtcatcgtcttctgtcatcgataccgataagcagccgtttccctgtccatccggggcg agcacgcctgctgctcaaaaggggaaaggccgctgtataggaggtatccgttcacc tgatttcaagcacaaggttgacaacaccgtgcccgcaccgttacgtctcaagcaggata ttcgagacttctgtgtatcagcccgaccgactctgcatgtgtaacccaacgaaagt acctctcaagatgccaccgtgtgtaacgtgacccgctcgaactcctgcggcggttga ggccattgggtatccctagaacaggttcaggttgagcgcaccaataacaaccgcatcat gcgtggcctgccccaaacaccctggagacatgcagcctgtgtagtgtgtagggattgggg gtccaggggggcgaagccaccctggcgggtttgggtgtccccaaatctccattccc caaagggttttgggtctgtgactaggcagggcaaaatcactgatgttccacaccgctttt tcagcccatcatcgtatgcatgctatcactatcaccgagtgaggatttccccacc ccaactagcccgaaggagcgcctgttaagtactccacttgcctag	gtcaagtcccccatcctgacgtgtcagggt ggaggtcgggaagctagacttccaagccg gaaactgtgacgacagagggttgaataaac aaccacgttatcagtaagcgttaaaagacc tacctgcgggtgtctaccagccgtagccc tagaaccgttcagttaaacaggtgtagggg agaaactagtgtgagcggaaagtaccgg ctgataactgtgtcaggtgagcatcactg ggaaaccagagcccacggggcataaaa caaatagaaggaaacgacatccgccatgt catac
722-723	Ga0183746_00132 GENOME_ACESSION: IMG_3300037962_SF_3 300037962 GENOME_ID: 309636 CONTIG_ID: 131	atggcgggtttctgactggacaaaaaagcaaccgctgatgccgtgagtgaaaagcg cgctcgggtgtgtgtgctggtgtgtgtccgtgtgtgataagcgctaccgttcacgatt cggtcagaagatcgctggtggcgagcgcaaccgttcgctggcgatcgaccctg gcagcaaaacgaccgggtgtgctgacgacgaaccgacagtcagcagcggcatg tgctgtgctgtttgaactatccatcgcggttctgagatcaaaaggcaattacagcaacg cgccgcttctggcgctgcgccgcagtgcccaacctgcgtaccgcgccccgcgttca	gtcaactacccccgcctgaaggcgggagct tgtgaagcaagcctagttgaccagggttaa gcggcgctgacagccgtacgttgataaca ggctgccaaagactaccacaggtgttccct cagtcctgtgctgtaaaagtgcaagatcatg ctggcgagggttaaacgccgaaggtcttg

SEQ ID NO:	Name	IscB sequence	ωRNAs
	SOURCE: JGI DATE: 2020-04-07.	ataatcgcatcaaacctaaggctggctggcaccgtcgctgcagcatcgggtggatacgtgatggcgtgggtcaatcggctatccaaagtcggacccatcaccggcatcggcaggagtagtgctgttgacacgcagaagttggaaagccccgagatcagcggcgtggagtaccagcaaggcaccctgctgggtctacgaggtgcgcgaatctactggaaaaatggggcggtagtgctcctactgtggcaccgccgatacggcgtggagatcgagcatatggtgccggtttcacggcgccgctgaatcgctcagcaatttgacctgaagctgcatccctgtaatgcgagaaaatagccagacactgaccgatttttcgccaccaaaaaggcctgaaaaagcgccctcaaggctaaccggcttatcagctgacgcacagctggaactgtacagcgccagctcaagcaaccgatgcgatgccagtggtcaatgcgaccgtggcgctgttctccgcgtcaaaaccacggattaccggctcggctgagcagcggtggccgaccaaatacaacggcagcgccctcaacatccccaaagacgatgccctggacggcggtgtgtcggagcttttggttaagctgtatgactggacagacggacactgacgatcaaggcgatggggcgccggcagctaccagcgacgcgcttaacgaacacgggtttccgctggctatctgatgcggcagaacaggtgcacggcttcaaacggcgacatggtgcgtcgatagtagccaccggtaagaaggccggtaccatacggggcggtgcgcgatacgaacacggcagtttcaacatccaagccgaacacgggtgacgtgcaaggcatttccataaatactgcacccatccaacggagcgatggctacgctattatgacacacttcagcaacctaacaggagaaggcgggacggcggtggcgtaa	accgctgccacaaggcaggagccgggtatcgacattcccaggggagacggggcgtaagccctgcgacacaaggccccgtaaggcgctattaaaggagaagatcgcatggcg
724-725	Rock porewater metagenome BRHa_1004791, whole genome shotgun sequence GENOME_ACESION: LADL02.1 GENOME_ID: 1462 CONTIG_ID: 3045 SOURCE: WGS DATE: 2019-01-27.	atgcgctggctgctaacttggcgagaggacattactccgagaggaggtagactctatgttagtctactgtatcaacaagcacggaaacggtaatgccctgcaaacacagaaggcaaggaggtcgtcgaagaggcagagcggaaggtcataggcgccctttacggctcagttgtttacgggtcgaagcgtttacagacaacccgtcagccctggagttgatacgggagccaagtacgtgggggtggcgctgttcggacagacgaaaggaagagcaagaatactttgctgaaggcgaggtgctgagcgtgagcagatattcggcgcaagatggaccgacgcaggtctatcgacgaacacagacgtggcggaagacaggtatcgaaagccaaagattgataacagggcgtgaccggaaggatgttggcaccagcattcagtcagggttgcaggcagctaaagggtgaccgatttgcctccgacgtcgtccggctcagtagaagtgagaccgcgagttgacacggcagcaatggcaagaggagtcctgcgactgagggccctggcagttaccaacggggcgagcagttaccagtttgaacgtcaaatcgtacgttcgacatcgaacggatcgtatgcgacgtgaagccaaaggctcgttggaggttaccacatcgaaaaaagagctgatgcgggaacagatcgccccgtaacctgataacgctgtgtgaagggtgccacgatcggtccataaggcggggatcaagctgacggcggttctggaagaaaccaatccgggtatgccgccacacccaagcaggcaagcagcgccctgtgtgagccctacgggaacgactcccgacttccgagacgacagcgccgtgaccaaaagtcgatcgactgagatgggactatcgaaaacccatgcggcgatgcgctggccatagcggtacaggggtacccgtcgaacctgtggacacccagtttttatgcgctgtatccaaagggaactaccggctgttcaaaaggtcctgcagccatattccgtaacagagcggccggaactgtttggcttccggcgtcgtgacaaaagtatgcttggcgggtggtaagaagggttcgaaggtaaacggacctcgtgttacttcaatgtgagcacactgacggcactgtgatcagcgctcgtatcgtatcagaagagattgcgtctattggagaacaacatcgctcgtcgaagaggcaggccgtttctatgaaggacaccaggctgtgtcagcggtactgttttacttaggggtccgagatcgttgaggggatcttcaacagtaatgcccggtatttatcacttcaactcagggcagccaatatcagggtgctaaagcacgttccgttgacgatccaaagtgcgctaaggaggcgctcaaggaagccattaaagaaactgacgtgtgtgtagtcaccgggggctaggcccaa ctgaagatgacataacgaggaagcgccgccagcgccctgggagttggcctgactcaagatagggaaatggttgagcagatcgagaggtacttcagggaccgacacctgacaatgacgagaataactgaggcaggaatgtctctgaggggcgagggcgctgactaacgatagaggacagcacctggctcagttcgttactgtagtgatctcagaaggccctgtgtccttgccggggccacctcagaatggctgctatgtatgagaacaggcgacggcggttgtagagatttgcctatggccagcgagggcggttcagtggttctcaagcagatttcttttggcctgggtgagtcggagctagcaaggcgactagaggggatctgctgacggcagagggcagggacttaagattgctacaatggccagtgaggcgacggtaactctgcgcctggcgccgccagcgccgctaaagctgtaagaacagccgggcaagctgtccacgggcactggcgaaatcgtctacggggaggataccgataccctccaggggctgttggcgtggcgtgattgagcggttttaagatggccatcgctgaatcctgcaccggtggctcctggcctatcaacttacacccgtggcaggtggtcagacttctatagggtgattcaaggcctacagcaagcggttaagaaggtatgtgtaggggtaacaaagacatttctggaggatggggcgttagccctcagtcgctgaggccatggctctaggggtgatcaacagttgaagccgatattggtcctcgattaccgggttccgggccaagtgaggagcccggagaacacgtgggaactgtcgttctgtgctgtgctcagatgtactgtgtagatggacagcggttccaggagatcgggataaattcagctcagccgcgactcactgcctgggtctgtgtgaaatccctatgcaggaatacctcaaacctcccgatgagagacacgatgagagaggtgttaa	gtcaactacccccggactgaagtcagggtgtgtaaaagcctctagttagaccagcctaagccttgaacagaggctacgttagcagtagagcttagacgccttaccgggaataccggataggcgactacccctgggtgccccctcagctcca ggctctagaactccttggttaaatgtcctgagagggtaggacagtgccctgggagatgcgctgctgctacacattggcgagaggacaccttaccggagaggtagactctatgtta

SEQ ID NO:	Name	IscB sequence	ωRNAs
726-727	Symploca sp. SIO2E9 2E9_NODE_8, whole genome shotgun sequence GENOME_ACESSION: GCA_010692645.1_AS M1069264v1_genomic GENOME_ID: 325549 CONTIG_ID: 361 SOURCE: NCBI_Prokaryotes DATE: 2020-04-22.	atgcaacgagttccagttatctcaccacaagggttcactgatccaaccaaaccttcaa gagccagacgttgctccgtgaagataaaagcaagattatccgatgcagcgtaata ttcgtgtggctgttgggagggaaaagcaactatttctcggtccctcaagggtatagc caagccgaatattatcggcgtaacgtacattttgagaatcccgatagtaagaacctaac cccacgcaataccggaaagcgaagggtggaacaattacaccttgggtttgatctgga gaattttgtatgtctgaaaaagcaggacttatatacagaggttgggtgggtggtcacacc aaacagcgaagtcgaaaaacgtttctgtttacgacagtaactggaaaaaataggcgagt tttcccttaaaaaagtcagttactcaagcggttcaggttatgctgaagctga	gtcaactactcacggctacttcttgagccgt gagcttgaagagccagatttcaacaatagt tgtctaccgcgagagtcacgcatgtcacac acggcacaatgttctcctagtgtgctgtctg taagttccctggcaccgagcgctttaagaga ggacatgtttggcgatgagcggtgaaggag caagtaactttctctgtaggattatcctatgca acga
728-729	Ga0374055_0154238 GENOME_ACESSION: MBmetcoassembly_\$F_3 300035696 GENOME_ID: 306290 CONTIG_ID: 154237 SOURCE: JGI DATE: 2019-11-15.	atgtcgggtccagtattatgatacaataaagccattatgccttgcctggagaagcgag caagaaaattatggagaagggtcaagctaaaccttgcgtggcaaaaagggtatattctgtatt aagcttatgcaagagccatctactagaactatcagaaggtcgtctgtgtattgatctgg aagtaaacgagaaggctataccgtctctacagaaaaagccgttgccttaacgtgacaaca gatacaaaagattgattagaacacggtgaagtaaggagaagattctgtatggcaagg aggcaaaagagacacctatctgtaaacgcagatctaatgattaaaggaacagagtttttc ttccaccatctaccagagcaagatggaataactaagctacagatgcaagtttattatctta ttctgccataaataagttatgtagaagattttaaagctgctactaagaaggaaaaaac gaagtggaacagatcatttccacattgaaagtggttaaaacttgggtctataatcaggtag aaaaacttggagtttaagttatgattatacaaggatctgataccaagatagaagagacgc aagaggatttagtaaatcaaaaagaagttggactttatttggagacacataatgtagact ctcatatattgtgtgacttgcctttagtaagcaagttaaagccatatttggcatatggaagat agaatttctcctattattatagaagacaattgcagaacagaatataagtaaaaggaagaaag agaatagagtagtggctctacagtaagtagtggaaatgtcggagagggctctatagcttgtatag gaataaaatgtattatcttggaggatcatctaaaggaaagtaagtatacttataattac aggagagagagactga	gtcaataacctacgacattactctgtagcgtg tgcggagaaaacttaaccaggtttttctcgatca agactaaaattgttactagacagatatata gttgatcagctctcgttatgtaagagggttaa agctacatagagaatgttctgtagttcttttc cctcgggtcagctatgtcgaagcgaggttgaa caaggatgcgtaagcatatttttagaaaag agaaaaaaaatgtcgggtc
730-731	Ga0310147_000821 GENOME_ACESSION: K73_2_\$F_3300034683 GENOME_ID: 292989 CONTIG_ID: 820 SOURCE: JGI DATE: 2019-08-09.	atgggctgtatcaagcgggggagacctaaccaacctgtatgtgtgatttaccttgaga aaggtggaagatggtatattatatacaaaagcggaaaaccgctatgccgacaaaa cggcatgacagggtaaggatattgtgaacaagaaagcaatgtgtatgccagagcaaa accgtttactatcacgcttttatgacagcacacatcacgcaggatgtgcatctgcat atgacacggggcgacacatcagtcacacgcgttattgacagcagcacaaccgatgttt tgtactcaictgtgaatcactgcccaacaagagcgtgccaaagctgatgaaagaacgca agatgtaccgtatgataaggcgtcatacaggcgccggcaagaaaataaggcgtgctatt gcgaatcactatttccgagcaccaagaagaggtgtacagccttgaactaaggagcct ataacagcaaatgtatgtcaagcgaagcaagcgggttcagcaacaggaaaaagggccg aaaggttgcgttaccgccgacagcccggcagttattgcagactcacatcaactatttcaata aagttgcgaagatactgcctatcaggaaaggtggtgctcgtatgtgcaagttgatagca aaagcttgagaatcccgatatagcaggaagcagtaacaaaggtacgtgtacagcta caacaacatgcgtgagtataattatgcaaaaacaggaaaggttaaatgcctgttatgggcaa gcgcaagatagagcatttgcacatatagtcccgcgtcaaaaggcggcagcgatacct acaaaaacatagcggggctatgcggcaaatgccatgaaaaagtcacaaagatcctaa ggccggacgaagcgtggcgagaagcagccggcagcgtaaggatgtctgatcc gagcactactcaataccattatgcgtatctctatgaatatcttaagtcaaaagctcggcgaag agaacgtagagatagcttaccgggtacgaaaccgagacaataggcggcaactggcgctt gagtaagagcattataatgacagctatgcgttggcgcttatggagtcgggcatataag ccgtatagagaagataaaaccgtatgagtataagcagtagacagcgctataaccgtatgtt acagatgcacaaaggacagactatatacaaaagacagcaagatagtgccgagggaac aggcacaagaaaaccgagcagggaagagccgtcgttgaagaataccggcaggagata aaaggaaactgcaggaaaaaggaagcgtcacgtgccatattcgggtgaaagtataca gggcagccaagcggatgagaacgcccggcaagacgtaccgataacaagcgggaagc agcgtattgtacaagggccaacgctttattgtcaaaaggcattttacacaaaggccaatcgtt gttgttgaaggacatgacggttatgtgtcgtccggcagttgcagactgatgacaggaa tacaggcatagtagtcttga	gtcaactactcctcaataaattgaggagcata cagccgcgttgatgctatgttgcactttagg atgcggtagttgaacaggctgctgatacag gctgcgcactccgggatgccactctcagttc cggacgctgatgggtgtatcaaacgggg ggagacctaccaacctgtatgttgattt accttgagaaggtggaagatg
732-733	metagenome genome assembly, contig: NODE_3892_length_550 3_cov_2.770374, whole genome shotgun sequence GENOME_ACESSION: OGDPO1.1 GENOME_ID: 4288 CONTIG_ID: 3891	atgtcaaccggccttcgggcccgaagcaggcggcctgcgggtatccgcgaaggatgta ctttgaaagtatttttgaacatgcgaggtcagcccttgatccgtgctcgcctgcaaaag gcgaggcatctgctcaaggcaggcaaggccgtggcaggcgctgaacgccgttcacg attcaactcgaatcgtacaggtgaaacgaagcagagcgtgacgttggcggttgacgc cggcacaaaagatgtcggccttccgtacacggaaaggaagaggtcttctgctccg aagtcgaactcgcacaagacatcaggagcttctgtgctgctcgcgtctcgcagcgtcgtg agcggcgtatcgaagacgcgtacccgtgctggcgcttttgaaccgtgtacgctcga aacacaaggagtggttcccttccgttgaacacgcaattcaggcgcacatctcgcgca ttgaagcgttttccgagtggttccgtacagcaagatctgattgaaccgcatccttcgac attcagaaaatcaaaaatccgaagtcgagggtacggactatcagcaggcgaccagct	gtcaataacccctgcctgaaggcagaggctt gaaggagcctttattgactagcctcagcgcc cctcttgcagaggtgctactgttgggtggaa tgtataggcaccgtggaatgtacatcctagt ccacgcactgcggcctgtgattaaaagctcc gagaggtaaagcgggtgtgcaggcacca aaccctgtccaacattggcgaaggatgtca accggccttccggccgagcaagcggagcc tgcgggtatccgcgaaggatgtactttgaaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
	SOURCE: WGS DATE: 2019-01-27.	<p>tggattctggaacgtgctgcgagtagtctttccgtgacggatctgtccagcactgtc gaggtcgtatccaagatccgatcccaactgtcatcatcttgagagccgaagacaggc ggcgtatgcgcccaacaacctgacacgctctgcgagacatgccacaaggcatatcacg cagggaaggatcaagttgaaggttggtcgggcacgtctcaggggcagaagccttcatg ggcatcatgctgtggacgctacttgaccgattcgcaaggcacatcctgaactgcctgtc gaaaacacctacggctatctgacgaagcacaagcgattatccttggcttgcgaagacg cattgcggcgacgcttctgtattgccggaacctgaatgcgtgcgacgaggagaaatc ctaaacagcgccagacgcgcaagcacaacctgcagattcacaaagtgtacggctctc gaagacgcttacggacggcaggaagatcggaatccgaaagctcaatcaaacaccgcat ctggtcaagaacttcaggctcttcgacaaggtagatgtcttgggcaaacggcttcatctt cgggtcgaagatgtccggctacttcgatgtcggaggctggacggcgtaagctttctc cggcatcagctacaggaagctcagcttctcgaagagaaagcacctatttaaccgaact tagaaacgaggacggcgcttctccctgtctga</p>	
734-735	Ga0310695_10005296 GENOME_ACESSION: RumRJGDNav2_\$F_330 0032007 GENOME_ID: 280889 CONTIG_ID: 5295 SOURCE: JGI DATE: 2019-06-13.	<p>atgggtccagttttagataagaacttaattccattgacccgtgtaagaacgtagagcagct accatgatgaaaaaggtagagctaaaccgtattggagagatgtatctttgtattatacta caaaacgaacctctgtagaaactacagcgatgtattgttgaatcgtatcctgttctaa acgtgaagggaattacagttgcaacggaacacgagtagtctcaatattacgtctgaagcg attactcatgttaagacaacgtagaactagacgtacattgaagaagatcgctgtcgaac gaaaaacaccatactgtaaatgtcgtgaaatcgaaaaaataaacaagacgataaatt gccatcatcaacaaatctcgttgggagctgaattacgtatacttaaaaaacttaacaaat acttctattaccgatgtttcagtagaagatgttctgcgaacacgataaacgtgcgtcta aatggaataatattgtttctcttttagaaactgtgaaatcgtattttacaaagctatcgaaggt ctaggtttaatcgtattcaagtggaagggttacgagacgcacgaatggcgtttacacgctg gatataagaaaacgtctgaaaaacttaaaaaatcggtgggaagcgcataatgtatagatgc attgtttatgtgaaatgttttaggttattgtataaaaccagtaaaatattgtgtctactatcgt ttttacaaagtcataagacgtaattctatttaacaaactattctaaacatggagctagaatag gtatgtgtgtacaaatgtcgttaggttaagaagaatcgttagtaagatatcctaagtatgg tttaagtttagtgggtgataacgaaaggtaagctatcgttacataatatacacaacggtag tcgtcttgcgaacgcgaataaagtgacttaacgatgttttagttacaaatttaaaatgggttt tacaactgatagttccaaagtgtgtaaaagcaggctcttcccatgactaa</p>	<p>gtcaataaccaccgactaaatcgtggcgtt gtgaagaaaactgaaggtggcgttcaaat aactcacaagggtgaagaagagactagattc gctgagtggaacaacacagcttgaagga attatcgcagatggaaagccaacagtgga cagaagacgtattacggaatccctcagtt ctgtattcaacgtcgcgtataatcgaagg aattcaactagaggtgcagttactgctgca cactccgaaaggagaagatgtttatggctc ca</p>
736-737	Human gut metagenome isolate term5_stool term5_stool_scaffold_21 28, whole genome shotgun sequence GENOME_ACESSION: QWCB01.1 GENOME_ID: 14510 CONTIG_ID: 2128 SOURCE: WGS DATE: 2019-01-27.	<p>atgggtatgtaatatcaatagatgaatgaaccttaattgcctaccacaaaggcatggtaaaagt agaaggctcttaaggataaaaaagcaaaagtgggtagaagagaaccttttacaataaag ctactctatagaccagaacaaatgtgtgacaagattgtaccccttggtatagatacaggctc taaatatattggagctgccgtagatcaaacggcgaaatattatagcatcagaagttaaaa caaaagacgatgttaaaaagaaaatgtagagcgtgataaactatagatccaacagaaga tatagaaaaataaggtatagaagcctgatttctaaatagaggttaattccacaaagaag gtagatataacctaacgctgtgaagtaatttaattctcatgttagagaaattgaattatcaaa agcatactaccaataaacaagctagatttgaagttaggtcaattcgtactcacttgatgaa aaacccaagttaataatcccaagataagacattgggtttacaaaaaggaacaaattac gttttttcaaatagtagagaacatgccttaaataggaataaataatattctgtcaatgtgtgga gttaaaaatacaagggttagaaggtcatcataatataatgatcgaatggtgtactgatgact tggataaatttaaaccttctgtgaagaatgtcataaaaaagctacacgctggaataataacta taaacagaagaaacctaaaagttaacttaattcagtgatgcaactataatgtcaacttagaa gtatgttattaaaaagataccagagcgaatagaacataatggctatgtaactaaagaaaat aggataggtctgggtcttccaaaagaacactttgtatagcctgtgttattgtactggagg ctatgactttgacttgccaaaagagatcttgttaaaaagacatgtcttaaaaggagattatca gctaacaaaagggttctgtagtgaagaagattcctcaaaccaaaatactggcctttaga aaatgggataaggttaaatatctaggaggaataatacttttaagggttaggcctctaggg gcacttgtgtctagaagataatataggcaataagggtgatttttctcatatgccaagagggtt caaggatcctaaacttttaattgcaaaagaattgtgcaaggaaaacactttatgtcaaa gaatagaatacatccaaaggcaagttta</p>	<p>gtaaactacaacgccccttaagggaatgttgc ttaggtttatcaaacacaaaccttgggata aagggtttacagactcagtagtagaaatgtt tactacgataatattgatgcacaccttgggtt cgccccagaccattgctctgtgttagcattt aaacagttgcgataattaaagttgtaagcaa cagtgatgtctaacgctaaactaatatattc gtcagggggaagggttaaacattctatgtga aaagttagtagaaaaacatacagttcgttat gtactgaattaaagtttaaaacataataaagt aagaaaggagtgcccttatg</p>
738-739	Ga0111054_100026 GENOME_ACESSION: IMG_3300008541_\$F_3 300008541 GENOME_ID: 28547 CONTIG_ID: 25 SOURCE: JGI DATE: 2019-01-27.	<p>ttgaaatagagaagaacttaaggttaattaaaatgggtatgtacaatcaatccggggcaa gaaattaatgccttgcagaaaacgaagctcgacatcttcttaagaaaaacgagcaaa gattattaattatgaaccatttacgattaagctgtgttttaaatgtgaaaacaaaacacagcc gataacgcttggcgtagacacgggttcaaaaggtgttaggtcttcagctacaacttcaaaa aagagctgttttctgagaatttgaacttagaataaacattgttaacttatttctactagact aaacttagaagaagtagaagaagcagaaaaacaagataccgtctgtcgtagatttttaaac agaagaaaaacaaaaaagaaggatgttacctcatattaaaaataaattaaacgctc atgtaaaaaccatagacagtgatttaatttacttccagtgactaaaaataattttagaaactgc aaaatttaattgtctaaataataatccagaataaaagattatacttctgcccctcaaaa aggatttggcaatgtgaagagcttatccttgcagagataattatcaatgtcaatcatgttaa aaagaaaaatgttaacttcaagttcatcatatcgagctcgaacacgtggagaaatgct ccaaacaatttaataacattatgtgaagaatgtcatttaaaatatcattccggagatttaagt taacttttaaaagaggaaaacttttagagatgcaacctttatgtcaattttagaaaaagact tccaactcaattaaagggaaggtatttcatctatccaaattgaagagacttccggtatataac</p>	<p>attatttaaatcatgctaatgcattttatta ttaaatgattagcctaagcttatttgactac gttactaagaataataggtactcttggagt gttactcctagctccagattctacgatattg ttaaatgttctgatgaatatgaacgtgctgt atgtgcgaccgtccaattttttagtagatgc taaaccttagaataacattggcgaagggt cataaccttctgaatatagaagagaactta aggtaattaaatgttatgtacaatcaatc cggggcagaagaa</p>

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aaaggcgaatagagaaaaagcagggtctcccgaagaaacatagatatgatgcttgggcta tttcaataatcctaatacgacagcttgggtctgagtggtggaataagcaagtcgaaa gcataacaggaaaatacataagctactcctaaaaaggaggaaaacgtgatttgaac aaagtcataataaactcatgatatgattatgacaaaaataaatttaacaatgaaattt ctttatcattgctagaagattaaatggctgtttacgttaaaaaatttaaacaggaaacttgc ttgataaataagcaaaagttaattcattttattcagttagaataatagttttattagaagaa ggaacaatcaatatgaatgctcgttaa	
740-741	Ga0272423_1005921 GENOME_ACESSION: MtNewZealandMG_2_5 F_3300033168 GENOME_ID: 280697 CONTIG_ID: 5920 SOURCE: JGI DATE: 2019-06-13.	atgtctacgttgacgcaagacagaagaccaccaaggaggtctcctcagtccttgcct ctggaatcgtgtcagcagacaccgttggggtaacgacgaaacggggcagggacgcg ggcaagaccgcacctggcagcgaacatggtcagggggagaatctgcaccgattcgcc agcgcgcagatcgtcaccgggtaacccccggaagtgggagtaatcccgaccatcagt gagaacggttcagccgagcgtgttcgtgctggacaagtgcgaagcagccactcatccga ccagccccgcgcgggcaaacagtgctcggggccggtcgcagcagagtggtccgac tccaccggttcacattcggttgggtggacccgacccgtgagcactcagcagtggaagg gtgcaaatcaagatcgaccgggtgagccgggtgactggcgtctcgtgtgctcgggtacc accctcggaggcgggtccacggatgttcgtgacgagatccagcagcgggacagacag atcagtaagaatcctaccctccgtgcccgttaccggctggagcgggctcagaaacct gcgacaccggggcgcacgggttccaaaccggatgacaccgagggctgtgttgcctccc tcgctgcaacaccgggtggactcgacaagctcagtggtgtcccgttgcgcaccatcgc gcccgtcactcgtgtggtgaggaatagtcgggttcgacctacagaaattggtcaaccc ggagatcaccggcaccgaatatcagcaggcgaccctcggcgggttcgaggttcgcgaa tacctgtcgcgaagtggcaccggagctgtgctactgcgaggtcagcgggggtgttgc ccgttcggtgcccgtgaaatcagacatcgtacatttcgcccgcgtcaaaaggcggctcgaatc agtgtctaattcgttgcctgtgctccgtgcaaccaggccaaagggttccggggtgtg gcccgtgtgtcaccgaccccgccgctgcccgggagtcgctgctcaagcccagcggc cattggttgatcggcggcgggtcaacgcaacccgttgggtgctgtacgagggatcaaaa gcaacgggctgcccagtcacaccggcagcagggcggccgaccaagtggaaccgagt gtctaaccgggtaccgaagtcacacccctgcacgactgtgtgctggtggtatcactgg tgttctgctgggtgcccgaacgtgtgctggtgagccacatccaccggcgccgacactac cccgccaccgtagcgacaagtacggattccccggctgcgggtgactcgacagaagcg gcacttcgggtcggcaccggagatcatgctcgcgggtcgtgtccaccgggaccaa ccgggtacctatgtcggacgctgagccgtgcgcggcagcggcagattcaatcaccact accggccagatggaagtccgtcaccgtccaaaggcatccatcaccgacacttccgct gctcaaccgagccgacggctggaggtacacgagaaggaggaggagcgcatctaa	atcaaccacctgcctgaagggcgggg ctgtaacggagctatccgttcccgaatgag cccagccgctagggtgtgtgaccagacca agacatcaagattggaggcgaacagatgt ctacgttcaggcgaagacagaagaccacc aagggtgcttctcagtccttgccttgaa tccgtgtcagcagacaccgttgggtgaacg acgaacggggcagcgacgggcaaga ccgcacctggcagcgaacatgtcagggg gagaatctgcaccgattcggcagcgcga gatcgtcaccgggtaac
742-743	Ga0116158_10014977 GENOME_ACESSION: AD_UK052_MetaG_2_5 F_3300009783 GENOME_ID: 21284 CONTIG_ID: 14976 SOURCE: JGI DATE: 2019-01-27.	atgggtccagttttagacaaaaataaaatccgttgcctgtcagaaaagagagcaa gaaaactattggaatgtaaaagagcaaaaccttttgggtataaagcttcttacaatcattc gcaagaagattcttctggaacactatgcaggatattgtgtgagacttgaccgggaagt aaaatgaacggaatgactgtaaaatcagaatcacacactctttaaattcttcagggttaagc aagatgcgacgtaaaagaaaaagtagaaaaagagcaattatagtagtagtagacgg cgtagaaattgtcctatcgtaaatgtaggtataaaagaaaggtcgagagatgccgccat ctacaaaaactagatggcagcagcatttaaacatggtcaaaccttgcctaatctgtatccg attactcatgtttctgtagaagatgcaaaagctattacaaaaaacatgcacggaaatggaa tgtgaattttagcccattgaagtgtgtaaatcatgtttttatcagaactgaaaagagttat aaactttatactttcggcgatagcaaacataggggaacgtaatgctttagggttataaaa atcaaaaaataaactagaaaaagtgtttctgcacattgcgtagattcttgggtattggtctaat aaagtattgtgtggcagctcaaacctgaatatacaaaaggttattggttactccgttaatt tattataaacgacaactacaggtttttaccagtaaaaagatttaggaaaaaacacggagg aacatctacttttggaaatcaaaaaaggaacgttagtaaatagtaagcgtcagcgactttctat tattggaggatccgctcaggagggtattgttcacactttaaagtgataataaaagatttac tcaaacgcgtaaaaaagaagaacttactcttcaacaactttaaatttatataaaataa	atcaactatccactttaggaaattttaggaa aacctttaccaagactaactgaatttaaca ggagaatagcttgcgaataataaccttc gaaagtattgtctgttgaaagaaactaagat gataggttggtagacaaaacacataga ggatgcttcccaagcttcttcaatgtttc gagtagcgaaggaaatttaaacagatgc gtaagcataatttttaggttcca
744-745	Ga0207997_1002358 GENOME_ACESSION: AntAce02UKSPAdes_2_5 F_3300025736 GENOME_ID: 22957 CONTIG_ID: 2357 SOURCE: JGI DATE: 2019-01-27.	atggtaaacacaacagaagttatcaagagagatcatcactagagatacttctc agtctctggctctatgattagtaataaacaagaggaaactcttagtcataactaattcaaa aaccttagataacaatctcgaagagaaccaacaccacaacacaggtgggtttaagacca aagtgtttgtaatagcaaaagtggaaggttttgatgccactactccaagaaaagcaag gcactttttgaacaacagaagcaaaagtagtaacgacaaaacattgcaatcaactt aattgggattgtgaagaaatagtcgaagagtaatttaggaattgatactggagtaaaa caattggttatcagtaacatcaaaaactaaagaattgtttctggagaatttgggttaaga aaatattagtaaaaaaattagcagatagcaatgtatcgaagaaataaaagaataaacctt ggtacagagaagcaagattctgtaatagaactaaatcaaaacctaagggttggttagcac catcagttcagcataaaattgattcgcatatcagattgattaataaaataaaatttgcattcca ataactaaagttaattgtaatacaagtaattgtgctcaaaagtacagaatccagatattg aagggaagtgaatatcaaaatggtcaaatgaaagatttcgagaatgtaaatgtttgttaga caagagataataacattgtcaaatctgaagaaaaagatgataaaattgttagatgtica tcataatcaaacacgaaagatggtggaagcgatagaccagataatctgataaacattcat caatcatgtcataagaatttcattctgtaaaaatcaaacatgtatttgcacccaaaatcat	gtcaatcacctcgcctaaaggacgaggctt gagaagtgttttcaagagtaattgttgatta gagaggttgaataatggttaaaacaacacag aagtattcaagagagatcatcacttag agatactctctagctcgtctatgattagt aatgtaaacaagaggaaactcttagtcata ctaattcaaaaaccttagataacaatctcgaa gagaaccaacccaacaacaggtgggttt aagaccaaagt

SEQ ID NO:	Name	IscB sequence	ωRNAs
		tcaaagaaacttcaatgatgaatagttatgtgcaagattgaaatatttagtagattgtactga aacatttgggtatattacaaaataatagaaaagaattaggacttgaacacacattataa tgatgcgtttgtatfttctgtggaaactaatcaagaaagatgtcaatcaatgtttcaaaaca aattagaagaataatcagcaattacaacaaatagaaggaacaaaattggcaatcag aaaagaacgatataaaatacaatctggagatatttttatcacaataaaaaatttaattgt aatggaatgttaattaggttaatatgttaagtttgaataaaatataatataaaatgtca aaaataatgatattaaggattattattatgttaagggaatcaaaatataa	
746-747	Ga0101770_1107140 GENOME_ACESSION: IMG_3300006674_SF_3 300006674 GENOME_ID: 27981 CONTIG_ID: 28184 SOURCE: JGI DATE: 2019-01-27.	gtgcaagattgtatattaggaattgacgcaggaagtaaacataggaatggcagttataa ctgaacaaggtaattgtctttatcgtgccgaagcagaattaagacaagacataaaagaga atattgaacaaggagagacgactcgcacgtgcaagacgtaataagaaaacaaggataga aaaccacgttttttaacagaaaaagaaaggaagcgtgttaccgccaagtatacaatca agaattaatgctcatattagattggtaattgatattgtaaaaattttacattataagcaataag agtagaataaggcgaattgatacacaggcattgtaattacagatattatggatagaat atcaacaggggagaaatgcatggatagatagcgttaaaagaatattgtaaaaataagagata attttactgtcattatgcaaaattaagaccagacatacctgttaattgatattctgaagtagac catattataccaagaagcaagggtggtagcaataatcaagcaatttag	gtcaactaccccaacctatagaggttggggc ttgtaaaagctcaagttgactagactaagtgc cgaataagcactacgttattctccatgtagac accgtagaatgatccctagtctctgtctg tcgtggctctgtaaacagtcctgtgagtag ggacagtaaccacgtgtgtaaggaggaat aacattgtcgaagggaataaactccgaaa ggaggcctaagccaattgtta
748-749	Ga0307373_10069487 GENOME_ACESSION: OX2MG_2_SF_3300031 672 GENOME_ID: 280771 CONTIG_ID: 69486 SOURCE: JGI DATE: 2019-06-13.	atgccggcagctgtctcgcggcgagatcaagcacaagcccgccatccacgcaaaaga tgcctgcggcgggcggtacggcgccggcgccgctggccaacctgcgcgacgcg cccgacgtttcccaaccggcctccgagccgtgtgcatcgtctgcggcgcaacgccca acacggcgcaagcattgcccggcgatgcgccgagggcgggcgtaaccggacgagtt ggggcgcggttccgcctagcgcggaccctgcggagtcgtgtggacagcaccg ctctgggtgcggcggtgcgccactggggccccggtagcggcgccatcagtgacgttg ggccgctacgacaccagggccctccaaaggccccgagatccggggcgctcagtagaccag cacgggaccctgctcgggtacgaagtgcgggagtagctgctggagaagttcgggcac gctgcgttactgcgtggcctctcccgcgaccctgtgctcaacatcgaccatgtgtgcc gcgcagcccgggggcgaccgaccgctgtccaacctggcgcttgcgtgccgctcctgc aacgaggcgaaaggacgacgaccgccgaggagtgggcgccgctgcacgct gcattccccgttggcgctcgccggacggcgggatgcggccgctgtggggcgct gcaggcgccgttgcggcgccgctgcggcgaacgccaccgctggggcgctcggcg cgtcctgggggagacggggctgccccggaggcggggacggcgctgcgcaccaagg ggaatcgcgcacgctcgggctgcgcggagcccgctcccgatgcggcgctgcgtgg gcgccctgacggcggtgacggcgctggggcggggaaggcggtcctcgccatcgcg gcccagcgccgg	gccccgttacccttgatgcttaccaggccc agcgccagagggtggcagccgttgaggggc gctcagttgcccgaaggtccaagaccac ccccggatgcttctcagtcggcgctctg aacggcccgagcagacaagccccgttg gcagggcgacgagcggcgccggatg gccccggcggtgacatggcgaggggag cggatcgggagcgccagcccgateccgtc acacagggccccgaaggcgccccgaa gaggacgtatcgagggaaggggcgcg gccccgtg
750-751	human gut metagenome genome assembly, contig: NODE_229_length_1448 6_cov_2.676301, whole genome shotgun sequence GENOME_ACESSION: OQLS01.1 GENOME_ID: 9921 CONTIG_ID: 228 SOURCE: WGS DATE: 2019-01-27.	atggaagttagcgaaactatagacgtttaccatcgtttacacaaaagatcggcaggcg agattcgacaatcgttaaatcttctaaagaaaaggacgaattgtccaaagtattttacagaa acggcaagctaccataagagttatfaatcggcttaataaaggataaataacaaattattg gctggaagatgtttctattgataccgggtataacagatgggtataagctttatagctggca atatcaaaaatccaaagactggagcaaaaatcgttaagcaaacgatttctaagagatgg cgtgaagtgtatggaatgcggaaaatctaattgtagattagaggttcaccataaagccaa gaagacgaaatggttctaatacgttagacaatctgattacattatag	tatataaatgtaataagagtgacatatgaa atctcttatctatagcttaacaataggtctcaa ggctaagtgtactgtactgtcgaagagacatg ttgcagataggaactatgttaagttagaaggt aaagacacacctttagatgtaattctcagctc gaagctctgtgagtacaacacaggaaacaa tgctaattgtcgtcattgataacagggaac acatactctacttgactttggcacgaagag caattctccgaaaggaggtgtcagaatga caagt
752-753	human gut metagenome genome assembly, contig: NODE_877_length_3110 6_cov_4.049725, whole genome shotgun sequence GENOME_ACESSION: OWGA01.1 GENOME_ID: 11972 CONTIG_ID: 876 SOURCE: WGS DATE: 2019-01-27.	atgagcacttgcgtttgtgttctcagtaatagtgatgaacgcctgatgccaacatccgtctt ggcaagggtgcggccactctcgaagacgggaaaaggccaaattgttaagcaccatccgtt accatccagctgtttatgacagcgaacgaatgttaacccatcgaaatctgtgaggac gtcggctacaactacattggaatcagcgtgaaaagtaaatctcatgaatatgtgtctgctca gtatgatacattgcaagacgagaaagaacatcacgatgactccgtatgcacgcgctac acgcagaaacagattgcgctaccgtaagcggcgcttcgataaccgaaagcgacaaa gggtggcgttgcactctctagaaacacagaagacgtgaatatcagctttatcgaacagta tgtatctgaattccgattactacgcaacgggtgaggtgattcctttgacacgatgctgctg caagctatccagaaggcggaagcgaaaccggaaggcgtagactaccagaagggtccg cggtataactagctacactcgtgaggcgggtattctaccgtgatgattacacctgccaggt tgtggcgtaaaattacggatgggtccatctgcacatgcaccacatgttctactggaag gcagacatggcaatagtctcagcgaagcttataacagtagtcgagaagtgccatacaca gctaaccatcaaaaaggcggaagctctacggattgggtgaagatataaagttcgccaat ctttctgggtcggcattcatgaacaccgtgcgctggcaaatcgttaattgactttacgtgct tttgaaagtcgttgcacattcatttgggtgcatgacaaaagaaaggcgattgctct tcacttgaaaagagtcataaacacgatgcgtatgcaatggcgactttcatccagttaac gctgcgctgttgacattatgaaaagggtgaacgcaataaccgattctcgaaggtttat gactcgcagtaattgacactgcactgggtgaaaaggcaagtggaagcgtctctttaac ggaaggattaacgtatcaataaaaggattctgaaaacctgcacaagtatcgcggaag cggttatataaaggacatcgcccttagtgcgcaagaaagtgaacctcaatcgtgtgatt	gtcaacaaccacgcctaatacggtcagtcg gttatagacgggacttgcgggaaacccgt aagctgtgttgtagcctaagtcgtcgtc cggtagcaggaaactacgttgtgtactaata atataggcacccttactatgctccacaagtg taagctctgcggacggctcgttaaacatctc gaggttaggagaagtgcaacgtcatgctc gagaggctaaaacagtacaacaacattggc gatgtggaccacaggcgcaagtcctgact tatagtattactattttacgaaaggagtatc ttgcatgagcacttgc

SEQ ID NO:	Name	IscB sequence	ωRNAs
		<p>tggtttctttaacggcgaagtcctgtgttcacatggcactcataccaataaaaagggtcgtgt aaatgtagaattcaaatgaccgcaaaaacggtaaaaatccgcgagccttaaaaaact caaaattgtaaaagcagcagatccatacatcccgatgggaaaaagtatcttaa</p>	
754-755	human gut metagenome genome assembly, contig: NODE_939_length_1954_1_cov_1.798317, whole genome shotgun sequence GENOME_ACESSION: OZ0R01.1 GENOME_ID: 8944 CONTIG_ID: 938 SOURCE: WGS DATE: 2019-01-27.	<p>atgtggctgtgcccagggtatggaacacacacaggacatgtccctggcttacagcaaa ggagatacaatgaattatgtttatgtactgacagcgggggaagcccgtgatgccgacc agacgggtacgggtgggtacgcaggacattaaagtcgggcaaggcgaagccgtacgta ctttaccgtttaccattagactgatgtatgacccggatatacaaaagatacagagccttacc ctggggatgacccgggaagacacacatcggcatggcaactgccgatgaaacaggca gatgtctgtattctccagtcggaacacagaacccgggtgatcccgagactcatggaaa aacgccggcagcaccacagggcatccagaaggaggagaacgccttcggaggaaacgg ctggccagaaagcttgaaccactatgaggatcctgggaacggatgctccctggctg cgaaaaaccagttagggtgaaggacatcatcaatcacaaatccaggttcaacaaccgac gccgtcgggaagctggctgacaccgacggcaaccagctcttaggactcatctgaac ctgtgtgaaaaggtgtgcccgatccttccatcagcgggattgcactggaaacaaaccg gtttgcatattgagctggaagcaggcggccatctggaatccggagtggattaccagcg cggcccgtttacggatctgcagttatcgcgaagccctggaaagacttcaggacggaa gatgtctctgtgtgcccgaacgcgccattgaacatgacatcatctgtgtcccgatcaaa gggtggaaagcagactatagccaatattggctgtgtgtgaacattgccatactgtga cataccagcagactgctgcggagaagctggaacacatcaaggcagggcagaacaaa acgtatggagtcttatccgtactgaaccagatcattccgtatctgtagaggctctgtctaa gaaattcaatgggaacatccgtctgtatccggttggagacaaaagcagttccgggacg aaaactatatacgataaggaccatggcattgatccctactgcattgcagtaaatgggcaaca tcccaagaaaattgatgtcccggaaacccattttcagatccggcagttccggaggcatga cagggcacgaattaaatcccagaccgaacggacttaccggctcgatggagagaaagta gccataaacccgggaaacagttggaacagaaaacggattctctggaagactggtatca aaaaatggcagcatattacggccgccaacaggcagaccggatgcgttccaagctgcag gtacaagaggtaccaggcgtataacacccaaacccgctgctcggggaacagtttt tgtatatcagggaaggcatacatcatgaccggctagctgacaggaggacagtatttcag ggccgcagagatgtgacaggaaaaacttttctgccagaagtgcccaaatctgtagatataa tcagggactggtatatgtataa</p>	<p>gtcaactaccatccgctaagcgaatggg ctgttaaaagcccagggactttacaggtgag cagttgagcagaaccgtgacgcagcaggtt cgccctatcactcgggatgacgccaagttc cggacactgagatgtggtgtgtcccaggtt atggcaacacacagggacatgttccctgg cttacagcaaaagagatacaatgaattat</p>
756-757	mouse gut metagenome genome assembly, contig: NODE_281_length_4622_5_cov_26.608620, whole genome shotgun sequence GENOME_ACESSION: OFHM01.1 GENOME_ID: 3964 CONTIG_ID: 280 SOURCE: WGS DATE: 2019-01-27.	<p>atgtcatatattgctttgttatagcaaatgataacagtagacttatgcctacaagaaccccta aaaaagtaagaaggctactaaaagaaaagaagctattatcttattgaaccttttaccata taagcttttatataaaagcgaaaaatgcacacaaagacatagagctatgtgtagatatagg ctataatcatcaaggaatgtctataaaatctagaaaaacagaatttgttccgaagaagga catttttattagatgaaaaagataaacatcaagaacaattaaataaaggcgtgctagaaga aacagactcgttacggaagcctcgtttgataatagagctatccctaaaggatggatagc accatcacttagacataaaaaaagcaggaattaatgtaataacttaataattgtgaagtttt acctattacaagataactttagaagtaggttaactttgatactcatgctattaaaaaatccct gaaaatggagaattttagaaggattgactatcaacatggagatacttatggctatgattcc ctaaagagaagctatttttctagagataattacacatgtgccatttgaataaagggaataaaa gataatgtaatacttagaatgccatataaattattataataaagataggctaatagacca ggtaactttttaacagtttgactaattgtcacaattctaagaacctggcatagatggcgtt tatggggcttaaaacctaaaacagattctttaagagatgcagcttatgaatatgtttaaata ttcattaaagaagacttagaaaaagtttaagaataaatttaagaacacctattatatt acttatggagctgttactaaaagacaagggttaaaftaataatatacctaaactcatgcaaat gatgcttattgtatgggaagcttttagacctaaatataaaaataaaacagatactttaaaaa gataagacgaataaacgaaaatttagaataattctgacttaaaagttcttgatacaagaa caaatgaatatatgaaagaaatttagagggaccccaaaaaacaaaggtaaaagagta aacaactccttaaacgaaaatttatagaggaccccaaaaaacaaaggtaaaagagta taagaaaactcactataataataatcaggagatatagtttaataaaaaggaatgaaagg acattatctttgaaaggtattataaatttaggttaaacagctaaaattatttagataacaaa tacatatatccatctacctaatttagttataaataatgtaattggatggatagaacaat aa</p>	<p>taagaaacagtagattcttatattataaatta tatctactacggttatattagaataacaaattat tttaattttgtaattagattaccaatagatactt cacaagcttattgctctaaggatgtattataa aaacttcaggtaagggaagcgtgactatata ataaaacctaatataacattggcgttgtgaac cacactcgaagagtagatactttaataaag acttatcttttaagaataccgtaaggaggt aaataaaattgcatatattgct</p>
758-759	Ga0256404_1004876 GENOME_ACESSION: RumenRJG_01DNA_2_5_F_3300026549 GENOME_ID: 34936 CONTIG_ID: 4875 SOURCE: JGI DATE: 2019-01-27.	<p>atgtctgtgacagttatcagtaagaccggcatccgtctgatccgacaagtgaatatcggg cgccggcatctgcttaaatcaggtaaagctgtcattgaacagtagcccgatcttcacaat ccgttttaacagaagaagaggacggcgatacacagccgatcgatgacagtgatactg gctaccagcatgtcgggtatccatcaaatcagagaacatgagttgtacatgctcagtat gacatgctgagcaatgagaagaacggcatgacaagtccgcaagtaccgcagagaa agaagaacagactcgggtacagaagccgcggttcagaaacccgaagccgaagatg acaaaaggcgaagtactgcccctctctgcagcatagagttgataaccagacaattgtgt tcgactcgttctcgaagtaattgccaaatcacagtgccacatttgagatgggcaagttcga caccagttgtgacaggcaatggcagatggcagctgctcccccaagggaaggactacc agcatggcagtaagtatctgtatcagacagcgcgtgcagctgtatttggcagagacaact acagtcgccagatttgcggcgttccatcaagaatggtgccattctgcataccatcatatt gggttctggaagtcataccgttccaaccgtctcggcaatctctcactgtctcggagaagt</p>	<p>gtcaactaccacggcctaagccgggtgg gcttgagtggaagactttaaagctcgttga ttagcctaagtccttgagcactagttgccg ctagctacgtgggatgctccacaagtcaca gccccttggatctgtgttaaaacctctaa gggtaggagcgtgagcagatgtattcga ccagcggacaacattggcagtggtgactac cgttactgcgactcgtgtgatcgtaagtcga gatgccgcaaggtctgtatgactcggaaag tgtaaactcttctccgtaaggagacctgaa ggcagcaatgctgaccggaacgcgaatt gaaaggaggcatcagcatatgtct</p>

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gtcacacggcaagaatcacaagcctgtgggaaagctctggggcatagaaccaaaagt aagcaatcttccgaggcaacgttcatgtccgcagctacgtggcgatgtattacgcact gggtctggcacatccagaggtcgatccacattcagttatgtgcaaaaacagcggcagt tcgcaaaagacggcatattgccaaagacatcgaatgatcttactgcatcggcagttc catccgaggcacccgtcggaagaagtgtttgtctaaaagaagacggaacaaccgtgtt ctgacaaagtctatgatgccaaatatacgcacatccgtgacggcaagaagaagctgtgtg ctcagttgcatgtgacaggactaatcgtcgtgaaagcagacacacggaaaagaacgaa cgtatctatcgtggacaaaaggtttccaaaggcaagacatctatgaggaaacagcactatt tgtttcagccaagcgacattgtgtttggcaaaatgccaaatattcagtttaagggcacacat tgcaattggcacaagagtgtttggaaacagggaatcggtaagttggctgatttgaga attattaaaagacaaggaggctatgtggctcctcccaccgcctaa	
760-761	Ga0164242_10000399 GENOME_ACESSION: OricomgcompostMG_2_ _F_3300012942 GENOME_ID: 33846 CONTIG_ID: 398 SOURCE: JGI DATE: 2019-01-27.	atgccgtgtcgggagaaacgagcacagctcttctggagcgcgggagagcccgtgtgc atcgctgtagaaccatttaccgattagtgatgttcgttgaagactcggagcttcag ccgggtggagctaaagctggatcccgggtcgaaggcctacccggatggcactcgtcgcg atgatcatgggtcaagcattgcttgaatctaccagctcgatccgcggacatctgat ccatcgcaagctcctgcacgagcagcattccgacggaatcggcgtatcataaaacgc gatatcgccctgcggttccataaccgtagcgggtcgaaggctggcttccgccagctc tgacgcatcgcgtggatgataccctgattgtggcgctgaatttcaacgctcactccgct cacgaagctgtgtcgaaagtatcgtctcgtacccagctgatggacaatccggacat tcaaggacgcgactaccaacgcggcagctctgtggagctgtgagctgcgcgaatcgctat tcgcaaatggggctataccctcgtctactcgggggtgtcggcatttgatggcgatggc tcattatggaattggatcatttctggccaaagtcgaagggtggcagcgatagcccccgca atcgctgcgggctcgtcctgatgtaacggcgaaagagcaatactctccgctgat ttctcattgatgaaccggagaagcgagcttggatagaggccggaactcaagactccgctca aggatgctgcggccatgaatgccacgcgtacaaagctcgtggaagccattgagcgactg ggcctaccggcgaacgagtagcgggtgtgtcgtacgcgtggaatcgcaacgattcg atgtccgaaaactcatcgttggacgcgtctcgcgggcaatgtcaacggagtcgaag gactggaaggggaaaccgacacaggtgattacgtgcatgggtcgcggtagggtactctc aactgccaacgataagcacgggttcccgccgagataccttgcgaggcataagcgtcattt ggcttcgctactcgggacttgggtcgcatttcaaacctcttaagaaatccaaagtagag ctggagcagcactcaactccatataccgcatcactgtgtcgtcgaagagacttccgct cttcatcagtggtctaaagtactgcgttactatcaaggtgtagagtggtgcaacgtatg gcgggtatcattttccaaaataatgaaaggtgttatgtga	gtcgtatcccattcctcatgcaaaaggagaac ctagtggcaggtttccgaacctagtttcatca ggaaaccagactcagttgtgacctgtgtcaca ctacgataccttagaatacatagggactccg ggatgcctcctcagtcgggactctgcggct tactgttaaacagctgtatgggttaaggcag tgcatgaagcgtataaaacctaaagatacct tgtcagagaggcgtcggctccctgagaccgg cgtcacaatgaggccccgtaaggcgatcgg agaaatcgtgagt
762-763	human oral metagenome genome assembly, contig: NODE_381_length_3267 8_cov_10.933728, whole genome shotgun sequence GENOME_ACESSION: UPJM01.1 GENOME_ID: 16567 CONTIG_ID: 380 SOURCE: WGS DATE: 2019-01-27.	atgttggtatgtattataaaacaaatggacaaacttttattgccaacggcacgcttggcaa gggtcgtgattattataaaacaaagaaagggttctgtgtagagccgttcaatc aaactctttatgaaacagaaacagatgtgggttcgaagagtgtattgtgtgtgatacagg ctcacgacatattgtgtgacagttgtcagtaatgacaaaatgttataccaatctcaaacg aattaaagagtgacataaaacgaaatattgatgtcgtgaaatgcatagcgtatcgcga gaagtgaaaaaacagcgtatagaagaagcctgatttctaactcgaagtaattctacaaaaag taatagctgtccaccttcagtcacacacagggctcagctcattatgacgaaattgagttct gtaaaaagatacttccagtttctgatctgatttggaaataagtcagttcgatacagcttfaat gaagaatccaagctgtatcagtgagaaggttaaaacattgggttatcaacaaggtttcaat tatgtttattcatccagacgagaagcagttcttcagatcactacacttgcgaatgtgt ggcaaaaaagattgtagactagaagtgcatcagcttcaactcaaaagcaacgggtgtaca gatgatgaagagaatctcattacattgtgtgaagattgttcataa	gtgaagctatgctcagctaaagactgacct cttctgttagtattagtgttcttcacaagt cttaggttagagaaatctaccactgttgaatt gaagaaccataaggttactacacacgggtg tcgcttagaccgtgtgtctgagttctatcatt aagttggaaggatagcattgagttctgtgtg tatagattaaaaattcaattcaacttgcagaa aggaagcccaattcttaacttggtaatagag ctaagataggcactacagctctgataagtgct gtcttatagagtaaaacttatcatcaacgaa ggagcgaagcgtaatgttg
764-765	Ga0318466_10005777 GENOME_ACESSION: GoaFecpsamplesv2_F_3 300031555 GENOME_ID: 280328 CONTIG_ID: 5776 SOURCE: JGI DATE: 2019-06-13.	atggtatagctgataagcaaacggacatcctctcatgcctactgtaaggcatggcaag gttcgtcgttactgaaagaaacaaagcaattgtatgataaaagatgtccgttcacatcaa actaacttgcacacgcctgacatgttcaggaaatcagctcgcgtagatgccggatc aaaacatacggctgtgtcagctacaaccgaaacgaaagaattgtacgctcggaggtcga gcttcgcatgacataagcaagctcctcgaagcaagcgtgaaataccgacgtcgagaa gaaaccgcaagacaaggtatcgcaaaagctcgttcgaaacagagtaaaagcaaaagcac aaaggatggcttcaccgtcagtcgagggcgaagatcggaaacgcagctccgattatcga gaattgtatgcaacatacttccgacaaagatacactgttgaagtcgcatcgttcgacacgc agaaaatgcaggatcccgagataacaggcattgagtagcagcagggcagctcatggg ttatacgtatcgcgaataccttgcggagaagttcgaccacaatgctgctactcgcgcaag cctcaaggtaacgggtgccgctcgaagtcgaacatttactccaaaatcgagagcggt atcgaaccgcatcagaacctcggctgtcatgccaggtgcaacgaagcaaaaaggt aatctcacctgtgaggaattcggatccggaagtagcgaaaaaagctgaagccggcat gaaacatgccgctcatagcatcatcgtcgtgacgtgtacgaacgcctcaaaagcca tctacggcgaacggcttcatctgacttaccgagcagacttatactctgacatgaggca ggactcgaagagagccacataaacgatcgcgctgtatttcaggccatttctgtcgcgag cctgccgaagaatacttaccggaagaaagtgccgccacaaccgtcagataccaa gctgactattcagaagggcggatccgaaagcgcacaggcacccttacgaagtttaag ggattcaggttttgcagaaggttaaaagcgaagcagcatgtgtgacatccacggcaga agagtaaaagctcattctactcaaaacacttgaaggtgaaagcttgaatcgacca	gtcaaccaccaccgcttaggctaaccgctt gaagcgggggctgtgaaagcccgggtga ctagcctaagtcttagctgactacgttacgag agaatataagtgacctcggagtgtaatacct agctcctcgaataacggatgtgtattaaaca gttctgacgggttaggaacagattgcatacg caaaacctctcgaatacattgcgaaggtat cactacagtcctgaagtactggcttaaccgcat aaaacttcaaaaaattcagaagagctgtg ccttatg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		agtaagataacgcataagggcatcaagtgcttatataacagaaaggagaacggcgctcctacccgcccttag	
766-767	Ga0118733_100054452 GENOME_ACESSION: Marsedof8samples_F_3 300010430 GENOME_ID: 32450 CONTIG_ID: 54451 SOURCE: JGI DATE: 2019-01-27.	atgcctttagaccagctaagggcagacatttgcctaaaggctgtaagctaagggtcttaacagattaccctttaccattcagctattacaagctactgggtaaacctaacaggaaactaattttagggttagatcctgggtctaaacactaggaacagcgggttaggttgattaaaactactaagaattttacgcttcaatgtaacttttagatctgattatgaagaaactaaacaaagaagcagttacagacgaactgaaggagtagaaaattagatccagaagatgtaattttatgggatgtgtaaacctgttaattgaaatttacttgggttaggcttaacagaattaaattttaaataataaaggcaaatgtgaattagaagaaagaaataagaaaatgggttttagtttggtagaaaacaggggtgttacctcctcgggtcaatcgaagtagattctattatgaagaaatttaattatatttagcatattacctgtaaccacgtaactatgaatttgcagcttttgatatacctaagcttaaatacctgatgttaagggaattgaataccaaaagggtgatgtatggctatgaaataacta aaagctatatattacaagagataattcaaatgtcaagttgttaaagggaactaaaggat aaagttttacaggttcatcacataattcatagaaaacatggaggtacggataaacagcta atttaattactttatgctctatttgcctatgataaattgcataaagggaattaaaactttaaact aaaaagaaatttaataactatagatgctactcaagttatctattatatacaaaaagaattagaa gttatttatttaataagaaaaaataaatactaaagttttatagaacctatggttattatfaa aaagtttaaaaggaaattacttaatttaccataagatcactatttagatggaaattattgtactta tcctaagagggaataaatttcaacagatcaaaacctaaagattttaaatttctataaaaaagt atcagttcctaaaggcgattacaacaaactaaaggcagtcataccaaatctctatgcct acagggaaaattcatggttttaggaaatttgatactgttaggtatttgggtaaaaatttattcat cagaggtagaatgagtagtggctatgctaatgaatgaattgaacaaaaggtaattaaaga ttaggccaatgcctaaagttgaaaaaattaaaaaataaatgctgggaaaaacaataattgtc gattcatccctcccttcgtttcactcaggaagggtcttctgcacaggaaggctaa	gtgaacaacccccaccttcgcctaacggctc agggggggctcaccagattaaagctaaag aaatttaagctacatagtttagccagcgag gtaaaaggaaaccgcctcagttcctatatact tctaaagctaaactgtattgtcagaggaagt cttaaaaattagctgtaacagggcattgagt agagcattacctaatttattaggagagaaca aaacttgaga
768-769	Ktedonobacter sp. 13_1_20CM_4_53_11 13_1_20cm_4_scaffold_5 68, whole genome shotgun sequence GENOME_ACESSION: GCA_001919995.1_AS M191999v1_genomic GENOME_ID: 114122 CONTIG_ID: 157 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	atgtcaaaagtgtgtcctcggatgatcaacaacagcccttagaccagtgcatccgggc cgagcacgcctgtgtcctcaagaggggcaaggccgctgtgatcgcaggtaccctttac cctcatcctgaagacgcagtggtgactccctcgggtgtcgtcgcctcggctcaaaactgat ccaggggcggaagacgagcgggctggccctggctgtagcagcgcagcggggaagtgtgt gttggcgccagaaactcggccatcggggagccagatcaagaagcgcataagatgcccg gctgtggcgtgcggcgcaagagacgtctacgttttcacgcggtaccgcaagccgcgttc cacaaccgcaagagttccaggagaaaaggcggtctccgctcgcgtggaagccgg gtttccaacatgctgacctggtagggcgctgaggcggtttgtgccccattgaggtgatc agcatggaactgttaagttcgtatgtagcagccatgcaaaaccggagatcactggcgc ccagttaccaacaagggaacgcagtggtgatgagacgcgggaatcctgctcgaag tggggcgccgatgcgctactcggggcgtaggatgttccctggagattgagcacat cctcgcgcgcacgagggcgacgcagtcgctgagcaacctcaccctggcctgcgag ccgtgtaacgtcaagaagggcacgcaactcatcaggatttctgaacaagaagccaga cgtgttgacacgcatcctggccaggcgaaacggcgctcaaaagctcgtgccgctc aatgacacccgctggcacctcttgaacggctgaaggcgaccggcttaccctggagac cagcagttggcggtttgaccaatacaacggggcgaagcgcatctgcgaaacgcata tggctcgtatgcagcgtgtgtcggccagagacccccacagcctctggaacgagccagg tggttcctttgctgacgaagccacggctcatggcaaccggcaaatgtgcggcgctgatg agcatggtttccatccgcatagggcagcgggaagaagggtgcatcttggctaccagacg ggtgacctgtgtcggcggtgtgtccacggggcgagggcgggagcgcagctgttgg cggtgtgtgcagcagccagcggtcttctgatttgaggaccaaaagctggagccagg ccgggatcagttatcggtattgccgtccatccatcgaacgatggctatctgtaccagca aggagggcggtcagtcagtcggccccaacccaatccacctga	gtcatgaacccctgcatgaatgcggggggc ttgcagcaaggctcatgctcgtacaagc ccggaaatgaccagactcaggttgaac agagcctacgttaggagcgaatccataggc acgtccgggtgcgacgccagccggaccg ctgcggcaatcgattaaacagggttagcggtttaccctagtgctgttcgctcaaacgcctc cataacctgtcaggcgcaatcaccggg cgcttccggagggtcgcaagagcaaga ggaccccatcatgtaaaaa
770-771	Ga0272449_1005167 GENOME_ACESSION: YNPB0241_2_F_3300 029977 GENOME_ID: 281651 CONTIG_ID: 5166 SOURCE: JGI DATE: 2019-06-13.	atgggtgacgttatttcaaaagatgtgtaaacacttatgccaaccaaagacatggcaaag tgaagagactctaaaacaaggctctgtcctaaagttgttagaagagaaccatttacaatccag ctgttatatgatactacaacctatacacaaactgtcacagttggtatagacattgggtcaaaa acagctgtgtatttcagcaataacggaaggaagcaagcaatattagtgctgaaattgaactca gacaagatattaaagattatttctggaagaaggagcatagaaattcagaagataacc gcaagagaagatagagaaccgaagatttttaaacagcgctagacatgagaattggcttg ctcaaaggtgataatgaaagttgatgcacatgtcagactgttaacttttagccaaaata ctaccagttacaaaagttgtacttgaagtagcaccatttgatcacacaaaatactgaacc tgaagtagggagcgaataaccaggaaggaccacaaaagggttttggagggttaga gaattatgctgtggcggtgcaggatataaactgaactgtcaggcgaagaaagggtgtgtgg aagtacatcatattgttccagaagtcagggtggaacggataatccctctaactgtatagtg ttaactgtgagggaacacaaagcaatacatagaggggaaaattgaattccacgcagtaga ctgaaaaagttaaaatttcaaaagacgcaagttgtgtatccacaataggatggcatatagt gaacaaattaaaagaacaatatgatttcatataacatacggtagtatttcaaaaagcaaaa agaatagaatgggattggagaagacgcatagaacgatgcgtttataattgcagggtggt agtagagatttaacagacatctgagtggtattttgtaagttttcagaagacagaaccg ctctgttacaaaagcaacccaattaaagggtgtaaaagaccagtaaatctgtcaagga agtacatggttttagaagatttgacaaagttgagtagcaaggttagaacgggaatttttag	gtcaactaccacccacctgtagaggtgtgtg gcttgtgaaagccatagttgactaccctcag ccattggggaagggttaacttccatcgggcta agtttagactgtcatgacccctgggatgct gctcaagtcacggctctgtcgtcgcattta aacagtcctgagtgtagggggcagtgtagac agacatgagaagccagcttaacattgggga taggcacctaaccttaacaaaggaggctt accgcatatg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gattaaggagtagtgatatttgaataggttctctgactggggagaagatatgtgatagtgtaaagcacagcaaaacttaagcttttagaaaaagcaagacattatgttcaaaaggagggaagagcgcatctctcctctctggaaaaagatggagctcctgcgctcaattatga	
772-773	Human oral metagenome isolate term3_saliva term3_saliva scaffold_55 62, whole genome shotgun sequence GENOME_ACESSION: QWCI01.1 GENOME_ID: 14517 CONTIG_ID: 5557 SOURCE: WGS DATE: 2019-01-27.	atgttagtatatgtgttagataaaaaacggacaacccttatgcctacacataatgtgtcctaaa gttcgagttttataaaaaaaatagagcaaaagtagtttcaaaagtgccgtttacaataaaa ttattatagaaagtagacaacattttacacaaccgcttactttaggaggtgatacaggttctaagt atgtaggtagtgcagttatcaatgatgtaacagcagaagttgtttacgaaagtaattagag ttaagagatgacatcaagcttaaaatggatagacgttagagcggtttagaagaagtagaaga aacaaattacgttatcgttcgaagaggtttaataaccgttaaatcttcaatatagagactgtt acacaccaactcttatccaaagttacaaggtcatacaagagagataaaaatttattaactcta tactccaattagtgatatttttagaagtgaggaggtttgacacacatctattacaagaccc aacgttagcatatcgtaaagtggtttatgcacaggagaaatfatccaacaagagaacttt aaacaggcaacaaaaagcaagagacgactacaagtgctcaatgtctgtgtaagaagaact gttagattagaagtcctacaccttttactagaggttggaagtgataaggttagtgtaaac ctaattactctatgttccaactgtcatcctgtgacacagtgctagaagaacagttattagcat ttacagaagagatttgtaaaaaggcaaggatattgtgaagtatgctactcaatgaatata ttaagacacgtgttgcaacgagaatctctgatgctaaatgacctatggtttattacgaag gaaatgtgtagagtggtttgggttagaaaaatctataatagatgcttctgtatcgtatg agaggagttttgttaaaaatagcgactcaataagataagaagaatgtgttgcaaaaag gtgattatgcctatcttagacacagtagctagaagaatga	gtgactgtttgacaagtatgattggagatggtgataccgagcactagttttactagtaat caccatacagactaagttactagaaatgaat actacgttttaaggatcatgacacctccagat gccacctcagttgggagcaactgtcgtatg tattaaagttggattgggtatgtaagtcctgt gtacatagtgtaaaaaggctttataacattgtc gagagggaagtcggtttttgataggttaaca gtatcaaaaatcgcattaccacctgtaaaaa ggtgagttagtttaatttttttaaaagaag gagtgtcattatgta
774-775	Lactobacillus equi DPC 6820 NODE_4, whole genome shotgun sequence GENOME_ACESSION: GCA_000504525.1 Lactobacillus equi_DPC_6820_genomic GENOME_ID: 55485 CONTIG_ID: 157 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgcaaaatcgggtttttgcatcaataaacatggtgaagctttatgccttgaagccaaga actgctcgtaaattgttagcccaaggtaaggccaacccattaaagaagaaccattacc attcaattttgtatgtgtgcaagtggtataagcaacctatttaacttaggtgttgatagcgggc aacgtcacattggacttgcgtttacaagtgcaagataaaatgtttcttcaaaagtgaaagtcgaat tgcgccaagacgttaagaaatgctgtatcacagacggaactatcgcagaagtagacgaa atcgcaaaactcgtatcgtaaagcagcatttttaaatcgcataaaaaatagaaaaataat tggttgccaccgtcgtgtacaaagcaaggttaacacacatttaattggaattaaacgcattgtt agccgtattgccgaaccagaattacacatcgaagttggtaaaattgatagcaaaagatg aaagatccaccattaaagtggaaggttatcaaaaagtggtattgtacggctatcaaacgg ttaacaatatgttttagctcgtgaccaatataagtgtaacgtcgaagaaaaagggtggc aaactaaaaatcatcacatttattacggttccttaggtggcactaatgtaccatcgaattttt aacagtcgtgtcggattgtatatacgaataacacgcgtgaagcgcgcaaatattatgaa ttatgtaagaaggaaaaagaaagtaactaaagcctttgaaggcgcaacctttatgaattctt cgtagagagattatgggacgctttcctaatcccaattttagatggtgcaaaaccattt acaacggcgcaaaccttaggcttagcgaatcgcaccacaacgatgccattacgattagtggt tatcaacttattagcgaagaaacgaacgagtggtgtgtgtgttgcaatttcgaagaaaa aacgttactacacgaggcaacagctcgttaaaaggacgcaaaaacagaatatcaccagt aaacgaaatgccaaaacactaaagaaacgaacggaatttggcttaattgattatgcagaa ttaagaatagtaaaactaaaggctacattatgttgcttattgtccaaaggttctaattgtaaggt tgcgagatagtcataatcattatgttagcgtcacagataaaaaattacacatctataaaaaat tattttaattcatatacaacaattggaataagacagtaatacgcgagaacaatatgttat gtttag	tttaagtctgctcactttgatgagattctgga gcaatccagaattttgtctagactaagcag tcagcttcggctgactgggaaactacgttatt ttatgcatcatccttaggatgatgccctagtt ctaagctctatggaggctctgtaaacatcgt aaggttaatgcgatagtaacacctgttaacaa gctaaaaatcattgtcgaagggttaaccaca tgaactgttagttctctgatagttgcgtgata aggtacatgctacttataaaaaatgaaagg agaaaaatcaaatgcaaaatcgg
776-777	Ga0335394_10066399 GENOME_ACESSION: MAT03spassembly_SF_3 300032456 GENOME_ID: 280657 CONTIG_ID: 66398 SOURCE: JGI DATE: 2019-06-13.	atgcaacgagtagttggtgctagacaagagcaaaacccgctgatggcgtgtcattcgggtg cgcgcacggcaattgctcaagcaagggaagcgcgcttctgtgaataaccctttcacc atcatcttgaagaaacgggtagtggcggtgttcaacctgtcagtatcaagatagaccgg ggagtaaaaccactgggatctatggtcgtgatttcaaaagggggcaagcgctgatt tgggctggagagtttaaccatcggggcgcaacaaatccgagataagttattgtcgcgcgt caggtattgggcttgcattggaaagtcgggacaggagacgcacgaaatttaacgggt aaaaacagcactatccaaaagcgcatgtgattgacgcgatttgttaggagaaggtgggg gacgagtcacatcagtcacaaatcgtcgcggttcagattagggcgaccgggcatcaat cagcccaattgtccgcgttgataagtaggttttccacgaaccagcgcaaaacgaagg gcgctcactttggcttccaacgggggatatagtcaaaagcaatgttcaaaaagggttaa aatagggcacataattgggaaggtcagtgtaggttctactgtttttcaatatcactacc ccaccgcaacagtgagggtatcagctatcgtactgcacaccgattcacaattccgacg gctatctttaccagagaggagagcgctttccgctaccacaagggtagcgggtcct cgcgccaagattcatga	gtcaacgaccactggcagagggcggtgg cttgcgataagcaagccattgttgaccagcc tcagctcctctgtgagacaccgttatttgggt cagcataccgacgggtgactgaaccagcc cgtcgtctatcgtcaccgttaaacagccga gggctagggcaagcgcggtggcgatgaga agcccgaaatcaccattggcgagggtcattta ctgccgcaagcgagaaaaatgaggttaact catgcaacga
778-779	Ga0115617_100462 GENOME_ACESSION: IMG_3300008733_SF_3 300008733 GENOME_ID: 28627 CONTIG_ID: 461 SOURCE: JGI DATE: 2019-01-27.	atgccctgttccacggcaaaagcgcggtctgtcttaaaagagaagaagctattgtgaag agggagaacgccttactattcagctgacgattgcaacgggtgagaccaaacagccgggt gggcttggcggtgatccgggtacaaatctgctgagcttccgcatcaacggacaagg ctgaactttatgcatccaagtgcgaactgcgacaggacatcaccgatctgctcgcgtcg tctgtcgtttacgacgggctcgcagaacccgcaaaacgcgtacccgcgcgccaagattc aacaaccgcatccgacaaagcgcaaggctggttgccttgcgtcgggtgaaaacggaat caacgcgcactgtcgcgcatagaagcggttctcactgcttccagtcacgaagatcac cgtggaaacggcgcttctgacatgctgctggaagatccagacattgcagggaag agtaccagaaggggcgaacaactcggcttctggaacgtcccgagtgctgtcttttagag	gtcaactacctcggcctaaggccggggct tgaatagcctctagtgtgactagcctcaggg cgtattttgcgggactacgttgggtgagaata tataggcaccggggatgtagatactagtcc cgcgctctcggcgcatggtttaaagctcc gagaggtaggagcggtgtgtcgtgctacga aaccccttcaacattggcggaaggtacaaa ccggtcgaaagaccgagagacaaaactt gaga

SEQ ID NO:	Name	IscB sequence	ωRNAs
		acgggcacgtttgtcagcattgttacggcagatcgaaagaccgggtcctaattgtcatca tctggaaagcagacgcacggcgaggattcggccggcaacctgattacgctttgtgaga cgtgccacgaagcgttcacgcggcgaaatcgctgaagacaagcgcggaactc gtccgtgcggaaagcgttcattgggaatgatcgctgggagggtttgaaccgctaagggc gtcgcaattcgtggttggaagtgaacaacacctacggctaccgactaagcacgcacgga tcgtgaacggcatcgacaagtcgcaattgtgcggatgcttttgcattccggcaacctcgg cgccgaaaggctatgcgaattcttccagaagcagacgcgtcggaacaatcggcaga ttcaacagcttccattcacaaggcgcatcgaagcgcaatcaggctccctttgagat caaaggcttccgtttttgacaagggttcctgcaagggagaagaaggcttcatttccggcc gtcgatcaccgggtactttgatgttcgaaagcttgatgaacccgcaattcagccggcatc agctacaagaagctgcgtcgtggaagagacgaacctatttaacagaaattcgaaaa gagaggcgcttctccccccccccccggcgggggggggcgcgccgcccgc gcgggagaatcaaacagagaacgcttaagaaggatttccactgtcggatcgggttgc actccggcgcaaaagtcgtgactcgtcggcgcacctccgataccggtttgtgtt tactgaaccgatttgaagcccccgttgcgattaaggctgagcctgaacgcgtcaacga cacgcgatggcgacgacgttgataaagacggagcccgatttgcgacgatcgagcac ttgatgagtcgctgcccgtttggcgatcgacaactgctacatcgatgctgacacgtg agatcccgattatgacggctccggctctacgtttgtttgatcagagctgccggtatcc aggaacaggatgctccgcgaaattcgtccgtggaagaaaaacgttccatccacgtc ggcgacaagtggtggttccctgaaccttatgacggctacaactgtctttgcaattgacttc ggtcatccggcgattgacgaacacgctcagtttggtaagggtgacttcaacaaggaaaact atatcgaagcggttccacgcccgtacgttcggcttggtaacgacttagagatgctcgtg ggcattggcgttgcacaaaggcggaacattagacaacgcgattgtctcgtatgattccac gcttaaatccggcggtctcgtctcaggacgagtttgcacaaacacaagctcttgatg ccatggcgacgttgcgttttaggtcatccgctggtgacgactaccgtccctcaagtc cggtcacgaatcaacaacaaacttttgcggacgctgttagctgacccgaaactggga gttgttgtaataccagatgaacacagtgctccgaaggcttttaccgaagcggtaaaaga acctcgaactaa	
780-781	Ga0209607_1003300 GENOME_ACESSION: AD_des_42_SF_3300025 847 GENOME_ID: 21342 CONTIG_ID: 3299 SOURCE: JGI DATE: 2019-01-27.	atggtatagcgaatttcaaaagatggacagccgcttatgcctacaacagacatggcaagg ttgctgactattaaagtcacaaacgaacaaagtaattaatcgtctccatttacaataaaac ttatatacgaactacagatttcatgcaggaatttattgggagtagacactggttcgaagt atgttgggtttgacgtttactcaaatgacaaaatttatataatccagctagaacacagaga tgataattagctcaaatggacgataacgcgggtatcgtatgatttcgtagcgaagaaaa acccgctatcgaaacctgatttttaaacccgtaaacattcaactaagctcaaatcgaactcc acctcaactaaatctaaagtaaatgccaatatacaagaatttaataataaaatcaattct ccagtgactaaatttaatttagaggtagctcaattgatactcattgcttaaaaatcatatgct agcaaacgaaaaatcaaacattggggctatcagaatggatccttatatggtttgcacaaata tgaaggcttatgttttaacaagagataagtagcactgtgcagcactgtaaaactaaaaatggt accttgcatacgcaccataataattatcgtcctcaaaaggagggttcagatgaagaaacaaact gattacctctgcgagcttctcacaataaattacataaaggcgagtagcaagtttcgatt ctaaatagtaggtaaaagaaagcctaatttaccgctacgcgactcaaatgtcgatttgaga agtcaattgcttaaatattatgtcaacgcaattgaaacttatgctacataaccaaagaaaat cgacaaaatctaggtttgcttaagaccattatatacgtctgctgatttgcgtacgaagga aataaatttaagcctaacacgcaaatttttataaaaagttgtgttagagcgacccgtgta ctgctaaatttagaaaaatacaattcaactacgtagaggtaaaatctgtgttttagacaa tatgacaaagttaagtatttaggtaaaatctatttatacaaggacgaagatctgcaggtact gtgcttggatgataatttttaaacccgatttatttatacatgcttaaggctgttaaaaa ccaagctcgtcaattgcaaacgagtaagttctagaataatctgttatttgaaggaggat aggtgcgattcctgtactaagctaa	gtacaaatgtgctcggctaaagaccgaccac ctttggactaagctgaatttactagactaag cctagagaaatcttggctacgataattaggtc atgacacctttgtgtcgtccctagaccattg caactgtcgttaacgttaatttgaagaaata tagcactaaaatccggtgctgtaagtataaaa agcccttttatactgtcgaagggaagtcttatt ctttgggtgtaattagattaaagataagaact acagtttggtaagtactgattatcatttaaaa cttaactaattcgcgaaggagccatttctatg g
782-783	soil metagenome genome assembly, contig: NODE_51_length_30035 _cov_3.493647, whole genome shotgun sequence GENOME_ACESSION: OFHI01.1 GENOME_ID: 3960 CONTIG_ID: 50 SOURCE: WGS DATE: 2019-01-27.	atggtgaaactgaaccaggaagatcgctacaccgcagaacagaagggaacaggaaaa catggtacagggaagccaagattcgataaccgcaagaaggagaaggatggtttgcacc atcaatccggaataaactgtatgcacatcaaggttttaaaccaagtcagtgaggattctgc cagtatcggaacaatcattgaggtagatccttcgatacacagaaaatgcagaatccgg agatctatggaattgaataccagcagggtgaactgcagggtatctcgtgaaggaaatacc ttctggagaagttcgagagaaaatgttatatgccggaacaaacgttccgttggaaat agagcactatttctaaatcccgcctggctcgaacagggtttctaaccttgcatttctgt cataaatgcaaccaaggaagagatgactgcaaggagttcggacacccaatgt catgaagaatgccagaaaaatctgacacagacggcatttatgaacgctgctcaggtgga aactcaccaaactcactcagagccatcatcttttgatacatcaccacacacagagat caggctgggtcttcaagagtcacacaaacgatgccttcgtactgctgctggcactac ggcccatgcaacacctgacaattcggcagaatcattggtgaagcgggtgaagagggttc ctgatgtgagattgaagtgcactgccacaacgatccgggcttctgttccaatgcaatc gccgggattgaggcaggagcccacagggttgacaccactgtgtacggcctcaggag aggaacgggaatccgatcactcaccattcggaatccttgaaggattacacaggga cagcaggtggagcgaataaactgctctcagctatgactatgcttgaactgatccatg	gtcgggtgacgtgtcatcacagaccgaag atcgacatgcacccctcaatgaagatcaaga gacacctgaggtgtccacagaagtcagag gatacatcaggacagaggggaaggga gaagctgcagaaggtcagttgggacttctat gaccgctgcatcggtgattgacagtctca ggaatgagtgtagagatcgccaggatt tcggcaacctatcaggcttggcatgattcc ggttatctgcacatcgagatccgccgtctc ggagaaggagggaacttactct

SEQ ID NO:	Name	IscB sequence	ωRNAs
		agaagatgggcatgattctccagcacaactgtccccacccggcaggaacgtccag acgcacaccgtggtacgcacgccgattctctgatgtattccaggcgggacttctcc gtgaacgtataccggggaaggagcatgatcaggaaaatactctccgcaacaatatgga tgttggggaagacgaactcaggaagatgctcagatataaagaatgaggcgggtgaga caggcaggcactccacgttgatgacatcctgagaatagggtgaagtcatggcaag agtgattgagctgggacacatcattgcaggccctgtgcaggcctgatactgtctatctg gggcatgaggctataaagatcgagaacccgggacatcaggggacatccaggaggctga cgaagcagagtggcgtgattccctctacaaccggcaacaagaagagcgtgtgcat gacatcaactccagggaagggtcagaggcattccgaaggctcattggcacagcggatgt tataattgacaacctgggtccggggcgggtggagaggcaggcttccatttgagaaga ttccgcaatcaatcccgataatttacctgtcgtcaagggtatagggaaggccttat gagaaggagggaagccctggactatcccatagggtgcacagcggcctgcatatgatc aggcctcaagggaagccaatgagggtagacgcattcatagttgatattgtcggcgcta tgttcggcgttatcggcgttctaacgcgctcattgaacgagaggcaacaggcaggga aataccttgatgtcgggatgttgaacccctggccgtctttatggggcagcacgtggcaac cgcacagcttaagaatgtcagcatggaaccccttaataaagggtcgtcgtgggaat ctatgatttctcaggacagaggatgacgttgagggtattcattgtttacaaccgatccgc agtgaaggcattctcaggggcctttccatggaagtgtcggcaatggcgactacgaaa caaacgcagcgagatttgacagtagggacacgctgattcctcaatcagggaagatt ccgtaatgagttctccgatgttactggcattctgagggaactcaacataagctatgcactc ctgaacgcacccgggatcttctcaatgatcccgatgaaagacaagatgtaaccga aacataaattggcagaaccataagggtgcccacagaccctctcgggtcgatacagagat cagatcccgaggaaacttgggaacacacggaagaagtgaataactctcggaagaagaa tcgggacacagttga	
784-785	Ga0116227_10003227 GENOME_ACESSION: SITSphegellanicum_5_F _3300009709 GENOME_ID: 34953 CONTIG_ID: 3226 SOURCE: JGI DATE: 2019-01-27.	atgccttgcctccgacggggcgaaggagcttttaagaagtgaggcgcagtttta cagctgcagccttttacaatcgtcttgaaaaaagagaagggggggaagtacaagagat aggggttaaaatagatccgggaagtaaatgctcgtatcgcactgtcggaaattttcc ggggaaagcgcgggtgtgtggggagctaatctgagcgtaaaggcacggcaatcagat catcccttgcggcagaagcgttataagaagaaccgaagaagcagaagacacgcac cggaaaagaaggagacttccccctcttacaagaagagtcacgggtttgcaacgg gagattcggtaaaagcgggtgtcctaaagggttaacaaagggaacacactccggaga gtctccatccattatccggaagtgtttccatcgataccctgagggaagtagacggat cagctgtagattttgtaaaatctgcaacacgcagacgggtatcaatattcacaacaacaa gggagcgacgcttctcccgagctgaagtcgaggtttccgctcgtcagttgatga	gtcaactcctcctctctaaagagagggctt gtaagacaactaagcgttaagttgaccgggc taagttctgttaagtaacaaagctacgttag agaagagtagataagcaccttgaatgctg ctccgttccaaagctctgctgttaccggttaa cggatataagggttaagtcagtcggtaaac atacaaaacctctcataccattgccgagga gaacattaccgggaagggggattgca ggtactgtcactagtaaaaaagagaaa cgattgcaagg
786-787	Ga0393264_0001341 GENOME_ACESSION: LB_025metaG_FD_\$F_3 300036981 GENOME_ID: 312532 CONTIG_ID: 1340 SOURCE: JGI DATE: 2020-04-07.	atgcaacaaagctctcttctatttcaagagagacctgcattaaagttggacataggtggcgt attcgtgctaaaccacgattacacacactatgctacgtctctcaaaaatgcagtgagtttct cttaaaaaatggccacgcggccacctaccgtcattaccccgctgtaattcattctaaagaag caacaagctgacaatcgcttgcctgagcgttatagcgtggaactgaaattgaccaggg cggagacacagcggctcgcgttattgcccataaagtaacgaatcaatacacagcc ttatgggggtccaacctcaccacaaaggcaatatcgtctcgaagactagaatcccgtc tcagtattagaagaggccgacgcaatagacagacgcgtatagccgtccgtcaagc gaatcgtgctaattgcaagaaccgaaggtcggctaattgcttccagtcgggtcaattctaggt aacatcattacctgggcaaggagactgcaaaatagacacacacattcacagtattgcttacg aatgggtaaagtttgataccagaaaatggataaccagagataagcgggtggagat caacgtgtgagttagagggtacaggttaagagagtacttctgagaagtataagcgg aagtgtgtatattgcgaagtggataaacaagcctctgcaaatgaacacgtcataccgctg caaaggggcgtctgtatagagcctctaatttggcgatagcgtgcgaaaaatgcaatcaga aaaaaggaaacaagccattgaggagtgttctatcacacgaccagagcgtctcaaaaag gttaaagggtcaactgaaaaaaggactccaaccagccgctgtgcaaacgataagga acgcaaccataagctgttatgcgaattaccgggttaccagtgctgctatggtcaggtgg ccggacgaaaatgaatagagtgaaacaaggctatctcaaaagaccactggatagatgcca tctgcgtaggtgagactggggagcgtgtgacgatgccacctgacatgcaaccgttgctta tccggcgatggggagaggcaaccgacatttcaagagactgacagctacgggttccct agagtaaaagaggttgcaattgatgtaaggaggtaatggtagatgcaatgccaaagaa aaaagggaagcgggtgggttcaaacggcgtgcaaacggcgacatggtttgttagaca caaaactgggttagagtttaaatcgcgggttatcagttatgacccaaaaagtaagtcaggc ggcttaacagccaggcatctaccactggcaagaaaataagtgcccggtccgcccgt gcaagattttacagaagacggaggtgtggttaataaccaggaaactgaaggcg taa	gtcgtgaacccctccattcatcgggttctga gttttggttgaggctcaggactcttataaac acagcgaacacagctcagccagaaaaaa cgtcgctacgttattcgtgaaaatagtaa ccgtcggtgacttccagctcactgcttga gagatacttttaaacagctacgtgttataaatt tggtagtggcagtggaagtatataaaaaac cagcattaaacattgacaggagaccattgct gctatgcaaatgcaaacgaaaaataacatg caacaa
788-789	Ga0127503_10276577 GENOME_ACESSION: WCWIAAnnotation_\$F_33 00010154 GENOME_ID: 281607 CONTIG_ID:	atgggcaagcaaggagtggtgagcggagcagcatcagagccggccagtgcttcggct gctcgggacggagagcagcatcaaacagggtgacgggtgtgcccgcctgctatga agcagggaagcggcgcttaccggcgctacccgttaccgtgactcctcaaggggca gccggaacgattggcgcttcttggctatctcaaaagcggcgttaagatcggcgag cggtaacacgacgcgtgggcatgtatgaacggctcaccgcccgtggctcggcgt ggagggtggaagtggagggttgacgaatacaaccgggtgaggcgggaacatgtg	gcctgggcacacacgatcgtgcttgcctg gtcggccgggctttgtctcaggtgtaatacc gcttgcaagaatgagagcaaacgctatcc agtgcaagcaagtactcattgaccgcttct gatcgtctgagagaataatctctcaactat ccggactttgtggaagactgctgactcgtc

[illegible]

191

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ctcattacgctgtgtgaaacctgtcataaggcataccatcgcggtgaaatccaftaaaag caaaacgtggacaagctctgcgcgatgccgagtgatgccataatcggtgggagatt ataaccgtgctcaacgtgagtttacgaatgtgaattgacgtatggtatattacaaaacaca gccgtatcaataatactatgtccaagacacattgtgctgacgtttttgtattcaggaacgt ccaagctgaacgcctaccatggttctatggagttcggctgcaacgtcggcacaatagatct cttcataattaaagcctacgaagggtgtgttagaagaagtaattggcgctcattggatt ggaaaatcgcaattgcaacaatagactatgtgaatggaatgataaaggcatttatctct ggcagtcaaaatggagaccatacttaaaaactttgacgataagtatatcatgacacca aagcatccgtcaatgcaaaagacagtgaattcattcgaagaaagcgtgggaagtatgataa ttgaacaattaccataaaaattgtgtatag	
804-805	Ga0257072_1000039 GENOME_ACESSION: GoafecGen1Rep3v2_2_\$ F_3300023707 GENOME_ID: 280336 CONTIG_ID: 38 SOURCE: JGI DATE: 2019-06-13.	atgttcgtgtacatcattacaaaaacggacagccgctaagccctgaagcctcaaaa gcaagaagaagctcctgaaggctggcaagcgagggtgtgcaaatatgagccattcactatc aagctgaaattggctcagctggctacaacagccgataactttaggtattgacgccggct ccaaacataattgggcttccgtgtcaacagaaaacaggaattatattgcttggaaacagt catgctgttcatgatggcaaggctactatcgtgaattctatagcaaaacgcagagagctt cgacgttaaccgcagaaacgtgaagaccgttatcgcgaagcaagattcttaaacatgtt catcgaagcacaagggctgtgcttgcacagcgtagagaacaaaatccatgtccacct aaagctgtggcggtatgcacaaagatactgcccatacctaagtagtggggaagtgc ccagttcgacattcagaagattaagaatcctgatattccgggtcgaatatcaacaaggtg aacaacttgggtggcgaacgtcagggaaatgtattattccgcgacaatcacgaatgtca gtgttgcgaagcagatctggagatccgatactgaatgtgcatcattgagtcgcgaaa aactggcggaacgcaccaaaacatctgataactctctgcgaacattgccatcagagcta tcaccaaggcaaaattctttac2taagtcgataatcgcggcatgagttccgagatcgag ctttcatgggcatatcgcatgggcaattctacaatagaattaaggcgttatatcaagacgtc aaactaacttatgtttacatcacgaaaaatactcgatcaagaataacattgccaaagcac atacagctgacgcttactgtatcgcgggaacgtcaaaagctagacggcctaacaatgagt atttacgtaagcaagttcgagcgaatacgcgaactacatcgagaagtccagcgaag gtggatctgtcgtattggctcaagcagggcatttcgtaagaggcttttcttgaacgatact gtgatggcccaaaatcagcagtggtttatcgcagcgatcgtcaaaaggcgagttttgtct aagacatcttgatggcacaacacggaaattgcgcttcaaaaataacattttgagacata acaattcttattatcgagaggagagaaggtggcgcttaccctccaccctgtag	gtcaataacccccgcctagcggcttgcgct tgaggcgggagcttgcggaagcaagctct gattgactagcctaagtcttaactgactacgtt gcttctgttatctcaccgtgggtgtttatccta gcttagcgtctgagcaggctctgtaaaaat tctgagaggtaggaacggtcaacctgagga cggcttggcgttaaccagacaagcgattac aacattggcgaaggtagaattactctaatttg aagcttttgcctggaaaggaggtacacgtt atgttc
806-807	Ga0075017_100018482 GENOME_ACESSION: AleBraStandDraft \$F_33 00006059 GENOME_ID: 22526 CONTIG_ID: 18481 SOURCE: JGI DATE: 2019-01-27.	atggtaccagtatgtataaagaaaagtgtccattaatgcttctgtgaaaaagagcaag acttcttatgaaagaggcgaaggcgaaacgtattggcaacatggctcttttctattaggc ttactgtggagccttcgctcgtcattatcaggatgtcgtcttgaattgatccggatcaa aacgggagggatatactgtactgactgagacgcaagtgtgaatcaatattactacagatac tccggatcaagtaaaagtaaggttgaaacccaggaataatcttagacgaagtcgtcgtca aagaaaaactccttatcgaaaatgtagagcaaatatgtcttattgaaaaacaaaggcggtt ccaccatctacaaaagcaagatgggataagaagctgacaatgattcgttatctcagac atgcttcttattacaagataaattgtagaggatgtggcagctaaaaccaaacatgaaaaa ggaagatggaatagatcttttctccttgggaagttggcaaaagcatalttttatgaaaattgtt ctaaaacttatccagataatttcttgatgaagaccaaagggtttgaaacaaaagctcatcgag atattagggaattttgagaaatctaagaaaagttagagtggacatgggaagctcataatgt agatttcattctctagcggagatggtacttaataaacaagttaaacataatcgtggactttta agattgaagtgtgatgaatttttcaagaaggcaacttcacgtccaagtctctattaaggagg tattagaagccttatggcaaatctattacacaaggaataccaaaagggaatgattgtgaagt atattggcgaagcaagtcgtgtaatatgaattgttatcttggaggagaatctaaag gaaaagtgcacattgttaattataattggggaaacgcgcacatcccgatcaatc	ttcagctaccataatctagataagtagattatg tattgctcacatagaagtaattgttaaccaa acgttatgatgtgacgaaagtgaattgattag acaggcgataactgaaaagtatcgcaaaa acatatagtagaatcaaacccgttattggta gggttaagatgtagagaattgttcgctagt tcttctctctacgattctagtgtcgaagcg agattaacaagtcctgtgagggga
808-809	ERB_1001085 GENOME_ACESSION: IMG_3300003453 \$F_3 300003453 GENOME_ID: 308444 CONTIG_ID: 1084 SOURCE: JGI DATE: 2020-04-07.	atggtcttcttggacaagaaaaaacgaccgtgatgcttctgtgaaaaacgagcaa gattattgcttgaacgcggcagagcgcagagtagacacaaaatgtaccggttcaccatccgctt ggtagatcggttgaagaggattccgtattgcaagaagaacggctcaaacctgatccgg ttctaaaacaacaggtgtgctgtcacgttggatggagcaacaggcacaagagcagtttt ccttggcgaagtattcacaagccaggaatcaaaagcaaaactagacatcgacgcgag cttcggagatcaagcgcaatcggaaccccgctatcgcaagccgcgtttctcacaaccg caaacgcaagaaggatggttgcctccaagcctagaagccagagtcattcaaacgatg aatgccgttctaactgattgattcataccaatcaccgcccgttagttagtaaacacgtcaa atttgatacacagttgatgcaaacgcagagatttcgggtgttgatcatcagcaggcgcaa cttcaggtctacaggttgcgagtagtctcttagagaagtggggcacgctgtgtgtatt gocggagcgactgatgtccacttgaggtcgaacacattgtcccgaatctcgtgtgtgaa gcaatcgggtatccaactcaacttgcctgccagtgcaatcaggagaaaaggcaatc gaacagcagaagaatttggcccccataatccaagcgcaggcgaaaaaacactcaa ggatgcagcaacaatgaccttcatcggcgagagacgagcgatgacgagctagatcac gtgagcgagatcaggcggaacactgttccaaaggccagccaaagggaattgcgctt caacgacgcaccttcccgccatctataactcgcgttgcgcctctcgtgtggcctgat gggcccgttggagcgttcgacacgcgcttctcagcgcggcttgcacgacgttaaga ccagcacgccgcgcgagtgagcgagcaaaagttaggcttcgccacttccactccat	gtcatagacccacggcgcgacccgggg ggcttggagggtgcagaaaggtacgtgaga gtcgttgcgtagctgaacatctatgaccag cctgagtgccggagcccgaaggacaag ggcggcgaagcactacgttccgtgggtcat ggcaccgccgggtacttctccagccccagg ccctgccgcctgggttaaacgcaagctgg gggtgtagcgagccgtgccgagcgcg gaaagcccttggaaatgggcgaggagag acgcggcgaggtccgtccgcgctcac caggcccgtagggttatccaaaaggagc ggaacgatg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gcccgatgatccggcaccacagtgagcgaagtcgggagagcctgaatggcac attcgaagctatgaacgtgtgagtcgcccctgcatgatgacctgaaaagcttgcgta gtcgttggccacagatgcacactgcgccaatgtcacctcctttcgtctag	
810-811	human gut metagenome genome assembly, contig: scaffold97368_3, whole genome shotgun sequence GENOME_ACESSION: UMJL01.1 GENOME_ID: 15909 CONTIG_ID: 18423 SOURCE: WGS DATE: 2019-01-27.	atggtttacgttaattacaacaaggacagcacttatgccactgaaaggtttgtaaa tgagaaggctattaaagaatggtcttgcctcatgttatgacctattccattacaattcaatt ggattatgacacaactaacttcaatcagccataaagtttgggtatagatgctgtagcaagc atatcggcatatcggcaacgacaagttagaaggaaattgtatgcagcagatgtagaactta gaaatgatattgtagaagaactatctactcgtagagaacaagaagaactcgtagaata gattcgttatcgtagggtcgtttcaataatagggtttcatctaaaaaaaagggttggttag caccatctattgaaaacaaaatccagactcacttaactgtttagagaagatacataagtttc tgccaataactaattatagtagaactcgtctgtttgatatacaaaagattaagaatccaa gtatatcaggtaaaagaatacaacaaggagaacaactcaactctttaatgtcgtgagtat gtgctcatcgtgatgtgcatcagtgcaacattgcaaaagtaagagtaagaccctatctt gaatgtgcatcatattgagtcgagaagggttggtgaaattcaagtcaaaacctcataaca cttgcgaatctgtcataaagcataaccataaagggtgaatagaactgaagttaagcgtg gcacatcattcagggtattctgcatttatgtctacaatgcgatggagcttctacaagaagtga agaatatctatcctaattgaagtagacttttgggtatatacaaaagcacactcgtatcactaa tgggctatcaaaaagccactatgttgatgcaagggtgtataagtggttaactctaccgctaaa cctcttgaatattttctatcaaaaagaagtgcgttgcataaacagacaataacacaaggc taattcttgaagggtgggaggaagaactcaatcaagcacatacttggtaaaagggtat agggtattgacttagttgaatacaaaaaggattgtattacataatttgaagaaggtagt gggttctttagattaggaacttgacggtacaaaagtaaacagggttctttagttgtaag cagatcggttagatagataaagaaaacaataaattgaaaagcgaatgcaaggcgca atttctccataaaactaa	aaccagaataattgtcaactaccacagg ctaaagacctgtgggcttgaaaaagcccaa gttgattagctaaagcactcgggtgctacgt aggagagaatatatttaccagggtgggtgt ttgcctcaagccccctgctctaagggttagt taaacaattcctgtaggtaggataggtgtg ctaacgaaaacctctcataacattgacgat gagcatttaacggagaatccgacttatgt aaaatttttaaaaaaagaatg
812-813	Ga0074432_100015 GENOME_ACESSION: IMG_3300005683_SF_3 300005683 GENOME_ID: 27814 CONTIG_ID: 6 SOURCE: JGI DATE: 2019-01-27.	atgtcagttattgtcctatagaagcgaagaacctctgatcccttgaactgagaacggg caaggcttctactagaacggaagaaggcgttcaaccgacgggttcgttcacgattc gcctcaaaagacgcgtaggcggagaagtttcaaccctgcgcatgaaagcttgaccggg gtcaaaaggtcactggaatcgccctcaccgtgaagtggtgaccgaaacgacccacccta tgttctcgcgggaattgcatcaccgaggggcatccatcaagaagtcactcgagcaacgct ctggttatagacgtcgacgaaggctcggaacacctctgctaccgagccacagatttcca acagaacgaagcctaagggttggttgccccccagctctccatagaggtagagacgacc ttgtctgggtcaatagggttcgacgactggttccggctcagtcgaatctcgtaggagttggt cgattcgacctacagaagcatctcaatccagaatctctgtattgagtaacacaagggtg gagcttcagggtgacgaatgacgggagattttgttggaagtggtgctggaggtgct tactgtgggtggaagggttccactcaagttgagcacattgtaccgaaggcttcgggtg gctcaaatcgtgtctgaacctgacgttatcgtgtcacatgtaaccagaagaagggttc gaggtctgtcaagacatttctgcgaagaagccggaggtgctcaagggttctgtgtgt gcgaaggcgcgaattgcggacgcagcggcgtggaatcgacaagggtgtgttctatag atacactcaaggctacagagctacgtgtcgaacggctagtggaggtcaaacgaagtgg aataagaagccgatttagtataccgaagactcacgcactgacgctttagtgtgggtatatt aaaaggcatcaaaaactggttacaacaacgcttctatctattctacaggtcggggagtc taccagaggacgagggttactgttgggttccacgtgggtacctaattcgaaaagaag tcggcgttcggattcaaaactggcgacttggttcagcaggtgtactgaaaggtaagaacg tgggaaccacacgggacgtgtagcgtatcgggtgactggaagttttaatccaaacaa gtcacgggttaatacaaggtatttcatactgttactcgggatgttcagcgtggcgatgg gtacagggtacgagattgggtcttcccaactttcaaggacgggtttttgccccgaagaac caatga	gtaaacactcctctttagggaggtgcttt gtcggggagacatcaaaaacacctgtttac cagggtatagcgaagactacgttgtgattag ggatccagaccgacgtccgaatgcttctca gttcagacctctcgaagacatcttgcagaca agccagagtaggcacgaaacggaagatg tcccaacaccggatcacacaacttccgagg ggagacggctcgttaagggtccgcttactag gccccgaagggtcgcgaaggagatatgt ca
814-815	Chloroflexi bacterium isolate Amazon FNV 2010 25 1 contig_75_703380_lengt h_11508_multi_3_in_2_o ut_2, whole genome shotgun sequence GENOME_ACESSION: GCA_003175725.1_AS M317572v1_genomic GENOME_ID: 161678 CONTIG_ID: 303 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	atgcaacacgtgttgggtctgatcccagcggcgtcgttgatgccgtgtcggccgca cgagcgcgattgtcttaacacaaaggcaagcagctgtgtcgcgacgtctcccttcc atcattttgaaggagtcacaaacagacgctggcgtgaagccatgcgcctcaagattgat cccgttagtcagacgacgggttggccctgtcaccgctgcgacggacacctctgagc agggtcacgggtatgtgtgtgggcagcagaattgacctcggggcagtgaaatccat cgggacgtgaccagccgtggcaaggctcccggttcccgaggtttcccatacctggtta tcgagaggcgggtaccacaacggacccgtcccgggggtggttccgctcactg gaaagccgtgtgcacaacgtccacctgggtccagcgtcgtgcgggtgtgttcccat cggagccatctcgtttgagcgggtacgggttgatcgcagttactcaacacccggatatt gctggcatggagtacaaacgggggacctcgcaggatcgaggtgcgggagtatctct tgctgaaatgggggtatcgtcgcctattgccaccagcaggcgttctaccaattgtgtg ggaaattgaccacatcatgccagcagtcggggaggaagtgtcgggtctctaattcgc ccctcgcattgcacaaatgaatagtaaaaaggagaccagactgctctgagtttgggca cccggaggtccagcggcagcggagagacacctgtatggatgccggccgctgtcaaacg cactcgcagcgggtgcacacggcctcctagcgtttggcatcccgtggagaccaaca gtggcggctgtgacgaaatggaataggacacagcatagcctaccaagaccattggctt gatgcttgcgtcgtcggcggagtactcctacgttctcgggggtggcaggacctgtcc cgctcttgatcacggcagcgcgtggcagcggcgagatgtgcttaatgaacgagcat	gtcagataccgggtcccagaagggtg ggcttgcgcgagcaagccatctgaccagttc ccttctctcgtggaggaggagccgttgac gaggagcgttcagagtcgacacctgggatg accgatccagtcaccaggaactgcaagcgc acggttaaacagcgggtgggtgccgagtg ccgtgcgagtagtaccgtcgtcaaccagg acgaggatcaccttacctgcgcttccga gccaatgccaggcagcgtcgcggtg tatggcttcccttagccagggaagaagatt ctgatgtgtatgaacac

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ggcttcccgcgtacgagggccaaggaggccagccgagtcagggaattcaagaccggc gatatggccaaggcgtgggtccgagtggaaccggagggaatccattgtggcaaaag tggcagtcgaaagcgcgcggctacttactgtggctagtgtgcccgtatgcccctcgccta ttgccggctccttcagatgcccgtatgagatataccaggagatcccgctgggag atcttcccagtcagcaccgggtcaggcggcgggtagcggaggagcgtgtccgcgaa gagaccagagtgatcaggactga	
816-817	Ga0123338_10037050 GENOME_ACESSION: grotaG_2_\$F_330000968 6 GENOME_ID: 37706 CONTIG_ID: 37049 SOURCE: JGI DATE: 2019-01-27.	atgaaagaaaaccagagatatgctttgtactagataatagggtattatccacaac aaaagaagaaaagcatgggtataagataagaaaggtaacgcaaaattatacaactaa aacctatgataattcaattaaattataaggtagacaatacagacaatagtaacatttacgtag gattagatgtaggtgaacaactgtgtgtaggtatgctcaaaatgcaaaacaagaataa agtaattttaagggtgaataaagcatagaaagatgtaagtaattatggagcaaga aaaaattatagaaaattaagaggcagaaaaaagatataggcaagccagatttgataat agaggtagtagtaaaaaagagtagattacccttcaataaaaaacagacaagatga aataataagattataataaattattaaattactagatataataagtaatagtagaagat gtttcatttgatatacatgtattacaagtggttataaacatataaatgggaatacaaaaca cttatagggtatgataaaaataagaaaagcagttattatgagagataagtttactgtgcaa atgtgtgtgtagtagagacaagattagaagcacatcataaacacaaaaagacaaaa cggtagaaaatacatgtgataacttaataactttatgtacgaactgtcat	tataaaaagtataagtaattgttagtggtttca agtgcagagtgagtgagctacaggaactg agacgcacaatgaactacgttagtgaaaga agttaaacacttacccttggtatgtccatcag tctgaagctctaagagttttgaggaaagaa gtaaacctaatgagccttgatacagagggtat atgaaaacctaacccaatttaacatgaactt atcgttgacattgacaagaaggaaaatacttt taggaaggttggaataatcctatgaagaa aacacagag
818-819	Ga0187784_10040265 GENOME_ACESSION: 1015_SJ02_MP15_2_2_\$ F_3300018062 GENOME_ID: 20522 CONTIG_ID: 40264 SOURCE: JGI DATE: 2019-01-27.	atggcctttgtgcttgaccggaacaacgaccgctaattgccgtgcacagagaacgggc acggctgctcctcgaacggggccggcgccgctgcaccgctgtatccgttcacgatcc gcctccgcgaccgggtgtgtgaggacagcgctccagccactgacgctgaacttga tccgggagtcagaccaccggcattgccgttgcgcgctggcagaggcaccgccgg atacagcagagacagtgatctcgcgaactgtgctaccggagcagatcaggtctgtc gtcagatgcagcaacgcgcaattaccgccgacgacgtcggacagccaacctccgtta ccgagctccccgttccgcaatcggcgtcagcggtgaggtggctcggccatgcctc cggtaagggttggaacgtggtcaatgcagcgggagtgatgcgcggctgcaccca tcacacgcgttgacttggaaactgtgtaactgcatacaaaaagctccaggcttcagaga cctccgggagtcggtatcagcagggcacgcttgcggctatgaagtcgtgagtactgt tcgaaaagtggggcggaagtgtgctactgcggtgccgaggacgtcccgctccaaat cgagcacattgtcccgagaagccgacacggttcgaatcgcgtcggaacctgacgttg catgtgactgttaacgagggcgaaggaaccgaccgcgggaattcggtaccc ggaggtccaggtccggcgtgacctgccccctcgaggc	gtcatagaccccgctgaagccggggg gcttgcctcgcagcgtctatgaccaggt caaggtcctgagcggccgacaccgcggt gacggactacgttgcagggaagccaaa cgacccactccggatgctccccagtc ggactctggaagcgggttcagacaac ctttgggtgtgacgaaacggatcgccgc cgaaggcgccctgcacattgcgagggg gagcggctcgggacgtccgccaagccc gtcacaggccgtaaggcgctgaagga ggtaagatg
820-821	Ga0373634_0000655 GENOME_ACESSION: RO12ediment_MG_FD_ \$F_3300036715 GENOME_ID: 312626 CONTIG_ID: 654 SOURCE: JGI DATE: 2020-04-07.	atgataccagtagtatataaaaaaacattctgcccgtcagaaaaaggcga gaattattatgtacgtataaagcgagccctattgggtataaaggattattttattttta caaaaagagccttctcaagaagctatacaaaagatatgtataggattgatctggatcaa agatgagtggttatagcgtataaactctgttaaacacacattaatgaacttcaggtgaaagc cctaatacatgttaaagcggcagtggaacaacggaataatgcgtcaagcagcaggaag tagaaatactccgtatagaaaatgcagatttaataagatctgtaggtaaaagaataaccacat caacaaaaagtagatggttaacaacatcttaattattagattaacatcaagatgtatcaa taaaagaagttgtgttgaggacataaaagcaagacccttagagggaagcgtaaatata atgttaacttttcgcatagagttggaaaaaactgttttataatcaagttatgacctat cctttaaaacatatcaaggatatgacacatacaatgagcgtcaagattaggtataaaga aaacatcaagaagtagataaagtatttgggacacatgcggtatagattgggcatattgt aattttagtattaggtggtataaagttaccagagaacaaaagggttaacatatttagagccatta gtattttccaggaggcagttacattgattatgaccataaaaggaaatattcgtaaaaaatc ggatcaacgatgagccttggaataaagagggggaacattatgtcaacataagaatattgg atattgttagttgtgggactattatagcaggctaaagttacataatacttttaattataaaa gattaaactcaagcgcaaaaattagaagatttaaaaaataaaacgatttaaaatataaaatg gtgttttaagataa	atataattatacaacgcctagcaataggtgct ttatgaaaactttaaccaagactaagtaagt aaatggagattagcttagacaagatcgcgct ttgtgcggctcgttggaagtcacaaagtc aggctgtgttagaccattatataaggata cctcttagtcccttttccatattgacgagta gcgaagagaaaataaaacaggtacgtaag tacaataaaataaa
822-823	uncultured Bacteroidales bacterium isolate RUG14530 genome assembly, contig: RUG14530_asm_132, whole genome shotgun sequence GENOME_ACESSION: GCA_902800075.1_Rum en_uncultured_genome_ RUG14530_genomic GENOME_ID: 348289 CONTIG_ID: 131 SOURCE: NCBI Prokaryotes DATE: 2020-04-22.	atggtttacgttcttgataaagaaggagccgattatgccgactgaacggcatggcgaag gtaagaaggcttttgcgtgatggcaaggctagtggttgcagactccagccatttaccattcg gcttaattatgatagtgacacattataaacaggaggtatgccttggcatcaaccttgatcca agcatttcggagtgtagcttcatcagaaaagaaggagctgttctctgcggaagtcgtgt aagagacgatactgttgagaacatctctaccggaagggaactccgtcgtggaagaagggt ccgcaagactcgtaccgtatccgagattcaaaaatcgaggtaaaaatgaggatggtt ggctcccgatgattaacagtagatagatgcacatgaagatgattcgtaatgtgcatgata tccttctgtcagcaaaacaatcatcgagatttcgaattttgacatacagaaaatccacaatc cttccataaaaggattagtagtacgaagggcccccagatgggattcggaaatgaaggg aatacgtgctttggcgtgacaatcatgtatgcgcgaactgttttggcaagtcgccagatcct gttctagaagttaccatcgaagaccggaagcggcggtgatgcgccgataatcct atgacgctgtgcaagacctgtcatcaggcataccatcaaggcaaaatagatttgaaggc aagagagggtgcatctgatcgaacgcatccgcagtaataaaatgaagcaaaccttatt gacagagcaacgaaggagttcacaacgtcagctttacgtatggctatattactaagaata cgagaataagtaattgtattgcaaaaatgatagttccgatgccttctgcatagccggttgc atttctgcaagcaggttgccttactgtttaaggtatagatgttccgcagacataacagatca	gttgacacccctgcacgtggctgttgactca agcgtgtacaagaagcggactagcctcag tcccgggacacccgggaactacgttggg ggtgaattgtataggacacccggatgctc cccaagtcgggactcgtgtggccagtggt taaaaggagcgaagctacggtgctgctg tatatgaaccaccccgcaacattggcgatg ggcgcatcaacaggcgcaagcctgtattaca acatagttgttaattataaataagagaatt atg

[illegible]

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gcgaattcctattccagaagcagactcgccggaacaaccggcagattcacaactctcca tcctcaaaggcagcctgcgaaaacgcaaccaggccctttgaggtcaaaagcttccgt cttttgacaaggtagcctaccagggggaagaaggctttangaatcaccgtggaac gacctcttcgacatgcagctgctgaagaattcagacattgcaggaaagagtaccaag aggcggaacagctcgcttctggaacgtccgcgagtagcttcttcagagacggcac gtttgtcagcattgtcacggcagatcgaaagacccggcttaattgtatcacttggaga gcagacgtacggcggtgattcggccggcaacctgattacgctttgtgagacgtgccata aggcactccaccgcggtgaaatcacgctgaagcgaagcgcgacaatcctccgtgc ggagaccttcattgggaattatgcgctgggaggttttgaccgctgaagcggtcacattc aaagctagaggtgcgaaaacacctacggctaccgaccaagcacgcgggatctcgaac ggcattgccaaatcacattgtgcagacgctttctgcatcgccggaatctcggcgccgaa aggctcgccgaattcttccagaagcagacgcgcccgaacaaccggcagattcataa gctttccatctcaaaaggcggtcattcgaagcgcaatcaggctccctatgagatcaagg cttccgctttttgacaaggttgcctgccaggagagaagaaggcttcttcttggccgtcga catccgggtacttggatgtcgaagcttaacggaaccacattcagcaggcatcagctg caagaagctgcgctgctggaagagacaacattttaaagaaattcgaaggaga aggcgcttctcctcctcctgaaggcagggtctccgcccctaa	
830-831	human gut metagenome genome assembly, contig: NODE_2506_length_13891_cov_2.537872, whole genome shotgun sequence GENOME_ACESSION: UXR01.1 GENOME_ID: 279330 CONTIG_ID: 2505 SOURCE: WGS DATE: 2019-06-06.	atgggtgatgtacaggacatagatggtaaacggatgatgctctacgacaaggcatgggaa ggttagggcattgctaaaagacaacaagcggtcgttgaacacatgctctttaccatc aaattgatgtacaagacatcagattacaacaagagattgtttaggcgtcgactcgggaa ccaagcatgttgggttcagctacgacgaaagcgaaggagctttacgaaagtgaggttat tctaagaagtgtgttggatcttctatcaacaagaagggaattaaaggagcactagaaga agcaggttaggtatagaagcaagattcaataatagggtaaaatccaagaaggatgg atggattgctccatctgctccataagattgattcctatgttagaattatcagtttttatattc tatactacctgtctcaaaattgattgttaggttagcccaatttgatactcaaaagatcaagaa tccagagatatcaggtaaagatgatcaggaaggtagcaattaggatttggaaatgttagg gagtagtctttagcaagagacgggcataaattgccagcattgtgaagggttaagcaaaagat cctatccttaatatccatcatattgagtcgaagagataggaggagattcaccatccaattta attactcttggtagactgtcataaggaaattcataaaggaaatatacaattgaaagtgaagc agaggcaagtcactcgtgacgcgacctgcatgggaatcatgaaatggaaatgtgtacga ggagttaaaatccagatacagataacgtttcgtgacgttcggatacatacaaaatataatc gtataaaccatggaaatgaaaatccatgtatccgacgctttgtgatttcaaggaaatttaa ttcatgtaggcttggatatttacaacgtaaattagttcgtcgccataaccgtcagattcat aagatgaaatattgaaaggaggaattgaaagcgaaaccaggctccttttaaagtttttg gattagggttattgataaagtgtgttcaaggagaagagcattttattacgcaagaagcg ttctgggcaatttaataatcgggataattaaggagaataagaagatgtatcttgcaaga aattaaaatattgcagccggggcggttaaaagcttgaaaggctgtagaggtatggaagc ctgatatagaaggagcgtatgtatagatgtcgccgcatccacagagaggtttaccgactg tatgcttcaaaacggcggaagaaatgtatgcgtagacgtaggaactaatcagcttgc atacagcctccgtcaggatgaacgtgtgtatgcattgaaaaaaccaataccgctatgtta cagaggatgatatactgagaagcggtttgttaacaatagacgtcgatttttactcgtg aaaaaggtgcttgagccgtgtataagcttatgcacaaagaaatgacgatagtatgtctga taaaaccgagtttgaagcagggcgtgaaaaagttggaaaaaaggcgtagtcagggtg caaacgggtgcatcttgaggttgaaccgatattataaaatgtcgtgagcattgtgtttaca tacttgacttgaattttaccctgtagaggaccggaaggaaatataagatcttatttatat gaagaaagacgatgaagctaattgataccggatagctgcttgatagggcaagtgatgtgt aagcagagccaattctgcactttag	gttgaaattcatttaattttatataatctat aggatcagggttattagcctaagtcttgaata aagactacgttattggagaataatagttacct acggatgtttatccaaagtcgtagcttaagg taggtgattaaacagggtattgtttgggttc caggttgcctatataaaacctcaataacatt ggcgatgggtactaacagggttttgcctga cttatgttgaataaacattgaattgtttgtaaa atg
832-833	Ga0307928_10013462 GENOME_ACESSION: AceLakmetagen232_2_\$F_3300031227 GENOME_ID: 279923 CONTIG_ID: 13461 SOURCE: JGI DATE: 2019-06-13.	ttgaaaaacgtaaacagagttttgttttaaatcatgacaagcacctcttagtcttgcac ctgcaagagcaaggcaactttgttttaaggtaagcagcaatttttaagaaatgacctttta ctattattttaaagaaccaaaaagtttaactcctaagtttcaaaatctcaagttaaaatagatcc tggaaagtaagaccacagggttagcttattctttaaaggaaaaataaaggatggatctta atctgggctgctaatttgaacatagaggacaatatattaaagacttttagaaaaaagaaa aaatgttagaagaagtagaagacatagaaggttagatataagacctcaagatttttaata gaacaaggaaacgaagattgttagctccttctattttaaagtaggattaataatgttattgtc ggtgaaaaaactcaggaattttgttacttaagatctgtattatggaagtagcaagattta atgcacaaaaatgatgaatccagaataagagagtagtagtaccacaaggaaactttaca aggatatgatgttagagaatacttattagagaataatcacagaacttatttattgttctaag aaaaatgtactttggaatagaacatgtattcacaaggtcacaaggaggaagtaataagag tatctaatttaacttagcatgaaattgttaatacaaggaaaggaatcaaccaattgaattt ttcttaaaaaagggaagcagaacttctgaagaataaaaaaagggttaaaagtattcttga agatgcagcagcagttacattattagaaaaaggcttagaagggaagttgattactgtattg aaactaaaacctcaacaggtagtataactaagtttaatagaacaaagcaatattacaagaa agatcattggatagatgcagcttgtgtgggaaaaaagggtggcagaacgttaagtattcc agataatttcaaccacttttaattaaagctatgggacgaggacgtagacaaatgtgttgg agataaatatgatttccaaagagctaatactaaaaagtagaataaaattttaaagggtttca	gttaatcattacgtaccagaaattataattct gtgactgattaagtcgacgtaacagactaa gtacttagactacgttaaggagaatata taggtactttggaatgcttccagttccaaac tctacgattaggttttaaacagagataaggtc actgtcagtgaaatcctaataaaccttcca ttaacattgtcaggagaacataacagatttt atctagtaggaggaatcctaattgaaaaac gtaaacaga

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aacaggagatattgtaaaagcaattgtaacagaaaggaagaaagttggacttatgttggagaatagcagtttaggaacaagggtactgtgatatactaccaagaagtttaattcaaggcataagtttaagatattgtaaaatgtctacaaaaaattgatggatattctataaacctttaactggaggttag	
834-835	Lactobacillus harbinensis DSM 16991 H627DRAFT_scaffold0014.14_C, whole genome shotgun sequence GENOME_ACESSION: GCA_000425885.1_AS M42588v1_genomic GENOME_ID: 51904 CONTIG_ID: 13 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	gtgggtgactgaggtggttcaaaatagaatattgtctcctcaatcaggatgggcaagcattgatgccctgtcgtccgtcaaaagtgtcggcatctgttcattgtaatcgtgtatcatagtcagccatctcgttcaccatccagttaaagtatcagtcgggtctgttctacaggacgtgtcaattgtgtttgattcaggtcagcagacataatgtgactgtcagtaacgagcaacgaaaaagtgcttggcaaggagaagtgactctgcgacaaaatgtaaaagccttttagatacagtcacatgtatcgacacacacgtcgacagcgacacacacgataatcgaaatgtgtacgtctcgaaccgaaa gccctagagataatccgttgggctgtgtgttaccacatccgttcgacaaaaatcgagca caatatccaatggattaatcgagtgctgaatgttttgcacaaagctgatatactattgaagtggccaaatttgacgtacaaaagttgaagatccgacattcacggtatcggtatcaacaaggcgacgcttgtgtatgaaaatgtgaacaataatgtctagaacgtgatgaacatacctgccaaattgtgaacgcaaaattgatacggagacgaagaaacattgaagttgcacattca ccacatcgtttacaggagtaaaaggcgggtactaacgcagccacgaatctgctgacgggatg tactaagtgtcacacagatcgaaaccaatagagccggcgccctcttattgctgtattttgag caaaaaagacaataaccaccgttaaaaatgccacattatgaatatgttgcgtaatcgatt gcttaccctcttctgaggcgcatcttaccatgtgctacatcacacaaggcttgcgaaaaa agttgggtttggcaaaagcacattatcgcatgtctgtagcaatcagtggaattcaacaaatt attgaagagccaaaatcagttgtgatgtttgatcatgttcggacgaaaaaacgatattgca tgaagcaacggcacgcagggggcgaaaagcaaaagaatttactcaaaaacgagtgaa aaagaatacgaagaagaatgaaaggttgggtcgtgaacgattatgttcgaatcagtgatgga cgttgcgggtcatcaccgggtttagtgtctctggtatggtctatattaggatcgtcaggg tgggttggtaaaaagctcgtgagtcgtactaagcttgccttctctatcacacaggaacttg gaggtgtacaactttaccaacagatgtctatgacatgcaataa	gtcagttaccacggcgaaagccgagagc aaatatgttctaactgactagccctgtctag caatctaactaccagactacgttaagtagg tgcacaccttcagatgtgtccctagtctgtc g cactgtgagggcaacgtaaacagcggtga ggggtgtagccgcagtcgaccctgacatcat ccagtcgctctccacaggcagggacccta ttaacattggcgaaaggcaatccacacatg ttagtctcattgtcgtatgtatgacccctta ctgaaaggttgcgtgataaataaaaaacaa aggggtggtgactgaggtgttcaaataga a
836-837	human gut metagenome genome assembly, contig: NODE_185_length_6055_1_cov_7.280233, whole genome shotgun sequence GENOME_ACESSION: OVYF01.1 GENOME_ID: 11769 CONTIG_ID: 184 SOURCE: WGS DATE: 2019-01-27.	atgcgcggtcagcccttgatccgtgtcgtcgcctgcaaaaggcgaggcatctgtcctaaaggc aggcgaaggccgcggtgagacgacgaacccgttcacgattcaactcgaatcgccacc ggtgaaacgaagcagagcgtgtgacactggcggttgacgccggcgcaaaagcatgtcggc ctttccgtcgtcagcggaaaaaggaagaggtcttgcgtccgaagtcgaactcgcacagga catcacggggctctgctgctcgtctcactgcgtgtgacgacgtctatcgcaagac gcgctaccgtcgcgccgcttcttgaaccgtgtctcgtcgaacacaaaggcgtgcttgc accgtccgttgaaaaccgattcaggcgacatgtcgcgcattgatgcgatctcggact gctgcctgtcacaagatcgtgtgatcagacggcatccttcgacattcagaagatcaagca tccggatgtcagggcagcgactatcagcagggcgatcagctcggcttttgaacgtgc gtgaatacgttctcttagagacggtcacgttggccaacactgtcacggttattcgaaggac aagatcctcaacgtgcaccacctcgaaagccgcaaacggcggtgatgcgcccaac aacctgattacgctcgtcgagacctgccacaaggcctgtcacgagggaaagatcaagtt gaagggtcaagcgggtcaatcgttcaggggcgaaagccttcattgggcatcatgcgttga cattgtcgtatcgtatgcgcaaggcgcatctgttgcctgttgagaacacctacggctat ctgacgaagaacacgcgcacactctgggactccccaaagcgcattgctcgtatgcata ctgcatcgcggggaacctcaaggccgtacgaagaggcgtcgtcctccatcagcgacag gtgcgcaagcacaaccgtcagattcacaaggtgacgtgtcgtcgaagacgcttaagga cggcacgaagatcggataccgaaagctcaatcaaacaccgcatctgggtcaagaactca ggctcttcgacaaggtgagatgtcttgggcaaacgggttcattctcggcgaaagatgt ccggctacttcgatccggaggtgacggcgtaaaagcttcttcggcatcagctaca ggaggtcagcgtctcgaagaagagacactatttaaccgaacttagaaggaggac ggcgcttctccctgtctga	gtcaataaccctctgtgaagacagagcgtt gaagagcctttattgactagcctcagcgcc cctcttcgagaggcgctactgttgggtggga atgtataggcaccgtgggtgtatccttag ttcacgctctgcggcctgtgattaaaggctc cgagaggtaggagcgggtgttcaggcacc aaacccttccaaactggcgaaagatgtca acccgccctcggccgtgtaggcggaac ctgcgggtatccgcaaggagatgctttgaa a
838-839	SAR324 cluster bacterium isolate NORP136 Contig_source1382A_8018, whole genome shotgun sequence GENOME_ACESSION: GCA_002401295.1_AS M240129v1_genomic GENOME_ID: 135030 CONTIG_ID: 170 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	ctggtacattagcgaggcacacattaccggcttccgggattctctcaaggagaaaa gatgcaacgagttatctgtcgtcgtacataaagaagaaccccttacccttccgacccaa gagagcaaaacaattgctaaaacaaggtatagaagccccgttcccttaatagaacactg cctcaaggctggctagctccatccctgtcagtaggggtttcaacatagaacatgggttg aaaaagttatgtgtttgtcctcagtcgaagcgatctcaagagttggtacgttgcacatgc aacagatggagaacccggaatttccggtaaggataaccagcaaggcacttttagctgga tacgagacaaagcaatatcttggagaatggaatcgtacttgtgtctattgtgagcgca aatgtaccttggaaatcgagcatacttccaaaaagtaaaaggtgatcgaatagggtt agtaacttaacttgcctgtgtgtctgttaataaaaag	gtcaactaccccccttaaggagagcgtt gaatagctcatgttaccagcttaagttttcg aaaactacgttaccttgggtatcttacctgg agtgcctgccagctccaggctctaaggttgg atgttaaacagactgaggttataggctcagtg cattcagtgtaaaaagccctggttaacattagc gaggcacacattaccggcttgcgggatt ctctcaaggagaaaagatgcaacga
840-841	Ga0395987_006804 GENOME_ACESSION: 713CPN11X4metaG_FD	atgttacgagtacgtgtgttcaaaagacggcaaacccatcagccctacaattccatcca aagcaagaagatgatcagagatggcgtgtgctacgccgaagcgacaaagctgggca acttctacatccaaatgaacatccctgtcggcgagaaagattccccacgaaacatcgccg	gtcaactacccccgttagccatccgggg cttgaagcgggggttgcagggtcagctcc cgcacgtgaagagctgactagggcacgaga

198

SEQ ID NO:	Name	IscB sequence	ωRNAs
	SOURCE: JGI DATE: 2019-01-27.	gccgtgctcgtagatcacgtaagacacgctatcgccctgcaagggttcttaacagaacta agccaaaaggctggctccacctcgtagaacaccgcttagtaacgttatgactttgtg catcgcttgcgtaagcttgcataactactggaatcgcgtagagcttgaagttgatac caggctatgcaaaatcctgaaatctctggtatttatcagcaaggtgagctagcaggttat gaaatgcgtgagtatcttctgaaaaaattaatcgacttgcgtgtatgctggagctaaaaat gtaccgttgcgaagtggacaatgttgcctaaaggcttttaggttgcaagtcgacttagcaa tctaacacttgcgtgcgacttgcacaaagctaaaggcttctgctatcgaagaatact tgaaggcgaagccagcgtactcaagaaaattaaagcgcagctgaaagcggcacttaa ggacgctgctgctggaatgcaaacgctgggaactttggagacgcctgtcggctttagg cttaccagtggaatgtggctcagcgccttcaacaaaattatagaaccaagcaggcgcta cctaaagctcattggttagatgctgcttgcagtaggtaataacacaccgcttgaagttaatt aaagtggagccactttatataaagcttaccgacgcggctccaggcaaatttggcaaatgg atgcttctggtttcccaaacgttcacgaactaaagaaaagacaaggttgggtttcgtactg gcgattgggtaaaagcaatagttcttaaagtaagcatatgggggtacatattggaagag taggaactagaagtaaacccagcttgcgttaggttaaagtagatgaatttaatttaattgta tttacttatgcagcgtaatgatggttacgaatatagttatgggacgtag	gccgtgctgtaaacagggttaattccagtgct gcaggcttaacaagccattataacattgtcta ggtagctgggcttaggcctatatacccgct aatcacggagagattttgattaaagt
848-849	Lactobacillus salivarius strain AH43324 AH43324_contig_69, whole genome shotgun sequence GENOME_ACESSION: GCA_002079805.1_AS M207980v1_genomic GENOME_ID: 120315 CONTIG_ID: 67 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgatgcagaaaagaaaagttgattatctatctagtagactccaacgacacaccact catgcctacttcacgttttaggtatggcgctgctgtgttgaagacaggacaggcaagatgg tacggaaatagtcgcaagactatccaatttcttagaccttactactaactcaaaaacta acactaggagtagatgctggttttctactaggtttatcagttgtaggtaacagcgtgaatatt atgcagctgaaaagctaaagaaacagaaaagatgcgaattactatgcgaagagaattaa gacgtactcgaagagagcgtttaaagacatgtaaggctagatttaataatcgtagacgcaa agacgggttgctagtcttagtattcaaacacagacagtagattttacgattaaagaaatcaaac gcttatatacgttttaccattactaattatgctgttgaagttacaccgttcgacaacaaaa actactaaatccagatattaaaccttggcaatatactcaaggcaagatgcacggttttaaga cgattaaagattatcttctagctaggataattatcgtgatgcttttagatggaaaacaatatcc agctagccaattaaagagttcatcactttagtccaagaaaagatggtggctctaataacca gataatttgggtttgttatcagatgtaaatcacaatcaagctaaccataataacggatcttag ctaagtgaagaaaatcgtcaaaagaccattgattatcgtggagcttattttatgagttttt ggcaacaagattaaagcattatttgaacattatacaaacacaaaggttatctaaccgcta acttaagacaaaagtataagattgaaaaatcacttaaatgacgctttgtgattgctggcg gaactgatgtaacgcttagaatgaacaacgtttactcaaggcaaaaactaagaacaaca accgaagtttacagaagttctatgatgctaaatattgtgataccgagatggttaagaaaaa gactggttaagaaattaaagtcaggtagaattaaacgttccaaagagcttaattatgataatt aagacagtttagaaaagaaaaagttaaaaaggtcgtgtttcgttagacgaaatcactac cagttaaagaccacatgatgtgtgctaaatatacaagactaataagattgaagagtcgcta gtgtcaaaaacaaatggaaaagtagttagattcaaacgtgtaagactgtttcaattaaaagtg tagtttagttatcatgtgaatggaaattagaaaagaaaatggaaaataatttag	acaagttacagaaccgccaactaaagtag gtgactgtagaataacaagctggttgattag cctaagtttaagtaactacgttaggggtgaat agataggtactttggaatgccacctagttct aaacatcatgtttaattattaaacagagctga ggttcaggctcagtggttaacattaaacca acctataacattggcgaagggtacacacgtg ccaactttatgttaggtgcgaattatgattgaa aggagctggcatgatgcagaaaagaaaaa agttgattat
850-851	Ga0233437_1033046 GENOME_ACESSION: SI_122_August201_7_\$F_3300024259 GENOME_ID: 35158 CONTIG_ID: 33045 SOURCE: JGI DATE: 2019-01-27.	atgttagtgttctgatggacaagaatggtcgaatgggacatccgactagtaagtgtggga tgattcgacggaaattaaacaagggaagcgaagattgtacatcgcttcaagacacga ttgttgaagatctttgatctagttatgatgaagaagacagtcgattgtaattattctag ggattgatcctggtatattgaacattggttactatgtgatcaaatgtctgatgtaaatgt atgacattcttctggagaattgaaaacgagaacagagaagatcaagggtttgtattagac aggaaaatgtatgaaatggctgtcggagacatcgtcgaaacgtgtccaaagattacat gggtcagtaaaagtttcgacaccaagatggaagaacaggaagaacatcagtttcaacc aactatatacatcttattcagacgcaatttgaatttgattaaagatacttcgatttgaatt tcgatgaaataatcagtagtctcaagatgattcacaagaagcactgaatccagatatta aaggcgttcaataccaagagggtatagatggtttcgctaatactagtgcatagtatttag atagagatggttataagtgtaacatcatgtggcgaacagatattgttgaagctcatcata ttgtcgagagaactgataatgggtctgatcgtccgaaaactcttgaactgtttgtataagtg tcacatgagatcatactggtagaagaaggtgtcctgtggcttaataaggttaatagcactc aatttcgagattcaggtgtattgaattcgtgcagatccagcattattccgattgttacaagatag taaatttatggttaagaagactttcggtagtgtactaaagtgttcgagagatcttgataatt cctaagacacatcgaacagatgcttctgtatttctatcgaacaactgagtgaagatattaa atttgatgacttacaggtaatgttgaattatcaacaattccggagacacaatcgaaagttt gtaaatcgatttgaagatcgttaagtattattagttgggttcgaactgttcaagcaaggaat cggaagagacgaagtggtcaagataagaagatgatctatcattagaagatttccgtcag acaacttaattggttacgagaatttaacgctaaagctggtggagtaattatgaataatcct catacttatgttccagacaacgatccggaatgaaatttctgttaggttaataataagttc aagaaaaaattcgaacaattattagtagtaataatccagcaacgggtgttctatgaaaa gattgagaaaaacggttatttgatacgtttacacaaattcgaagcatggagacaactcttt gtgtaattcgggaatcgttccgtttga	gtataatcatgcaagtaaaaccgtgtgattt cgaatagcgaagcattataaattatagattatg caactactgggagcatagtagtaccaggttga cacagacaatgaactagatgttaccggtga gtgccactccagctgctgtgaattattttt aaaaactggaggttgaatccctcattgtacga aagtaggtaacgaaatcgtatattacttagc tcagtcgagggagactgtttcgactaacctt aagtaactcgaacaggagtttcaaatgtta
852-853	Ga0209175_10000039 GENOME_ACESSION: J19CDNASPADES \$F_33	atgtctaactatgtttttgttctgataccctcaaaaacctttaactccctgtaagccatcaat agcacgaaagtgttaaatgttgcgaagcgcgctgtgttcaatggtatccatttaccatca ttttaaaaaaggaggttgatgccaaactcgaaccactggaactgaactgaactagatcccggtct	acgctgcatctgtccagattcctgacaagc aattaaataggagactactgaaacccatga caattcgggacaatccatgaattcgaatgaag

SEQ ID NO:	Name	IscB sequence	ωRNAs
	00027781 GENOME_ID: 30210 CONTIG_ID: 38 SOURCE: JGI DATE: 2019-01-27.	aaagtacagggaatgcactaaacaaggcaacaaattattttgctgccgagttgattca tcgaggatacgcgattaaatctgcttctgattccctgcgcgtattcgacgttcgcagcga accgaaaaactcggatcgtcaagctagatttcttaacgcactcgttcaaggcgtggtta actccctccctacaacatcgagtagaagacgacgatgattgggtgaatcgcgttcgtaa cttgacccattggatctgtgtcgaagaattggtagattgatctgcaaaagatggagaa tcctgaaattcagggaattgaatcagcaagggaattacaagggtacgagattcgagaa tatctacttaataaattggagagaaaatgtcttattgtcgggcagaaaatgtgccgttaca agtggaaacattaaacccaaagctaaagggtggaactaatgaattccaatcttctcttc ttgtgatgtgttaccagaaaaagggaactcaagatattggcaattcttagcgaaaaag ccagaagttctcaacgaattttgtctcaagctaaagacctcttaaggatcgccgcgg ttaattcaactcgggtggcgctgtttaatcgactccaagaacagggttacctgtctaacg gggttcaggttgattaaatattaatcgatgcgattaggtttacccaaatcacattggtta gatccgcgtgcgtgggaaaaattgattcactagagattctgacaactcaaccattatgat caccgcgaaaggatgggaggtgcgcaaatgttcaaaatgacaaataggtttccctcg aaagaattatcgggctaacaacaaaggtaagaatggaaaagggtgatactgtttctgta atcaaaaggaaaatggcacaataacagaaaacggatcaaaacgggttcgattcagggt taattttgatgtccgattagatcgctcaaccgtgatctgttatctcgaaatcacctaaggct attcatcgcttgatggatcagattatacctttgtctaa	tcagggtcaggatcagaccaccgttatcgg gaagcgttaagttcctaccttgggatgcttg ccagtcccaagcccccttaaccgaacagttaa acaagcaagggtttattcaagctagtgtctgt ttggatagtagcagctgataacttagccgag gcacacattactgaaaaggctcagacacaac catgtctaact
854-855	Ga0306922_10010477 GENOME_ACESSION: sta12C44000082v2_\$F_3 300032001 GENOME_ID: 281679 CONTIG_ID: 10476 SOURCE: JGI DATE: 2019-06-13.	atgcttctcagtcgccagcttccgagttgcgcgtcgacagaagccagggtcagcgca aacgggagcacaacagacggcttgcgtcaacctgttcgaggggaaaaagaaacggcg cctcaaaacgccgagcgcgggggttaaccgcacgtttgtattatcaacgcagggcatcc cttgatccgtgctcaatccagggcgaggattctgattcggaaaggccgagccagggt ttaccggcttttccgttcacgattcagttgatcgacagggtcagggcagacattcaacc gtcgtataaagcttgaccccgagcaaatcagactggttggccattgtacgcaaggac ccaacgattcgaccagacaagcgtactacacctcgcggaaataaccaccgttggttaa cgcgatccgcaaacacatgatcaaaaggcgatgttcaggcgtcgcggagaaatgct aatcttcgatatcgcgcaccgaggtttgataaccgaaccaagcgcgaagactggctgcc gccgtcacttcaagccgtgtcgacaattgtggcctcctggctcaaccggtacagaaaact cgcgccatcacatcaatctacgtttgagtcgggtcaggttcgacatacaggccttggaaa cccctgcatcgaagcgttgaataccagcgtgggagttgttcggagcggaaactcggg aatacctgtttgagaaatggggcggaaggtgcgcgtactcgacgccgaagcctggc gctggaggcggaacacatcgtgccagggtttcgggcggtatgaatagggtcagcaac ctgacgttgcctcgcgaagtgcacaacagagaagggtcgacgtcattatgtgtttt tagcagatgacccctcaagactggcgcgatcttgattacacaaaaaacgttatcga cgcgccgagcggtaaacgtttaccgaaatccattaacgggtactgtacgggactggc ctggagggtcagtgctatcggggcgccgaacgaagtcaatcgacccggcttggcatt ccaagaccatgcgttgacgccgttgcgtcggggaactgtcaaaacttgagggttg gaactgtccattctcgtatcaaaagccaggacgcgggaagctaccagcgagctgct tgacagttttgggttccacgaggctatctgacgcgccagaaggctgtgaagggtttcaa acgggtgacttggttaaggcgacaatcccaaaaggaaaattcaacggcacacaccaag gccggctagcgataaggggcagcggtttttgtaatcagagctcagcggggaacgtc ggaaccaactggaatactgcaaacatctcatgcgaacgacggttacacctacgaaatc aaaccatcctcgccaccaccaactcttctagactgaaacattccagttatgagtcgaa aaaccctgaaacaagaacacgcacacaactcagtaataacaaaaacaacaacagtta tggcaaatctga	gtcaacgaccgcccgctaaaaagcagcgcg ctgtgagctaatcgaaagccatgtttgac cagaacaagtgaactacgttgaacgaca agcaaaagactaacctgggatgcttccca gtccagggtcttcgagttgccgtgcagac aagccagggtcagcgcaaacggggcagc aacagacgcttgcgtcaacctgttcgag gggaaaaagaaacggcgctcaaacggc gagcgcggggtaaccggc
856-857	Ga0393278_0001403 GENOME_ACESSION: LZ_38metaG_FD_\$F_33 00036985 GENOME_ID: 312543 CONTIG_ID: 1402 SOURCE: JGI DATE: 2020-04-07.	atggttccagttatgacatgaagaaattccgttaatgccttgcagaaaaacgtgctag gaaacttatggagaaggagatgcaaaagcatattggaaggtggcgtattttgtattttt acaaagagagccttcggcaagaactatcaggaagttgtgattggaattgatccagggtc aaaaagaactgggatcaccactgcaacagaaaagaaggttatctgtaatacaacttttgata ctcaagcggagtaagaagaatgttgagacaagaagaattttgaagacacagacgt tccagaaaaactccttataggaaatgtagaacaatagaaggataggtggcgtgccacca tcaacaaaaggcaagatggggagcccatcttaaaataatagattttggaaaaaacttattcc tctcactgtttcttggaggatattaaagcagaacactaaaataaattgtaggaaagtggaa taagaattttagtcattggaagttgaaagaatggtttgagaatgaggtttagtctgtgg atataatttctataaatttccagatttgaacaaaagcacaagagtttatagagaattcat aagacatcttcaaaactaagagatacttgggatgccataatgttattacattgcctttgc gaatttaacagtaggagatattaaacctattatggaattcttaaatgtgaatttttcaaatgga gtagaagacaaattcatgttttaactctgaaaaatataaagagcaagaaggcattcctaaf cttaaaaaagatgggattagaaaacaataggaacaacaagatcttaggactaaacaga ggaacattggttaacataaaaaattggtttaacatattgttgagggaacatcaaatgagaa aatttcttcatgatgtgaaaactgaaaaagattaactcaacatgctaacaagaagatt gtaaagttaactaatttaagatggaggacgaaattcctccatggctaaagccatgggt ttccttgcgtaaaaccaaataga	gtcaataaccaaggctttagccttgggcttg tgaggaaaaatcattcaccagaactaattcac aagggcagaagagactagattagccaag ggggacaaaaaccgataaaggcaactag ggagatgggtcatgttagacaacatagatct tcaggatgtcagctagctttagctctactt ctccagtatcgaagctgtatttactaagaa ggagtaatccaatggttcca

SEQ ID NO:	Name	IscB sequence	ωRNAs
858-859	Ga0272428_1004076 GENOME_ACESSION: FingerMtnordMG_2_\$F_3300031520 GENOME_ID: 280175 CONTIG_ID: 4075 SOURCE: JGI DATE: 2019-06-13.	atgcgaattccagttgttgattcaaatcaaaacaaacttatgccaccacaccagcaagag cgcgaagtggaattcaatctgcaaggcagttaaacgctgttcagattgtggcagtttta cgtgcaactcactgtagaacatctggacgcgatactcaggatattgtatcgggttgac cctggcaagaagtattcagggttgcttcagctgtgaagattcagctgtacacagcac atttaattcccttttcaaacggtaaggacagaatggatgcacggcgactgatcgacg tggacgtagagggaagaatacaaccgccagattgaattctcaaacgcgctcatctca gaagcgcttggagaatcgtcgtcaagccaaactccaccattatctgtgccaatcgccag ttagaactcaggatcgtgctgaactctgtaaaatcctccggttacagaaatccgttacga gtacgtccgagctgatgtgattaacgagtgacgcaagaaagcaagatccggcaag gggttttagcgctgtgatgtgtgacagaatggatgttacaacagttggagcagtttgcgc ctgtcgttaaaagtgaaggctatcaaaactctaaaactcgtgaacatctaggattgaccaag aacaacacagagaagcaaaagctgaattgccactcatcggttagacggcgtaagtatt gcggcatcgactttgtgaatcggaaatcacagagtaaacgttgatggagctaatc gtgttgggactgtttcattacactgcacatcttctgtattcgtccctccatctccg ctgtcaacttcatctcatgtgcctccaagagtgaggttagcgtgaagtgtggctcta caactcgtcatgtgtgctgaagcggtatttagtcaactccccaaaaggaaattgttatgtg tcaggagacaccgaaaacagatatccgtcagcagcgatagctggaacggctagggc agatatcgctaaaaaagtaacttttagccgtcgcctcaatggaattcgtttcttag	ttcaccaacctcggcctcaaggcgcgagga ttgtccgaaccaattcggacaacgtaagagg tgaatagccattgagactcatttggtagaaa acttcaaacacttctctagtcggatttccctt aagcctgattgttcaggcgttgggtcaagc caagacatcttgatcgagttgggttaagaga cttaaacgagtaattgggttatacaaaatgc ga
860-861	Ga0315294_10016080 GENOME_ACESSION: YL17G13_40_MG_2_\$F_3300031952 GENOME_ID: 281645 CONTIG_ID: 16079 SOURCE: JGI DATE: 2019-06-13.	atgttagtatataatttaataagatagcaaacctttaatgcctactagtccttaggaaagcta gattgcttttgaagaaaggaaaagcaaaatcatagaatgaacatttactattcaactta tatattgtagttatgttatagacaatcagttacattaggttagattcaggatcaagaatgt aggtttagcagcagtaactgaacaaggaaaagtattatattctgtcgaagtagaacttcgg caagatattaaagaaaattgaacactgaagtagtatgtagacgcgtagagcgaagaga aaaacaagatatacgaacaaagattcttaaatgaaaaaaacggaagggttggtacca cctagtattagatccagaatagaatctcacgtaaagctagttactgatgtgactaaaatacta ccagtaaaaaatatttagtagaagtaggattattgtatgttcaagctttaatgaatccaaata ttgaaggaaaagaatacagaatgggtattcttaagagatgatagttgtaagaatataataa aagtagagataaataattgtcactacaagattacgttctgatatactttgtatgaagaa gttagaaatagatcatatgataccaacgagtaagggaggtactgatagaccaacgaattta gtttgttctgtgctgcacataatagaatcaaaagtaatatgattatgaagaattactatga agcgtctaccttaagatagaatcctttaagaaaccgttttatgaatgtagtaagagtcact tagtttctgtgtagcaaaacttagaccagtaagataacttatggatatactgactactctgaa gagaaaagaatttggattagaaaagaatcatacagatgatgtatcgtatacaaaatc cggcctaagaatataataggaacagttatcagatataaacaagtagtaagaagaagaga tctttacatcaaatgactcctttcttagtaagaaggaatccgaattcgataagattaaaa aagaatactaaagtagttattgttagaaaacttaaatgggtgttgagagataaagttagata ggtaatacaagtaggatttatccggatttgccttccactttgatgtagtagatattaacg gaaataataaagattattaggaaggaaggtgatgaagatcagctaaaaatactcagtt gatattcgaataataaattggcaatgtgttttaattgttcataa	gacgtagctgatactattatgtgcttaggat cgactcaacctcaaaagtgttttaggcgtc gattgtttaaagctacttcaataagtagacta cgttaacaaagaaatgaagatacctacgaat gatgctctagtcgttagcaaaactatcgtagct ctgttaaacagttctcaagaggttagagaaca gtcaactacatgtaaaagctttgttaacattaa cgaaggggcaataacttctgaaaggaagat atactatatgtta
862-863	Ga0136617_10020436 GENOME_ACESSION: AntDrymUQ3222106_4_\$F_3300017789 GENOME_ID: 23166 CONTIG_ID: 20435 SOURCE: JGI DATE: 2019-01-27.	atgtcaaacacagctatttctatagacagtaataagacacccgtcaatccagttcatccagc acaagctcgcaagttactggattcggggaagccgctgtgttcagtttatccattcacgt tgattttgaagcgagtcattgaaatccaaatgtatatccgttaacactcaaaatagaccctg gttcaaaagttactggcattcgttagttacgaatcaaggcaatgtgtctgggcaatggag ttgcagcatcgcggacaacaataaaggaagctctctgtcatcgttagagcggtagctaga gggagcaggaaccgcaatactcgttatcgtccaagcagcattctcaatcgtaaaagacc agatgggtgggttagctccatctttaagacaccgcaattttgacaatagaacttgggtaaaac gactgcataaagtttgcaccacttgggtcaattgtcgaagacttgaagtgcaccccaa gcaatccagaatccagaatcctcgtgattgaataccagcaaggaaactttgaaggttatg agtcccgtagtacctgttgaaaaatggaatcgtcaatgcgcttattgtgtgtcaagga cgttccgctcgaaattgaacatccaacaaaatcacaagggtgttcagaccgtatttcta atcttgcgttggcttgcacaagtgcaatcaacgcaaaaggaaatagagataaaggactt cctcaaaaggtaagtcagatgtgtgaaccgctgtttgaacaagccaaaacgaccctaaa ggatgcggcatcagtttaactcaagtcgggtgggcaattgttcaatactttgaagcttttggatt gcctgtgagcatcggcactgggtgtcaaaacaaagttcaatcgattcgttttctgtggcca aaagctcattggattgatcggcttgcgttcccgctgtcaatgccataaactgtttacaa ccaagattctcaaggtaaaagcaacagggtttgtgtgtcggcagcgatgtaaacagaca agtttggctatccgcaaaaacatcgccgttgcgtccaattctgggattctgtacagcgcat atagttcgcgctaa	cggtaaaacctaccagagttggacatttatt gacaactctaccgaacgtgcagactaagcg aaagctcagatttttgagtcagacccaacg aatgaacggcagttgtgtctgtcgtgattgc attgttggcagcgattaacggttgattatcc gcgtgtggcagtcacaacaagctcttaaac attgtcaggcgaacttacttagcaataaag ttgattgaataaatcatgtcaaacac
864-865	Ga0376462_0006212 GENOME_ACESSION: Beaver3Rectum1_2_\$F_3300035507 GENOME_ID: 304486 CONTIG_ID: 6211	atggaccgaaaaggtcactactcggtagtagaaaggaaaaatattatggttaagctctag ataagagaaaaaacattaggttttaaacccaaaaagcaagaaagctcttagaatc taatagacagtagttcacagaatgtatccatttattataggattaaagatgtggatactag agactgcgatacaaaaagaatttagattaaaattgaccaggttagtaaatatatacagggtctt caattatagatgacttaggtaatgtatatttctgtcgtgaactgaacatagaggctctgtgt aaaagataagcttaaaactagagcaggtgctagagaataatagaagaagcagagaaact agatatagaagacctaagtttgaacaaaataggtgaagaaaatgcacaaatataatag	gtcaatgaacctcggctgaagccgaagggc atgtaatcgaaggttgataaattcgttttcc gggtgaattgagcagactttcgttatctagtt atgatggactgggtgtcttccagcacctat ttcatcggcgaatctaaacagggtctcgag ggaacttttacagactgtgtaaaagcaaac ccagtgattaggttaaccaattcgagagttat

SEQ ID NO:	Name	IscB sequence	ωRNAs
	SOURCE: JGI DATE: 2019-11-15.	actcagctagagaggaggttggtgctcctagcatatacaaaagcatagaggataacataa ttaatttcataaaaagtaactctaaatggatataattactctataaggttgaggaggttaact ttgatacccaaaaatagataatgacaaaattaaggttggtgagatcaacagggtacacttt ttggctatgagggtcaagaatatttattagataaatatgggcgttaattccaatattgtatga ttcaacagagagacgcaattgtcagggtgataacatattagaaatagaacatatgatacag tcaaaaatggtggttcaaacctgtttaaactgtactctagctgcagtaaatgtaacaaa agaaaggtagtcaaaacctagaagattggttaaagttctgaaaaatcaaaaccaaact gataaaaagagagctgaaaatatagcaaaaattcttaagatgggaagccttttagaggt atagatatgcatcatgtgctaatagctataaaaacaaagctaattaggaattaaaaatttag aaaacataaaatctgttgaatgtcatacgga	taaaaaagctgagataacattgtcaggaga gataaccttcgggttacgtaactatggaccg aaaggtcactactcggtagtagaaaggaaa aatattatg
866-867	TPA_asm: Spirochaetales bacterium UBA4673 UBA4673_contig_503, whole genome shotgun sequence GENOME_ACESSION: GCA_002405195.1_AS M240519v1_genomic GENOME_ID: 135224 CONTIG_ID: 68 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atggtttatgttcataattgttttggttaccgtcatgccatgcactgaggcgaaagctaaa catctgttaaaagcacataaagccaaggttgtaactttacaccgttcacaatacgcccttaa cttcgtctgtgacgataattactcagcctgtaacccctgggttagacgcaggctacaaaacta taggattatctgcattccactgaggataaagtaactctttgaaggagaggtaaaactctgtgag gatalaataatgccctctcgaaggagacgagcgagaaaaacccgatalcgtgctcctcgt tttgataacagagttcgtctcaaaaacagagggtgtgtgctcctctgttagaaaatcgtat cgggaacgcaactaagtgtttattgctaaagtttgccgtctcgtccctgtttctagacagaaa gagagcaactaggtttttggaatgttagagatgtgactatggcgagacggccataagtg cagaagtgtcttttgtaaaactaaagacacagttctcgaagtacacacacgttagtacaact aaagacgggtgttcagatagaccggataatctaataacactttgtaagacgtgccatgaag cttatcacagagagagattaaactagataaagccgaaaaaaggattcaaaaggcgcaacg ttcatgggaactatgagaaagcacttattaagcgtctaaagagccttatggctcgggga tggtagaagtaacttattgtttattaccaaacaactcgtattgagaatgacctgtataaag ctcactacatagacgctagatgtatagcggaccacccctaaagccactcctactaagactat ctacagtttacttaagcatcgtctctataacagacagatagatgacttacaatccttaagag tggcgtaaggaaaaataatcaatgtccaagaagatttcggcttaggctctttgacactgtt caatacaacagagaaaatagctctgtgagggagacgtactagtggaagttttagcattc gaaaaccgaatgtgagaaaggtttctgaagcgtga	taatgagtgaggatatcaacaaccacgggtg caagcactgggggtatgaggagtgaanaacg actcatgctctgttgattagcctaagtgctct atgcactactttattagagaatataaggtact gttgatgtcgtcctagctgacactctacgg ttcgtgtttaaagagtcacaaagggttagaga cagtgatagcaactcaaaccttaataacat tggcggaaggacacgggaatgaagactcc gttaccgaagaagacgctggccttaattgtc gtattgtctctctctattctcattcagggcat gagatttagagagtccttaagccgcatcacc ataggaggacagcatg
868-869	human gut metagenome genome assembly, contig: NODE_9257_length_400_9_cov_72.398331, whole genome shotgun sequence GENOME_ACESSION: OOBZ01.1 GENOME_ID: 8412 CONTIG_ID: 9256 SOURCE: WGS DATE: 2019-01-27.	atgatttacgtaagaagtaagaaggcatggttattaatgccacaagcgttgtgtaaga taggttatcttctcgtcagcggaaaggctcacgtagtcagccgtgttccattgtgtacagtt ggattacgatagtaccactacacgcaagatgtgagctctggcattgatcgtggctcaaa cacattggcgttctcgcaagttccgagaaagagagatgttcagcacaagtcgagtt agaagtgtatgtgaaattacttctactctgtagggaggtgagacggaatagaagaaacc gcaagacacgctaccgcaaggctcgttttgataatcgcaagaagaagatggttggctag caccaggtgtgaacaaaagattgagagccactgaaggttattcgtttgggtcataaacta cttctgtcacgaagaccacaatcgaggtcgtcagttgtgctcagaaaaatcaaaaacc cagacataaagggtgatgaatcagcaaggcgcaacagatgggattttggaacgtaagg gagtagctgtttggcaaggatggcgataaattgtgtcactgcaaggaagcaagagcaaga cccaattttgaacgttcacatttggagagccgcaagacaggcgcaattctcctagcaat ctcgtaacgctttgcgagacctgccataaagcttatcatcgtggagagttcgattgaaaatt aagcgagggaagttccttgcgtgat	tttaagttttggaatcctttaattaggttacgc gattatcattcaatcgtccgaagcggattag ccttagccccgaatggaattagggaactac gttaggaatgaatgcattagcagcagcaat gttcgtcaagttctgacctctcggtcgatg attaaaagaagcgaaaacacggtgtgtc gcaagaacacattctacatttgcagtg gcgacacacccactcgtgtggagatttat ttattaatttaattttagttatg
870-871	Ktedonobacterales bacterium SCAWS-G2 chromosome, complete genome GENOME_ACESSION: GCA_004208415.1_AS M420841v1_genomic GENOME_ID: 252008 CONTIG_ID: 0 SOURCE: NCBI_Prokaryotes DATE: 2019-05-30.	atgttcgtgtatgtttgaattgtcatgggaaaccatttatgccctgtcagccacacaaggca cggctgctcttgaaagcaggcgcaaggcaaacattgtacggatggaacccctaccattcaa ttgcgcatggtcaggttccgggtacaacaagagatatcactggcattgatgccggaagc agacacatcgaggtggcagcaacgacagagcagaccgttctctttgaagcggaggcaa agcctcgacggatattccagcaactgttggcgacacgtcagcaatttcgagcagccaga aggagccgcaagaccgctaccgacatccagatttcgcaatagaagaagccagatg gctggctgtcctcctcagtcacaacaccggctcgtgcccattctcaactcaatcaagatggt tcataagctcttgcagtgaaacaaacgacgattgaggtggtcagttcgatattccagaag atccagcatcctgagattgagggcatcactaccgcaaggccacaactgggcttctg gaatgtcgtgagtagctgtctgtggcgatgccaccgggtccagtggtgcccaggcca cgtcgcaagatgcagtagtcaatgtgcatcactcagagtcgaaagacgggtggggat cgtccagagaacctgattacctgtgtgagacctgtcacgattcattaccggaggcacc aagagcacaagatcaccggaagagcaacgggtttcgggatgcaaccagatgggcat cattgggggaagatctatgagcagcagctgctcagtttcccaacgtccatctgacatag ggtagatcacaagcagtcctggattgaacaaaattggagaaattctatctgtggatg cccgtgcatcagtaggaacccctggtgtgctcgatggcacccttacctgatgaagtt cgttcggcgcaataaccgcaactgcacaaagccaccatcagaaaagggggcaacag gcagcgcaataaccgctgccaagttggtcattgttttcgacttttgactgcgtgagctatc aaggagacgtcgtgtgttttggcgagaaagttcgggctactttgatttgcggctccctg gatgggaaagatccatgccagcgccagtcacaaacagttgaaacgagtcagaggg cttctgctgtttaatcgaagaggagcggcattctccctccctcaaggggcggtct cctgccgtgaaatcatga	gtcaatgacccccggagagatcggcgag gcttgaagacgttcattgactagcctgagt cctttactgactacgttcgagcgggtcacgat acctccaggtgtctacgtagcctggagccc tatcgctgtgattaaaagttctgattgggtag gaacgggtgtgagacccgacaagccgctt gaacattggcgaagcgtacatacatctgaa aggatgatgcagttatgttc

SEQ ID NO:	Name	IscB sequence	ωRNAs
872-873	Ga0180433_10023929 GENOME_ACESSION: SS_ome_13_SF_3300018 080 GENOME_ID: 35197 CONTIG_ID: 23928 SOURCE: JGI DATE: 2019-01-27.	atgcagaagttatcacaaagaatacatgcacaccaagatgctcaaccagctgtgtgc aactgtgattctgaattaacacagagaagaattctcagtgatcagaattaaaaactgtgat aacagctccgaggtgatcaagctcaacatacagagcggctcattacaacaaaagtgtat gtgcttaattgtcggggaaaaccactaatgccttgcaagtatgctaagtctaaagaatgtgt tagaaaagggtgcagcaaaaagtcattaaaagatcctctttacaattcaattaaatttgattgt gaaaacaaaagttcaagatgtaacattaggaattgatacagggtattccaatttgatttca gcagtatcagaaaaagacgaactgatatcaggagaaagtattatagaaaaatgggatgact aaactgtattcaagcaaaagctatgtaccgaaggaatagaagaatagattgttgatcaga gaaccaagatggaaaaatagatcagctaatataaaaaagaggttggttaccctccttacttt gagaagatttaaaactcatatttcttaattaaaaagattaaaaaattattaccgatttcaaaat tagaattgaaatcgcaatttcgatattcagaaaaattgaaaatcctgacattgagggaaacag gatatacaacagggtcattgtatcaatataagaatcgcatagcttatttaactactaggga aaaggaaaatgtcagttatgtataagaatataaaaaaggtgacgggtggcgactacat cataattgggggaaattaaaaatagaccacaagattggcctttagtcatgaattcagatcat aagaaattacacgcaaaacatgaagaatattgattacaacaaaagaaatcaaatcataca aagattctacgtttatgaattatttagaaaaagatttatagttgtttgaaattacatatgaa ataaacttttcaaaatagatgtatttagttatgaaaaatcctcatgtaattgatgtttgtat agctggaggagctaatcaaaaagatgttcacaaatttaaatagacagaaaaagaaaa ataacagggtgtctcaacttaacagaaaaggatttaaaccttctattagacgacaacgttatt ctttacacccaaaagatttagttaagattatggagaaatataatgaagttaaaggaaatcaca gttatgtgtcccggttaatttaaaaaatagtttggaaataataaacagtcagttaaaa aattagatgaatggaaatttcatcaaaaaactttaatatggaggacagctctga	gtcaactaccaaaggctaaagcctttggcct gaatgttaacattcaacgcaagagttgatta ggaggcataaaaattatgcagaagttatcac aaagaatacatgcacaccacagatgctc aacagctgtgtgcaactgtgattctgaattaa acagagaagaattctcagtgatcagaattta aaaactgtgtataacagctccgaggttgatc aagctcaacatacagagcggctcattacaacc aaaagt
874-875	NODE_5238_length_591 7_cov_2.09499 GENOME_ACESSION: mgm4784220.3 GENOME_ID: 238330 CONTIG_ID: 5237 SOURCE: MG-RAST DATE: 2019-01-27.	ttgagccttgacatcctctctgattctacatgattgcataggttatttctgcgagactcctaa gctccgtgcccatcgcccaaaaagcgttttctaatcgcaatcaaaaaaggttgcgcatct attaaagcctcgcggttgatggaaattcgattgttaaaagagcttgcctctatatttctgtca gttcaattgtttatgaagctgtcaagcgtgatgtagacaagacatcaggtagaaaaagcgc taagctgtgtaaaaggttctcctccgttatgctggcagatttggcgaattgctcagcttga ggcgattgcgcccgtcgtaaaacgtcaagcgtgcgcaaaaagacggttaaggaaacagc cagatcagaactcacttaggactgattaaagacaagcaaaaataaagctaaagcaagtcct gaaactcatgcagtagatggtattgctgtagctgcgacccagttcactcagtatcggattac tcacaagtttggcgaagattctggcgcactggatgggtgctgtagagattacgcccgcacc gtttcacgtcatactgcactgagttactcgtcgtccttgacattcgacaatgccgaca aggcgccgggaagcgaaagcgcaaaaggcgggtactgtgaatccgtttgatacagagcag gtgacaaaatgaaggtcaagacaaaggcggaagtcattacggcggtgggtggaggtttc accgataccgaaaaagtcgaaaaaggttagcgtttacgacaaaactggcatcgcttgg gcaggttggcatcaagcaaatcaaatatttggaggtctcaacaagctatgtgtagcttga	gcaaatatattcctgttctcagcaagctat aaaccgtccgggagctacgtgatgcccc tctctgcagcgtgctggcgagttgatgcgt actctgggttgcactcagcgacactcc ggatcgccctgggtaccgggttcaattagg ttgccaaatcttcaaaagcgtgtatctgtg gtgtgcaaacgtttccgtcccccgcaggtc tgcgaactgcaccaggtccagcccttcac atcggtgtgactcctttgctagctcttgcat ccccatcgca
876-877	human gut metagenome genome assembly, contig: NODE_6773_length_596 8_cov_3.051265, whole genome shotgun sequence GENOME_ACESSION: ORJR01.1 GENOME_ID: 10538 CONTIG_ID: 6772 SOURCE: WGS DATE: 2019-01-27.	atgcgaggagaaccgctgatccgtgctcaccggcaaaaggccggattttgctcaagca aaagaaggcggttgttcggcgagaactccgtttacgattcaactaccatcgctacggg tgaaccaaacaacctgtgactcttgggttgactgcggctcaagcacgtcggcctatca gcgagcaccgacaagaatgaggtgtacgcttcgaaagtcgagttgcgactgacatcgt tgatctgctctccaccaagcgtgaactgcgtgttcacgacgcaaccgcaagacacgcta tcgtgccccaaaggttgacaaccgggtccatagcaaaaacaaaggttggtgctccttcg gttgaaaaccggattcaagcacactgtcagagtgagtggtttgcgacttctgtccg gtcacgaaaatcattgtgcaaacagcgccgtttgataatcagaaaaatcaaaaatccggata ttgaaggcgttgggtaccagcaggagatcaactgggttctggaacgtgcgcgagtagc tgctctggcgtgacggacacgtctgccaacactgcaaggcggaagtcgaagatccgggtg ctcaacgtgacacatctggaagcagaagaccggaggtgacgctccaacaacctga ttacgctctgcgagcctgccaagcgatgcatcgccggggaatcacgctcaaggca aagcgccggaagtcctctctgagagacattcatgggagtgatgcttgggcgttttca atcgctcaagcaacctatccgagcttgaagtgaagaatctacggttatctcaccaa acacaagcgcatcatcatggcatcgccaagacgattgcgggacgccttctgcatcgc aggcaacctacggcaaaacgtgcgggaagtattttcagaagcaaacccgcaaac acaaccgacagattcacaactgacggtttgaaggcggtgttcgcaaaaagaaccaat ccccctacgaggtcaaaaggttccggttggtaaggtgatctgcaaaaggcgaaagg ctttatcttcggccgagagcgacggcgaggtttgatgcccgctctggacggcactc gcatttctcggggaatcagttgcaaaaagctcagactgttgtaaccgcgagaaacttttta acccaactgttaaggaggaagcggaattcctccccctgctga	gtcaacgaccccgccctgaagcccgagggc ttgaataaagcctgagttgactagcctcagtt tccgaaaacggaactacgttgggtgggaat acataggcaccgtgggatgtcaatcctagtc ccacgctctgcggtctgtgttaaaagctctg cgaggtaggagcgtgcccagatagaca aaaccccttcaacattggaaggattaca accggccttcggccgaggagataaaactt gaga
878-879	Ga0376486_000912 GENOME_ACESSION: B11_SRoWell667_FD_ _SF_3300035524 GENOME_ID: 311392 CONTIG_ID: 911 SOURCE: JGI DATE: 2020-04-07.	atgttagtatatgtattacatcaagatggaacaccattaatgccttgaaccagtaataagca agattattgtgaaaagtggttaagcaagagtggtggaagaacgccattcacaattaaat taacgatgataccactaaaaatactcaagaactaacattgggtattgatgtggatctaaa actattggcagtgcaagatgagatgatacaaaatagatatactattatcagaagttacagtt cgtaagatgttaagatcatatggaacaacgtagaatgtatagaagaatagacgaac aggaaaacaaaggtatcgcaaacctcgttttttaaatgttaaaaacttattcaagacaatcg ttatcgcctattataatctaaatataatagatcttaataaaaggttatgtttatataataaat attaccaataaccaatcttattattggaatgggaacctttgatccacatgctatgcatagacc	gtgaactccccaccgctaaagacgggtga gcttcaattttgaaagctctgttaccagac tcaatatctagaatagatactgattatcttg gatgtagataccttggattggcgccctcagatc cttgctctatctggcactgtaaacatcgctg aagggaaggcggaagtaaacacacctcga agcctagataacattgtcagagagatgctcg attctttatgtgaacgatagagatacgca

SEQ ID NO:	Name	IscB sequence	ωRNAs
		agagggttatgtggcatccatgtgttatataaaaaggattacaatttgggttcaataatcaaa ggcatatgtattatcaagagatcaatatacatgtcaatattgtaagaataagaataaagatcc acatttagaatacatcatattgtatataaatcacaaggtggttctgtagaccagacaattta ctactctttgtaagactgtcatgaaaaattgcataagaatcaaaftaaacttaataagtaa gtaagatcaacatttaaacatgcgcacacagatgaatgtattgcaagatgattagaaaat atataccagattatacagaacatattgatataactaagaacaattagacaattttttgct agaaaaagctcattgtatagatgctgtatgcgtagaagctactagtatatacaaccagag ttcttaccagatgctgtattttcaagaaatgtattagtaaaagtaattatcaattaaactaaag caaacattctgaaaaagaatgcctaagctaaattcaaggatttaaaagatgggatact gtattataataaacaccgtctgtttatcaaggtagaatgagtactggatgctgtattatg tgatattcttgtaataagtatagttttaaacctattcctcaaaatttaataaaatgaacggatat cagctagaaagtcagtgattatgatagagggaatcatgtaa	taaccctattaaattaggagattaattatcatgt ta
880-881	human gut metagenome genome assembly, contig: NODE_98_length_85073 _cov_10.354925, whole genome shotgun sequence GENOME_ACESSION: OIYE01.1 GENOME_ID: 6139 CONTIG_ID: 97 SOURCE: WGS DATE: 2019-01-27.	atggcatcttaccgaaaggaggatccaatgcccggagtatactatgtttaggcaggagac gggagtcgcagatgccgacacacgaaggagacatgtgcagaagtactggataccg gcaaacgcacgcacgcgaacatgtccctttaccatacagttactttatgacaataacccc gtccctccagcccgttactattgcagaagatccggggcggacgaatcgggatggcagc cctgtccctgaaagggaattactgttttctgcccttgcgcagacgagaataagagatc gcaaaactgatggagaccgcaggaagtaccgctgtccctccaggtctggagaaagg aaggcaagacagcggctgtgcgatacgggttggcagcggctttaaagccgggatgat gcggaagctgccaagatgcagcagacaaagttcatcactgtgaaggctatccggaaca cggaaatcccgttctgcaaccgggaacgggacgaagggtgtgttaccatccgtgcaaa ccatctgttgatcgcacatccatctgtccataagatgcagaagtttctgcccacacc gatgtggcactggaagtcaaccgggttctgttctgtctgtggaagatccgtccatctgtgt gtgacttccagaacggaccattgaaaggatcgaataccgggatccgcagctcatgac ctgcaggtggaaaatgtctgtctgtgaaaaagagattgagcattaccatcatatctcc ccaaaagcagggggcgttccaacacccttggaaacattgccggactatgtaaaagctgt catgacagggtacacaaagacacattgtatgccaacgactggaagacctgaaaaagg gactggacaaaaagtattgtccctgagtgctcctgaaccaggcgggtcccgttctctgcc agaagctggtacaggagtttggaaaagatcatgtcttttactgtaccggcaggataggg caagagtacgtaccgcccgtgggtaccagaagacaaaggaaaccagctccacgaagt ggaatcctgtgtattggacttctggcattggaaaaagttccagaagctgccaagcttcc catccccaccggatccctccagttccggagacaggacagaagttgatcgtatcacaggt ggaacggacatacaaaactggacggaaaaaccgttgcagaagaccggaagaacggac ggaacagaagaccgattccctggaagaatgtgttgaaaaacaggtaaaatcttctgggaa gaaggaaagccaaacgaagagaagccggctaaccgtgtgtaaaagcgtatccgccgtta caacgatccggaccgtctgatgccaggtgtgttctgtaccaggagtcgggtacgtc atgcgtggacggcactgcaaaaggccctacctgcaggctgtcggatgggaagtaaaag atttccagtaaaagcaatgtaagatcttaaaacaaaataccgggttagcttctgtgtctga	gtcaatcaccacgactaaagtcgcgggctt gcaatagcaggtctgtgtctttagccgctggg aattgagcggagacgtgatgtctgtgttccc gtaaggatggcactccgggatgttttcca gtcccgtgacctgcagagcagacaccaa gtctaggaagactgtttacagccatgcaca ggcttaccaccgtgtcggggatcccgctga tccccgtggactgcattggcatcttacgaaa ggaggcatccaatgccggagat
882-883	Ga0114934_10005458 GENOME_ACESSION: 4SBtaG_3_\$F_33000110 13 GENOME_ID: 21023 CONTIG_ID: 5457 SOURCE: JGI DATE: 2019-01-27.	atgaaagatgtaaaaaacagagtcctgtgattgatactaatcagccccatgactccag tacaccacgggttagctaggaagctaatacaaacaccgtaaacggctatcttccgtagac agcccttcaccatagctatgcgttccgagatgacgagaaacctaaagaacataccctta agatagaccctgtcttaaaacacaggtctagctgtatgatactacaactaacacagt agtatgggctgcaactctgttacaccaggagacggcgattaaagcccgactcatctcca gatctcaataaggagcactctgtagatctgtaagtcagggtatctaaaccagggtcga taacagagcaaaacaaaaggatggctaccgccgtctctaatgtccgaatccataaac tatgacatgggtcaacagactacgtaagtttgcacatacagaactgtcagcagaacg tgttaagttcgtatggccttaaggaagaccctaccatacaaggcagtgatatacaaca ggagatctctacagaaccaatctatgggagtacactaggagagagaccatcgtgcctgc cagtactgtgtgccaagaatgtaccactagaacgtgacacataaacctaaagtcaag gaggactagatgtcaaaagaaacttagtactagcatgtaaacatgtcaaccaagctaaag acaacacagatgttaggttcttcttaagaacagacaatgactcttaagaagctttaat cgttaagtcgtaaggttatcagatgcagctgtctgttaatgccacagtaacaaactcttac gggaactctgaatacaggacttccagtagagacaggaaacaggagcacaactaagctt aatagggtcagtcagggtctatcctaaagatcactggatagatgctgctgtgtaggtgact ctgggagagttgttactcttaatagctccatggcaccattagttaaaagccatgggtcac ggttaaccgtcaggtaactatactgacaaaataggcttccctaatagtaaacctaggagca ctaagaggtcctctctcagtaggttatgtcaagactggcgatatagtacaactagacgtt actactggttaagatcaagggtatgtcaaggatgtccagagataggatcagatataca tacaatataaaactcttattcatcatagtagacggacggtaaggcggttccagttaggc atgtcactactactactcttcatgatggatactatataaccacaggagtcgttag	tatacagtaagtaatatagtcattgttagtag ctactctactattagataagtggttgcctata ggactaactcaaccagctcgtcctaagg agcagacgatgtgataagatcatacctagg ggtgtccgccagctcctagctatgtgtgt gcgtgaatgacctaaggtagcactatagt aatcaggtacacataaccttcatgtctgag cgaggctaaccagctttacagtacacaacc cgaaaggagtttaacaatgaagatgtaa
884-885	Ga0377182_011632 GENOME_ACESSION: B610_RodWell662_FD_ \$F_3300036940 GENOME_ID: 311426	ttgagattagcaacaaagagaactacgtgtgtacacagagttaaagaccacacccggg ggtgcttctcagctccggcctcttgaagcgcaggcagcagacaagccgacgggttcg gcacgaacgggtctgtctgtgacccgggtgtgcaaatgtcccgaggggagagttcag gaaacccaagcctgaaccgtcaccgcgcgaagcggagcgcggcgttaagccgtgtatt gtattgggtaagaaccatcaaccttfaatgcccgtctatccggcacgggctaggaagctg	gtcaaaaaacccggcctaagggccgagg ctgttaagacccctcatgtgaccaggcaag tttttgagattagcaacaaagagagaactac gttgtacacagagttaaagaccaacccgg ggtgtcttctcagctccggcctctgaagcg

SEQ ID NO:	Name	IscB sequence	ωRNAs
	CONTIG_ID: 11631 SOURCE: JGI DATE: 2020-04-07.	cttcgacaggacgggcccgtgtgttgtaaacgcttccgttcacaaatccggcttaagac aggattggagggaaccagccatgacagtaaaagattgaccccgatcaagacta ctgggttgctgtgtctggaacagggtgaacacacacggtgtgacctgagcgaac taaccaccgagggctggaaatgacagaagatggggcagcgtgacgctgttcgccc tggaagacgaagcaggaatcttcgttaccgccagccgcggttcaacacccaagcagg cctaaaggctgggtgccccaagcctgcaaacgccgttgacaatactgtcttgggtca acaggattcgtcggtctgcccgtgtcttcgaccacagtagagcgtgtccgattcgacat gcagctcatggagaatccgaatatttcaggcgttaagtatcaacaaggaaacgctcttgggt acgaactcaggagtagatcttgagaagttcaatcgttctgtgtgtattgcaacggtctgt caaaagacccaatcctagaggtagaacatattgtgccgaaacccaacatgggga gcgcggctctaatcgggtttcaatctgacctggccctgtcgcacctgcaacaatcaaa gggaattgtcagcccgaagagtggtctaaaaattttgtccaaatcccgaagagaatcga aagcaacgacacaaggcgttgcctcaggttctggaaggcaaacggcccactctggcac atgctgctgtgtgaacgcgactcgaatagattgtttttgatctgctcaacacaggtctc ctgtcaggccagcacaggagggcaaaaaagtaaacgccattgttgcaaaatccc aaagcacattgtctgtatgcagctgtaccgggtcaagtttcaacactttatgatggaacca gcctgtctctgtatctggctatggccgggggaagctacaagcgaacccgtttaacaaa acatggatttctctgtgattcctgatcgccagaaagccgtcacggcttcaaacggc gatatgtcaaaagccgtgtgccaaggaagagcgtggaactatgtgtgctgggtt gctgttcgtgtcctccggtgttttaataataaaacgagactcagaccgttcagggaatacc gtggaagcattgccgtctctcgttaaggccgatggctacggttatgccgaacctcgtt cctctctacctgaaggaaaggcgtctccgcgaggaaaggatga	caggcagcagacaagccgaggggtcggc acgaacgggtctgtcggtatgccgggtg gcaacatccccgaggggagagttcaggaa acgccaagcctgaacctcacc
886-887	Ga0311301_10001791 GENOME_ACESSION: Sb_comMetaSPAdes_\$F_3300032160 GENOME_ID: 280981 CONTIG_ID: 1790 SOURCE: JGI DATE: 2019-06-13.	atgtcgcgcgtgtgtgtctgaccacagagggcgccattgatgccctgactccggct cgcgcacgcctgtgtctcaagcaacaaaggcagcagctgtggcgccgttaccggttac gctgatcctgcgggtcgcggcgtcagatgcccattgacagccgttgcgcctcaagatcga cccaggcagccgaaccagcggcctggcgtgacaaatgatgcgacgggagaagtggt atggcgcgcaagactcaccatcgccgcgagcaggttcacaaagccctccagaagc ggcagccgtgcgagggggcgcgcagccggcgcacacgcgataccgcaagccgcgt tttcaatcgccgacgcccgaagggtgtggttccccgtcgttctcagccgggtgcgca acatcgagacctgggtcgcccgcttgatgcgtgtgtgccactcggggcgttctgtacg aggtgtgtcgcttcgataccagcccttcgagaatccgaaatgcaagaaactgcctacc aacacggcacactggcgggttatgaagtaaaagtagctgtgtcctaaagtgggacac cgctgcgctactgcaagcaaacggcatagccctgcaggtcgagcactcgtgcccga aggcgccgggggagtaaacgcgtgagcaacctgacgttagcttgcgaacgtgcga ccggaagaagggaaccgcacttcggaggaattcggttctctatctcagccagg ccggcctccctcagagacgccggccgtcaatagcatcaggtggcgctgtacga acgctcaagtgcagcgggttcgcatcgagacgagcaggggtggcgccaccaagt gaaccgcacgcaacgcgacatccccaaagcactgggtgacgcggccaacgtcgg tcaatcgactcctccgcgactgtctgagcagcgtgcgaccgtgtctatcgggcaat gggacagcagagccgtcagatgtgcccagtgatgcgaaggcttccacgcacaaa ggccaagaaccgagcgcgaaacacgcgttccgacagcgcgataatgtgcgagcgt cgtcccgccgacactgaagaacaagggtgtgcatgtggggcgatggcagccagagc caatgtgtccttcagattgccagcggcgtgggacagtcaccgacattggctatcgt ctgcaccgcctgcagcgaatgacggcgtacgtatctgactcagagcataaggaggt gcagcttccctcggttccaatga	gtcacgtaccctatcgctttgaaacttcac atgcgatgggtgtggaagcaagtcct ggacgtgaccagcccaccaggtctgctctg gtaaccgtttccatgaatgcataggcactcc gcgatccgatccagtcgaggacgcagcg gtcaacgattaaacagacagacaggggtcg agtcatgtctgttgcaaccgaaggaaacc acgggaaaacagggcgaggtacacctca cttcacccggctcaggtatgaacgactctca ttacgagacccccgaagggtgaaagg aacgaacctatgctgcgc
888-889	BBAY79_10002962 GENOME_ACESSION: IMG_3300000980_\$F_3300000980 GENOME_ID: 27489 CONTIG_ID: 2961 SOURCE: JGI DATE: 2019-01-27.	atgtcaaaatfttattctgtaatagatacaacaaaaaggcactagaacctgctctctac catagctaaaaagctactgaaggcaggaagaaagctcggttttaggcaatatcttctacta tattcttaagaaggtttagagccagagagaataatcatgcaaaacttaattagacc aggctcaaaaacactggtatagctattctgcaagaagataaacttatctggcgacaga gttaactcatcgaggtcaaaagattaaagacatcttgaatctgtatgactactaaagcgtg gacgcagaggtgagaaaaactcgtatcgtcagcctagatttctgaatcgaagaaagaga aagcttggtagcaccagctggagcactgggttctgactatgactcgggttaagag attaatcaaatattgtccaattgattctatagctcaagagtttagcagattgatacccaaaa atgactgaacctgaaatcagtggaagaatfatcgaaggcactctttaccaatacagag gttagagaataatctttagagaaattcaatcgtacctgtgcttattgaggcgaagatact cccttagaagttgagcatacaagctaatacgaaggtggctctaatcagtagtataatct ggcaatagcttgtgtaccttgaatcaagctaatacctaatttgatattaaagatttctggctg gcaaacatctgtactcaagcgaatattagcccaagcctaaaggacatttgaagat	gtcaggaatctctgtagatagcagctccgtt aaaaagacggagcttgaaaacttagtacctg accagttcagctcttcaggactacatgtaa gagcaagagttaaagacctaccttgagtg cgtgccagcccaagctctagaaccgaaaa gttaaacagatgtatcagattaaatcagtg ttttgtagatttaccactcttaacatgacg aggcaaacataatcttactggagaaagac gcaaaaatgtcaacttt
890-891	Candidatus Diapherotrites archaeon isolate CSed10_239 CS- Sed10-C837, whole genome shotgun sequence GENOME_ACESSION:	atgtgaacaaaaatgcagaagtttaggaagagaatacatatactactgatactctc tagtcggtagctctatgtctgtcttaaacagagaggaaattcttagtgagcaagacttaa aaccttttctaacaactccgaagagacaaatcccaacatacagggaaatcagaacttgag agtagctgatatgtgtatgtgcaaaacatcaaggggcaacctttgatgcttgcacaacac aaaaagcaacaattactaaacaaggaaagcaaatagttatcaaaagaaaacattca ctatacaactaactattgcaacaggagaacccgcaaccaatcactctcgcatagatg caggaaagtaagtgattgggttctcagcgataactgagacacaagaattgtgagtggtg	gtcaatcaccctccctgaaggagggtt gagccgtgaggttgagagaattgggtgat taggagcgaagaatggaacaaaaatgat gaagttaggaaagagaatacatatacct actgatactctctagtcggtagctctatgtc ttgtctttaaacagagaggaaattcttagta gcaagacttaaaaaccttttcaaacactccg

SEQ ID NO:	Name	IscB sequence	ωRNAs
	GCA_003559395.1_AS M355939v1_genomic GENOME_ID: 245282 CONTIG_ID: 102 SOURCE: NCBI_Prokaryotes DATE: 2019-05-30.	aattagaactaagacaaaatgtgagtgaaaaactaatagagcgagcaatgtatagaagag gaagacgtaacaagtattgtatcgagaaccaagattcaacaaccgaaacgagaagaa gggttgctagcaccaggtatatacaataaacttgacacacaccaacattaattaacaaaa taacaaagtgttacctattaaacatagtcgtggaagttgcaaaagttgacacacagaaa ctacaaaatgcagacatagaaggagccgagtatacaaaagggtcaaatgagcgggtataat aaccttagagcattcatatttacaagagataaatacactgtcaaatatgcaaaagcgaagg aggaataacttgaacacaccatattattcagagaaaagatgggtgtagtaatagacccgat aacttagttacattacattttaaagtgccacaagattttcactcaggaaaaatcaaacacaa atttactaaacaaaaaagtttcaagacacaaagcgttatgaacaatgtttgacaagattag ttgataaaaaaacttacgaacacacgtttgtttacataacaaaagaaaacagacaaaaact tgaattagaaaaatcacacgtaaacgacgcattcatatagctgtgtggaagtatacaaca agatgtataatcagtaatttgaagcaataaagaagaacaatcgttcctgcaacttaaca ggaaagggtttaaaccaagtattagaaggcagagatacaaaatacaaatcaaatgatttagt taattttgacaacaaaattatacatataaggtgtgtttacaaaggaaacttggtttaaatta acagatgatttaggaatacaattaacaaaaattataatgtgtagaggaacaaaatagg aaaaggattacaataatga	aagagaaccattcccaacatacagggaatc agaacttgagagtatctgatatt
892-893	Nocardiopsis sp. CNR-923 NODE_20_length_84413_cov_159.688_ID_35, whole genome shotgun sequence GENOME_ACESSION: GCA_001942255.1_AS M194225v1_genomic GENOME_ID: 114820 CONTIG_ID: 65 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atggctacgttccgcacaggacgaaaggcccccctggcgtgcttctcagcgccagg ccctggaatccgcgacagctgacactccccggatcgaggaaacgaaggcagggag accaccggcaggtgtgaccgggtgcggaacatgtgcgaggggagaccgggtaccgct cacctggcgggttcgggtgacccccgagccctcgccggggcgaggagaaggccgtg aggccacacacccacccctacgtgttcctcagacaacacggcacacccctgcag ccctgccccggcgacggcccgcatctcctggcgaaggccggggcgccgtccac cggcacaccccttcacatccgactcaagaccgcacccgcccggactccagatcg acggcgtcgagatcggtgtgcaccccggtccaaagaacaccggcgtccgtgttcac ggaaagaagccggcagcgccggggccgtgacagatccagctcgaccaccgcccgcg cgaccatccgaagaagatggggcagcggttcggctaccgagcgccgcccggcgtg cgaacctgcggccaccgcccggccgattcaacaaccgcacccgcccgaagggtgtgt ggccccgtccctgcgacaccgaccggacaccacatgggggtgggtgcacggcgtgac ccgactcgtcccgtgcgctggcgacgaggtggaaccggaccgcaaccagctgc ccaagtcacacacccgtgacgcgggtggcggcgaaggcgtgagggatcaccgaga ccgtggacacggtcctgtgcgggggtgcacggccgtggatgtatgccgttaccg caccgaccggcaggggttccccggctgcgggtgcggctgaccacgggttctcgggt cgccaccggcgacgtgtccgcgcccgtgcccaccgggaagaaggccggaaacc ataccggccgggtgcggtgcgcctcggggagcttaacgtcaccaccgcccgcg gcacctgcaggggcgtcaaccacaggcacgtccgctgctcagcgagccgacggct acgcctacaccacccgggaaggaggaaggcgtttctccccggccctga	gtaaaggatcctggctcgtgaaagaccggg ctttcaaccggcgtgggtgaggggccgttt accagaccagccatcaccatgaggaggt gacctcgtatggtcagcttccgcacaggacg aaaggcccccctggcgtgtctcctcagcg ccaggccctggaatccgcgccagctgaca ctccccggatcggaacgaacaggggcac ggagaccaccggcaggtgtgaccgggtgc ggaaatggtgcaggggagaccgggtaccg cctcactcgccgttccgggtgcacccccga
894-895	Streptosporangium sp. 'caverna' chromosome GENOME_ACESSION: GCA_003203775.1_AS M320377v1_genomic GENOME_ID: 162526 CONTIG_ID: 0 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgactacgtttctgtgagtgagaagaccacaaagccgtgtctcagcggtgtgctc tggaaatcggtgagagcagacaaccccgagggtggggacgaacggctcagcacc cggtgtgcccggcactggcatggaacatggggcagggggagatcggttcgggcggca cccgcctggaccggcgtcaccggcgggtgtctcggcgtctgaggggagcggacgtg agggtcaccggccgtgttcgtctggagcgcacacggccaatcgctggtatcggtccac ccggccgtgcccgtgcgctgctgcggcggtgctggcggtgtgtggcccggcacacc ccgttcgtcataaggctcaaaagccgacggtgagccgactccgcatcagggtgtgca ggtcagcatcgaccggcgagcaaacacaccggcatcgcatattcaccgaatacggc gggagccggaccggcgtgtacagcgtgcagctgaccaccgggggtgcgagatccga gacaagctcacctcgcggccgattgcggggcgccggcgtgcggaaactcgcg tatcgcgcgccccggttaacaaccggacacggccgaagggtgtgctcgcgcgtcgc ttcgccaccgctggacaccacatgtcctgggtgtcgcgctgaccgggtggcgtccg gtcaccgccatccacgtggagaaggctgccttcgacacgcacgccctgtcggcgggtc atccgctcgaaggatcgcaataaccagaggccacccctgcgggatacagggtgcgcga gtacctgctggagaagtggggtgcacctgcgcctactgcggcgccagaacgtcccc ctgaacatcgatcacctccaccgcgctcccggggcggtctgaccggatcagtaacct cgtcctggcgtgcatcccggtgaaccaggcgaagaacgccactccgatcaggagttc ctgaagagcaggccgcgctcctggcgaagatcttaagcggcggaaggcgtcgttgc gagacgcggcgccggtcaacgccacgaggtggcggtgtgtgcccgggttgacgcg accggcctatccgtgactacggcctcgggcggcgacgaagtgaaccgctcgcgc actggcgcggccgaagtcgcacacgctcgtatgcgctgcacgtcggtgacctgagacgg tgactcctggccgtccatggtgtgtgtggaaggcagccggcgcgccacatttc cgtaccggcaccgacgctacggtttccaggtcgcgctgccccggatcaagcagg tcaagggttcaccaccggcgacgtgtcgcgcgaacgtaccgaatggcaagaagac cgggggtcacaccggcgtggcgtccgctccaccgggaagattcaatatcaccacgc ggcatggcaccgtccaaggaatcgccaccggccacattcggtactccaacgagccga cggctacggctacaccactcaatcagacggccgactattcgtgtttctccccggcct gaaggaccgggttcacgctggagggtatcgatgacactccggcatctccccatga	gtcaacgacccggcctgaaggccagggc ctgaggtgcccgtgtgacaccatggccc cgcgttgaccagcccgatcagctcga ggaggtgcttttgatgactacgtttcgtg agtgaagaagaccacaaagcgtgtcct cagcggctgtcctggaatcggtgagagca gacaaccccgagggtggggacgaacgg ctcactgaccccggtgtgcccggcactg gcatggaacatggcgaggggagatcggt tcggcgccgaccccgctgaccggcgctca ccccga

SEQ ID NO:	Name	IscB sequence	ωRNAs
896-897	Ga0307928_10027040 GENOME_ACESSION: AceLakmetagen232_2_\$ F_3300031227 GENOME_ID: 279923 CONTIG_ID: 27039 SOURCE: JGI DATE: 2019-06-13.	atggttaacttttttattgatggaagaataacatactgcacccaacaaagaatcagacat gatttatcgctggctgcgtcaaggtaaaagcaaaagtactgaaaggtgcttaaacagg acaaccattattagtagcaggtatttaaaacatttatgaaaccaacaaaatgtaactgtgaatt cagaatcggaattgatcctggatacaaacacattggttactgtattataaaatagatgtgga gaaacagacaataattgcattatttctggagaagttgaaacaagaactcagaaattaca aaaaatgttagaacgaaagatgtatcgccaaaatagagacataacagaagaaaaaat gttaagagaaaaatttaactatgtaagtttcgaaacactgtatggaaaaacagagcaaac acaagtttcaaccaacacactggcatttgtaactctcacacaactctttaaagtgatct tcgacagaattccattgaaacagagtaaaattacatagagtacaatacattgatatacata aagttatcaatccaagcatttacaatggcaatacagaaaggtccacagtagcaggttga aaatgttaaatcatatgtacgatacagacaattataaatgtcaaatatgtaacaaaatgtt ggaaaaaaatgaatcatgtacatcacattatacagaatgatggctaaatgatcgcc ctgagaatttgattctattgtgtaactaatgtcatgtctgttcatgctggaagagttgcatgt ccaatatcaacagcaaaaagttcagagacatgggtgtattaaattcctgtatgaaatatta tttgaagagtagtaaaatgtaattcagtaagaatattatggacacattacaaaaatgttc gaaagaaatgtgaattgaaaaatcccatgcaaatgatgcgaaagttaacgctgtatgtga ttcaaatgggttaccgaagagtcagagagtagattggagtgattctaattgtataaat ttcaacacagagcagaaggcatgttcgaaattgggtacagagatagggatcgaaagtat tacatgattggaaatccatttgcgatgcaggaatagaagaaacgatctggacaggag aaaatgagtttgaagagtttcgaaagttatatccgaaagaacaactgaatgcgaaacctg ggagacaataatcgaaaaaataacagaataatcttattcaaacctgggtatataaattat gttcagaaggaggtgatactattaaaggctggcactcacacaataaagttgttggaga acgccttggacgaattcgacaaggagattgtgaaaaaggtttaaacagttgtggaatgtg attgtttaa	tcaagtaccggggcgcagagcaccggctt gggatagcgcagagcgcttatgtacgc cgagaactttgcttctgggcttgacagaca ctctgaccagatgttaccctgagataccact cccagttcaggctctggaggtaaccagtc ttattgctgaccagtaggttaacgaagatcgt attacatctgccaagtcgaggagaccac gtctcataaggagatccgaccaatggtaa cttttctattgatggaagaataacatactgc acccaacaaagaatcagacatgattatcgc ctg
898-899	Ga0209749_1010998 GENOME_ACESSION: EPASupSA47SPAdes_\$F _3300027958 GENOME_ID: 25534 CONTIG_ID: 10997 SOURCE: JGI DATE: 2019-01-27.	atgagagtactggtttatcaagcataaaggatccctcatgcccctgccatccggcacggg ccaggagactactaaaaaggacaggctgcagattcaggcagtagccctttacaatta tccttaaaaacaggcaggcaggtatttgcagccctttgaactgaaatgatccaggcag taaggctacaggtatgacgctgttctgctgattgcaaaaagggtatgggtcttctggccct ctgaagtaactcaccgggtcagggcagtcagagtagtcttgcggaacgcagggcaat acgaaaaaccgccgaaccgcaaaaccggatccagagctccccgttcttaaacccga caaggcctgaagcgtggttgcaccatcgcttatgagtcgtgtgtacaatatgcaaacatg ggcagcacgcctataaaagtgtgcccgttaacagacattcatctgaactcaataaatttg acaccagcttatggagaatcctgagatcacaggtgttgatgaccagcaggcactctgt tcgggtatgaaatcagggaatatcctctgagaagtggggagaaaaatgtgtattgtg taaggaacacgtaccccttgagattgagcatattataccagatcaagaggcggtagcag tcgaataagtaattgcccctgttgaataatgtaacaaagataaaggcaacatgaca ccggaagaatttggttatcccgaagtcagaaagcagcgaagctgccgtaagggaatgc tgacgctatgaacgcaacacggtagctataggcaggggcgctcaagcaaacaggcttta tgtggcttctggtctggaggaagaacaaaattcaaccggattagtcaaaactaccgaaa gagcactggatagacgctgcatgtatcggagagacagagagagggttattatccctgc cgggacaaaacctttaaattactgcagccggccatggcaacagcgaatgtcggga caataaataatgtttccgatacggcacaggacaagataaaaagtattttgttttga ccggagacatagcaaaagcaatagtcgcaaaaggaaaatgtcgggtacttataggtga gcatagccataagagcatcaggatatttgaataaagaatggagcgggaaagcgaaattg tacaaggaaattctcataaataattttaaaccagtcagcatttggatgatacagatatgaga caggcaatgcattctcccggcaataaattccggggcttctgcgtggaaggatcgtga	gtcaactaccgggttttaaaaccgggcat gaaagagagttcacataattctcttcatgcc catgttgaccaggtacagccacggatcataa ggctctatggggctacgttattccggctatgac accctgaagtgtctgcccagctttaggctctgt cgtcaggcattaaacaggcataagggggtta gaccagtgtgtgtacacaacaagccggaa taacattaccgaggcaaacattaccggcgca agccgagaacaaggagtaattctttattgag a
900-901	Ga0376086_0000584 GENOME_ACESSION: LAC_S2_MG_FICUS_ FD_\$F_3300035686 GENOME_ID: 312530 CONTIG_ID: 583 SOURCE: JGI DATE: 2020-04-07.	atgtccaatgttttctttaaaccacaaataaacagcttcttaaccctatccatcaggctgg gcaagaatgctgttctcagggaaaagccgctgttttgcagctatccctcacccttaata ctgaaagaggaggtgttaaatcctcaagttgaaccactccgcatcaagatcgaccttggc gctaaaaccagcggactagcaatagtcacgattgcacaggcgaagttgttgggtagcc gaattgcagcatcgcggttttcaatccgggatgctttaacctcacggctcaactccgac gggtcacgtcgcaatcgcaaaactcgtaccgtcaaccacgcttctcaatcgacccgac caaaaggatgtgtaccacgtccctcaacagtcgagttgtaaacctgacttgggtcaa gcgctgtctgcgtatgccaaattacagccatatctcaagagttggttaagggttgacacc cagcagatggaaaaatgcagaattagcgcaattgggtaccaacaaggcagcttggctgg atacgaatccgtgagtttctgcttgagaatgggtagagcctgtgcttactcggcgca aaagacaccaagcttgaaattgagcatatccaacctcgggtcaaaaggcggttccaatcgt gtcagtaattgtgtctagcttgttcttgcgaaccagaaaaaaggtaaccaggatattaag gagtttcccaagggttaagtcagagctactcaagcgaatttggcacaagctaaagaccc ttagctgatacagcggctgttaattgcaaccgatggaattgtatgaaaactcaagaaaa caggcttacctgttgaggcgggcacaggtgtgcaccaatacaaccgaagcttgcga gggttagagaaaaccattattgggatcggttgcgttggggttcgacaccagaacag ttaatcacatcagggttaagccaatactcattgctgccaagggtcatggaaccctgaac	gtaaactaccggccaccaagctgagtagacag ctatgttgggggttttaaaaaaagtcacag agtttaaccgactaagtagcttgtgtactacgt tttagaagcccatcacacctgggatacgaag ccagtcctctgctgtggtgacgattaaa caggcatatcgggttgaagccaggtgtcgtca gcgcgacaaccttcaaaacattgtcaaggc taacattaccgcaaggaggacaggaatg tccaat

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aatgccgcactgacaaatgggggttccgggtcggttactgctctagaaccaagtttcacaaagggtttcagacgggtgacatcgttaaagcagttgttacctctgggaaaaagattggagttacgttggctgtagccactcgtgccacgggtagtttcaatatctcgacacctgacgggctaactcaaggcatcagcccaagttactgcatacacattcacaaaaaggatgggtattcctatgtatag	
902-903	Synergistales bacterium isolate INTA.AUR.055 contig-100_2301, whole genome shotgun sequence GENOME_ACESSION: GCA_009929915.1_AS M992991v1_genomic GENOME_ID: 322850 CONTIG_ID: 232 SOURCE: NCBI_Prokaryotes DATE: 2020-04-22.	ttgaccagcggggcctgggtccaggaccaggcagacgttgcacacaggttcaagacccaccgtagagcgcttctcactctacgctctggaagcccaagcagcagacaaccaacgggtgggtacgaacgggttggcgcaacgccgatgtgcaacagccggaggggagattccgcactatgtcggaaagggttacggagacggatgctttatcctgcaacggattacggaaacctttacgggtccacctgtttggaggtattgtcatgtatgtgttcgtactggacaagcacaagaagccctgacaccgtgcccgcgtcaagggcgaggtatctgctggctcggggcgggcggtctccacaagcgtttccctttaccatccgtctcaaggaccggacggcggaagagagcgttatcaaacctgtcgtgtcaagggtgacccggggcgccgggttaccggaaatgccattgtccgggagacaggcaaggtggaccccgctctgattgccggaatcgaactggaaacacaggggaacgctatagagacaacatgacgaagcggcggggctaccgccggagacggcgagcgcgaacaccgcctatctgccccgagattcgacaacagcggaagccggaagccgggtttccggcgagcgtggcaccggatcgacacagaccgtctgtggatgcggcggttacacgcatagcgcctgtctccgattcagcgtcgatcgaatcggtcaagttgatacgcagaagatgtggacccggagggttccggaaagaggtaccagcaggagagcttgagggtacagaggtcgggaatacctgtctggaaaagtgtgtcgaagtgcgcctattgcaatgccgggaacgttccgtcgaggtggagcacatctaccgagagcgagaggcggttcggacaggggtgtcgaacctgacgtcggcgtgcgagagggtcaatcgggccaagggagcgcgtcctgttaagaggtttctgcatgacaaccggcgcttctggacggatagcggcgacgcgaaagctccgctctcccgccgtcgggtcaactgacccgcaacgcactctttggcgagatgctgctgtttggctggcgtggagacggggagcggcgactaccaagtacaaccgacccgcctcgggctcccgaagagtcattgtctcgtatgctgtgtgtcggacagtctctcagcaaaaagttaacagacagcgtctccatgttgcgtgtacagagcgggacgatactacgaacactgacggacaatacggattcccagagcatacctccctcggggtaagcgctttttcggattcggcaccgggtgatactgccggacggcggttccgaaaggaaagtacaagggaacgtggacagggcggggtggctgtacgggaatcgggtgttgcactgtccacagggaataacgcagacggcaaaaagagagggttaatgtcaaatggatactgcataatactggaacgaacaatgggtacgaatacagcgttaatagcgggtgtag	gtcaattaccggccacctaacgatgtaaacgcctcggtttgaggtgggggcttgaanaagccagattgaccagcggcgctgtgtccaggaccagcagacgttgcacacaggttcaagaccaccgtagagcgcttccctcaactctacgctctggaagcccaagcagcagacaaccaacacggcgtacagcgcttccctcaactctacgctctggaagcccaagcagcagacaaccaacgggtgggtacgaacgggttggcgcaacgccgatgtgcaacagccggaggggagattccgcactatgtcggaaagggttacggagacggatgctttatcctgcaacggattacggaaacctttacgggtccacctgtttggaggtattgtcatgtatgtgttcgtactggacaagcacaagaagccctgacaccgtgcccgcgtcaagggcgaggtatctgctggctcggggcgggcggtctccacaagcgtttccctttaccatccgtctcaaggaccggacggcggaagagagcgttatcaaacctgtcgtgtcaagggtgacccggggcgccgggttaccggaaatgccattgtccgggagacaggcaaggtggaccccgctctgattgccggaatcgaactggaaacacaggggaacgctatagagacaacatgacgaagcggcggggctaccgccggagacggcgagcgcgaacaccgcctatctgccccgagattcgacaacagcggaagccggaagccgggtttccggcgagcgtggcaccggatcgacacagaccgtctgtggatgcggcggttacacgcatagcgcctgtctccgattcagcgtcgatcgaatcggtcaagttgatacgcagaagatgtggacccggagggttccggaaagaggtaccagcaggagagcttgagggtacagaggtcgggaatacctgtctggaaaagtgtgtcgaagtgcgcctattgcaatgccgggaacgttccgtcgaggtggagcacatctaccgagagcgagaggcggttcggacaggggtgtcgaacctgacgtcggcgtgcgagagggtcaatcgggccaagggagcgcgtcctgttaagaggtttctgcatgacaaccggcgcttctggacggatagcggcgacgcgaaagctccgctctcccgccgtcgggtcaactgacccgcaacgcactctttggcgagatgctgctgtttggctggcgtggagacggggagcggcgactaccaagtacaaccgacccgcctcgggctcccgaagagtcattgtctcgtatgctgtgtgtcggacagtctctcagcaaaaagttaacagacagcgtctccatgttgcgtgtacagagcgggacgatactacgaacactgacggacaatacggattcccagagcatacctccctcggggtaagcgctttttcggattcggcaccgggtgatactgccggacggcggttccgaaaggaaagtacaagggaacgtggacagggcggggtggctgtacgggaatcgggtgttgcactgtccacagggaataacgcagacggcaaaaagagagggttaatgtcaaatggatactgcataatactggaacgaacaatgggtacgaatacagcgttaatagcgggtgtag
904-905	Ga0310695_10017382 GENOME_ACESSION: RumRJGDNAv2_SF_330 0032007 GENOME_ID: 280889 CONTIG_ID: 17381 SOURCE: JGI DATE: 2019-06-13.	atgacagtattcgtgagcagcaaaagacgggtacaccccttatgccgactttcaatatcaaaaagggtacggcgtatgttaaaagacggctgtgccaaaatctacgggtatagcccgttcacaaatacagttacagtatgagcagcagcactacacgaagaaatgaacctgtgaggacacaggtgaccattatgtaggattttactaaaatcaagaacatgaatatgtttcagggcagttatgaccattgacggatgagaaactgcgccagatgatccccgtaacctccgagagcaggcgaatcataagcgtaccgcaaacggaggttgataacaggagacacctatgccggaagcgataaatgtttgtccatccataaggaataagatcaacaacattatccatagctgcaaaagtataaagcgtccccataaaggatatttacttgaatgcggctccttgacacagacctacagcggtagagggagggcttctgctcctaaaggtaaagtatttcaaaagggttcaagatatgtttacgatactttaagggaagcagttatgccgtgacggatacaggtgtatgtgtcggtaaggtatagaggacgtgcgtgactcgcctccaccatttaggtataagacagggcaccacacacacgcgtgctggggctgtgactgtatgcacaaaatgcacacaccggcagcacataaacctgcccggaaaactttatgaacttgagccgaaactcaaacctttcaagctcgaggtttatgacatgatgaggttccagctcataaaagacgccaaggtttgttccaacacagagggtacacatatgtctacgggcataatacaaacgcgaaagattgtcgcgtcgaatcagtaaatccacgccaatgacgcttattgcataggtttttcagatcgcgcacaggacggataaagcattataaaaactccgcccgaacacccgcatccttgaataattttatgacgcaagggttatagatgtaagggtgtacagcacataaagggtcggaattatcctgtaaccgcacaaaccgaagtatttcaaggataaacgaaacaatctccgtatataccggggaataaaagtgtccaaggacacagggaacataagaacaaaaagatacggcatcgtccccgtgatatgttctttaccaaggcaaacgatataagtcgaagggttgcagcatatggcgcaatacaccatttggaaaggcacaagcggtcaaaagtaaaagacgtaaaaatcattaaacacactggcggtggagacggcagcatga	gtcaataacccccatctaaatgcattgcattatagtgagccttgaanaaaggcttcgcgcctgacttttcaacgctgattgattagcctaagtcctgaagcttccggaactacggactacgttaccgggtgaattcctgggagggcgaccgtttgtaccaggttaaacaggcactcgggggtactccacaagccccggcaccacttaggtggccggggaatgaatacaatacacgacactcgggtctgtatttaaacatccgtgaacagcggatgtcctgagggtagggtaagtatacagacaataaaaccaccggataacattggcgatgtgacccgcccttttggcgagaacccgggatgctttatatccgggaagccgtaaggcaactgttttaagaaggaggcataagaataatttatgaca
906-907	Ga0197142_1000319 GENOME_ACESSION: GBS60_MetaG_2_SF_33 00020153 GENOME_ID: 26410 CONTIG_ID: 318 SOURCE: JGI DATE: 2019-01-27.	atgtttgttctgtatgttgacgcgacccaaagccactcatgctaccactcccagcaggccaagacgtggatataatctggcaaggcaactccttctggaagaaggcggtgttttgcgttaggtcctaatgttgacccgtctgtcgtcctcaactcagccgattgccgtggcgattgacctggcagtaagaagaaggctggagcgtcaagtcctcaagccatacctacctaacaatccaagcggatgcagtaactgtgggtgaagaccacatggaaacacggcgaatgcgtcgtactgcaggtgtacgcaaacacctgcccgtcaacggagagccaatcgagctaggggtgatttccgcttaccaaagcgaggtggcagtggaacatctgtctggcaagctgggtgacaaattgttccccatcactacctttgtgttagagacatcaaggcgaagacgaaggtagcggcggtggatgttcttctcttggaaagcagggaacaatgggtttaccagca	gtcaagtacccccagctaaagctgggggcttgcgggagtagagcctagcagctggaacccgcaaggcgcaagaaagactagcctcccaagcaacctgttctacaggggcggttgaaagggaactgaaccaatgcttccctagctgcaaaaagacgtagcgaatgtcttccctagttgtctacctctctgttggctagtggaagggtgtacacactccgaaaggaggcttacttaacaat

209

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gtgacagaagatatcagacacattgaaatcttatcaacagactgaagaagattcttccaat aacacagttattattgaacagctaaattgataatcaaaaaatgaatcctgaaattgag ggcattgggtatcagcaagggtctttatgaatacaaaacatgagagcttatcttatggctc gtgaacatgaaaaatgctgattatggtgaaagtttaaaagaaaaatcttcacataacatc attgtaagcaaaaggcagaatcagggaagtaacagagcagagaatttagctatactcaca aagattgccatacaaaatcacataagcaaggctgaaattatcaaaaccaaagagtataag cctaatactttatgctcataaatacaaaaaagtttagacaagaataatccagatgtaaaaaat cttatggaacatcacattttgtaataaggataaatttaggattagaaaaagccatgcaac gatgcttttgtaattgcaaatagaactatgcaggaaagaattatccttggaacatagaaca aaaacacagacataatagagcgatacagcttaacaggaaggttttaaaccttctataaga acatctgtgtacaagattcagccaaaggattgtagcagttaatggacaaggttttctgtt gtaggaaatcacagaacaaagggtcttatgtaaaagttaaagattaccacaaagtcattcctac aaaaataatagagatgtatataattcggagggttagtatggaactaa	
916-917	Ga0310696_10000167 GENOME_ACESSION: AusTroGonzalMGv2_2_-\$F_3300031993 GENOME_ID: 279930 CONTIG_ID: 166 SOURCE: JGI DATE: 2019-06-13.	atgggttatgtgtgaacaagtcgggaaccccaatgccaactagacggtatgtaagg ttcgccgtcttatcaaaacccgattagctatgttaattcagcacaacccgtcactatcaaat attatatgatacttcaacaaggtacaaattgttaattcaggtgttgatgcgggaagtaaacat gttggtctttcggcatctacaacacaacagattgttgaagctcaactagatttacgaacc gacatcactaaaaaafagcaactgcgaacaacaaatcgtattggacgacgctaccacaaa acacggtatcgagaagctagatttaataatcgtatcaaatcaaacgacataattgggtag cgccatcaataaaaataaattcgatagccatataatgtggaattactgaatatgtcttttta ccaatataaaaaataaattgaagtaggacaattgatactcaattattaaaagcaaatgat aataatgaactgctcccaggggacgtgactatcaaacctggtgtgcaactcgccttttggga atacagcgcgaatatataattacgcgcgaacatcacactgtcaatgttgcgaagggaagaa gtggcgattcgatactgaacattcaccattgaaagccgcaagattggcggaaatcgacc agataatttaattaccttatgtaaaacatgccatcgcaactatcacgctgattttaaaattac ctagtcgaataaaacgacccaatcatatcgggatgccgtacataatgaatatgtttcgaaa ccacatgtataacaattaaaacaaatgcttgaaccaattattcaagtaaatgtaacatagg atataactaagcattgccgcatatgaataacacacataataatagatgctag atgcattagcgaacacccattagctaaaccgagttcaacataattggcgtatcaaaaaacga agaacacataatagacaattgcacaaatcaccattcttaagggtgggttaagaaaaata atcaagcaccatagaaagtatttgatttcgattattgacgctgttaatatcaccaatgac ttgtttgtaacgggacgccgttaaccgggttattgcaataaacaatattgataaaactata cgggacaattcaattagttataaaaaattctcatatgtatcacactaattctaacctaatgg aggaaatgatgtatcattcatctaataccaactaa	ataacggcccatagatgtataatctaaaaat gaacttgaaactcttgctctggaacactttt cgttgattagactcagtgaaagcataataat gctgaactatgttaaaagaccatgacaccc tgagataattggccaagtctctcgccctgtcg aacgatattaaacacccagatgggttaagg gtagtgtatcgttcataaaaaatcttttaacctt gtcgtatgtcaccacagggaattacccctg cattaccattaaattaggataaaagtagatg cgaaatg
918-919	Ga0257048_100529 GENOME_ACESSION: IMG_3300023543_\$F_3300023543 GENOME_ID: 309626 CONTIG_ID: 528 SOURCE: JGI DATE: 2020-04-07.	atgaaacagataaattgtactttgtatcggcgaatggtcttggccttatgcctacaaca ccacaaaaggctcgaaattctttgaaagaaaagaaagcctttgtaaaagaaaagtaccatt cacgatacaactatctacaagacaggaagtgcaacacaacaaagattgggaatcg ataccggttctcaaaatagagattggcattactgacaatcaagaaaattgtcattctaat cagaatataagttgcttcaacaaatggaanaagacagcctatagaacaagaaagactt atcgcaaggaagacgatatcgaaaagtaagatcggaaacccctggctataagcctcata caaaatatactatgtaaaatgctcgtaaaagaaatggttaaatgacacactggatgaa gtcgaaaaatgaatatggaactaacagatgcgaagatggtctaccacctctgttcagca aaaagttgaccatataatcaattattgacagatactggaagctttaccaataatacaaaa ctgacaatagaattggacgttttgatgacacaatcaaaaatcccaattttcaaatgaa gaatatcaacaggggacgcatgtatggctatgaaaatgaaggcatatgtacttcataagt atggctatagatgccctgtatgtggacaaaagtttggaaactgaaagaaaaataagacgat tgttctacctcaaatgcatcataagcatttcaagagcaaaaggactacagataacccggac gagtatatgcctgtgtgtgataaatgtcataatgcagcagctcataaaaatggtggagttct cgatgctataagaaaagctgaagctaaaaacattcgaggtatgagagatatgacacttatg aatatcgttgcctatcgacttatggtattataaagaagcatatttcttttggaaacataac aatgccgacagaaaagaattgaaatgggttaattctcattccaatgatgctgtgactata gcattacatagtgataatttttgcggaatgtagaataaatgatacagaacaaactatttatt acaagcaagtcagaaaagaaaaacgttctctcatgaagctacaccgagaaaaggtcgt aaacttccaaaataagaagctaaacgtataataagaatacaaaaagagttgttctaaag gaaagaagtattctttatagacaaggttgcataatgaacaaaatggatggataacagg gtttacaggaactcagcttatgtacaacatatgatggagaatatctaatctgtaggaaa gaattataaacaataaagcttcttcaacattggaattcgttcaagaaatacaaatggatatta gaaataaaaagaaaatga	acaagtaaccgagctgtgaaaagcttggctt tagctgccctgctctgtagggagcatgacta gcctcagtttatacaccacaattcgtaaagcta cgtttatagaatcaacactctaccagcgatg ccctaactgctatgggttggagctctgtataa gactgaaagtaagggttgcgcaactccatgc atcaagttctatagacattggcgaagggcaa cgcataccgtcttgttgaagacaggagcg aataccgaggatgcggatgagacagtgc gtactcatcatacttttatataaatacaaaag gaggcatagttcaaatgaaaacagataaat tt
920-921	Clostridium estertheticum strain FP3 scaffold_9, whole genome shotgun sequence GENOME_ACESSION: GCA_011065935.1_AS M1106593v1_genomic	atggtcgaaatattctttgtatggactatctgccaacagtaaatgctcgcacaaaaat aaagcatattatcttattctgtaaaaaaagcgttaaatgcttaataatttccaatgtaatac agttacaaaaaacagtttaaggatgacaaagatgatgtgtaaaaaattatcttgaattgat gatggaaagtaaaaatgttgggtcggaaataatcaaaaatgtaaaacaaaagttagaaca attttaaaggaaactatagaatttaaggcaggtatgttctaaaaaaatgactgtacggaaggg ctatagaatgtatcatagataccataaaaggatctgtaaaatgagattataacagaagcg catctaaacgcaagaatagactgttccaactatactacaaaaaaacagctctatattaagg	atatacaaatgtgaagtgatgttgccttc aagcctaagtgtgtgtttttaaaggga aatagcacgatgaactacgttagatgatcgg ttaagacgtaccttagatgtacttcagctg aagcactacgatgttagaagaacataacc taaggcattgtatgttagacacatcaaac tcaatacaccacatttgacattggcaagaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
	GENOME_ID: 327298 CONTIG_ID: 84 SOURCE: NCBI_Prokaryotes DATE: 2020-04-22.	gttgtaataaactgttaaaatggactaagatagatgctatataatctgaagatgttcttag acataaggctatgtgtgaaggttaagctctatacaaatggcagatcaaaagtcaaatag acttgacaataacatcagattagctgtatttatgagagatggctttaaagcgtagactgtaa ttctaacaccaaaacttcaaatgcaccacgctaaccacaaaatagtggtggagcagatag tatttataatgggttaactctatgcgaaaaatgccacatgaaaacttttggcaagagtgct gatgatggatggatatttaactaagattaaaggtaagaatctatgtcttacacccaatgca tgttatgcaaggtaaaaaatctgcgaagtgaacttgaaaaaatagcaccgatatcttga ctacaggagctgatactgcaaatcatagaatagactggaatatagaaaaagctattctaa tgatgcttttagtggtttgtgacactgaaatgaagctactgatacaacattaaagactggtat ataagagctttacgtaagaatcaagggtgatacagatacattatgaggtatttaatta agagattatgttaatatatacaaaagcgaatggcataagttatattgatatacacagcattg tatccagtttaagaacagtttaacatgacaactaaagatgataatgtacttaaaagatacgg attaaaaagctttagccttattcaagacaaatagatacgggtttcataa	gaaaaataccactaggaggatgactagag atgttcgaa
922-923	Acidithiobacillus sp. GGI-221 contig_1300, whole genome shotgun sequence GENOME_ACESSION: GCA_000179815.2_AS M17981v2_genomic GENOME_ID: 41877 CONTIG_ID: 1551 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgctgctgccgaaaaaagccccgataaacttgcggaggcacactttaccgcgagagc ggagactgaaaaggtagcgacatgcagaacagatattgtattggacaaaaatggca accgctcatgcccctgccatccggcaggcgatgagcgggtgctggaagggaaggcgg cgggtgtccggcgtactcttctgcatatcctgaagcaccgggaggggtgggtgaaaagc aggaaattgggtaaaactggatccggcagcaaaaaccacggcattgccctgggtgct tcgttcgcccggcgtggaccacccgttatctggcgccgaactggcgcatcgccggc cggccattcgcaaggcgtggagcagcggagcggcgatcgacggcatcgccgggga aacctgctgacgggaacccccgtgtgataaccggaccggcctgctgggtgtgtgc cggcgtctctacagcaccgggtggacaccacgggacgtgggtggcgccgtgtggcc gatggcgccgtaaccgcccctctcgcaggagttgttgaagttcgatctgcaggccatg gagaaccccagatataccggaacggaatcaccaggggaacactttggggctacgaag tgaaggaatatctgctggagaagtgggccggacctgctgtattgagcggcggatg gtgcccctggaggtggatcatattcatcaaaagtgtaaggcgccaccgaccgggcag ccaatctcaccattgctgcccacgactgcaaccaggaaaagatcaacagccccgtgat ctgttctgaaaacgggcaaggacggcgccggcgatcgagggcaacgccaaagcg ttcgccggcgaagcgcgaagaagatcgcgcaacggaagatccatgagggaacccg gttgcagcggatacagcgtcaagcgaagccccgctg	gtcaaccactccccctaaagccctaacgggt tataggcggagctgctgggtgaaaccaaag gctcctggttgaccaggctcagtcaccagaat accgggactacgttattggggtcaggacac cctggaaatcgccgcccaggtccaggctctgtc ggataaggattaaacagcaagctggagta gtgcaagccgtgtccttgggtatgctgctc cgaaaaagccccgataacctgcccaggc acactttaccgcgagagcggagactgaaa aggtagcgacatgcagaacaga
924-925	Ga0105758_1004307 GENOME_ACESSION: IMG_3300007742_SF_3 300007742 GENOME_ID: 28195 CONTIG_ID: 4306 SOURCE: JGI DATE: 2019-01-27.	atgcccgtgttaccggcgaagcgcggcttctgtcttaagagaagaagctattgtgaag aggagaacgcctttcactattcagctgacgattgccaaggggtgagaccaaacagccggt gggtctggcggtgatccgggtacaacatgtcggactgtccgcatcaacggacaagg ctgaactttatgcatccaagtgcgaactgcgacaggacatcaccgatctgctcctcg tcgtgctttacgacgggctcgcagaacccgcaaacgcgctaccgcgcgccaagattc aacaaccgcatccgaccaagcgcgaaggctgctgtcctgctgggtgaaaaccgaat caacgcgcactgtcgcgcatagaagcgggttctgactgctgcccgtcacaaagatcac cgtggaaacggcgttttctgacttgcagctgctgaagaatcccagacttcagggaaga atacaggaggaggagaacagctcgcttctggaacatccgcgagatgttctttgacaga cgggcatgtttccagcattgttacggcagatcgaagaccgggtgcttaattttcatcatc tggaaagcagacgtacggcgagattcaccggcaacctgattacgctttgtgagacg tgccataaggcccttcatcgcggtgaaatcacgctgaaggcaagcgcggacaatcgtt ccgcgcggaagcccttcatggaattatgcgctggagggtgctgaatcgctgaaggcgt cgcatcctgagctggaagtgaacaacacctacggctatcgactaagcacgcacggatc gcgaacgacatcgtaagtcgattgtgcagatgctttctgcatcgccggcaacctggc gccgaagcgtcgcggaattcttctccagaagcagacgcgtcggaacacccggcaga ttcataagcgtttcactctcaaa	gtcaactacctcgccctaaaggccggggt tgaaatagcccttactgtgactagcctcaggc cgtcatttgcgggactacgttgggtgagaata tataggcaccgcgggatgtagatactagtc cgcgctctcgcccgatggttaaaagctcc gagaggttaggagcgggtgctgctggtacga aaccccttcaacattggcgaaggatcacaa ccggtcgaaagaccgagagacaaaactt gaga
926-927	human gut metagenome genome assembly, contig: NODE_3246_length_455 7_cov_1.626388, whole genome shotgun sequence GENOME_ACESSION: OHYD01.1 GENOME_ID: 5476 CONTIG_ID: 3245 SOURCE: WGS DATE: 2019-01-27.	ttgaaaaaagaagccgcaggcgaccggcagggggaaggactgccgccccgtccggtac ggtcccggtaccggcgaggatgccggacggaaggcgtgaagccggaaaggaaaca tatgtagcagttgtgaccagctactgggaagccccgatgccgaccagtgcataccggg cacgtaaagctctgaaaaaggacgtgcccgttatcttaaataccgtccgtattttaccatcc agctcgtggaccgggcagacgggggaagtaggaaatggaatataaatccgacaccg gatccgtccatgtgggatcagtggtgtgacgagaaaaaagaacttttaagcgaacagc gtgaccttctggagaatgaaccggaacatcacaatgaccggaagaagaaccgcaggac acggcggaaccggaaacggatccggagcccaggtttgacaaccggaagaagaagc cccaggagggccatgaaaaatggctgccctcaccctcctccacaagatggaggtcca ggtacggctgttcagggaattctgcagggtgtcccggtagctccgctgtccttgagatg gggaaatttgacacacaggccctgaagccgtatcgaagggggaaccgtaccggaa ggggaaagatacagaagggggaacgctatggcagcgaataccctccggctgccgtat ttcctgcccggacggacatacctgcccgttctgcccgttccgttaaaggacgggctatc ctcatgtacacatgtcgttactggaagggggacggacaaaaccgtccggcaacct tgccctgctctgtgagcagtgccataccctgcaaacatgggaagaacgggacatctgta cggaaaggaaacggagtttaagaccttaaggatgatcttatatgacatctgtaagggtg atcatgctggacgagatcaaaaagccgccccgggggtccaggtgtcc	gtcaattaccacggccatatctccctccgg gagaatgaccggggtctgcagaaggaaag ataacttttattatccggaatgacgtaagcct gggtgattacttcaggccagctctgtatgtca ggcgccccggctgctgaaactgcagggtga acggctccgttataaaggaaatataatagcac caggggatattctccagctcccactgtctct gcggcctgctgttaaacatcccttttgggtat cggggaagtgtagccggcataaacctttat aacatcaagatggagaacacttgaaaaaa gaagccgcaggcgaccggcaggggaaag gactgccgccccgtccgtacgtccgcc gtaccgggacgggatgccggacggaagg cgtgaagccggaaggaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
928-929	Ga0311351_10027103 GENOME_ACESSION: II_bly_15_\$F_33000317 22 GENOME_ID: 280479 CONTIG_ID: 27102 SOURCE: JGI DATE: 2019-06-13.	atgacatcaaccaagcctctgctggttcttagatgtaaatgggaaccatcatgccga cttcaggcaagcgtgctcggaagttctggcctctggcgtgctcacatggtacgagta gacctggggtcatccagatgaacgacacctaccaagaagactgtaccctcacagacatg atcgtcaaggtagatcctggcagcaagtataccgatgtgctggtggcgatacagcctgag ccagggtgcttgaaggtaaccaacctgatagagttacagcatcgtgctgactgattgca tgcagctgaccaggagagccgcttttgcagaacaggcgaaatcgcaatactcgtatc gtgctgcaaggttttaaatagaacaaagccaaaaggttggtagcccatctttgatgcac cgtgtcatcagcaggtgaattgtgcaagcgtttgatgaagtggtaccgcatcagtag ctggctgtcagtggtgaagttgacatgcagaagatgcaagacgccagcatccaaggt aaggagtaccagcgaggtgagctcttcgagcgtgaatgatggagtacttgccttgaag tatgaccatactgcgtctattgtgataccaagatagcccgtttgaaaaagatcatgtactg gcacggagtcaaaacggatcaaacaggatctcgaacttaacttctgtcgtcagaccctga accaggctaaaggaacctgtctgtgcagacccttctgctaaagatcctgtcagattagc tcgtatcttaaacagctgaagacacctctcagagatgccgagctgtgaacgctacgcg taaccgactgttactgaaatgatcaacttggcttacctgtctctacaggacaggcgctc agactaaatggaaatcgacagacttggatctcctaaagacacatgcactcgatctatatgt atcggagatgtaaaaaccgtatctgactggcaagacctcatctgaagcaagtgctccg ggagaggccggtatgctcggaccattaccgataaatacggctccctcgaactctatgtgc tcgtaagaaagtacatggaggtttcaaacaggtgatctggtaaaagcttctgcctgtc aaaagaaaagctactatggggtagcttctgtgcgagctagattagaattttctatagattgg gctcggtagagatgctatttctgtagaagtaataattgtaaaattactgcaaaagccgac ggatacaaatatttctcaagacatacagctatacttattcaaggaaaaatgtattgtcagtta g	gtccgctaccactctgtgaacaaggcagg atgctaaaaacatcctgtagaccagctcagt cctgaaaaggactacgtttgtatagtcagac acctggagtgctccctcagctccttctgtc tctgtctgattaaaaagtcgggaggttagg gacgggtgtggcagacctagtaagctactca aacattgacgagaggaattttctagccggt aaaggccctccgttaaccaaggagaaaaa tgacatcaaccaagccttc
930-931	Marine metagenome k99_2163173, whole genome shotgun sequence GENOME_ACESSION: SPKZ01.1 GENOME_ID: 279200 CONTIG_ID: 16826 SOURCE: WGS DATE: 2019-06-06.	ctgagtattttatgagagtttttagttaacaacatggtgaagttatgacctgtaagc cgcgcaaggcgaagatcttctcgtgaggggcaaggcaacagttgttaagcgaatccg tttacgatacagctcaagcatggtctcaacgggctacaagcaagatttaacgttggcgttg atacgggacataagagttggtatttctgtctatccgcaacgaaagaggtcttttctgctg tcgctaagatgcgcaatgatatactgcacaaatgactacgcgaaaaatgtatcgccgtca aaaacgtaataagttgcgctaccgagcaccacgctttttaaaccgttctgcgagcgcacg aaaaggagggtgcccccttattcaatgaaagttgatgcgcatattaaactaattaatc agcttaaatccttactgcctatcactaaggttatctcggagacaggtacgtttgatagtgcta aaattaacaatccagacatcacaaacgcgcagtatcaacaaggcgttcagtaggattga gaatgtaaaagcttattctggtcgcgtgatggatacaagtgccaaagtggaaagaaagg ctgctcgataagctacacgtacatcatatagtttttagatcacaagtggttcagatgcac caagcaaccttaaacactgtgcgaaaaacatcatgcagcgttcatgacggaaagtgtg cgctggccgataaagaacacaagagctggaacttgcacagacaatgaacattattcgtc ca	gtcaactaccgcagctaaagcaagcggg ctgtgaagggtgacttacgagataggtcttcg actgaataggtgattagctaaagcttaattga ctacgttctttaggttatgacaccagtgatg cctcccaagctcgtcgtctgtcgttggatt aaagagttcaacgactcgggtgtcgaagtgc aaaaagcctttacaacattgacgatgggaat caaccgcgaagggtattgtaatttttaactg agtattttatgaga
932-933	NODE 4_length_65995_ cov_8.35354_ID_126919 0 GENOME_ACESSION: mgm4799991.3 GENOME_ID: 238386 CONTIG_ID: 3 SOURCE: MG-RAST DATE: 2019-01-27.	atgcgcgtacctgtactcattcaaggaaacacgttgatgccactactcagccaag gccagaatccttatcaaaactggaagcgaagccttactggaataagttgggcatcttct gtataatcctaagcagagaagttgaaccgataaccgacaatagcagttggaattgacc ccggatcttcattcgaaggttggtcgtggtaccacaaacacatcttgaaaggcat gtctgaggcaccacattatatacaagaaggctgtggaagttcgagaaacgatgcgtcgcg cacgaagacacagaatctctgtagacgtgaagcagcatttgacaatagattgaggaata agtcttcgtaccacataaccttagccaggttgaacgcgaaactaagatattgaatca gttgctcaagggttgcctatctcagatgtgttgggaagatgttcaagccaagtccaaga agaattgtaagaagtgaacctatgcttttagtccgattgaagccggtgaagaattggtctat actgaaatcaggaacttggactgaagttgtatctcaacaaggatcgaacaaagattc ttcgcgagaagtttaacttgaaaaaatcttcgagaagagtaagcagacatttgggtctcat gcagttgatgcgtgggtgttagcagccgatgtgttggcgcaatcaagccaacctgtttg gaattgatctattggataccgattaggttgacagaagacagttacataggttcgaacaa caacaggaggttgaggtcgtcttattggtggaacaaggatcaatggattgaccagagga acattggtcaagcacatcaagtagcgttcacacatagagcggaaactctgaagaacaga gtttcgttgcatagcgttaagacgggtaaaagggtgacacagggcgcaaaagatccagga ttccggatcttgaccaggatattcgtgcggggcacgctctccacgaataaattcgtg ggtttccgctacccccgcgccccacggagtttcatga	tcagatattcgtcgttggttaatttcccttcg gagttgcgaagcaggcttaggttttcgatcct acgaiaaccatcggtgctggtggtcccgag gagaggtacgtctaggggaagtacacgaa cgtccttgcagccacaatgcttgcgaaggt ttttgccccctagtctactgtttagtcgcgtta gcccttcacaagccccggtctcaggccgg ggtagttgac
934-935	JGI24025J20009_100178 87 GENOME_ACESSION: Cruiolmetagenom3_2_\$F _3300001749 GENOME_ID: 24902 CONTIG_ID: 17886 SOURCE: JGI DATE: 2019-01-27.	atgcaaaagtatatgtattgaacaagacggaaacacctttgatgccttgaagaagcaa aggcaagacatttaataagagatgtgaggcgaaaagttagtaagcgttaactctttcacaat tcgcttgaaattgggattgtgagaaaacactcaaaactattcttaggagtagatagtggt attcgaagattgggttttcagctatcatttaataccaagaatttaatacagtggaagttcaggt aaggatgatgtatcaagaactatcagaagaagaatgtatcgaagaacagaagaa gtaaattggtatagaaaacctaagttgacaatagagtagtaacagtaaaaagaagaactg gttagcaccctcaatcagcacaaattggatactcatafcaaaaataataaaggttagag aaatattacctaattcagatactgtagtagaagttgcttcatttgatatacaaaagattgaagat ccggatattcaaggtaagattatcagaattggagaacaaaagggtactataacatcaag gcatattgttaaacacagagataattacttgcacattgtaaaaggaaagtctaaagacaa	gtcaatcgccccctcctaaggaaggggctt gagtcgtgagacttgagggttaactggttgat tagaaggcatagaaatgcagaagttatca gtaaagttaaagaacacacaggagatactt ctctagtctctgctctgaagttctgtattaa caaaagggaactcttagtatgcagaacata gtactggctgataacaactccgaagaaac atacgttatcagtggtggcggtcctaaagcc attacactgatatatactaaataaaaatga gaatctattatgcaaaaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gattttgcaggttcacatataagaggaaaaaagaaaggccaacaacagaccagaa gagtttaattacagtttgaataacttgccatgagaacaccacaagggaattattactattgct gttaaacccatcaagaca	
936-937	Ga0182027_10092596 GENOME_ACESSION: 712metIIAssembly_\$F_3 300014839 GENOME_ID: 21119 CONTIG_ID: 92595 SOURCE: JGI DATE: 2019-01-27.	atgttaaagcagcagtttaagttggagagatacatgcacacctacggatgcttactagtc gtggctctgtggcttctgtttaaagacaccgaaagggtcagtgagcgaagcctaataaa cccaatttaaacactccgaagtgaccaaccccgcaaggggagaccggcttctgagagta ccagtcctaaacatcgcatgcaaccgttaattgccaacaacaccacaaaagcaaaaa actgctacgacaagcgaagcaaaagcattcaagagcaccattaccatataaattgctt tacgctactggcgaacaaaacgccaataacactaggcatagaccccggtacacaaaac cattggattcagcgagtaaccgccaacacgaattatagctggcgaagcaccctcc gcacagatataccaaattgttaaaagaaaagcaatgtatgcacacaaaacgtagcc gacaccactggtagcaggaagcagatcaacacggcaaaaagaacagaaaaacagct accgccaagcctacacaaaactgtattcacatcagactgcccagaaactgcaaa gaattctggcgataacaaaagtaattgtgaattgacgctttgacacccaaaaatgatg aacacagaatcagcggagtagaataatcaacaggaacattacaaggctacgaaatccg agaatatctgcttgaataatggggaagaaatgcgcatactgtaaaaaacaaacatccc gcttgaatagaacataatcccaagtcagagcgggaacagatagcgtcaacatct cacgttagcatgccatgattgcaaccaacaaaaataatctgaccgcagccgagtttgg atacctagaataatcagcagcaagctcaagaacactcaacaaacgccattcatgaacgt cataaacgcaagaatcaagaattacttaattgcgaataccttacggttacgtaacaaa aacaaccgcaatgcacaaggttagagaaaacacacgtaaacgacgctttaccatagcc aaaggcatagaacaacaacgaagccttaacataatgttaactcagcgaagaagaacaa ccgagcactacaagtttaatagaagggctcaaacatccataagaagaaaaaggcaca tcttccaaagcggagatttagtcagatacaaaaaagaattatgtatgttaaggcgtcttc aactacgggatagggcaaaatgatagattcgaagaaaaaaactcaattcaaacgtta aaaactgaggttgattaaagtagcgaaggggctacaatttcacgtaagcaattacgcc caacatgaatgttgggtcttctgcaaaattagataaagtcgaagattgtggaacgctta a	gtcaactacctgctctgaaggacgaggtt gtctcgtgagggataaggttaaggttgatta ggaggctttaaataacatgttaagcagcag ttaagttggagagatacatgcacacctacgg atgcttactagtcctggctctgtgcttctg ctttaaagacaccgaaagggtcagtgagc gaagcctaaaaacccaatttaacactccga agtggaaccaac
938-939	human gut metagenome genome assembly, contig: NODE_3479_length_552 9_cov_6.161409, whole genome shotgun sequence GENOME_ACESSION: UESJ01.1 GENOME_ID: 15157 CONTIG_ID: 3478 SOURCE: WGS DATE: 2019-01-27.	atgtcgggttctgtatttctcaggacaggaagcaccggctctgatgccgtgccgaccgc ctcgggcgcgaagactcctgaagtcggggcgagctcgcgtgagactcttcccgctc acgatccgcttgggtgagtcggcttctcgaagattctccgttcagtcctcctgtaagat cgatccgggttcccgtagacgggctgttgccttctgagcggagcagagaaggctcatc accatgcgctcttctcgtcaatctcgttcacgcggcggaatccatccgtgatgcgttgacg gctcggcgaactctgcggctcggcgacggcggaatctccgccatcgtgccccgaggtt tctcaaccgcacgaagccgagggatggcttctcctgcgttcgacaccgcgttgata ccgcaaccgcctgggtggcgaagctcgtcaaacctcgcccgcttaccgggagctgtgga ggagctcgtcaagttcgacgcgcagaagctcgagaacccggaaatctcgggcaccga gtatcagcaggggacgctcttcgagtagaggtgcgcgagatctcttgagaagttcgg caggaaatgcgttactcggagcgggagacgttccgctcaacatcgalcatgtgttcc cgaagcgcgcgcggttgcgaatcgatctcgaacctcgtcttctcgtgaactgca accaaaagaaggatcgaacctgttgaggtttctcgaataatcgccccgaagtcgtga tcgcatcaagcgccggctcaagacttgcgtcgcgcacgcgcgaacggtcaatgcaacg cgttggctcgtcttcaacgcgctcaaggcgttcggttctccgtcgaacgggttcgggc gcactgacgaagtcaaccgccatcgttcggcgttccaaagagcactggttgatgcg ctctcgcggggcggtgaacgggtgcactaccccaggggatggcatcctcaag tccgttgcacgggacggggagctatcagcgcacgcgggtcgacaagtacggcttccc gcgcggctgcttgacgcgcaaaaacgcgttcacggattcgccacggcgacatgtgc aaggccgtcgtccctcggggaagaaggccggaacctaccggggccgtgtgctgttc gtgcgcgagggttcttgcattcagacgcctgaaggcaaggtggacggcatcggttgc gtcattgcccactgcttctttaaagcagcggatagcgggtgcttgcggccgcacct cattcctctcctcgtga	gtcaatcaccctgctgaaggcagaggtt gagcgtacaggccggttaccagtcagtc gatccaaaaggagaactacgttgcagttag gttacaagaccggcccggtgcttctca gctccgggctccggaacggcagaagcag acaagattgggtattcacgaaacggtctgc cgttgcacgtccggactgcaacattggcga gggggaattcaactcttgtgtatgcacggg ggtaaccccgttaacgtaaggtttacaaa acatgctgggttct
940-941	human gut metagenome genome assembly, contig: NODE_3969_length_350 8_cov_1.320319, whole genome shotgun sequence GENOME_ACESSION: OQCF01.1 GENOME_ID: 9680 CONTIG_ID: 3968 SOURCE: WGS DATE: 2019-01-27.	gtggttaaggcgcacgcttaccattcagctcagcattgcgaccggcgagtcgaaaca gccgggtgagctcgtggttgatgccgggttacaacatgtcggacttccgcatcaactga aaaggctgaactttatgcatccgaagtcgaactcgggaaggacatcaccatctgctctct gctcgtcgtcgtttagcgaatcccgccgaacccgaaacgcgctaccgcgaccga gggtcgaacacgcacatccgaccaagcgaaaggctggttgcaccgtcgttgaagaaac cgaatcaacgcgcatttgcgcgcatagaagcggttctccgactgctgccgatcagaa attaccgtggaacggcgctccttcacatgcagcggttgaagaatcccgacatttcagga aaagagtaccaagaaggtgaacagctcggcttttgaaacgtccgcgagtatgttcttita gagatgggcacgtctgtcgcattgtcagcgagatcgaaagacccgggtgtaattgttc atcacttgagagcagacgtaccggcagcgaatcgcggcgacaaactgattacgcttgc agacgtgccataaggcgttcaccgcggcgaatcacgctgaagccaagcgaggac aatcgttccggcgcaagcttctatgggaattatgcggcgagggttttgaaacgcctaaa ggcgtcgtatcctgagctgaaggtgcacgaacctacggctacttaaccaagcacgcgc	gtcaactacctggcctaaggccgaggtt gaaaaagcctctagttgactagcctcagtc gtcgtttgaggactacgttgggtgggaatgc acaggcaccgcggtgagcagatcctagtc cgcgctcgtcggccgatgttaaaagctctg agaggtaggagcggtgctgtcgttga ccccctcaacattggcgaaggatcacaa cggctgaaagaccgaggagacaaacttg aga

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ggattgcgaacggcattgtcaagtcacaatggtgcagatgcttctgtattgcccggcaacttggaaagccgaaggcctacggcgcgaaattcatttcttccagtaac	
942-943	Ga0066665_10044143 GENOME_ACESSION: GraSoiStandDraft_32_SF_3300006796 GENOME_ID: 27025 CONTIG_ID: 44141 SOURCE: JGI DATE: 2019-01-27.	atgtttgtacccgttgttaataagcaaacacaacacactgatgccaccacagctgccagggcacggcgctggatcaagagcgggaaagccacaccgttttgaagcagggcatctctgtcgttcggctcaacgtcgagcccttccagcaacatccaaccggttgcgtgcgaatcgaccaggttccaaacgagaaaggttacagcgtcgtgtccgccccatagctacctcaacctgcagctgataccgtggactgggtcatggaggcggtaaccactcgccgccagatgcgccggacgcgccgagggcgcaatactccctgtcggcagccacgctataatcgcttgcgaaaacaaaagagctcccacttctacgaaagcccgttgcgaatggaaactcgccgtgtgccgttggcttatacttataccatcgagcccttgggtgcgaagatatcagggttcagacagcaggcaagcggcgatgggacgggtttttccgttgaagtcggcaagcgggtgttttatgcggaactagcgaggcttgcgccagtggtgacgaaacaggatgggaaaccaaggagcttcgagaccaacttggcccaagaaaacaaagaaaacggcgagggtgtgggaaggccattgcgtcgacgcctgtgtctggctcattgtgggtaggaggcaactcaaccggacaataccgcttgcgttgcgtgacgccttgcactggtatcgtcgacaattacaccgacgcgaaccgaaaagggtgggaaacgcaagccctacggagcgacctgagctctggc atcaagcgttggcaccctgggttaaacacccacattatgactgtgtatgtaggggcaccatgcgcgggaaactcagcctcatgcacccgagagcgttaggagactcacaaaggggccaaacactccgaatgccatctatcaagctgttacgttggaggacgcgactcctcccaatccgctatga	gtcaagtaactgctccccagaagggggcttgtgagcacgaacgactggccatcgaggagggcccaagaccaaacacgagactagcttggcccaagctctggctacagggcactggaagggaactgaaccatctgggtgtggcagcccgcaagacgatacgatgcctcctatgtttgtatcgctctgggtgcagcgcgaagggaacgt agaaaactccgaaggaggcttatcgcatat
944-945	Ga0216255_10000297 GENOME_ACESSION: IMG_3300021179_SF_3300021179 GENOME_ID: 311738 CONTIG_ID: 296 SOURCE: JGI DATE: 2020-04-07.	atgagagttttgttataatcagcatggtgaagcattaatgccttgcataccccgaaaagccaaaatactgcttagagcaggaagcgaaggttatttctgttcaccttttactattcaactgaaatcaggctcaactgtgtataaacaggatttactctggcggttgacacggggcataaacgaagttggcttgagttgtttcagaacaaaggaaagtttttcagcgggtggcttctatcggaatgatatttccaaagagtgactgagcggaaaatgatccgtagaacacagaagaaccgattgaggtatagacagcctcgttttctcaacgaaagccagcaccagacaaggtagggcttgcaccttctgtcaatgaaagtgatgcccatgcaaaatgatccgtcagttaaaaatgttattaccaataacaagggtgtatcgaacgggtacgtttgatacccataagctaaaaaatcctgacattaaagatgagcaatacagaaggcgcttcaatatgttttgagaatgtaaggccttatgtcttaagctgtgatggctataagtgcagtggtggcaaaaaggatgtgcctctacattagaggttcaccacattaaatccgctctcaagggtgttctgatgcacctgaaaatctcattacattgtgtcaaaacatcataaggcactgcatgacggaaaatfgagtctgaattacaaa gcacaaaagttttaaagctgccactacgatgaatgtgatccgttctcagttgttgaagttgtt cctgatgcgatagaacttttgggtatgtactaaagcaaatcgctatcaacatgacatcga taaaagtcatagcaatgatgctttgtcatagcggcggggtctttcaggaaagagagttag agcggatgggtgagctttaaagcgaagaataaccgttctatccagaagaaccgaaaagggtt tctccttccattagaaaacaacgatatgcatacaacaaaagacgtggtgaattttaaag gaaagcaataccaggcggttaggcatacagaacaaggagcatactgaaaatgtccgat ggacttaagctattgtaaaatccatgaagcaaatgaagtggattttcatcaaaaaggtatt atttacgcatga	gtcaactaccactgagctgaagactcaggggcttgaatcgtgagatcagagatagcggtagcaatcgtgaatagggtattagcctaagctgtcttagagcactacgttttttagttatagcaccctcggaatcctcccaagtcggttgcctcttcttgggtatataaagatgtaacgactcgggtgta ctaagactgacaagctattacaacattggcgatgggaatcaacctgtgagaggattagttt taattaaactatttaaggttttaaaatgaga
946-947	wastewater metagenome genome assembly, contig: NODE_9891_length_3035_cov_2.017785, whole genome shotgun sequence GENOME_ACESSION: OVQI01.1 GENOME_ID: 11618 CONTIG_ID: 9890 SOURCE: WGS DATE: 2019-01-27.	atgttagtatatgttttaataacaaaggaacccccaaftatgccttgtagtcagctaaagcta gaaaactactaaaagatggaaaagctaaaatattaaaaggaggaattagaaaattaaatc aagctaagtatgatacataactttagattgttataaggtaaaaatagagaataaagaat attttatatttggtagaagaataagtggttttttgatatttagaaaattttaaaggcgcaaaagt caacaaaggtagtattgataaagaattttaaattagttttaccagaaaaaatatattaatt gagaggagagaaggcaattcctcagctaaagcacttag	ttataagtaactaaccccttgcataagcaag agccttgaaaaagcctaagttgaccagtctt ttgggggttaaacccactccgttatttagtt atcataccttgggtgtttttgtcagcaccatg ctctatgttggcaacgtaaacagtctgaga ggttaggaacagtcgacccaatgtgtatgtc cacaatacagctattataaccaatctaagaca accattaccgaaaggagattatatatttatgt ta
948-949	Ga0306925_10047498 GENOME_ACESSION: flu12C44000176v2_SF_3300031890 GENOME_ID: 281676 CONTIG_ID: 47497 SOURCE: JGI DATE: 2019-06-13.	gtggcgtcagcggcgacacgtacctcaatgtgctagcggatagcgtcgattgggtaaa agacgccgtggaacacgtgccatgttgcgtcggccccgcagaaccgaaatacacctt gcagagcgaaatcgcaaaaacagagctaaatgtccatttctccaagcaccaggcgcgctggcaactaaagctgcgcgtagtcgacgttccgccgtatgtacccattaccgggtatgc ggttgaagacataaaagcatctactcacgggaagaaacgaacgtggaactcggcgcttct cgccttggaaagtaggttaaaaagtggtatgtatccgagttgcgcaaacctggcaccctga cgttaaaggctgggcagagaccgagaactgcgcaaggcgttgggtgttgaanaaggt agcgataagatggcggttgccttccgctcacaatgtcgattcctgggtattggccaagg atcggttgagagcggcgaagcggcgagacacgacgactcttctactgtataaagcgctt caattcagacgccgagcgtccatctccagaacccggcaaaatcggagtgctcgttag tcatgggggtacgctgagcatgggtttaaaccaggatgcgtcgtttaaactccgaagca cgggctctcgaccgtagcgggaacgatgcaaggacgagtcagcctgcacagcttgcac	gtcaacgaccacgaccaaagccgtgggc ttgaggggagcacgaaacaaacgatgcgaa ctgaagttaaactcgagactagcttgacaa gattcgaggcgactcgggtcgttgaagtca actgagatcactcgggtgtacggcgcaa gcatgcgggtgcttcccaagctccgct caatctgtgtgacagcgcgaagggaatgc atacacagatacaggactgcttaccgcga atctctgtagttccgcttctgctcaggatgg cgaaccatt

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gggaaaaggcttggccgaacgcaaacgcgaagaacgaccttttaaacgaacaa ccttaattttcaagccgtccgcaggaaaggcattcctccatgccggaagggtgggtt cctgcctaaatctctatga	
950-951	human gut metagenome genome assembly, contig: NODE_33_length_11320_6_cov_37.108961, whole genome shotgun sequence GENOME_ACESSION: OG0X01.1 GENOME_ID: 4582 CONTIG_ID: 32 SOURCE: WGS DATE: 2019-01-27.	atgacaaattatgctttgtattagataactaatgtaaaacaattagcaccacaacaaagaacaa aaagcatgggttcttaattcgtaaaaagcgtgcaacattgggttaataatccaatggtaata caacttaataaagaattccagatgaagacatctgcaaaagataaaattcgtttggaaattga cgtatgtggactccatgtatgggtgttcattagttcaaaaaatgtcaactaaaaataagttatt gtttaaaggaaattattgaacacgcgaatgatgtaaaaacatcttatggaaattgacgtgaat atagacgttatcatcgtttatcacaaaagatatagacagaatagggttaataatcgtatgtcttc taaaagaaaaggaaattgcaccagtaattttacaaaaagacaagctataataaaggatt atcaaccaactaataaattgataaataattactgattattggttagaagatgtatctattgatat tagagcattgacggatggttataaactctgtggttggcaatacagaatctaataagattgg atgaaaattattgaaggctgttattcttagagatggttgaatgatggatgtggcagat caaattgtaaattagaagtacatcataaataacagaagattaaatggttctaatacattaa gtaactctattacgtatgtaaaacatgtcatcaaaaacagagggcaagaagaacagta tatgaacattattttgataattgaaatcttctgataagaaaaatcttaattacgcttaa	tagataaacatactcaaaagtggaacacaata agtcttttgacatggaftaacaataggctc aaggctaagtactgctactatggaacata tggtgcagatatgaactacgtgtatagtaag gtaaaaacacacctgtaaatgtaatcttcatt tgctgctctgtgagtgccaatcaagaacat agctaattgctctgctatgataacagagaata catgtcctctatgacattgccaagaagaaa aattctccgcaaggaggtgtcagaatgtac aaat
952-953	Mine drainage metagenome contig00009, whole genome shotgun sequence GENOME_ACESSION: AOMP01.1 GENOME_ID: 99 CONTIG_ID: 8 SOURCE: WGS DATE: 2019-01-27.	atggaaaagggtgcgaatggcggtttatgtactggacaaaaaagcaaccgttggat gccgtgctcgagaacggggcacggctgctactggagcgtggacggcggtcgtgca taagattatccgttcaccatccgtctcaaggacgggtcggtggggcactgacggcgct cgcgctcaaaatcggctcgccgacggcacaagcggatcgccctgggtgcggaatcc gagacgggtgacgtcggtgacgggaggtacacgggtggagcagctgtaatacctga ttgacctggagcatcggggcggtgctgacgaagaagctggaacagcgttccaacat cgctggcggtcgcgccacgggaaacgcgttatctgcccggcggttcaataaccgc agccgtccggagggttgctcgccgaagctcgagcatcgggtggacaccacggagt cctggtgcaaccgattaacgtctttggcgccgattacggcgatcagcaccatgtccacc gggtcgatcgcagaactacagatcctgaaatcataggcatcgagtataacagggc accctgttcggaatagagggtgcgcaatacctctggagaaatggggacgcacgtgcac ctattgtgacggcgtgacaaaacctgaggtgggaacacatcctggccaagcggaatg gcggcaccgatagatcagcaatctgacgttggcctcgacggccctgcaaccaggagaa gggcaagctgttctgcccaggttcttgatgacgggcaaaagcggttccgcccgtttga acggaacgcgcggcatttgcctacgacggggaaagcgaacattgatgcaagaagcgc gcagaacgcaaacagcagcaggtgacccgattggcgcgcatagcagcgagccaa ggcgccactcaaggatgtggcagccattacgctaccgcccggcgtactgaggatgt tggaaaggcagagattaccgggtggagatcagcaccgggggfcagccaaatggaatc gcacccgggttcagtatacgaagcggcacgcccgtggacggcgtgcatgggtcaggt ggatcgtatcgaggatggatcgtccattggcggtgcatgtgacggcggtgtgca gtaccagcgcaccaatctggacaagaagggtcgccgtgtgtttttccggcgagaa acagcaccacggatttcagacggggacatggtgcccggccgaagtggcgaatggcgt acataaagggtatccatagggacgggtggcgtgcttgcgtgggtcattcaagatgcg cacgaaaactggtctgcacgatggcgtaaaagcggaaagattgccggtgattcagcgttc agatgggtacagctattttcgggagcaggttgcgtga	gtcaattaccccgccctgaaggcgaggct tgcaaggcggtacgaggctaggttgatca tggcgaccgtgtacagctgtgagacgttgg agacgtttataacaggctgtacagacgact ccgggatgctgcccagttccgcgacactg caagggttgggtcatgctgcggaaggcaaa agcccggaaggaccaatcgtcaccgcaag gtgagacgggttatagacccccgaggg gagacattgtacaagttgtaccggggccg taagggttcaccccgcaaggcggttatgg aaaagggtgcgaaatggcg
954-955	Ga0401356_0239 GENOME_ACESSION: 20190111_54_FD_SF_33_00036521 GENOME_ID: 307143 CONTIG_ID: 238 SOURCE: JGI DATE: 2020-04-07.	atgtctaactctgttttcgttagacactaataataaaaccactcaatctctgacaccggggg ttgtaggggtttactcaaaagccgcaaaagctaaagttttcagaggtttccgtttacgatta ttctgaacaaaagctctgatgtcttctaaacggattttggcacaagcacaacgaccttgaa agacgcagccggcgtcaactctaccgtagggcgtttgtgaatcgggtcaaggaaatcgg gttacctgtctcaacaggtagtggtgggaaaactaagtacaatcgcattcgtttgaagtac ataagcggcactatatacgtatgcggctgtgttggggatgttgagcaattagaatgcaac gaatcaaccattgtcatcaaaagctacaggcagtggaactcgtcaaatgtgtcgcacaga taaatgttggttccgagtcggtacgtccccagaaataagtttgcaaatgctttcagactgg tgacattgttaaggctgtgtcacctctggaaaaaggttggaaacttatgtgggtcgtgtag ccgtgcaacaactggcagcttcaacattcggcagctaatgattgattcaaggcattag ccacaatatggcagcctagatcatcggaaagatgataccaatacagcttctctaa	ggctataagccgggggcttgaatgaagccc acacctgaccagaagaccgttgaaaactg gtaccggttattagcaaggttttaagttctca ccttgggatgcgtgccagttccaagctctga aacctcagcattaaacaggctctacaagggtt aagccagtggttggagatcggccactaat aacaacttcgaggcaaacgttacgagttttc aggacagacgacaacaatgtctaactt
956-957	human metagenome genome assembly, contig: NODE_1669_length_17466_cov_67.5101, whole genome shotgun sequence GENOME_ACESSION: ODJC01.1 GENOME_ID: 3042 CONTIG_ID: 1668 SOURCE: WGS DATE: 2019-01-27.	atgccttgttctccggcaaaagcacgcttctgcttaaaagagaagaagctattgtgaaga ggcgaaacgctttactattcagctgacgattgcgaccggcgaggccaacagcccgtg agtctaggtattgatccgggtacaacatgtcggcctttctgcatcaacggaaaaggctg aactttatgcatccgaagtcgaactcggcaggacatcaccgatcgtctctgctcgtc cgcgttacggcagtcgcgcgttaaccgcaaaacgcgctaccgcgaccgagggttcgac aacgcacatcagcacaagcgaaaggtggctgctcgtcggttgaaaaccgaatcaa cgcgcactgtcgcgcatagaagcgggttcttcgactgctccgggtcacaaagatcaccgt ggaaacggcgtcttctgacttgcagctgctgaagaatccgacatttcagggaaagaata tcaggagggaacacgctcggctctggaacatccgcgagatgttcttcttcagagacgg gcatgtttccagcattgttacggcagatgaaagaccgggtgcttaatttcacatcgtg aaagcagacgtacggcggaagattcaccggcaacctgattacgctctgtgagacgtgc cataaggcccttcacgcgggtgaaatcacgctgaagccaaagcgggacaatcgttccg cgcggaaagccttcagggaaattatgcctgggaggttttggccgctgaagcgtcgc	gtcaactacctcggcctaaggccgaggctt gaaaagcctctagtgtactagcctcagctc gtcgtttgaggactacgttgggtgggaatgc acaggcaccgcggtatgcagatcctagtcc cgcgtctcgtcgccgatggttaaaagctctg agaggttagagcgggtgtcgtcggttga ccccccaacattggcggaagatcacaaac ggctcgaagaccgaggagataaaactga ga

SEQ ID NO:	Name	IscB sequence	ωRNAs
		atcctgagctggaagtgaacaacacctacggctatcggactaagcacgcacggattgcg aacgacatcgctaagtcgattgtgcagatgctttctgcatcgcggcaacctcggcgcc gaaaggctcggcgaaattcttctccagaagcagacgcgtcgggaacaccggcagattca taagctttccatctcaaaaggcggaattcgaaagcgcaatcaggctccctatgagatcaaa ggcttccgcttttgacaaagtgcctgccaggagaaagaggcttctttcgccgctc gatcatctgggtctttgatgttcggaaggtgacgggaacccgaatttcggcagcatcag ctacaagaagctgcatctgctggaagagacgaacctatttaacagaaattcgaaagga ggaggtgcttccctccctgctggaaggcagggtctcccgccctaa	
958-959	Ga0213833_1005875 GENOME_ACESSION: SWATP21_2_\$F_330002 1419 GENOME_ID: 35212 CONTIG_ID: 5874 SOURCE: JGI DATE: 2019-01-27.	atgccaaaacaaaaacatagaaaccttgggcagcaaaagagtccttgccttcaaacaca aagaagccgttgatgccttgcctatctgcaaggcgcaagacgttctgaagcaagggaag agctgttctatcagcgtttcccttaccattttgttaaagaccgtgaagacggctacata caacccattgaatgcaagattgatcctggtgccaaagagacaggcattgcattcgtagcct ctttcaagcgtgggaacgtttgcatttggccgctcatctgaacacaaagggttcaagtc aagcaagcactcgaagcggaagacatttcgaagacggcagacgcaatcgaagacc cgtcacaggaagcccgtttgacaactgcgcaaggcccgagggttgggtgcctcaag tctacaaagctgtgttgataacatcagcaacttcacgttaagtatcatcgtttgttcttc actggcttaccgttgagcatgtaagtttgacatgggtgtaagtcaaaaccagaagtg ccggtattgagtaacacagggaacgcttcatggttaccgctcaaggatactgctcga gaaatacaagcgttctgtttactgcaacaaaacagatgtgctctggaagtcgaacat gtcgtacctaagagtaaaaggcgtcgaatcggattgcgaatctggtgttggcttgcga agtgtaacgaaaagaaaggcacaaggtatggaagcttttcaggaaaaagcctgaac gcttgaagacaatcaaacagggtctcaagaagccactccgtgcagcggctgcggtcaat gccacagcaacaagattgttagagagcttcgagccttcggttccagtcactacgaca acaggcgcacaacaaaattcaaccgaaagggttcaaggctaccccaagcacatggc tcgatgctgctgctcgtggagaacaaggcgaatctgtcgaattactctctggttctatctac acatcaaggccatggggagaggttcacggcaatcttcaagccagataaacaggcttc cctgtacaagccaaaggcgaagcaaaagtaagtcgaagggtttcaaacaggagaca ttgtcaagctgttgcctacaggaagaaagtgggtacgcacatgggtcgtgtcgcag ttcgaagtagcggctcatttaatacaaaaacaaaggacaaaaccatacaaggcatcgttta caaacactgtcagttattacacaaaatgatgggtacgctatgactttaccgactga	gtcaacgaccccatggctgaagccagggg cttacgagagtcagcccgtatgttgaccagg gtcagtcgtcaaaaggcagactacgttcgagt ggaacatataggtagcgtatgttcttcca ggccatacgtctacgataggtgttaaacgg ttacgggtgatgccagtgccattgttagac aaaacccactcagaacattcccgaggagac ctttaccgcggaagcggtaacgttagacc tcaatcattgaggtcattgataacaggagt tggcacaatgccaaaacaaaacatagaaa ccttgggcagcaagagtcctt
960-961	contig_288372 GENOME_ACESSION: mgm4547164.3 GENOME_ID: 237431 CONTIG_ID: 54383 SOURCE: MG-RAST DATE: 2019-01-27.	atgacacatgccaaaacacagcaaggagggcattcaatgcaggaaatacattatgtgtg gcgaagatggcacccttcatgccaactttaaaggaaagcccggtgagaaccttatt aaactgtgacaagcggcagagtggaactgtaccgtatgtacacagctgacatagag ggccccaagaaacacagccgttatgtgaggtactgacccgggaaggaccaatcag gaaatgccgttatgaccaaaaggcgggtagtgtgtacaaagacatgtgaccaccaga aataaagacatagcaaaaactgtggcggaacgaaaaatgcatcgcgaacaatcaagac gaggtgagcgtctggcagcaaacgtcttcccaaacgtctgggacaacaaccaaacac cttgagagaaggattctccagggtgtaagaacccgttaattgtcaagatattagcaaca cagaagcacggttcaataatagaaaacgtctcccggtatggaatactccaaagtacaagg cacatgatacactatcatcaatattgattaagcgaatctgcaaaatactccggttatcat tggacacttgaacgaataaattttattcatgttaattggaagacggaactgtacgtgggag agacttccagaatggagacttaaaaactttactgatgtgtacgaatatgttgggaaccaac aaaatggaaaatgtatttctgcggaaagccaatagaacattaccatcatatgtaccaag acatagtgaggttcaaacgtcctgagaacattattggcgtatgcaaaacatgccatgag gaaatacatactggacaagagatatcacagccattgtgtgagcataagaatatgcagca ctttcggtattaaaccaggcgataccgttcattgaaatggaactcaaaaataattcggtaat aacittatgactgtacgggatatgaacatcgaactgcgtcaacgtttccgattgcaaa agaccacgacaacgatgccgtatgcattgcatcctaccagcgcacccagattttatcag gatacacaaftaactcatcaagttatgcagttcagaaggcacaaccggcagagaattaft cacaacgtgaacggacataaagcttgacggcaagccatagcaaaaaatcgtaaagcca agatttgagcaaaaagataaggcgttctgtgatttaattgtcaccggctgaaatttcaag gctgacagtaatacaagtcgtcgtgttatacaaatatggaccgacttatgccgggaac cgaattctgtatgaaggtcaacgtcatcatgagcgggcagcactcgaacggttctttg ctgagagcggtagacaaggtaacagagagttcaaaagcttccaaatgtcgtgcataaaa caaaacaggggacttctatcgtctcctaa	gtcaataaccaatgactaaagtcacgggctt gtaagagctgtcatgtgcaatttagcatgttc gtgatggcgggaagtggattatgagcagag ctatgaccgggtatfaccagtggaatggaac tagtccgctgctctgttaaccggccgag gctaagggaacagtttatggcatggaagag ccttacaagctatgacacctgccaaaacac aagcaaggaggcatcaatgca
962-963	Ga0209345_10000187 GENOME_ACESSION: Cruoilmet7SPAdes_\$F_3 300027852 GENOME_ID: 24899 CONTIG_ID: 186 SOURCE: JGI DATE: 2019-01-27.	atgttagtatacatactaaacaaaaataaacctttaatgccttgcacaccagcaaaagc gtctaagctttttagggagaaaaggcgagaagttgtcaaaagaacaccttttactattaaat tgaatatgttagttcgggatagttcaaaaagctatattaggaatagatacaggatatcaa ctcataggttttcagctattatgaaaaacaagaactaatctcagggtgaagttaaattagatt gtatgatgcctaaacgcttacagtagagctatgtacagaaaaagtagaagggaataggt tatgtatagaacacacgttgagagaatagggttgggtcaaaaaagaaaggcgtgctac ctccatccattcaagaagataatgaacacattttaaactttatagaactgattaaacaatgt tacctatacataggattataatagaagtcgggaattttgatatacaaaaaataaacatccg gatattcaaggaaaagataatcaacaaggcaatattgtatgaatacaaaattatgaagttat tcttatgagaagggaacatggaaaatgccaatattgcgggaaggaaattttcaaaaacaaat	gtcaactacccttactaaagggaagggtttg aatggttaacattcaacgcaagagttgattag gaggcatgagtaattatgcagcagtttgaa atgaatacatcacaccttgagatgcttca ccagtcaccaagccctgtgattgataataac agagattatagggtcaggtgatacagatttaa aaactttccaaacacccagggtgaatcaa tccatgtagaggagtagatcttattgtta

SEQ ID NO:	Name	IscB sequence	ωRNAs
		tcatcacatgatcatcattatataccaaaaatgaaagcgggtacagataaaacaaacatct atctttactacataaaaaatgcatattagattacataaagaaaggggtgtataaccaattgaaa aaaaacagcaatacaagattctacattcatgaataatcgtcaaatggaatataaaaaag atattgattgtatataacttacggagcatatacttttctgataggggtataaatggaatgga gaaatctcatgtgaatgatgcttttataatagctaaaggtgagatccaatccagatgctattct tttatataagcaaaaaaagaagaataatagatgtttacaagtgaatagaaaatftaa accatctataagaagaaaaaggtataacatagacctggagatttagttaaatacaatgga aatttatacgaagftaaaggaatacatctctatgggtactcggattagatttagattcatctg gaaatattataatatactactaaaaaattagatgattgatatatcatcaaaaaactttgata tggagagcagcttga	
964-965	uncultured Clostridiales bacterium isolate RUG12885 genome assembly, contig: RUG12885_asm_36, whole genome shotgun sequence GENOME_ACESSION: GCA_902783575.1_Rum en_uncultured_genome_ RUG12885_genomic GENOME_ID: 346814 CONTIG_ID: 35 SOURCE: NCBI Prokaryotes DATE: 2020-04-22.	atgttagtctatgtaacagcaagggcgagacggccctcatgccacatccagatgcggc aaagtcgcaggtgctataaacaaggcaagcaaaagtattaaagcgtgtccgtttacc attcagctgctgtatgaacctgaacacgccgttatccagcctgttgaggcgggtacgata ccgggtcaagcacaacggatatacagcagttgcagttctatccggagcgcagaacaaa ggaagctatgcatcagaggttcagatgcgtcaggacataacgaaaaatcttctgtacgg cgtgagatgcggcgcgcaaggagaacccgcaaacagataccggaaaccaagattt gacaacgggtgatcccaagcacaagcgtggctgccgcctacatggaaaaacaag atccagacacatgctcatgagctgctgtatgtgtgtcaatctccgtgaccaaagtaac gatagagacagcatcctttgacctcagaagctaaagccgatctcgtatggcttaagca accggaaggcactgattatcagaagggcagcagctcgggttctggaacgtcaggagat atgtgttattcagggaatgccatattctgcatgctgcaagggcaaaaggaagatcccat tctaaatgttcatcattgagagcaggcaagccggcgccgatgcgcaataaacctgat aacctgtgagacatgccataaggcataccatgccggaagatcaaacgtccctcgg atatcaaacgcaagcaactttccgggatgccattttatgggatcattcgtatggccttc tataacgtcataaaggacgaacttggcaccagagggcatcagtgatgacatcagg ctatctgaccaagaataccagatacaataaaccttcaaaagacgcactgcattgatgcc aggtgtatagcgggatatccggagcgtgatccattaggctattctctacaacagaaaa taagatgccacaatcgcagatccacaagatgacataggcaagcacggcatcaggaaa gccaatcaggcgccttatctgtttaaagggttcagactaacgatatagtacagctaaag gcagtgatgggttcacgggagacgtatgaagatcgtttgactggcgacattaaa caatcagaagttggaatctccatcaaaaataagatttttctgttctgacacatactt acagaaggagaaaagcgtccccaccctccctatagaggaaggggtaccgcgcttg tctgctctga	gtcaatcaccaccacttaacctgacgggttg aagtggggcttgcaaaagctctgattaact agcctcagtgcaacggccggcgttaggcc ggattccattagcactacgtgtccagtcac aacacctcgagcgttatgctaactcgtg ctctgttgcctgttataatgttctgatggta ggaactgtgtggcggatctaaaaagctctga caactgtggcgaagcattacggcccg gtatgtccgcttaagcgttaaacatcct attctacggaaaggagccgtatcagattgt a
966-967	sediment metagenome genome assembly, contig: NODE_10398_length_35 81_cov_1.964041, whole genome shotgun sequence GENOME_ACESSION: ORXB01.1 GENOME_ID: 10886 CONTIG_ID: 10397 SOURCE: WGS DATE: 2019-01-27.	gtgcgttgcagctccaggccctgtcttgcggcagtaaacaggcataggggtcgaagc cagtgctcagcacataaaaagcctggataacattggcaggcaaacattaccggcgaacg ggagcgtctttatagcaaaagcttctgtcttgataccaagaagggcgtcttactctgc catcctgcgaagcggaggaactgcttaagcagggaagggcagctgtgttctgatactat cctttaccatcatcttgaaggggaagttctcagccagcactgcagccattgaggttaa aggtcgtatcctggaagtaagatagccggcttgcagtagtaaatgatgcaactggcagg ttgtgttcagctgagatccagcagcgggagatgagattaaagatagctagatacccg gcggagcgtccgaaggcagcaggaacgtgaagacgcggatcgtgaccacgggtt ttaaacccagaggagaagagggtgtgtacctctccctggagagccgggttgccaa tgtgttaacctgggtaagctgataaagatactgtctatcaatccatcctgatggagc tggcaaaattcgacctcgagaagctggagaatccagaatccagggggttgatgatcag cagggaacctctttgctatgaagtaaggaatatctgctggagaagttcggacacaagt gtgcctattgccacgggaaatctggggaacaaatgctgaagttgagcacgttatccctaa aaatcttaaacatggtccaaaaggcacagaccgaataagcaacctgtgtattgctgtga gacctgtaacaaagccaagataacgatcaaccggaagaatggatgcacgcttacagg ctctggaagaccgcttgaccaggagcgtgccaggaaattcccgctgttatgcaacagc ttaagcaaccttaaggacgcgactgccgtgaatgcgacaagatgggtctttaccggc aattaaaggatagggactgccaatggaagccgctctggcgagcaaacgaaatataat aggtcaaaattaggtctacaaaagctcactggaatagatcgtcgtgtgtaggcgtcagta cccagagcagattagcattataaagtgtgactcttgagataaagctgtaggcattgg caagcggcagcgggtgggaccgacaatatgtttccctatccggtatgcgcctaaagt taaaacatttatgggttacaagacaggtgatttagtaagcagtaattccgaaaggcaaa tacgccggggttatatcgttaggatagccatagacataaagcgaagctttaagctaaagt gcttcgatgtcatcctaagtacctgcggctctcaaggggctgatgggtacgaatatgtt ttgaaaaggaaaagcagctctctcccgagcataaagccggggctcccgctgc catga	gtgcgtacccacgcataaatcgggggc ttgtgatgacagccggagcgtaccagc ctgcagcaggccgttaaggccatcggtc acgttatccaggtcatgacacctgggggtc gttgccagctccagccctgtcgttcggcag taaacaggcataggggtcgaagccaggtg gcagcacataaaaagcctggataacattgg cgaggcaaacattaccggcagggaggc tctttatagcaaa
968-969	Ga0172379_10011290 GENOME_ACESSION: GroweLOW334metaG_2_ \$F_3300014208 GENOME_ID: 27073 CONTIG_ID: 11289	atgatttactatatacgaagattgtgatgaaatgtaggacaccgacgaagaacatgata gtaatcgtcgtggttaagagtaaaaaagcgaagataaaggaagatcgaagacgtgg atgttaatacaaaattcacaagaagattgatctgagaagactatccagcacaattcgtatc ggattagatccaggatacgaaaacgttgattgctgggtctcaagttgactgaaacgaag tggaaaagatacttgaaggagaagctaaagtacgaaccaacaagtcagacagaacttg actgaaaagaagatgcacgacaaaatagacgccagaacagacgaagaatgtgcttc	gtcaattaccagcgtttaaagaccattggg ctgcgagctagtatagtcgaatgccgttta cgggattgaacagacgtagtactaacgatt caccagcgggtgccactccagctcgttgc ctggaggtagttagatctaactcaagtcctc attgctagctattaggttaacgaagactgat

SEQ ID NO:	Name	IscB sequence	ωRNAs
	SOURCE: JGI DATE: 2019-01-27.	gtaaattgactcatgtaaatttagacacccattttggaagaaccgtaagaagcataagtttc aaccaacacaaatatacacttaattcagtcctcatcaacatactaaagaagattttcaactag taccttagatgaaagtcagtagtagtagaataatgctaagttcgattcacaaagataaac gatccgagatcaagaactatcagatcaaaaaggattgcaatacgggtacgaaacgtg aagaactatgtcagcaagagacaatacactgtcaaatctgtaagaagaagaagctc gttgatcttcattgcatcacactctacccggtctcagcatggaacagatattcctaaaaact taatacctctttgtaaacctgtgcacgatagagttcataagggtcaagagaaggtacaaaa cctaaattgaatacttttgcgctagtgagtagtaattcaattatgaaagagctttatgaaa tcataagctcggagtgagtgctagtaagacttacggatatgttactgatactcatcgaag agtttaggactagagaaaactcattgcggtgatgctagttatcgcctttgtgacgaaga caacgtattgagtagacggtagtatatcgacaataacagtcataatgttttgcacaattt cgtcgacataatcgtattttacgaatcaagttgaagatcgaagtataagattaatgtagt gttgttcttggaaatcgaatcgacgagaaggacaagataaaaaaacatcataact gagttacaacagaatacggatatacaagattcgttttagctcgagggtataaaatcac gacgaaataataagatatgttatttgcacggagacatgataaaggtagccacatcaa aaagaccgagactttcgttagttacatcgatactgttaaagggttggtatcgacacaagga accatacaggagtaaacctgattaagaatataccaatagatatgtgagcaagaactca ataatggtgctcgtaatcgataataattttga	tatcgttagccaagtcgaggagacttatttt gtgcatgcacaaaaccgatcaaggaaacaca aacatgatttacta
970-971	Ga0120380_1006999 GENOME_ACESSION: 3300011989 GENOME_ID: 239578 CONTIG_ID: 4011 SOURCE: JGI_Dark_Harvest DATE: 2019-01-27.	gtggcgaagagcaccacccctcgcaagaggagaaatccggcggtctatgacgggggaa gtaattcacgaaaggagatacaaaagtgtggtcagtcctgtcagcatctgga aatcgtgctgctcctacaactgcatacaaacgacggaatcttaagagtagacgtgca aagatctacgtacagaccgctgttcacattcagctcaggaccgtgaagaagggtgct acacagcctgttgagctgaagatggataccggtgcacagcatatcgggatcagtgctgc tctgaaaagcacgaatttggacacggcgctacgatgtgctgcacgaaaaagaaat gcatgatgagcgaagaaaaaacggtacagagaaggcaaaaacctcgttatcgccg acacgcttcgataaccgcaactacggccataatcgcaaggagaataagtggtttccccct tctttgaaacataaagaaaacatccacattcagctgcagaaaagatctgtgcgtagtcc cggtgacagacgcttattcagatggggcagctgcatattcagggtattgaaagcctatga agcagcgaagccattcctgtagggggacgactaccagaagggtgagcgtatggttatg ccactctcgcgaggcagcttctgcccgtgacaattacacctgtcaggtatgcccga aactggatgacaagcaccatccgatttggaaatgcataatcggctattggaagcaga tcattccaatcgatgagcaacctgatgacggttgcacaaatgtataccgcagcgaac cataagcccgaggagggctatatggtctggagccagagctccaacattcaaggcg cgattcatgaacacgggtgctgggtggatgtttgcacagctgaaaaagcattgcccga tgtcagatgccatgacatagccaatgct	gtcaaccaccccgccatttactcggtaaat ggacggggcgttcggggaccagaataccc ccgtagcccggttgaaaagccagttacatg ttcgaggagaaatatacaataggcacttcg aacgctgctcaaatctgaacactcggggc tgcaataaacatctatatggggaagggaat gtgtacgtctcggcgactgacagcaacat cgccagaatgccgtaaacctccgataac agtggcgaagagcacc
972-973	Ga0256407_10008120 GENOME_ACESSION: RumenRJG_04DNA_2_\$ F_3300028886 GENOME_ID: 280890 CONTIG_ID: 8119 SOURCE: JGI DATE: 2019-06-13.	atggtgtatgtattgggacaaaataattacctttaatgcctacagataattataggaaaagt gactattattaagaaaaaagaagctgtgttcgttaaaaagaacaccgtttacaattaaattat tacattgactaatgtatataagcaaaatatttcattaggaatcgatcctggaagtaaaagtat tggcttatctgcaaacacgtcgaacaaagaaattttgctggaacgggtatattgagaacg atgtgactaaactgattcaactagaaaagaacatagattatctagaagaagtcgtaaaact aggtatagaaaagaaagggttttaaacgtacaaaaatcagaaacctggatggattccac ccacatatagacaaaagcctgatgttcatttaagataaattatgataaagatttttacc tataactaatttgcctaggttaaccgtatttgatactccctagagccctacgaaatgtatg acatttagacgctctgttttaataagaagataattacgaattgcaatattgccattggtctctg gagataaaactttacaatacatcatatgaaactagacaacctgagggaactcttttga aatttaataacattatgtaaaacctgtcatgctaataatcatagagggtgaaatagatatctac aaacaagaacgaaatccttaaggcagcagctatttatgaacattttaatactagattagat aagtatttaaaactagagggaatgataatatagacataaccaacggcttaatacatataa tactagaataaagtacaatctccaaaagatcctggtagatgcaagatgtattagtggaa atccattagctataatgtgtgatcatttttatcttttagatcagtcgaagacgaaacagacg cattcataaagatacaattcaaaaaggcggaactagaagacttaataatcaccacaaat atgtttggtttcaattgtacgataaagttcattataaaaatgaagaactgtttattatgtctagg agaaaagatggagaattacttttaaaactataggtggaaagatgagatatgaatctgtttc ctataaaaaattgaaatcctgaaccaagaaaagctattattgtacacgttatacaagaaa taactatatttgaaggagatttttaa	gacaggggttagcgattcgaagtaattattat attactagttaattgagatattacttctgctgac tacctaagagattattactctactgttacttaa gaatatataggcactgttattattgctctag tcttcagctcgtcgattgcatattaaagatct gacgaataggatgggtgtatgatatataaaa accttaagataacattggggaaggcgcaatt acatcttcgagatgagtcacttcagttggt gtgactgttttgaggtgatgtattatg
974-975	Ga0315286_10038107 GENOME_ACESSION: YL17G10_0_MG_2_\$F_ 3300032342 GENOME_ID: 281637 CONTIG_ID: 38106 SOURCE: JGI DATE: 2019-06-13.	atgcagtagctgttctgtgttgacaagaacaagcagccgctggacctctgccaacctgcc agggcaagacagcttctcaagaagggaagagcaaccgtatttcgcccgttatcccttcacg atccggctcaaggaccgggagctggaagagctcaccatgtccaccagggtcaaggt cgaccccgccagccggaagcgggtattgcaatgtgcagagaagagagtgcaaggt gctttggcaggtgaaactggcacaccggggactggccatccagagtgccctggatgac cggagagccatccgacgagggcgccgaagccggaatgcccgtaccgcccacctgc cttcgacaaccgagggcgggggagggttgctgctgctccacgctggaagccgggt ggccaatatcacaacctggtaagcggtgaggtgtgtctgctccgtggaagcatct caatggaaactggtagaagtcgacaccagggctgcagaaccggaaatcgggagtg gaataccagcagggcgagctgaggggtacgagatccgagaatcctgctgaaaaat	gtcaacgacccacgcataaatcggggg ctgtgagaaaccacaagccgagttgaccag cctcagccaccagtcgaaggactgacggg gctacgttatcgggagagttgaagaacgca ggttggagatgcgcgagccagcccaacctc tgcacactggatggttaaacatcttgagggtct aaagaagtgcaccacagaagtgctgaccg ataacattggcagggtcacattaccgccga atggcgagaaaggagtaatccctatgca gtac

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gggggtcggaaagtgtcctactcggcgggacaggagtgccctgcaagtgaggcacat cgtcccgaaggaggcggatcaaaagggtcagcaacctgacctatcctgccag gagtgcaacctggagaaggcagccgcacggcgggaagagttcggcaccggagatc caggccgaagccggcagccctcaaggatgcggcgctgtgaataccaccggttgg gagctctggcagggctgtctgagagcggattaccggttcattgcggacgggagga ggaccaaaftcaaccgcacccgttcgacctgcccaaggcccactggatcgaatcggcc tgcgtgggtgaagta	
976-977	Ga0224423_10002744 GENOME_ACESSION: 3300021431 GENOME_ID: 241590 CONTIG_ID: 2640 SOURCE: JGI_Dark_Harvest DATE: 2019-01-27.	atgacagtatatgttaattgactctgcgaacacctcttatgccagcagaaggtcggca gagtcagacatatgtcgtactcgggtgaagcgggaatagtctgctatttccgtttacaatc aaactgagaagaaaagtggagcgaagtgtacacacaacctcgggttggtgttgacac tggtttcaaacacgtcggcataagtatatcaactcgaagaagagaactcttccgttatcatt ccgtcatcgtcacatgaagtgaagaagaacttgaagaagaagaagatagaacc ggaagaagacacagaaaagttcgtatcgaagccgggttcaataaccggtcaagtc gaagaagaagagatgattccggcgacatctcgtcatatgtagagtgctcataagaaga tatagaactgtgttgagattcattccaaagtctgtatagaattataaatctcgaatcgga gaattcgacacgcataagatgcgtgactcagatgtagaaggagagatgtatcagaagg cgacttaaccggcttcgacaacgtcaagcccttcgtagatggcgtagcggaaacatag tcagcattgtcagggaaatccgggtgacaagaagatcagagttcatcatatcaagcaccg tgtcaacgacggctctgataaccagcaaacctagttgtctatgtcacacatgccatac aataccataacggagaatacaattgaagatgtccaatttgaatttgaagactattgactctc ttcggtctgctcgtcgatgaatatagtcagaagatagaatattcgacgaagtgaagaatag ttccagagactgacgtgagaagacatacgggtacatcacacggaaaaaccgtgtgata agtgaacttgagaagtcacatacaaatgacgctctgattatctcaagaatttcaatgctgt ccggaggaacagactatagaagttaaagcatatgagacgtcataacagacagattcataag aaaaatcccaftaaagggtggatccgaaagaaaaaccaagcaaaagcaftttcataaagg atttgcctcaatgatttgcgtgttgacaaccaagcaaccgggtttataaccggccgtat gtctagcggatagtgactatcaagacgattgatgataaagattcagaaaaacagatc gttctatgaacgaatcagattgataagaagagcaaaagggaatgatttatgattacaaaa ataa	gtcaccaactgtctaagtaaaagacttagcg gcttggttctttagggaactgaagccgatggt gactagcctgagtgaaaggagatagccgg atggaatccggcagactccgaactacgtta ctggagaatataacgataggtaccgatgga ataagtcctagtctccgctctacgctcaag agttaaacagtcctgtacggtagggacagtg cattgaagaagaaccccggaataacattg gcaagggtctaccaacaggcgtaacgcc ctgcataatccctaacgggataattataaag ttaaataaaattataaactctcaatgaca
978-979	Ga0224508_10013895 GENOME_ACESSION: SF_Jan12_sed_USG_8_5 F_3300022413 GENOME_ID: 35052 CONTIG_ID: 13894 SOURCE: JGI DATE: 2019-01-27.	atgtcgaatgcagattggtgatcagcagcgatacaaaaccttgaagccgattcaccca gccgtcgtcggcgatgtcgaaatccggcgaagcggcggtgttccggcgatccgttc acactcatttgcaaatccgggggtgaccaggggacaagctgaaaaacgttcgggtgaagatc gaccttgagcagcaagaccactgtgtctagcgtcacttgttgatgacgcgtcgtgtgtgggc gcagagctgaacatcgttggcaacagatcaagatgcactcgaaaaacggcggtcgtt tcgtcggggccgtcgaagccgcaagcgcgtatcgcaagccccgttgcgataaccgg cgacgtgtcggcggtgtgtccgcaagccctctgcatcgcgttgaaccaccatgacc tgggtgcaacgcttgtgtcgtatgcgccgttggaaatctcggtagaatgtgtcgggtt tgatagcagctcatccgcaacccggacattgaaggcgtgactaccaacaggcggaac tctggcagcaagaggtccggcgatgtgttcaccgtgcggatagcctgtgcctact gcggcgcaagcatgtccacttgagcttgaacatcatatccgcggagcaaaaggcgg gtcgaacgcgcccaaacctcacagcatcatgtgtctcgtatcagcggaaggga atgaatcgattgagacgtttctcaagaccaagccagcgttctggccgcacagggccc aactcaaggcgccccgaagatgcccgcgtgtgaacgccacacgtggcgcttagg ggaggaactttgccgcacgggtcaccgggttgaagccggaacggagccagacggc atggaaccgcaaacgacagggttccaaagacgattgggtgacgccgctgtgtc ggtcagtcgacccccgacgcgttgcataatgtgggtgacgcatccgttcagatcgtctgt gccgggcatagctcgcgccgtatgtgttgcgacaagataggtttccacgcacccgc caaaaggctctagccaagtcaagctttaaactggagatcgttgcgtccgtgtgtgcc gagtggaagaaaggcggtatctatgttggccgggttgcgtgcgaagctctggtcatt caatattcaaacgccaaaggagactgttcaaggtattggctggaagtgtgtatgtctgtg catcgtgcagatggctatctcacagcttggagcatcgagctccatgctatag	gtcagctaccgctcgtgactgaagtcgagcg gcttgaagccgagagctgaccagactc agtcgcgttcagtcggtctacgttagggc cgaaataggtacccctgggtggccgaacc agctccaggcctacgggtcaggttaaca ggctcagtcgggtcaagccgtgctgaca cgcaaacgtgcccactggtgtcgaggga gcacatcacctccttaggagagaagaccat gtcgaatgca
980-981	Ga0272428_1003415 GENOME_ACESSION: FingerMtnordMG_2_SF_ 3300031520 GENOME_ID: 280175 CONTIG_ID: 3414 SOURCE: JGI DATE: 2019-06-13.	atgaacgggtgtgtatgtcctttaccagatgggggtgccattgatgccttgcctgcgtgat cgctcgtttgttactcaagatggaaaagccaaagtcgtcgtagaacgccatttaccatc aagctgtatacacagccagtgaaacccctacacaggcattgacgcttggcatagatac ggcagttcagtgatgggtcggctgtgctgatgaaaacgggaatgttctctatcttctga ggtggaaatccgcaatgatattggcaggattgtcaaaagagcatgtctttgatgcaacgatg atcgccacgcgagggatcatacaatcttccgacaaccactgttctcctcaaacgatgtg tgcctgatggagattaccagcaaacgaagggaattcgcagtgaaacgcaatcaca gggaagattggggatttctgaagtcgacaaagtcgttacttaggtcaggaatactcat caaaaggagaatgtctacaggctacgtcatcttattgatggagcttctggaacaaagtgct ttgaacctctcccaagttgacaagatgaaaagagtgagcgcccggtcatcatgggtga tgagacaaaaaacatgccaaagtttcatcctctatcactga	gtgaactaccacgggggctaagccctcgtg gcttcttcggaagcctgagttcaccagactg tcaccagaaatggtggcaccgttcgagagg tcatgacacctgcgttgacgatcagatcg ctgctctgtcgcctgggttaagaaggactg agggaagggttcggtgattccaggctcaaaaa gccccttgaacccctgtcagatgaagccgga ttctctcgtggtcagacgaagagatagc gccacctgcgtgaagcagagcattttctgaa aggaaattcttctatgaacggt
982-983	JGI1684J13235_1001204 GENOME_ACESSION: PonA232FremontCA_3	ttggtaaaagtattgccaaagatggatgatttacaacccacaacagacatggcaag ttagaagattacttgataacaataaagctgaagtgatgttaagaaccttttactatcgtt actttatgaagttgaatcaaaaaaaccaaaftaagggtttattttgatactgggggtaa	gtcaactaccacccctaaattataatcaagat ttagagggtgcttgtgaaaaacagtcgaagcc actagttgactactcaagccttagttgctact

220

SEQ ID NO:	Name	IscB sequence	ωRNAs
	CONTIG_ID: 2131 SOURCE: JGI DATE: 2019-01-27.	ggcccgaactcctgggaagcggagctgcccgcgttcacagattatccatacacga ttcggctgggtgaccgtaagaaggagctgaagcggcgggtgtttaaagttgacca ggggcagtcacacccggcattgcttaaatcggcaagaactccacacagaaaacacc aagccgtactccacttggccgagctcacccatcgcggggcacaagctcgtccgcgcta gctcggcgagcggcatatcggcgccggcgccgacccgaacccctgctaccgggct ccgcgatttcttaaccgaacgaaggacaagggtggctgggtccaagcctggcgctacg cgtcgataacatcctctcatggcagcagcgttactccaagctgggtccaatcactcgatc gaaatagaatctgtcagctgttgcagcagaaagataaataccagagattggaatcg gctaccagcaaggtacgtagcgggctacgaggtacgcgagtatctctcgaaaagt ggcgccgtaaatgcgctactgtgacgcaaacacgtccacttcagatcgaccacatgt tcctcagaaccgaaaaaccgttggcttcaaaaggggtctaacagaccaagtaatcgac cgttgcctgagctcatgtaataagagctaaaagcaaacaccagtcgaactcttcttgcg atcaacccgaccggcttgaatacattcttcgcatacaaaaggaccgctgaatgccgcg	cgcatcggcaaggtaaaaggccgggacga gcgagacggattcgggcaagaaaggctt aaagcctcgcgtgtgtgacagactgc
990-991	Ga0181296_101698 GENOME_ACESSION: IMG_3300013886_F_3 300013886 GENOME_ID: 28693 CONTIG_ID: 1697 SOURCE: JGI DATE: 2019-01-27.	atgaaagtatttgggttgatccgatcaaacaccccttaattgcctatgcatgctgcaaaagc ccgcaaatattaaaactaagaagccaggtgaaaagccggaagccgctgttatcaa cttcttttgcggccgaaacatttgcacagcctgtttaaagttggatcagcaacggcg caggagagaccgcatagctgttgcaggaagagttaaagccggaggttactctctt agtgcgagagatcccttgcgacgacgacatttcaggcgcttaaaagttcgtcgacttca cgaaggcaagacgcggcagactgcttatcgttaagccgcgttatgataacagggctca gggtgaagtgcatttgcggtaagaacggcggaaggcaaaacacgtcaaggc gcacagggcggttaagccggagataaaatcagcacgaacccctgctaccccttgc cttaaaagcgaaaaagactcaataaagagttatataaaagatccggcgctggggccccg gttagtcttgcacaatccaatgggaaggttgcatttgcataaaatcgacccgctcgcat taatgaagcgggttatcgaataatccgtatataaacgtgatacagcaaaagcgcctctt attttgaatatggtcgcagacaaaagaggggatacagaaaagattatccacgtgtt gttattgcgataaagaggaacggaggtggaatcgaacatacttccccaaagttaaag gagcggtgatagttgcgaaatcactgcctgttaaggatgcaacaataaaaaa ggcgacatgacccctgaagaagctggaatgacctgttatagagcccaagccgttca cctgtccaggggtctacaagtatatcatgctgtcagcaagcaagaattacctgttgagc aatacaattctttggaatacctgtagtttacttatgacagtttacaagctggcagagaa aaaggttggaaatgataaaaagattatgacgactcaattttaaattccgctgctgatac acgacagaattagaccgggctgcccataaaaaagttgcgctttttatctggctaac gcctacaaaagcgcagacagataattacgcttccattatctccaagcaagagaacgcc aaaaggttctccagcgaagctacgggtgactttgtttacactttaagactctttagaag tcaataaagcatgcttatcatgtgggacagagctaaactcgtcctaaaagctatfaagaa aagcgaaagcatccgaaaaatgccgtctgttgggttaagaaggagatatcatataactt tgcagtcagcaaaaagattacagtagagctacgtccttgcagtagcaac	gcagctcccatgtattgtgagccagcttca gcgacaacccccagtaatgggaataagg atgagctacctttgatgggttgcgaaggtc caccttgggtgtgttcagctccaggtc tggaaactttgagtaggaaactcgcctac gttttggagggaaggaatccgagataattca ggtacacctcaaggaaggttaccattaa agtttagcagggagcgtttaccgtcttita gaaagagcgagtagctattataagcccat aatcaaattagagttgataatggggtactcat ttatgaaa
992-993	Ga0335394_10010380 GENOME_ACESSION: MAT03spassembly_F_3 300032456 GENOME_ID: 280657 CONTIG_ID: 10379 SOURCE: JGI DATE: 2019-06-13.	atgcaacgtgtattgtgacgacaaaaaagcaaccgctacccctgcatccggca cgagcgagagaactgcttaataagggggaagcggcggtctaccgtacttttcttacc atcatcctgaaagaacgtgagggcggtgatacacagcccatcgcttcaagatcgatcc cggcagcaagcgactggcatggcttttagtcgacagcttcaagcgtggcaagcgggtg atttggcgagcgcaactcaccatcgcgggcagcagatcgcatgctcctgtcaag cgggcaactgcgacgttctcggcgcgacacacgcgctatcgacagcgcgctt gataatcgccgtcgtccagaggggtgctgcggcgttcatfaaaagccgctggaga acatttggacatggcgatgctcctgaaccgtgctgttccatcgccagtatcagtcagga attgtgtaagttgatagcagttgatgcagaacgccgaatcaggggcggtggaatacca gcagggcgcaactgcccgggtatgagtgccgagtagctactggaaaagtggggggc caaatgtcctactgtaatgcgaaggacttaccgctgcaagttgagcatatcaaccccaaa ttgcgcgcgcgagcaaccgtgtcagcaacttaaccctcgttgcacgactgcaacca ggagaagagaacacagacgggtggcggaatttgggttcccgaattcaaaagctggcaa aagccctctgaaagacgctgcggcggtgaatgccacacgatggcgctctaccatcgt ttgaaacaatcggactgccattggaggtcggcaggggtggcgaccaaagttcaatc cacaacgcaagggttacgaaaaactcactggctggacgcgacgtgtgtggcgagagt ggcgaaaaagcttatatcacgctagcgatacaccgttgggtttaaagcaacgggtcac gggtcagccagatgtgtcgtaccgacaagtgtgttccatcggttaccgtctcgttca gaagcgccatttctgttttcaaacggggacatggttaaagctatcgtcctgtcgtcgaag gttgtggagcagcatagtggtcgcgtggttgcgtgcgacaggcagtttgcacatcaca acagcaaccggaagaagtgacgtatcgatcatatcaaggttaccatcacgccga cggctacaattataagaaggacaaggcgctatccctccccatgcctaa	gtaagcatccccacgctaaaggcgggg cttttagcccaatgcttaccgactacgct cgcaagagcgctacgttagcgcgcaaatat aggcacttcggaatgctacgacgttccgaa ctctgcggtgaacgattaaaaagatgcttg ggagcaaacagtgctgttcgaccaaaccc tcgtataacttcttaggcaacattacacc gtaagggaagaaaagttgaggttaactcaaa tgcaacgt
994-995	human gut metagenome genome assembly, contig: NODE_621_length_2590 2_cov_19.039652, whole genome shotgun sequence	atgcgcggcaagcctctgatgcggtgttgcggcaaaaggcgcgacatgctgaagg cgggcaaggcgtcgtcgtcgtcgaacgccgttcacgatcaagctgacctccac ggcgagacgaagcagagacgttgcgttcgatgcaggcgcaaggcacgttgg catttccgccacgacggaagaaggaggtcttgcgtccgaagtcgcgttcgacagg acatcacgggacttctgcccgtcgtcgttccgacgtcgaaggcgcaatcgaaag acgcgtaccgctccccgcgttcaacaatcgcgttcgatcaagcacaaggatggctt	gtcaataacccccgcctaaaggcggaggt tgaaagagcccttattgactagctcagcaaa cctcctctgggagagagctacgttgggtg gaattgacaggcaccgtgggtgtttatcct agtcccacgctcgcgtctgtgtttaaagt tctgagaggtaggacgggtgatcgacaaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
	GENOME_ACESSION: USZF01.1 GENOME_ID: 17063 CONTIG_ID: 620 SOURCE: WGS DATE: 2019-01-27.	gcgttgcgttgaaccgcattcaggcgcacatgctcgcgcacgatgcggtctgcaga ctgctccccgtcaccaagatcgtgattgaggcggcatcgttcacgttcagaagatcaag aatccgactattgaaggcacggactatcagcaggcgcaccagcttggattctggaacgt gcgcgagtagcttcttttcagagacggatcatgttgcacgactgcatggtcgttcgagg gacaagatcctcaacgtgcatcatcttgagatcggaacacggcggtgatcgccaaa caacctgacacgctgtgagacatgccacaaggcttatcagcaggaagatcaagtt gaagggtcaagcgcgggtcaatcgttcaggcgggaagccttcattgggcatcatcgctgga cgtgcttgcaccgctacgcaagacgaccctaaactgcctgctgagaacacctacggc tatctgacgaagcacaagcgcattgcttggcttgcacgaagacgcattgcgccgacgcc ttctgattgcggaaatctgaaagcgttgcgaagaggatcttcttccagcaacaga cgcaaaacacacacgacagatacaggtgttcgattctcaaggcggagtcgaaa actcaatcaggcggccattcctcgtcaagggttccgctattcgacaaggtaagaatcgg cggacagattggtttgtttcgggcgacgcgttagaggcatattcaacattcgtcgcttg acaaaactgtgatcggaagacatcaattgcaaaaaactgagcttctcgaacacgca aaactttttgattgaactacgaaggagtaa	gaaaccccttccaacattgacgaaggatgac aatcgcccttcggccgatcaacaccggcc ttcgggccgagcaagcggagcctgcgggt atccgcaaggagatactttgaaa
996-997	JGI11958J13698_111217 4 GENOME_ACESSION: Elkslonscriptome_39_\$F_3300001229 GENOME_ID: 25896 CONTIG_ID: 673 SOURCE: JGI DATE: 2019-01-27.	atgttacgagtagcagtttatcaaaatcaggtaagccgttaatgccaccaaacctagcc gcgctagacgttggtaaggatggcaagccaaagtggtacataatgacttagaatgctt tgctattcagttgaccttggagacaggagagggaagcgcaaccatagccatggggatag accctggtaagggttactcagggatcgggggttcagtaagtcattcaccctcggatggg acacctagtctaccgttaagacgggtgaaagacggatggaattacgacggattatcggt cgagctagacgaggaacgaattacgggaagttgccctactctgaacgttgcacatcgt caagctcgggttgacaatcgtaaaaagggttaactgccaccgtcaatccgagctaataag caactggaattgcggtatgtaaaagagttgttttaactgtttccgattagtgccattcattg agttggtcatggctgacgtagataaaaccagtgggcgcaagtcggctcggcttgcggttgc gtttctcccagtgatgtaggtcagaacaaatgcttaagtgttgcgtgagttacgacc gtgataactaccaaggatggcaaggacgggaacggaaccagtcagcttaggcaat ggttaagattagccaaagacaagaagacaatcgaaccagacaccagctacacaagc cgtgtatggtgtaaccttagccgctgttgattacccgatggcaggaatggcactctgata atgccaaagcatggtgactggcaaggcagtgctgaagttacctcagcaccatttgcataat ccgtagaccaccaattagccgtagacagttgcatctgtgttccatcaaaagggggtaaa cgccgcaagtatgctggcagcagttaccgcatgggtttaggaaaggcgacaaggttat agccgaaaaagctggaaaaacttacgtcggctgggttctggagacaccgagaacaa gtttcagttagcgggtattactggaacgacttgggcaatttagtgccaaaaagtcaggt gttgcagcgaagcaggggattaatcgtgcttccctcagctgcaatctcccttatt tgaagggtcgaattga	ttcaacaaccccacgctactagcgtagagc gtggggatgcccgaacacgaacatttaa accgttgaataccgctcttagtctcagcttgg cacagacctccggatactctcagtcggga ttaccttaaaacttttgcgggtgtgttgg acaagacatcttagctgagatggcggaag ggactaattacttactggagggttatcaca gtttacga
998-999	Ga0206349_1775808 GENOME_ACESSION: DieMetAnnotativ2_7_\$F_3300020075 GENOME_ID: 280099 CONTIG_ID: 2787 SOURCE: JGI DATE: 2019-06-13.	ttgagtaacgtttttagtgcacaccaacaacagccattaatacattcatcctggaaa ggcacgatttctcctaataagagggaagcagcgggtctgaagcatttccgttcacaatc attttaaagggtgaaatttccgatccagtggttcagaaactgcgaatcaagatcgaccag gctcaaaaccacggggatagcgtatcgaacgaccagtcggcgaaagtcgtcttgcga gctgaactatcgacagaggtcagcagatcaagaagacgctggacgacgcgctggtg tacgtcaggcagcgcaaccgacacactcgtacagaaagccaggttgcacaatcg acgcaccaagaagaaggctggttaccacatccttgcagagtcggtacagcaatatcat cacgtgggtccagcagcttgcgaagctcgtccatcagacagctgtgagcctggaattatg gaagttgacatgcaactgatgaaaacgtgagatcaggggattgagttaccaacaagg tacattagcaggatagaggtgcgtgagtagtcttggaaaagtggggacggaatgtgc ctattgcaacaaggatagcgaacttcagattgagcacattcatcctcggccaatggc ggtaccaatcgatatccacctctttagctgcgagaagtgcaaccgtgcgaaaggaa acc	gtcagggaccgcatccctaaagggtatg ggcttgcagtgatccggtaattggattgtac aagcttcacgttccagactcagtagggcaa ctagctacgttagaggcgaaattagttacca taggatcgaggccagtcgaatgcgctacg gttgcaggttaaacaggtttacaagggttagt ggcagtgcttgcacgccaaacctcccta acattgtcaaggccaccattaccgggtaac cggagcgtctcaaggagcagcacttga gtaac
1000-1001	Calothrix parasitica NIES-267 DNA, nearly complete genome GENOME_ACESSION: GCA_002368095.1_AS M236809v1_genomic GENOME_ID: 133458 CONTIG_ID: 0 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgcgtgtacctgtaattatctcagtgcaagccgttaatgccgactaagccaagtagag ctagacgggtgataaaagaggtaaagcagtagataagtggtcaaaagttaaacctatttta cgtacagcttttaaaacctgattctggcaataaaactcaggatgtagtagtaggtatcgacc caggtaagcaatttagcggaaatagcagtttatcccaaaaagatactaccaagaggagta aaagacgtaaaatagcgcgacgtttacacagcagcgatcgcgaaggtgattatgtcg aggctattaaagctgataaaactaccgaggggtggtaagtgagatactaaaactcaag tatcaatttctgatgctaattggaagcgtctgggacagtttaggtatctaagtaaaactct acagcgtctccagggtgtaattagtaacttccaaaaatacaagtcacggcgaataa	gtcagtaactcaggtaagacagacactgag cttgcggagccaatccagcaactgaagt attgactagcccactgagccgttactcggt aagacttccgaattgttccctagttcgattat atctaaactcgtttggctgagtgctgacggc aataatgcgggaagaactctacactgcattt tgccttaagaaggacatcttagtagcgttg ggcgaagagacttaacattacgaaggat tatctcttatgcgt
1002-1003	human oral metagenome genome assembly, contig: NODE_37_length_11894 9_cov_9.025561, whole genome shotgun	atggtttatgtaataagcaaaaggacagccaataatgccacagagaacatgcacaa ggttcgtctacttctaaatcaataaggcaagggtgtcacaagaacaccattcacaatac agttggtcagtacaagcaaacgtacacacaggagataacgctcgtgtggatgcagga tctaaacacgttggtctatcagcttcgacagaaaagaggagttatfttcagcagagcttc gcccacgaaatgatgtcgtaaatttgatgtcatcaagacgagagttacggcgttctcgtag	agtttgatcagtaaccaacgacttttagctgtg ggattttacaatcccttactgattagcctaagt tccatgagaactcgttatctgagaatatata ggcactttgggacgtgagtcggaatccga actctcggcgttcagttaaacagtctcga

SEQ ID NO:	Name	IscB sequence	ωRNAs
	sequence GENOME_ACESSION: UPUN01.1 GENOME_ID: 16847 CONTIG_ID: 36 SOURCE: WGS DATE: 2019-01-27.	aagtaggactacagctatcgtcaagctcgttttctaataaggattacttaaacacaaag gatggcttgaccatctgttgaagtaaaatctggaatcatattcaaggtattagttgataa caaaacttttgctattaaaactatctgcatagaaacagcagaatttgacttacagagactca aggctttggaagcaggagagcctattccattgtgtaagattatcaactcggagagatgta tggacactacaatgtgagacagatgttctacaccgtgatggctattctgccagtgttgcg gtgctcatagcagagataagaaggtaaaactccattgttaccatctggaaccgccga agacggcggtaacgtcctgacaacttgataactcttttgaagattgtacacagggtat catgctgggactgttgactaccaacgacaaagcgtagaagaaggtctactcgcgacgc tacctttatgggaatcatgcgaaagacactgatagaaagcgttcacaatatgttctagaca taaatattttagcatttatggctacattacaaaactgtgcgcgagaagaaaaataacc aagacacataatagtgatgcctttgtatagcaaaatcttgatgcagaagactagaaa aggctttgtgatgttccgaaacgacaacataaccgtcaaatcacaatgcaagatcaat aaagggtgcacgagggaagatgaaccagacaccaaatttgttgcggctatcaacttttg atagagttatgttctaggacaggaaggtttatcttcgaaggcgttctagtggatcgtttg acattcgtaaactaaatggtgagaaaaacaaacaaatatcaactataaaaagttaaagca cttagaaagccgcgaaggcattattagtttctattga	ggtaggacagtgctgtgaagcatataaaacc ttggaataacattgctacggacaactaac ctagattactaggagataaacctgctttgat agggtcgtgaacaaggagtataataatg
1004-1005	Halomonas pantelleriensis strain AAP genome assembly, contig: Ga0075190_101, whole genome shotgun sequence GENOME_ACESSION: GCA_900102875.1_IMG - taxon_2654587879_annot ated_assembly_genomic GENOME_ID: 183751 CONTIG_ID: 17 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	atggcgggttttccagctcgacaacacaaacggccctgatgccgtgcagcgagaagc gcgcccggctcaacaagctacgttttccgcgtgactacgtgatcgacagaacagatc cgtggcttcagaccggcgatagggttcaggccatcgtgcccagcgcaagaaagcc ggcgtgcatgtggccgagttgccatccgcaagaccgggagcttcaacatccagaccg aacagggtggcgttcaggcattgctgctgcctgcctgcctgcctgcctgcctgcctgc ggctacgactatcaccagacaccaccccgtaacgacaaggaggagcgggacgggc ggtagcgtga	gtcaactacctctccctgaagagagaggctt gtgaacacaaagccgggtgaccagggaaa gctgtagccaaaccgctacgtttgcaacag gtcgcaagaccaccgccaccgtgcttcc cagcgggtgggctctggaaggtcagaatcat gtcggcgaaaggttaagcggcgaaggctc gtatcgtacggcaacgtaggagccgggtg cagacattccgaggggagacgggcccgat aggcccgcgacaccaggcccgtaaaggc aacgacacggaggaaatcgtatggcg
1006-1007	Ga0136257_102499 GENOME_ACESSION: IMG_3300010222_\$F_3 300010222 GENOME_ID: 306173 CONTIG_ID: 2498 SOURCE: JGI DATE: 2019-11-15.	atggatcaaaacagaatctgtacttttcgcagatggcagcgaagcgatgccttgcac cagcgagggcacgacaactgctcagcgggataaagctcagtatatcgataccagcc gttcaacaactcttactgaacgagaagatggcgatactcaggatgttagtctcaaatga ccctggaagccaaacaaactggattagcactgttggggcatttccaaagaaggaaactcggt aatitggcgacgaaacttgaacacagagggcagatcaaaagaagctttgagggaaga agagacagattcggcgacgagaaggacgagaaaaactgttatcagagccctgggt cgaaaaccgaacgaaatcggaatgggtgtgttccaccgtccgttcaatcgcattgataat attagagaatggcgaaacgactaacgagccgttctcgttgcgcaattaatgcgag acggttctgttgaactcagaaaattcaaaatccagaaatcaaaaggaaaccgagatcagg acgggaccctgaaaggctacgagctcaggagatctctcgcagaagttaatacactcttg cgtatctgtggcatcacagatgtactttggaactcgtacacgtaaaactgaaagccgt gggggttctgactgcactagtaattagtgtgtctgcacggctgttaatcgggacaaag gaagtcagcccgtcaggaatttgtgaagacgaagaaaactgagttggatcaaaaag agacagatgaaccatgaaggacatgggtgtgatgaacagatccagtggaaagtgtg gtgaagtacttgaggagaccgggtacttgaagctaccactctggaggagaacaaat ataatcggaccgatcaaaactataaaaaagaccactggatcgacgctgttgtcggag aaccgaatgtgcataattcaaaacatctcgtgttgaataacagagcgaaggttaggg gtgatcggcagatgtgtcagtagatgccaacgggtttctcgcacgtctccgaagcaata caaacgcatcgtgattccagactggagatctcgttcggcaggttgcctgaaaaata caaaactagcggcagcgtatgaggaaagtcacgattcgggtctaatgattcttcgcgat aaacacgagagaagaacaagttgatggaattaaactaaaatactgcgagcctctcagcg agctgatggatactcctactcacaagaaccgaaacaaactacgcactcctcttattgaa	gtcaacgacctccgagtaaaaccgggagct tcaacagaagctcaggttaccagcctaag ctatgctgaaaggctacgactacgttacggg caaaaataacgataggcaccctggagtgcc ctccagctccaggtctgctgcttctggat aaacagactgagggtctaagtacgttccagt ggcaacaaactgtccgataacattggcga gggagccctgacctgcaagcgcgacattc gcgtgaagcagagactgtgagtaactacta ctaacaattaataatcgaacacatagcgat ggatcaaacaga
1008-1009	Ga0137385_10003252 GENOME_ACESSION: SagtaG_30_\$F_3300012 359 GENOME_ID: 35417 CONTIG_ID: DATE: 2019-01-27.	atgtcacagggtgttgccttagatagaataagcaaccgctcaatccgggtgcacccgggtg tgcggcctgctcttgaacaggggcaaggctcgggttaccggcgctacccgttcacgct gatcctcaagcgtacggtggagcgaccggagcttaccgctgcgctcaagatcgatc caggcagccagaccaccggcctcggcctgtcaacgatgcgagcgaggaggtcgtgt ggggcgccggaagtgcgccaccgggggggacagatcaagcgcgccctggacaagcg gcgtgcggcgccgctgtagcaggagcgagcgaagaccaggtaccggccggcagcgt ttcggaaaccggagcctcgcacggggggccctgcggccctcctggagagccgggtgt gcaacggcctcactgggtccgctgtgatcggtctcgtcccgctcaccgccatcagc ttggaactcgtcgggttcgacacgcaagccctcagaatccgcagatcagggcctcga gtatcagcaggggacgcttggggctacgaagtgcggaatatgtcctgtcgaagtga	gtcaggaaccggcccttagaagggggca ggctgtctcaagaagaggaaagcctcact cttgaccagtctcagccaccagtcttctcctc ggaagagggtgacggggctcattcagg gcgaatgatagggacgtccgggtgttca ccagcccgacgtcgcggagcagactta aacaggttgaggggttaagccagtgtgt gctcatctgaacccgctcaggaactgtgag aggtagatcaccggcgcaagcggagcg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		atcaccagtgcgctactgcgacgcgcgctgtcccggttgaactcaccacgtgcagccgagagccaaaggcggtccaaccgcgtcagcaacttaacactggcctgcacttctgtaaccaa	tcgcaagagcaaaggaaagcaaggaaatggctcgtatgcacag
1010-1011	Ktedonobacter sp. 13_2_20CM_2_54_8 13_2_20cm_2_scaffold_1 0398, whole genome shotgun sequence GENOME_ACESSION: GCA_001914965.1_AS M191496v1_genomic GENOME_ID: 113894 CONTIG_ID: 13 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	gtgttgaagttcgacccagggtgctgccccagtcttccgtgctgcaacctctgacttaaagaactcgggtccaggaaagggtggtcagagaaagtaccgcgtgctatccccgtcaggggagctgtttcactccgaaaggagccttacctgatgaagggtttgtttatcacaagaaggaagcccttgatgccacgactccacggcgcaagagtggtgctgaaggcaaacgtgccccgtctcgtgcgccaagagcctttaccattcgttgcgctttgccacaaggcgcatgtgcaagcggcggaagggtggcggtgatactggtccaaaggcgtgggcattgctgcatcgccaatggcgaggtgtgtctccagggcgaggtccatctgctgtagatcacacggagaactgacccaacgacgccagtttggcgcaatcgacgtgctcgaagacgcgtatcgtgaagcgcgtatgacaatcgacggcgacctgatgctggttcccccttctctgcactccaaagcagaggcgaccgtcaaggcggtgctcgtcatgcccctcttttccggttgctgggtcagctgcaggttgccgttttgacacccaaggatgcaaaacctgacatcgcccacttggagtagcagaaggcgaaactgcaagggtacttctcgtgtagtatgcttagcaaaagtggcagaggacatgcgctactgtgaggcgctgtaggtgccccttggaaactcgagcatattgtccccaggtcaaggggaggaagcaatcgggccaagcaatcaccctggcctgcatgacctgaaccggcgcaaggggcagcaaacgcagcgagggttggcttccagacgtgcaaggcgtcttccctgaaaagccgggtgtccaggacttgaaggcgtc	gtcaatgacctccccacagagtgaggaggctgtgagggggtcggccctcgcaagccggattgaccagacgacacctccgagccagtcgcggagaaagcgagcgtcgtggctgtgtctgaagaggtggcgatagaccacagtggtgaagttgcacccagggtgctgccccagttcctgtcgtcaacctctgacttaagaactgcggtccaggaacgtggtcagaggaaagtaccgcgtgctatccccgtcaggggagctgtttcactccgaaaggagccttacctgat
1012-1013	wastewater metagenome genome assembly MBR_assembly, scaffold 53832, whole genome shotgun sequence GENOME_ACESSION: CXWK01.1 GENOME_ID: 678 CONTIG_ID: 53831 SOURCE: WGS DATE: 2019-01-27.	atgttagttccagtcagtgatttggacgggaagccccgctgatgccaaagcagaagcacgcgcgtatgcgctgtatgcagcgcggagacgcgacccccgttttggcgtaaaaggagtggtgtgtattcggctcaaccgtgagccgaggtgcccgaatcttcagggcattgtcgtaggagtcgataccggctcgaagcgggaaggttattcagtagtctcggagggcgacacagattgccaacgtccaagccgacgcgataacgcacgtcaaaagcagcagtcgaagcacgacgtaacgcacggcgctctcggcgacacgtaaacggcacaccgtgcgtgcccgttgaaccgtggcgctacgtcggcaggaatgggtgctcccccgctcaacgcgggcacgtgggaagtgaactacggaacatccggttctgcccggcgtgtacccgtcactgacgtggtggtcgaggacatcaaaagcccgacccgcaaaaggacgaggaggacagtggaacgggtcgttcagcccgtggaagtcggcaagcagtggtttctacgtccaagttcagaagacatcggctgtacttgaagcaaggctgggaaccggcgagattcgcaaaacattggccctgtcaagtcgtcagacaagatgttgaacgttggagcgcgcatgtgtatagcgttgggtatcgccaacgacgcgctgggacagccacatgccgtctcggcgacacatcgtgctggtcaccaaatcgtcgaagtgtttacggcgacagttgcatcgttccagccctcaaggcggtgaacgtcggccttacgggtctacgggtctctaggattcaagcggggtccatcgtgcgacaccgaaatgggagtggtgtctcgtggagggacctccagggacgcatcagcctgcattcgcaggaaaccaaaaacgtctatgccaaaacgccagagttggggacattacgttctcgcgtataatgataggagccagcttatgtccctcctggaatgttgggtgtctccacgtgggaatctga	gtcaataacccccccccaataggaggggcgttgggaagcgaagaccctcgcaaaccgagataacaccgttgactacacaacactgaatgtgaactgttacttggcggtggcagcccaaacgacgctagagatgccacctagtccctagcctcggatcaggagtggtgaaggtgcgtatataaggagcgtaaggccat
1014-1015	mouse gut metagenome genome assembly, contig: NODE_2358_length_222 09_cov_9.968132, whole genome shotgun sequence GENOME_ACESSION: UWSU01.1 GENOME_ID: 279226 CONTIG_ID: 2357 SOURCE: WGS DATE: 2019-06-06.	atggatatactactgacaggagggaggcgctcatgcccagacagacacgacgcatcgtgctgcatcctgctgaaatccggcagggcgtcgtgacgcgctgacccgttcaccataaaattgaactacgacactacataaatgtacggccggtcatcctcggcatcgaccggggcgacgaacatagggctgtgcccgttaacggagggcggaaccgtgttaccggaggtgctggacgaggaataaggacatccctgtctgtagtggcgcggaaggcatctccgtcagggcgacaggaagcagggcggaagagagcgcgacgtggcaacgggaacggcgctgaaggacggaaagatcgagcgggcgttccctcgtacgggaaggacagggcgtacatgtaaggtgatcaggaaataaacaggcgcggttctcgaacaggaaacggcctgacgggtgtctgacggccacggcgtacagctactgcggacgcatfataacctgtaaggaggtgctgggttctccccgtttcaaaaatcgtgctggaactcaataattcgcttcatgaggctggacgacccctcaacgcacggcgacatgttccaacgcggcccgtgtacgctatgacagcagcgttgagtcgcaggtgacgcttcaggacgggaagtgcctgtcgtcggcgaaaccataacagtagaccaccagctgagggaatgcagcgtgacgggagcgagaccgtgaggaaaccgcggcggtgtgacggcatgccacagatcgtgacacggacgaagcgggcgcggaaaacttgcacgtcggcgggcatgcggaagaagcatgacgcgtggcggtgctcaaccagatcataccgacctcgtggaagggtgtccgtggattatgacgtgtccgcaacggcggggtggagacgaaggagttcagggaacccacaataatccgaagggtcaccacctgcagcgtacggcggtgctcgtcactggaaaactttgaagtgtgttccaatgaatgttaccacatatgccagttcaggaggcatgacagggaaggcatgtgaacgtgagatgtacaacaggaaactatgtcctgacgggaagggttgcggcgagaaacaggcacaagccatggggcagaaggcgacagccctggaagaatatacgaaggcggggagaaccgaccgtgaagggtgaagcatcccgaggcgatgaaggacatggcaaggcattacccggcgtccaggttgatataacggcaggatgaaggacgtcgtgaagcggcgctcggggagctactgggtcgacgacggttcgaaaagccttccgtccgaagcggatgtgacactgaacaattcggcgctggtttgtatcaaacagtagcttaa	ataaaatgattcgtgacaaatattccgttacataatgaccgggtgttgtaggtcatfcttctgaccacatccacgggtcgggtacaggcatattctgacggcgtcgtgtgccccctacggccgtgaccccgggatgtgttccagtcgcagcactgcagagccggcagggaagccgggggagacattaccccgcgaccagcgcggggagttatcatfcatgtaaggagagacagtaaatgttatat

SEQ ID NO:	Name	IscB sequence	ωRNAs
1016-1017	Ga0223845_11796712 GENOME_ACESSION: 3300021387 GENOME_ID: 241586 CONTIG_ID: 38073 SOURCE: JGI_Dark_Harvest DATE: 2019-01-27.	ttggcgaagtggaaacacgtgcggtatgcacagaacccgctggacatatgccctggcgg aagacgtaagtcattgcaaccgttttttaaggaaaggaggcgctagtgctgacatga gagatgtcgcggtgatctcgaagacaggcacgtcactgatccgacgtcgggtaccg cgcaaggaaacttcttaagaagggaaggccgttatgaaaaataccgccctgtgttcac catccgcttacagaacgcgaggacgggtgaggtgcaggacatggaactgaaggaggat acgggtacgtgcatafcggtgatctgtctgtccgaaaaacacgacatatgcaccgtc agtacgacctctgagtgatgaggtggagaacacacgaccgtctgaaatctgcagg acgcgccgtacagaaaaacgttacaggaacccgccttgacacaggaagagctttat cacaagaacgggttttcgcccctcatcagggaacggcgtgacaggcatgtggaatctgt gcgtgagatgtgcggttaagccataaagcgtgcctacatcgaatgggacagtttga cacgcaggctactgaaagctgtgaaactggccttccgataccgaaggaggattatca gcacggagaacagctacgggtcatgacactcagggaagcggtatcacaaggagacgc catacctgtgtgtcgcggcgctccgttaaggaaaggccatcttcacgaacaccata tcgggtctggaaaggcgaccgcacgcatccttccaaactcgcacggctcgcgaa aatgccacacgcttcaaccacaggcctggagggaagctgtatgcatgaatccggt catgaagccctgaaggaggcgacattcatgacgaccgtcgggtatgacatgctgagga ggatgaagaatccgcccgcacgtggcattcatatgacatacggcgccgcgacaaa gctgtcccgcagaaatctcggcattgaaagacacacgcgaatgacgctgtatgccatg gaaaaatcacatccgaacacgcgtgcgccaggaaacgttcggaagagaaggcgaa acagccgtgtactgaaaggtctatgacgcgggtgacatcgatacgcgggatggcgtga agaagaaaggttcagagcttgatgataacagacgaaccgccgtgagacgaggtgtc tcccaaaaacgaacgtatctttcacgggaagagggttcggcgggacaccgttccatcag gaaagaacgtacacgaatcacaccgggaagcctggtaaaatcagacagtgaggtgatg accgtaaagggtatccacagcccggttcgggaaccaatgtggaattcgaaaagccgc gtcaaacggaagaaatccgctctctgaaaaagctgaaagtactgaaagttaacctgtt tcaggatggagcgggttaaacactttataa	gtcaactgaccacctccatgcgggaacgt cccggcgttcatcctgtggaagggtgctgc gggaagaaatcttctgagtcgggttattgac tgagcaggaaacaggtcctgggtgaacccgt tcctgtcgcactcacaagaggtgaagtgact tccatgttttacagacagaagttaccggaa gagaactacgttccggcgaaaatacaggc accatgggatactccacagctccatgcaac tgcgccgtggcattaaacatccctgagggtc aggggaaggtgtcgtcggatataaaaccgt cggataacattggcgaagtgaacacgtgc ggatgcacagaacccgctggacatatgcc ctggcggaagacgtaagtcattgcaaccgtt tgt
1018-1019	human oral metagenome genome assembly, contig: NODE_892_length_1639 4_cov_3.332334, whole genome shotgun sequence GENOME_ACESSION: UPUY01.1 GENOME_ID: 16858 CONTIG_ID: 891 SOURCE: WGS DATE: 2019-01-27.	atgggtatgttttcaaaatcgggcaaccgcttatcccaattcgtagacacggtaaggt cgaagattatataaagaagggaaggcaagggttactcatagagaaccgttactatcgg ctactagtgaactgaaagttaattatcagatttaacacttgggttgatacagggttcatca aaaataggtgtgtagtattacataaaagaagggttttatattatcagaagtaaaatc aggaatgatatctgttaagatgaacggagagcgaatatactgtagagcaagaagaac aaggaaaactagatatagaatacaaaatcttaaatctgtgtaattctataagaaagata gatttagtctactatgacaagcaagatttaacatcagatagagagattgatttatcaagt ctatcctgcctattaaatatttaaatgaactggaacgttcgatacagatttacttaaatat gaaggagaggcggtcaatcgtcattggggatagcagaagggtcaaaatcagagttcgt aatcaaaagcagctgttttaaatagagattatactgtcaatgttgcataaaactaaaaa ggaaacacttcatattcaccattgttatcgtcaaatgtgtggcgagatactttagataat ctaattactttgtgtcagattgccataaaaactcaccgcggagaacttaagtactttgaa agtaagttagctgttaagaaaaagggaactcaaacacgtacgcagatgaatagtatt aggtttcaacttttaaaacattatccagaagcaatcgaacatttgggtttataactaaagaa aatagacaacttataggcttagagaaaaagcattacaatgacgtgttgaatagctactgg atgttaataaagcctaagtttttagtgatgtgtgtataagaaaatattgattgcaaaaggt cgttaccgattacatcaaggtcagcgttcgaatgaagctcccaagggaagggtcga aggttttttaataaggacattgttaatatagaggtgtgaatttttaataaagggttaataa gtcgtggtgttattgtcattgatgaatattgatggcattacacagagatttgagaatccga aaaatgttaataaacaactttaaagaggtgtccgcaaggagtacaacaagatgtataag ccagaaaatcattccaaatataagcttaa	gtgacgtgctcactcagtaaaaggccgagta cctttaaggtagcgtgcgtcaacagattgag ctctcagaaatgagggtacgatagtgaga gaaatacatcacacctgcagttgtcgcctc agactgtcgtctgtgactactaattagtcg agaggaaagtctaagctctgtgttagtagt ttaaanaactcaattatcattgtcagaggaa gtcggattctattctgttgcagagataagat acgcactacagttgaaaacgaaactgtgtgat aacatataactgttttaattcaagaaggagc cgtagatactatg
1020-1021	Ga0256407_10001717 GENOME_ACESSION: RumenRJG_04DNA_2_ \$ F_3300028886 GENOME_ID: 280890 CONTIG_ID: 1716 SOURCE: JGI DATE: 2019-06-13.	atgggttatgtacaagactcgggcaaccctctgatgccgactgaaaagtacggcagc gtgaggatcatgctcaggaccggagggccagggtgtggaagtcattgtccgttaccata cagctcagcactgagaagcgacgctataccagcccgtgagccttggggcaggtgcg gttcccgcgtatcggcctgtcagcgacgagggagaagaagaactctgtgtcgtgcc ggtgaacttaggactgacatcgtggatctcctgtcaacacgacgcgaatcccgcgtacc agacgtgcaaggctccgtcacagggaagcaggtttgacaaccgcgtcagcacaaga aggaaagctggctccgccatcggtgaggagcagaatggacttccacctgaagatggt ggactgggttcgaggatccttccgtgactacggctacttttgaggtcggctctacgac atccagaagatcaagaacccgacatctcaggtgagcagtagcagcaggcgcaacagc tcgctttctggaatgtcaggagatgtgtgttgcggcgacggccacaagtgcagcact gcaaggcgcaagtcgaagatccggtttgaatgtccaccacatcgagagtcagaagacc ggcggcgatccccgaacaatttatcaccctctgtgagacctgccataaggcataccat agggggagagattgatcgaagtcgtgtgtgcaacagcttcagggtgcagcagccat caatgtgggtgaagaacgctgtgtaccgcaaggctatacagcttttgatgctgtgtgtct gcaggacatacaggtatgtcacgaagcacagcggataaatgcaggctcggagaacga cagctacactgactataggtcatcagcggttaattcggcggaatggtttctgacagtgta ttcgcgttcggcagatacagcagcagataatccagatcccaaggcaaaatctctgaaa ggctgtcgtctgaaaaagaaccagggcaccctatctgtgttcggctaccgcctgaacgat	gtcagccaccaccctgaaggatgagct tgtaacagctccagctgattagacggcactc taattgtcagccgttacggcgcaatatacag gcaccttggatgttctcgaagtcgaagct ctgcggcaggtgttaaacagctcctggag gtaggagacagtgctgtcggcacaacaacc gcccataacaacgtcagtgaggagcattac cctcctctgaggagtcattctgccaagg cagattttttaacgactaaaaagtaagtat g

SEQ ID NO:	Name	IscB sequence	ωRNAs
		attgttctctcaagagtaaccgggtatcatcaccggcgccaggagcagcggtatttgcctgaagacttggaaaccggagataggatgcagctgfcagctacaagaggttgcctgtttacaagctgtacagaactgtatgttttaacaaaaagaggaaatgtccggcgtttccfcagccccgctgaagcaggagatccacgccgaatgaatcatga	
1022-1023	Ga0223824_10002447 GENOME_ACESSION: 3300021254 GENOME_ID: 241583 CONTIG_ID: 2462 SOURCE: JGI_Dark_Harvest DATE: 2019-01-27.	atgtatgtatgtttatgttttagacaaagacggcaatccactcatgccacaaaaagatttggacatgtacgtaaagcttttgaatccggacagggcaaaagctgtatccactaaaccttttgaatcaactttgtacgagtcacataaatttactcagtcactttatggcggaaactgaccccggaacaaatattgtggaagctgtgtcaaccagaagggggaagttgtctatgctgcgcgatgtcatcacaaggataaagatatccaaagttaagactaatagagccgctcatcgagagcttctcgctgtggcgaacgactccgtgagaaaacgcagagcgaaaaccaacggtactttaacgaattttcccgacgggcgaagctacctaataaagatgggttctatgtgtgctgaaagaattattaactgagggcgaatttgcataaccgaaacgtcccgctggatggttaacaccaacagcaagacaatgtgtgcaaacgcataatataatcatcaaacagatttgcataatcttgcctgttaacaaactggacactggaaatcaacaatttgcctttcatgagaatgaatgatgctccgtaaaaggcatggatttgcagaatggtagaatgaaatgttcccaatgtaacgagatgtctgtcattacaggacggacattgcgtcctgtgtgatggaagattgacattaccacatacgtacctcgtcataaaggaggtagtaatactccgggaataattgttggactatgcagtgagtgctactcgtgaatacatcaaaacaaagttacccttgatgatattgtctttaaagaataatgctggtacgtccattgttaattgtccatgccatataattacgaagagtattataacatgtttagtgaggagcattttcacgtttgtgatggctatgtacctcaacaaaaagagcagaaaaataatgacaaaagacatttctgctgatgctgttattgtcagcaattggcagcaatgtcaacttaataatgatattggaataatctttgaataaggcaatacgaatcatgataagctatcgttaaataacgctgaacgaacctataaagtaggcagaaaaggtcgttgccaagaaccgaagccacgttttgagcaagataaagagtgctgtataagtgattgtatgaacagctttgcgacgaaatcggtatcggaagctcgtatagcattatcaaaagttagagttataaagctgtatcgagataaacgatactaaaagaattctggccgggtgcgatatctgtttcaagataaagatattgtccctaccagcgttacaacggacaattatcacagagcttggatattgacaaaaa gaattttcagctcgaattgcacaattgtacaacgaaatctttagatatgtttag	gtcaactacctctgctttatgcttcgatttgaagcaggggctttttcagcccctagcagacgtagttagcagagacatgacgtacaggaatgaattccgaaacaccgggggtgatgccaagcctccagctctgtgctagctacgccgagctatggcaatctaactcgtgacgactgtcacgaagacttatctcaaggagatttcaaatgtatga
1024-1025	Ga0307929_1001023 GENOME_ACESSION: AceLakmetage1330_2_\$F_3300031697 GENOME_ID: 279922 CONTIG_ID: 1022 SOURCE: JGI DATE: 2019-06-13.	atgttaagttaatcagtagtgcgttttaaatcaagaggtaagccgttaatgccttgcctcacaagaaagcaaggcttttgccttaaggcaaaaaagcaaaaatagttgctcatagaccttttactattcagctaaattatgccactggggaactaaacatcaattaccatagggcgttgatgctggttacaaaaatgttgggatctcaattgttaagctctaaaaagaattttatcaagtgaattcaattgtcgaaggacagattgaacgaacaaaaagcgaagatgtaccgaaataacagaagaagtaattgagtagcaggaagccaaggtttgataaccggaaaacgccccaaagggtgggtggcacaagatttggcataaatttgatagtcataattaaagtttattgaacacttaaaatcagctttccgattactgaagtgatttgaagtggtcactcttgatattcaaaagattaaagaat acagataaagtgagactgaatatcaaaatggtagcacaagatttctggaatcttccgggaatatgtttccatagagattaccatcaatgccagagctcaaaatgtcaggaaaaggaaaaacaagacaaaaatcagataatgaacacaccatattgttttggaaaaagagatagatccaataggccagggaacctaatcactcttgcactaaatgtacacactaaaaaccataaaaaaaagggtatgctgtttgttgggaacccaaggttaaatcttttagacctgaaactttcatgtcaacggttgggtgaaaaatgtcaatcaattaaaatgcaagcatacttatgtttatcaaacaaatcaaaaaagaattatcaaaagccttgaaaaaacctattactgatgcttttattgttactatgtgactcatcaagaacgaaccgacccaataatgtttaagcagaaaaagagaataaccgaagtttagaaaaggttctatgatgctaaatattgacattcgaacaggtacggtttcaaaggtgctgagttacattctgttagaacaaccagaacaagaatctcaatgtgagaatttaagaaaatagaggggaaaaaaagtcaaaaggaaaaagagtgattcgaaggcagagatacttttcaaccgcatgatttattgtgtacgaaaacaaaattggaaagtaattgtgtacacataataaggagcttctgttagaattacaattggfcaacaaacgttttagtcgatctccaaagaaactcaagcataaattacataattcattattttaaaccacaggaggtggcaattcctccctga gctaa	gtcaacaacccctgagctaaagactcaggggcttgcaagtaacttttgaagcctggttgaatagcctaagttttaaactacgttacttttagaaaaatataggcacctttagatactccactgtctgaagctcgtcggttagttttaaactctcttgagggtgaaggagaatgagctgacaaccaaaactcttgaacattggcgaagtgacaaatattcttcggatgaggacaggacttgaga gtaccgtcaattttttagaaaaagtaggtaaaatgttaagtaattca
1026-1027	contig6504_4380_Ch2-EY65S_meta GENOME_ACESSION: mgm4530144.3 GENOME_ID: 237411 CONTIG_ID: 1016 SOURCE: MG-RAST DATE: 2019-01-27.	atgggagatgtaggatagttccaacctacacgtcaccagcactgctgagggaaggaggaaaactccatagtggtttgtgtgaacagggataaaacaccttttagcacctgtcacgaggcagttgcaaggaaattgcttaaacaggaaagcagttgtacacagaattttcccttcacgataaggctcaaaagacagaagatacatcaatgtttaaaccagactacagactcaaatcgattatggaagcaggcatacagggaattgcaataatcaaaacaaactctgaagtaattttc atgatgcagctacatcatagggacagacgtaaaaagaaaatagatagagaagacgtgcattc ccatgtacagaagaacagaaaaacaagatacagaaaaacaaaggttttgaacagacgaagagatgaagattgtttaccactacacagagcagggttaacaacattgcaacgtggatgaagaagactttgtatgttctattacagtgatttcttatgagaacgtcaaatcgtat acgcagttgctaaagaaacagagatttcaggtattgtatgagatagcagcaggaacgttgcaagggtatgaggtcaaaagaactctgctgagaagtttggcagaagatgtgtttactgtgtgtctacaacgtaccacttgaagttgagcatgaattccaaaatcaagaggtggaacaacag agtagataatctgttatagctgtcatgaattgaatcagaagaagaagaataagacagcggaagatttggatattccagaatctgtcacaagaacgttgcacaaagcatttaaggactgtgcag	gtcaactaccaccgtctgtagagcggaggttgaagagccattggttgaccagcccaa ggaaactgacctgaaggacaagggaatgatgttcctacgttatccctgtcaggcactctgggtgctgccaaagctccagacactgcccgtgcagacttaaacagtcctgagtgtagggac agtggctgcacatggcgaagcagggaataacatggcgatgggagatgtaggatattccaacctacacgtcaccagcactgctggagggaaggagaaactccatgatg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		tagtcaactactagatggaagatctacaagggtttgaaggcaactggactgccagtaga atgtgtacagagctactacaaagatgaacaggctgagacttggttacctaagaccat cattttgatcgtgtatgttaggacattctacacctgacaggatttggttcaagacaggaact gttttgacgtgttagcaaaaaggtagaggcacaagacagattgctgtctgtagatgatg gtttccacgagggtatagaactaggaagaaatgttttatggctttcaaggcgagacata gtaaagctaacgttccgaaaggcaaatcaaaaggagtagtgacaggatagtagcatg tagaagaagtgatattttgatataaagaacagagctggaagaaggattgctcaaggattt catacaataattgcaaaagtagttagcgggtttgacggatattgttagtgagggaacaaca aaaatatctggcacatttctccaacctgtagagggttggagcctccatgtgccagtag	
1028-1029	metagenome genome assembly, contig: NODE_193_length_6736_2_cov_18.951298, whole genome shotgun sequence GENOME_ACESSION: OGJK01.1 GENOME_ID: 4439 CONTIG_ID: 192 SOURCE: WGS DATE: 2019-01-27.	gtgggtgagcagaagcatggatggatagctgtgcacaaggcaggcactgccgattttccc agttcggcacaatgcagagtgtccaccagcttagggaaacagtacgacaaggtgtgca caggtttacaaccgggtgccgtccgatgctggacggcggtactgcacgcagcattta gaaacaaaaaggagggtcacatgcctggatagacagactatattgtcttggcagggt atgggaaaccgcagatgccgacaaaagaaaagatatcaaaaagctgttgcgtgct ggtaaggcaaggattgcttgccttgggtgcgtttgataccggcgagcccagggtt acctatgtacataaaatcgtacagttccggagacatgacgtgcaattatcaataaccaga gggagaggacatacagactcagtggaagatcattgcaaaagaccgcagaccacgggt tgaacaaaagtgcctacactttccggctggtatgaagaacagggttaattgtcatgccc gaaggaggcagaccggatgcgtgcacagcttaccgttgagaaaagcacaagacgttac aacaatccggaccgtctgatgcctggtgccctgttcgaatacatgggtgaatacatgtatt gtccggacagttgtcaaacggccagatcttgcgtgatgtgataagaaaacaaactac ccggcaagaagaatgccggatcataagacacacaggggcgttattgttgcgtag	ataaggtaggaaaggcagggggaagaaga ggtgtgattatggcagaagaacaggggaaa aagaaaaagactccgatgtatgacggaccg atcggcggtgtccgtccgtgtatgacgtcc cgtcatgtggggagcatctgtttattccggat atctgcgtagcgtgcgggcagaaaataat gggatagttaatgtgt
1030-1031	Ga0310691_10013239 GENOME_ACESSION: 176DNAHGlopp2v2_\$F_3300031994 GENOME_ID: 279761 CONTIG_ID: 13236 SOURCE: JGI DATE: 2019-06-13.	atgactgtattcgttcaggacataaacggaaacccgtgtgatgccgacagaacgtgcgg catggttcgccgcagaataaagaaggcaagatgtctattgtctcatatgaaccgttcacg gttcgactgacatacagacaacgggaagaattatgttcaggaatgtactctcgcattgact cggggacgagtcatactgggttgcgtcactactgaaaagaatgaacttttcagtgtcga agcagaagtgaaggacaagggaagtaaaagcaaaactaaagaaaagaaaagaaaacag aagacacaggagatgtaagaacctgcgtcacagacaagcccgggttcaataaccggact cacagcaagaaggcgggtgggttccctcctactgtttggcagaacataaacacatacact cttttgataaaaagagtgagtttctatttccagtcagtaaaaataatttgaagctgctgactt tgatacacagaagatgtgcaatccagatatctccggggaagaataaccggcacgggcaga tggaaaggccatctgaacacaaggaggtttgtatttggagagacaactatactgtgcca atgtaataagaatgcttctacacgaagtgaaggttcacacacatacatatggaaaa caaacggaggttctgacagacctgataatcaggtctgtgtatgtgagaagtgtcacaaca agattcacaaaacaaaggcacaatctgccagagaacctgaatatcaagcaagactgctt tgaagttaagggtgcccggactgatgaacacataaagtgcgaagcagtcagggaatc aggaaaacatttctgatatccgggtgaaagtgcatacggatacagaacaaaatctgtca ggatgtttcataacatagagaagtctcatgcaaatgacgcatacgtcatttccgaaatac aatgctgattcattcagacagaatctatttgcacagacagaacctgtcgcataagagacaac ttcaggactttgcaccgagatcaaggagagacaaggatgggaaattcaaaatgtcaaga aaagagagaagaaaaggagatgtgataaaaagatgtaaggagataaaggagatat acggcttcacgaagcgaagcctgtcaaatacaataaagtatttactataaccgggtctc agagcgaccgggaatttctactcagaacaccaaagacaaaactgaatcaattgattca gtttcacacaagaactcaaatgataagaaaacatacaaatcattgtatatagaaaatat caagaaaagataa	gtcaccaaccggcgggtgtgaagaaccgg cggcttggctgaccggaagggtgcacacag gctgatgtgaaataccctaagtctacggaga cagccggctgatgaccggcacactccag cctacgttactcagaataatagtcactatg gatgttttccaagtcggactctgaggca gataactaaacagttctatgcgttagggaca gtgagtcgtacatacaaacctgaggataacat tggggatgggttccaccgggacaaatccgt tgcctccgttactatgagattgtttccattaa attaaaaaaaacaaaagaaagaaatga ct
1032-1033	Ga0117908_1060975 GENOME_ACESSION: IMG_3300009314_\$F_3300009314 GENOME_ID: 311702 CONTIG_ID: 60974 SOURCE: JGI DATE: 2020-04-07.	atgggtgtcgggtgttgcctgtaggttaacggagaatcttttctgtctgatatcaaacat agaaatccagtggttcgaacagaagatgttttcaagacattgtgcaagtccgagca gaaagacgcagagcagcagaaacctgtcatagatggcagaaaagcaagattcaata atagagcatcatccaaaagagctggttagattgcctccaacttaagatgaatgtcatgga agttgttagagttgtcaggcaaatctctcccaatttccatatacaccgttgaagatgttgc gtgataattaggagattaagcaatcctgatgtggaaggagtgatatacaacaatctaatc ggctggacgagaatctaagacttgcctgctgatgagagacaacttactgtcggaaat gttgtaaaaaggaaatcaactcacatctcatatagtttggacttcaaaagtgccaaa gatagattttacaacttgataaccttgcgagaattgccatgaaaaggtacactctacggg agaaagcggcaaatcaagctcaaaaggcggcaaaagtaacgagcgtcggaaacgaag tggagagggttaccgatggatgttttctga	gtcaatcaccctctgattccgagaaatcag aaggggcttgcgtagacaatgcgcaagcc ccaggttagaccagcccccttactacagag gtgaaggagaagtgtgtgtatcacacc ctgggtgtgtctccagttccgatccataag gacacggcctgttggcttattcgaacaatt ctgtgagggttaggaagatcgaataagcgt aaaaacacagagataacaggggcaggag gaatttacttttacagtaggaggttcaaca aacataatgcgt
1034-1035	Ga0117908_1013265 GENOME_ACESSION: IMG_3300009314_\$F_3300009314 GENOME_ID: 311702 CONTIG_ID: 13264 SOURCE: JGI DATE: 2020-04-07.	atgttacaagagtagcagattatcaccaagacggaaacaccactgatccgtgcaagcc agcaaaaggctcgttaagtactacgtaagtgaaggctgtcaagcgtggacaagagaag gtgtgtcttatattcaacttacttgggatagtaaaacgtactacggcgtgtgttgggtat tgaccaggcagtaaatcgtggtgactgctgtgctaactgaccaggagattgtgacctc cgcaatggcaatattccagacattacgaggaagtaaaagacaggagaataatgcgga ggctcgcagcgacgtgaagacagacggcgcaaggtacgtcggaaagatacgaaga aggctggatggattagccctacaaaagaccaaagtggaattcagactaacgctaatta ggagataacctgaagtatacttacttacttctgtgtggaagatgtacgggtcaatcact acaagaagcgtgggtgaagcacttctgtgttgaatcggtaagactatgctttatc	gctaaaattgaccttaactgcacttgacag catgaatggagatatccccccgaacttacc gctatcgactgaaaggcgactggataaa cttggaataatcgtgggaagacaaagtacg gacgataatttggagagaatattttcgaattt gtattggcaagtgaagtaataacccctga attcaatttaggggtgttgggaatgtacca caagggttaactgtgaactatgactacgtt

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gaattggagaagttggcactctataagttcgagggttgcaacaaggaactaagg gatatagatgggtgaagaagcagtagcaagacaagttagtttgactcacatcg gtatgctgctgttaatagcaggagaagtaatagggtatgtagggtattataatgtaccag agtctgggtattcaaacgccctaatactacgaagcgttcattacacttcgagaaccgca gaaagggtgaataagacgggtacatgtgtggaacatgggcattaggtattaggaagaca ctgtgtgtatctggaaggatggcgtatctgcacaggtgtttcaacgaaaggtcgtatga tctgcacgatattgcaatcaagcaagcagtagtaacaaagtgcgaaggtagaagattt gatgttactataccaccagactattacgcggaaggatgtag	agctacgtgtttgagccaagggtggagcgat atagtcgggtacaaaag
1036-1037	Ga0137372_10037833 GENOME_ACESSION: SagtaG_40_SF_3300012 350 GENOME_ID: 35427 CONTIG_ID: 37832 SOURCE: JGI DATE: 2019-01-27.	ctagacacgaacaagcagccactcaaccgggtgcacccaggacgtgcacgtttgtgctt caagcagggaagctgcccgtgctcaagtagtactaccatttactatcctcaagagcagt ggagagcagccacaggttagagcccttgcgcataagctcgaccggcagccgcacc actggcatagctctgtcaacgaccggagtgccagggtctttttgcggctgaactcacgc accgagggccagcagatcaacaagcgctgcagaccggcagcggtacgcagaagc cgagggctgcacacagcagataccgcaagccagatctcgataatcgccgagggccac agggtgtgctggcccttccctcatgagtcgagtagacacacgtgtgagctgggtacac cgactccgaaagctctctccgatcacagccatcagcatggaattgttcgcttcatctgc aaacgatgcagcagccagggtacgagggcatcgctatcaacaggggacgctgtctgg atacgaactcagagaattttgtctggaataatggggcgctgcctgcagctattgtggaa gcagaacatccccctccaggtcagcacatggtgtgcgtgccaacggcgacacgaac cgctgcagcaatttgtgtctctctgtgaacctgcgaatacacgtaagggtaccaggac	gtcaggaacccaatcccccttttaggggatgg gcttgtttcggcaagctcacctgaccagact cagtagccccgtagctacgttcagagcga aataggtacgtcgggggtacgttccagccctc gaccgctacgggtcaagggttaaaagggtgta caagggtcaagccgggtgttgcacccaa accgcttatgaacattgcaaggcacacatta ccccgggtcacacgggaggtcgtgaagagc acaccaaggagttatccagtgtgtttc
1038-1039	Ga0137380_10001296 GENOME_ACESSION: SagtaG_34_SF_3300012 206 GENOME_ID: 35421 CONTIG_ID: 1295 SOURCE: JGI DATE: 2019-01-27.	ttgagcaactgtcttcttagataccaaccaccaggcaftaaacctgtccatccggcg gagcacgcctgtctctctgggggcaagccgcgtgttctgtcgtatcccttaccat catcctgaaagggtgtctgcacatccctgtcttgccttgcgtctcaagctggacca ggcagcaaaaccaccggaatagcagtagtcaacgatgcgacgggggaagtcgtctttg ccgccgaactcacgaccgaggcgagccatcaaaagagcgtctggatacgcgacggg ctgtccggcgagccgacgggcacgaaacacgctaccgcaagccacgctttgacaa ccgaccccgcatcggaatgggttcccccttctcgcgagagtcgtcttccaacacct cacgcaggtcaagcgccctatgagcctgtgtccatcacagctacgtcaagaactgt caaatgtgacctgcaagccatggacaagccggagattgcccggagccaactaccaaaa gggacgctggccggaatcggaactcgagagtagctgtctgaaaagtggcagcgctgat gcgctatttgcggcaagacaacgtgcctctgcaaatgagcatattcatcccaagca aggggcgacgaatcgctgacgaatctcagctcgctgtgagccctgcaacatgc gaaaggcacgagggatcaaggagtttctgcgaacaacacagacgtactcaagcgc attctcgtcaatcaaaagcaccgcttaagatgcttccggggtcaatgccacacgttgg gaactctcccgcgcttgcgaagcattaggttccggtagaatgtggcagtgaggagctg accaagtcaatcgtctctcccaagcactgcagaaggcgactggaattgatgccgctgc gtgggcaagagcaccaccgtgtctctctccatggggtagtgcacttctcatcacg gtctaggtcaggaatcggaatgtgcggcacgaataagtagcgttgcctgttctg gcatcggaacgccgaagcaccactatggatatcaaacggcgatatggtgcgagcg gttgcacctcagggcgacgcgttggggagtatgtcggacagtgctcgtgcgtgccact ggctcctttgatattcaaaccaaaaggacgggtgcaaggaattagccatcgtttctcac acccttcatcgccgggagtgctatcgtttacaaaaggagaggtggtatga	gtcaggaaccccgggctaaagccacgggg ctgtgtgtaaacggcggaacctgaccagtc tcagccggggcaaccagctacgttgggagc gaatgcataggcacgttgggggtgcgcggc cagccccaacctctgcgacggatgtttaa cagagctaacgggttaactcagtgcttccg tatcgaacccgtcacgaactttgacgaggc caccattacctggaacagggaggtcatct tgagcaac
1040-1041	JGI24023J19991_100091 25 GENOME_ACESSION: Cruiolmetagenom1_2_SF _3300001750 GENOME_ID: 24900 CONTIG_ID: 9124 SOURCE: JGI DATE: 2019-01-27.	atgtccgttgatgccaaaggcagtcagcgagctaagcctgattgccccctgaagggtg ctcaacaggcccacgggagcctggccggaataccggcactaatccccgtgctccacaag gacggcacttccctcacacctgttaaaccggcgaagggccgcaagctgttgaagggtg gcgtagcagagaagcgttggaaacagctagcccaattctatattcagatgttgggtgaaac aggcaaggagagccagagatgtggctggccgacgatcccgctccaagtatgacgg attagcgatagcctcgcaacaacaggttcagtgccggtcatgtgagctgcccacag gtattgcggacaagctaccaatcgtatcagttgcggcggttaggcgttcaagctac gcccggaggcccaagcgatttgataatcgccacagccgggagggctgattgcgcttc acagaaggcgaaggtagatttgcgtgaaggctattcgtgagttgtcaggaatctatcc gatcaccggcttgtgtgaggatgtgctgttaaccattaccgcaagcactggggaaa gaacttctctaccgtcgaatttgggaaggcatacctctacgaggagttaagaagctggg ccagctcaagctctacgagggtgggagacacaagaggagcgggatgtcaggggct caagaagatcaagtcgaagccaagcgctgctggagagccatgctgtggacgcggtta gcaatgctcttaggtggctgggaacgcttacttacgggtccggagtttgggtgtttta gcaccgaacctacgagcggaagcctccatcttcagaatccagctaaaggcgggtc cggcgagtgcatggcgacagttgtttagcggttcccaagaatacaatctgcaatttga aaagcaggctctatcgacagggggttaccagggttagattgagcctgcacgatctat cactagaggcaaacgagtcaggcggaacgcaagattggagattaaagctgctattc aggcagactatattggaagagggttctcaagggttggcccttgcgctgctctcccc atggataa	gtcagggaaaactggctagccgagcttcc aatgaggatataatcgtttatagatagct aatgtcttttagacctgttctaatattgtgt gggttacactagagacgtgggttaaatcc ataataccaacattgtcaaaccttccccg ctgtgttagaaggagaataaatctaacagt agagtagtagtcaacgaccccatagctaa
1042-1043	Ga0079367_1013592 GENOME_ACESSION: IMG_3300005782_SF_3	atgcagcatattttgtgctggacgccgagagaagccgcttattgccccctgccatccggcc cgggcgcgagagctgatgagcaaaagcaaggccgcgattccgccaagtaccggttta ccatcatcctgagtcagaggagcgaggcgcatgcgaagcactgcggctgaagattga	acatgaacgctctcactcgcgcgcgcgc gaagcgcggcacagtcgcgctcatgaggc cagcctcagcagccaacgcggctgctacct

SEQ ID NO:	Name	IscB sequence	ωRNAs
	300005782 GENOME_ID: 305361 CONTIG_ID: 13591 SOURCE: JGI DATE: 2019-11-15.	ccccggcgcgaagacgacggcctggcgctgctgagagagcaccggcgctg gtgtggcgccgggaactcagacatcgagcttcgcgatcaaaagaagatggaggac cgacggcgccggcgagccgccacgcggaagctgcgacatcgcccgccc gtttgagaatcgaacggcgccggaagatggctggcgctcgctcgacgctgctg cgaggggaccatcacctgggtgcgagattgcaggagctggcgccgatcacgcacgtg agcttcgagcaggtccgcttcgacatgcagaagatggaaaacccggagatcggcgct cgagtaccaacaggcgacacttgcgggctacgagctgcgcgagtagctgctggagaaa tggggccggaaatgcgcttactgcggcgcgaggacgtcccgctgcagatagagc atcgtgccgaaaagccggcgccagcaaccgggtgagcaaccttaccgtctatgcc ccgcgcataatagagaaggcaacaggacggcgcgaggatcggcgatccggagg tcgagcgcaaacgcgaagaagccgctgcgcggcggggatgctgaacgcaacggcg tggcgcatctggcggtggcgctcaaggagggtgggattgccgctggagagcgggagcggc ggccgcacgaagtacaaccgcacggcgagcgcgatacccgaaaggctcactggattgatg cgcgctgctggcggaatcgggcgagaagatcatcatcgacagcgatcacggccact catcatcagcgcgaaggacatgggagcagacaaatgtgcgcgacggcgatcacgg cttccccagcagcataagagcaggcggaagctgcactatggcttcagacgggggac atcgtgcgcgagtgctgccgaaggggaaatagcagcaccacgtcgcgaggtatcg tggctgcgcgacggcgagcttcgacctgaagaagatggggagcggcgagaagatga cggcgaaactggaaatactgcgaagcggtgcaccggcgagatgatatgagtagggcg aggcgcaatccctgaggagctgtggaggcgcggaagcggtgcgcgctgcgcgcga cggcgccggagcgtagtccggccctccggagcgcaacggcgctccatagggcagc aatggtag	tacggcgcaatgtataggcactcgggggtg cggttccagccccggacaatcgggcgagc ggctaaacaggcacaaagggtgtagggcag tggcgctggcgagaacggcccgataagat tggcgaggcaaacgttacccttcggggg agaaaggcgggtaaccgcatcgacat
1044-1045	Ga0134857_007242 GENOME_ACESSION: IMG_3300038321_SF_3 300038321 GENOME_ID: 311752 CONTIG_ID: 7241 SOURCE: JGI DATE: 2020-04-07.	atgcctgttcgcaaaagaagcaaggttactataaaacaaataaagctaaataacta actataatccatttactatcaattattatgccaaactgggtgaacaatacaagaatgtaattta ggaatagatagtggtgctaaaaatgtgggttttgcacttccacagataaagtaataata aaaggagaaatagaattaggcaagcgtttccaattattactactaaaagaacatttag aaaaagcaggcgaaaccgcaaaacagagatataggaaagcaagatgggaaacagaaac acgaaaagaaggctgctaccccaagttatcaatcaagaattgataatcaaaataattgg attaataaatttatgtcattattgcctaatgttaatttaattgttgaaattgataacgg ctaaataataaatccagacatacaaggtgaagaatatcaacaaggaaatttatgaatat gagaataataaatcatcttatagtaagagaataaacaatgtcagttttgtgtaaga gtatgatggtaattgggtgcatattaccatataaaacaaagaaaagatgggtgaacaaat aaagtggataatttagcacttgatcacgaacaatgtcataagattatcatttaggttattaa aagttaaaatfaaaagacaaaggattataaagaacagcattatgaatatataagacaa cagataattcaaaaggttaattgtaaaataacttatgtagtacacaagggtgatagaat agtttaaaattagataaaacacattataatgatgctattgtataagtggaatfaagaaagta aagaaaaccctgaacaattattaaaaataaaacaatttagaaaaagaaaagaagttgca tgaagctataccaagaaaaaggcagaaaaagaaaaataactcaaaagagaataagta aaaatactaaatgttcaaatgggtttatcttaattgacaaagtaaaagtattagataaggtag gttgataagtgattttgtaattgaggtgttatataaaagataattgaagataattatataact ttagaaggtaaaactataaaacagtaggttaagaattagaattaatatgccataataata actggcaatattggagattgtgtgtttcatctcatgaatga	gtcaactaccacgactgaagctgtaggctt gtagaagtaattctattaggtctaaagttagt acttaagtcttaattgactactgttagataat atcacaccataagatgataccaagcttatg caactgtgttaggctctgtaaacagttctgtt ggtaggaacagtcacctaatttggagatttt acctcaaacatatactaacattgaggatgggt aaataactctgaaggagaacgaacttga ga
1046-1047	Ga0224514_10000173 GENOME_ACESSION: SF_May12_sed_USG_10 _SF_3300022217 GENOME_ID: 35059 CONTIG_ID: 172 SOURCE: JGI DATE: 2019-01-27.	gtgaaattatactgtaataggtaaggataaaaaaccccttaattgcctaccattccatcacga ccaagaagatggattaaagaaggtaaagcaactccttttgaagaaggaattttctgtgt aagattaaataaagaactgtcaaacgaaaaattacagaaagtgtgtgttattgatccag gcagtaggagagaagcatttacaattaaatccaaaagtcatacatatgcaaatattttacc gatgccgtttattgggtgaacacaatttggacaccagacgaatgatgagacgtactaga cgacagcggaacacacatgtaggaaacgaagatttaatacaaacctaaaaaagatttt ctatcacatccacaaaagctagatgggatataaaatgagagtttgaagtggataattgg aatatatacaataacagattttatagtagaatatcaaggcgaaaaccaaagggaagaa aaagtggattgtatattttcttccctttacaggttggaaaaacttggtttacaaggaaactgaa cgcttaggaaactttcattgaacgaggttatgaaccaggcctcaagggacaaattgg atttagagaaatctcagataaaaaggcagaataatttgaatccataacatagactcttgg gttttagcaaatgagatcgaacaggacatgaagcaccagataataaagaattattgagat tgggtcctttgaggaaattataggcgacaattacatatgtttcagccatcaaaagaaatata ggagacgggtagggtgaacaattatgttggattaaaactggatctataggaaacatat gaattatggattgtctttataggagggtattgttaggaatagaattagctacatgaaccaag agcacataaaaagaattttcaaatgtccaaattggagacattcatatttaacaataata gttggagaattatagatttggagagagaagatgataattataggttga	gtcaactatccaaaaattggatctgagaatcg agaccgacgaggtgtgacagattaaacag gagactagagtgaccaaggtcagaatgga cgtttgaaggaaactgaactatctgtccgtg gtagggtgaagacgttatgaggtcttccct agcttatatcctctctgatggctcagctcgaa gggaaaatacactctgtgaaaggagactt atagtgaatttatacctgaata
1048-1049	contig_283691 GENOME_ACESSION: mgm4547164.3 GENOME_ID: 237431 CONTIG_ID: 53603	atgctggtatgtattaaactcaaacggacagcccttatgccgacacgcgctgcggc aaggctctaggtctattaaatcgaagcaggcgaagggtgtcaagcgctgtccgttcacc atacagttttgtatgatacaacgaattgtgtacagccgacgaccttggaaatcgacgcagg atccgaacatatacgaacatccgcatgtacggagagaaggaactgtacgacatccgaga ttcagctgcgtaccgacatcaaaagatcccttgcaggccgacagataccgccgatcca gaagaacaggaagaccggatacaggaagccgagattctgaaccgcgtacacgcgaaa	gtcaactaccaccacttaaaccttaattggg ttttgaagtggggctgtgtacgaaggttacc ttcgtataagctccgctgactaccctcagctc ctccgagagactacgttactttgtcatcaca cctcgggacgttcgtcctaattcgcagctctg tggctcgtcaattaaacgtctgacgggaag

230

SEQ ID NO:	Name	IscB sequence	ωRNAs
	\$F_3300009670 GENOME_ID: 21291 CONTIG_ID: 3560 SOURCE: JGI DATE: 2019-01-27.	gttctaagcagacagggttcgcgtggtgaatgaaatcaaacggaaggtgtgttctat gggtcttacgcaccgcggccagcagatcaggaatgttggctgtctcgaaggaatca ggagaacagaaaggaatcgcaaacagatacagaagccccgttctgaatagat gaggaagaaaggttggctccgcgtctcgcagatcgggtcgatccgtgacaacttg gggtgagcaggctgcagaggtttgcggcgtctcagccttatccacacactgtgaaagt gacttgcagaagatggagaatcccgaaatattccggtgtcagatcagcagggaactgt gcagggtcagaggttagggaatatctactgaaaagtggggacgtaaatgcgcctact gtggagcagagaacgtacccctgcaagtagagcacatccaccgaaggaaggggtg ggagcaacagggtatccaactgacgtctcctgcgaggtgtgtaatacagaaaaagga acgcagcctattgaaatattcctgaaggcaggccctcgtaccttaaacgtatcctagctc aggccaaggcaccgttaaaagacggcgccggcgttaatgcaaccagggtgggtctgtgta cgaagggctgaaagatacagaactccagtgaagcagggaagcggcgcccttaccag ttcaacaggacaaggaagcctacgggaaaggccactggatcgtgctgtgtgtcgtg gggtgtcaggagaatcgggtccatcccggttagggatgcaacccttaacagtttagca acaggccacggcagccgtgatgaccagggtggacaagtacgggttccaggcag gtatcaaaagaaaggtgtgtcgtcgttccaaaccggagatattgttaagcggct gtgcatcagggaataacgagggcactataccggcagggttagccgttaaggcgaaga ggcttcttctgtgattgcaactcccgcggaatgaaacggctacaaaactgtagc ctactacaccgaagggacgggtacaactacggggcgatccccgtccaaccaagtcct gtga	gttattgcgggttatgacaccctgggatgcgtg ccagttccaggtctgtcggcatgagttaaa cagttctttagggtacggaacagtgttatgc catgacaagccacataactgtggcagggc aaacattaccctggaaaggggaggtcgtga agagcagaaggttcatgtcttaatgcagc at
1056-1057	Pseudothermotoga lettingae TMO chromosome, complete genome GENOME_ACESSION: GCA_000017865.1_AS M1786v1_genomic GENOME_ID: 39811 CONTIG_ID: 0 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	atggtatattgtatttcgaagtagatgaacattaatgccaacaaaaagacatggcaagt aagaagactactaaaacaggtcttgcaaaagtgttgaagaagaaccatttacaattcag ttgctgtatgacacaacgcactatagcgaacacctattacagttgggttagacattggctcaaa agtaatcgggtttcagccataacagacaacaaagaattattttagtcagaaagtagaactg cgacaagatataaagaaactgttttagagagaagagatagaaagaaacagagaatat ggaaaaacaagatataaagaaacaaagatgcaaatcatgttttcgacaataggatggca catagtaaacaggttgaacaacaatcagatgttgaataaacttttggcagatcacaaaa gcaaaaagaacagaatgggttagaagaaacacatagaacgatcatttgaatagc aggtgttagcaaaagattgaacagagctactgaatgttactttgggaatatttcagaaga caaaacacgtcttaccacaaagctaatctaatcaagggttgaattaggcaatgaatactg taaaaggagttaaagggtttaaagattcgacagagtaagatataacgatcaaataggcat gcgctggatacttga	gtcaactaccgcgcctatagaggcggg ggcttgaaaaagccctgattgactaccctca gccacataagtgctacgttagacaggtaca ccctatgatgtctcctcaagttcaggctctgt cgtacaggttcaaacagctctgagagtag ggacagtgtctgcacatacaagcctgtct aaacattgaggtagcaccctacccgaaa aggaggcttaccgcatatg
1058-1059	contig_5238951 GENOME_ACESSION: mgm4547164.3 GENOME_ID: 237431 CONTIG_ID: 422111 SOURCE: MG-RAST DATE: 2019-01-27.	atgcttacctactactatcgagacgggtcaccctgatccgacatacaacatccgta aggtgagacacatgcttaaaagacggcgtgccctgctcattgtcggtcataagccgggattca cgatccggctgacctaccgcttccacagcaggaaggaacgctgtacagccggtaga gatctgcgaggacgcgggatacagcacatcggaagtctcggttaagtcggagaagcac gagttcccatgagcagtagatctgtcttggacgaaaagcagcgtcatgacgatcag cggagataccgccggatgagaagaacggcctcgtatcgcaagccccgggttgata atcgccagaaagcagatcaatggctcgcgcgtcgtggagaacaagaacagcgcgc atgccgatactgtccggatgtatgcaaaagctcctcttaccgagcgtacgtcgaaat cgcgacattcgacacgagcactggaagccaaagcgtccggaaggaagcaccgg aaggaaaggattatcagcatggccacgctatcagatcgacgaagaggaagcgggt ctttcagcgggacggttatctcggcaggtatcggaaggaactcttcaagacggagc gatcctgaggtccaccatcgatactggaaaaagaccattccgaccgtatgagca atatgacaccgtatgcaccattgccataccggcccaaccacaagctggagggaag ctctacggctgggaaccgaagatcagaccgatggcgaggcggcctttatgaatgccgt ccgctggcagatcgtcgaagaattcgacggacttccggacctggacgtccatagca cctatggggcgtggacggaacttgcagacgacagcgggtgcaccggaaagacacacg ccaacgacgcctattgtatgggttcttccgtccgaacatcgggggaaggaaagttctgta taagaactgcgctgtgcacacgcgtccttgagaattctacgatccaagtacatcga tcttcgagacggcgcaaaaaagcggccaggagctctcctcggggaggacgaaccg gccggaaagccgtattcgaagaagatcagcggatgcaccgggtcagaaggtgtcc aaaggccgacgttcgatccggcagcattacgccatccagccgggatgcgtctcc ggtagccgaaacagaagtgcgtgacaaccggaatacacactatgggacgaggctcct gatcggcggaaagtctgtcgtcgttaagaatatagcctctgcgatatcgggggcatg gaccttaagcaaaagaaggaggtccgcgcatata	gtcaataaccacgactgaagtcggggctt gctaagagcaagctccggcagcagtcgag ccctcggggcattattgattagccccagtag cggcagcaatgcggcactacgttaccgc gaatatacagccaccggcgatgcttcaca agtcaccgctctgcggctccattaaaca tccctgatgggaagggaagtgtgcggg caaaaaaccggggataacatcggcgatgt gaacacaccgaaagcgtatcttcggatgc cgttaagtagcaggtatccatagcaggaccg gcctccagcagcgataccgtatgggag ccgtaaggcaaacatttctgaaccatttca ggaaaggaaatcaggtatgctt
1060-1061	Ga0247841_10015468 GENOME_ACESSION: LaCruzMarch2015_16_ \$F_3300029286 GENOME_ID: 280592 CONTIG_ID: 15467 SOURCE: JGI DATE: 2019-06-13.	atgcagcagttacaagcaagttaaagaacacacctcggatgttccccagctgtgact ctgtaagtccgattaaacaaggaagaattcctagtctacgggacacagtactgacttg taataacctgaagggaatctactcaacatacaggagggcagaaccacaaagctctgc tgtttgtatgtgttgggcataaattggctctccaattatgccaactagtcctatgagagccag aaaactgttgaggtcgtgtaagcaatggtagtaagcaattccatttaccatacaatcaa tagttcctataggcaataaataagcaagaatagtagttgggattgatagtggtataagaat attgttattctgttaaaacttcaaaaaagaattgtttagtggcatagttgttctgaaaacaa aacgaaggaaaggttatctgaagaagaatgtatcgagttataaaaggggcaaatgtg	gtcaattgcttctgcctaaaagcaggagttg cgcagtaatgcacaacgagtaaatatgtaaa agtatttgaagagttgattagggcaggga gaaagcaagatgcagcaggtiacaagcaaa gttaagaacacacctacggatgcttccaa gtctgtagctctgtatgtccgattaaacaa ggaagaaattctatgtacgggacacagt actgactgtataacctcgaagggaactac

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gtatagaaagccaagatttaacaatagaagaaaagtagtacttggcttccctccatg aacgcaactatgatgttcattgttaattgttgacaagattagaagttttaccaatagccaa aactattgttgaacggcaattttgacatacaaaagatttaataccagaaatfaaaggca aggaatatcaacaaggcaatattgatgtttgaaaattfaaaagcgtttglaattcttagag agaaaggcgagtgtaattctgttgtaaaagaaaaggcaatgatgtttggcgtttccaca tattaatggccgaatgactagtcttaattctgttacaatttggctttgtacattctaagtgcca cgacaagatacaaaaaagaatttagagaagtcattcaatagtaacagggaatataaga aattactttttatgaatataaattagatagattccagaagatttagattgcaaaactactacg gatataacttaccgcaaaaaaatggaattgaaacttccaaaactcacattaacgatgctt ttgtaattgctgttggaactaaccaaataggtgcttggccatgacggtaattcagaaaaga aaaaacaacagaaactctacaatgcaatcgcaatgttttggcccttccattagaaaacaaa gatatactatcagcccaagatttagtaacgataaataataaaaaatagcattgttgcca ctagaattatggcgaatatgtttatgtaagaataaaatgctattaacattgaatttttct gtaaaaagatacaaaaacattttatgaacaatagtttaatttcaagggaagcaagagttta a	ttaacatacaggaggcgagaaaccacaaag tctctgctgtt
1062-1063	Ga0070741_10034480 GENOME_ACESSION: CoaSoiStandDraft_11_\$F _3300005529 GENOME_ID: 24556 CONTIG_ID: 34479 SOURCE: JGI DATE: 2019-01-27.	atgacctgacccccctcgaaggcaagggtctgtctcaagagggaagcgaaccac ggagctgatgggaggccaagcaaggattgtgctcatctggggagaccttggccatc accgatgtcgtgggtgaagatgtgctgcccgtgacaagggaaggaaggtggcaagt ggaatggatcgttttagtcccgttcaagtcggcaaggaccatctctatgggatgctgcaaga gatgggggtgtgcttctcactgaggtgacgggtggcaaacaaaggagttaaagacctgta tggtttgaagaaaacaaaaagcaaatgaagcaattttgactctcatgccgtagacgcc tgggtgctagcagccatcacgggtgccaaagagccacctgtaggcagttgtgga tgtctgccagctatctccacaggcgcaattgacccgactgcaagccgcaaaaggcg gagagcgcaagccctatggggggacgcgttcgttgggtacaagcgtggaacctgtgt cagacacaagaaatcggcctctgtaccgttggcggcttgaccgcaagcgtgcaacgg tgagcctgcatgactatcaaacgaacaaaggctgacgcagggagcaaaagtgaggg gggtccaaaccttaacgtggatgcatccgttccctgttgggtcagagagccacatcaga agacaggcaagggaaccgccatcccaagcgcgtgcgacaggaagacctgcttcac cccacatctctcggaaaggggtcctccgaggtcgagttga	gtcaggaaacccacggcgaaagccgggg gcttgcgaagggttagcgcaacgttaagttt ctgtctagtgggtcttctctgaagcaggc agcattgtagagctacaagaacgtccagata ccaccctagtctggaccgcttctaccacgc ccaagtggatgctagatcaatcgtagct aagcccacagtgccaccacgaagggtcatt gtaccctgggttgaccaggcgaaagga tcct
1064-1065	Ga0163155_10003458 GENOME_ACESSION: OliLakLVI9SP4G1_2_\$ F_3300020180 GENOME_ID: 33728 CONTIG_ID: 3457 SOURCE: JGI DATE: 2019-01-27.	atgtttgtccagtagtagattcttaataaacattgatgccaccacgtgtcaaggcg gctagatggattaaattcaaaaaagcaactgtgttttggaaacgtggaattctattgtgaag actaaatgttgaaacctcagatagaaggttcaagaattgcagtaggaatagaccaggt tcaaaagagagagggttctactgtaaaatcagaatctcatacttttctaataatacaaacccat gcaattgtgttgggtaagagagcaatggaacaagaagatactgagaaggacaagac gttcaagaatacggccctatagaatgtagatggaatagagcaataggtgttttagttcc aagcacaaaatcagatggcaattgaaagtgaagatttgcgaatgtttatgaaattattccc aattacacatttcaattgtgaagatattaaagcagtgagtaaaagagacaagaaagtgg aatgtgtcttttagccctattgaattgtgtaagcagtggtttattctgaacttagaaaactag gaactctaaaagagacacctggattttcaacttatattgatagaatgccctaggtttaata agacaagaagaagaaactaagcagtggaattgtatgctcattgttagacagttgggttctgc aatcaaatagtgggtggacatacaaaaccagataatacaaaagttctagaataattcca ttgcaatttcataacagacaattgcatgttcagaatttttcaaaagtggtgtagaaaaagtt ttggcgttctatgagtgaaaggttttaaaaggaagtttagtaaaacacatcaagcatgg actatgttttagttgtgtgcatcaaaagggttcataagctacacaataaaacagtaaca aaagaatatgtagaatgcaaaagaacaagatattaacttctatctataacacttggagg gtgacaattctccacagataaatcagtggttttctgttcacgattgcatga	gtcaataaccacgaataaatcgtgggtt gtgaaagtagccactatcaaacacaaga ttaaactattgactagacattaataatagg tttctgaactgttgtagtcaaaagatttag gatgcttactagtcttcttccctgaaatcc aatgtcgaagtgtatgaaaaagagttta
1066-1067	Ga0315289_10016130 GENOME_ACESSION: YL17G11_40_MG_2_\$F _3300032046 GENOME_ID: 281640 CONTIG_ID: 16129 SOURCE: JGI DATE: 2019-06-13.	atggtagtttttgaattacaagcatggggaggcattgatgccttgcagtacaagaaaag caagattgcttcttaagaagaaaaagcaaaattgttaattataaacatttacaattcagt tgctttatgggttcttgggtataagcaagaaccaaattaggtatagatataggggcaaaa catgttggagttgctataactctggaacactgtgttagtaaaaggcagattgatcttaga caagatgtttcaaaattattagaacaagaaaatattagaagaaatagaagaaatagaa caactcgttatagacaagcaagatttcttaataggggtgttcttaaaaagagaactgttta cctccctcaattcaaaagtagaataaataatcagtgatgtggaataaataattttataatttatt acctaatgtcgaattatgattgaagtagctaaattgtatattcaaaaattgagaatccgga aatttctggaaaagaatatcaacagggaactttatatgaatatagaatagaattgctattta atagctagagaaaaaggaattgtcaattttgataaagaatatcaaaaaataatggtct ggagattacatcatatttctggaagaaaaagatagaccggcagattgggcattattgcat ttgcaatgtcataatgaattacatgctaaaaactagaacatatttgcacaaacaaaaatct aaatctataaagaagctcttttatgaattataagaaaagggtgttttagtatttttctgaa gcaaaattcagctatgaaatataactttcaagatagatgcaattagatttaggaaagtca catgttaatgatgccgtagcaattacagggaataaaaaataaaataaaattctgattctatat ttttataatcaatttagaaaaagaaaagggtttacatgaagcaacagcaagaaaaag aaaaaatggtaatatattatcaataagaattcaaaaaataaacttttaaatgggttaatt taaatgatcaagtaaaagttttataaaattggatttttctgattcacatctggtgcttct atataaaagataatttggaaaatatataattaccagaaaaatctataaacagatcattt	gtcaattacccctcctaaaggaggggcctt gaacggtaacgtttgagagtaattgtgact acactaaggctttagacactcgttatttaga tgatgatacctacagatgataccaaagtctgt ggctctatcgaggcgctgtaaacaaagg aaactctagtcacactcatttagtaagctt ataacttgggtgaggaacaaatccgaaa ggaggaaaactaatggta

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ttaaatataagaaaaataagttataataaactggcaattcatccccacctgaaggaagg ggattcttgcagaaataggftaa	
1068-1069	human oral metagenome genome assembly, contig: NODE_4988_length_5676_cov_2.188756, whole genome shotgun sequence GENOME_ACESSION: UPQLO1.1 GENOME_ID: 16747 CONTIG_ID: 4987 SOURCE: WGS DATE: 2019-01-27.	atgtatgtcgtctatgtattaaatcgtcaaggcaagccgctcatgccaacaaacgatttgg acatgtgcgcgtatgctaaaaactggcaaaagcaaaagccatttcgacaaagcctttcgt cattcagttgcagtagcagtgccacggattttgttcagccctttatggaggaacagatcctg gacgcaccaatattggcgaggctgtgttaacaacaaaggcgaggctgtgtacgggct catgtcataacacgcaataaggaaataccaaaactaatggcgatcgtgccacacatcgt aggcgcttcgcgcgtggggaacgtctacgccgttaaacgtcgcgcaaaagccaatggtag aatgacatgcttccagaaggacgggaagctatcaggatataaagatggggttcttgaaata aaaggcatcatcaacacggagagtcgttcaacaatcgaagccctgcgggatggct gacacctaccgcaaggcaatgtattcagacgcacttaaacatgggtgcagaacatctgcaa gattcttctgtcacagattggacactagagtacaaccgttcgcgttcacgaactagag gatggttctgttcgtggatgggtttccagaacggcaggctcaaggctacgccaataaa gaagactatctatgcccgtgcaaggcggtgttcatttggcggtgctcctattgaacat tatcaccatatacaaccacggcacaaggagggttctaatactcctgaaaatcttgggact ttgctcgtgttgcagatgacacacagggcaaacggaaattagacaaaattggcaaa tacaanaagatgccggaactcttatcgttaatatcgtattccattcatttggatggcatag tgaagatgttcgggaacaaaggctcatatttgaaggcgcaagagactccgaactcga ggcgcaacaatattcaagggaacacattacagatgcagctctgatttcagggtatt	gtcaataaccaccgcccacaccaaaggt gattaaggcgtggcgttgaanaagccagc gttattgagcagagacaaataatgacacgac aggattaaagttccgaaccacttgggtgatg ccaagcctcttgcctggcgtagtcgtgcc gagtcattggaacacacattgatcttatt gtcattaagactatctctaaaaggagatttt gtatgtatgtc
1070-1071	Ga0395764_001082 GENOME_ACESSION: AusTroGonzalo_MT_8_FD_\$F_3300037731 GENOME_ID: 311391 CONTIG_ID: 1081 SOURCE: JGI DATE: 2020-04-07.	atgaaagtatttggatcggtatgaacgacagaccgctcatgccaaccaccaaagaaag gcacggatattcttcgtgataagaagccacagtagtaaaaaagtgcccttcacaatca aactgaattacaaaacgggtctgcacacaaacgggttatatggggatagatacaggat ctcaacatatcgtgtatccgttttacgtgaaaatcgtgaaggtatatactgtttatcgaa aacagagtatcgttcgcacaaccatgaataaacggaaactgatagaatctcgtaaaac ccttcgccgtgtcgaagatttcgcaaaacggcgtatgcacatcctaataatggcatttcata cgaacgggttatgttaaaaaagcgataaatcgtaaaaacatacgcacacatggaaga aaaaatctgtgaagtttacatcatccagacaggctgtgaaggttacccatcaatccagca aaaagtgtatcacacaactactattataagacctataaagaaatcttcagacagcattac agcaaatgttaccatagaagttggacggctgatgtatcgaatgaaaaatcctgaaata catggggagatgtatcaacaaggatcgcaatatgacctagagaattgtcgcgctatgtat tcgagagagatggatataaatgtcagtggtgcaaaaagaaagccgggacaaaaagaa gatgggtctgtgttaagattattgcacatcatatcattttgtgcacaagggtgcaaccgatg atcctgacgggaatgattacaatctgcaataaatgtcatacaacaaagaatcataagccggg cggcattcttataaatggatgtagcaggtaagaagatggctcgagatctgtgatgcc acatttatgaatatacgaagaaagcgtttgttgatcttccggattgtcattttacttatgt aattttactaaggtgaatcgggaaaaactgaaactggataaacacatgcacaacgatgca accgcaattgcattgtcatatgtttttatattcttaagatgcgatatctgtatatgataatgaa gaaactgtctatatacagcagggttcgtaaaaaagaaacggctcgttcacgaacagacccca cgtaaaggcaaaaaagacaaacaaatacaaaaacagggtgagaataacaaaaatacaa agctctgtaactgttaaaatgaaaagaatgggtgaggttaaacagtaaggatgaaacaa acatatcgtatatgacataaaacataaagtattttgataaggatcatacaataatagagaa gggtgggtcacatcattttccgaatctcatgtcgatagaggatgtgaatgtgataat aaaaaaatgaaactctgaattgttcctgtgactgaagtgcattttgcatcataatagt aattggtaataaggtcccaaatggagttgccacgatataaaacgaataaagtctaa	tcggcaattgaattaccgagcatgtgtcga tatgcaaaagtgttctattgtttccttcggg acaagttgactagccttagtgaagcttactctt cggagatgtcggagactagattctgattatca atactctaactgttagtaccgaaggggt gggttcaggaacggatgttactgggagta ggtaaacaccgacctgatgcaagaagtagt caggattttggcgatgggtgattacacggca gtgaaggtggctgtatgtccacctaatatg caaaggagttactatgaaa
1072-1073	Ga0373632_0063621 GENOME_ACESSION: RO12ediment_MG_6_\$F_3300035181 GENOME_ID: 306458 CONTIG_ID: 63620 SOURCE: JGI DATE: 2019-11-15.	atggcagttttgttttagatcgtaatggcaaggcattaatgccgtgcaatgagaagcgcgc tagcttattgctaacacgcaatcgtgcgcgctgcatcctgccattctcgattcgaa tcattgaccgtcagcgtcttccgtgagtttcaagcattacgcaattagttagatccaggca gtaaaacgacagcagcgtcttagtaagagaatcagataacggcagcattgctgtcgtga atctgttcgagttatcattcgcggacgtcaaatcagtgaaagcattaacagcaagaagcaa tatgccgccgccgtccgtactgctaactcgcgtatcgcgcaccgcgttttctgaatcgc ggaaataagcaatcgtgctgttagcccaagcctgcagcatagaatcaacaccagtagt gcttgggtaaatgattccgtaaatagcgccgatcagcgtatctcgcaagagctggtga agtttgatagcagcagatggcgaccagaaatctctggcggtgaaatcagcaaggca ctttgctaggttacgaggtacgcgagttattgttagagaattcaaccgcacctgcgtttact gcgacgctaaagatacaccatgcagattgagcacatccatcctaagctaattggcggc actaacgcacatgtaatttggcgttagcttgcggccctgcaatcagcgcaaaagccgca caggatattaatgtttttggcaaaagatccgggtgcgtctaaaacgcattttagcgcaagt acaaagcctttgaaagacgcggccgggtgaatgtactcgtatgggtttgttcagcgca ttaaaagcgactggattaggtgtggaatcacatccggcggtgaagactaaatacaaccgc agtcgcctgaatatccaaagacgcatgcccgtgatggctgtgtcggcgctgtagaat cgattgaagattggaataagccgacattgaatatcaaatccatggcgccggtgtgctatca acgcacgcgcctgacggcctttggcttctcgcgggttacttaacgcggttaaaatattc aaggctttcaaacggcgcatatgtc	gtaaacacccccgaccttaaggctggagctt ttagaggaaagcgtaaaagttcaagtttacc agcctaagccgtgtcatttagcggctactgt gtacagatgtaaaagacctacgtcaggatgc ttccttagtccgtgccttagaagtagcgacg cagacaagcgacagggtgaagtagcaaacg gtctgtcacaagataagcaattatcgaaagct gctgtacaacattggcgaaggagagtgat tgaactgtcgtgagacagggttaacaccgtc acaagaccgtgaagggttgacgatctgcaa agacagataattattaaaatagagattatg gca

SEQ ID NO:	Name	IscB sequence	ωRNAs
1074-1075	Ga0137373_10002777 GENOME_ACESSION: SagtaG_42_SF_3300012 532 GENOME_ID: 35429 CONTIG_ID: 2776 SOURCE: JGI DATE: 2019-01-27.	atggtcttctgtatctcaagcgacggagcgccgctggaccctgccatgaggccagggc caggaggttctgaagaaggcccgccggtgttggcgacatcaccattcacgac cggctgaagaaccgcacggcgctgaaagcgacgcacgatcatcggtgaaatcg atccgggtagcaagaccactggcctcgccgtctccaggagaggaccggcgctagt gttcgccgtgaactcagccaccgagggccaggtcatccggacgcgctgctggcgcg cgggcccctccggcgtagctccgacggccgacaccccgatcggcctgctggttcg ataatcgccgcccccggcggttggcttgcggcgagcctgctgcaccgagctgacc acggaaacctggctcaatctgctgcgacggctctgtccaatcgccgcttgcagtgag ttggcgcggttgacactcaactgttgacaacccggagatcagtgagtgagtgaccag cagggcactctgctcgtttacgaaatccgggctacttctggaagaaggggccgctg ctgtgctactcggcggtgaccggcgctacccttgacgtcgagcatatctgcccgcgat gcggggcgcgacgacggctgtctcgaacctaccctggcctgcgggactgcaacca agcgaaagggaacaaacagccgaggttggctatccccacatccaggcgaggcg caggcaaccgctcaaggacggcgccgcatgaacagcgctggttggcgctctatggg cggctgctgacgacggcgctgcccgcgaagtcggcgacgggtggagggagcggtg aaccgagcgcgctcaacctgcccgaaggcgcttggagtgatccgagtggtggcg gtgagcagccggacaatctgctgtgttggcgaggctccgtctgctatcgccgcaa aggcgacggctctcgcgagatgtcgggacgaacggcgagcgctcccgattcgccac aaactaggcgagaacgggttctgggttcacagaggtagtgggtggcgggcggtg gcttactgggagcgggcgggcactcagctggcggtgtgtgtgctgcgctccggtt ggtttgaccttaaacgaagcaggtcgcaagcgggcgatgacatcggtattccggt atggtgcagcgggcgatggatagcctatgcaaccagaaaacgagcgcggttctctg a	gggtgccagggtagttttctgcagctctgga tgtctgggcttgcgggcaaccgcaggc ccgtcctgacctgaccaagtgtagccaa cccgctcctgctggggaggccatgacactcc gggctgctgcggccaggccggactctgt cgcccgtcatgagccgtggggtaagcg gcgaaggtgtgggtaagacaacgctctg cacatgctgcggcgacatcacccccgc gaggggagagggccccggaacggggcat aggaacgaagcaaatg
1076-1077	Ga0373625_0005570 GENOME_ACESSION: YL18G_12A_MG_FD_\$ F_3300035701 GENOME_ID: 312749 CONTIG_ID: 5569 SOURCE: JGI DATE: 2020-04-07.	atgaaagtattgttttaagtcagataaaaaacctctgatccctgcaatccagcaagagca agaaaaatgcttaatacaggaagcctctgtgttaagcaatacttttactatcattctaa aatcgtgttggtcattcatgttaaaacaagtgaatatagattaaagattgaccccggtagc aaggttctggtgattggcgatagatctgcaaaaggtgaagcggtatttgcgtgaggttaa agcacagagggtatactatcacagattgctaaagcaagacgtggtattaggcggtcaa gacgtaatcgcaaaacacgttaccgaaaagctaggttcatgaataggcgacagacggga caatttagtcccaatacagattgatgaagaaaaataataaaggcaaggcggtgcttc caccaagtttatgcagtagaattgataatgtggaacatgggtgagcagattaaagaaact ctgtctataaaaaacataatctatgagctgacaaagtgtgatacacagggtatgcagaatc cagataattcaggtgtggaatatcagcaagggtgaacttgcgggatacagatgaagagagt atttgttagagaaatagagaggaaatcgcttattgtgcccgaagatgtgccttggaa gtagaacatgctcattccaaagacaagggtggtatgcgtaggggtatcgaaccttgaatct catgtctgaagtgaacatgagaaggcaataaaagcattgaggagttttgaagagca aaccagaccaattagctataataaaatcacgttaaaagagccattaaagatactgcctct atgaacgctacaagatggcgattgttaacaaattaaagcattaggtgccgattgagat aggaaacaggagctactacgaaataatcgttcaagacttgattgccaaaggaacactg gataagcggcgctgctgttgattaaagcacaccgataatctgataattagtaatacaagc ctttatgtattaaagctgttgacatggctctcgtcaaatgtccggtttatgttgaagtgt aaaggtcagaagtcggaatctgcaaaagtaagatgcccgggtaccgaagatcgaag ctaatcattaaaaatttatgctggcggttcaactgggtgatatgtaaggctgttgcctct cgggcaaggaaaagggaacacatgtaggcaaaagtagccattcgagtagtgagatttc catgttataataagacaggtgatcatgaccatattcatgtagattttgtgaattgatacaac atgctgatggatattcatataattattaa	gacgacaaaagcttgactcaacatagccg agaaatcgggatagtaaccagactacgcc tgaaccagattcagggtactgttattatgct atgataccctgagatgcgctccagctcaa gctctatctgtatagcattaaacaggacgatg gaactaaggccagtgattatacatgacaagca tggataacattgtcgaggcaaacattaccc agaaatgggagacgccccttgcggggcatt ttgaatatgaaa
1078-1079	NODE_67_length_48612 _cov_5.28557 GENOME_ACESSION: mgm4784118.3 GENOME_ID: 238329 CONTIG_ID: 66 SOURCE: MG-RAST DATE: 2019-01-27.	atgtgcaaaagtcttataaattgacaccgaaaaaagaccgttaagcccaatttatccagcca agcgaggcaattgttaagaacaaaaaggcgagcagtttttaggcgtttcccggttactttga ttctcaaaagatcacgtcctgattccccgatctcacctttaagattaaaaattgacctggtg caaaatttaccgggaattgcactggtcaacgattcaacaggaggtcgtatttgcagccg aattaaagcatagagggttggcaattcgagattcttaacgtcaagaaggaacattcgag aagtagaagaatcgtaaaacacgttatcggcccaagattcttaacagaaactagacc gttaggatggctagcgcaagcttacaagtcggttgggaattaaagcttgggtcga gaaactacgcaaatcgcaccaattgcagcaattagccaggaattagtagcttgcgatctg caactaatgcgaatccagatcttcaaggttaaggaatacaacaaggtagacttgcgggt tatgaaactagagaatacttctcgaataatggaatagacaatgcgctatttggcgtaa aggatgttccacttcagatagaacatatttaccggagcaaaaggaggtcctaactcaat tacaactctgactttaagttgcgaaaaatgtacaaaaagaaaggaaactaaggctcttaaa gaattctcaaaaaagaccagtcgaaggttagaaaaaattgaaacaagcaaaaagacc attggcagatgcagcagcagtaattgcaactagattgacagatattgtcaattattcaca aaaaggatgggtattctatgcgtcatagctcctgcccccggttctctcaccgccaact ga	gtaaacaaccaccaccaattgctaagagc aatatgtgtggggcgcttttaagaagctctagt ttaccagctaaagtgttaactactacgttaa ggcaagtgttaagacctaccaggggatgc gaagctagtccttgcctagaaaccaacga ttaaacaggcttaagggttaaacagtgct gtgctgtgagattaccgaccttaaacattga cgaagctaactaccgaaaggaggcact ttatgtgcaaa
1080-1081	Gloeocapsa sp. PCC 73106 scaffold_00137, whole genome shotgun	atgccaaagtagctgttaatttcaaaagacaacatccctctaatgcccgaacaaacagtc gcgctagaaaatgattaaagaaggaaaagccctagggaattcaacgacctcggtattt tctatatccagttaacggaatccccatctagcaatcaactcaacctatttcttaggaattg	ttaccaaacctgacctaaaggtacaggat tgaccaaaccaattgtggtgacttaagcggt gaaatggccacagagcctaactatctacag

SEQ ID NO:	Name	IscB sequence	ωRNAs
	sequence GENOME_ACESSION: GCA_000332035.1_AS M33203v1_genomic GENOME_ID: 47767 CONTIG_ID: 136 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	accccggtaaattgtttcagggaataggtgttcaatccccctctatacccttggacagctc atctagagttacccttcaaacaggttagaaaaaggtgagagaagcgtcgtataatgcgac gagcaagacggggaagacggattaacgtgaagcaaacctttgaactgagagcgcacg ccaaaagcgggtctctaatcgaaaacaatcgaagttaccgcaagtaataagcaaacg tcaactagaataatagagtagtttctgaactatcctaaatctatccaaftactgtatatactttg aatagctgaaagccgacgtatgatttaacgtctgtgtagaaaaggagctaaatcaggga gggttttcagctgtaattgtaggacaaaactggcgatcgaacaattgtctaaactgtctc ctgttcatacacgctttgttgggagacatctaatctagaacaafttaggggttagaaaagt caacaaataaagccaagcaagtcagagagctatgcgaacgatggtatcgtcttagctt gtttccatttttacagtacttaccttttactactacctccatggacatctatggaaagga aaagttaacttaacaacagcaatcttggctgtaataaaagacctccagtcagtcgtcgtca acttcatctaatgttccagtcaaaagggggagtaaggcgaaaatattggaggaacaactac aaaattcgggttaagaaaagggtgacttagctttacccccaaaagggttggcttgcagtg gacagaccgaaaaacaatactgtcagcgatgctaaactggaaaagggttagggcgagata agctctcttaagtaaaatattccgacgttctacgggttaattgttagttgctag	acctccgaatactctctagttcggattcaatc taaaccggattggtacgggtgtctttaaaga caggacatgggttaggttggcggaagg acacaaactctctctgtgagattatctccca tgccaaga
1082-1083	Human gut metagenome DNA, contig sequence: In-M_001490, whole genome shotgun sequence GENOME_ACESSION: BABF01.1 GENOME_ID: 148 CONTIG_ID: 1488 SOURCE: WGS DATE: 2019-01-27.	atgggtgtagtactgaacaggatcgccagccccctgatgccctgtaaggaggcaaaaggc cagaaggtattaaaaaataaagaagagtagataaaagagagccgtttacatcaaa ttactgtttgactcgaaaaccggatacaggatattacgttggcatcgatccggaagta aacatctgggactgtcggcaacaacagacaaaaggaaattgtatccgcagatgtggag ctgcgaagtgtatgtggaactgttatcaaccgcaggcaatacaggagaacacgcag aggcagattaaagatactgtctccgagatttgaacccgggtcttctagaacaaaggct ggcttgcctcgcgtcgaacagaaaatgaagctatttaccgggtggtcaacgatgtatg caggattctccgggttcgaaaatcattgtggaaacagctcttctgatgtacagaaaatag agaatccggatataagcggagaaggataccagcagggaagccaattggatttttgaat gtcaggaggtatgtattatggagagataaccatgaatgccaatgctgccatgggaatca atggataaaatattaaatgtacatcatatagagaccggaaaaccggcggtgattcaccg aagaatctgattactttgtgtaaaatggccatgatcaatcatatgggaaaaataatccga acttga	ggcgtgtaataaccaccacgaaagcgt gagcttatataagcttttattgactagaccg gtgtataaagcactccgttatgtaagaatata agttaccagtggagctacagcctaactccact gctctaagaatgggtattaaacgttctgacg ggaaggaacaggtgtgctgttagaaaacctt atataacatcgtctaaggctgcatgcttcgg catacaggaggtttaaatacatg
1084-1085	Ga0307928_10009122 GENOME_ACESSION: AceLakmetagen232_2_\$ F_3300031227 GENOME_ID: 279923 CONTIG_ID: 9121 SOURCE: JGI DATE: 2019-06-13.	atggtaaaacaacaagaagttatcaagagagatacatgcatctacagagatacttctct agctctctggctctgattagtaattgaacaaagaggaaactcttagtcataactaattcaaa aaccttagataacaatctcgaagagaaccaacccaacaacagaggtgggtttaagacca aagtgtttgtaatagcaaaaggtggaaaagtttgatgccactactccaagaaaagcaag gcatcttttgaacaacagaaagcaaaagtagtaacgacaacacatttgcaatcaactt aattgggattgtgaagaaatagtgcaagagtcgaatttaggaattgatactggagtaaaaa caattgtttatcagtaacatcaaaaactaaagaattgatttctggagaatttgttgaagaac aaatattagtaaaaaaattagcgtatagagcaatgtatcgaagaaataaaagaaataaacttt ggtagagagaagcaagattctgaatagaactaaatcaaaacccaaggttgggttagcac catcagttcagcataaaattgattcgcatacagattgattaataaaaataaacttttattcca ataactaaagtaattattgaatcaatttgaatgctcaaaaggtagaataccagatattg aaggaaagtgaatatacaaaatggtcaaatgaagatttcgagaatgtaagatgtttgttaga caaaagagataaaactactgtcaaatctgtaagaaaaagatgataaaatgttagatgttca ccacataattcagaggaagatgggtgaagcgatagaccagataatctgataacattaca tcaatcatgcatagaagaattccattgggtaag	gtcaatcacctcggcctaagagcagaggctt gagaagtgattttcaagagtaattgttgaatga gagaggttgaaaatggtaaaacaacaacag aagtattcaagagagatacatgcatcttag agatacttcttagtcttggctctatgattagt aatgtaaacaaagaggaaactcttagtcata ctaattcaaaaacttagataacaatctcgaa gagaaccaacaccacaacaacaggtgggttt aagaccaaagt
1086-1087	Ga0126373_10059781 GENOME_ACESSION: PanSaiMetPlot_11_2_\$F _3300010048 GENOME_ID: 34137 CONTIG_ID: 59780 SOURCE: JGI DATE: 2019-01-27.	atgccctgccagcctcgaagggcagcccttctgctcaagcaggcaagcgcaaaagtga aaagcatgtgtccttttaccatccaactgctctatgggagtagcggctacacgcaagagg tctccttgggcatcgtatgctggcacacagcagatcggcgtctcgcgaacagagcagg ccgtctgttcgagggcgtgaggtctccccctgtacagatgtgactgaactgtcggcaacgc gccgggcaatttcgccggccagcgctcaccgcaagcagacataccgcaaaagcgcgctt cccgaatgtgcactgtacatgtgatacctaactaagctcaccgcatcgtcaccgaattgg agaagtcacacatcatggatgctcgtgtatcagtggttaacccgtgtcgcagagtgagg gagcttctacatcatgaggtgggtccgacgaataaccgcaactccacaaggcgact atcaggaaagaaagacagcgacaaaaaacactgcatcaaaatcgtccagggtctc gcctcttcgactcgtcgtcgtcaggggaagccttgccttctgtatcggcagacgcagtag tggctattttgatctt	gtcaaggaccgttccctcaaggggatgg gcttgcgtgagcaagccttgtttgactagcct tcgtctcgcgaagacccgttggatgctc aggataccctacgggtgcttactctagcctgt agctctatcgtcgtgattacaggttctgag ggcaggaacggtgttgcagcagaacaag cgcactcaacattggcgaagagtagcagac tccgaaaggaggtagacgttgtgttc
1988-1089	Ga0120401_1040793 GENOME_ACESSION: IMG_3300009431_\$F_3 300009431 GENOME_ID: 311717 CONTIG_ID: 40792 SOURCE: JGI DATE: 2020-04-07.	atgaatgggttcgatacgaagctactcaacccaagggtttatgtcagaacaaaaatggta atcctttaatgccatgcaaacccgcaaaagcaagaaactattgcgagacggttaaagcca aggtagtgtctcggctccttttactatcaactcctctgggattgtgaggactcagttcaatc tgttaccttaggtattgacaaggggagtcacataacaggctttagttgttggaaacgaac agaatttaattgtagccaaatccatcaccggacgacatcaaaaggaaatggacgctc gaatttgagcactaaagggtgaaaagatagcatctacaatttgataactcttctgagattt gtcatgacagagtacatcaaaaaggcgaaaggtggcaagttaagatcaaaaggcggttaa ggtagtaacaggtagtgatggattttctgtagaggattgtcgaaggacaaatgaaaagccgc ttgcctctccgttcgctccgagcaagggttaaaactttaattgtacaaaagggttagaaaaat agcctccttatcaactgtgtatggttatcagactcagcttaccgtaagtcactttcttggcca	tcggtagtttataaaaaccgttaaactacaaa agccgacatggagctaaacagtggtgcatcca gagactctatctggatgcagaagttagatag tcaccgcaaaagcgggacactctacgggata tcaccgacagatgttgttcaggtttctgct tgaaaagggtgacaccgctttcaagacaaa ttggtcgtctgtgggaccgttattaaagggtta ggcgaaaggtgaaggtgtgcggctctgtaaa aaaccgatctaacagccacgaggagcaaaa aaactctacattcaggaggtcataggtacg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		agacgcatgacattgatgcgatgttaattgccactctttatactggagaagtattaggtacc atcgtgac	ttatgaatggttcgatacgcaagctactcaa cccaag
1090-1091	Ga0101770_1001685 GENOME_ACESSION: IMG_3300006674_SF_3 300006674 GENOME_ID: 27981 CONTIG_ID: 1591 SOURCE: JGI DATE: 2019-01-27.	gtggcagtttctgttagacagacacaaaaaacggttaatgccatgtaccgagaagcgg gcaagaaagatgcttgaagccggtcgtgccgtgatacacagaatgtatccgttcacaata aggctaaaagatcgaaccgttgaggaaagctctttgcaaccgttgaggctgaagatcgtat cccggtgcaaaagataccggactagcgtttatagtgagaagagccaagaagaggga gctgtaatatggctatgcaaccttcataaagcaggcattaaagacaagctcgtatctcg cagagccatcagcgaagcagaagaacagaagacagaagatacagagcccccacggtt cttaaacaggcataacgagaagtgagagcttggcgagaaacgcccaacatggaaaa cactactcagggcaatgcaatgaggctaagaactttgtagataatgggcataggaattatt gggtggccgcttgcataatgctcgtgttatgcaaacgataatacagtaaacagctcag aaagttatgccaataagggaataatctatggagcttggagattcgatattcagcaaatgg agaatcctgataatcgggagttgagatcagcaaggtgattgttaggctatgaggttaaa gaatatctgcttactaacaacaaattattcctgcgttactgtgggccaagagacgatattctt gagggtgagcacgtatgctcgaataaaccaagcaggggccaagggccaacacagg cttggcaatctgtcatagcttgcaggaggtgaacatggccaaggttaattacaaccgc aggagtggtgaaatattgaagaagtcgagtagcaagaagatcagactagagctaaag aatgttagtaacatcctgaaagtgtgaaaacgaccactccccgatcctgcatgatgaatg ccacacgatggaacttttgcagttattgaagagcactggctcccatagaatgcgggtca gggtggaaggacgaagaacaaaggatagatcatggcctaccgaaggcatcatcagaca agaagaaggctatcactactacgatgctgtgtgtgggacagaacacccggacaga ctggactttaagacaagtacgctgagatgtgtccgcaattgtagaggaaacaaggcag atgtgttccctgacaaatattgggttccgagcaagcatcgagagggcacaagcaatact ttgggttccagactgtgacattgtaagagctatgaaacctaaaggcaaggatgcaggag agcatatgggtagagttacgttagatgtatctgtgtgttgcacataagaacacacaagg ggagattgttgcataagcgtaaattataagctcgaagttgattgaaggcgcatgggt acggatacggcaaaactttaaggggataa	ttcaactaaccttttaaccactaacgcgcaaa agttgtaggccaactaggccgaaaggcg aggcgaaaccagcctaaccgaaacttag ggctacgttatcctgtgtcatcaccttggaa tgccttccagttccaagctctgtcgtcaac gttaaacagatgattgtagcgaagtcagtc gttgagcgtaaaaagccaggataaattgg cgaggagagactcacgaaagtgcggtta caagtcctgtaagggaattttaaggagcgt gttaaaagtgca
1092-1093	Ga0114977_10001865 GENOME_ACESSION: S_130805_MF_Meta_2_ \$F_3300009158 GENOME_ID: 35224 CONTIG_ID: 1864 SOURCE: JGI DATE: 2019-01-27.	atgaacagagtttttttagatacaataaacaaccactagctccttgcactccagcacg ggctagacagttgttagatacaataaggtcgcgcctatcgatataatccattcacgatta tattgcatagagaagtagaattagatgtaacagataatctattatcaatttagaccagggtg cagtgactaccggattagctattatcggccatttccctaacaagggtgatgtgtatttttgg tgcagaaattaccataaatactaaagctattaaagctaaattatgcccgtagcgggttgc cgccggaagacgtagccgattacgttatgaaacaaagatttgataacagaaccagac cagaaggtggttgcaccatctctgaatcacgggttaattgattaccattttgttaataa atttaaacagttattatctaatgctaatatgtgcaaatgaattgccaaattcgatacaca aaatgaataatcccaataatcaaaaattatcaataccaacaagggtttatgaaagatttgata atacaaaagattattatcatagagatggcgaatctgtttctattgtgtgctactgatact aaattatttaagaacatgttgcgcctcgcgcaattgggtgaattcggtaataatcgtgtt tatcatgcatagcttgcattacaaaaataaataaacagttgatgaattcttatcagata aaccagaagttttagataaattgaaaaataagaattgcctcagtgggcggtcgcgccat gaactcaatgcgaatcgtttgatatcagataattactaaattaaattgaaattgggtttac tccggataccaaccagttataatagaatatgttgggttatatacaagaacattggattgat gcagctgtgtgtgacacatgttcaggatataatcctataatgatacctgtgaaaaatt aaagctatcggcaaggctcgcgcgctagtgataaacgacaatatggattccatgc atcaataaaaaaactggtgaacaacaagccgctggaaataaaaaacgcacatcattgattt caacaggagattttgtacacctatccgccaccggaaaataaccggtgaatataaaggac gaatagcagctatttctaaccagggaattttgacatttaaacataaaataccaattccaaca attgttaaggcgctacctaataatgcaaaatattattatagcaattataaaaatttccgattaat acaacatggggatagttatgaatacattaa	gatgcaataataagttttgttaggatgttttta aactatcttaattgttaaccagactaagttatt agatttatctaataactcgttatgaagaaaa gttaaaagacttacctatgggtccttccagc ccgtagctctaaagtagacagaactaaacag gggctaaatcgatggcaacagtggtgtac atgaactatttataagtttgcgagagagc agtactaggtgggagaccacattttatatag gaaaaatgaacaga
1094-1095	Ga0137384_10058259 GENOME_ACESSION: SagtaG_31_SF_3300012 357 GENOME_ID: 35418 CONTIG_ID: 58257 SOURCE: JGI DATE: 2019-01-27.	atgtcacgcctgtgtgattgtggatgcccaacggcgctcgttgatgccttcacgcccgcgc gtgcgcgcctgctttaaagcgacgggaaagctacagtgctgcgaggttccattgctctt gattttgaagggaagcgaacaggaaacagtcggctacccccctacggctcaaaccttgatcc aggctcgaagacgacgggctggctgtgtgtaacgatgcttctggcgaagtcgtctggg cagcggaatcacgcacgcggggatcaggtgcagaaaaacctgctgactgcggaag tgtgcgtcgcgcgtcgcagagacataccagatctgccaaagcgcgtggtcgaatc gtcgtcgtccaaagggtggttccaccttccctcgttcacagatacaaaaatgtgtc ctgggtgtacggctcgcgcgtttgtccattggcgcgatctggttcagctgttgcgtt tgacatgcaactgctcagaatctctcgtacgagcgatacactatcagcgcgggacgct ttggggcacggaagtcgccagatcctcctcgccaaatgggagcatcggtgtgcctattg tactgcgactgattacccctggagatcgaccacgtgtgttcacggagtcacggaggca gcgaccgtgtgtgtaactgtgtcattgctcgcggcggtgaacctggccaaggggggac aagctc	gtcagcgaccccgacgctgaacgcgcccggg gcttgacatgtgcaagccccgaagctgac cagccccctccttgcaggcgaaggaggag ccgttagagagaagaatcagagccaccggc gagtgactttccagctcgcgcgtcgtcgtat gcacattaacacgcgtcatgggttcaaac gcagtggtgcaatgtatgaacattttcata actggcgaggaacacctaccgtgacgtt ttccggcgacggcgaacgcaacgcgcc gcctcataacaggcagcgttggaggaggt atgtcacgc
1096-1097	rank12_10070934 GENOME_ACESSION: IMG_3300001484_SF_3	atgatgaacgaagcagttacatcatcggttaatacggtaaacctaaaatgcctgtttcag aaaaggcagaatcaggcatctgtttaaagacacaaagcaaaagtggtgaagaattcag tcttacctactctgcaactctgtacgaacagactgtgtctacactcaagatgtaaatcttg	gtcacaacggcctaagctaaagacttagc ggcttggctcctcaacgaatcaagctgatg gtgactagcctaagtgaaagacatagctc

SEQ ID NO:	Name	IscB sequence	ωRNAs
	300001484 GENOME_ID: 27520 CONTIG_ID: 2351 SOURCE: JGI DATE: 2019-01-27.	gttttagatgagggtccaatcacatcgtgtggtgcaacttcacacgggaagagattctgtctctgaacttgaatccgaagcagtgacatcaagaatgtaaatgagcaagaacaaacagaaagagacagacgtaacaggaaacccgttacagaaagccgagattcaataatcgggttgcatcgaagaagaaggaatggctgtctccaacagccagacacagagaagacacacgtataatctgtgaagtataattcatcttcccaatcagttgtcagactttcgaagaaggcagattgatacccaatgatagaaacccggaaatcagcggcactgactatcagcacgggaagatgtatgacatgagaactcaaaaggttcgttcgtacagagacaactatacgtgtgctatctgcccacacaggggaaagggtgatgtgcatcatatcataccggtctcaaaggagggcactaacgtgccgacacgggttttctctgcaagaatgtcacgaagattacataaggggaaagtcaaaftaccagagaatatactctgctgcaagaatctaagaatctgaaagatgcagcgaatgaacactatgtcaaaacgactcatcaaaagaatgtgtgagacatttcccaatgttatcacacgacggacgtttgttctatgactaaagcgaagaggttaaatacaatatacaaaaagtcacatgcggcagatgcgcgtgtgataagcggccagccagaagcgacaccgctcgatacacatactgcttgaacaattgagacgccataacagacagatgcagagcatcagccgagaatcagaaggcacacgacgggaacccggcatggcaagagcaagaagacggaaattgggatatgttcgaagagagtgtaaggaaatcaaatctgtttcgggttccaaaacgaagcatcgtttgtatgaaggaagaatggatgataactggctaaacagacccgggtcttttctgttaataataaagacaaaaagagaaaataaattcaatcaatatacaaaaactcaactataaaacctcaataataatcaattgtaatatagatatacagaagattataa	tcgaactacgttactggcgaatacataggta ccaagggtatgtcgtcctagtccttctgtctacgtctgggtattaaacagtcctgtccgttagggacagtgtattcagaataaaccgctgga taacattggcgaaggcgcatctacagagaca gtatgtctgtacttacagatttaactgttttt aaaaaattaataaacaatgatgaacgaag ca
1098-1099	Ga0116142_10030893 GENOME_ACESSION: AD_UK033_MetaG_2_\$F_3300009685 GENOME_ID: 21279 CONTIG_ID: 30892 SOURCE: JGI DATE: 2019-01-27.	atgaaaatggtgtttgtgtagataagaataaaagacctttaaataatgtgtcatctgcaaaag gcaagggaagtattgtctgatacaaaagctgtagtaattcttattatccctttgttatccggttta aagaagaaggttaatactctatagtagcagattatcaataaaaaatagatcccggtgcaa aatggacaggtttggcaatattgcagaatgaagttaattgttaacttttaggagtattaaac ataaggcatttgaggtatgtgataagttaaatccagaagctctttaaaggcatgggagga gaagcagaaaaacagatacaggcaaccgagattgataatcgtggaacgctagaaa attaggtagaataatgcttctttaaagatcaagggtttgataatattcaaaattggaataagaag ttacagaagttctgccctatagagagaattgtttatgaaaatgttaatttgatactcagttatt ggaaaatcctgacatagtcggaactgaatatacagagagagagattatgtgatttgagata attgaatacttacgagagaaaacgggctttaaattgtcttattgtgaaaggggggaatga aggaataaattagaagtggagcatatttcccaagaagcaggaatggaaagcaatagata agtaatttaacctgtgcttgcataagtgcatacaaaagaaaggaatatgaccgctaag aatttggtaaacctgaagttcaaggtatgctaaaaaacctttaaagatacagcaatggta aatagttctctgtaaaagaattgatgaagaa	ttcgaaaactactatcacaaatcgagagtg aaaaactctactatgtaccagccataagtc gagtactacgttagaaatgttatgacacctta cggatacttctcagttcgtagcaactgtcaa gtatttaagaatgaggtctgtattgagactaca gtataatattatgaagcatacctaactattg cgaggggagatatttcgaagaagtagctgta cctgttgataataaataacacagagattatct gaaaagataaggagtaaatgaaatg
1100-1101	Ga0180432_10045612 GENOME_ACESSION: SS_ome_17_\$F_3300017989 GENOME_ID: 35199 CONTIG_ID: 45611 SOURCE: JGI DATE: 2019-01-27.	atgcagaagttattagcaaaagtttgagaacgtaccagagaatgcttctctagttctgtctct acaagttctgatttaaacagagaggacactctcagtgatcagaacatacaactgactaata acactcccgaagagaaatctccaacgtacagggagggtcattaaaaccatttctgtacgt attatctaacaagggcaaacctttgatgcctgttctcaagccaaatctcaagaatgttaaa aggaggaagacacatgtttataaaaagaactcgttactatccaattgaatttgaatgtg caaacagaattcaaccaataactgttggaatagaccctgttattctcatattggattcagtt gcactactgatacaaaaagaactcttatcaggaacagttgaactgataacaatatgaatatca agacttgatgacagacgaatgtaccgtagaataagaagaacagactatgtgacagaaa accaagggtttaaacaataggacatctacaagaaaaaggatggtttgcaccatccattcaa agaaaacttaacactatgttaacttctttaaanaactaaacaaatttcccaataaacaag ataattgtggaagcagcaaaccttgacattcagaaaattaacgatccaacataaaaggga gaacaatatcaacaagggttcttggatgtgagaacttaaaagcatatttaattcacaga gaacaggggaagtgtcaactatgtgggaagaatacaataatgttggtgcattgttaccac atcatctcaagaatgatgttggaacagacaaccagacaattgtgcactctctcatgata aatgtcataagaactgcacaagcagaagttattttctcattgaagaaagccaacagttc aaagcagagacctttatgtccatgttgtagatggagactaacagaaggttaaaagagata atttctgatattaatgtacatttggatataatacaagataagaaggcatgagaacaatata gagaaaacacattcaaatgatgcatttattatcggggaggaagtgaacagggcagatgtt tgcctgataataatgttcaaaaggaggagaataacaggagtttcagaagaacagaaagg ggttcaagggttctgttaggaaagagagataaagattcagtcgaagatttggftaaaata gaagggaagtgggaagagacaaaggggacacattgcaagggaagagaataatggtt aatgggaacttattagttatgaagaagtagagatatataatgtaggagtttaatatg gaggcggcgaattcctccctgcctaaaggcagggttctcctgtcgaaggagatga	gtcaactacccctgcctaaaggcaggggctt gctccgtgaggagcaagagttaattgttat tagggagcttttaaggagacttaaatgcaga agttattagcaaaagtttgagaactaccaga gaattgtctctagttctctgtctacaagttct gatttaaacagagaggacactctcagtgatc agaacatacaactgactaataacactcccga agagaatcattcccaactacagggagggtc attaaaaccatttgtgta
1102-1103	Ga0311361_10020611 GENOME_ACESSION: IIIBly_16_\$F_3300029911 GENOME_ID: 280449 CONTIG_ID: 20610 SOURCE: JGI DATE: 2019-06-13.	atgcagcagttagaaaagagagatacacacacctcggatgcttcacaagtcctgtagc aactgtgattcattattaaacagagagggaactctcagttgaatggatttaaaacctcttct aacaatcccgaagtgaattacgttccatagcaggggcttagaccactaatcatgtattt aatttaagtaagtaacctttgatgccgtgtacgcttgcaggcaagaaagaaactattgaaaa ctaataaagcaatagttgtaaaatttaaaccttatacttaacttaatttgaatgtgaaaaa aaaacacaagaggttacttttaggaattgattctggaagtaaatcattggttctctgcccgtg cagatgataaggaaactaattgtgtacgttaatttttagacaataaaacaaagagtaggctt actgaaagaaaaatgtatagaaaatctgtggaataaaccttggatcgaacaaagat	gtcaatccccaaagctgaagcaagcagtg gcttaaacccgtgagggataagagcaattg gttgattaggagcaataaaagaattaaaa gcaaaatttatgcagcaggttagaaaaagag atacatacacacctacggatgcttacaagt ccgtagcaactgtgattcattattaaacagag aggaactcctcagtgtaattgatttaaaacc

SEQ ID NO:	Name	IscB sequence	ωRNAs
		<p>ttttaaataagaaaatcgaaggaatgctaccaccatcaattcatagaaaatatgatact catttaaatacctaataatcggattaagtaacttttaccataaataaggtattgaagttggt aatfttgatattcaaaagatagaaaatcctgacattaatggaattgaatatcaacaaggcaat aagtttggtatcaaaaatacaagggtcttctaattgttcagagaaaaaggattgtccaatta tgttgtaaatcatttacgaaaggaatacagccacgttcattcattgcaagcaagaaaaag attctggttctaattcgttaagaatttagttatattacataaaaaatgtcatgagagattacac aaaaaggggttaaaftaaagactccgaagcttataaagatgctctttcatgaataatcatg aaaaataggtttgtcaaaatattcctgatgtgaattaaacttgcgttatataacttttataat agagtaggcttagacttagaaaagactcattataatgacgctttgttaattggtgatggttca gttcaagacagaattaagccaatagaaattatagaaacaaataaataatagacacttc aaftaatagaaagcattttaagcctgccattagggaacaaagatactcaataaggccaaa agacttggtaattgttaaaatcattctgttatgtggcacacattgtaaaaggctgtatg ataatattagaaaataaataatcagaatacaatatcgttgtgtaaaatagaaaagg gtatcacttgggtattggaataagataa</p>	<p>tcttctaacaatcccgaagtgaatttacctgtc catagcaggggcttagagccactaat</p>
1104-1105	<p>Ga0163148_10007436 GENOME_ACESSION: OliLakL V19MP7P2_2_\$. F_3300020203 GENOME_ID: 33724 CONTIG_ID: 7435 SOURCE: JGI DATE: 2019-01-27.</p>	<p>atgtctaaccaaatttctgtctagatgccaaacaaaagccttaacccctgccgccctatt acggcaagaaaattgttagaagccaataaagccaaggtatttagacaggttccattcacaa tcattgctcaaaaagggaattactgccacaccgggaacccgattgagttaaagcttgaccag gttcaaaagtaacgggcatagcccttaagcaaggaaacaaagtcattctgggggtcggaa ttaaccaccggggcggcgcaatcaaaagcaagcttagagcttagacgttcccttcgtcgg ggacgtagaaccgtcacaccgttatctgcaagctcgattcctgaatcggactcgttcaa aaggttggtagctcctgacccatgacaccgggttcggactactgaaacttgggtcaagaa attgattaagtttgcgcctatcggctcaatcgttcaagagttgattcgttgaattgcagcag gttgagaaccgggaagtcacaggtgttgagtagcagcaagcgaactagctggctacga agtcagagaataatctacttaacaagtgggacagaaaatgcacctactgttctgttgagaac gttccgctacagattgagcatattcaacccaaaagcgaagggcgttccaatagaatctcta acctgtgcttgccttggagaatgcaatcaaaaaggggactcaggtatgcaagact tttggctaagaaagtcgaccttttgaagcagatcttggctcaagctaaacgcccttgaag gatgctgcgcgtgtgaattctaccgttgggttggtaaacaccctaaaagaaacgggtct acctgttctactgttagtggggactaaccaagttcaaccgaacccgcttaggttgccta aaactcattggctggacgagctgtgtgggacaggtttagacattggaagtattaacct ccagccctgtctaattaaaggcaacgggtcaggttctagggcagatgttaggacagacaa gtttgggttcttcccggtatgtgcttaggttaagtttgcgaaggttccagactggcgac atcgtgaaagcgggtgtcacgaaaggaaagacattggcaccattgtttaggtgggtggc agtggagacaagagggtatcctaataatctatccgctcgtgggttattcaggggcatcgt cacaaatctgcaaaacattcacaaaaggatggctacggttacggattttaa</p>	<p>gtcagataccagccctaaaggactgagctt gcagactgcccgactctagcgagcgggtat aaaccaagctaatctgaccagcctaagcctt aactggctacgtttttgagtcacgacaccgg gcgaatcgcaagtagtctccctgctcgtg ttgaaagttaaacagctttaggtcactgaga cagtgctttagccttaacagctcttaaaaca gtggcgaagctaaccttaacgtgaaaacagta gagacaaaaccatgtctaaccaa</p>
1106-1107	<p>Ga0272448_1029175 GENOME_ACESSION: YNPCB0191_2_\$.F_3300 031463 GENOME_ID: 281650 CONTIG_ID: 29174 SOURCE: JGI DATE: 2019-06-13.</p>	<p>atgaattacgttccggtcatctccagtaccggttaagccctgtatgccctgccaccggcc agggcacgggagctcgttcgacggggcggggcagtagcaggttccggaagggtttt tctacatccagcttctgacccgagcgaggggcaaacccagcacgtgatgctggatc gaccaggttccaaaagggaaggtacgatcaaggacacaaagcgacgttctga actattggcggtatagcgtttaccacgttaaggaggtgtgtagctacaggagaagatgc ggcgaggcagggaggtacaggaagccccctgtaggaagcctcgggctaacaggagat gtggcggaattccgcttccattctggctaggtgggttggaaagcttaggtgtcaggtttt tatgctccctgtaccctatcgacctgttctgtgtcgtgagggacatcaaggcaaggcttcgaa ggggcagaagaggttggaaagcttcttcccccctccagatagggaaggaatgtttta ctgggaactccgcaagctcgcgccgggtgacctctgttaagggttacgagaccgtcaa atgcgggagggagtaggttgcgcaagctcctccggaaggacgagagagtttctgtg cccactgtgtgactcctgggtcttggcagcatccgccgttgggtggatcgggtccagaag acacctggtatcgtccggtcgcgcccttgcgggttctgtaggcgatcctgcattcgcggc agtcagctaaaggtatgtacgcaggcggcatggcgggacagtcagcttggggtcctcg gcggggcacgaggtgtccaccgcgcttcgggttctgctatgtcggcggttatagg gggaccgcttagcctacacgggtgagatggcgggcggctgactcagggggccc gccaggaggtatcgtacgggtgtcgcaccttgccttggaggggttgggtgcccggagaaa agggtgaaaggaggtacattctatgagaatcaggtgacttaccggacactgtgtgag gttgaatcttccgggggattcaattcataacggctga</p>	<p>cttgaagaccagatgtctcgggtatgtgt tgtggaagtggggaatcctgagcgaggga ggaccacttgcacaatgttaacgagatcccg agactttgcaactatgttttgatgaatgtca agtagcccgactgaagtcggggctgtgtg aggagtagcggaatgccgaccggaac cgcaagacaccgtgcgaagcgttgggaaaa actagcctggccgagcaaccggaccaggt aatg</p>
1108-1109	<p>Ga0224414_10007183 GENOME_ACESSION: GoaFecdilinmedia_\$.F_33 00029705 GENOME_ID: 280326 CONTIG_ID: 7182 SOURCE: JGI DATE: 2019-06-13.</p>	<p>atgcaagctgcgtatttggatgagcaagatgtgtgaacggctcatgccgacaaagcg atctggccgagttcgccatttattagatggaaggcaagatcataaacacagacc gttcaccatacaactggtataggtgaggtacatataccaacctatggattgtgtcagg atacagggtatcagcatatggaatcactgaagtcagaagcgacgaatattgtgagt aagaacgaatcctgtgaaagacgaaacacggtcatgatgatcagcgaagatacag gcgagacaggaggaaccggcttcgttacaggcgaggtcgtcagcaactgcaaacg gcgggaaggatggctggcgccgagcctgcagacaaagcagatgcgatgtgaacga gatcatcagatagcggcagcgtccgataacggatatacatgaactgggagaatt tgatccggcactgtctgaaagcgtcggcgaagagagccgggtcccggaggatccga ttaccagcatgctgcatctacgggcaggaacgtcgtggaagcgggtcttcaagaga tgggtacaaatgctgtatagtcggaagaggagtggaagaggggcgatactgcatatgc</p>	<p>gtcaataaccccgccgattcaagaatcg gacgaggtgtgtccagtaggataagcctgat tgataagcctgagcttgaacagagactac gttatcagtgacgataagcaactcacctta ggatgctccacaagttctaagctctgaggat atgcattaaacatctctgagggtaggagaag tgtgcataattttaaaccctgcataacattggc gatgtggaccacgcaatggcgagagat ccgtacatgtcgaattatccggatggccggga tggaaaaaaggaaaggaatttccggatg caaaagctgc</p>

SEQ ID NO:	Name	IscB sequence	ωRNAs
		atcatgcgctgttctggaaggtcggcacggcagtcagatcgaatgagcttgcactgtat gcgagaaatgccatacatccgcgaaccataaagaaaggcggttctatggggatacga gccccaaacgggttgcagcttctgagaatgcgacttttatgaacaccgtccgatggaatct acaacagagataaaaaagcagttgggaagatacaatgtccatataacatccgtgcaaga catcaacagagcgaaaactgctgaatattggaagagccactgcaatgacgcatattgc atgggttaactacagaccgggagacagagcagagcagacatttcagaaaagtgcg aggaaacaccggatcctgagcaaatctatgacgcgagaatagtgatatacgcgatgg gaagataaagagcggatcggagtttaggaagttagcggacaaccgcagggatcag gagcagcgataaaaagtctgaggcaatacagaggactaaagtactaaggcgagggtg tccgaaagaacgcagcactatcagataagaccgggagacattctgctgtggaaaaaac gccatataaggcaacaggagtgcatgtaacggaaacagggtgctcctgcaaaaacaaa agagtgctcattaaaaacaataacaatacaaaaacatataaggagatggcagtttctcca cgctga	
1110-1111	Ga0182027_10069116 GENOME_ACESSION: 712metILAssembly_F_3 300014839 GENOME_ID: 21119 CONTIG_ID: 69115 SOURCE: JGI DATE: 2019-01-27.	atgaaagctttgttaattagtaagagcggtaagccactaatgccaactgtccggcaattgc aagattgttataaaacaggtgaagataagtgattagggtaacaccatttacaattaatg gttggtgagacacggaggtatatacgaacagttgaagcgggaatggatacggcgggga aaatattgggtgtgcccgtataaacaatggagaagtattatgtcagaagtattttga gggatgatactaggaagatggtacaaaggaaaatgtatcgtagccatagacgatata aaaagactcgttatcgctccatgagatataagctatctcgaaaaaacaaaggacg aatggctccaagcattaaagttaaggttgattctcatttgagagagaagagtttatggaatc aatattgcctaatgataaattgacaaatagagattgtaagtttgattgcaagaagataataa tctgacattaaagggaagataatcaaaatggctcgaagattattataataattaatg cttatgttgtaataaggataattataaattgcaaaaatgtaagggaacaaacaggacaat agattgcacgtacatcatattatagagaagattgtggaacggatggcccaagtaata gaataactctttgtaaaacctgtcatgatcgaatgcataatggagaattgatataaaggga gtaaatcaaaatctaataatgcgtcgtgaggtgtcaatgataagtttcaattgcagaagtattt tggggaacataaagaacatttggatagagactaaataaagagaacaaattttgga ttgcagaagaacattatttgcgtggtgtctattgttgaatgataatgaaaaagtaaaag taagtaattgtgtttatataagagacatgttctgctgggattatagacagacgaatggca agagaagtgaagaagaattcctacgggtaattgttggattgagaaaatgatttgaatt aaaaataaaaaagggaattgggtggtttcggcaaaaaggtaaggatttttgaattatc tgagataaattgaaattcaatgtcatgctattataaaaaataataaacgtatttct gcgagatcgaccacattgtttctatgtaattcaaacagataa	atcgtctaccattatcctaattcgttagaataat ccttcggggtgaagcagatcagaccactaat tggttcgttgattagtagccgttatcaat gtagaacccatcagatccgtccagttgga agctctttcgtggctctgttaacagttctaa gtaaggaacagtaaccacagcgtgaaggt tagataacaggtcgagggaaggtcggttc tggattatccaaaatcagatcgcataact ccaagttgtggaagataaagatgaaa
1112-1113	Ga0307430_1002475 GENOME_ACESSION: SalMarWE160130MG_2 _SF_3300031337 GENOME_ID: 280957 CONTIG_ID: 2474 SOURCE: JGI DATE: 2019-06-13.	atgaaggctttgtactcaacaagcaggggaacccgtgatgccacgacgccacgaaa ggcccgctcgtcgtgaagcagggcaaggccaagtgtccggcagggaagccgttcacc attcagcttttgacggatcagcggctacacaagcggctcaccctcgccgtcgcacgg gggctatgcaaccatcggaactctgcccgtacggagagagaagaactcatcgccggg caggtgcacctcctgcccgggtgtcacagcggatcagcggagcgaagatgtaccgca agctgagggcagggcgaagcgaatacgtgccccgggttcgacaacccggcgtgaagc cggagggctgctgctcccccgtcgggtcagcacaacgtgacggccactgaaattgac acacgcgtccgctgcatcttcagtaaacgggtgatcattgaagtcgcgaatttcgata tccagcagatcaagagaccggacattcgggggtcaggtaccagcagggagaacagg cgggggtctggaacgtgcccgggagatcctgtcacaggcagggcacgaatgccagaa ccccgagtgaccaagaagagcgttctgctcaagcgcaccacctgggtattggaag aacgactacagtgacaggccggacaacctgatcagctttgtaccgagtgacacgca agacaaccaccggccggagggtctctacggatgggaaccgagagtcagcccttc aaagctgagacgttcacgacgggtgagatggcgttgatcgaatgcagtcggagcggga ggcgacgtatgggtacctacgaagagcaagcggatcgcgctaaagctggagaagag ccaccacaacgacgcgttcgtgacgcgggcggggcggaaccaacgcggcgagg cggcggtttcgaagagatcaagaatcaccccgtcgcgtccaggtttacgatgcg atctacgtcgacacacgtacgggagagaaggcatalcgggaaggagctccactgcggc agaacgacgcggaaacaggaaacctgagcggggaaaacctgagaagatcgggggca gaagggtacgaagggcagagctctatcaggcgcaacgactatccgatccgtccaga gatatcgtgttgcgacggggagaacacgtcgtgaagggaatgtcagctatgggaa acaggttgagctgtatgtcctgatggcaagccacgtacgtcgtacgagcaaggccg tgccgctgagaagaaggagggttatcgcgtctgtgttga	gtcagcaaccgcccccgaagggttag gcttcgggagatacttcaagcccggctg atcagccctcagccaccagcttcgggtgat ggggctacgttcattgcgaatacataggcac tttgggatattgccaagtcaccaactgcg ggccaagattaacatcgtcgagggtaggc gaagtgttttggccgacaacccgtttgaac attggcgatgcaaccactcccacaaggga ggcaagctcttctgttgctgcacactatcag gagcagcagatgaag
1114-1115	human gut metagenome genome assembly, contig: scaffold32196_3, whole genome shotgun sequence GENOME_ACESSION: ULVP01.1 GENOME_ID: 15549 CONTIG_ID: 49215	atggcagtatgttttaagtgcaaacgggtgaacggctgatgccgaccgaaatfacggc aaagtgcgcatctgctgaaagatggccggcgggtgattgcgaagcgaacccattcac catccagctgacctacgataccgtatcattacccagccatgaaatgtcgttggcac cggctatgagcatatcggtcagcattaaagacgaaggcaaaagggtcgtatcgacg aatatgattgtcagcaatgagcgtcccccgcctgacagcgtgacgagcgtaccggc acgcgcaaaaaccgctggtatcgggtcgtacgggttcaataaccgtgtatccagcaa aaagcccgttggtgattgtcctccctggataacaagggtgagcgcacacctggaattat ctcccgtatctgtctgtatcctgtcaaggatgtctttatcaagccgcgacatcagata cacagctccttgcggcgtggaaggcaggggacgttcccgaggcgcaaggactatca	gtgcgtaccatcaaacccggtttatccttggc gatgtggaccacagtaaccccgaatgagc ggcaagtgcgtattaccttttgggttgcattt gaattatcatttctgtattcgttgatcccc cgccaaatgacaggtgactataccatggctg ggactacatcaataagcgtgtcgttgacgt ccccaccaatgacaggcagctataccaa gaagcatctcacattaaattgttcggttg actgtccccaccaatgacaggcagctata

SEQ ID NO:	Name	IscB sequence	ωRNAs
	SOURCE: WGS DATE: 2019-01-27.	gcatggccgcagtagggccacgatacgtcgtgggaagctgtctttgaacgggaccac tatacctgtgtattgcaagagaggtctgaaagcggcgctatcctcatgcatcacgc ctactactggaagggtctgcatgggaacagcatgaggagcttgcgacctgtcgaga agtgaacacacctgccaatcacaaggagggcgggaagctgtgggctcgataagcc tctcggaaatataccggcgaagcgtttatgaatagctgcgcgtgattctctatcagcgt gcgatggctcgttccagggtgctgcgggaagtacacatgacctatggcgtcatctcgaag cgtgttcgaccaacctcggcctcccgtattctgcgtacggatgcctactgcatgggc gagctcgtccgggaagccagatgcgaaacagagctctccagaatacagggca	cctgaccccccatgtaaacctgcgatgata atgttgacgtccccaccacaaatgacaggca actataccctgacggatgaatgccttacctgt ggcaccgttgacgtccccaccacaaatgaca ggcaactataccggctgcgaaagggaagaa agaaa
1116-1117	Ga0190313_1000813 GENOME_ACESSION: 4870071011_MG_2_\$F_3 3300022188 GENOME_ID: 20967 CONTIG_ID: 812 SOURCE: JGI DATE: 2019-01-27.	atgatcgtattgttagtagattcagcaacaatattggctatcccacacgcaagtgtgacat gatcagacgattattaaaaactgtggcaagccaaaggtttttaaaggcggactcaagtacagg tcagcctatattagtacaatatcagaagtttgataagtcacaaacaaatcgattgcgaatt cagagtcggaatagatcctggatacaacatacgggtattgtgtttacagataatagtaa caaagttataaagttgtttcagggcgaacttgaactagaacatcatgatgttaaacgggt tagatactgtcgaatgtatagaataatcgacggggcaatcatagaagaataataaac gcaaatcaaggtggctaaattaaacatcaatttggagaatcgacgcaagcacaaatt tcaaccgacacattggcattgataattcacacagcaatctattagggttgattttgataga ataccattagatcaaatgaattacatgtgaatataaattgttattgtcataagataaataa accgaatattcgaattggcaatatgcaaaaggtacacaatattgttccgagaataactaac tatattgtcgaatcgagataaattatcaatgtcagatatgtaaacagatatattggctgattacc aaacgaagtacacatatagtttggaaatcaaatggagatcagacagtcacacttaattta ataattactttgttcggaatgtcataaaaaattacacgcttgaataatgatgtctactaaatc aggatcagtcataaaatcgtgatcgtgtgtattgaattcatgtatgaagcatatgtttgct gaatgatgaaggtattataccaacgcaagatacatatggatatattacaatgtctccggaa gcaatggggattagaaaagaccatgcaaacgatgctagtgaatagcaaatatgtgattc gaatggctttatggaagagcttagacaataatcctcaatgttcggatgaagatgtaactatca attttaaacagtatagaagacatgttcgaattgggtacaacgatataagatcgaaaatatt ataattggtgttttaaaatcctgaaaaagcgtttgcgtggaatcgaatcgtcgtatcgggc aagctaagaacaagtgtgttagtagaattaaaaaataatctaaaaatgcattaa ataaagttcagattattgtaaacctggtggaaggttatgaaacaaagtgaagaagat gacattcgtcgtgtgataatcgtgttcgaaaggaactgcaactgtgcaatattatga gctgggcataaagttaataactgttcaattatgcaaaataaacaaaatattgtaccaaa aatattaacaatagcgggaatgtgtgaattag	ccgatcattaacccggacaaccttttaattta tgcggcgaaaataattaccaatacctaagtcta atgattaggccgcgacagacgtattgatt aaatgttaccctagaatgccattccaggttc caggctctggaggtgtcgtgtcttaattat tatttaataattagctgctcattgtcga caaagttagtaacgaaagtcgattacattt aatcaagtcgagggagaccaatctaatctttt aagttaatctctatcatgagatttttaaggag atacatgat
1118-1119	Ga0272429_1007931 GENOME_ACESSION: FingerMtsudMG_2_\$F_3 300031449 GENOME_ID: 280176 CONTIG_ID: 7930 SOURCE: JGI DATE: 2019-06-13.	atgtcaacgattatgattcaatcccggtgattgtgtcgtatgggacgccgtcacggcga ccgtgccgtcgaaggcgcgcaaacctgtcaaaaggaggcagggccacgggcttctgga acaagctcgaccagtagtcatcccgctgacctgtggagacacggacggcgccgctcg acgcggaactggcggtcgaccccgccgcaaggtatgacgggtacgctcgtcgtcgtc gcaacgagaacctcctgaatatcaaggttgatgtgcccacaaagagaagatcgtgcgc aagatcgccgaacgtcggcagctacgccggcgctcgttccgcacatgccgtcgc gaccacagcgcttcagcaatcgtcggcgtcctcctgggtgatcgcgccgagtcacac gtcgtgtccaggtcgcctcaaggtttgaccgtgtgtgtgtgacacctaaccgatcacca acgtcgggttgaggtatgtgcttcaaacacggcgacacaccgttacggcaagcacttc agtaccgctgaaattgcaagggcagggtgcgcaccttctggaagaacgaggggcgt tggctcgcgcgttcaggggctacgagacgaaggcgtcgcgcgagggtacggctaca agaaatccagttagcaagaagctgacgtcctcgaagccactgttccgatgccctcgt tggccctggcgggtgcggcgtgacgcctctcgcgcggggcggtcgtgtgattgat gatcgtaccgcgcgtcggcgtcgtcgttaccacacacgcagcccgcgggggtggc ctgcgcgcgccgtacagacgggggtgtgttggcttgcgcaaggattgatgacgg gaccccgccgggaaggttggaacaactcgtcggagaggtgaaaggaggttatcgctatt acgacaccgacgggaagcggaatcgaccacaaagtgttag	gtcaagaaccgcccgttgaaggcggcgcc ttgaatgagggtatcaccgaccgcgcgcg gcgaacgtcgcgggagttttgactagggg gcatcatccgatgcagacgtcccgagggc tacaagaacgtccgactgttcacgaggcc ggacccctcgttagaaccttcaatagcac gtlacgtttcggcgacacacgtccggcg ggaccccgacgcgagaagacatccgaaa ggatggtgtatgtcaacgattatgtatca
1120-1121	Ga0137383_10000411 GENOME_ACESSION: SagtaG_26_\$F_3300012 199 GENOME_ID: 35413 CONTIG_ID: 410 SOURCE: JGI DATE: 2019-01-27.	atggtgttcgttttgacagacacaagaagcgttgatgccctgtacccccagcgggcg cgactctgtcgcgcggcgccggcggtgtgttcacgtgacggcgtttgtatccgtt tgcgagaccgacgacccaggacagcgtgtctacaagagagcgtgtcctaaatgaccc tggcagcagaaccagcgccatgacctgtgtcggcgagaaagacgacccagcgggag aggtgcatcatgcgtctttgtctcgaagtgcagcatcgaggagagctcgtccatcgag gcaagcagaccacaagcaatgcccgacggcgtagacgagtgccactctcgtcacc gcgcgccacgttcgacaacgggctgttgcgaagggtgtggtcggccctcgtatgtg ctccctgtgtgtaacatctcactggcgagcgatagccgatgggtggtcggcgtc gcacctggacgggaggacgaacgggtggaacggggcgcttgggcatcgcaaaa acgatgccaacgacgcgtgtgtcgtggagagatcgcgggcgtaagctcttgaagct gaagacgtccggatcaaggcaacggccgaggcgagcattgtgcacacactggac agcagaaggctcccgcgcggtacaaagatcgctcagaagcagggtcggggctttaaag accggcgacttgggtcggaggggtgacagccgctcagcagcgaggtatccac ctgggtcgtgtgcccgtcgggcttctggtcgtccgctgggcaagtgagcgggat ccacgcgcgttatgtgtctgtgacagaggcagatgggtatgagtcgaatggaccg	gtcaacgacccagcgtgaagcgcgggg cttgagagatctcccgaggcccgctgtgtg tccagcctgagttcctgatcccgaggtagg gagacgagggaaactccgttcgtccggtcag gacaccaggaacgcctcctcaatcctgg ccctgtcgtccagcattaaaagccggtcgg ggtgtcggcggtgtgctggacaagagca agccggaggaacattggcgagaggagctg cctgggaacccgtccaggcgctcaga cagccccgtaaagggtcccgaagggaac agaaaggctcaacagcatg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		agcagcagaacaggagctgttccccacccctgccccaaaggaaaggaacgcctgttc ctccccatctttgaaaagtgggcatccgcaggcggggattggtga	
1122-1123	human gut metagenome genome assembly, contig: NODE_181_length_4155_7_cov_9.331606, whole genome shotgun sequence GENOME_ACESSION: USVZ01.1 GENOME_ID: 16979 CONTIG_ID: 180 SOURCE: WGS DATE: 2019-01-27.	ttgagagtattgttttaacaaacgaggccaacggtgatccgtgctcaccggcaaaag caaggctgcttctcaaaagagaagaagcaatcgtcaaaagacgcacgccttttacgattc aacttacgacgcgaacgggccaaggaagcagcctgtaacgcttggcgtggaattccggc tacaagcacattggtctttcggccacaacccgaaaaggccgaactctatgctcgggaagtc gaactgcgtcaagacatcacggatctttgtctgcccgtcttcttgcgccgtctcggc cagccgcaagacgcgctaccgttgcgcctgtttcgacaaccgggtggcaagcaagcgc gaaggctggctgtctcttctgttgaaaaccgattgcggcgacacgtcgcgtgtagaa gcgggtcatgaggtcttgcgggtgactgccataaccgtcgaacacggcggtttgacac gcaactgctgaagaacccggacattgccggagaggcgtaccaacaaggcgcaacagct cggattttggaacgtgcgggaatacgtgctttccgggacggacacgtttgtcagcactgc cgggggaagtccaagatccgatactcaatgttcaccacattgagagccgacggaccgg ggggcatgcgccgaacaatcctacacgtgtgcgagacctgtcacaaggctttccatcg agggggaatcgaactcaaatcaggcggaggaagtcgttcaaacggaaaccttatg ggcatatgcgttggacgctctttgagcgtttgaaaaggctcaccggaaactcagagttc gcaacacctacgggttatctcacaagcacaagcgcatctcacacggaatcgcaaaagagc cactgtgcagacgcgtactgcatgtcgggacaaccttgggtcaaaagcgaactggagggcttt ttcttcaaaagcaaacccga	gtcaactaccccgctctaagagcgaggctt gataaaagccttgggtgactagcctcagacc gccccaaaaggcggtctacgttgggtggga atgcataggcaccggggatgtcaatccta gtcccgctctgctgggtcgcgggttaaaagc tctgagaggtaggagcgggtgccgtgaaca ggtaaaccccttccaacattggcggaagaatt caaccggtcgaagaccgaggagtcgaat cttgaga
1124-1125	contig_3041208 GENOME_ACESSION: mgm4547164.3 GENOME_ID: 237431 CONTIG_ID: 335256 SOURCE: MG-RAST DATE: 2019-01-27.	atgaaaccagtatattgttaggtataacggccaagggttatgcttacgaccccaagga aggcaagaaaactagtggaaacaggaagcgtgtatgcttttaagactccgtacacaat caaatctctgtataaaacaggagtaatactacgcccacattcgaaggaacagatacggg aagccagcacatagcgttgcgtaaacacggatgagaagcttatacaaaaggaagagt atgcacttcgaagcacaatggacaagcgtgctcttatcgaacaagagccttctacagaaa gggaagacgttatcgcaaaactagttaccgtcatccgaattcagaccacatacaaaac gagctctatgtgaaaaaccagttcgttataaagcacttaactcactggaagaaatgtaag aatgtattcgcatacaatgaaaagaaggatggcttccgccatccatccagtcgaagtgtg atatgacgatgcaaatcatagacagataaagatgatacttccgcaaaagacaaggtttcg tcataagtcggaagattcgcatacgcaagaaacacagccgtccattcaggggagat gtatcaatcaggacgtctttacgacgaggagaacgtcagagcctatgtgtttgagagaga cggttacacttgcgaagtgttgcgaagaagggccggttctaaccgaaaagatggttctaca gttaaaacttgcggcacatcatatcgaacttgcgtccaaaggagccacagataatctgaca ggatgataactgtgtgacggatgccatacaacgtaatacatcaaaaggaggaaattct ttatcagtgatgcttgatgagaagaagtcgccaagaggatcgcagacgagacgtttat gaacatcattcgcaggaaattattcaaaagagtttcccaatgacgaattcacttaccgggaact tcacggctgcgcagacagaagacgcttgcgttagagaagacacacgcaaacgacgctgt gccattgcatgtcaggagcgtcgttgaagataatctgatacaacctattatcaacaagt aaggcataataaacgttcaacttcatgaggctacggcggttaaaggacgcaaaagaccca accgtgacagattcgaatgcgaagaacacaaaatctgtcacaattggcgctataata gcaagaaaacattttatgtatcagacaaggtttcttacaagagacaggtgggatgataag cgggttactaaaacggcgcttattgtaaggacaagatgatactatataagtatcct gataagagttacaaccagataaatttaagcgatttctgtgtatcacaacgacaattg gctgattggagcagtgcttccaataggcaagtag	gtcaactacccaccactaagcctaagggctg aagtggggggcttataaaagccaacattta ggcaggcagtgaggagaattgtactagcctc aggagttagatgtatttagtcagactaactac tacgttacagagatgaatacggccaacatgc tctcaagtggataaaggcgttgcggcaagc gatgtctgatatagggtaagtaggacaccga cccaggagcacgaagtctctggacattggc gatgagcaatcgtgctgtggacacgggca gaaccgtcagtgctctcggcaaatcaccatt cttagaggagggagcacaagccttgatga aacca
1126-1127	Ga0210056_1025515 GENOME_ACESSION: CryGey1480SPAdes_\$F_3300025022 GENOME_ID: 280077 CONTIG_ID: 25514 SOURCE: JGI DATE: 2019-06-13.	atgcagaagtatttagtagttaaagacgcactcaggtgttccacaagctgtgagct ctgtaagctctgtattaacagtcggtgtgaggtaatggacaggtgtacaggacatacaact ggctaataacaactccgatgtggaccaacttgggaacaggaggaaacacacttttagtagt gttcacactgaacaaaaatggtaaaccctaatgcttgaacactgcaaaaggcaaggca tttactaaaagaaggcaaaagcaaaagtagtaagaggtcagcctttactattcaattactata tggttcgggaactgcaattcaacctatttcttggatagatcctggatgacaagatag gattttctgctatcagataagaagagtttaatttcaggagaagtcagattaagaagaagat gtatcagcaaaatcaacagaagagagaatgtatagacgaaacagaaggaataaattatgg tataggaaagccagatttctaataagatctcaaaaaaagaaattggttagtccatctatt aggatataaattggatagtcattagcctgataataaaataaaggaaattgcttatatca gatactatagttgaaatagcctctttgatgccgaaaatagtaaatccagagattagtagt attaaatatcaacaaggagaattacagggatataaatacagagaattattatgaaaaagt gggtagaaaaatgtcttactgttgtaaaaaagataatccattagaagtagaacataataac caaatcaagaggtggaactgatagggtatctaactaacttcttctgtcaaaaatgtaac tcaagaaaggaaacagaacagcaaaagaatttgatatacgaataacagaagcaag gaaagaatcactaaaagcaactgcatttgaataatgttaggtgaggttagttaatatatt gaattgtgaagtggaacttggatataaaacaaagcacgatagaattaaattagggttagag aaatcacacagcaatgatgctttacaatagcaggtggtaataagcagttagggtcgggtc aatatattggaataggtatagaaggaatgcttcccttacaacttaataagaagaagattta caccttccattaggaggcaagataattacagccacatagtttagtaaaatggaacaat	gtcaaccacctctcttaaggaaaggaaactt gtgaagcaagtcctcctgtttagtaggggt attgaaaaatgcagaagttatttagtagtta aagaacgcacctacggatgttccacaagctc gtagctctgtaagtcctgtattaaacagtcgc tgtgaggtaatggacaggtgtacaggacatac aactggcctaatacaactccgatgtgacca acttgggaacaggaggaaacacatttttagta

SEQ ID NO:	Name	IscB sequence	ωRNAs
		caattattaagaattagaggaaatcttctaaggggcaagagtagttttagagaataagaa atcagtcagatagatagtggttaattatataatataatgagagatggcaattattccctgat ataaatgacggaggttcccttgctacaaaataa	
1128-1129	Geitlerinema sp. FC II Abyss71_314_len:29491, whole genome shotgun sequence GENOME_ACESSION: GCA_002286845.1_AS M228684v1_genomic GENOME_ID: 129498 CONTIG_ID: 158 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgtctaatittgttttttactcgcatacgaagcgacagcccccgcaccgtgccatcccgc acgggctagggaactgttagccaagggaaaaagcgcaatttacaggcgctatccgttca ccatcggttttaaatcgcgctgtggcgacgtgcccacatgtaaccaactcaagattgaccc cggctcccaaacaccggactcgcctcgttaaatcgacgcagctcgtttgggtgctggg aattgcaacaccgagcgctcgcattaaagcgaaattgcaacccgctcgtctgtcgtc ggaatcgccggaaccgcaagacgcgatatcgcaagcctcggtttcaaccgtatcccg ccgaagggttggtgccaccgtctctaaaccatcgctcgaaccacgatgacgtgggt gaaccgttgcgaaaactttgtccgattgtgtcgtctcgaagaactgtcgggttcgata cccaaaaactccagaaccccggaagtgtggtgattgagaccagcagggtgaattgttg gctatgaagtgcgcgaatacgtgttgcaaaagtgggtcgaaagtgcgtctatgtccggc ttgagaacgtgcccgttagaggctgaacacatccagcccaatcgaaagggtggaagcgt cgcgtctccaatcttaccttggcttgcgtccgtgcaaccagtctaagggaaccgcgac gttcgagagtttctgttggaacccctccgttcgagcgcattctgagacaggcaaaaa ccccgttgaagatcgccggcggtgcaacgcaactcggttggaactgtacacgcgattg aaagagacgggtttgccgtcgaagtcggaacggcggttttaaccaagtcaaccgcac ccgattgggttgcggaagctcactggctcgatgccgttgtgtggagagacgccat cgtcgcatttagcaactgaaactccgatagcgtcttgagcaaaagtcacagcacgagat ttcgaccctcctacgaccgttacgattccccgcgccgtgagacaaacgaaggcgcaa gtcaacggcttacaggcagcgacatcgctcgagcgatcgccgaacgggaaatc gcggacagtggagggagcgatcgccgagtgcgaaaaaacgccacctgcttgc gtccttttgaaaaaacagatcgatttaaccgtcaaacacaaatccagatcgttcacaa acaagatgctgtaatacgaataaactgtcggcggttccctccgacgctga	gtcaaccaccgacgccgacgagcacgg gtacggcgccggcgtgaaaaagccccaca gttgaccaggctaaagtcctttacagactccgt tatcgaggctacgacaccgataagttagcca gcttgttgcctgtcgtcgtcgttaaacagc cctactcgggttaaggcagtgccgaccgacc tcacaagcctcgataaccttgcggaggcaa acgttactgccttaagcgacaggagtaga cagaatgtctaattt
1130-1131	Ga0307954_1000334 GENOME_ACESSION: OrgLakmetagen437_2_\$ F_3300031391 GENOME_ID: 280790 CONTIG_ID: 333 SOURCE: JGI DATE: 2019-06-13.	atggctgtttttgtactcgcagacagaagaaccactgatccgtgctcggaaaaacga gcccggctttgtcactcgcagggcgccgggtgtgtgtgctgagcgatccgttactattc gcttgaagaccgctgctggcggtgtgtcagaagtcgtctcaagatcgacccggc agcaaaagaacccgcttagcggtgtccagagtcagcgcccaaggaggagcatgtactgt gcttatagagctgactaccgtggccatcagatcagcaaaagtgtgaccagcgccgg ggatttcggagtcggcgccgtggctgctcgtctacccgtgcaacgaattcgcaaccc gacaaagccaaagggttggctggcgccgagcctcgagatcgagtgatgataccaccg agcatcgtaaccgtctttgtaccttgggtgccgtgtcgtcaatcagccaagagctggtgc gggtttgacctgcaacagatggagaatcccgaatcagcgccgttgatgatacgaagc actttgttgggtacgaagttcgggagatttttgaaaaaatggggccttgaatgcgcttat tgtccgtgacggacacatccactggaaattgaacacattgttccgaatcgaacggcgcc tctaaccgaatcagcaatcgtaccctcgctgtcagactgcaatcaggaaaaagggctcg caactctgcccgaatcttccaaacgtccagtcgcttgaagacaaacagccgcgcat ggacaacgtgctgattcagtcgaagcggcccttgcgggacgctgcggcggtgaac	gtcaactaccccgccgtgaacgacggggct tgtagagaacacattcaagccaggttgacc agggaagcgggacaccaaccgctacgttt gccacaggctgttaagaccactccgggtg gctactcagcaccggacatagaagctg ggatcacgctggcgaaagggttaaacggcg aaggtttcagttgcccgcgaagcgggagcc gggtgtcgacattccgaggggagagagg tttggcctccgttacaggcctgttaaggc actatttgaaaaagatatggct
1132-1133	Ga0376687_0042770 GENOME_ACESSION: AusPinMetagenome_11_ FD_\$F_3300036817 GENOME_ID: 312296 CONTIG_ID: 42769 SOURCE: JGI DATE: 2020-04-07.	ctgctcaggaccgtcagacagccggtaacggggccctcaccgagggcacagaag gagtcaccatgtgtttgtcgtcgaacacaaaaagccgctcatgctgtaccccc aggcgcgacgcctcctgttcacgaagcgggcggtggtccatcggtcagccctt cacgatccggcgatcggacgcagcgctccagagagcagcgtgcaaccggtgtg ctcaagatcgatcccggttcgaagaccacggccttggccgtgcgcgggtcgaggaga cggatgagggggaggtgcaccacgcgctgcatcttgcgagctcaccatcggggag aagagatccgagcgctgtcgaacggcggtgtatccgcccgagggcgttcggc caacctgcgtaccggcagcccgcttcccaacgcgtcgggtgcggcggtggttc ctcgtcgtcgtcagcggatcgacaacgtcgtgtcctggacttccgctaccggcgct gggtcccgtgtgtcgcgtggaggtcgagcgctcaagttcgacaccagaggctcg agaatccggagatctccgggtcgagtcacgcgggggaactggccggttgggaagt gcgatctacctgctggagaagtgcgacaccggtgtgtcctactgcggccgacggaca cgcccttgaactgacatatacagtcacgcgtcgggtgttccgaccgggtgtcca acctggcgctgagttgccacgcgtgcaacgtcgcaaggagatcacaccggcgcg agttcgccatcccaggttggcggtcaggcggaagcagccctcggggatcgggcg ccgtcaacgccacgcggttgcgtctgcgacgaactgcgaagcttgcgtcggctc accagcctgtgttggcgacgcagcgctggaaccggcgcgcttcgcatcccaag accacgcactgagtcggatgtcgcgtggagcactggcgaggagtcagcgctccc ggcacaacactctcgcgatcaagcgacgggtcgtggcgctacagcgcaccaatgt ggatgagcatgcttcccgttgggtatttgcgtcccaagcaggtgagggcatca agacggcgacgggtgtccagttgctgacgggtacgactacgactgtgtgacac agggtctgagcacgtcgagtgccgctccccgacctgtgtccacca	gtcaacgatccccacgcatgaatcgggg gctgttccccgggtgtatccgtctgtgtg tcagccgaggtgcccagccttgtggcaa ggcgctgatgggtaccagcgttgcgtcgt cacgacacccgctgtgttcttccagcactc ggcgctgtcgttcatcgttaaaagacctgtc gggtacgggtcggtgcgtgagcccaac aagcgagcgcacactggcggaaggagc gtcctgagcccgctgctcaggaccgtcag acagcccgtaacggggccctcatccgag ggcacagaaggagtcaccatg
1134-1135	Ga0172375_10015289 GENOME_ACESSION: LakKiv1m_2_\$F_330001	ttggcatcaagacgttccctcgccgcggccggcggaacccggaagaccgctatcggc aggcaaggttttctaaccggaagcgcaatcaggcgctgctccgctgattggagagc cgggttctgaacgtcaattcctgggttaaccgcttggagcgattagcggcgtaagctcg	gtcaaccacccacgggtgaagccggggg cttgagagatcagggctgaggttaccagc cacagccaggggcgcaaggccatcggtc

SEQ ID NO:	Name	IscB sequence	ωRNAs
	3137 GENOME_ID: 31118 CONTIG_ID: 15288 SOURCE: JGI DATE: 2019-01-27.	atctccctggaactggtaaagtcgacacgcagaaactccaaacccctgaagttccggc gtagagtaccagcaaggcgagcttttaggtatgaagtcgagagtacctctggagaaa tggggccggaaatgcgctattgcaagactgcgaacgtaccgttcagattgaacacatt gttccgaaatcggggcgggactaacgggtatcaaatcaacgtctggcctgtgagtc tgcaatcaggcaaaagtaattctaccgcccggagttcccatctggcatccaatct caagccaggatgccgctcaaggatgcgctcggtcaatgcacagagatgggcactgt ttaatcaataaaagctcgggctctctgctgagatgggcagcggcggacggacaagt tcaatcagatgcggcaggatcccgaaggctcactgattgatgcgctgtgttggg gaatctgggtctaatgataatctctcttgggcaattcctgtgcaaatcaaggcggcgg tcacgggtctggcaacgatgtggaacggataaatatggcttcccggttcgcatgcaccg aaggccaaatcttcatggggtatcaaacgggacattgtgcaggcgaacatcccgtaa	acgttagtcaggctcatgacacccccgggatg cgtgccagttccaggctctgtcggccagcat taaacaggcataggggttgagccgggtgt gttggcccaaaaagcttgattacattggc gaggcaaacattaccctggaaacaggagg ctcattgtgagcaag
1136-1137	Ga0180732_1000856 GENOME_ACESSION: KR11_01MetaG_2_\$F_3 300020171 GENOME_ID: 30715 CONTIG_ID: 855 SOURCE: JGI DATE: 2019-01-27.	atgaaagtttatgtaattcaaaatcaggcaaaccaatgaatgccgacagaagatttggaaa ggtaagaagattataaagtcggggaaagcaaaagtggttaccgcaaaagcctttacgatt cagcttctttacgagactactgaaatgtgcaaccattgattcttgggggtgatacgggagc gaatgacattggagttgtattcaaaaagaagatggtgaacctgtatttttaggtgaattaga aacgaggacaatagaagttgctcaaatatgaaggatagatgcgaacatagaaggggcaa gacgcagacatcgacagaaaaagcgaagagaagagcaaaagcagcaggatcatttt tggaaaagaaaaataaccacattacggctgtgaaagaagcaataaactgtcaaacatcaaa gcccgggaatgtacgatttgaataatagaagcggcggataaatggcttactccaacttg tactacttattacagacgcatatataatttataaagaagatagcaaaaatactaccgattgct atagtcattttgagatgccaaatttgaattgcatagaattataatcccgatgtaaaaggta aagattatcaaaatgggagaaaagaagatatgtcaatccgctgaatactgtctttgccg agataaacattcttgcagttatgcaaggtaaagagtgaggagatgcgcttaccattgtt atttggcaaaagtgagaatggaagcgatactcctgagaatttagtgacgttgtgtgaaaagt gccatgaaaaagttcacaacaatcagaagcagacagaaaaataaaaggattatttgaag gaatcaagaaaaagtgatgctcatgctacttcttaattcggttttacctaagttgtttcaatg gttaaaaagcacttttgaataatgtcaataagacctatggatagcaactaaagagaacgt tgggaataataatttgcacaaatctatgtagttgagcttatttaatagctatttggatgataccc cccccatgatttgacgagttgtgaatcgttctgttcaaacagttcagacgacataatagg cgaacataaaaaagacaagaagaccgaaaatattacatcggaagaagaagtcgctgtt aatcgtataaaaaggacaggaacaaacctttgatagcttaaggatttagtcacaaaatgtg cagaacatgagactgttctgaatctgtgactgtaaaagccagcagactagaccgaagagaa gcaccaagccctttgggatggcgatgtcgttaattcaggggagggattcatatagtaaa gggatttacaggaactacttgggatttgtgatgctgctgatggcaataacacaaaaaac ataaaagaagcgggaattggttaacaaaatcaagggaattgtttgtatttga	tcatcatatagctgtgacattagcattgaag ctgttgatgtgctgttgagaaaagtttaattt ctgttaaaaaatattaaaaagcattcagctttt ataaaaaatttagtcaacaaagggttagacca caggctaaaaaaggagttgcaggtgaagag caccttgtgattttccagtcgaagtagctgtt aatctgtactaagtcagggaaaccaaagc accgaaagtcgctttacaggagttttcatg aaa
1138-1139	human oral metagenome genome assembly, contig: NODE_518_length_2044_7_cov_4.380541, whole genome shotgun sequence GENOME_ACESSION: UPWB01.1 GENOME_ID: 16887 CONTIG_ID: 517 SOURCE: WGS DATE: 2019-01-27.	atgtctatgagtttgacatatgttttagataagaatggacaacctttaaagcgaactcaaaat gtggcaaaagctatagactactaaatctggaaaagcaaaaggtgttcaaaagagaaccgtt tactatcaaatattatataaacctgaaactcatatagttcaagatttaactttaggagtagata ctgggtcaagtaaaattggtactgcagtagtagataatgatgcaaatgtttattatgcttcta agttacaatcagggaatgataatttctaagaatgatgagcagacgaatttcatgtagagcta ggcgaactagaaaacttagatagacctgtgaagatttagtaactgtaaaaactcaactaa gaaagatagattttaccgactatgggtgagcaaatcaattcgcatattagagaaattgagt ttgtaagcttatttacgttgaagacactagttattgaaactggtacattgtatccacatcta ctggagcacatagaagatggaattgcttttaacaaactgggttatcaaaaaggatcta actatgttttgaattctagagaagcttctaataagagacaactacactgtcaatgttg tggagctaaaaatactagactggaagtcatcatattattatcgttccaaaggtggtagtga tgaatttagttaatcaattacgctatgcgagaagtgatcaaaacttctcatgatggaaagtt aaagaatttgagtcaaaacttagtggtaaaagaaaaggtatattaaacatgctactcaaat gaatagttatcgaattcaactattgagacactatccagaagctattgagacatttgggttcat gactaaagaaaacagacaatcgtccgatttagaagaatcgatgtaaacgatgctattatc atctctacagggtgtattacaagcctaagtataagacagaaggtgactacaagaagaat gtataccagaggtgactatgctgttactgtacgtcgtgacaaggaaaaaagaataaggt aggcaaaactactaagccaagaataacagaccagtgatggttgaagagcatgataa agttgaatactgcaatactatagcttttgaagtcattgaggtttgtggttaattgtccgctta tggacattgacggtaaatatttgaagtttagagagcgatattgaaaagctgatacagctc agtaaaagacctcaagagaatatacagtagaaaaaattgctgtgtacaaaagtcactttct tgtgcacatcatga	atcaaatgctctgtatcatcagatggtgag cactcaggtctaaaggttaacagactaagct agtagaaatactagactacgatgattgagtg agatacatcacacctgtagtgtccgcctca gactgctgctctgtggctaagtattagtaga gcaaaaagtacaaaaagagctcgggtgactt agcttaaaaaccttctcatattgtcgagag gaagtcagattcttaactgttaacagagtt agatagcacaaaggcttattacaatacctg cttacagcataaaattgtgaagtcattaga aaggagtcagtcattatgatttgg
1140-1141	Ga0307930_1006096 GENOME_ACESSION: OrgLakmetageno46_2_\$F_3300031600 GENOME_ID: 280812 CONTIG_ID: 6095 SOURCE: JGI DATE: 2019-06-13.	atgatttacattgtagacgaaggaataagatagccatccgaccaagaagcatgatag taggtcggttaataccgaaacgaaagagcgaagattataaaagattgaacaaggtatt atgatagttcaattatacaaaaggttttttcaaaagcgaacattgattgtgaatttaggat aggtattgatctggatagcaaatattggaattgtgcatatttaaaatccatggaacattatc acaaaactattataaggagagggcatgtttacgaacaaaaaagatcaaaagaggactgta gccaaagaggtgatcgccgtatgcgaaggtatctgtcgaagaataattaaaaagaat atggaacaggtcgagctaagtcaaacatcctgtctggaaaatcgaagtaaacacaat tcagcctacgcatcttcatgtaattcaatcacatgagaatgtgtggcaaaattataaaatta	aacctaaaggatttgggcttcagttaatcgc aagctgatttaattcaggttgacagacgtt taataaaaggttcactacggatgccactccc agtcctgtagcaatggaggtgaagcggagcat tgcaagtcctattgtgatttaggtgaacg aagatcagttacttttgattaaagtcggggg agacgaaccgaactattgaacattcttagtt

SEQ ID NO:	Name	IscB sequence	ωRNAs
		attccaatcgaacaagttaaaatccatttgggaataataaactttgatattcataaaatgattaa tcctaagatacaatcttttggatcaaaaagggttacagtttggttfgaaaacgtgaaaagt tatgtgaggaaatcgagataaattcaaatgtcagatttgcataaaggatgttggtagtatcag aaacgagggtccatcatataaattggaagagtagtgggtcagatcgtcctgaaaattta atactgtgtgtagaagtcgccacaacaagggtccataaaaagggttaaaattgtccagaca aatctatttcagtaaacggataccgtaattgctggagttttaaactcagttatgggaacaatat ggagtaggtttgaaaatcaatgtccagttcaagacacctatggttatattacaagtggtgttc ggaatcgagacaacattattaaatctcatgccagtgatccagcgttatgcatttaatgata gcttgggaagtcataaacattgaagattattcttgggaagattataattcaattaaatgaa tcaattcccgacatactgatactcattcaattaaagacataccgaccgtagatatacaatttt aaattctaatttacctgggaaaaagaaaactgtagtggcttggaaatcgaaacgaagggc gggcccaagatcctaagaagcaaacgttgcggagttgaaacaatcttattgattgagga acgcttaactcaacagttatagttgcaattccagggtcaaaaggtaagagaagttgacac tgattacacgatccgaaaaggtaggtgtcaatgtgggtcgaatgattaaagtttgaag gtgtgcaaaacaaaggtaggtgttacttttgaataatgattaaacccaaagaattagat acatttggaaactaaaaatgtcataaaaataaacaattgtggattgtgactttattaa	tgcctcgatcaatgggtttttattaagaggttaa aatgatttac
1142-1143	Ga0182741_1016627 GENOME_ACESSION: eDN5th3BELigOPv2_\$F_3300017649 GENOME_ID: 37689 CONTIG_ID: 16626 SOURCE: JGI DATE: 2019-01-27.	gtgcttctcagccccggacccttgaaacgcgcgacagacacggctggggtagcca cgaacggggcgcggtagacggatccgcccacaaactgggtgtcaacatggtcgag gggagaccacccgcgagggtggtggcgcacgccgatctgttccggcaggtcgggca gcagggtgctgggtgaccgcacatcacagctcaggtgttcgtctgcacggcaagggtc ggccgtgattgcccactcatccgcacgcgcagtgctgcgcaagggtcgggc gcgagtcacccgcctgacccgttcacatccgactgtgtggatgtggacggcaccgac cccgggggtgtgtcgtgaggtggagctcggcatcgtatcctggaagcaagaccacg gcatggcctgttctgacacgcacatccggtaaccggacggcaggtgtccctgatgaac ttgtccaccgaggactcgcgatcaagatgtccctgtccaaggggcggcgtgcggcg gggccgcccgttcccgaacctgctgacgcgcacccagggttcgacaaccgcacccg caagcccgcgcgacggctggacgtgtgtgctccaccgtccgtgcggcaccgggtgtg caccactgtggcctggctgacggactggcagggctggctccgatcaccgggtgcat gtggagtcggcgcttcgacacccacctgctccagagcccagggtcagcggtcgtc gctaccagcagggcacccctggcgggactgaggtcaggtgatctgctagcgaagta ccagcaccgggtgtgtctactgtgacccaccgggtgtgtgctgaacctgacacgtc gcccgcgtcgcgtggcggcagcaaccgggtgtccaaactgggtaccgctgctgccc ctgtaacgaggccaaggacaacctgcgggtggagcagttcctggccatgaccagcc cgactggcggcggtgtgctgggttgaagaagccgtgtgacgcccgcagccatga actccactggcagccctgggtggtggtgacatccctcagcggtgtcgtacccgtcactg ccaccgggtgtgtgcaccaagtggaaaccgacccgggttcggcgtcccaagaccatgc gctggacgcccgtgtgctgggtgaggtcggggaggtgtccggatggccgagcaccac cctggcgtgacagcaaccggcgcggtcctctaccccgcacccgcagcgacccggc acggcttctcgtgctgactgactgggtcaagcgtcaccacgggttcgacccggga gacctgtgcccggcagtcgtgtccacgggttaagaaggcgggcacccacttcggcagg gtcgccgtccgcccaccggcagtttcaacatcacaccctcgcgtggcaccgtccagg catccaccacgtcagctccggctgtgcaacgagcagacggctacacctacgcaacc atgaaggaggagtcgggacgaggggcagcgccatccctccccacggctaaagccg ggggtatccccggcgcacccgatga	gtcagacacccacggctgatccggggg cttgaagccctcagcatcagcagctgacca gaccgagacacctgaaggaggtgactac gttgcacacgagcgaacagaccactccg ggggtgttccctagccccgacccctgaaa ccgcgccagcagacacggctggggtagcc acgaaacggggcgggtgagacggcatcc gcccacaaactgggtgcaacatggtcgag gggagaccacccgcgagggtggcgac gcccgatctgtccggcaggtcggcgacg aggtgcgg
1144-1145	Ga0164242_10000581 GENOME_ACESSION: OricomgcompostMG_2_\$F_3300012942 GENOME_ID: 33846 CONTIG_ID: 580 SOURCE: JGI DATE: 2019-01-27.	atgagtggtttctgtttggataagaagggttagggcactgatccgtgagtgaaaagcgc gcacggctcctcctgacccggaggaaagcgtttgtcaaatgatgcagccgtttactattc agttgaaagaacggctgttggaggactcggagctacaatccgtcaggttaagtcgatc ccggttccagacataccggatggcgtgtgtagagacgcggaaggtatcaagtattgc ttgaattgtatcaactgatactggtcgtgagatgattcaccgggaagctattgagggcagc aatgtatcgaacaacgcccgttcacgcaagacgcggtatcgacaagcagcgtttctcaa tcgccggaacccaaagggtggttgcgccagcctgatcagcgggtgaatagcaca ctgagctgggcgttgaattccaacgctgggtaccgctcacgaagctcgtgttgagcgc aatcgttggatattcagaagctccagcggccggagatcaagggtatcgaatccaacg ggttacacttttggcatggaggtctgggaatacttactgaaaagtggggcatcgtcgt tgtgtactgcgaggtcctgaccgttaattgacgattgatcatgttaccctcgtcagcg gaggcagcgtcgtgagtaacctgtacggcctgcgagtagtgaaccagtttaagg gcaataagcccgttcaagaattctgaaaaggcatccggatcgttgaagcgcatcttgc aagggttaagcaatcgctaaaagacggcgtgtgttaattcaacaagggtacaagctca ttgaggttttgaacaactgaagcttcaattgaaccgacactggagccatgactaagt gaatcggcgtggttaaatgtccaaaacccactcattgatgcgtttcgttggagat gtgcacccgttctgattgattggcaagccactcaggtcattgcatctacgtcgcg gtatgatcaacgcgtcatcttagataggttgggttccctaaggctaatcaacgcgcatc aaacgtccgtacggattcgaactggagacatagcccaagtttttcagaggcacatgtg aaacgtcagttcccttcagatctccaagatgcatagattgacagttaaaatagatgattc	tcaatgaagcgtcttctactgacaaactacac acctagggttccgaacctagcttaccagga aaccagactcagcgtgaccggtcagcgt acgataccgtgaaattaggtaaccctgct gccctctcagcgttgggtctacgggtgtat gttaaacaggcctacaggtttaaagcaggtgc atacatcgttaaacacgggatactttgtcg agaggacctttacaggtaccgtaaggattca tggagaagaacctatgatt

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ttccaactggcgcggcgtaaaaagatagtgaaagctatcctaccgctatcttaaatgaagc agcgcaataacggctatttcattacgctgcaacgtttctaa	
1146-1147	Ga0209697_10006283 GENOME_ACESSION: AliMusmetaSPAdes_2_\$ F_3300025316 GENOME_ID: 22533 CONTIG_ID: 6282 SOURCE: JGI DATE: 2019-01-27.	atgaaaatatcaagagataagcaacagggtgcagacaaaccagagcagaaaaccgctgt gggtattcgtccgcaatattggtgggcaaaagcaatcaaccccaccaagcctggcaaacgtta tcgtcttgaagagacggtaaaagcagtcgggtttgtgcaaaagccatacacgatccagat gcttgacactgcggaggatctgttcagccgtacagactcggactggacagtggttacag taatgttggggcttcgggtatcaacgaacggactggagttgaagtcctttcaatggaattg aacttcaaaaaggcgcaaaaagatcgtaacaccgatcgcaaaactcgcctcacggcag aagaacaagaagtgccgacatcgggcagctcgggttgataaccggagaaggcgcaaaa ggcaaacggctccgagcattcagcataagctggacactatgtcagaatagccaccgc ccttatcaatgaaaagtctgctcctacagccggcgctgtgtagaagggtcagttcga attcagaagatcaaaaatcctaatatcgaaggtgtgactatcagcaaggtgaccaagctg gcttttggaaatcctcgtaatatgttttcacagagatgagcacaatcgcaaaatcctgatt gtaaacacaaaaaggagaagaaccttcctcaagtgcatcacttggcttctggaag gtgatagtactgaccgtccggtaactgatcacgttgtgttcttccatagggcctcag aacatgcgaagaaggaattttatagggatgggaacaaaacttaaaaacttccgtctg aaaccttcatgagcaccatctactag	cccgcagtaaacccgcggcttgagggtggg gcttgttacgagcagtagcgtatttcgtagca agctcgggttgattagccctagtagcgtgtcgt gatgacattgctactacgtttctgagaattctc ttccagcagcatcggaagagtgtaataggc accttggatcctccacagctcccaggcact ggagaacaacgaatctgttgcattccgggt ggtgtcaaatgacagctacttgcgtttagt gattaaacatcgttttgggaatggcgaagt gttctcaaaaaacctcagaaaaacattgg cgaagtggaccaatgatgccttcgggcac ccacagagggttacgactctcgttataaaa tcaaaagaaaggccgtctatatgaaatat caagagataagcaa
1148-1149	Ga0137371_10000193 GENOME_ACESSION: SagtaG_43_\$F_3300012 356 GENOME_ID: 35430 CONTIG_ID: 192 SOURCE: JGI DATE: 2019-01-27.	atgccatgtcaaccacgcaaacagcgcctttactgaaagggggcaaggccaaagtgg gaagatgtgtagaaggaaacaatcgacaactacacaaagcgacatcagaaagggtgg caagaggcagcgtaacactgcacccaagtattgtcgtgttttcgcttctgttattgtgtga agcatcaaggacagacgtcgtttgtgttggaggaggagttctgttacttcgatctacgt accttctgatgtgccaagtccatgccagtgccagttacaaaagtgtgctgtgttccaga aggcttcggcctgtgtatagaaggaggcgccggttctcctctgtctcgttaa	gtcaagaaccaatccccatgatcggggac gggcttgcagaagctcgttgcactagcctg agtcaaacactcgttgggtgagtcagatgac ctatgatgcttactatgctgtgactctatc gtttacgggttaaaagtctgaggggaggaa ggctgtgtgaagtgtaaaaagctcattcaaca ttggcgaagagtacgtactccgaaaggag gtaaacgctttgttc
1150-1151	human gut metagenome genome assembly, contig: NODE_4561_length_458 8_cov_3.294066, whole genome shotgun sequence GENOME_ACESSION: OIZI01.1 GENOME_ID: 6169 CONTIG_ID: 4560 SOURCE: WGS DATE: 2019-01-27.	atgtatgtagtctatgtattagatgaagaaggcaaacgccttatgccacaaaacgcttcg gcaaaagtgcagcggatgtctcaggggacaaacttgcaaaagtgtatcagtaaaagcctttcgt aatccagcttctatataagccaaaaacataatcacccaaccgttacagggcgacacaga tcccggctgtaaaaacatcggcatgtctgtatcaatgacaaaggagaaatccttactcc tctaccactgaatcaaggaaaccaagaatccggaacttatggcagaagaagaaagctca cgtcaagcatcacgagagcggaacgtctccgcagaaaacgcccggcgcaaaaat aaaacaacaacaacttccgggaaggcgcaaaaactctgggttatgaaaacggagtcct gctctgaaagatatcatcaatactcagcgtctgttttaataacaggaaagcgtcccagaaattg gataacgctacagtaagacaatgcattcaacacataatcagcctgtcagggcagatagc agattctgccggtaacagactggagatagaacatacaaaatcgttttatgaaaatgga agacggtactgtaaaaggtagacactaccgaatggcagactaaaaactataaaaatgt gaatgactatcttggcatctgcagaacggaaaatcgcaatctgcgacagcaaaataga acactatcatcatcgtacaaagaacaaaagggtggcagtaaccgccccgacaatataat cggactctgtgaatcctgccacgtaaagtacatagtgagaaactagtcacaagaaat cggcgcaaaagaaaaatacggccatcttccgtgtgaatcaagcaataaccgttcatctgc tccgagctggaacagctcttggtagaataattatatactgcagcgggttatgagacttat acatatagagaaatgtataagctggacaaaaccatgatatagatgcagcctgcataagct gctataccgcacaacatagaaacaccgatacaaaaagctgaagcctataaaatcaagca gtaccgttaaccataacaggcagataatcattgccgaagaagcgtacctacaactcg gcaaggaaaaaatagccaagaacaggaaaagcgtacagacaaaaagaattgtcttc aacgaatgtataagatacaaaaagaactgcagcaaaacagagcttcgagaatcat gaaaaagctaacagttataaaaagcatccgcgtttacaacaacatgaaaagattaaagcc cggcagtggttttatatagcaaaagcctgaaacaaaaaagctgaaaagccagagcaaaa gcataagcaagcaatcaatcacagcgggaaatatacttttaagaggtgctattaccaatgg aaaatactacaagccgaagactataataaaaacaatttctcgcaaaaaactgcagattt caatactcaagctctgctctacatctaa	gtcaattacaccacctaagaggtgggtgc ctgcaagcttcttgcaggcaaaaagggttaa ttgagcagagacatgatgcagcgtcactac ggggtaatgccaaagccccgtaccctgggtta cagctgcacctagtcitagcactcttactatt ccctcttaggggaatgaattatctatttttaag gagaaccaaaaatgtatgta
1152-1153	Ga0079226_10027342 GENOME_ACESSION: PoultitStandDraft_2_\$F_3 300006879 GENOME_ID: 34654 CONTIG_ID: 27341 SOURCE: JGI DATE: 2019-01-27.	ttgcatacaggcgaacagaccgacttgggggtgcttccctcagctccaagctctcgaacce gcgtcagcagacacggctgggtgtagcttcgaaacggggcgccggagacggcgggcg tgccccgctcgacaaccgggtgtgcaacatggctgaggggagaccgggcgcaaaagccgg cgtggcaccttctgggtgcggccgtgagggcgaacccgaccacgggggttgcgtgtgt gttcgtgtgggaacagaccacaagccgttcagccgaccaccgagcggcggggcacg gaagttgttgaaggcgggccggcggtgtgcaccgcccgttaccggtttgtgaltcagg gtcaaggaccgcagcgttggggcggtcctgtgtgcagcgtgtacaggtcgggacgacc ccggctcccgccacaccggcatagccgtatcaccagagaaggccactagtaaaaggagt ggtgtgcgaccgggtgtgtgtggtggcgagttgatcaccggggccagcgaatcagccg aaacctgtatcgcggcgccgctgcggcgggggccgacgcagccgcaacctgcgata ccgcaagccgaggttctgaacctatccggcccggtgtgactgtgcggagccaac gctcaaatggcgaagcgactgtccgaccgtgcgaacaccttccccgtccgagcgtg aacgtgtggcgccggcgacggtgccccgtcgtgcggcaccgggtggacactc tcgctcctggcggaacaggttgacgggtggcgccctgtcaccggctggcaltcagga	gtcacttaccctcgtgtaaggcggggggt tgtagccccgttcggcaggactgagtgacc agactgagaacaccttgaaggagggaat ctacgttcatacaggcgcaacagaccgacc ttgggggtgcttccctcagctccaagctcga ccccgctcagcagacagcgtgggtgtagc ttcgaacggggcgccgggagacggcggg cgtccccgctcgacaaccgggtgtgcaat ggtcgaaggagaccggcgcaaaagccg gcgtggcaccttgggtgcggccgtgta

SEQ ID NO:	Name	IscB sequence	ωRNAs
		attgtccggttcgatctgcacgatgcagcgcggccggcatcaccagtgctgagtacca gcaggggcacccctggccggctcagggtccgcgagtaacttctgtccaaagtggaaccaca agtgcgcctactcggcgccgtccggagtcggcccggtcgggtccgctgaacatcga ccacattcaccggggcagcaaggcggtcccaaccggatcctcaacctggccctggcg tgcgtggcctgaaccaggccaagagcaacatccgggtggaggacttctgtgtggca agccggccgtgctcgtcgggtactgctgcagcgcaaggcgccgtggccgacgca gcggcggtgaaccaccgggtggcggtgtttacatgttgccggacaccgggcttcc ggtttaccggccctccggcgggcgaccaagtaaccgacgggtcacgggcaccccc gaaagcgacgcgctggacgcgttatgcgtgcggcttctgaccgggtgaagtgtatc cgtccacgacctgtgtatcgctgcggcgacgcggcacctatccccgaccgggtc ggacaaacacggcttcaaccgtgcacgtgacccgaccaaaggcgactacgggttc cagacgggagacctgtgacggcggtacccaccggcgcaaggccggatccat atcggcacgggttccgttcgagcacgcggaatgttcaaacatcaccaccggcgccggaac catccgcgacatccatcaccgacacgtccgctgacacggcgggcgacgggatacacc taccactgcacatgcacaccacgcgtccggctctctccaccggaacgacaggtacc cggcgacagaagaatatga	
1154-1155	Ga0315295_10106663 GENOME_ACESSION: YL17G14_0_MG_2_SF_3300032156 GENOME_ID: 281646 CONTIG_ID: 106662 SOURCE: JGI DATE: 2019-06-13.	atggctacgaccaccacacgtcagaagaccaccagtcgttctccagcgggctgc tctggatccctgtcagcagacaccgttggagtacgaacgaacggggcaggaggct cacgtcacgggaaaccgtggcggtgcggtagtgcgtgcaggggagacc gtacgcgacgacgcccgtgcgtacggcgacggaacaccgggtcacaccggaagt ccgtcaccgggtgggggtgacccaccaccacaacactgaatcaacaactgaacaggt cggttaccggcggaagggtttctgctcaccggcgggtgaaccgttgatccgtgcc acccgcgcgggcgacgtcaactctcgaagggcgccgctcgcgtggcccgatga tccgttaccatccgtgtggtcgtacgactgtcggcgacgtgaggtcgtggtgtgtg gtgaaactcgtacccgagcaaggcgacggggtacgtcgtgctcgtcgtatcgt acggcggtgacgtgctggtggcggtggaagtgcgccaccggggcgtcagatcca ccagaaactgtggcccgccgctgttgcggcggtggcgggcgacgttaactgcg tcatcgggcaccgggttctcaaccgtgcacgtcggaaaggcggtggcgccgtc ctgcagcatcgggtggacaacgtcaccggctgggttgaccgttgcggcggtggcacc ggtcaccgggtatcgaatggaactgtgctgttcgacaccagctactggagaaccga acatcagtggtgtggagtatcagcaggcgaccttggcggttgcaggtcaaggatga ctgctggagaagtggggcgcaagtgcgctactcgcacatcaccgggtggcactca acgtcgaccacatccaccggcggtcaccgggtggaagtacgggtcctcaacctac cctgtcgtgttccctgcaaccaggacaagaacaccggcggtcgtgagcagttcgtac cgatcccgacggttggcggtcctgtggcgccgcaaacggccgctgctgacgc cgccgggtcaaccaccggctgggcactgtggcgaaactggctgccacc	gtcaaccactccgcccgaaggacggagct tgtgggggcttgttcccgaaggttctgac atccgggtatcgggtgaccagcaagccat cactgaaaggaggcgtacccaagatggct acgaccaccacgtcagaagaccacca gtccgtgttctcagcgggctgcttggaa tccctgtcagcagaccgttggagtaacga cgaacggggcgaggaggctcagtcac gggaaaccgtggcggtggcagcggtgtg atcatgttcgaggggagaccgtacgcgac gatccgctgcggtacg
1156-1157	Ga0265297_10033497 GENOME_ACESSION: Munlanwell13791_2_SF_3300029288 GENOME_ID: 280701 CONTIG_ID: 33496 SOURCE: JGI DATE: 2019-06-13.	atgtattcatagtagtgaagaacaatcttcagcatccaacgaagaacatgacatga ttatcctgctgatcacaacggaagctaaatcattggaagatagatagttcaagtatta agaaattcagacatcctaaactatcactgtagattatcataggtctagatcctggataca aaaacatcgggtatagtttttaaatattataagaatcagatccagaatattctgaatgtga agtactaactagaactctgagattactaagttatctcagaagaagaatgtatcgtagaa gtaggagaagcaaacaccgagagaatattctgagaaatttgaagagcaaaattaaa gctcctagatggaagaacagggaagaagaacatgggctcctactacatgcatctgttt caaagtcactaaatcttctcagtgcatctcaatagagatataatcaatcagagatag tattagaacattttaaattcattcctaaagctttgattcaactgtctctagctggaagtat cagaaggtcctcaatttggattcgaatgttaagcagatcgttagagtagagataatta taaatgcagatagtggtgagaattatttaagcttagcgtccatcacatccaagagagag ctgattggtggaagtatagaccgagaatttagttactctgtgtcaagtgtcatttactcct tcatcaactctagctgaatgtccaagacctctaaagcttccaatgagagatagtgga gtattgaaacagttgatgaactacttagttaattatatttccagcttatactactggtatct gacactgcccgtctgagacactatataacattgagaatcacacgttaacgatgtctaaag ttagatgctttatcaaatagatctagaaaattcaattgtcaagactatccaactgtcaat cttaaacagtttctgtcgtcacagagaactgtgttcagagataggaagatagaaaataca tctgtgacggatttacgttgcgttgaatagaaaatctagatctacacggctgagtcgaag ccatctttacaagagtttaaacagaatctccagaagaaaagtagtagctaaacctggtga gaataatctactttagaactaattctcaagctaaagttagacctggagacatcttaaacatca gaacatttaattgttctcaagcagtggtgcttctactcaaggaactgttaacatctgaaactg aaattaaattttaaatacagaattgtcgaatagaaacattcaggattagtaaccacc cgctaa	gtcaactaccaccgctaaagcaggtggctt gaaagagccaacgtagctttagttgtggg ttgaacaggcactgggtgtgaaacacacc caaggatctgactagtttccatgggggt gccactcccagctcctcatcactggagta acagatagtaatacttaagctcctatgctgt catagtaggtaacgaagcagattactacta gaccaagccgaggagactaaagacagg acctgatctataggtcactaactaggagaaa cagatg
1158-1159	Ga0256842_1000027 GENOME_ACESSION: 4559240metaG_2_SF_3300026512 GENOME_ID: 20948 CONTIG_ID: 26	atggttttttattaaagcaagaataaaactccttagcccaactagcgaggtctaaagctag aattttactaaaaaaaggaagcattgtgtcacaaggtttatccttttacaataagattaaaa gagaataaagagtgtaaaaaatttttgaataaaatttgatgtggtcatcaggttacagg agtagctatagtagatcaggttaattgtttttcttgcagagatagtcagtagaggagcagt aattaaaaagcaatggaactcaaggagagcaataaagaagaagtagaagaatagaaaa actagatacagagagccaagatttgataatagaactagacctgaagggtgtgttaccacca	gtcaataacctctccaatccttgacggatag ggaagaggcttgattgaccagactagttg ctagaataagctaacacgataatttatgacc acaccttggagtgcttccagctccaagctc tgtgtaggctcttgaattgggttaagccct gtgaacctaggagcgtttgcatagcaaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
	SOURCE: JGI DATE: 2019-01-27.	agtgtgaaatcaagagcagacaatgtaataatttgcataaaataatcgtaaattatacc tatttttccgtagctgtagagaaggattatcacatcataaaggctaaaaaacaccaaac aggcatcaattgttcaagtgctagaacttatactataaaaaggttaagtagaactttgatg taatatgtaggagaggggtgggaacttacgcaatagaaaagagcttggtttaccaaaac aacactactatgatctatgtgtattggaaaaataataatgaaatagtaacagataaa ttacttagaattaaagcacaagggaaggaagccgacagatgtgtcgaatggatgattt ggatttctagaactaaaagctaaaagctctaaaatagtaaaaggtttccaacacggtgaca ttatgaaaagcaatagtaagcaagggaagaaagataggaacttaccttggaanaagtgctg taagagctagtgggaattttaacattactaccacaaaaacgactattcaaggtatttaattata aatactgtagaacaatacaaaaaggagacggatatgcatacgcaatcgcacaatcaaat ag	caagcattaatatcattgtcaggagagaggt caaaagaccgttaccactgcaaggtgagtta atttgaanaaggattttaaatg
1160-1161	Ga0209123_1000186 GENOME_ACESSION: BarGra031ASPADES_14_FD_SF_3300025718 GENOME_ID: 23742 CONTIG_ID: 185 SOURCE: JGI DATE: 2019-01-27.	atgaacggagtagtattgtcatcagtaataagcaaacagccgttatgccgacctgccggcc aggggcacgggaagctctcatccggaggcgaaggccgctgtgtccgaagttaacattac gatcatctaaaggatcgtgcaatcggggtgtattcaaacgggtcagggttgaatcgtatccg gggtccaagggaacaggcatcgcttgggtcaacgaagtcacaaagttgtcttctgtt atgtattgtgtcaccgtggattggccatcagcagatattagcatccgttagggctactcc ggagtgcccgagaaacgtaacacagataccgcaagccgggttggtggcgaacacga caaaactgtagggttggctggcccatcgctattacacgggttcatacagacgatgacct gggtcaggagggctttcttgcctagcggcgctgctgaactctcaagagctgtttaagttt gatctccaaaagcttgagaacccagatataagcggcatcgaatacagcaggggcacgct ggctgggtatgagttgcgggaatacttttggagaagtgccacggcacctgtctctactgc gatgtcgaagatattccactccagattgagcaggtaaaagccaagcaacggcgga ccaacgtatcagcaacctgacattggcggtgcgaacctgcacaacggcggaaggaaa actctaacatcgaagtttttctggcagggtaaacagatcgctgaaaaagatcaaggggca gctaaaggcaaccgctcaaggacgcaaccgcagtcattgcaacgctgttgccgtttgttcg agtccttgaataatacaggactcgctgttgagacgggatctggagtgtagaactaaattcaa ccggacgattcaaggatcggcaagcagattggatcgacgcagcctcggttgagatct cgggcccagtgatcatatcccttggcgctccacccccctgtgcgcaaggtctgtgtc atggcaaccggcagatgtgtggcacagataagttcgggttccccatcagggcacggact gcccagaagcaatttttctgtttcagactggcgatattggttacggcaacgtccctaaag ggaagaaaattggaatacaccggcggtgtgtgtgttcgtccaattggaatttcgattt caaaccgggtacaggctgtgttgcggggatcgccacagacactgcacatggttccatca aaaggtatggctatgcatacataataa	gtcagataccccacggctccggggggcttg aaaaagctctaattccgaccgccaagcca ccaattcaacggatcggttgcggggctta cgattaacgggtcatgaccggaggatgcat tgccagtatcccgatggctgctataaggaag aaacggcggaagggttaaacggcgacactga ccagcattgacaagctgttaattcattggcga ggcacacattactcccacagggaagacg gtcctgaaggaccaacctatcaaggatgt cggatgaacgga
1162-1163	Ga0373956_0001933 GENOME_ACESSION: WV9ome_14_FD_SF_33 00035119 GENOME_ID: 312721 CONTIG_ID: 1932 SOURCE: JGI DATE: 2020-04-07.	atgtttgtccctgtttagacagagaaccagcagcccatgatgccgacgacgcctgcccg cgctggcgctggatcgccagcgggaagcgacgcctttctggagagcggtgtctctctg cgctccggctcaatcagggaacctcgtcgcgcgtcgtgcagccggtagcgggtgggattg accggggtatgaaacgggaaggctacagtgctcatctggcgccacatatactatctcaat atccaggccaaaggcagggtgggtgtcaagaagcagagagcagaagcacgcggat gcgcggaaactgcgggaaccgcgaagacgccttgcaggcgaacacgtcccacccgag gcagagcaagaagacactgcggccatgcagcaggccaggtggcagttggaaactcgcg gggtggcccgttttctatgctcgtcttccgggtcagcatcatcgccgggtgaggacgtggct gctcgaaccaggccgggcaaacggcggttggaatcgtctgttttgcgcgtggaagtggg gaagcactggttttagagagatccgcgcgttgcctcttggaaactgtgcacagcta cgaaaccaaagctgtcgtgagcagttggggctcaagaagacggaagaaggttggcc gaggtctgggaagcgcaattgcgttgatgcctgggtgcctggcctaccgtcgccgtcggtg caacaccgccccggtatcaggggtgtgtgtcttgcgccactatctggcatcgccg ccaattgcatcgtttcaactggaaaaagggggcatgcgcaagccctatggcggaacac tctcaagaaggatcaaacgaggcgactagtcagatccgaggttggggcaaggccct ggtgggtggcacgatggatggcagactagctgcatgatccaggaaacatacaagcga ctgactcaggggcggaaggtgcaaggtgacgtcgtgacgggtgttgcgatggagaa catgcttgcctcgcctccatccccctcccaagaaaagcgtgttctctcccggcttcaaa agcgcgggttctggcaagtcgcgttggatgagcgttaactatgtgtgaggcgagc gagaggcaagcatgcagaagcttgcgtggcgttcgtcgtctgttgcataaacgcgt gtggagctcaaaaactgcttcgctgttaccgtctggcactgtgctgtcggcatg gtctctaggggggtgttactttagcacgtccagaggactggttccacctacgagctg ccacctcatctgatgtctggacctctgtcgttgggttgggtcggctccgcttttttgcata tggcctgggtgacgtcgcaggagagcctgggttgccttctgtcggcggttatgacgcc tgccgttgggttctggtctggcgaaatggcgctgttcgacgcagcagatctaccctacc tggtctggcgccaggctattcccgcaattgtgatcgaccattcctggttctgtgatggg ctatggcatgggaactaccgttgtgtctcctcctgtgtgtgttcttccctatgatcact actgctctgggtttcagacgcttgatcaggcgctgtgtggtatgcagcgcgctgagagg ggcctcttgggtccaatgctcaccagatcgaaattgccttggccctgcctgcgatcgt gcagctgtgcgactggtctgactctatcatgattggggcgctggctgggtgaattgtga ccaacacccatcagggaactggggccctgtctcaaatctcaagaacacagatataet	gtcaagcacccccggcatgcatgccgggg cttgtgaagccttccgactcgccagcggaa gacgagcaagcgaaagagcgagactagct tgtcaagctctggcaacggggcgattgaa ggaagctgaatcacctggcgctggcagcct ccagaaacagcagagcgctccctaatctt gttccactcgggtatcagcagcgaaggat tatctaccgaaaggggcttatgcacat

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gcccgtgatgttcgcaccgtgctggctgctgcaatcatgccccggggtctttacgggat cacctggggcctgaccaggatactgagattctctacacag	
1164-1165	Ga0316622_100118550 GENOME_ACESSION: OWC_Aug_OW2_C1_D 3_\$F_3300033416 GENOME_ID: 280756 CONTIG_ID: 118549 SOURCE: JGI DATE: 2019-06-13.	atgctcgtgtacgtgcaggaccaggaaggaaaccgtcatgccacccgccgtgcg ggggcgtgctgcccgtggctgaagagcggccggccgcgtcgtgcgcgcgaaccg ttaccatccgattggtggaccggcggggtgggtacacccagcacttcagggccggcgt ggacctgggcacagcgcacgtcggggtgacccgtcgtgtcccccagaggaagtgttc tccggcgagttccggctgcgcacggcgtgagcgggctcctacggagcggcgccgag ttccggcgggcgccggcgtagccggaagtccggcaccgccctccgcgttccgcaatc gccgcccgaggagcgaactcggccctcgggtggggccaggtggcgagacactc cgggtgtgtgcactggtggaggccctcctgcggctggccgcctggagcttggagatcg gcaacttcgaccgcaccagtgtgtccatccggcgtgcgggggtgggtaccaaca gggagagcagtagcgttcccaatcccgaggagtagctctgtggcgggaccgcca cacctgccagcggtgcccgggagcgtgggtgatccgcgcctgaccgtccaccacctg cggcagcgccaagagcgcgggagcaaccggccggccaacctatcacgctctgccg gacctgccaccaacggcaccaccaggaccggccgtcccccttg	gtcaaccaccaccgcttcagcgggtggcct gtggggagcccacaagccacgggttgac cagccccagccagggggccacaaggccc caccgggctacgttcaggaggcagacac ccggctgtaaacgctccagggccggcga ctgtcgggtaaagtcaaaagccccgctgg tagggggcgtgcttaccaggcgaagcct cccgaacatcgccgaggagcaccacacgg cccaacggcggttaccggcgaaggggt agaaggaaaccagcagtcgtctgtac
1166-1167	Oscillatoria sp. PCC 10802 genomic scaffold Osc10802DRAFT_Conti g7.7, whole genome shotgun sequence GENOME_ACESSION: GCA_000332335.1_AS M33233v1_genomic GENOME_ID: 47782 CONTIG_ID: 2 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgaagggtttgtttgatacaaatcgaagacctctcgtacgcgactcctctagggc aaggaaagctgctcaaggaggaaagcagctgtttttaggctgtatcccttactgtaattc ttaaacgggctgtgacagtgcagcagctcagcactgaaatagatccgggtt ccaaaaccaccggattggcattgttagcgaaggcagccggcgagctgtttggcgga gaattgacgcatcggggtttcaaatccgtgaggcattaaatagccgaaggtgaaacga cgcaaccggcgctaccgaaagactcgtatcgcgcacgaagattcaacaaccggcttc ggaaagctggatggttggcccaagcctaacagcagggttgagaatattgttcctggg tgcggcgctgcggcgttttgcgcaatcagcagcgtatcttcaggagttagtcagattgac acgcaagttatccaaaaccagagataagtggttgtaataccagcagggtgagttaca aggctcgcaggttcgtgaatactgcttgagaatggggcggaagtcgcttactgtgg ggccaaagagactcccctagaagtcgagcattatccaggagtaagggggggaagc aatgtgttttaacctcactttgcttgaccgccaccgtgcaacgagaaaaagggaaccgag atgtagcggattttctgtgggcaaacagacctactgcagcgcacccctccgctgcca agcgcccttgaaagatgccgctgctgttaactcaacgagatggcgctgtatgaaggttg aaaaatcacagattgccggttgagcaggttcggggggagtgaccaataataaccgaaa acgcttggttaccaaaaactcactgcttgacgccgcctgtgttgccgagtcacactcc gaaaatttggtatgctgcaaaatcgaagaccactgttaataaaagcgacaggagagg ctgccggcagcgcttaaccagacaaaaacgggtttcttattccataagtcagagct aaaacttatcagggtgggcaacgggtgataatcgacgagctgacattccgaaggga atacaccggcatccacagaggcggtgcaatcggtcaagacggccagttcaaaatc aagtgcccaagaagcgattcagcgtcaattacaagtatttaacccaattcaaaaag gggatggctacggttatagctctag	gtctgaccttaactcgtgacaaggaggtggg ttggagcgcgcgggcaaatcgcgacccact aaaggcgtataaatcgttaattcgttattta cgtcagccgaaaggacgggagtcggagtt cagcctgggttttagaacgacgggctacgt tggggaagtcagacaccacaggtgcggt ggcagctgtggtcgtgctcagcattaaa agggttagtttagtggaagccaggtgtgttg cttaaaaaagcttccaaaccttaacgaggca aacattaccagcaatgggctgacctgtgag ggttaattgtcatgaag
1168-1169	Ga0101763_1138 GENOME_ACESSION: IMG_3300006673_\$F_3 300006673 GENOME_ID: 308857 CONTIG_ID: 137 SOURCE: JGI DATE: 2020-04-07.	ttggcgaatggagttgactacatagggtcacaaataaagactgcactcgaatcgtc agctgttcgtcgaatcgccgaatcgaataacgcgacgagaacctcgattcaaaaat cgcaagcggccggaagggtgttaccgacagcctaattgacagagttttggagtcaaa acatgggagatcgcttaactaggttagccctgtaggttcgtcgttaagagttggttaa ggtttgactcagaagatgcaaaaccgggaattatcgaggatcgaaatcagcaagga gagttgttaggtatgaagtgcggaatatctactagaaaagtggggcgcaaatgcgcc tactgtagtgttctgacgttccaacagggaaccggagcacaacaaaatttaacgtgtc gttttaaatctgccgaagaacattggattgacgcgcctgtataggtcaggttgatactgta aagttagcaactaccacacccctgttattataactggtacggcagtcacaggtcat tcaaatcgacaaatcagcgttctcgtgaagggtatcaagtcgaagctcctgcaaggatt ggaagactgtgacattgttaattgtgttctgataaaaatgctgattaaaggcgtcag gctaaaaactgtcgtgcaaaaactcttgcggcatccggttaa	agctaccccgcaagggttaacgttatccgg tcagcacacttgggttacgtggtagcgac aagctctgtcgggtgtcgttaaacaggcaac cgggtgcgtcattgcccagtcggcattcg caaaaagccgtgtatccctcggaagccaa cgtgttgagacgaacaatgtctaattac
1170-1171	uncultured Nitrospira sp. isolate Nitrospira sp. RBC083 genome assembly, contig: MAG083_64, whole genome shotgun sequence GENOME_ACESSION: GCA_902500745.1_RBC 083_genomic GENOME_ID: 337206 CONTIG_ID: 63 SOURCE: NCBI_Prokaryotes DATE: 2020-04-22.	gtgcagacaaggcaggggtggccacgaacggccccggcagctacgcggggttgca acatggtcagggggagatgtgccgcgaggcacacgtcaccggcgaaggcgtgaaca aaggagtaatccaatcaatcacaggtttttgtattgaaccaagagacagccactgatg ccctgccaccccgccaggcgcaactactcaaaaaaggagcgggctgtcgtccatc gcgctgtacctttgtgacccgtcttaagaatcgcaaaagccatcgtgcagccggtgc gccttaagtcgatccggcgagcaaggctacagggcatggcgctggtgcgacaccagcc ctcagactctattaccgtgctgagccctttgagttacagaccgtgggtaccagatcagtg agtcaatgacagcccggtcactgcgtcgtcggcgagaggcaacctgcgctaccg cgaccccgctgttataatcgctcgtcgtcgggttgcttccaccagcctacagca ccgcatagataccacgctggtcgtgggtcaaacgccttcaacgcttggccctattagaac atcagttgcgaactggtacgttcgacatgcagggatgcagaatccagaattgacggc accggtaccagcaaggcagcgtgactggttacgaagtgcgtgagtactgtcgaaac atttggcagcgagtcgtctactgcgagcccaagacacccactacagattgagcatat ccagcccaaggcactggcgccaccaaccgaattagtaaccttgcgtggcggtgccgg tcttgcaaccagaagaaggctgcgagcgcatgaaaccttctgtgcaaaagccgga gataccggccatatacgaagcagccgagcgcccaactcaaggatcgccggcggtt	gtcaaccggccctccctaaaggaggagc ttgcagaaaggattcccgtaagctcgaggt tgaccagcaagaatgggtacgtcccaat ctactgttcgactagagctaaagaaccgac cctgggatgttctcagtcggcgcaatcg aaggcgccggtgcagacaaggcagcgggt ggccacgaaacggccccgccagctacgcg gggttgcaaatgtgcgagggggagatgtgc cgcgaggcacagtcaccggcgaaggc gtgaacaaaggagtaatcaattcaatcaca g

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aaactcaacgcgtgggcacttacccaggcgtgctgtcaaccggcctaccgctggagc tggcctcgggtggccgcacgaagtacaaccgcgaacgccttaattgtgccaagactcat gccccttgatgctgctgtattggcagttccgtgtactgaaaggctggcggaggagcact ttgcacctccgctgcacggggcggcggcaggtatcaacggaccagactcagcaaaagc gcttccggttgctactgtatgcctataaacgagtcacggctccaacaggcgtatct ggtcggaggcccacgtgcccctcaggcaaggaggcaacctaccaaggcgcgatagc agtacgcgtctcggcagctcaacatcaggacagacgaagaagtattcaaggaafttc gtatcgatattgcaaacgtattcagcgtggagatgggtacggatcattgtgagacctcaa atgacaagcccaaaagcacagagtatgcggcctgcttgaatgccgcgtattccatcca caatga	
1172-1173	Ga0209639_1000561 GENOME_ACESSION: EPASupHS05SPAdes_\$F _3300027566 GENOME_ID: 25530 CONTIG_ID: 560 SOURCE: JGI DATE: 2019-01-27.	atgcatgtattgtattggatagaacaaaaaacattatgccgtgtcatgcccggtagc acgaaaactcttaggcaggggagagctgcaatctataggcagttaccatttgcgataatt ctcagagaataaacaatgtgcagaaccacaaagtacgcatcaaaatgatccgggt tcaaaagactacaggactgtgtgtgtgggaaaggataataacgggtatcgataatgg gctgtagaactcaaacaccgtggtcatgccaataaagaattattagataagcgtaggggc aatcggagaagtagggcgaacagaaagcgtgatatagagctgtcgggtccttaacga gcggaggcgtgggggttggctcccaccttcttgcgaagtcgggtccaaataacgtaac atgggttaaccggctttaggttggcacccttattagtggtgtcctaactggaactattaa gatactcagcttatacaaaagctcgtgagatttccgtgtgtgagtcacgaaggcaggtgca aggctatgaggtcagggaatctcttagagaagtttggccgaaatgcgtgtattgctggg gagacagagtcaccgctgcaagtagagcatgtcatcccaagcatccagcgggtgtcc gatagggacgaatcgggtaagtaactgtacgttagcgtgtgaggtctgcaacaaagctaa aggaaatagtcagccgaacgattggtgggaaagctgcaacagtcacgatagctaaag acaagatacgggctgaaacttgcgaaggtctgaagcaactaagcaaccactcaag gatcggcagcgattaatagtagcgcgtggggcgtttatcgggtgttgagcaggttagga cttccgttagaagttgttcaggcggccttactaaatttaaccgaacgaacgtatttacc gaaacgacttggctagatgcggcgtgtgttgggaaagtagcggggaacagatcgtctt tctgatggccaactcgtgctatctgtctaccgggcatgaaacggcaacgatgtgta cggataagtaggttccctattaaacatgctccaaggcgaagtcgttcatgggtttca gacaggtgacatagtaattgctgtaattcctaaggcaataaagggaatgcacacagg cgtgttagcgtaccggttcagaccgagctttaaactgaatgattgtgatgtacacgaag tatctaaggattattcacagagctgatgctacgttatgagttgtcttgggggtgcaggt tctccccacagatgaatcgtggggctcccacctggcgtgtgattgtggagattaa	gtcaactacccccaggctaaagccggagg ctgtaggagagaatcttataagctcagttg accagcctaagcttgcacaagactacgt tatcagggtcatgacaccttgaatgcgtgc cagttccaagctcctgctgttagcattaaac aggtgtactagggtcaagccaggtgtgtaag cataacaagccttgataacttggcggagca aacattaccctcgaaaggggagattgcaga aatgcat
1174-1175	Ga0209347_1003246 GENOME_ACESSION: AutmicBR23SPAdes_\$F _3300027640 GENOME_ID: 23478 CONTIG_ID: 3245 SOURCE: JGI DATE: 2019-01-27.	atgttagtctacgttatcaacagcaggttaaacgttaagccctgcaaaccaagaagg caaaagaggctgctgcaagaggcagagcgaaggtcgtgaggcgcacgccctttacgg tccaactgctttacggctgagcgggttacagacaacgcgtcagggtcggagttgacacg ggagccaagtagctgggggttgcgctgttcgaacagacgagaagggaagagcaaaag aatactctgctgcaagcgcaggtgtcatttgcgagcagatattcggcgaagatggaccgg cgcaggggcctatcgcagagcaagacgtggccgggaagacaggtaccggaagccaag atttgataacagcggcgatgggaaggtgtgtggcccccagcatcagtcacgggttg acggcacgctgaaggtatggagttgtcttgcgaggtgtctcccgttgcctcagtagaagt ggagaccgcgaattttgacacggcagcgtatggcaagaggagcttgcgactgagggccc tggcagtagcaacggggcgagcagtagcaggttgaacgtcaaatcatacgttgcacat cgagatgatatcagttgtccagtgcaaggcctaagggtcgtccgttggcgggttcaccac ctacgacaagagctgatggcggaacaacgcgtcccgtaacctgatcagcgtatgtga agagtgccacgatcgggtccatagggcggaatcaagctgactgcgttcttgcgaaga acgaatctccggtatgccgcgcacaccaggcagggaagacggcgttatgacagccc tacggggaacaactccgacctctgagacgacagggcggcgtgaccaaaagtcgacgactt gagatgggactatcgaactcatgccaacgatgcgctggccatagcgggtacaggggt accctgcgaacccgtgacacccagttcttattgctgtgttcccaagggaactaccg gctgttcaaggtgctcgcagccatattcgaatcaaaagcggcggcagctatttggcttc cggcgctggcaaaagtatgcttggcgggtgtcgaaggggttcgcaaaaggcaaac ggctccttgggttacttcaatgtgagcagcgttgacggcactgtgatcagcgcgtcgatttc tacaagagatttgcgtttattagagaacaacatctctacttgttgagaggaggcagggccg tttctccccggactga	gtcaactacccccgactgaagtcaggcgtt gcaggagcctctagttgaccagcctaagcct tgagatagaggctcagttgacgtagcagctt agacgccttaccgggaatgtccggataggt gacctaccctggggtgccccctcagctcca ggctctagaactccctgtttaaacagtcctg agagggtagggacagtcgcttgagatgc gctggctgtcaacttggcagaggagacct aactccgagaggaggtagactctatgta
1176-1177	Streptomyces sp. SID12501 NODE_3_length_364520 _cov_60.6328_ID_5/0, whole genome shotgun sequence GENOME_ACESSION: GCA_010550245.1_AS M1055024v1_genomic GENOME_ID: 324827	ttggctctggaattcagtcagcagacaacccgggagcaggacgaacggggctcg gacggcgttaaggcatccgggtgtggaacatgtgcgaggggagatcaccggcagtgac ctactgcgggtggcgtcacccccggaccatcagggtcgggagccggcgtgaggccg acccccccgtgttctcctggaacagcagcggcagcattacagcccaccagccccgc caggggcccgaagctcctggtatccggccgcgcgtgtcggccggcacacacggttc gtgatccgtcgaaggaccgcacggccagcctcggagggtcagcgtgtcgtgagctgg gcatcgacccccggtccaagcacaccggccttggctgtgttaccggccgggacggtga gcggccggggccgggttcgggtccagctcgcaccacggggcgccgcgatccgcaaga agctgtgagcagcgtccctaccggcggggccggcgtacccgcaacctgcgctacc gcgcacccccgttctcaaccgaacccggccgaaggcgtggctggccgctgcctcca	gtaatgggatcccccgctgaaggcggcg ggggctttcagcccagccggctgatgtgcc gctgttaccagcaccagcaactgatatgga ggtgaccgaagttggctacgttcacgcag gacagaagaccacccatctgtgttcctca gcagttgctcgtcgaattcagtcagcagac aacccccgggatcagggaacgaacggggct cggacggcgttaaggcatccgggttggaac atgtgcaggggagatccggcaggtgca cctactcggcgtggcgtacccccggaccat

250

SEQ ID NO:	Name	IscB sequence	ωRNAs
		caccactatgatgcagtatctattatctcggctgtggccatccggtgatttgaatctcggg gggtggacggatttttagtacatcccctcatctgattatgcagttccgacatcatgatcgcca gattattcattgccaattcgagcgacatataaagttagtagagataacggaaagttagttc cgtggtaaaaaatcgaaagccccgttcgagcaaccaaaatcaatgcctgcgctcaatgt gtggtacgatgcgaagtgaagcggagcggcgagcacaaggctcgtctggctctatctc agctcgttgtagtgaagcaatcgacgtttacaagaacccccgtcgggtatgccaggaa ccgtattccgtacgaagacattttagttatgcaaggcagcgtctgtacggcagttat tttgcgcgacgtcgaagggaaggtgttctcttcaaaaaatcgaaagtttggccgt cgaggacttattatctgtaa	
1184-1185	human gut metagenome genome assembly, contig: NODE_551_length_2749_2_cov_18.9512, whole genome shotgun sequence GENOME_ACESSION: OBX001.1 GENOME_ID: 2388 CONTIG_ID: 550 SOURCE: WGS DATE: 2019-01-27.	atggtttatgttttcaaaagggggaaacctctcatgccatgtacaatgttattgcaaga ttgctgttaaacaggacaaaggcaaaaggtaaaaagaagatgccgtttatgattcagctga cttatgatgctacgcaatataccaagattgtacgcttgggttagacaccggagcgctcat attggtgctgctgttagatgctgacaacgtgtttgtacatgtctgaacgaaatcag aaatgatatacacagaaaaatggacagcagcagagcctatcgtaaagtaagacgatac ggaaacgagataccgcaaacagatggatgaaccgtaaaaattccaaaggga cagggttctctacaatgggtgagtaaacatcacatcagaagagattgaaattgtaaa atctattttgcttattacgaggctgtgattgagacaggtacattgattgccatttaataa acccaatgttatcaatcaaaatacaggcaggtgggttattcagcaaggacgtgattatgt tttgcataatcaaggcaaaagtcttaaacagagacgttatcttgcgaatgctgcagg ggaaacggaaagattcgaagctgaagttcaccagattgtaccgctcgaagggtggctc taacgaagaataacgtgattactcttgcatacctgccattctgacttcatcatggaat gatgaattaaaggtaaacggtaaacaaaaaggtaactgcgatatgctactacaatgaac agtatcagaacacagttgttaagctttatccgaagcgtattgaacattggatatgtaca aaagagaatttgcacatgctcaggtctgccgaacgcattgcaatgatcggttatgattg ccagtggaggaaatcaggtgaatttcaaacacatagcttatgcaggaaaaatgtatcc ccaaaggggattatcagcaaacaaagggtatatacgaagtgcagccactataacgaaa aagataatcggatttcgcaaatgtgaaggtacaataccttggtaagaatattttataagg gcaggatgtctactggatatactgtacttatgataattggaacaaagtagattttcgtat tatccggaagggttataagacacaaatgaagaactgtgaagaaattacggcaaggaa cgggtggatgattcaggaaattgctatctga	gtgaactaccatcccctaagggaatggg ctccattaaatgggtaccagactcagcagc cagaatgactgctacgataattaggtcatg acaccctcgggtgacgcatcagaccgtgct ctatcgagtatgttaagtgggtggtgtaag caaacccctgtatatactcgaagagccct tttatcattgtcagatgaatcggaacaact gtatgtaacagtaggcatagtagcacca cctgcttttaagcagagatttatcgaggatt tactcaatgg
1186-1187	Ga0310136_005540 GENOME_ACESSION: M271_2_5F_3300034520 GENOME_ID: 292992 CONTIG_ID: 5539 SOURCE: JGI DATE: 2019-08-09.	atgacattggcggagcccaatttaccctccgaaggagagattgcagagatgtatgtattg tcgttgacgccgaaggaaaccgcttgcctatgtccacagtagtcgaggaaactg ctgaagcaggggcgtgctgcgagtcgtgaggaatattcctgactatcgtacttaagag acgaaggaaactcaccgagcattaggttaagattgactccggctcaaaagcaac gggaatgataatattacaaaaaagtcatttgggcagcagaactgaacaccggg gacacaagataagggaagattgctggagaggaggcagcttcgccgaagcaggagat accgcaaaagaaagtatcgcaagccgaggttgacaaccgcaggagaccggaagggt ggctcccgccgacttgaagcagggctcccaacataaacttgggcaaacgggtta ataaagctgtgttcaatctcagcaatattcttgaagcttgcgaagttgcacccagaatta cagaaccggagataaccggtatcgagtaccagcggggcagctttacggctacgaag tcagagaatacctgttggagaatggggagggaagtgtgctactcggcaggaaacga cgtaccgttagagcttgaacatatagtgccaaatcgaaggcggcaccgacaggggt ccaacctgacgctgcttgcacactgcacacagaagaaaggcaatctcacggcaga agaattcggttattcagaagtcaaaaaaggaaggtccattgaaggatgttgcgc agtcaatgcaacacgctggcggttatgtgcagactgaagaaactggttgcctgtgga gtgcgggacaggaggaatgacgaagtacaaccggagcaaacgtgggttgcggaagg gcattggacagacgacgctgtgtgaggccagctacgctgaacacgtcgggtgagc attaattctgtcttgcaggtcaagcagttggcagtcaggagacaaaggtgcataaca gacaaatattggtttccgaaggcgtatgcaaacagacaaaaacatcagggttttcga cgggagataattgaaggcagtaatacgaaggggaatgcggggagccataattggc aggatagtcataagacagggcgtcgttcgggctgaacggcttgacgtacaccgaa gtatctgacgattctcagcaggcgatggttatgattacgccttttagcaatagagcga aa	gttgactgctaataactgcagttggaggcgt ataagcctcagagtcacacgcctcagcca ggggcgtgaaggctatcgggtcattatcct gggtcacgacacctgggggtcggggccag ctccaggctctgtcccgctgttatgagga gctgacttgcgaaggctcgaaggcagctgg catgaaagccaggatgacattggcgagg ccaactttactccgaaaggagagattgcag agatgat
1188-1189	Ga0187878_1000231 GENOME_ACESSION: June2016WEW_17_1_5_5F_3300018005 GENOME_ID: 30401 CONTIG_ID: 230 SOURCE: JGI DATE: 2019-01-27.	atgaatacgcgagacagaagactaacaccaccaacgcctccctaattgctggatggga tgaagccgtacatcgcgagggtagcggagttaaaagctgacacagcttctgttttatctc atggcgaaggagcgtctaaaaattcgaagaatcaaaaggacattgtagcttataccca acaaagatcttaagatttgggaatagtagaataaataagaaatcaaaatcaaaaaa gaataactatagtagtccgttttatccagacggaaaaccaggaatgcctacatcta aagagagctataaattggctaaagaaggcgaagccgaaagaaagaaataaattgaa tgtatttgcattaaattgaagtcttggcctatctatagaatttacacagatagattgtta agatcctgggagtagacacaggaatagctgtgagtcgaagaaatgattttgattcata tatgttagaattccgggatataagaagggtcgaacccgttactgttatcaatagacata gaaagaaaatagagaagtacacatactattgtatagagataacagatagaaggagg ctaagaagatcaaggagacatagaattgtagaagaagatgaagaatggctaaatag aactaagaaggtaagatagcaccatcgtatttagcaagaacaattgaattagaggc	gtcaactactgccccagcagatcagctttag ctgagagcgactccggcttccggcactc tgaaggggagcgtgtgactcagaatata acgaagaattgactagcctgtaggagaat atctaactatgaatacgcgagacagaagac taacaccaccaacgcctccctaatgtctgg atgggatgaagccgatcatcgcgagggtac ggagttaaagctgcacaca

SEQ ID NO:	Name	IscB sequence	ωRNAs
		<p>ggftaatgaactgagtaaatgtatccgattacaatcataggattgaagatgtgcttcaat cactggggagataaagatgtacaagggtcaatacttttcaagtagaaattgtaaga atttgcattagatagactgaaaaagatagctcttccaataataaaagggtatgagact gctagaagaagggaacaattgaaactgcttaagaaggagataagacaaaagggtcaat taaaatctcatgtaactgactgactgcaatgggttcaataatctgaataactaatcttct agtaataaatttcatttcaatgtaatacaaggcctaataatctagaagagctctattgacag aacagcctaataaagatgggtattacagaaagagctggaggtcatattctcatacccaat attcaaaaggattgagaaaaggagattatgttgaaagcaatgcccgaatttaagaagata tatagaggatggatacaggatatactgatgacagaatctatatctgatttcgattggaat caatctccatcttttctgtagataatattgattattggatagaaatcaggattgataaact gagattgggatggattaaagatactatagattttgcaattggatctaacaattgatgc agaaaataagataataaataatgaaagtcataataataattatagagatgaaaaagaaagct gaaaaggatgcaattaaagaatcaacaacaagataagacaatagagaggaatag atgatgcttggttaa</p>	
1190-1191	Ga0268280_1010517 GENOME_ACESSION: saktG_18_SF_33000282 98 GENOME_ID: 37736 CONTIG_ID: 10516 SOURCE: JGI DATE: 2019-01-27.	<p>atgcagaagttaacaagcaaaagttaaagaacatacctacagatacttctctgtctgtagctc tataaattctgtattaaacaagaccaaagtcttagtgcagagaataaagtactgactgttaa caattccgaagagaatcttcagaaacataaactgggtcagacttgagattctgaattgtt ttatgtacttaataaagggggtttagctcttctgctcttgccttgccttccaaaggaagaggct cttaagggttagggaaagctcaagtagtaaaatgtatcttttactatacaattgactaatgct agtggagaagcaaaacaagagatagattaggagtagatgagatataaacatgttggg atttcagctattagtcagataaaagaattatattcagcagaggtagtccatagaacaataatt tccgaattattatctgagagaaaaatgtatagaaggaataagaataaattatggtatag aaaacctagatttaataatagggaaaaaggttaagggttgcctatctccatctgtacaacata aaatagattctcatattagaataatagataataaagaataattattaccaataacaaaattatt atagaaactgctcaaaattgatattcaaaaaataaataatcctgaaatattgtaatatagaatc agaatggagtagcaaaaagattttggaatgtaagagaatattgtattatagagataatcac actgtcaaaaattgtaaaaagaataataaagtattaaattgtacatccttgaaactaggaaa actggtgtaataagacctgataatttaattacattatgtttgaatgtcacaaaaagtatcatt agggaaaaatcaaatagatataaagattaaagaataattttagtagtagacatgtatgtctat aattaggaatagactaattgataaattagattaaagataatagaagagacatatggata tatacaaaagagtaaaagaatggagaataagatagaaaaagtcataattatgatgctttta tcatttctggaggatttaataaatttagagtagtatttttagtaactcagaagaggaagaa taatagaggg</p>	<p>gtcaactactccaggctaaagctacggagct tgcagagtaactctgcaacaagcaagtgctga agcacttgcagaggttgattggaagcaaaa aggaacaacaatgcagaagttacaagcaaa ttaagaacatacctacagatacttctgtgct tctgctctataaattctgtattaaacaagac caaaagcttagtgcagagaataaagtagctgac ttgaacaattccgaagagaatcttcagaaac ataaactctgggtcagacttgagagttctgaat att</p>
1192-1193	Ga0326511_10041793 GENOME_ACESSION: 0_2496_0518_D_2_SF_3 300031867 GENOME_ID: 279726 CONTIG_ID: 41792 SOURCE: JGI DATE: 2019-06-13.	<p>atggtttatgtgcagacattatgggaaggcactgatgccacagagagacacggaaa gggtgaaaaactctgctgacggcgcattggttctgtttgattgcgtgaaccgttcacaattc gcctaactgatgaaagcacctcatittatatacaagaggttcttgggaatagatgcaggaag ccgtcatattggtgtgctggctaccacagcagacagagaacttttctcggcacaggtgga attgctgacaaatattcagaaattgctgcaaatcgtatggaactacgcaggaacacgcag gagccgcaaaacccgctatcgcaaacgcgattttgacaataggagcagatgataatggct ggcttttcccaagtaccagacagaaaggttgatacacactcttaggtttatccgcatggtgatg gatatactaccattagcaaaacgacaatcaggtgtgcaaaattcgatgtccaaaagataa aaaacgatgcaatcgtcgggattgaataccagcaaggcagcaatttgggttttacaatgtt aggggatgtgtttacacgcgatggttcaatgccagcattgcaaaaggaaagtcacaaa gacccagttcttaattgttcaccacattgagagtgaagaataggaggtaatgcccacaaac aatcttgaacctgtgttaaacatgtcataagaatatacagaaggggagattactctaag atttatgcgcggtgtgattttctgtgacgcagcagcgcgatgaatgctatgcgatgtgctgt ataactcggcaaggatgagttccgtaattgtcatctgactatgggtatatcacaaagcat actcgaatcaggaaacgggaataaaaaaatcacatacggcagatgcaagggtgcatacgcg ggcatccgttagcggtagcacaaacggatgtgtatatctttaaacagcgcaggcgacaca acagacagattcacaaagtgtgcaattctaagtgaggagataggaactcaaccaggctc cgtatctgtttaaaggatacaggctattcgcacaaagtcttctaattggcagaggccttttat tactgttagacggcagagtggttcttttgcataaagaccattgattggaagtcctttcaga gggggtctctgccaagaaacttttctgtaatacagtagagggtttttaaatttctaataaaa agagtttgacgaactataataatcaacaatgagtaa</p>	<p>ctcctggctaaaccacaaaggttcgctcgg ctgctcttaactgattagactaaagtttagaa cactacgttggttgagaatgtataggcactcc gagatataatccagtttcggacactgcgg acggcgattaaaagaccgaaaggagatgt ggctggtgtgtgctgtaactgaaacctctcca actttgctgaggggtacaccagggagtaaf cccgacttacacaaagttgtttaaanaaag caaaaacaatg</p>
1194-1195	Ga0334886_1000752 GENOME_ACESSION: GraShuR1_MG_19_SF_3 300033176 GENOME_ID: 280356 CONTIG_ID: 751 SOURCE: JGI DATE: 2019-06-13.	<p>atgctagtctatgttcttaacagccacggaacacactgatgccgtgtaaaccacagaaa gcacgaagactattgaaagaacagaaggctaaagttagtaaaagaacacccttactatt caactgtgtacggttctcttgatacaaaagaatgtatattctcgggtgtagatgccggaag caagacaataggagtatggcctcaactgagaacaaggaagtgttgacgaagtagaat atcaggagaaacatgcttcatatttggagacggaatagcggatattgacatcagaaa gctcgacggaacgaagtgcatgctcagcaagctataagaacaaagtgttgaagtaa aacaacatcttactatgtgaaggaggcaggcgggttctcccccacagctaaagcagg gggtctccgccgattaa</p>	<p>gtcaaatccccacggctaaagccggggg ctgttaaagctctgtttagtagctgagtg cttcgagcactacgttatcgcaaatgtatag gcaccgtaggatgcgttcccaagcttacgc tctcgggttggtgttaaacagctcctgatgg gtagggacagtgctgtacagagaactgt cgaataacattggcgatgggaagattactcc gtaaggagggtggactatagctta</p>

253

SEQ ID NO:	Name	IscB sequence	ωRNAs
		agactgaatgataagaatgtatcgaaggataaaaaaagaacaattatggtatagaaaa ccaagattcaacaatagaagaagaatgaaagattggtgctccttcaattgaaagaagat atcaaacattttaactttgattgaaaaaattaaagatggtgccataacacaagtatag ttgaagtagcaaaattgatattcaaaaattagagaattcagagatacaaggagaagaat caacaaggtagattgtatggatatacaaatacagtaagtattctgaaaacagttcagaaaa tatctgtccattttgtaaacgaaagcttgagtcgggtgaattctaaagcaactcatcagatt tatgaggagtgattcacgaagaacagacagaatagaaggattatatttcataaaaaat gtcattgttcgacttcatgaacaaaaacgagaaaaaagaatttcaaacatcaagatcgtaa gtaacacatcaattttatgtctattataataaaagattttgaaagatattcctgatttga agtaacttatggaaatgtcaccctttgtgtagagaaatatttaggattagcaaaagtcaca ccaacgatgctttgtattgcaaaaggaaactcacaaacagagttagcttttgaataa ttcaaaacatagaacaacagaatacttcagagacaaggagaggatacaaaccaagc atcagaaggagtagataattatccagccttttagatttgattggatcaaggaaatataat atttctaaggaatgaatggatagggagatagctttatgtgattggagaaaaataaac atctctcaacatgaataatgttgaataatattttagtaaggcagtttatcatttgtaaaaa attaa	cccgaagtgaattctatctcaacatgcagaga gacttaagtcaaa
1203-1204	Ga0376445_000045 GENOME_ACESSION: PitchLake_TarA_s_2_\$F_3300035492 GENOME_ID: 306414 CONTIG_ID: 44 SOURCE: JGI DATE: 2019-11-15.	atgcaccgcgtgttcgtcctgagcagcaccgcagaaccgctcatgcccgtccatccggc acgggcccggcgctgtctgtcctccgggaaggcggaaggctctatcgccgggtgccgttc accgtcatcctcatgagcgcgcggcggagggcgagacgcaaccgctcaccgagcagctc gacccgggagcgaagaccacgggcatcgccctgtcgcgtgttccagcgcggccgc tgcgtgtcgtgcccgtgagatgagccaccgcgcgagcgaatcgcggcctcg acagccgcggcgtcgtcgcggcgggcgccaccgcaagaccggtagccgcgca ccgcgtttgacaaccgcgcggcgccggcggtgctgctgcccgcctatcgaatccagtc gcgtgcacaactgtccgcctgggatgccaaagctcgccgcctagcaccgggtgacagc catcgccgtcgagacgtgctgttcgacacccaactgctgagcaccgggacatcgcc ggctgtagtaccacagggaacgctgctgggctgcgagctgcgcgagtagctgctgg cgccaccatcacacctgcgcctactgctggggcagtcgaaggatccgatcctgga gcgcgagcagctgtgccgcgtccgcggcgagcgcgagcgggtggccaacgcgg tgcgtgcccgtccgactgcaatgtgccaagcgaacctgctgcccgggacatggatc gagcagctgcgtcgtcgaagaagccattgacgccaccgcggccggcagatggag aagatcctaccggctggcgccgagcctgcgcgatgcggcgcatgaagccacc cgctacgccatcggtgcacactcaagccgcggcctgcccgtcagtttcagcagcg gcggccgcacccgtttcaaccgcaccgctcagggtaccacaaagcgactgcatgga tgcggcctgctgctggtaggaaggcagcaccgtgcgcctggatccggcgccaagcc gtgtagatccaggccgtgcggcgagcggcggcagatgacccggccgaccgcta cggtctccgcgcaccggtgccaagcgctcaagcgctgctgggcttgcagaccggc gattgggtgcgactctgtcaatcagcggcaagtacggggcacctacgtcggccgct agccgtgctgtagcgtgagcgttcgacatccagctgctcgtgacgggtcagaagatca agattaccgcggcgagcgcgggtcaccctactgcaacgaggcgacggctatgccta caccacctga	tccatcgccgcgctggtagcgtgaggtt caggcgacagcccgacactgattcaacca gccccagcccttcgggtacgttaccgg gaatgccataggcaccgggatgcttcc agtcccgggctcgtggtgagctgtaaca gcacggcactgggacaacgggagtcg gctcacgaaaaccccgatgaactgggc gaggagacgttaccggcgaggggag aaccggcggtgaaccgactgattgaga caggagccgacgatgcaccgc
1205-1206	Ga0255344_1006041 GENOME_ACESSION: wastewater_treat_13_\$F_3300028564 GENOME_ID: 281794 CONTIG_ID: 6040 SOURCE: JGI DATE: 2019-06-13.	atgctagttccggctgtagatcaggttgaacaccactccaactactaccgccaaaa gcacggcgtgtgtgaaagaagacgttgctcgaagcattttcatcgtggcagcttctata tcgtttaaagaaattgttacagaaccgcacataaatatcattaggaatagatccgggttc gaaacgaactgctttacagtgctacaagaacccgtgtagtattgaattgagatgatt ctacaaattcacaaaaaataattgagaacacgactatgtatcgtcacaagcgcttat cgaacacaccgtatcgtgaatgtagtggaatcgcttagtcttcgtatgatgacgtg ttccgccctctacttactaggttgaaagacattaaagattgattatgtttgtaaaaaa ctccaatcactaaagtattattgaagatgtagctgctgttactataattctaaaaaga gattagtactcaaatgtattcatgttggaatgctgttttaccattacaagctggaaa gaattattttataattttagaagaccgtgggtagcgggtgtataagaaaaagggttgca gacagcgaagcatcgtagaacatgggtataagaaaattctgataaacttcaacaaa tgggaatatcaatgtatgattctcattgctgtgtagatgtattataatcgaagattagac ctgttcggaaatttaatttatacaatttctacatttctagacgcgaattatttcggatgttg gtaaaatcagaagaagacatgggtgcacaagaacactaggtcataacagagcagcttg gttattgtaaatattgaagcgtggtcgtattagaaccaatgattatgtttatgagggtta catgaaagtgaagaagaatcagtagtatttattcttacaaggtaagaattgggcc aaagtttcaaatgaatgattgtgtgattgaacaaatttactgtatttaataattatattcga acgaaaggtag	gtcaactaccaccgctgaatcggtggcgtt gtgggaaaaccgttagccaggactaacac aattaaaaaagactagactgtcaaatggg ttctttgaaactccgctgggtgagcggcaa cgcacggacagcttggaagaaactagcta cctgtcgtggcagacaacatgtgacgtttg ggtgtccgtagcctgattcctc
1207-1208	Bacillus sp. REN51N contig_15, whole genome shotgun sequence GENOME_ACESSION: GCA_000815345.1_AS M81534v1_genomic GENOME_ID: 69819	atggattttgtactaaccaagaaggaaaaccccttaattgcaaccaaacgcaccgtaag gtccgtctatgctaaaaacggcagaagctaaagtgtggaacgaaacccgtttaccatt caactattatttaaaacggaagactatgtcaaccatattttagggtagacagtggtttt atcatatcggtatctctgtctgacagagaagaagaagtgtttcatcagaggtttcgttac taaaaggaaatggtggaacgaatgaagaacgtgttcctatcgaagaacagacggct cgtctcgttaccgaaaacccgttttgataaccgaaagaaagacaaagatggttagcc ccatcattcaacataagctagatgtcaccctgcgttattttagtgggttaagctatattc	gtcaataaccctcgactaaagtcgaaggctt gaaaaagcgtttgagagtcgaatgcaaga ccatcggtctgtgtgtgattgacccatg ctctcttaggagactcgttagagtgaaat aggtaccctagaatactccactcgtttaggc tctacgtttgtgttaacatctctgagggt aggagaagtgggacaatgttaaacactct

SEQ ID NO:	Name	IscB sequence	ωRNAs
	CONTIG_ID: 6 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	cctatcttaacactattatgaagtcgctaattttgatacacagaagatattaaccctgatat tcaagggttagaatatcaagaaggcgtacaaaaaggtttataactaagagaatacattt tacaccgtgaccatcatatgtcagaacccgaattgtaaaacaaaagcaagaaaaa gtattggctctcatcattatctccggccaatggaggagtgacagctcctaataacctc attactctttgtataagtgctacacccacgaaaccacaaagggttttaagagactggag accaaagggtgaacgcctaaggctcgcacctttatgagcatgggtcgttggaagttagtc aacgctttggagtgtaaccacacatatgggtacttgaccaagcttaaacgcatagaattga aataagagaagacacatgccaatgatgcgtttgtcattgcaggaggaaacgaacacacg taagagctcaagtcacagtgaggcaggttcgaagaaacacccgttcttagataaggt ctacgacgctaagtacattgacgctagaacgggagaaaaagcagcgggacaagacttat ttaatggctgaaccaagcgaataaaaaatacgaatggagagaacctccgaaagtatcgtc aggaaaaaggtatctaaagggagaagaccatagctactatgcgatactcttccagccaa gggatttagtgcgaagtgaaggttacaccgctttgtgttaggcactcaaaacaaaggcgc ttacgttaaaactaaaagatagaaaaaagtagcgaaacccgtacattaacgttaattaa agtggaaaaaggattcgtcttctagactga	ataacatggggcaaggtggacctacagccttc ggctgcattactcagtgtaaggagttaaaaa aagaaaatg
1209-1210	NODE_3452_length_333 31_cov_8.192224 GENOME_ACESSION: mgm4491404.3 GENOME_ID: 237261 CONTIG_ID: 556 SOURCE: MG-RAST DATE: 2019-01-27.	atggcggaaggattacaacggacgcgagtcggaggagacaaaacttgagagtattgt tttaataaacgagggcaacccgctgatgcgtgtcaccggcaaaagcgcggttactgct gaaagagaaaaagctgtgtgatgcgtcgtacgccattccaccatccagctcacgcagg ccacgggtgaacacgtgcagccagtggaagttggggcgtggacgcgggttcgaaattcatc ggcttatcggcttcgaccgacaaggctgagttctatgcgtcagaagtcgaacttcgtacgg atgtagttgacttattgtcttcacgtcgggaactgcgttagggcgtcgcgaatcgttaagaca cgttaccgtgcaccgcgggtcaataatcgtgttcacagcaaaaaaagggttgggtggca ccgagcgttgaaacaaaattaacgcccatcttaccggttgacgctgtaaggcaacttc tcccgggtgacgaagatcgttgggaaacggccagtttcgacatccagaagatcaagaac ccggatatacgaagggcggaactatcagtggggagagcaactaggtttggaatgtccg cgaatacgttcttctgtgatgggcacgggtgcctgccaatgccacgaaagtcgaaagcc ccgattctgaatgtccaccatcttgaagccgttaaaacgggtggggacgccccgaataa cctgatacgcctatgcgagacgtgccataaggcgtatcacgctggaagattgtctgaa ggcaaacgcgcggacaagcattccgagacgtgcctttatgggaatcatgcatggacat tttcaaccgcctgaaggcacagtcgccggagcttgatgtacgaataacctacggctatct gacgaaaaacacccgcacccgacacggactggacaatcccatcatacggacgcttct gtattgccggcaacctcggggcaaaaggcgtgggtgagttcttccaaaagcagact cgcaagcacaatcggcaaaatcacaagatgccatcctcaaaaggcgggtgccgcaaacg acagcagggccccgtacgaggtaaaagggtttcgtgttcgacaaggtgcggatggca gccgagagcattcatcttcggaagacgcaattccggctactttgacatccgtttgtgaaa ggcgagaagggttctcctgccatcagttacaagtatcttcagcttttagaaagtcgaaaaac tttttaactcagcttattaagaaggacgcgattctcccgactgaagtcgcgggttctc cgctaa	gtcaaccacccctgactaaaggcagaggct tgcgaaagcaagcttagttgactagcccca gtgaggagaatccgggactacgttggtgg gaatgtataggcaccgcgggattgcaatcct agtcccgtgctctgtggccgtgattaaaag cactgagagtgatggtcgggttgcgggta tctgaacccctccaacatggggcgaaggat tacaaccggcagcggatccggaggagacaa aacttgaga
1211-1212	Ga0079226_10038291 GENOME_ACESSION: PoulitStandDraft_2_\$F_3 300006879 GENOME_ID: 34654 CONTIG_ID: 38290 SOURCE: JGI DATE: 2019-01-27.	atggccaccagtggtggccgcgtgacaaccaagccaaccggtgaggttgcgtggcggg gttgtgcccccccaagtcgctgcaccacagccgcgggtgttctgctggataagaataa gcggccgttgagccgatgcggccggatcggggcgcgaagctgctggatgcggccg ggcacgtgtgcaccgcgcgccccgtttgtattcgcagggtcgtatggtgatcccgca ccgaggaggtgattgtcgtatgggtgcaggtggggtgcacccggatccaaggcaac gggtatcgtctgttcatcaccacccgagggtgcggccaggcgcgtcgtggtggg gagctgggtcatcggggcctgcagatcaacaaaggatgcaggcacgtgccagttgc gtcgtgggcgcgggtcacggaactcgcgtcaccggcgccaggttcgataaccgtac ccggcccgccgggtggcccccctcgtgcccaccgggtcgataaccacctctcg tgggtgaacaggctgcgcagatgggcaccagtgatggcgatccacattgagcgggtca ggttcgacacccacgccatggccagccccagggtcagtgggatcagatcagcaggg caccctggccgggtatgaggtccgcgagttactgttgctaaagtcgaccggcggtgcg tgtactcggggcagtggtgtccgctgaacatcgaccacgtccaccgcgcagcaa ggcgggctctgaccgggtcgtgaatcgtgtgctggcctgcacgggtgtaacaggcca aaggcaacatggctgtgcggagttcgtgaccgacctgcc	gtcaaccaccccgccctgacggacggggc ttgtgagcctcgtactagatgggctggttgac cagaccgagagacgttgaaggaggaaatct acgttgagcgtgcgacaggagaccgacc tgggatgtcttcagtcacggctctcgaat cagtgccagcagacagcgtgggtgagcc acgaaacggggcactgaggtcgaagctga ttcgacaccggcggtcaacatgtcagagg ggagatggccaccagtggtggccgctga caaccaagccaaccgggtgaggttctggc ggggtgtgtcccgccccagctgcgtgacca cagccgcgg
1213-1214	Ga0194044_10012301 GENOME_ACESSION: JunL4417m_MetaG_2_\$F_3300021074 GENOME_ID: 30309 CONTIG_ID: 12300 SOURCE: JGI DATE: 2019-01-27.	atgataaccctggctatagattacgggtcctcaaatgtgggtatagctctggtcagaataac cgaagcggggaatgaaccgctgttcgctggaactgttatcctcgtatgaggaagctgaa ggaaaaagtgagagacgagggctggtcattagagggttgagaaggacgagaagacga agaatcgacgctcagagaacttggtgaagccctcctcgggcttaggcattgaggaggaga caaggctgccagaatagtcgggtctcgaataggcgtgggtataagttctctgtttcagac ccgaacgaaacgaaaaggtggacgaggtgaatcggcataccgatgcactagagaa cagttttccaccaactggaacaagaactcgaagagattctgagcgaccgtgaggttgc gacaagcactctcgtctgtgaagaataactcaaccgcaagggggaccgatacgtga aatcagattgataagaattgataatcggggagcagccgctgcgcatgggggattgca acaaggtgaccccgaggcgggacacgctactgacgatgcaattgccagcaactgggt cacgtatttcagatgcaatcaagacggaaccccaagcttgaatgctcaaccaaac gggttgaactcgttgcattccaagaatctccgagggcgtattgcaacaacgacgatt	atcagcacgcgccagctcagatcgaccga ccttattagctcgtatcagatcgtcac cgatctgtacataattagatggtctgcata agccatgaaaaccagaccgtttcctctgtat aggaaggctccgcttctctgtacggcttgt aatatcattcaaatcaagactgtccagattcg tgccccaaacttcttagcatcttggggc gatgtcgtcggcacgctcgttatttatcta tccgaattatcagattctccacatccccg ctg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ctcaagaagatactgcgcaggaggcgaagaagagctcaagaatcttagagccga gttaccaagcaccgaacctgagacgtctcaggcgacgcgtggaagtgtggaaaag ggtattctgaacaccttgagaacagtgccggcagggaaccgttactgcagggaacacag caagtcctactgtggaaaagtcttggaagggaagccaccagagtcaagtcaccattgc agattctgacattattccagacgggaacagattgcgttcagcaagctctggagatacata gaggcgcgactgcttctctcgtccgaagggaattgaccgaattgtgtgaacgaactg cgttcgaccttttagccggtaagcgaagaaaatagagacgcctccagcgaaggtgtg gaaaatatttatcagtaggccccatgtatgggtcccaatgaaaaggagatgctgcgaa aggaattcggagggtatgcgttactgtgtaacccgtctgatacactatggaccgaga ccacatctgccgaggcgagatttctctcgtatagctatctgaacacactaccgcttgc cgacttgtaattcggaaaaatcggcgagctctctccaggctcacttcgaataagttag gacgcctattcaatgtacaacagctac	
1215-1216	Ga0310914_10061055 GENOME_ACESSION: GREbulKMGAN108_2_ \$ F_3300033289 GENOME_ID: 280317 CONTIG_ID: 61054 SOURCE: JGI DATE: 2019-06-13.	atgcaacgaaaaagttaattccgatactagactggacaagattcaagactcgaagtttgg gtcgtttgaagtcactgaacactctgggagtggtatcccgaaagatatcccgatgcttc ccaagtcgggcttccctctggcgctcagcgtcgatgggattgcagacacctacacagca agggcttaccgccaaatgtgtgattgtctctgctcgtgactggaatggagaaccattat gccgacgtgcgcgtcgcgtctagacgttgggttaaacagcacaagctacgccgttct ggctcaatgtgtgtgtgtgtgtcggttgcgtttgagccgagcgacagggaataagcacg aagtggtgtgcgggattgatccaggaaagcaaacggagcggtacaccgtggtcttccaaa gagcacacctatctgaatgtctcagtgatgccatagattgggttaaggatgccgtgggat cgcgcaagattctgcgacgagcgagcggaatcgcaagacacataccgaactaaca gcagaatcgagcacgcggaggcatcccgccatcaaccaagcacgttggcagctcaa cgtcgagtggttaaccaactcgaaaagattgttcgatcacagggtacgtgtcggaaga cgtaaaaggcgactacgcacggcaggggcaggcgctggaatagctcttctcaccgttgg aagtcggcgaagagctggatgtacggagagctgttaaagctcggaacactgactcctaaa gccggctcacgagacggctgaactcgaaaaacttccggcgtgtgcaaatgctgcgggtga caagatgtcgatgcgcttgcatacataacgtggacagttgggtgctggcaaaagatgc acttggagcagcaaaaggacgcccggagaatatgcaactgggtgcacgaaggccactc cgattcaggcgccgacgttacttgcagaatccaacaaaggcggtgtccgctgtag tcatgttggaaacgctgagccaaggatttaagcggcgtcactcgttaagcatccaaagca cggcgtgacgagcgttgggtgacactaaagcagaccgtatcagcctgcacaaagtftcag gaaaacgccttgcagaacgcgaaggtcgaagacatcattgctgcgaagaagctcg atagctttccggccgataccaaaagcggcattcctctgccgctaagcgacagtttc ctgcctaagatttatga	gtaaacgaccgccagctaaagcaggcgccg tttttagagcagcaaaccaagctgcaacga aaagttaattccgatactagactggacaaga tcaagactgcgaagttttgggtcgtttgaag tcaactgaacactctgggagtgatcccgaa aagatgatcccgatgcttcccaagtcgggct tccctctggcgctcagcgtcgatgggattgc agacacctacacagcaagggttaccgcca aatgc
	Ga0256871_1000783 GENOME_ACESSION: RumR/JGAnnotation_SF_3300025804 GENOME_ID: 34932 CONTIG_ID: 782 SOURCE: JGI DATE: 2019-01-27.	atgactactagaattatctatgttttaaatgctaaccggcgaacccctcatgccgacgacgt ttggcgaaggttcgtcgtctgatacaatcgtgtgaagctcattggtacggaaatagtcgca cgactatccaatttaacacccgggttggaaacatcacccaaattgcattgaaggtgtgac ttgggaaatcacttgggatttccgtgtctgcacaacgactaaccaagaactttacagcg gaatcagtcacgggactatcaaggtgaagtcgaacgaacgtcaaacggcgagaatat cgtcgaccagacgtaatcgtttgcgtcaccgcaaaagtcggttcaataaccgtagaagaat ctgacgggttgggtgcaccaagcattcagcattatctgattttacgggtcgtgaaatcttac ggattcaaaagtgtttaccatcagtaaaagtcattcttgaacacgttctcgtatgttgc gctgaccaatttcggcgttcgtcgtgaagactatacaaaaggccgtctgcatggttatcatt cattaaagaatactatatacgtacgcaaaacgggattgacccgattgatggacgacattat ccgctatccgaaatggtatgtcatctctcaatcaggtagtcgaaggcggtaccaacagtc ctgacaataccattttgctgctcgtaaaaaccacaatactgccaatcacaatacgggtgtt ttggcagatcttccaaacattatcaaaagcagtttagtcaataactaaagggtctttttaatga atgtaatgcatttacgattgccaaaacggctgaataacaaacggttgcactgacgtttggc tataaaaccgcccagcaacggcaattatcgggtttgagaagaccgttaacgatcttgc atcacgccattgatgcgtttgattgccaatggcaacgccataccaaatcgtatgacgaa cattattcatcgtgaaaaacatccgtaataatcgacgttggaaaaattttatgatgcgaa atattattccaacgttgatggcaaggtttacgccgtaaaagacttgggtccggcgatacca atcgcaaaacaccggaacttacaacagtagacgatgcgaacgggttgaagaatct aaaggccgcttcaatccgtcgtcagcattatcaattcagccccatgataaaatcttatg gcagggaaaaccggttggagactgtctgggcacgatgagcaacggtataatccgtattgtt caaatggataaacaacgaaatcatcagcaccacaagaagtacaaattgttatcatatagt aataacttaattgaaactattattcaatga	gtaataactccggactaaagtcgaagctt gtaattctgtccatctttttgacagatggaa ttattagctgtgttgattagccttagttcaaatg ctttgaactacgttacttaagaacattacaat aggtacctgtgaatacttctcgttcacagc actacggttagcaattaaacaggactgcaga gttatgtccagtgttgaactccaaccttaa gataacattggcggaagagaacacacaacc gcgattctttccggttgactatcgtttataa agattgaaaggaggtgtgtcggtatgactac tagaatt
1217-1218	Ga0223826_10007721 GENOME_ACESSION: 3300021256 GENOME_ID: 241585 CONTIG_ID: 6980 SOURCE: JGI Dark Harvest DATE: 2019-01-27.	atgctgacatcgtattggcaaaagacggttacccgctgatgccacttataagatcaaca aggtgcgccgtatgtgaaggagggttaaacgcgggcccataagcccgggtt taccatccgcctctgtatgaagcggggaagggtgcgcagcctgtagagggtttgtgaaga taccgggtatggacaattgcgtatcggtaagtcggaagacacaggttcacacctga ggaatatacgttcttccggtatgaagacccgtcacgatgactgccgtaaatcacagaag gacaagaaggaaaccgcatcaggcacaggcgctcgaagttcgataaccgggaagaagga taaggcgtgatcgccccgtcactggataatagggtacagccatgtggatatacgtaac catgtataaaaaagtgctgccagtaacggatgtgactgttgaagtaggaactttcgatacg	ctatgtttccggcaaatatgttttagagcttcg gcatctgcactaaacgtataggcacttcggg gtgctccaaagccccgcgctgaaacccct ttcatcagggggttgatgtgtacgaaagtac aacccctgcctgtgtgtttaaaccctctga ggtcaggaggcggtgcatacagatatataaa ctgcccgtataacgttggcgtatgtgaaccc gcctgtaaaggcagagaacccccgcggcg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		cagggtgttggaagccgtagaggctggcagacccctcccaagggatcgggtaccagt atgggtccgcagtacgggtgtatagctccgggaagcggcttttaccgggacggataca aatgcatctgctcgggaagccgctattaggaccacgcgattcttaagatccatcacctt ggatttttaaaaggcgaccacagcgacagatgggaaattggccacgggtttgtcgaag tgccatacaccctcccaaccataaaccgggaggaactatataaccttaagcccaagctg aaaccgctcggggcgagcattcatgaatgccgtcaggtggaaaatatacaacatgggt aaagagcggaaacctgggtttatagctcacatgacatagggcgcgggtacaaaacgcgaa cgggttaaggcgccatattgggaagacgcacgccaatgatgcttactgcatagggttttcc acccaaaacataaaaagggcgtgtggcctataagaaggtgaagaggaacgaccgtat ccttcaaaagttttatgacgcagtatatatagaccggcgaggggaaaaagaaaaag gcgcggagctttctgtaacaggacgaaccgctctgtccgagggaacacagccggaat gaacggccgttccgaaaggagaaggcgtcaaaaggccatgtaacgaccgaaagggg cgtacacagctgaaacggggaagcctgtactgtataaagggaaggtaatgactgtcc ggcacacatacgaataaagggaaggtaacgtagaattcacgaaaaggcttcagatga caggaaatccgcaatttaagcaaaagtaacattatcagcctatgtatcaatccgggtgg gtgcgataaagctga	aacttaccgcggaagccgtaaggcaaatg ttactgtaaaccatacataaagaggagc atcggaatgctg
1219-1220	Ga0209656_10018899 GENOME_ACESSION: BogForECP1SPAdes_4_ \$F_3300027812 GENOME_ID: 23899 CONTIG_ID: 18898 SOURCE: JGI DATE: 2019-01-27.	atgcttccctcagtcgccgctcttcgagttgccgccgagacaagccatggtcagcgca aacggggcagcaacagacggctcaggtttcaacctgttcgaggggaaatcaaacggcg ctcaaaacggcgagcgcggggtaaccggcacgtttgtgtatcgaagaggggaatcct ttgatgccgtgttcaaatgccagggttaggtttgtgttcgtaaaggccgagccaaggttt atcggtctattccgttcacgattcagttgatagacaagcatccggcgacgttcaaccagt cgctataaagttggacccggagcaaatcagctggtgttgcctgttcgagagcatgc agacccgaccaacaacacgtgtgtccctcgggaatagcgacccgaagccgagc gatacggaaacacatgatccagaggcgatgttcagcgctgcggcgaaactgctaac cttcgataccgcgaccaaaggtttgataaccgaaccagacgcgaagactggtgcctcct tcgcttcaaaaggcagtcgacaatgtggcttcgtggttaaccgggtaccgcaaacgcga ccaattacctcaatctacgttgagtcctcaggttcgacatgcagcccttggaaaacccg acatcgagggtttagatgacacgggggtacgttttcgggtcagaactctgggaatacc tgcttgagaagtggggcggaagatgcgcttactgcgagtcgggaaggcctgccccttgag cggaacacatcgccccaaaggcttgcggggatccaatagggtcagcaacctgacgc tcgctgccacaagtgaaccaacggaaggatcacagccgggttcattgtttctggcaa acgatccgtccaggctggctcgcattctcagtcatacgaataaaccttatcagtgccg cagcagtaaacgttaccgaaaaggccactaccgagtcagtcgtcgcagacggcttagag gttactgctcagcggcggtgcgcaaaagttaatagtaccagctcggaaattccgaaga ccatgcgcttgacccgctgtgtcgggtgaactctcagaacttgagagttcgaacattcc cgtcctttgcatcaaaagctaccggcgcggaagtacacacgacccggctagacagttt cgggtttccgaggtatctcacgcgcaaaaggcggtgaaagggttttcaaacgggtg actgtgaaaggcaacatgccaaagggaattcaaaaggcgtgcaccaaggccgatt ggcagtgcgagtgctggagctttgtatccagagctcagcaggaacgtcgaaacga gctggaaacactgcaaacgtctatcggaacgacgggttacacctacgaaatcaactcac cggcgattcctccgccacctaaggaggcgggtccctcgttattaa	gtcaacgaccgccagctaaagcaggcgcc ttgtgagcgattgcaggcctgtgttaccag aacaagtgaataactacgttgagccgcaagc aaaagactaacctgggagcttctcagtc ccaggctcttcgagttgcggcgagacaa ggcatgttcagcggaacgggcagcaac agacggctacgggtcaacctgttcgagggg aaatcaaacggcgctcctcaaacgcccag cgcggggttaaccgac
1221-1222	Candidate divison MSBL1 archaeon SCGC- AAA382M17 AAA382M17_Contig_8_ C, whole genome shotgun sequence GENOME_ACESSION: GCA_001549475.1_SCG C- AAA382M17_genomic GENOME_ID: 234923 CONTIG_ID: 61 SOURCE: NCBI_Archaea DATE: 2019-01-27.	atgagtcataaaagctcgttcggatcatgagggcggaaccgctcatgcccacgaacc agcgaaggccagggacacctttagatgacgacagagcggaggtgttaagaggaagccg ttcacgattcggctgaagtaccacaaggagaataacggcgaggtgacctcgggg tggacggcggtactaccacgtcgggttctccggcgtcacggaggtgagagctgac cgggtgagctcagctgagagacgacgttccgataggctgagcgaggcgaggcggac tacaggagaacaggcgagcaggaagaccaggtacagagaacctgtttcagacaac aggggggaaggagacggctggctcggccctcaataaagtacaaatgagcgaaccac ctccagctcgtcagatgacggggagctactccgttaccggaaccgtggttgagggtg gccacgttcgacacggagaagctggcgacccctgagctggacgtgacggagtgctc cggagacgttcaggggttacgaggtgaaggagttacgtgctgcaagtgaggggagga gggtcgctactcggcgcgaggagcgtgccccctgaggtcgaccacatagtcggaa gtcgagggcggttcggaccgggtcgacaacctgacgatcagttgcgtggactgcaac cgggagaaggcgaccggaccgagaggtatggcgaccctgaggtacagaggaa ggcgggggaggtccctgaaggagggcggttctgtaacatcgtccgggtggaagatgtg aacgagctggactcgaccacacgtacggctacaggacgaagaagggaacgcatagac ctcggcctggagaagtcgacgacaacgacgcttcgtgatagcgggcgcgggggac caggagcgaactggaccctaccgctcaccggtggacgcgggaacaaccgcaagctg gagatcaacaggaggggtcacggccggagcgtccctcgggagcgctacgagttccag ccgggtgatttggtcagagggtatgacgtctgtaccgctcaggggttctcgtctctat ggcaggtacgtccgagcctttgacgaagaagggaacgaggaatacttgagagcggtat ccctcaaacctgtgaagtacggttaaggctctgctggcgaggagaaacagtcgaatt ccttcccaccgaatccgaagattcgaaggagctcctcgcgacaaccagatggaatg cgcgagagattgtgttaa	gttgactactcccactaagctcgttagga aggagcctgaagggttccgatccgggag cgggagactagccttggcgatgtccctcg ggagtgggggcagatgagctacgttcggg gctgagtggtcgaagacatgttgatgcc tcccaagtcgcaactctgcaaggcaggtg gacgaaacacggcgttatcccgatgtgtg caacatcacctgttcgagcccccacat tggcgatgggaacttactcctgtgtgtcg ggagccgggactcgagagtcggcgcat ttaacataccatttaggaggtgtacaaatga gtcaaaaa

258

SEQ ID NO:	Name	IscB sequence	ωRNAs
	SOURCE: NCBI_Prokaryotes DATE: 2019-05-30.	gagagtatgtcttagcaagagaccatcacaaatgtcaacactgcaaggcagatggttga aaggatcaagttacacgtccgtcacctcgtctcagaaaaggaggagtaaccgccctg acaatctcatcattttgtgtgaaactgtcacgcagcttaccaccgtggcgagttgaattaa agaaagcaccaaaaggctatgcccagctagtgcaatgtcaatcatgctgacccctgc ttgaccgactcaatgaatttggtgacaagggtggagacaacgtttggctatttggcctaaa gaaagcagactcacaattgactgccccaaaacagcatgactgatcccttcgtcattgcag gcaatctaattggctgacagattagactttcaagagctgagaaaagcatgtgcctgccataa cagacaactccacaagcaaaatttctaaagggtgcatccgcaaaagccaatcaagctcc aagagaagttcatggtttgactgtttgacaagggtcaggtagaaaacaaaattggtttgt tttggagaagaaccagtggtctatttgccttcgtatgtgacaggagaaaagctcaata agggggagctatagtgccaaaaaattagctcgtacaccgagcaaaatcagtaacacac aatacgcacaatcgctccaacgggagcatag	
1229-1230	wastewater metagenome genome assembly, contig: NODE_1534_length_122 24_cov_158.353768, whole genome shotgun sequence GENOME_ACESSION: UOQZ01.1 GENOME_ID: 16109 CONTIG_ID: 1533 SOURCE: WGS DATE: 2019-01-27.	atgtttaaacaattaaactttaataaagtttagtctggatactaacccgaaaccactgatgc catgtatccagcagcagcccaaaaattgcttcaagtggtgctgctctgtatttgcagat ttccattactattatttgcattgacgaactggggagaaagcaatctacaagacatagaaa tcaaaatcagaccagggttctaaaaccactggtgtagccttagttgtacatggtgccacagg ccatgctgtggccttctgtgcacacattgagcatagaaacaaatattaaattgctctgtattca agacgtgctatttcagagaagcaaaagacaacgttaaaacacgttatcgccaagctagattt ctcaaccgtactaagcctaaaggatggctgccaccatcgttggtaagcaaacgtgagaat attttaaacgtgggtgtccgtttgctaaattaaccccccttagcaagttgcaacttgagacag ctaagttgacacitcaaaagctgagaaacctatgattaaagggtgtgatatcaacaaggt aaaatgtttgatactgataagaagccttacttacttgaacgtgaaaaactattactgtattt actgtggcatacatgctagtcaagccaaaatggagattgaacacgttataccatagaagtaa aggtggtacagatagctgaacaatctgtattatcctgtgaacctgtaatcaggccaaaag gtaatcaagatgttgaacacctctcaaaaggtaaaccaaggttattaaaacgtgtgaaaaa acatttagatacaagtcacaagatgctgcacatatgaacagatttcgttggatgtgatga acaatctccgagcagtgcatgcatgcatggtgcaacttaaaaattggtttgtttgcttaact actaaacaaaatcgtttatcactgtgattgccaaaagatcattgattgatgctgctggtttgt acaagtgatggttagtactgtcaagtagaaccaaaatttaaacccgttattatcaaaagcagt gggtcagggtcaaggcaattctgtctgtatggacaataatgggtttccaagaactagtcga aaacctagaagtaaaaacttcttgggttttaaacaggagatatggttaagccgttataact gaaggagctaaaacaaagctacgtcaagctacatgtatggccgtgtagcgggtgcgtag tacagggttacttggatgtaaaaacagaaatacgaatcactatgcttataaacattgttaa accaattcacttaattggacggatagttatgcgttaa	gtcaacaacccaccctaaagggtgaggtct tgcaaaagcctagtgaccagccttagttta agctacgttatcttagttatcacacttagagt gcttcaccagctctaaacccctgtgcaggctct gtaaacatctcttgggtgaagagaagtaacc gtaatttggcattaaattgcaagccttagataa cattggcgagggtgaacgttacaacagtgatt tatcactaggtccgtaaggaaaaaataatgt ttaacaattaa
1231-1232	Ga0099741_1041 GENOME_ACESSION: KloWWTsluDR11_LA_2 _SF_3300008024 GENOME_ID: 30893 CONTIG_ID: 40 SOURCE: JGI DATE: 2019-01-27.	atgcaaaatcaagagttcagtggttgaatccagatggctgccccgctatgccaaacaaac ccagccgcgcagctcgttgcctcaaggaggcgaagccagcaggttatctataacgatctg agcatctctctattcagctaatagaggagccatcggtgtagaaacactcagcaggtggtgt taggtattgatccagaaaattatcacagcaggtatagcagcacagacagctagagccactc tatttatggcgacactgcaattaccttccagacgggtcaaaagatcggatggagcaaacgccg agtgtatcgtcgttgcagcgtatcgttaactgcccgtcgtcgtcccgccagattctgaat cgacgggtcaaaaagattaccacatcaatcaaggtcaatcgtcaacttgaactacgggta gccaaaagaattgtgcccgtatataccatcactctgattgtttacgaggtggtcaaggctgc tggttccaagttcttcccctgtaatggttgacagttttggtatgcttgaacttggagaaa ctgcgcccaactgagcagaaaataggttgggaactctcaagttcgactcaattgggg ctggaaaagcagaaaaatcacaaagcgacactatccctcaaacatcagtgatgacgg cattgctctagcagccagtcattcctaactatcaacaatggcacactaaaatgctcatg gagctaattgggtgggtttttgtcgggttactcctgctctatttttctgattcgcagaccac aattaaaccgctcaactgcactcctggttccggctattgttggaaataggcgcaagtatg gcggtacgacaactagggcatggatgaagaaaggagatctatgccaagcagaacagcg aagtcgtgtatctattgttgggtgagtgccgacagcaaaaacaaatcagtgcttaatt tcggctggaacgtattgctcagttacagcctcaaaagttccatttgattcagcgtcaacg gtgttactcgttgcgtcagatggcaagttgcaaggttgacggcattgagtcatacaaccttg a	gtcaactacccaccctagtcgaaacgacc gaggtgggcttaaaagaaagccaatcttt taacgcaagagtgactagaccacttaca acatctagttggcagcacttccgaatcttc tctatgtcggactatctcgaactgattggttt cagtggttgaagacaagacatcttgcctag atgtgttcgaagagacaataactgcaaaag gattatctctatgcaaaatctaa
1233-1234	Ga0376514_000350 GENOME_ACESSION: B38_SRoWell1209_FD_ _SF_3300035040 GENOME_ID: 311412 CONTIG_ID: 349 SOURCE: JGI DATE: 2020-04-07.	gtggaaggatggttgcgccttcattaaaaagccgggttcacaaatcgaacacctggacc aacgggttgcgcgttttgcataatcagggcaatttcaatgaggtgtgctggtttgacatg cagaaaaacagaatcccgaataatccgggtgtcgcgtatcagcagggggaaactcatggg ttatgaggtgcgggaataccttctggaaaaatggagacaacgtgtgcttattgcggaaa aacagatatccgttggagattgaacatctggtccaaaatcgaagggtgttccaacag agtcagcaatctgacactgcccgtgtgagggcgtgcaacaggaagaagggaacaagcc cctagaagagttccttcaaggaaacgggagctgttgaagcgaattcaaaagcagttctaa agtgcctcgaaggatgcgggtgtccgtcaacacagccagcagtggtgcttaccgtacctt gaaaaaaatcggactgcccgttgaacaggctccggcgctgcaaaaatcaaccggga cgaccaggggacttcacaaaacgacttggtcgtgatccgcctgtgctgggaaaagcac acctgaaaagattttcagatcgacaaaacctgtttgattgtaaggcagacgggtcatggc	gtcaactacccctcggtgaagaccgaggg gcttgaaaaagcccaaggttaccagctta agtgttcgagtactacgttagatgaaaac aggtatcctgggtgctcgcggcctcagg ttctacggcaagtgttgaacagggtctaaag gggttaaacagggtctaaagggttaaacgggt gctgcttgcgcaaacccgggtcataacattg acggcgcaaacattacttggaaacaagag gattttaaattgagt

SEQ ID NO:	Name	IscB sequence	ωRNAs
		tcaaggcagatatgcagggttaacaagttcggattcccgacacagcaaaagtaactga	
1235-1236	human gut metagenome genome assembly, contig: NODE_3498_length_12905_cov_12.405058, whole genome shotgun sequence GENOME_ACESSION: OWGD01.1 GENOME_ID: 11975 CONTIG_ID: 3497 SOURCE: WGS DATE: 2019-01-27.	ttggcgaaggatcattaccggctcaatagaccgcaggaaactaacttgagagtatatgtttgaacaaacgaggagaaacgcgtgatgccgtgttcgacgcaaggcacggattcttca aagaaaagaaggcgggtggtgaacggcgacacccctcaccattcagctcagatcgc cacgggagaaacgacgcagcccgtagccctcggcgtcgacgcccgggtacaaagcacgt tggactttcggcctccacaaaaaggcgaactctatgcgtcggaatcgaacttcgtca ggacgtttccgagcttcttgcgggtcgcagtgctcgtcgcacgcccgtcgtatccgcaa gaccggtcaccgtgctccggttcgacaaccgcaagagagataaggatggctgcgc cgtccgtggagaaccgcatcgccgcacatgtctcgcgttgaggcggctcgtcgcgc at cctgccggtgacgaagatcatcgttgaacggcgctcttgacgtgcagcgtctgaaga atcccgcattcagggtgtcgtatcaacaggcggaacagctcggctctggaacgttc gtgaatacgtctttccgagacggacagagtgccaacactgccacggaaagccaag gacaagggttcaacgtccatcattgagacggcgccacggagggcgacgctccca acaatctgacacgctcgtcgagacgtgccacaaggcattccaccgcgcgagatcgag ttgagagtgaaacggccgcatcgtttcaggcggaaccttcatggcgctcgtcgtc gacgtttctcagcgggtcgtcggcgagggcactccgaaattgaggttcgaaacacattcg gttatctcagaaacacccgcatcacacggaaftgaaaaactcactgcgcggac gccctctgattgcggcaacctcaaggcgaagcggctcggatactactctttcaga aac agacccgccgccacaaccgacagatccaaagctcagatcctcaagggtggcgctcg caaacgccatcagacggcgtagaaatcaaggatccgtctttacgacaaggtcgttcc caggggcaaacggccttcatcttcggccgtcgttctcgggatacttcgacattgaacgtt ggacggggagcggatttccggcgcgtcagctacaagaaactcgtctcttggaaaaac gtcgcacctatttgattgagctcagacgcaactaa	gtcaataaccctgtctaaagacagaggctt gatagccattgtggccttcgagccttgattga ctgacctcagtcgccgtacggactacgtt gggtgggaatgtatagacaccgcggtatgt agatctagctcccgctcgtcggccggcg gttaaaagctcgtgagagtgagagcgggtgc cgtcgtgtaaaacccctccaacattggcgga aggatcattaccggtaaatagaccgcgagg aaactaacttgaga
1237-1238	Ga0256829_1001599 GENOME_ACESSION: 131447meta_G_2_SF_3300026518 GENOME_ID: 20587 CONTIG_ID: 1598 SOURCE: JGI DATE: 2019-01-27.	atggtaagagtgtaagcatagacggtaaaaactttatgcctactaatagacatggtaaagt aagaagattgctaaaggataataaagcaaaagtattatgcaaaaactctttacaattcaatt gctttatcagacagatgatgttactcaaaagataaactagggatagatacaggatataaat ttacaggattgcttttatagcaataataaagttttacaaaaaggaaactattgaactaagaca agatgtttcatcattatctttaagaagaacattaaagaaggtagaagaatagaaaaa caagatacagaaaacaaagatttctaatagaacaaagacctgacggatggttagccccat caacacaaaagtaataatcaatcatatcttgaattggatagatagatttactaaatcttccaaa ttaccaactaaaagtagaaatagcaaaatttgaatttcaaaaatataatctgatataga aaaagagttatcagcagggaatatgtatgatatgagaatataaaacagatttatttag caagagaacatggaactgtcattatgcaagaaaaagaagaatgataaatggcatatcac accatatagttccaaaatcaaaagggtggaagtataagggtagataatttagctttactgcat gaaagttgtcataaagaacttcatagaaaaatgataataaaaatctaaagcccaaac agtataaagatgcaacatttatgaatatacaaaaatggaactgtcattgatttaaaagcca aatatcaagataaagatttctttacgtttggttatatactaaaatagacagaatgatttagg attgaaaaaactcactacaacgatgcaatagcaatacaaaaagagatagtcattgaaaaa taaaagcaatcctatttatatacaacaagttagaagaaaaaaagggtttacatgaagcaa caccacgaaaaaggtagaagaataaaaaatacaactcaaaaagagaagcttcaaaaataca aaagggataacagtaaacataaaaaaatagctttatgataaagtaaaaaataatca aataggttatgtcagtgctttacaggtaaaatggtttatgtatagatatacaaggcaacta tatcaaacagtcataataa	aatcacctccccataatcagattatagagg aggcgtgttagcaatagcaagataaactaatg gttgactagactaagtgccctgcaagaggt caaacctacattactgtttacacacctttgg atgatacccaagtcgaagcgtctgtgctgata ctgtaaaacaaaggaaactcttagtcagat caggtaaaaagcagtaataagattgtcgatg ggttaattactccgaaggagctttacaagatt atatataaagacaaaagagatg
1239-1240	NODE_22_length_170560_cov_21.0839_ID_3751648 GENOME_ACESSION: mgm4799989.3 GENOME_ID: 238384 CONTIG_ID: 24 SOURCE: MG-RAST DATE: 2019-01-27.	atgggagaattactccgtaaggaggtgggttatattgtagctcgttattaacaggcacg ggaaaccgctaagtcggtgtaaaccgcagaaggcacggaagctattgaaagaacagaa ggctaaagtagttaaagaacaccattcaccattcagttgtgtacgggttcctcgtggataca gacaagatgtaattctcgtgtagatgccggaagtgaacaaataggcgtatcggcccaa ctgagaacagggaagtatttcagcagaagtcgaattgagaacagatattgaggtcgaac ttaaagtcaaacgcaactcctcattcagggaatgcagactttatgggcataatgcgatgggc ttctacgataaactaaaagaactgtattcaaatgtgtcctaacttttgggtacataactaaa cacgctcgcattaaacacaatctcgaaaagtcgaccggtataaatgccaggtgcattca ggaaccatcgacaaaaaagcgaattgtgtgactcttcaacaagtagcgaacaaa aacaggcaattgcacaaaacaaatccgaaaaagggtatccgcagagagaacaaagctc ctggatacgtacaggttatcagctgttcgataaggtgaatacttggtaggggaattgtttgt gtttggacgcagatcgtccgataatttgattgctgactcttgacggagaagtcgtcagtc gttctgccagtgtagtaactaaaactgtgataaagagcttcatcattattgttgaaagg agggaaagcgtcatttctcaccgcgtcaaaagcacgggtgtcctga	gtcaacgacccccggcaaaagccgtgagc tagtagctcttgtgactagcctaagtgtttcg agcactacgttaccggcaaatgtataggcac cgtagatgcgttcccaagcttaccgctcgc ggttagtggttaaacagttctgatgggtagg gacagtgctgctaaccgagaactgtcggat aacattggcgatgggagaattactccgtaag gaggtgggctatattgta
1241-1242	Viral metagenome NODE_4741_length_29778_cov_7.47405, whole genome shotgun sequence GENOME_ACESSION:	atgcttctcagtcgccaggtcttgaatccgcgtcagcagacaaccgtaagggcaagga cgaacggggacacggagctcgttgaaacatgataactggtatgcaacatggtcgagag gatcgtttcttagagttcacctctcgtggagacagtcacactgcaacttcggagcaatccg agggagcagaagcacggggcgttatagccccaccataagtagccgcatcagggtgtt cgtgcttgacaagcacgggaagccgctcatgccctgtcatcagccaggcgacgcaaat tacttaagtcggacgagcagagtcaccgactggctcgtcgtgattagagtggtgg	gtcaactaccccgctcgtgatggacagggctt gcggaaggtttccggaagccaggttgacc agaccgagatattagaagaaggtgagaca aaatatctacgtgttatagagtaagaggac caaccctgggagcttctcagtcaccagct cttgaaatccgcgtcagcagacaaccgtaag

SEQ ID NO:	Name	IscB sequence	ωRNAs
	SCL101.1 GENOME_ID: 279143 CONTIG_ID: 341 SOURCE: WGS DATE: 2019-06-06.	acagagaaatagagcagtggtgaggtcccgagtcaccataaagattgatccaggtcca agcacaccggcattgtctgiccagcgtagatgagccggtataactatggcgtggtg cgatccagctcgaccatcgccggcagcttattcacaaaggatggaacagcgggcca ctaccgccgtcgagcgcctggccaatctacgttatctgtctccacgggtggagaatc ccaccggcaggtgtccggcgtgctgggaaaaatgcaaaacacaaagggtattgc ggcccggtgcatacaaaagcgtgatttctgggataaacgatacgcagtagcgtctccac caagcctctccatcgagtcgaaccacacagcgtgggtaaatcgcctgtccagatgg gctccggtcagcgggtctggcgtggaactggtgcgtttgatacaggtaatgcagaac cccgaatttcaggtgttggatccagcagggcacccctagcaggttacgaggtacggga atactattggaaaaatgggacagaaagtgcgttattgtggccagcgtgtgtccgctt aacattgatcacgtgtgtctcgtatccagagggggactaatcggatatcaaatcttactct gtcctgcagaagctgcaacgaatccaaagggtctgaagacgcagaagtttgggtcaaac gacgtctgggagaacagaacggagagaagattgctcgtgaaggtaccgctcagggcaaa agcatcactcaaggatacagcagcagtttaattccaccgttgggcatatggagagaactt ctgaaaacgggactgccagttgaaacgggtacaggtgggcaaaccaagtggaaacagg aaacgattcggtatccccaaagtctcacaccctgtatgctctctgctgttgagaggtgataa tatcggtacagtgcccaatagcgtgtgtatcgtgtgtacagcagcagggcaagcatca gagcagcactagacaaatacgggttttctccgtcaaggttgcacgaaccaagactca tcacggcttgcgcaccggggattttgtcggggcgtgtggtcgaaggcgaagcacaag gggtgtccatgctggccgcgtaatcgcgcgtcatctggaagcgtcttgggtgaagtag atgggatttctgcaagaactgctcggttttacaacgggcagatggttatgataataaga aaggaggaaagcgtcttctcatttccgcccacggacgggggttccgagcgcgataa	ggcaaggacgaaacgggacacggacgtc gtgaaacatga
1243-1244	Ga0207433_10050431 GENOME_ACESSION: YelNatRA01SPAdes_SF _3300027863 GENOME_ID: 37535 CONTIG_ID: 50430 SOURCE: JGI DATE: 2019-01-27.	atgggtacgttgacacgagcagagaagaccaccgggggatgcttctcagtcacctgc tctggaacccgtgccagacagacaaccccggggtagggacgaaacggggcacggggc cggtacccatcccgggactggtgtgacacatggtcaggggagaccggggatgtgtg tgctgtgcaccccggcgtaaccgtcgaagaccgacgaaccggcgtgagacggagag cggggctgttaccggccctgtcgtgtaccctgtccgtcgcgtccgggtgttctgtc gggggtcagcgcgtgtgtgacccgtgcatccggctcgtcccgaggtgtgtcgtc cctgcggccggggccgggtggtgcacacaccccgttcgtatccgctgtatcagacag gagcggcagcagtcgtcaccacccgctcggcgtcaagatcgaccaggtcctccgc cacaccgcatgtgtcgtccgggtgcacccggagggccgcacccacggcttgttc cgtgacgtcgtatcggggccggcagatcgcgagggcgtgacagccggggccg gctaccggcgtcggcggcgcagcggcaacctgcgtaccgggcccgcgttggcgtg accgccacccggccgctgcgacgctgtggggcgaatgccatccacggccgctgt tctccgtccctgcgcagccgccaagaccggggatgggagcgagagatgccggc ttgccccgagcctcgcaccgggtcgcagggacgtgtcgtatggtggccgctgtgc ccgctggggcggcgtcgtcggcgtgtggaactgtccgcttcgacctccaagcc ctcgaggaccccggtatcgccgggtatcggtaccagcagggcacctcgcgggttac gagatccgcgagtagctgttggagaagtacagccgacttgcgtactgcgacaggac cggcgtgctccgtccaagtcgagcacgtccggcccgttccgctcgggtcggaccgg gtctcgaacctgtgtatcgctgtatccgtgcaaacgcgaaagactcccgtccgtc gaggagttcttgccgccaccccgaccgctgccaaggtcctcgcggggtcgcga agccccgtcgggacgcgaccgggtcaacgccaccgggtggccctgcaccgccgc ctccaggccatgttccccgaccgagtgagcgtcgggtccggtgggcgaccaaataca ccggaccggagccgctacccaagaccacacctgcagccctctgcgtcgggaag gaccacgcagtcactgtaccggcgagttgttatagctgtcgcgtcggccgag gcgtgtactcccgcacctccccgacgctacgtttccccgcctacaacgacccgc accaaaactgccacgggtacgccaccggcgacctcgtcagagcggccattccaaccg gcaagtacaccgtaccataccggagcagtgatgtgtccggaccagcggagcgttga cgtccgacccctaccggcgagtcggcgcaaccggcgacgtcagtcgtctccttaa cgagccgacggctggagatgtcccgccaggaagagacattcaatgactcatga	gtcaaccaccccgccctagcggatggggg ctagtggcgaagcgattcgtccaggtccc gtccgactaggcgtgttgaccagaccgag ccatcgaagaggcgaacagatggctac gttgcacacgagcgagaagaccacggg ggatgtctcctcagtcctctgtggaacc cgtgccagcagacaacccgggtaggga cgaaacggggcacggggccggtaccatc ccgggcactggtgtgacatggtcgagg ggagaccggggatgtgtgtcgtcgcac ccggcgctacccgtc

SEQ ID NO:	Name	IscB sequence	ωRNAs
1245-1246	Ga0207433_10084486 GENOME_ACESSION: YelNatPaOPGAMG01_F D_\$F_3300027863 GENOME_ID: 312031 CONTIG_ID: 84485 SOURCE: JGI DATE: 2020-04-07.	gtggaagactccacctgacagccggctcggctcaagctggaccggcagcaagacca cgggcatggcgtgggtcgggagtcggaggaggtctatcccgacacgggcgaggtgc agcgacggcgcatgtgctgaagctggccgatctcagcaccggcgacacgtgatccg cgaagcgctacgcagcgtgctgtttcggccgctggcgcggtgcgaacctcgt caccgcgacggcgctctgcaaccgcagcgccggctggctggctgctgcccacctct gcagcaccggcgccgacacgacgtggctgggtgcgaggttgacgctgctgccc aatcacagcgtgctgcaggaactgggtgcgttcgacatgcagttgattcagaaccgga gatcagcggcggttgaataccagatggcacgtccagggatagcaggtgctgtagtac ctgctagagaagtggaccgcacctgcgctactgccagccacaacgtgcctctaca ggtcgaaacatagttcctgtgcgaaggcgcgagccatcgtgtatccaacctcacgt ggcgtgaggaccttgcaacacagccaaggggacgcaggtatgctgcgcttctggc gcaagatccgaagcgcttggcgctgctggcgagggccaaggtccgcttcgcgat ggcgagcgtgataaattattacaaagctgagtttgacaacttacctgtatgacggagt gaagcaatgccctacggcgactattcttggctgtaattttggtagaggtgctcttcg ggcgagaatccatttctgctgggtagccgcttcggcgatctgcatcttcggcgaaaa gagccgatttggcgtgggtttagcttctgcgatcgggtgcttccgtattggatccgat ttgaaatatgggtcaaatcggcttaggatgcatttggctcagaaccgcttcggcgat tgggttggattcggagcggagtgcgtattcggcgatcgttgcgttctgggtacagaacc gcttgggtgagcgtgcatcttggcgggcggtgcgcttccggagaatccgctattggt ttcccgagctggaacagatgaccgaattgtatatgcgatcaacgtcgagggtggccc ctggatcgaaggttggagttttcagcggcgattgacgatttctgcaaaagttcgagtc aatggcgcggaactcaagtcgcgatacctatcgggtgcatatgaggtcgtcgaaagt gtgcccggaaaaggtggaatcgtga	taccataacgacgcatcaggagagtagggg actgtttgtgacggggcttttctgcgcaaac cctatgtggcgctcctctgattgcgaaacg aagtcgctgtgtgatcagttgcggaattt cagcatagcgtatgcgaaagtcgggttagt ggagcgatagaggtgcaatgggtcgcg caggtctgttctaggttcggcgaaacg ccttgcataatcacaatcgtcagccca
1247-1248	Cupriavidus basilensis strain CCUG 49340 contig_0000008, whole genome shotgun sequence GENOME_ACESSION: GCA_008801925.1_AS M880192v1_genomic GENOME_ID: 300461 CONTIG_ID: 7 SOURCE: NCBI Prokaryotes DATE: 2019-11-09.	gtggcagctcttctgttggataagagcgggaagccgctcatgctgcagcgagaagcg cgccagattgtgtcgtgcggccggcgcgctgcaccgctgaatcggcttcacga tccgctagtggatcggggcgagggcaacagcgtcttgaaccggcttcgcttaaac gatccggcgagcaaggggaaccggcgctggccatcacgcgcgaggtcaaggccggcg tgacgagacggctcgggatgttgcgctctgtctcgcggaaactgatccaccggcg cgagatcagcctgagcctgacgcagcgggccatgcgacgtgagcgctggcc gcaagactcggatctgtccgctcttcaataaacggcgcaagcccgtaggctgggttg cgccaagcctacaacaccgtgtgacagaccagatggcctgggtgcgcttccggcg gctcgcggcgattacagcgtcagcaccgaactgtccgctcgcacatgcagcgctgc aaaaccggagatctcgtgcgtcgaagtatacaagcgcaactggccggctacgagat ccgggagtagctctagagaaatggggccggcagtgccgctactgcgatcccgat acggcgctggaaatggagcatatcgtatcccgggcaaacgggtgctccaacagggtct ccaacctgacgctgctgctgcttgcgaatcgggaagaaggacgcgctccgctctct gatttctcgcgaaggaccggagcgtctcagacgcattctggccaaggcaaaacacac gcttcgagcgcagctgctggaatcggacccgggtgggttctgtaaacgcgctgaaga ccacggggctttcagtgagaccgggtcgggtgcagcaaacgaagtcaatcgcggacg attcgtatcgccaagacgcacgcgtggacggcgctgctgggtgagtgatcgactcag tttcgggatggcaaaagcccacgttggctatgaagtgcctggcgccggcgcttaccag cgacgcgactcaccgcatccgcttctcgcggctacatcgcgagcaagcgcg attcggcttcagaccggatgacatgctcggggcggttgcctcgtgggaagaaggctg gcagcgatattggcggggtggcgattcgcgccaccggctctcaacattcagaccgc gaagggtgctgcaggggtgtctcaccgctattgcaagatcatgcaacgcggcgacgg atatgttattcgtctgtagcgaatccacgaaggagagcgggcgacggcgacgctt cgctgcgcgctatccctaccggcctgaaggccgatgtttccgcgcgaactga	gtcaactacccgctctgaaggcgaggat tgtaaaggcaactgaaaaggcggttga ccaggcaaacgggtaaccaaccgctacgt tgtgaataggtgcaagaccacgctgggat gcttctcagtcaccaaccttgaagccccg gttcagacaagcgacagggtgaacacga aacggatcggggcgatcgccgttcgca acattccgaggggagacccccgaaagg gagggctcactagggcgtaagggaactt tttaggagagattgagtgga
1249-1250	Archaeon BMS3Bbin15 DNA, contig: NODE_38, whole genome shotgun sequence GENOME_ACESSION: GCA_002897955.1_AS M289795v1_genomic GENOME_ID: 236434 CONTIG_ID: 33 SOURCE: NCBI_Archaea DATE: 2019-01-27.	atgcagaagttatgtagagttcaagaacacacagagataactcctcaagttctctgctc tgtaatgtagggtattaaacagagaggaaactcctaattgtccctacaaagtactgctgat aacagctccgataggacttacactctgagcaggatggacaggaacttgcagttcctgtc ataaatatgcacaaacactttgatgcctacaacaccaaagaaaagcaagaattttttaa acaaaaaaagcaagcaaacactaaatcagctacatttatgaatatagtaagatggag actggtaaactctgaagtgtcatgacttatgctatattacaaacatgacaggatta agatgaattggaataatctcatgtaaatgatcttttgaattgctggggaacaaccag agcagaagtgtaccatataatgaccacacaaccagaaggaataataggatattcagac aatagaaaaaggttttaacctctatcagacgacagagatataaattacagcctggcgat ttagtaaaatataataatatttattcagagtaaaaggagtattcaattatgttaaatgggttag actaatatcttaactaaagcaataagatttattatgtaaatattaagaagtgagttggtta aatatgggaagggaatccaatttaaaaagatggaattag	gtcagtcacccagggctaaagcacgtctgt cccgtagggggcaggagcaattggtgatt aggaggcattatataatgcagaagttattgata gagttcaagaacacacagagaatactctc aagttctctgctctgaagtgtgattataaac agagaggaaactcctaattgtccctacaaa gtactggctgataacagctccgataggagct tacactctggcaggagtgagacgacttgc gag
1251-1252	Ga0209591_10003844 GENOME_ACESSION: MAT01SPAdes_\$F_3300 027850 GENOME_ID: 31815 CONTIG_ID:	atgaaagtttctgtagtcaacaagatggctgccttctgagccaccacggcccgcaaa gcgcgaatcttactgaaacagggaagcaagaattattgtcgcgaccattacgac caatatactacggcagtgctgatatgtccagcgataactctgggtattgatactggttat ctcagattgtttcagcgtgtgagtgtaaaagagagttgattggtggagaattaaagttt gcttccatctgtaagtgaacgcttaactgagcgaagaaataatcgcttgacgccgaat	gacaactccacacactgaacgggataccgt tgcagtggtggcttcttaagaatacaagaag cgtgtcaatagtctcagtaattatcaagcca gaattactacgttattcagaatataatggcac tttgggataggtttggcagccgagtttagt

263

SEQ ID NO:	Name	IscB sequence	ωRNAs
1259-1260	Ga0163147_10029708 GENOME_ACESSION: OliLakLV19MP6G1_2_ \$ F_3300020192 GENOME_ID: 33721 CONTIG_ID: 29707 SOURCE: JGI DATE: 2019-01-27.	atgcaacgtgtattgtgaccgacaaaaacaagcaaccgctcatgcctgtcatccggca tggcgagagagaagtgccccataaggggaaagcggggttaccgtacttttcttacc atcatctgaaagAACgtgagggtaccgtcttcgtcagaagcgccatttcggttttcaaac ggcgacatgtttaagccatcgtccctgctggcaagggttgggaacgcatagtggc gcgtggcttccgtgcgacaggcagtttcgacatcagcacctcttgggaaagcagcc ggcggttcacacgcactgcccgtgtgtacacattcgaggggttacactataagaag gacaaggcgctatccctccccacgcctaa	ataagcatccccacgcctaaaggcgggct tttagccctaatgcttaccgactcagccctt ggcagggtacgttagcgccgaactatagg cacttcgggagtacgcccagctcccgaactc tgcgggtgaacgattaaagagaatgcttggg agcaaacagtgctgttcgcgccaaccgtc gcataactttgtctaggcggtgactaccaggt ataaaaaagataaaaaaacagagaggc atgaagacgtggaagaaacaggaaaatac ccctcgaacgtggatgtagggggtgatta ggcgcttgcatacgttcgttaacgaaggt tcgctgccgaatccaccatttgcggcc aacggtcatttctcttattatcgaccgtttt cagcctgagcagttacgtctgggcaacttta ccccccgtgaagggaagaaagttgaggtaa ctcaaatgcaacgt
1261-1262	Ga0137385_10022959 GENOME_ACESSION: SagtaG_30_ \$F_3300012 359 GENOME_ID: 35417 CONTIG_ID: 22958 SOURCE: JGI DATE: 2019-01-27.	atggcgggttttctctagaccagcgccaccagccgctgatgccttgcagagaacgg gcacggctcttcttggcggcaacggggcgtcgtgcacggagagcgccctttgttattc gcttgaagaccggacgcgcgacagagcgggtgcagccggcgttgaactga tccaggcagcaagacgaccgggaggtccgtgacgtgtccaagaggagaggaagg cgaagtcacagcgcgcgaactgttccgagtgagattcgtcgtatctgctggagaa gtttggcgccgttgcgtctattgtggcgggggccagatggcctttgagatcgagcatgt ggttccaaagacggcggatcagaccgggtgagcaacctgtgctcagttgccat gagtgcaatgaggccaaagggaacagaccgcgcgagatgctggccacccccaggtg caagcgcaagccagaactccctgaaggatgcggcagcgtatgaatgcggcagcgttgc cgtggtggcaacgctgcggcaggtcggcctacccctgcacacgtggagcggcggac gcacgcgctggaatcgggatcgtttggatcgagaaggatcactgcctgcacgcctta tgtgtggcgagctagcaggggtcagctccctgcttgcacatcgtaccatccacaca caaggccggggcagccatcagcgcacgaatgtggacgtgtcggcttccgcggcgt accgaacgaggcagaagcgatcaggtttttgcacgggtgatctggtgcaagctgtg gttcctctcctcgaagacggcaggggtgcgtgtggcgcggttgcgtgctgcgaaaa gtggctcctccgttaggcaagatcgaatggatcaatgcgaagtacttggccttgatcca gaaggaaacgcttctcctccgcaataatcgggggcatccgagggcggaggtg ggtcaacatgcaaacactgtacaacgtatccagcacaggtggcgcgtgagggagccgg cgtgtaggagcgcgatttctcagtcctcaagggtgagcaaggagccctggaacagg aagatga	gtcaacgacccacgcgtgaacgcggggg cttggagtgaactgcaagccgtcgttgc agccttagcatcaatggtcagggcagcca ggtgcggatgctcgttggatgggtcaggg caccggccaatgccacgttggccgt ctccggtgtgtcgttaatgctgttcgggt atcgagcggtgtcgccagcagctcaagcc catttaccattggcgagagggttgcctggg tcagccggcccgagcgctcaccagccc cgtgaagggtgtcgcgaagagacagcaag aaagatcgtccaagatggcg
1263-1264	Ga0376681_0062128 GENOME_ACESSION: CalMonMetagenome_5_ \$ FD_ \$F_3300036805 GENOME_ID: 312378 CONTIG_ID: 62127 SOURCE: JGI DATE: 2020-04-07.	atgtcgaacgttttgcgtcgtacacaaacagcgcgcgtcaacccggttcatccggggc ggggccgtcttctcgaagaaaggcaaggcgccgtgtaccgaggttccctgtcacgt tgattctaacaggacgctgaaggagcgtcagctccagctgttgcgggtcaagatcagc ctggcagcaagacgagcgggatggtcgtcgtgaacgacacgagggaggtcgtgt ggggccgccaactgacgcatcgggggtgagcggatcagcgcgtctggacaagcgtc gcgcgtgtagaaggagcaggagcagcgaagacagataccgcaagccgcgttt cagaaccgcccagcgaaggcaccctgccccgtcttagaaagtcgggtgtgca acgtgctcagctgggtgtgctgtgatgacgtgtcggcgtcgtggtcattagccaag aactcgtacgggtcgtacacagggcgttagaaacccggatcagaggcgctgcatac caacaaggacgctcgtcgggtacgaagtgcgtgaatatgtcctgtcgaagtgaatcac caatgcgctactgtgacgacgctcagtcctcattggaactggatcatgtccaccagcg agccgacacggctccaaccggtgaacaacctgggtgcagcgtgtacatcctgcaatca gcgtaaaagcaatcaggatgtacgcgacttctcaagcatgagccggagcggttcac gcatcctggcgcatcctaaagcgccgtcgcgagacggcgagcagtcacatgacgcg ctggcgctctacggcggtcgaagggttggcctccagttgaatgcggcagcggc ggcagggaccaagtacaaccggcgggtgcgaggtgctcccaaacccactggtgtgac gcgtgtgtcgttggggcgagcagcccaagcgtcttgacgtgagaggcgtgaatccg cttcagatcatggcgacggggcatggctcagacaaatgtccgcatggacaagtacgg gttccctcgtaccggcccaagcaacacagcgcgtccagggattccagacggcgac ctcgtgcgcgtgtgtgaagaagtggcccgaagcagggaacctatgtgggtaaagtgg ccgtgcgtactcgtggcatctcaacatcaccagcgcgagggcgtgtgtaccgacattc atcatcgttttgcgtctcagcgcgcagcgtatggctatcctatcagcagcgaaagga ggcgcggttctcctcccgctga	gtcagagaccgtcccggtgaacgggacg ggcttgcgtgagacatgcagggccgctc tgaccagctcagccaccgcttcttctc gaagaggactgaggggctccgttgaag cgaatgcataggaacgtccgggtactcacc agcccggaacccctcgggactgcattaac aggacgagggagtaagccggtgtgtgtc catctgaacccgttcacaacattgacagg tgagcattaccctggaacaggagggccat ctggcgagaacaaaggacgatggatgt cgaac
1265-1266	human gut metagenome genome assembly, contig: NODE_7107_length_352 3_cov_7.120819, whole genome shotgun sequence	atggatattgcttattgtttggacaatcaaggcaaccgcttatgccaactaaaaggcttgg cggtgctggcacccttttacaggaggaagcagaagtgtgttataaacctttacca tccagctgaaatatgaagcactcacttttccaagatctatgtaggattgacctggac gaaccaatattggccttgcgttagtcaacggcgaaggtgaagtgtttatgcggctaattgc actacgcgttaaccaagaattcctaagctgatgacagaccgtcgcgaacatcgcaaaagc ctctcgtagaggacagcgattagcagtaaacgccttgcataagcgaataataactttaact	gtcaactactccatcctaaaatgagcctagct cattttcaggatggagcttgaacttgcgtttt agttgagcagagacaattacacagaag gattaaagtctcaaccacttgggggttttac caagccccctgccttggcgctatctgtgccg agtctatgtacttttactgtatattctgttagg

SEQ ID NO:	Name	IscB sequence	ωRNAs
	GENOME_ACESSION: OGQF01.1 GENOME_ID: 4615 CONTIG_ID: 7106 SOURCE: WGS DATE: 2019-01-27.	gagtttcctaatggccgtgaagcttcaggatacaaggatggcaatatgccagttaaagaca tcatcaacaaggaaagctgtttaataaccgcaagcgttctgctagatggcttactccaact gcaaacagctgtgtgctactcatataaatttggctaacacatcaacaagttatgccatt aagagctggactatggagtaacaagctcgtttatgcaactagatgatggctcgtttct aggagctgactttcaaatggctactctcaagcctatgctagagtgaggattatgtcttcg atatgcaagggggatgtgtgtcttcttggtaagccaatggataaaaaactaccattgc caccatctgaccctcaaaagtaaggaggcagcgataaggcatataaccgtattgtttat gcgactctgtcatggacagcttcatcaaatgaagcatggctagaagagaaggcaca cgcaaaaagtatgcaggaaacgtccataaataattgtctatgccctttatatagaaggctca gtgcagttatttggcgatgataatgtcatatattgacgggtttgatacagcccatctcgc gaatacatgcatatgcctaaagaccattttgcagatgtattgtatagcctgtattggcgct catafcgaaccaaagtagcacaacgacaagcatttgaattcatcaattccgttgcataaa ccgttcgctcatacacttcaactgaagaacttataggtacaaaggggaaatcgttgc aagaaccgtaccctcgttttgaagcaaaaggagactcttagtcagtggtgcatacaa atggccaaacaatatggagaggctaaagcccaaggatgtgtcacagcttgaagtgac aaagtctatgcgcgatacaatagcttaagagagctatgccagggtctatcttatatacc aaggtaagagcgtttatfaactggccagctatacaaaaggactatactatcgcgtttggc caaggcaagaaaaattccctgctaaggaaatgcaaaatcttaaggacacgacattagttt atgtgtaa	cattatgtttgacaagatatggcgattaaatt tattttatagtggatattgaatggatatt
1267-1268	human gut metagenome genome assembly, contig: NODE_287_length_7771 9_cov_9.382185, whole genome shotgun sequence GENOME_ACESSION: OVYC01.1 GENOME_ID: 11766 CONTIG_ID: 286 SOURCE: WGS DATE: 2019-01-27.	atgtgcaacacaggaaggatgtgcattgttattgatactgttcacagcatattgtgttg gattaccctgtggagacagagtgatttcaaggaatgaacatgcgtctgtctatggaga aacgttccctgttagcgacaagaacagtaaatgcggcggtggcgtagataccgtaaaatgac ggtatcgtaagccgaatggagacatcatagcaaacgcatgtattttgaaaaagcaaac gcagaggacagcattggagaaaagtaaaacaaccacacagtcaccaaaactaaaagg atggctgcgcgcgtcactgcagtcgaagtgtgacacatcatttccgtatcattgaccggtac ctgaaataccctccggtatctatcaagaatctgtgtatgcaggcggcggtgtgtat agcccggtatgaatgacccaacgatacatgtgagagatgtaccagcgtggaccgatgtatg atgcagaaaacctgagagcatatatttgcgaaggataactatcaatgcgcctgctgttaa ggcaaaagcggggactaccctgaaggcagatgggacaacagtaagctgtgtctcat catactctgttccgagccgaggggcaacagataatcccaatacatcatagtgatgc gaccactgccatacaaaaagcacatcaacccggcggtatctgtatttctggatgga aaacaataaaaagggtgcccgtggttaaggatgtacttttatgaacatctgagaaag cggttatttggcagatccacagccgcatttacatgggaatataacggcagctgacc gaaagcagttacgtcttccgaaagccatgcgaatgatgcgggtgctatcgtctgtttgg gaaagaggcgttctactgcaagaatctgtcagactatgcattacaacagatagcgaag tccaaacgctccctgcatgaagcaatcccaaggaaagggaaggaagacccaatcaga aggcagtcgtaataagaaaaatacaacacaggcggaatgatttaagtattggatagtg tacttgcgtatggaaaagcgtgttcatctgtatttactggaacaagtcctacttaattga taaaaatggctgttatatcacagcccgaaaaacataaacaatggccattatcaaaa ctgaagcgggtacatcccaatggttaactggctgcatggcataa	gtcaattaccaccgcttagatcttaagag atctgaagcgggggctgcaagagccattt gactagccctcagtgattttagtaagctgtgc cgggatgccggagggagatactgccagat gaactacgttatgagtgtcatgataccgcct gtgaacctagcagggactgttatccgggc aactgatcatggcggtgggaagccatacaga cccgaaggatcaagccctcatgacattggcg aaggaccttacgcagacatcggcgagg acggtttcattccctgtattatgaagaacaga aaggagactcttatgat
1269-1270	Ga0400264_0001320 GENOME_ACESSION: GliderLake422_FD_\$F_3 300037329 GENOME_ID: 311581 CONTIG_ID: 1319 SOURCE: JGI DATE: 2020-04-07.	atgtcggcttggtagtagacgaagcaaccactgatccgtgcaccgaaaagcg cgcgcgcttctgttggccgtgtgcggcggtgtgtgtgcgtgcgtatccgtttacaatc cggtgaaaaaacacctgggagcgtgtgattgcattcaaaacagtgcaaaaagccctta agagacgcctctgcggttagttcaaccgggtggcgctgtatcagacctgaagaaaat gggcctgaaggtgaagtggtactggcggtcgaagaccggcagtttcaacattcaga caaacagcgt	gtcaactaccccgcttaaatgacggagctt gcagagaacacctacatgccgggtgacc agggaagcgggacccccaccgcaccgtt tatcacaggtcgttaagaccaccgcgga tgcttccctagttggcgctctggaagactg ggatcacgctggcgaaggtaaacgctg aagggtccggcgccgcaagcgggagc cggtgatagacattcccaggggagagag gcttccgctccgtcacaaggcccgtagg gcattcattgaaggaaagtattgtc
1271-1272	TB_FS06_10DRAFT_10 00002 GENOME_ACESSION: UncGr_3300000233_\$F_ 3300000233 GENOME_ID: 36372 CONTIG_ID: 1 SOURCE: JGI DATE: 2019-01-27.	atgtctcagaactacgtcttctgtgtgatgcgaagcgtcaaccactatgccgtgcacgc caggcgctgcacgtgagttgcttaccagcaacgtgctcggcttactgtcgcacaccgt tcacgatcgttgaagatcgagtcgggtggcggtattcagccgctagatgttaaaactga tcctggctccaagacgaccggacttgccttggattatccacggcaagaacgggtcgcgag tcatctacgcgttcatctccatcatcagagtaacgcgattcaaaaagcactccttcagcga cgtagctatcgtagcaaccgtcgcgcgtcacacgcgtatcgcgtcgtcgttcttca atcgcactcgtcgtgacggttggctcccaactcgggtcagtcaggggtggatcacatcg tgacgtgggtgaagcgacttcagcgttccccactcaaatgtgttctactgaacatgtt cagttcgacctccaaagacttaaacatccagaacaacgcgcacgtcactccaggcactca gacaccttgggatggagatgcgtagtacacgcatggcacataactacacctgtcaat actgtcatgggtcgcagcggcgataaacgtctcactgggatcatactgcctcgtcgc gccacggctccgacagcctcgataatgccagcgttgcattgtacatgcaacagctca aatcgaattatcgtatagagcgttggatcgcgtcgtcgtcagacttccaacgcacaccatac cgcatccgtcgtatgcacccaaggtgcgtgcacagcgtcgcacgctacgcgatgcc gcccgtgtcaactcgtccaacacgcgttgaatcgtctcagcgttcgtccgatagt	gtcagtgagtgtatcatcagtcagggtcttcc ctgacctgactgatagatggatccttaccg atcctgtgatgaccagactcagtcacgcaaa tgactacgttcagcaggtcatgacaccgga ggatgtctgatccagttcctcgtctgtgt gtcattaaaaagatgagctcaacgtctgg tgtgacacacagcaagcgtgtgaacctt gtcaggaccactttacacgaagtgtgtga ggttaactcacagtggtttataggaacctt atgtctcagaactac

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gcgtacacctctcgtgggacacacatgaaccgaacgtcgttacgtgctcctaaggagc actgggttgatgccgcctgtgtgcacacgccacctcttgcctatcgaccaaccttaa gccccctaacgccatgccgtcgggtacgtgtcgagacgcatgggtgcgcaccgacaggt acggctccccctcaacacgtccgaagggttcatccaccttcgggtgtccgtaccgggtg acatcgtcacttacacctgatgtcagggtcgtcgaagggtgtcacggacacgtcacggt gacttcactcgtacatcgacaaggctctacgcgtgacataggtgggtgatgtcga catccgtcatcctcctccgatccgtctgttcaagaacagacgggtaccgttactact ga	
1273-1274	Ga0206350_10291578 GENOME_ACESSION: DieMetAnnotativ2_6_\$F _3300020080 GENOME_ID: 280098 CONTIG_ID: 1822 SOURCE: JGI DATE: 2019-06-13.	atgagcaaggtttttagtagataccaacaacagccattaaatcgtcaatccaggac gagcaagagtactacttgacacaggcggtcggcaatcttcaaaatgcatctttcagtat aatcctgaaggttgccatagaagatccattattgcagaggtgtgcacataaactgatccag gctcaaaaacgacaggcatagccatcatgaacgacgcaaccggagaggtgttttgcgtg cagaactctcccacagaggacgacaataatcatagatcagacgacggcggtgtgtgtg cgtagaggcagacgaaatagacacacacgctatcgccagcctgttggcagaatcgga aacgaccagaaggtgtgtgccaccatctctgcagagccgaatgccaatgtcatcatt gggtacagcggttgcgaaggtctgtcatctcacaacattagctagagcgtgtcaagtt gacatgcagttctagaataatccagagatcagcgattgagtatcaacaaggaaacta gcaggatcacgaagtttagggacttcttagagaagtggggacacaaatgcacctattg cggaaagcaagatatccaacttcagattgagcatataattccccgtccaagggtgttca agccggtatctgaatctctgccttgcctgcgagaagtgttaacagaccaagggaacgaa ggacatcaaggacttctcaaaaagaagccagaactgctcttaagatctttagtcaggc caagctccacttcagagctcgcagcagtcacacgcaagaattgcatctgag actagaagctctgggattaccaatcagtgccggtctggaggactcacaagtacaacc ggactagacgtgagcttcccaaacacactggatagatgtgcatgtgtggcaaggt actccagagtgatcgaatacagggtgtgttcccttagcaatcagggcctatgggac ggatcgagacaaatgtgtgatgaacaatacggcttccctcgcacagaccaagga gaaacatcccaaacacggatttaagacggcgatgtgtgaaagctatcgtgccacatca tctcaagcgtgctgtagcgcacattggggaatgtcagctaaggcgaagcgga	gtcaggggaccgatccccataagggggatg ggcttgcagtaatccttcggggtgtctccga gcttaccgttgcagactcagtagataacta gctacgttagaggcgaaataggtaccaga gggtgcgaagccagcccgatgctcaggt gtcaggttaaacagattacaagggttagt ccagtgctgtcaacgtcaaacgcttctaac attgtcagggctaccatcacctacgaaagta gaggctcctggagcaaaaaaagatgag caaa
1275-1276	Janthinobacterium sp. isolate palsa_998 73.20120500_P28.8_cont ig_1114, whole genome shotgun sequence GENOME_ACESSION: GCA_003133225.1_2012 0500_P28_genomic GENOME_ID: 159801 CONTIG_ID: 5 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	atgggtatttttacttgataagaaaacggcgccgtgatccgtgcagcgagaagcg ggcgccgaagctgtcgtattgccgcgtgtgcgggtgcacctgtgttcccgttcagtat cggcctagtgtgaccgtttctggaggacggagtcctccagccgctgctgttccgatga tccgggcagcgtcaccagcgccctagcgtttgcccgtcgaagaaggcgtatgaac gacatgagcgtatccgtgtgattaatgccccgttcacatcctgtgttggatggagt tgggtgaccgcggccagccatcaagaagtcgctgcacgacgttccagatgcggcg cgcccgccggcgaatctgcgtacgcgcgcggcggttcaataaccgtaccggcg aaggtatgctgtgcgcatctctgcagcgccggatgcacaccacctgtctgtggccgc cgccggcgccgtttggcgcaatcacgcacctgtgcgaagagcggggtgacttcgat atgaggaaagtgaataatccggagatctgtgcgctaagggtatcagcagggtacgctatt gggctatgagatgcgcgaatactgttggataaatttaaccgcagctgcgcgattgcgat gcaaccggcggtgcccgtgcagatggcgacatcgcagccgagggcagatggcggtcg aatcgcgtcagcgacctgacgctggcgtgcacccatgcaacaagaatgcacgcgatat ccgcgagtttcaaaaagaagaccgggtgcgcctgcccgaagatcctcgaagaagcgaag gcgcgctacgcgacgcgcgcagcggtcaatgccaccagggtggaccgggttcgagggc ctcaaaaagaccggccttccggcagaactgcgcggcgacacccaactggagccgc tcgctgttggccttcaaaaacgcgcgcccaaatgccgctgcgtggcattgtggg cggcataggcggcacgcgtgcgccgaccttgcgcgtcaaatgcaccggctgcggctc cgctgcaagacgcgctgaacaaatgattccggcggtgcgacctaccctccggcaa gggtatccacaaggccgcgtgtgcgttccgaatgacgggtaattcaacatccagaccg gcatcggccacgcgctgacgggttggggttctcacaagaacattgagtgatcattcagc gcgcagatggttatgtgtgtgcaaaaagcaattcaatcatcaatttaagagctcc gagcgtcgcggttgcgtcgtcgtccctcccgccgtgaacgcgggggtatctcggagc aacatctga	gtcaacgacccgcccataaggcgagctt gaaaggcgcaagcctgacgttgcaggttgc ccccggaaaagtccaagcaagcctgca cttgcaagtaggtacagagcgacggcggt atgcacgaaacggaccgttaaccaggcagc aaaatagcgttcccatagccgattacaaca tttccaagaatggcgccgctcctcaagg gcgcgaccgcagatcgggcgctgaagccc acggcggtcgcagattttgcggccggcacc aaccaggagtcgcatgggt
1277-1278	Streptomyces regalis strain NRRL 3151 P010contig20.2, whole genome shotgun sequence GENOME_ACESSION: GCA_001509475.1_AS M150947v1_genomic GENOME_ID: 96760 CONTIG_ID: 117 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	atgactacgtttcccgaggcgagcagaccaccaggctgtgtcttccagcagcctgct ctggaatcgttgggagcagacacctgggagcaggggacgaaacggctcacggacac cccgccggcgcgggcgccggcaggggaacatggcgaggggagaccgacggg caggggaccggagcctgtcggcgtcacgccaagcgcgctgccacggcagcgaagcg ggcagagaacgggagcggtgacgtctccccacatgagcagtcgtacgacggtggcgt ggggcgaagcagagttctgtactctccaaggaagggaaccgctcatgcccgtccacc ccggccgcggcggaactctcgccaaaggacggcgctgtcggccgacaaggtgc ccttaccatccgctcgaagaccgcagctgtggcgactcgggaagtcgacggcggtcag ttacgcatcgaccggctccaaaggcaccggcctcgcctcaccgacgagagaagaag aaaccagcgagacaggcaccacagtcagcattccggcgggagctgtctcagtcgaact ccaacatcggcgacagatccacaagtcagtcacacgcgcgccggctaccgaca ccgacgacgtccgcaactgcccgtaccgagcccgcagatcgaacaaccggaccaca ccccgaaagggtgctaccacctccagcgccaccgctgcgacaccaccttctcatgg	gtcagctgcacatagagaaggaggacca ctcgtccccgaccagaaggtgtcagtg accagccccagtcacatccacaggat caggagggtgccccgatgactacgttccc gcaggcgagcagaccaccaggtgtgtct tctcagcagcctgtctggaatcgtggga gcagacacccctgggagcaggggacgaaa cgggtcagggacaccccgccggcgcg ggggcgccggcagggaacatggcgagggg gagaccgacgggacgggacgggagcct gtcggcgctacgc

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ccagccgcctctgccgtacgacccgtcaccgagatacagtggaattcgtcgccttcg acaccacagcatgagcgcaggcagaccctcaacggcgtcagtagggcaggga ccactcgcggaaaccaactcgcgcctacctgcaagcccaatggaacaacgcctgcg cctactgcagcgcaccggcgctcccccctcaatcgaagcactcacacccgcagccg aggaggtcgaaccgcatctcaacctgtctcgcctgcgtccctgcaaccgggcca agggcagccaaccatcgaggtctctcgcgcacagaccgaccgctcacgaaat ccttcagcaggtcaagcaccctccatgacccgcccatgaacgcgacccgatgg cagctatcgaggcactgggcacccctggcagaccgggtgcacacctggacggcga ggcaccaggaaccgcaatgccatggggctgaccaagacgcacacctggacgct ctgtgcgtcgggcctctcgtacacgagagcggcgacgcgacgtacgttcccgagc aggctctcgtcgccaaagccaccggacggcgtcctacacccgacaccccgacc gctacgggtttccaggctgcggcgcccccgaaccaagcagcacttcgggtatgact ggagatctgtacgtgtgtcatgccatcggaagtgggcaggacactggaccggcc gcatctcggtcgagccagcggacagcagcctcagacccgatgggcagattcaa tgtctccactggaatctcgggtctgcaacggggcgatggctacggctacgacgc gatcggaatcgccacagtcacatctcgaaaaaccggttga	
1279-1280	Sulfolobillus thermosulfidooxidans strain CBAR-13 Scaffold_2_CBAR13, whole genome shotgun sequence GENOME_ACESSION: GCA_001280565.1_ASS. CBAR13.1_genomic GENOME_ID: 86729 CONTIG_ID: 2 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atggttttgtgttagacaacaatcaaaaaccgctgatccgtgttcgaaaaacggcgga gggtgtattggaacgtggtagagcgcgtattcacaagatgtatctcgcgtgtcatgctgc aaccaagaaaagggtccgcgaacggcgtgaggaaattcgggtatccgatatcccaatcca agcccgacaccccccaaggtatcgccgatgatgaacgcgactcgtgcgggtgttac gaccagctcaaggtacagaactttctgttggggagatgggggcacgcacgaaaat gcaacgtgattcagcacaacctcccgaagaacattatgatccctga	gtcaaccaccccgccctgaaggcgaggagc ttgtccaagctccatgggtgaccagtctcag ccaggaggtctgagacctcatcggtctcgt ctgttcgtcatgacacctgggtgtctctc cagctccaggctctgtcggccgattaaaa gcgagcttaggtatgtgaagcggtgtggc gggctgaacaagcggaatgaacattggcg aggagagactttttaagcgttaccagctcc gtaaggaggatattatattgtt
1281-1282	Ga0209615_100317 GENOME_ACESSION: 4B3metaGSPAdes_SF_3 300025075 GENOME_ID: 21011 CONTIG_ID: 316 SOURCE: JGI DATE: 2019-01-27.	ttgccaccggcgcaagcaaggccatctgtagtcacacatggtcaggggagcggcgcg cgcaaggcgcccgtcacaggccgtgaaggcgatccgggagcgccacagtcgta gcccgtgtgactcaagcgcgggtgtttggatagcaaaagggtgagagcgtatggtttt tgtttacgatcgaatgcttggcgctaatccgtgcagcgaaaaacggcgcgactgct gcttcccgtggcgcgcgctgtgcatcgctgctcccttgcgacccggctcgttga tcgcaagcaaaagcactgcgaattgcagccgtgcgcctcaagattgaccccgacgc aagagcaccggcatggcgttgggtgcgcgaggttgagcacgttagcgacccactggag aggttacacgagaagtgcagtgctcaactgtttgagttgggtgatcgtggtccagat cagcgaggcgctgacagcgagacgccagatcgccgtgcccgcagggggcgcaag accgtttatcgcgctgctgctgttctcaatcgaccaagcccccaaggctggcttgcgcc agcttgacgacatgagttgacacaacgctggatgggtcaagcgctacagcgctgggc tcccatcggtgactctctcaggagttgtgcgtttgatatacaaaagctcgaacccct gagattgctgggtgcgaataccagcaaggcagcttgcgggctacgaggtgcgcgagt acctgttgagagaagtggggcgcgcttgcgttactgtggcgcaaaagatgtccactg caaatcgagcacattcaaccccgagcgcggggcgagcaaatcggtgagcaacctc acgctggcttgcagtgctgcaaccaagccaaggggccaaggctgctgaggaattctt gtccaaaagccggatgctgctacggcgattctggctcaagccaagcgcccgttaaag gacccgcagccgtcaattcgaccgctggcgctgtttcaggcgctcaaggataccgg gctgcccgtgagcacgggtgagggggcaaacgaagtcaaccgcacgcgcctgcaa gtgcccagacgcacgcgttgatgccgatgctgttgacacgttgacgcgctgcgcg actggcgcaagccaacgctgaacatcaaatgcactggcgcgagctaccagcgca cgcgccctggacaagtatgcttccaaggggccacttgaccagaccaagcgcaattcac gggtttcaaacgggggatcgggtccaggcagtagtggttctggcaagaaggctgga cccatgtggggcgcggtgctgtgctcctcagtgagcgttcaacatcaagccggcag cggtttgttcaaggcatctcgtatcgccactgcacgctgttcagcgagtgatgggtac gggtatttcaatcaaccgatagcaaacagaggataaccagatgaagcctttttgaagt ggcgggcgagcaaatataaaataatcgatcgattctgaaaacacttctaaggaagac gattgattgagccatttgcgggacggggcggttttttaaacgtcattttgaagagtac tgatcgccgatgcaaacgaagatttaataatctttacaagcaagttcagacgaacggga aggattttgtgctatgcttcagccgtgttcacacctgagaacaacaccgaagccgattc tatgaactacggcgagaattcaatgcatgtacggaaaccgccaaggaagtcggttttct gtttattgaacaggcatttcttaacgggctgtgtcggtaactcaagagggaattcaa tgttcttttggcggtacaccaaacagcttcccaaatgttgcgatcttaactccatga aaagagcaagcgcgctgtttttgaggttgcagacttcaagggtcataatggaaaaggccga catcggtcgggtgtttattgcatccgcttatgcacattaacggcaaccggcaggtttta	gtcaatcaccctccctaaaggaggcgctt gaggagcaatcctgatggccgaggttga ccagaccgcgaaggagattttctaatca ccgttgattaaagaccgacagctaccaca gggatgcttctcagtcctgcttgaag tggcaactgcagacacgcttcgggtgaagca cgaaacgggttgcaccggcgcaagcaag gccccatctgtatccaacatggctgagggg agcgcgccgcaaggcgcccgtcacaag gccccgaaggcgatccgggagggcg acagtc

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gcagctacgccaaagacgggtcacactgaccgaccaacaggcggttagctgattgcgc aagaaactcattgctcgggtgttctgtgtagctattgcaatcacgatacggatttacacg gtcaatttactcggaggctaaaatcacatctttgatgtacagcgctttattcaagtgatgca agcaacaggaaftaaagcagccgaactactcgcactctatggttaa	
1283-1284	Ga0394881_0001061 GENOME_ACESSION: WSA_S1_MG_FD_SF_3 300037311 GENOME_ID: 312717 CONTIG_ID: 1060 SOURCE: JGI DATE: 2020-04-07.	gtgcaactagcccaacaagctcttaaacattggcggaagcgaaacattacgagaatcag aagagacgcaacaatgtctaatgttctgttctgtatgccgaaaaacgaccgctgacatc ctgtaccgaggtgtggcaaggtcactcttgaagcgtggaaaaggcggtattatcgtc ttatccattcacactcatgttgagcaaaacggtagcggcaagctcctaaccgctagaactca aattagatccagggtcaaaagggtgacagggtatgccatcaacaaggcaacaaggatctt tggagccgaattgcagcaccgaggacaccagatggttctctgcccgcgcatcaggcagtt ttgacatcgccacaaccaatggaagatggcgaggaaattagccacaataattgcaacca atccacaaaaaggacggttattcgtatgattctga	gtaaactgcccggttctatgctgacggcag gctagcgccaaccagcgggatacaaggcca caacatgaccagcctcagctcttcgaggact acgtttttgagctatgacccaagagtgccg tcgctagtcttggctctgctgctagtgttaa acatctgtagtgggttaaggaaagtcacact gccaacaagctcttaaacattggcggaag cgaacattacgagaatcagaagagacgca acaatgtctaat
1285-1286	Desulfovibrio sp. isolate SA7853_bin2 NODE_147_length_2377 6_cov_35.726614, whole genome shotgun sequence GENOME_ACESSION: GCA_009712225.1_AS M971222v1_genomic GENOME_ID: 315930 CONTIG_ID: 7 SOURCE: NCBI_Prokaryotes DATE: 2020-04-22.	atgtctcctcagtcgccggaacatcgagccggatggcagacacggcggttcggccac gaaacgctgccggcgagacagcgctagaacatcgggaggcataacgcccgtacc gcatagcccaacatgccggggagcgttccggaaaaggcagccggagccgtcac gggagaaggcgcaagccgactcccagcagggaagcgtgaagctcccaatcacaact atcccaagcaccgcgaagccgctgtctgtctcgcacaggcagggcaagcctttgatg cccaccatccggccaggcgaggaagctcatgaaagtcgggaggcgccgcatc aaacagatccgttaccatcaggctgatcgaccgtatccttgaagagtcggcgtccag ccgatccgctcaagatagatccggctcgaagaccacgggagtgctgtagtcggcga ggaagtcaaaacggacggcactaccatcgtccaccatgtattatccatcgaactcact caccgcgccagaagatccgaaatccatcaccgagcggaaaggctatcggccgctc ggcggtcagtgaaactgcgtacgggggcttacgttcgacaatcgaccaagccgga aggttggctgcccgtcgtcgttccctgtcgaacgctcatgctcggaaaaggcg gctgtccgggttggctcgtgtccgcccgcactgtcgcagcgggttcgcttcgacactca ggccatgcagaaccggagatctccggatcgagtagcagcaggcgaggttggccgg atacagatccgcgaatattgtcgaaggtggggccgaagtgccgtactgcgggg cagtgcatgttgccttgcagtcgagcatatccaccctaaggccaaggcggtcgaacc gcgtgtcgaacctgactcggcctgccaaccgtgcaacgaagcgaaggatcgatgcc cggtggggagttcctgcccaccgcccggacgtcgtgagcgtgtgctggccagggc aagaccccgttgcgcgacgacgcggcggtgaacgcccacgcgaacgccatctcttc gcgctgcgggagaccgattgccagtgaggctccgccactggcgagcaggaataac aaccgctcgcgttgcacccccaaaggcgactgtcggacgcggcctgcccggac aagtcgaccaggtgatcggtgggacatgcctgtgctgcatcaagcaactggacgg ggcgcatatcaacgaccaatgtttacgcaatgcttccacgcgctatctgacacg gagaagatggcggggcttccgcaccggggacatggtatgttccgatgttccgaacg gcaagaaaatggggcgccatcggcggtgtagcggtgcccggcagtggtgttcaa catccagaccaagcacgctgtcgtcaggggcatcaacgccaggcattgccgctgctgt cgcgggccgacggctacggttacggccgcatgcctcggccatcctcaggaggccg catgacaacgctggccaatattcccgggactcgccgtccgattcctcccgcggtga acgacggggcctcctcgaggaggttgggtga	gtcaatgatccctccctcaggaggggcctt gctgaggaggacgttggcagggccgtgttg accaggcgcaagttgtttagtttggcaaacg ttagaggacaactacgttgaattggcgcta aacaccgactccgggagctcctcagtcgcc ggacaatcgaaagccggatggcagacacg ggcggtcggccacgaaacgcgtccgggc agacagcggtcagaacatcgccggaggcat aacggcgatcacgcgcatagcccaacatgcc cgaggggagcgttccggaa
1287-1288	Ga0394874_0001047 GENOME_ACESSION: KP_S3_MG_FD_SF_330 0037309 GENOME_ID: 312525 CONTIG_ID: 1046 SOURCE: JGI DATE: 2020-04-07.	atgttacgagtagcagttttatccatcaggcaaacattatgccaacgaaagctagtc gtgctagacgttgggtgaagagggaagacacgggtgtgtgtaacacgaccttggatttt tcaaatccaactacttagatgcccagaactcagaacacgcaacctattgagttgggatt gaccccggtaaacattacacggcggtggcggtgagctgccaatttactttgtgttgg ctcatttacaactgccattcaagacgggtgcgggaacggatggagcaacgtgccatgatgc ggagaggcgtagaggcagacgaattaaccgcaagctaccttacaatcaacgcgctcat cgtaaaaacgatttgacaatcgtagaagtaaaaagtgtgctccagcattcgagccaac cgcaattagagttgcgagtagtagacgaattatcgtgatgttaccatcactacaatcgc ctacgaaatcattaaagcgcgtgtgataaaaggatttagtctgtgatgttggtagaaat ggcagttcgagaacttagaaatttattgccaagtcgagcaagttgaaggctggcagaccg caaacatcagacagcaattaggattgcacaacaaaaacattctaagggtgatgcaattc ctgctactcatctgcagatggtgtgtttaggtgcaagcggtttattcgtatggcatca cttctaccatcattggttggaaaggagcgttaactataactcctgcaccctttaccgta attcgccgtccccgggtatcgcgtgcaactccacctcatgttccaactaaagggtgga cgaggagaaaatagcggaactgtgacgcggcatggaatcaggaagggtgatttggc aaaacacctaaggtgaaattgggtattgagtggtgataccgaaaaagcttgcagtcga gtgacgctgactggcgtagattggggcggtttagccctaaaaatctcagttggtacacg gaaacacgggattaatcgtgttgcacaactaagagattgtactatctactggcatcgagcca gtag	gtcaataaccccgccctgaagaggacgg gctgtgaagaaccaaatcttagctgttgac tagcccgtttagcttaactctggcacgcactt ctgaatacttccccagttcggaacattgcaa ggctgttggtagtcgttgggttaagccaag tcacttggattaggtggcgaggggactta atactttactggagggttatccatgtttag a
1289-1290	Ga0376669_0038378 GENOME_ACESSION: CalSwaMetagenome_3_F D_SF_3300036830	atgccatgtacaccgccaagcacgccattgttcaaaagtgtgaagcaaaacaaa gcgtacaacacttgattgttctgtgcagctctgctacgaacaggagcccgacaacca gcccctagtgtccggtgtcgtacccagcgtcaaaagtgtgaaggtatgtgtcgttgcctc aagacaccgtgctcaatctgatgttgggaagcgctgatcatgtcaaaagacgtgttaaga	gtcaggaaacctatgctaaagccagggg cttgagaacggctagctcaacgtgaagttcc tgtctagcgggctctcttccgaggaggca gtagcgtatccggctaaggaaagcatgtcc

SEQ ID NO:	Name	IscB sequence	ωRNAs
	GENOME_ID: 312381 CONTIG_ID: 38377 SOURCE: JGI DATE: 2020-04-07.	cccgtaggactatgcgacgggctagacggcatcgcaaatggagaagaccacgaagcggtt tgacaatcgcccaaccgcaagaagagatcccaccatcgaccgaagtcggtgggaa gcaaaagcacgcgtgattgctcaactcaaaaagatcatgccactgactggtggtgta gaagatgtttgtgcagaacacgcaaaagggaatggtggaagtggaacggctcgtttt tccgatacaagtgggcgaagacatctctatcgcttcctcgtcgtgagatgggattacatgttc acttgaaaggagatcaacaacaaagagttcgagaaaaagtacaagctcaagaaagac gaagagcaaatcgaagcagtcctttgactcacatgccattgacgcttgggttatggcagc atcagtgagtggtctcctcaaccgacttgcacagactctctacatcgttgggtcag ctccacaggcgtcagatccatcgttgcagccctcaaaagggtggcgagagaaagccctg tcagggacccgatcccttaaaagggtgggcttcagtgga	gagcgccaccctaacctggaccactgcctg ccacgccgaagcagacgaccggaggggt ctatgccgtagagccatcgcaagggtca tcttaaccctagcagtgactaggagaagg aatcccttgaga
1291-1292	Ga0334819_000080 GENOME_ACESSION: 713E2X1metaG_FD_\$F_3300037088 GENOME_ID: 312158 CONTIG_ID: 79 SOURCE: JGI DATE: 2020-04-07.	atccccaccacaattggcggggcgaacgagcctaactgaagggaacgcgttacgac gtttcagcaaggcgctctctacatccgctgacgaacggctgacggggtcatacagg aagttgcgtcgggatagacccggctcgaagaagggaagcgttcacggtaaagtcgga tgccacactttcatcaatattcagcggcgctgacttgggtcaagaagcgggtgga aacccgcaggttctgagaaggtaagcgcagcgggaatacacttgcgtcagaac cggltcaaccgggacaatcggcggaattccgagtagaccagggtcgtatcgggtgga aggtgaggattgacacttgcctcttaagctctatccgatttcccgtttgtgtcgaagg attaagcgcgacaccacaaggcgacgcttgggaataagagttttagccactcgaagt ggcaagaagtggttcaaggttagagcttgagaagctcgtccgatcacactcaagcagg gtgggagacattccagatcggaacgcagcggggttgaanaagacgaagcgaagat gtcggagaggtttgacggcactgtgtgagacgttgatctggcgaacagttgggtcgg cgggaacatcaagcctgacaacgaaaagatgtggctggtaacgccgttcaattcagcc gtagacagcttcatgttcagaatcccgccgagcggtatccgaaagccctacggcgga acgcgagcctgggcttcaagcgggggaagctgggtacagcatcctaagtttgggacgt gctttgttggcgggtcgtcaaaaggcagagtcagcctgcacagcctttcaccggaaaa cgcttaaccagagtgccaaatcgaagatatggaatttgcattatgcaagctggaggg cgctcaattcctccgtggctaa	gtcaatgaccgtcgcctaaaggcagcggt tgtgccaaggccaagccaaggttgaagc acaactgaagcggagactaggatgtagta actgagcttctggcggtggcagccgaaa gcacaccgagaatgctcccaagtttccgt ccctcgggaagccagtcgggaaggggaa attgagaaagagaggttaactctt
1293-1294	Ga0172381_10002433 GENOME_ACESSION: Leawel6488metaG_2_\$F_3300014204 GENOME_ID: 31499 CONTIG_ID: 2432 SOURCE: JGI DATE: 2019-01-27.	atgaaaaaagtattcgtacttaacaaaaacggagaaccgcttatgccaaaccgaaaaatac ggtaaaagttagcgatttaataaagatgtaaaagctgaaatttttgcgatgaaccatttact attaatactatataatgatacaccagaattcatacaaccaattacacttggattgataccggt attcattcattggaatttcagctttaaacaattaaagaagaattgtttgtgtaattggaactt cgtaatgatatcaaatcatgatactgataaacaacagatgctgtagcttacgaaggatgag acttagatatcgtcgaagcgaattcgtatttagaactagaaaagataaattgttaccaccatc aactcaacacaaatttgatactatatacaacagattgataaactaattactagattgccaat aactaaaaattatagaaacggctaatttgaatttcagaaaattataaccatttaattacta gtgttgaaatcaacaaggtgaacaattagagatttggaaattatagagaataataattccata gagatgaccataaattgcaaaatcagattgtaaacataaagacgataaaaactataaacc ggtaataaacgtacatcatatataatagacgtaattggtgttcagataaaccaagtaattta ataacactttgtgaaaaatgtcatacagctaaaacatgaaaaaggattcttagtagattg gcaaaaaaacgggtttaaagttaaagggttttaagattctacatttatgaatattataaaacta gaattcctaatactacgagaattatatacccaatagaaattactaatacacacgggtatatt actaaagctaatagaattcataataaattgagaattctcaccataacgactgtttattatag ctggtggtactgacaaaataaataatcgaacttataatcttaagtgtgaacgtagaaac aatagagcattacaacatttagagatgctaataatattgattcaagagatgtgaaataaa aacgggatatgtataaattctgtagaacaacaaacgaataaatcactaatgtgaaatt tgagacaatatagaacaccaatattaaatccagatggaaccagaaaacaagttaccaaaag gtaataatagattagaaaaacgtgtataatttctgaaagggtgctaaaattaaaaactga aaattgtctgtgtaaacatttaagtttaataaaaaacaaattcttattagtgtggcagact aaccaaggtacatatactatagggcaaaaattctataccagcaaaagtattgtaagaaat aactaatagaaaaggataatagaaaaattataa	attaaaaacataccttatttttaaaattgagaat gaagtctcgtatgattacaattatgactcagg gaaacctcgttatttaagtcacacacctac ggtgctccactagcctgtagcaatgtgtttacc cgactaaacatctataccctatagcgagaa gtggaaggtaattacaagcttttataacatt gtcaaggtggaccacaggggttcgacctg atttatttataatttaaaataaaaacaaatga aaaaa
1295-1296	Archaeon IPdc08 DNA, NODE_102_48, whole genome shotgun sequence GENOME_ACESSION: GCA_006538205.1_AS M653820v1_genomic GENOME_ID: 289565 CONTIG_ID: 47 SOURCE: NCBI Prokaryotes DATE: 2019-07-22.	atgcagaagttattgtagagttaaagaacacatcaggagatgctcctcaagttccctgct ctgtaagttaggtattaaacagagataaaagctcagtggtccttacaagtagctggccaa taacagctccgatgaggacaacactctggcaagagtaggcaggacttgagagttcctgt gttaaatatgcatggaaaaccttaattgccatgagacctagaaggcgaaggtttccag aaacaagaaaaaggcagcagtagttcagcgaagctctttaccatacagtttaagacacctt caggcgaaacaaaacagcgttaaaattgggtatagatgctgtagattcaacaataggtctt agtggtataagtgacaagagtgaaattactctacagtgagtttaaccttgagaaaaagaattct aaactcattgagcagaaaaaattacagaaaaacacgaagaagtagattatggtataga aaaaggcataagccctctgtaagaagaagaggtataggttacaaccaatgatttagtg aaatataatgaagctctatgtaagtgaaaggtgctccatattggggaatatgtaacattg gtgaataaataagggaatatgtgtattatgttagaaaaatagagatggcctaattatgg taaaaggaatacaattctaa	gtcaatcaccaccattaaaaagggtggcct gcctcgtgagagaccagggttaacaggttga ttaggaggcattgaagaatgcagaagttatt ggttagagttaaagaacacatcaggagatgc tcctcaagttccctgctctgtaagtgaggtatt aaacagagataaaagtctagtgctccttac aaagtactggccaataacagctccgatgag gacaacactctggcaagagtaggcagga cttgagag
1297-1298	Ga0376652_0002900 GENOME_ACESSION: CalCamMetagenome_6	atgagcaacgtctttgtttatgataccgatctagacaactaatccagttcatccgggga agccagaatattgctctcatcaggaaaagcagcagatttaccggcttaccctttaccattg tgctcaaaaaggctgtagagacacccattgagcctttacgcatcaagatcgatccaggga	gtcagggaacctatgcttttccaagggtct tacaggaaatccgaaggttctgttgcttca cctgaccagactcagcggaggcccgctc

SEQ ID NO:	Name	IscB sequence	ωRNAs
	FD_3300036780 GENOME_ID: 312370 CONTIG_ID: 2899 SOURCE: JGI DATE: 2020-04-07.	gtaagaccaccggcatagccatcatcaacgatcagactggtgaagtcgtcacggctttg agctttctcacctggggaaacaatcaaacgacgccttgacaacggcgacggtgag aggaagcagacgcaaccgccataccagatatacgcaaccaagggtgcagaaggttaact ggtgtctacccgctgctatcacggccaatgggcacggctgcagaagaatgtgtctgac gacgagtagtgcttcacgtatggcaatcccaagcaatcggacgcaagtagcgctcaaa acggcgatatactgctcgtgctgtgtgacagaggcggaagcctattgggacctatgtggg acgtctcactgcacgagcgactggatcattcaatcacacacaaaaagggacaacgc aaggaaatggataccagattgcccgtgctctatcgcctggatgctacagttaccagaa aggagctgccattgtcttgcacatcagaatcagagcggtttctgcttccccaccat ccgttcagatagaagagtgccccgggtgtaagagccccagaaggaaggcgcttttct ccccctagcggcgccgcaaggaggtatccagagccataa	ctccgaagctacgttaggagcgaataggt acgttgggatcggtggccagctccccaccg ctacggtgaagggttaaacaggattacaag ggtagtgccaggtgctcttcacgttaaacgc tctataacctgtcaggccccaccattacctac gaaagtagaggtcacgatgagcaac
1299-1300	MLSBLCLC_10010460 GENOME_ACESSION: IMG_3300002220_SF_3 300002220 GENOME_ID: 27621 CONTIG_ID: 5321 SOURCE: JGI DATE: 2019-01-27.	atggcagattttgtattagtaaagatgcaaacattgatgccttgcaatgagaaaagag cacgattgtgttagcaaaaatcgtgcagttgtaaacgcgttaaacggtttgtaattcaac taaaagacaggagatcagcaactgtgtaattacaagacctagaagtcgaagattgaccca ggtagtaaaattacaggaatctgtcttctagaaacagagaatagtagtcaacgttttaaat ctatttgaattagaacacagaggacgactaatttctgcgaagctaaaagccagagcagca atgcgttagaaatagacgtaatagaacacacgctatcgcgaagctagatttctaaatcgtg ctaaacctaaagggttgtagcgcttctctgtatgcacgtgttgaactacgatgaattgg attaaccgacttattcgggtgcacctgttgaatcttttagctttagagcggtgttaagttgata gcagaaaatgcaaaatgctgataatcagggaacagaatacaacaaggacattgttcga gaaagaggtaatggaatacttttagaaaatggaaccataaattgcatgtattgctgattg aggatgttcttgaaaaagaacatctatccctcgcatgagaggcggtgataaccgcat atcattattggtttatcttgcaggaaatgtaacaggctaaaagcaatttaactttgaaagaa ttcttggtaacaaactgaattactaaagaaaataaaagctggtttaaagcgacctttaaaa gatgctgctgctgtaaatgcgactaggaataagctttggttgaatgttgaacacgggttta tctgtgaaaacgggaactggtgctcagacaaaatacaatagaaaattgtattccatacttaa gcacacgcgctagacgcggcatgctgagtgatattgtgtgttaagaattggactgtt ccacacttaggaataaaatgtaccggtagaggaagctacgcaagaacaattatgacaa atatgtgttttccaaagcatactttccaaaagaaaacgagtagttgttttcaaacaggaga tgttgaaaagctataatcaaaactaaaatacgtttaaaatagtcattggaaaattgttata agaacttctgttatttttcttagtaaaagaaaagaaagtttaccgtaaaatggcttaaat gtaaacttgcacaaagagcagacgggtaccattactataataaacaatacggttctatagt gttggtaataatttttaacaactggacaattatga	gtcaactaccctccctaaaggagagctt tgctctaagttgaccagcctcagctttaacta gactcgtttatctgagaataataggtactcc gggatgctcactcagctccggactcagct tgcataaataaacatctcagggttaaggat agtgtatgtgagatacaaacctcagaataa ccttggcgagggtgtttttacaagtcgcgtaa gcgaagagacaaaatcatggca
1301-1302	Ga0335055_0000121 GENOME_ACESSION: TYMME0r0171_2_SF_3 300034110 GENOME_ID: 281031 CONTIG_ID: 120 SOURCE: JGI DATE: 2019-06-13.	atggtagcagattatagatgaaacaagaacaccttaattgcttgcagaaaaagagcaa gaaaattattggaaggaaggaagcaaaccttaattggtttaaaggagactttgtattttt acagagaaaacctaatacagattacaacaacagatattgtataggtattgacctcagtagt aaaatgtcagccttaacgattaaagtcagaagctcagacaattaaaaacgttcaatattctgc cccaatttctgaaagaaaaggtagaataaaggtcagcttcaagaggtgtagaagaa agagaacacacacttatagaaaatgtagatttaacagaagggttaataaaagaataccac ctcaacaaaatcaagatggttgcaactttaacctaatttaagcatttttcaaaattatct atcaatctgtagcttttgaagacgtaaaagctaaaaaattaaaggagctaaaagatgga ataaaaacttttagcctttagaagtaggttaaaattggtttacgatgaagttgaaaaagtta tgcactttacttatatcaggatttgaacttacacttttagaagcagtttaggtttacacaag gtaaagataaaatgaaagttgcttttgaagtcattgttagacagttgggttttagtcaatca agtagttggaggacatatagaaccagataataaaaaataacatttttaaacctttaaacctt tacagaagacagctacatgaacattgtcctgctaaaaaaggaaattagagaaattatgga ggtagttatccttaggaataaaaaggaggtacttttagtaagcacaataagtggtgtttgtt tagttggaggacaagtaaggtagaatttagttacattgcttacaactaataaaaagattg acacaattagcaaaaaaagaaatttaaaaataaactaatttaaaatggaacagctcaat tcccactttatagaagaaaagagtaactttgagcagaattatga	gccaaactaccactctgaagagagtggtggt tgaaggaacacctaacgaggaacttaactta agtttaaacaggagaatagcttagacaagtic agaaatggacgtttgaagtcactaaaggag gtatgtcctcgaagggcagcaggttagttac aagatgcttcccaagtcagtaaccaataact atccgagtagcgaaggaataaaaacaag gtgcttacgcacattaaacagttaa
1303-1304	Ga0265338_10013913 GENOME_ACESSION: 42126metaG_2_SF_3300 028800 GENOME_ID: 279820 CONTIG_ID: 13912 SOURCE: JGI DATE: 2019-06-13.	atgtccaaagtgctggtggtggtacaaacaggcagccgcttgatcccgatccagg gcggggcgccaggtgtctcaccaggggcaaaaggcggtgctgcgcgtgtgccccttt acgttgatcttgaaaactgcttcccggaggcgcaacccccaccactgctctcaaaactg gaccaggaggacaaaacgaccgggctggtgctgctgaatgacgcgaccggtgaggtg gtcttcgcccgggaactccagatcggagtcaggaatcaaatcgccgctggttagcg gcgcgctgtcgcaggttccggcgtcagcgtcagacgcggtatcggcagccacgctgg cgcaatcgccggcgagcgggctggtgtctccctctcgtgagagccggctggcca atatccagacctgggtccggcgtttcgccgaattcggccgggttccgctgtggttgg aactgctgcaattcgtatccctctcgcatagaacaggccgtgtgagtgccgaggaatacc agccaggagcagctggcaggttctgagatccgtgctatctgctgaaaagtggcagcat acctgtgcttctcggcggtgctgcatctcccgctcggatccctatttcagctctccg ccagtagtggacaaccaagtagagaatctacccttggctgtggtgctggtgcccagg agccagaggcacgaggtatgctgttttctgctgaccaggccagatctgctcaagc gactccttgcgttgcagggttccgtcgaagatattccagcagtcattgcatgagcag ggtgtgtatgagcgtctcaatcctctgttggctctcgaatgtgagcagcgaggagcgc	gttgtgcactgcccagagcggtgggacc aggcgctcaccgtctgccaaagcgtgaag ttcatgaaccagaccaggttctgcggaact ccgttagaaacgaatagatagttacctacag gtgcgaggccagccggtagcactacggtg gatgattaaacaggtgaacggcagcagca agccaggtgcatccacaacaacggtticta acctgtcaggccaccattaccgcgcaa gcggaggcttctcggagcagaaggtata atccatgtcaag

SEQ ID NO:	Name	IscB sequence	ωRNAs
		<p>accgcgtggaatagcaccacgcgccagcttccaaaacgccctggcttgacgcggcca atgttggcgcaagtacccccggctgtgttcttcaagcaggttcgggtgctccagatcaa agcgacggggcatggagccgtcagatgtgcagtcagttagcgtatgggttcccgtcg ggcaccacctgccacgcgagcaagggaagccagtcgagggcaccgccagcgcca gaagcgctactttggatttcaaacggggatatgtccgcatcgttattccagtcggtcgg agtaaatacgcaggcacctatgtgacgcgcatcaggtccggcggaatggcaactttaa gtttcggctggcaggagtcagtatgttcgtcgaatatgggtattgccagatgatccag cgagcgatggctatcagatacttccagaacagagagcctgtcgaattcagggatgcca gtgtgctgtagagtcgctctgtcagcaccagagcagatag</p>	
1305-1306	Ga0063011_10003046 GENOME_ACESSION: IMG_3300003891_SF_3 300003891 GENOME_ID: 308497 CONTIG_ID: 1946 SOURCE: JGI DATE: 2020-04-07.	<p>atgatcaggaaattgtttaaagcaagcaggttaaggattacaggtggtggaccatcagga aaacctgtagtgttaattttctatacaaggagtttgacctgcaggacagtcacatcgag gtttatcactgtgttgacctggctatacaggatagggtttgcagtatgtgagattaagga taacagactgtgtatttctgtaaagtgatattggagacattatctgcataaagaaaag gatgacagaaagaggcctatcggagactaagagatatatcagtcgctataaaaaaga gaaggctatcagtaagacaggcgagaccacttaaaagtcaagaaaccaaggaacgt aaagtcaacacacaggaatttcagctacactcagacatgcagtagatattcacctgaacatg tatgcactgcttcaaaagctatgtccttgcctgtatatcagactgaaagggtcatagaagat aatacatttgacataaggcgaatgacatggggatctgcatatgtagagagtagtatcaaaaa tccccaggcagactaatacagaaaaagatgccttctgtgttggaacaggcgatttaca gaggcaccacttaatcccagggaacagaaacggcacggatgaagggaacacattatct atctatgtcgaactgtcatgaagacattcatgtcgccagggtatattcgtctgttaagggg gtaaccatgtggagacattaggaacaatgaatgcagtagcaggcatactgaatcagggt agcggactccgacatgttctcgtcgtacatggccttagcaagaaaaacactggggata gagaaattgcatggtaacgatgcagcagcagcggcgtggcatacagtcactgccaatga ggtagtgtgacactggcagttatcttctgaagaagaccccgaggcagataacaggggcaag ggtgcatgcagtcaggacagactttataagataggtggcagggtagtctaaaaatag aagaaaaaggacagaccagaaaggagaagtcattgcagagctatcgtggatgaaga aagagagtaacagtccttctgggttaagatactgaaccattaagggaatagtcct gcaatagctggagacatctgggtacacaaacaaactgttaaaagattttagcaaacgg cgttattgcaaggaaactatctactgcctgacctgaaaggaataacagggaacccctac attagccctgatcagtcaggagattgttgaaacaggggatagtaataatagggttaa</p>	<p>gtcaaccaccccgccctgacgggcagggg cttgattaggggggaacctggtgcaggtat ttccgggctaaccgtagtgggtgaacga gggctgtgctgatgcagaactggcgttca gggttacactccaggtcctgagagacagag gtgcactcctctgcaggagtaacccctcaa cctcacacagggtgcagcatcctcccat ggagacgaacatcctcatggatgccggga tcaggggcaccgtaagggtgtagaaagg agcag</p>
1307-1308	human gut metagenome genome assembly, contig: NODE_10685 length 34 28_cov_3.671806, whole genome shotgun sequence GENOME_ACESSION: UPOI01.1 GENOME_ID: 16693 CONTIG_ID: 10684 SOURCE: WGS DATE: 2019-01-27.	<p>ttggcgtatggaccagggggatgtcatcctcaaaataaacgaaaaacgggttttatagg atttcagaaccgatttattcattcaatatgaattataacagaaagagaacacttatgacg attttgtattaagtaaaagacgggaagcgactaatgcctaccataagaccggtaaaagtc gtcacctgttaaaagatgaaagctaaattataagtcaccagccattcacgatacagctt ttatagatacagaccggctatacacagccgattgaattctgtgaagatactgtgaccgcta tattggaattagatcaaaagtgaagaagcaggaatatgtcagtcaggaaggtgtgccactt gccgacgaaaaacagaacatgacgcacagaggaataatcggcgaacggagagg acgagcggttaccgaaaaaacagggtttgataatcgccaaagagaagggttggttg ctccatccatagagcataaaaaaggaagtcacacttagcatggatcgaaaaatattaaactga tgcccgatcactaatgcaacctttgaaactggcgagttgacaccagaacttcaggcta tcgaaggtggcagtatcttccagaagaaaagactaccagcaggggccaggtataata cggaacgctccgcgaagctgtgttcgtgctgacatacactgcgtattctgtggacg ttccgtaaaagacaagcagttcttcatgtacatcatgcattttactgaaaggcagcagtc gcagtcaggtaaacgagctggcaacggcctgtgaaaggtgccatacaccggcaacca tgcaaaagacggactattgtgggttatacccgggaacattcgttcaatggcgagc tgctacaatgaatatcctccgccaccggatcgtgaaggagcgaagcgttgatgagc atgtacaggtatccgttacctacggagcagatacgaagcaatgagacagttcatgggtct tgaaaaagatcatgtgaatgatgcctatgttattgggaatctacatccggaagaacgtgca gtttatatagatacaaaaaaacgacgtaacaaccgtatcctggaagatttctatgatgc cagctatgagattcgcgtgacggcagtatcaaaacggcaaggaaactttataatggccg taccaaacgtaaaaaaacaggaacactgagaacctgcataataaccgtagacggaag gttacaaacgggtcgtcgcaatgaagcgcaaaacagtttcaatgaacctggagacatc gtatccctgaatggggaactctgtgttcacggcacacacaccagtaagaagggtgct gtaaatgtacagttcacaagcctgccagaacggacgaaaatctgccgatctgaagaa g</p>	<p>gtcaataaccccgctgattcctcgggaaa tcagacaggcgtgtgaaagaagagaata gtaaccttctcattgttcataggcctgattga ttagcctaagcagaatgctacgttatccag gaatatactgtaggcaccgtcagatgctcca caagtgcgacgctctcgccgttatgattaaa catctctgagggttaggagaaggtgtgtaagc tttaaacctggagtaacattggcgtatgtg accacggggtgtcatccaaataaaccca aaaacgggttttatagatttcagaaccgatt tattcattcaatatgaattatacagaagaag agaacacttatgacg</p>
1309-1310	Ga0315284_10064466 GENOME_ACESSION: YL17G09_16_MG_2_SF _3300032053 GENOME_ID: 281635 CONTIG_ID: 64465 SOURCE: JGI DATE: 2019-06-13.	<p>atgatgccttgcctccggcaagagctaggaaagtactgcactgaaaaaggctaaagta gttaaaaaatgtccctttacgattcagttatcctatggatcttcaggctataagcaagaaatg gtcggatcatcgtgcctagctcttccactataggaatagccgcaagtcacaaatgagaagt gcctatattcttcagaataatagtagtaaggcaagacatttcaaaaagatggatggaaggaa aacatatagaaacaggcgtagaaatagaagacacagatagcgaagccagactttttaa tagaaaaacagaccgaaggcatactctacgggttaataagcaaaatagaatccatattagg gagattaatagattcaaaaacttttgcctgtttcaaatggtgtgtgtaaggcatcctaaa gtagaaggcatttttaaataatgtctcgtgatgaacgtgtgtaaaccttcaagacaac ttttgagagagatggttcaaatgttagacattgtaaaaaaaggagaaaagagcttcatgcc</p>	<p>gtggattacccgccctttagggcgggcctt cgagggtaaccttggggagtaattattcac catgactcgtgttttaggaaactaaaactcct tatttagcatgtagaacacctagaatgccgcc tcagttctaggctcttctgtgctctgtaaaca aggacgacagtcctagtcacacacgttgcg aaggctgaataacattgtcgagaggaggtg gttcaatgaaccacaacttaggaataggag ataaactactgtta</p>

SEQ ID NO:	Name	IscB sequence	ωRNAs
		caccacattatacaccgacaagcgggtggaagattctatagacaacttagttactttaga taaagattgtcatgtcaagatcatcaaggtaaattggaattgaaaataggagaacataaat tcagggggaaaatagatactgagttagctatactgagaaaaaatttagatttggcgaatagt gaaaatgtatatgatttcaagccaaggttaaagaaaactattaggggttgaaactctcc gagtaatgatgcttgcctattttggaagtcacgcctcaaacacttttttagttaaaatgtc cctaaaggagattaccaaaggactagagggtgagaagtcagcaaaaagtacctaaga ataaaattgttggtttaatcgttatgataaaatagatgataaaataatctattttatcaaaa tgagaatgtctatagattttaaagttgacggatattataaccgagagggtccgaaggtaa tagcaggggcgaacactaaaattattaggcaggagaaaatcatgtctcatagtgaacaaa catttcatccctaa	
1311-1312	Ga0194138_10000001 GENOME_ACESSION: TAR4v2_3_\$F_3300018 405 GENOME_ID: 35910 CONTIG_ID: 0 SOURCE: JGI DATE: 2019-01-27.	ctgaccagccttagtcccgtagggactacgttatcttagttatcatacccccgggtggccg gtccagcaccggcgctatgcaggcctctgtaaacatcacctagaggtaggtaagtcaac ctggagagccgattacggtgaagcttgataacattgacgaggaccacgtaactgccctac gggcacgtaaagggaacgatatattgtctaattccgtattcgtctaaatgacgataaaacg ccattaatgccgtgtaaaccggcagggcgagattattatataaaacaggtagggcagca gttttcagaaaagcagccctttactatcattatgaagaacagatagaaaacccgttttgag aacattgaggtaaatagacccccggtagtaaaaccaccgggtagcgtctagtaaatgaa tactaacgtggagctagatgtatctgggtgccaacctaagcaccgaggtcagcagat acgtgatcccttattgtcccgcgtcaaacggcagcggggcgacgcagtcgtaaactgc gctaccgtaaacctgatttttaaacggagttaaatctaaaagggtgggttggcccccagatt tacactcgtagatataacccctcacttgggttaaccgtcttattgcgtattcggcagtgac gtcggcagtggtgaactcgttaagttgatatgcaaaagatgaaaacccggagattgca ggtgaaggctatcagcaggaagctctttaaactatgaagttaaagagtatttactgtatcg atataaccacacctgtcagttatttctgtgtcatctaaagacccattttggagattgaaca tattgttctagagcatcaggtgtgtagcaatcgactgagtaacctgacacttctctgtcgga catgtaacggagagaaagggaactgtctttggctgactggaaggataaatgtaccacc gaaaaaaccttattgacatgaagcgattaaaggatttgcagtgaggacagaaacttaa gcccgcactgaaagatgcggctgccgtgaacgctactcgtatcgtctagtagctgagct ggactatctgggactacatcaggaacccctcaggaggacacattaccataataatcgtaa gcaacagggttaccgaaggaccattggatagatgctgtttatgctggagaaagggttgc tcacgtctatccctccggattacatcctttaaactataaaaggtaaaagttaataaccg ccagatgacaaagcctgacaaataggttccctcgaacgaagcctaaaggctctcccg agctttcggtttaagacaggagatatcgcggtctcccggttggcgtggggagggtgtgc atccggacctcgggtacttttaacgtatttaggaagagattgtagtggaaagcgttttagtcca tctctcttactgacggttacgaatattgtggaccaacgcacgtcaagatgattgttgataa gcatactgtaattaaagggtgtttgagagagtcataatcaaaaacttaattgacgccgatt cccttctcatagaagggtacattgaatcgtaaagtagttcatggtatccgtaggacagtaaa actagagaccatactaaactatattaa	gtcaataaccccccctaaaggggtaggctt gtaaaagcctaattgaccagccttagttctta acagaactcgttatcttagttatcatccctg gggtggccgggtccagccccaggctctatgc aggctctgtaaacagtcttagaggtgaagaac cgtcaacctggtagccgattacggtgaagct ttgataacattggcgagaccacttaactgct taaattgcagtaaacagggaacgat
1313-1314	Ga0070706_100065069 GENOME_ACESSION: KBSK5StandDraft_5_\$F _3300005467 GENOME_ID: 30504 CONTIG_ID: 65068 SOURCE: JGI DATE: 2019-01-27.	atgaagggtgttgcctatcacacccggaaagccgttgatgccaccacgccgaggcgt gcgcggatcttctcaaaacgtggcgcgacgcatggtgactcgtgagccattaccatt cagcttcgtttcagacgactacataccagccgggtaccgttaggtaggataccgggt tcccaaacggtcggcatagctgcaatcacgaaccaggaggtggctgttcaggcgagat gtgcctacgcacagatcagaaagccttccctcagcgagcagtagtcgcagaaac cgccgtctccgcaaaacgcgtaccgtgcggcgctggggcccaatcgtcgagaccct ttggctggctgctcctctctcgtcgtcgaagcgtgaagcgaccgtcaaaagggtacgttt tgtcgcgaggatcctgccattagtcacaaatcaacgtggaaattggcagctttgataccag caaatgcagaacccggaggtcctgggtttctaccagcaggccaccttcagggtat ctgctgcgcgagtatctgttcaaaaatggcagagacagtgctcctattgtcaagcgagc ggagttccattacagatagaacacctatccccaaatcacgaggagcagtgatcgcgcc agcaatctcgtgatcgctgcgatgcctgtaacaagcggaaaggaaacccggacggcag aagagtttgctatcccgaatccagggccaggtcgtgtgccactcaagatgctgcc acgtctcgtccatcaagactgccgttagatgctcttggccagcagttcggactagagcg ggtagcggtgagatattggg	gttaatgacctccccattcagtgaggaggct tgtgagagctggtcttctcaagccgattg accagactcagttcagagccagccaggcc gaggaagcgagcgtcttactgagcagtgaa gaggctacgatacaactgaggtttcagcttc ggaccttggtgtgctgccccagaccagg cactccaactcctggcctaaccagtggtgcg gttagcccgctggccaggggaaagtacc gcgtgtatctgtgtaggggagcagcttct actccgaaaggagccttaccgat
1315-1316	metagenome genome assembly, contig: NODE_3533_length_615 4_cov_2.708149, whole genome shotgun sequence GENOME_ACESSION: OGDP01.1 GENOME_ID: 4288 CONTIG_ID: 3532 SOURCE: WGS DATE: 2019-01-27.	ttgagagtatttattaaacacagaggtgaacccctatgcttgcctcccgcaaagg cacgcaagcttctaaaaagccggcatcgcaaaatcaaaagggtcaatccctttcaaatca attgaccggagccacagagagagcgtgcgaacccgggtgacattgggtgtgagtcggc agtaaaacgatcggttaagcgccagtagcgcaaacggcaactctatgtgagcgaagt ggctcttcgcacggatgtgagcgatttactgtccacgcagcgtgaatttcgtcgcccg cgaaatcgaaaaacacggtaccgtgcggcgtcgtatcgataaccgggttcgtctaaaaat aaaggctggttagccctagcgtgtgaaataagatccaagcgccatctcgcgcgtgca aaacgtgtgcccgttataccgtgacacaaatccgagtggaacggcggttttgatc caaaagattaaaaatccgacattgcagggtatattatcaaaagggtgatcaattgggg tttggacgtgctgagtagtgttaaacgtgacgggtcatgtctgccaactgcaaaagg taaatcgaaagataggatattatgtccatcatctcagagccgaaaaacagggggtga cgtcccggaactgtatcagcgtttgtaaacgtgccacaag	gtcaataaacccccgttaaaaggaggc ttgaaaagccttgattgactagccccagctg aagcattgtttcagactcgttattgggaat acataggcactcgtgtgtgcaatcctagct gaggacctcgcaccatcgattaaaagcac tgagggtgaagtgctggtgtgagttactta aacccctttgaacatggtcgaagattttac cgaccggcaacgctgtggaacaaaactt gaga

SEQ ID NO:	Name	IscB sequence	ωRNAs
1317-1318	Firmicutes bacterium isolate MGYG-HGUT-02286 genome assembly, contig: GUT_GENOME140265_8, whole genome shotgun sequence GENOME_ACESSION: GCA_902385415.1_UHGG_MGYG-HGUT-02286_genomic GENOME_ID: 303325 CONTIG_ID: 7 SOURCE: NCBI_Prokaryotes DATE: 2019-11-09.	atgtcagttgcggttatgagcaagacaggtatgcgattgatgccacaacgcgaatcgggcacgtaagctgctcaagtaaaaggaacgggtatagatatacaatccatttacgattcagctgacagaaagagagagcgggtgatgttcagacggtagaaccttgcgataccgggtatcttcattatggaacgtctgttaaattcggagaagcacgaatcttggcgtgcagattgatctgttaacagacgagaagcagaagcacgatgcgtgcaggtatgccgacagacagacagcagtaggagagataccgccagtcacgatccaataacgaaaacgtatgacggatgattgcacccgtctgagcataagaagatataccagacgatttcacgcatctgc aatgcaatgccgattaccaacattacattggaatgggaaatttgatacacaggtattgaa ggcactggaagaaaatagaccactgccacaggggaacggattatcagcatggagagaga tacggattgcaacgtccgtgaagcaggtttgtcacgtgacaaggtatagctgcatgtgct gtagaaaacgatagcagacgggtgcaattctgcacgtgcatataataatataccgaagtc gaggtggcacaatcgtatgccaaacctgctacagatgcgtagatgtcatactctctgc aaatcataagccaggtggtaggctgtacaactggaaccaaagattgcgtctcaaaagg tccaactcatgaccgccgtgcgattgggaaggtatgacaacaggtgaaatctaaagtttca gatatcaatcatcatatcagctatggagctgagacaaaggaacagcgcagggcgtttgata tcacaaatcgcacgtcaatgatgctctgttaattgggacagtttcatccgaagcatcgtca agaccagtgctataagaagaacgcaggaaacatcggtgtctggagaattctacagtc gcaagtatatcgacagtcgtgacggtaagaagagaagcggtcaggaaactgttcaatgg cagaatcaaccgtaatacagaagaagattccgagaatctacatcaatcacagacaacagaa agtatccgtggcagaagaacaataaggaaacaacattattcgtccagcctcacgataatt attgtgttaacaggcaacaattcgagacatccggatgtcactgtaattggggcaagagca atcctgctgccacagaaaatctgttctatcaaaaaacttagcatccacaataatgctgg tggatattcaaatcagattttacgtaa	gtcaataaccacgcccgagataaatctcg gacggggctgtgaaaacagatatgttttat aagtcacgtttttattatcgaaggcgttgacc gtaaggtcagattgattagcctcagttgtca gcggagccagtcgactgtttcatatagccag gaacactccagcacactggtaccggtaa aataggcactctgtggatgctcctcaagtc cagactctgcggcatgcgattaaacatctct gagggtaggagaagtgtgtatgcgttaaa ctaccggataacattggcgtaggagacaac cgactctgattcggactggctgccagtc cgaaggatagagaatctgtgcgagttgctt tgaggttcaagtggccggcacagaaggc gtaagccatccggttctccggaaatattca tgaaggagggcagcttatgtca
1319-1320	Ga0370526_00122 GENOME_ACESSION: 20181107_10A_FD_5F_3300035008 GENOME_ID: 307132 CONTIG_ID: 121 SOURCE: JGI DATE: 2020-04-07.	atgtctaactttgtattcgtgctgtatgacacaacaacacggtaagtcagtcacaggg caagcaagcgcgactattaaattacacaacagctgcgggtatcgccgctacccattcacaa tcattataaaatagcagttcccgaccgcacaatacagccctgtaaaactaaagattgatcca ggctctaaaactactgctggtgaggtgtgcaagatgataaagttatctgggtgctgagtt gacgcaccgagcagcaaatcaaggtacttagaatcgctgcgcgtattcgtcgca accgtcgcaatcgcaaaactgttaccgcaaacgcgtttttgaaaccgtactgaagaga aggatgggtgcccaaacgctagaaatctaggattgaaaatccttactgggttagggca ataatcgtatgtccgattacagggatctcgaagcgtgcaaaatcgaatctcaagc aatgcaaaatcccaagtttcggggaaagattaccagcaaggcgaactgcagggtatg aagtcagagagtagtactagagaagtgggttaggaagtgcgttactgtgtgctgaaa acgtgctcttttgaaatagaacacattcaccgcaaatcaaggagggtagcgtcgtttc taactcactctgttgcatttatgcaaccaaggccaagggttaatagagataaaagattt tctagccaaaagcctgacgtgctatccatgttttaagacaattgacatcctccccgtct ga	gctatccgctaaccgtggaacagggtggga caaccagccctttgttgacagatagccag ttctaagccttaaatggtactgtaagaagg tcaagacacctaaaagtgtttccagcttta gctctgtcgattaaacttaaacatcgtattta gctaaaggaggtgagttatcccaaacgct ttttaacattggcttcgcaagcgtctccctt tgggagagaaaataacctgagaaatca gaggacaaaatgtctaacttt
1321-1322	Nostoc commune HK-02 DNA, nearly complete genome GENOME_ACESSION: GCA_003990685.1_AS M399068v1_genomic GENOME_ID: 249565 CONTIG_ID: 0 SOURCE: NCBI_Prokaryotes DATE: 2019-05-30.	atgtccaaagtattgttaattgatactgaaaaagaccattagaccacatccattcagccca agctagacaaactaataaagaacggttaagcagcaggtttaccgctgtttccgttcaccatta ttctcaaaaaatcacattcagattatcagtaaacaccactgcgattaaagcttgatcctggag ctaaaacaacagggattatattgttcaacgatactactggggaagtattttgcagccga actaaacatagagcgttgcattagagatgcttttaacttatagagacagttgagaagta gcagaagaagccgtaaaactagataccgtaaaacagattctaaacagaatacgtccaa aagggtgttagcaccagcgttacaaagtcgtgttgaataatcaaaacatgggttaataa gctacgtaaaattgcacaaattgaagcaattagccaggaattagtagcgtttgatgtcaat taatgcgtaatccagatatacgaaggaaggagtagcaacaagggtattgactgtgtacga aatcagagaataattacttgaaaaatggagtagacaatgcgcttactgtgatgttaagac gttccactacagatagaacatattacccaagagccaaaggaggcagtaattcgaattacta acctgactctaaattgtgtgaaaaatgcaacacaagaagaattaaaggatatcaaaaggt cctaaaaaaagacaaaacaggtgtagcaaaaattggctcaacaaaaaaacatttaa ctgatgcagcagcagtaaacacaactcgttttcattgtagaggctctaaaaagcaactgg attaaacatagaacaggttctgtgtgattacaagaatttaactgtagccaacagaattctag aaaaaactcattgttagatgctgtgtgtgtgtaaatcaacgccaaattgattattaaag gtgttaaacattgttgattatagctaacggctatgtaggcaattatgtcgcacagata aatagcgttttccaaatcgtattgttcaagagctaaattcactttgttttcagactggaga cattgttaagctgtgtttaaactggtaagaagttggtaatatattgggagaattgcaac tctgtcaacagggagtttcaacatctctactaaaaacggattgttcaaggaaattagtcaca aatattgttcagtaattcacaanaagatggctactcatatacacaataa	gtcaaccaccaccactaatcggagtaccg atatagttgggagcttgaagagccgtattgt ccagcttaagtgttaattcactacgtttaaagt aaggttcaagacctaccagagaatgcgttag ctatgtcttctctagaaactaaaggttaaac aggtttaagggttaaggcagcttttgagat agttaccgacctgaacattgacgaagctaa cattaccggaaggaggacttatgtccaaa
1323-1324	Streptomyces lincolnensis strain LC-G chromosome, complete genome GENOME_ACESSION: GCA_003344445.1_AS M334444v1_genomic	gtgggagcagacacccccggagcagggagcaaacggctcagcgacatcccgccac cgcgcgggacacccggccgggaacatggcgaggggagaccgacggcgcggtgac gctccccgcacaaccccgcttacccggcgacggcgagcagagatcttctgtcc tftccaaggacggacacccgctcatgcccctgcccccgcggcgccgggaactcct cggcaaggagcgggcccgtgtcgtccgacagaatgcccttcccaattcggctcgtcga ccccgctccaaggggcaggggtctcgtccctcaccgagagaagaaggaagtcgacgc gcacggcgccgtcgtcaccgtcagacgcgggctcacagcactggaactccagaccg	gttcttgcaccccttcggaaaggaggtcact cgtcccccaacgcaggttccgagtgacca gcctccagtcacagatccccaagatcagg gaggtgccccgatgactcgttccccgcag gtgagtagaccaccagcgtgtgtcttctca gcagcctgctctggaatccgtggagcaga cccccccggagcaggacgaacggct

SEQ ID NO:	Name	IscB sequence	ωRNAs
	GENOME_ID: 168281 CONTIG_ID: 0 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	cggcgaccagaaccaccgatcatgcagcagcgtgccggtaccgacaccggcgccg ctccgccaactgccgtaccgagccctcggccagacaaccgaccgccccgagg atggctcccaccctccctgcgccaccgctgcgacaccactatcccaggcagcacgcc tctccgctacgccccgccaccgagatccacatcgaaacgctgccttcgacaccac gccatgagcggcgagaccctcaccgaagccgagtagcagcaagcacactcgcc ggaaccgaagcccgccctacccacgcccgaagtggaaccactcctgcctactgcg acgccaccggcgtagccctgaacatcgaccacctcaggccccgagccaaggagggt cgaaccgatctccagcaggtgagaacgccccctccatgacgcccgcctatgaaacg accgtttggcagctcacggaaccctcagcgacctggcaagccggtacacgctggt caggcgccgtacgaagtggaaaccgagcccatgggactggagaagaccacacc ctggacgcgctgtcggcgccgtgaatcatcagagcgccgacgccatgtgcggtt cccggggcaggtgtctcaacgtcaaaaccggtgcggatcgtacgcccaccact ccggaccgcttcgggtcccacggctccggccggccgaccaagcagcattcgggt acgtaccggggatcttctccggcccatgtaccgaccggaagtgggcagccatg gaccggcgcatctccgtccgcccagggggcaacacagcctaccacaccgagg gtcgtatcaacgtctccaccggaatctgagactgctgcaacggggcgacgatacggc tacagcaccggcaggaactctcagatcaacttctcaaaaaaccggttga	cacggacatcccgccaccgcgccgggacac cggccgggaacatggcgaggggagacc gacggcgcggtgacgtccccgcacaa cccggctacaccggcgcatcgccgca gcaga
1325-1326	Thermoplasmatales archaeon Gpl AMC_Cont1933, whole genome shotgun sequence GENOME_ACESSION: AADL01.1 GENOME_ID: 0 CONTIG_ID: 1823 SOURCE: WGS DATE: 2019-01-27.	atgtaccgtaggaaacaggagagacaactctgggtacaggaaccaagattatgaaccg tggaaacaaaagaaggatggctgaccaccagcatagatcacaaactacagactaca taagactgattgagaactgaaaaaataactacaatatccgatatattatgaggttcttc attcgtacagcagaagatgaagaacccctgagataagtggaatagatcaccagcaggga gaactcagggatagatcagagaataatctccttgagaagtccatcatatctgtgtat tgcagcaagaccggctgtcctctgaaatagagcatctcacaccagatcaagggcg ctctaatacagtagacaaccttgcctatctatgccatgaatgaaccagaagaagaacaa ccttaccgcagaagagttgcgataatccgcatctgagaagagtgcatgtatgcccftaag agatgcgggttcatgaacactgtaaggtggaactcactcagtaactggagcagatcat acatcgggtacatcacaaaggaatcgtattctttgggggttgataaagacccatgcca cgatgcatttgtattgtcggggaaccaaccaggccagggtattgccattacaatcaca cagagaaggaggacaaccgttctatccagacaacaggaagggttgcagccatcaa taagaaggaaaaatccccctacaacccgggtgatgtaggatcggggagagcaga tacgacgtgttgggttcatagttatggaattatgtaataaaggaatggtgaaagaa gatgaatataagcacg	gtcaactaccctccctgacggaaggggctt gttccgtgaggaaacagcgcaagagttgatta gggggcttgaagacatggtgaaaaacat gaaagagaagcagaagttagttggagaa atacatcacacctcggatgtccacaagt ccgtggcaactgtgatcatcattaaacaga gaggaaactcagtgatgtatttataaaaa cccccttacaacccccaatgtggacctaca gtcaggcatggctggacaggacctgagg tcccggttataacatgcgcacgaaccatt atgccgacaacaccgaaaggcacgaac actattgaaatctgaaagcaagggtgac agttcaaatcattcactatcagctccctat gcaaccggagagacaagcagccagtaat attgggtatagatccggatacaagcacatt ggattctctgcagtaacagaaaagaggaa ctgataatcccaactcaataagagaaatc
1327-1328	Microcoleus chthonoplastes PCC 7420 scf_1103659003802 genomic scaffold, whole genome shotgun sequence GENOME_ACESSION: GCA_000155555.1_AS M15555v1_genomic GENOME_ID: 40746 CONTIG_ID: 47 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	ttgagcaaggtttctgtttatgataccaacaagcaaccctcaatcaattcatccgggtag agccaggttactgtgtctcaagggaagccgccgtattccggcagatccctttaccata atcttaaaaaaggaggtaagtaatccagttgttcaaccattgcgagtaagattgaccag ggcgcaaaaccactgtagtagcaatagtaacgacaacactggcgaggtatttgggta gccgagttacagcaccgagggtttcaacttcgggatgtttgacttcacgtcgtcaactac gacggtcgcgacgcaaccggaaaaccggttacctcaaacctcattttttaaaccgcacc gtccgccatcatggtgccgccatcttaaatagtcgagtggtcaatatactgacctgggt aagcgtttgtcggcactgtgtcaaatcactgccatttctcaagaattgtcaggttgacac gcaaaagatgaacaagcagatattagcggagttggctaccagcaaggcaccctagca ggctacgagatacgagagtttctactcgaaaatggcattaggactgtcttactgtacg ctcgcggaaccaagcttgagattgagcatatcaagcttaaacccctactgattgtgccaa gggacatgggactcgccaagcgtgccgtaccgataagtggttttccacagcgttattg ctcccgaataaagtcacaaaggctttcaactggagatattgtgaaagcgatagttacc aaggcgcaaaaaagtcgggacttattgtggggcgtgtagctactcgtccctcggcagcttc aataittcagcgagggaacggattagttcagggcattaatcacaagattgttaatttatgcac aaaaaggatgggtattcatcgcaaataa	gtcaacaaccaccactgaccctagggggtg cagtgggggcttgaagaaagccctagtgtg accagctcagtagcttttctactcgtttttga gtcaagacacctcaaatcagcggcagttt gttctctgtcatctaacgttaaacaccattt gggaggttagtgcgttagtctaaacgctc aaaaaacattgactagggcaacattaccgc aaggaggctcattgtgagcaaa
1329-1330	Moorea bouillonii PNG strain PNG5-198 Ga0081470_101, whole genome shotgun sequence GENOME_ACESSION: GCA_001942495.1_AS M194249v1_genomic GENOME_ID: 114832 CONTIG_ID: 0 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgttacgagttccagttatttcgctgatggcaaacattgatgcctacaaaagcttctgc gctcgtcgttgcttaataaggtctgtatcgtttaccacaaatgatctaaatgttttctgtt cagttagttatcaaccatcagttatcagactcaggatcgtatgttattgacctggga aatctttttctggtatcgtctgtcagtaataaaagccactctttggacaggcgatttagtt gccgtacaaaaaagtcgctgtaggcatgatacaggcgaatgatcggaagaactcgc agaagtcgtcgaataatcgcaagattcctatttcaaaaggtctcatagacaaaagcgatt ctcgaacaggaggaaacaaaaggttcccccctcgatfcggcaaacggtcaactagaaa atcggttagtaagagaacttagtctttgtaccagtaagtgccattgtttcgaaatagca aggctaatgcaacaaaggtttttctcctgtgatgtgagacaatactggcgaatatctcag ttagaaaaaatagccccagtgactcagaacaaggctgggagacagcctgaagagag aagcgttagggctgatcaagacaaaatagacaaaagtcggcaaacggttaatactat gcagtagatgggattgctgtactcattttaccgacgcaaaaattattatcaccgc aatggaaaattgagtgccccaaaactgtgaagtaactaacgctgtgtttccgtaattag acgtgcccccataagtcgctcagttacacatttgcagttttcaaaaggtgcaaaacgctc gcaaatacggaggaacaaccactatgctatgttttcgcaagagagactacgttgaagcg	gtcaataaccccaacctcaacgcagtaggtt ggggctgtgtagacaaccaactaccatgatt gactagaccatttgagccgaattttgggacg aacgcttgatcactccccagtcgaagcctct tcaaaactgtgtgtcagttgtgttagacag gaaatctcaattcgggtgcgaggggagc agttattactcaaatcttttctcgtgag

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gtcaagccgggaaagcctaccgaggttgggtaagcgggtgagacagcaagacaagttt ccgtcagcgataatcaattggaaagaatcggaacattaccgcccgaagtcggactttt aaaacgttcgaccggcttaattgtaaccactaa	
1331-1332	human gut metagenome genome assembly, config: NODE_98_length_93064_cov_26.028363, whole genome shotgun sequence GENOME_ACESSION: OLRQ01.1 GENOME_ID: 7932 CONTIG_ID: 97 SOURCE: WGS DATE: 2019-01-27.	atgaactacgtttatgtattagataagaagggaagccattaatgccgactaccggttacg ggaaagtcgccggatgctaaagtcgtgacagggcaaacgggttcagaaccttcattta caatccaatttaattatataccaaaacaacggctgtcgagccagtgattgctcgatttgatc cagggcggaactaatatcggaatttctgctgtccgagaagatggaacctgtctcaatcttac acattgtgagacacggaacagagaaattccaagctgatgcagaagcggaaagaacac cggcagacatccagaagaggagaacgtcttctccggaagagactggcaaacgtcatg ggacagtaagtcaaaaccactggacgggatactccaggtgtgaaaagccgcttaaa gtcaaaagacatcattacacagagagtcgttttaaacacgtctgcttgcggaaggtggc tgacaccaacggcgacacacactgttacgaacccatgtgaacttattaaaacgcatgtgtca gatacttctgcccggtatgtttatggaagtttaaccggtttgctttatgcagctggatca accaaatgtaaaaaagtggtggattgttttcagcatgtccgctgttctgggacagcgg gataaaagaggtgtctctatccagcaggatggacgtctccctcttggtaaaacggac tattgaacattaccatcacatcgtaccaagagcataccgtggaaagtgaactgccgaaac atcgaggactttgtccggctgccacaattagtaatacttcacgggaagccctgaata aactaaagcacaagaaggtggatcaacaagaataacggcggaatttctgtttaaact aggtaattccatttctgtttatggagtttgagcagatgttccagacatacttatgcaacaa atggatggaatacaaaacagttccgtgaaaaatacaggcttgcataaagatcatagtgtag atgcttattgtatcgtcgaagtattctgccgaacccaaagtggagctccagaagatc gcatgagattgtccagtcaccgagcatgaccgtctcgtatcaaacggcagacggaaac ggacataattgccttggaaagaagggtgtgtgaagaatcgacataagcgtttcgaacaga aaacggattctctggaagatgtatgagaataaccggaatttttccaatctaactgtaa agaaatcgacgcttctataacgacttaagagagattctccagggcgatgatactttat cagggaaaaagatatatactaaagtggtggcgtcttctgtggaaacagataccgtatgctc ggatattggaatacaaaatttactgcatcgaagtgtacgattgtacaacagaacactggact tgtgtatgtaagttag	gtcatgaacaccgagcgacttctgctgctgtg gggcttggatgggcttacgccctgaaaaagc ttaagacggaaagctgcaaaagagtttatgag cagagcagtgacataccgacgcttcagggt gattccaagcctgaagctccgctataggtg tgcctatgtatgaaacacaacaggatgga ggaaagatccaagctcgtacttaccgtaaa ggagtaacttatgaactac
1333-1334	Ga0119967_10003500 GENOME_ACESSION: IMG_3300014060_SF_3 300014060 GENOME_ID: 278220 CONTIG_ID: 3499 SOURCE: JGI DATE: 2019-06-01.	ttgccgacatccccgaggggagcgaagccggaaggcttccgttacaaggcccccgaag gcagaagttgggagaaaaaatgctgtttttgatttgacagaagaaagaccgttga tgcctgtcctgagaagcgggcaagattgctgttgcacggggacggcgaaggatccat cgggttgacgctgttaccatttcgattgtgcaaccgcacgggttgagcaatccgtgtgagc ctgtacgcatcaagttggaccctggaagcaggacaacccgaattgctgtgtaagagac ggatgaacgggttaggcagacgcatgtgtcttctcggcaactgcacatcgtgggca ggatgacgggatgcgtcacacagcgcagggttttcagacgcaggcgacgcacggc aaacctgcgtatcgagcaaaagcgtttttataaccgcacgagacctgaaggtctggattgc gccatcttgcggcatcgtatcgagacgactgtttcttgggtgaacggctgcggcgctgg acaccggttgcagcgatcagccaggagctgggttcgttcgatacccaaaagcttcagaa cccggaaatctccggagttgaataccagaaggggacgcttttcggctcgaggtgcgcg aatactgttggagaagtggaaccgagcctgcgtattgtgtgtcgtcgaatgtctct ggaaatagagcatatccagcccaggagccgagggcgggtcggatcgggtatcgaaccta acgcttgcctgtgacaggtgcaacaagcggaaagggaataatccatcgaggagttcct ggcccacgatccgaagcgctgtgctcaggtcaaggcgacggccaaatcaccactcaa ggacgcagcggcggtcaacagcacacgggtggaacctgtggcgacagcttcgggatag gacctggaaatagaggtcggcacaggcgggtcggacgaagtgaaccggacccgtttg gggctacctaaaaaccatgccctgatgcggtttgtgtggggatgtgaaaagcgggtgcaa ggatggggcgtgcccgtgtttacaagttaagctgctgccaatggaggtatcagcgcac acgtcttgaccgttttgggttttccaagaggctgcctgatactgtaaaacgtgttcagggggt tcagatcggagacctgggtcaaggctgcgatacaagggtataaaggcaggaaattatgtt ggccgtgtgtgtgtcggggcttcgggaagtttcaacgtgcaaacggcacacggcggtgtgt gaggggatttcataccgctattgccgattgtcccaacgcgcagacggctatggctactttg tgcaacttttggaaatagcccttggaaagggaagattgcgagaggccgcctga	gtcaactaccccgctcgaaggcgaggct tggggaaacctgagccaggttgaccaggg ggagcggggaaatccgctacgttggcaaca ggctgttaagaccactccgggatgtctcct cagtcggcagactggaaggtcggatcat gctggcgaaaggtaaagcggcgaaggttc aagtcggcgaaagcgggagccgggttc cgacatccccgaggggagcgaagccgga aggcttccgttaacaggccctaagggcag aagttgggagaaaaaatggct
1335-1336	Polaromonas sp. JS666 plasmid 2, complete sequence GENOME_ACESSION: GCA_000013865.1_AS M1386v1_genomic GENOME_ID: 39632 CONTIG_ID: 2 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgccgtgcaccgagaagcgggccagcgtgctgtggcacgcggccggcgcgct gcacggctgggtgccactgtgatccggctgtgtgacgtcagggccgcgcctgcgactt tcagccgctacggtatcaagctcgtatccggcgagcaagaccaccggcgctgcgctgtgtg agggatgtggaatcggttgacgcgtccacgggtgaaattcatagttgggtatgcccgtatc aacctgctggtatgtgtccacggggcgccgagatataccgaagccttgacggctcgtc tcagatgcgccggcgccgaggagcaattgcgtgtcgcgacccaggttctgaacc gcggcaacaagaagtcggctggatcgccctccctcgcagcaccgggtggacacgg ccatggcctgggtcaggcgcatcaacgttggggcgccggtgagggccatcagctccga gctgggtgcctttgacatgcaggcactgcgaacccaagaatttccggctggagtagca gcaagggcacctgttgcgtatgagctgcgcgagttacgtgatgcggcgcaagcgtgtgc agggtttggcaccggcgatattgtgtgcgcgaaggtgccaaggcggtcaaggccg gggttcacgttggcggagtgccgtacgcgccagcggcatttcaacatccagaccat	gtcaatccccaccctaaaggatggggctt ggaactgaaaggagccaagccaggttgac cagggaaagcggtaatacccgctacgtt agcaacaggtcgttcagacacaccagcga tcttccatcagttcgtcgtctgaaggtctg atcatgctggacaaggttaaaggtccgaag gtcttgcagcagctcgaagggtgagc cgggtgctgacattcccaggggagcgttct ggagcgtcgtgagggcgacacagaccgtc accagggccgtgaagggtgacccgtgga agacggcctgctcagcagtttgcagttg aattattatgtcaggagttgtgtgga

SEQ ID NO:	Name	IscB sequence	ωRNAs
		caggcggaatctccggcgtgttcaggccatcagccataagcattgccgggtgacgc agcgcaacgacgggtatgggtatttctcaaccggccgatcacacaggcgtagca ggtatggcccaaggcatcggatgtcgcgcatccggcgctctacctcccccatgaatg gcagggttcacgcgcaatctga	
1337-1338	Fervidicola ferrireducens strain Y170 AN618_contig000036, whole genome shotgun sequence GENOME_ACESSION: GCA_001562425.1_AS M156242v1_genomic GENOME_ID: 99018 CONTIG_ID: 35 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	gtgcccggatcacagggtcctctcaagaggaggaaaaagatgtattgttactgcga caaatacggtagcccgccaccgacgaaaagggttcgacatgataaggaaactgagg aagaagggtaaagtaagaatcgttgggtggaggacgtccggcaagccaccggtagcag tttctctggacagggaaattgtattgtctccaaaacagttacaaaggaaactatcatagcgt cgaccgggataactacataggggtttgtgtgtcgaacaaaaatgggaaactgat cgcatactgccagggtgttttacaacacagaattcccgagattaaaggagttgatgacagaa agaagaagacaccgaaggacagagaatactattcgggtgcaaaaagagacattatc cgccagacatggtagggttttacaacaaatttaaggcgccaaagaaatagtaagggttagaga caagtcgaagccaccctcaaacatggtgtgaaacgcactcaatcttacaacaaaggctt ttaaggctcttcccccttccagcagacgaggtgtattgttcattgagggataacgtcttcgat gtcagggcaatgacatggggcaaacatagcggcacgggttaccagagatcacccagaa cgccagcagagaggaaatgcatccttgcgggacaaaagagaatttgcacaaacacca cctgatacagcggagagcggtgtgacggacgttcagagaaacctgtatacctatgca aagactccatgaagatgtacatgccggagggttatatccccataaaaggcatcaacg cagtgccgtgcaattggggacgatgaacgcaataatggggaactcgcgaaatgccgc ggctggagttgtacctgcatcgacgtggcacagggcg	gtcaactacccccgtttagaagtgagggt cttctgctgtaggggtgaactgctgacaggtc tctgacaggcagtagttgaacgggggtg tggcgacagagaacctgacgttctggagc acactccagctccgggagacagcggtga aatcccggcgcttacacgggtgctgccc gacacctcaggagagaccaccaccttcaa aaagggtccccgatcacagggtcctgcaa gaggagaaaa
1339-1340	uncultured Lachnospiraceae bacterium isolate RUG11561 genome assembly, contig: RUG11561_asm_92, whole genome shotgun sequence GENOME_ACESSION: GCA_902770305.1_Rum en_uncultured_genome_ RUG11561_genomic GENOME_ID: 345617 CONTIG_ID: 91 SOURCE: NCBI Prokaryotes DATE: 2020-04-22.	atgatcgagagtgtgtctgtttatcacgaccggcagaaaaactgatccgacaaatgcgt acaggcgcggaagcttttaaaaggcgacgcccggaaatgagtggtataaaccgatg ttcacgatccggcttcgacaccgggaggagagatgtgcagcctgtcgaatgtcctg cgatacaggctatcagatgtggtgtatctgtcaaatcagaaaagcatgagtagtgcac aatcagtagacatgctcaaatgagacagaaacccacaatgactgctgtaaataccg cagggcaagaagaaccgcaagcggtaccgcaaaccaagggtttgacagccgctctaa gaaaaacaaggaaatggcaccttccctgcgtcacccataatgtatcagcagactgtcaa agcacaccggaaatcgacatacatccagtagcgttacaagacctgtcttaccgacg tcttaccgacaggaaactgcatatcgcaaaagacacagcaaatgatgcttattgtatcggg cagtttaccgaaacaccgtaccagaggacagatattccagaacagcgccgcaataat cgtgtcttagcgaattctcagatcggaagtagctgacatgcgtgacggctccaaagaaat cagggttcgacgtttcatgcccagaaacaaacagatgcgaaagcagacgtactgacaaa aatgagcgtattatcgtggcatcaaatgtcaggcaggcagctgcatataagaaaacag cgttatgacattcaggcaggcgatgtgtcatatttcacgacaggatcgtgtgtaaaag gcacacaattgtga	gtcaactacccgcgactaaagtcgtagcgtt gtatagctacaagctcgttgattagcctga gtgcctgtgacactacgttaccgtgagctgc gtccggatgtccacaagtcggaccaact gggaactgtcgttaaaattgtcgtaggcgca ggcaactgtgcgcagctgttttaaccacg gataacattggcgatgtggaccacgtggct acggattttatccgcaagtagcagcatccgc tggctctcagactcgggaaacgtgaattat tgacttaagcaggaaggagcatcgcat gatcgacagagt
1341-1342	TPA asm: Thermotogaceae bacterium isolate Hy Vt-16 Ga0136651_10009681, whole genome shotgun sequence GENOME_ACESSION: GCA_011043375.1_AS M1104337v1_genomic GENOME_ID: 326424 CONTIG_ID: 149 SOURCE: NCBI Prokaryotes DATE: 2020-04-22.	atgggtgttccataatcatatggtgctgtgtgtagatgcaaaaggcaggtcgttattgcc aacacatccagcaagggaagaataattgaaacaaggtaaaagcaaggatatacaag atggtacctttcactattcaggtgaggtacgaagtcacaagaaccaaaggagaatttactat tggaaatagacgatggagcagagttcgtaggtattgtctgttaaaggtaagataaaattatat ttgcaactgatgtgagattgagacaggtatgcaaacgcaagattgacgaacgtaggtatg atagaagaacagacgaacagaaaactgttacaggcctcgagggttttgaacaga cgtcctcctaaaggatggtgcccgaagtgtcaaatatagaagaagatgtaattttaagag cggtagatgatttacgaaatacatgaatatgaccagagtcgtagtcgaattaggttgggtt gatagtcgaagcatggtggaagaataattaaagcatagagtatcagcaaccaga ctttgaaggcagaacagacgtgaacagggtgtgtgagagatggttataaatgtcaacat tgtgttacaactatcaaatgcagatacaccattatccaagaacaaagggtgtgtacgg acacattaacaatctgataacgtgtgtgcgtaagtgtcacaaggaaactacacgaaggca aatgggtattaaaaaagaacctaaacaatacaagtatcctgctattgtcaacaaggga atggtattgtacgacaaattagtgatagattcggcaagaataatgttaagtaacatttg gatgataacatcaaaagaacgaaagaattaggattgagaagatcattggttgatg catgttctatattgaatacaacaaaatagagacaagacctttttgattatccctaaacgtcg tagaaaagaataacaatccaacgaagaacacgaaactttcaagggttcaaacattg ggtatctgttaaaagctgttcgtgtgtgtaagaaaatgttaggtgttaattaggagtttaaga aacgaactttaacctcgtacatctttgatgataatttgagggttagttattcgaagcaaaa ttgtgtggagaccacaagggttgggttacaatttgaattga	tataacaattagtttttaggggtctgagccc gaggtgtccttggaaagtgggaagctcc cctgaggaaatgaactcagggtgtgaaaga ggtggaaatctaccacactgctcgtggtt gtccgtgaggacaggtaccttaaggtgtgtc tccagcttaagctaccggccatcgaggaa cgaaacgtagactgagatgtcgtatatacttt gtagtgcctacagagaccatgggacatggg cgaggggcaccataccccacaaggaggag gtacctttaaattgggtgttccataatcatat g
1343-1344	Ga0307376_10002476 GENOME_ACESSION: TR2MG_2_SF_33000315 78 GENOME_ID: 281014 CONTIG_ID: 2475 SOURCE: JGI DATE: 2019-06-13.	ttggcggaacccggcgttcggtagatgctggccaccttctcaagaaggaaaagccc gcaitttctgtacgaaccgtttactatccaatttactacgagaacacggaaattgtgccgg tagaaatgacgttggggtatgctgttcacgcatcctagtagcagtcgaggtgc atgacctgacgcggaatgtgcagatgtcttattcgaaagagattctgttcgcacaga catctctgcacagtgaagcgagatccgatgcccggggttaatcgcaaaacgggacta cccgttatcgtgaatgtcgtttcagaaccgcgttaaatctgtatgttcgggtcgggcaaaa atcacacgccgaagtcttggaaagagaaacgcaaaaacagtaaggggcacaacaa acgtttctaatgtctgacgagggttcaggaaatgtcagcatgaacgtccgagcgaaa	gtcaattaccccgaggtcaaggccgggggtc tggaggagaatcctgacagcttccgcttta gtacaggaacgaaagtccataagcgacagc ctaagtccaatcagactgcttcttcaaatat cagacagccggatgaacgcggcccttctc cgggtcagtagtccggcacaactttctgga ggctctgtaaaagcttttggggtggggcg gtcaacctcgggacacctgtcgttaccagg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		agggtaatcatgataccgacaagatcctaaccacgctgcagaacaaagtgtatacca tctgtagctgaggtcaagaattggtcagataatgctgtgacgaagatccgagtgagg tgaccgctttgacactcagaaaatggccaaaccgagattcagggtatggagatcaac agggggaccctgttgggttatgaagtaaaagaattttgctctataaatatgtcataattgtg gtactgcaaagggaagcaagacctgttttggagatcgagcatgtttataccgaagtc ccgggggtggcactgataagatacaactgtgtgatctcctgcagacgtgcaaccggg agaaggatcccgactgccgaggaatacgggttccccaatattcagaagaggcggc gaagttcagggttttctgctacagcgtcttactcaagctataaatgggcttatgcaagg agttaa	caactcgaagaacattggctgggaaatga acgggcaaaagcctgtattatccccgcga ggggaaccatggagatgtatttgcagtgc acaggccgaagaaggccgaacaagaa ccg
1345-1346	human gut metagenome genome assembly, contig: NODE_253_length_5728_8_cov_8.992662, whole genome shotgun sequence GENOME_ACESSION: OGYY01.1 GENOME_ID: 4822 CONTIG_ID: 252 SOURCE: WGS DATE: 2019-01-27.	ctggcgctggatgccggcagcaagcatgtggcctgtgacccacgactgaaaaggag gagctttatgccgagaggtcgaattaaggaccgatactgtggacctgctgtccaccgc aggcagaacaggcggaccgcaggaacaggctgcttaccgtctcgaaggttcgac aaccgtgtccactccaagaacaggatggctggcaccatcgattgaacagaagattca gagccatctgaaggcggtaaaggatgtacataaactgtccctatacaagatcaattgtc gaaactgccagctttgacatccaaagatcaagaacccggatattgccgggaagacta ccagggaaggcgaacagtaggcttttgaatgtccgagaatacgtcttatgctgggacg gacatgtatgccagaactgcaaggagatccggggacaagatcctcaatgtccaccac atagaagccgaaaaacggaggagacgccccgaataacctgacacctatgcaaca ctgtcacgataataccataaggagagatagagcttcacatfaaacggggcagctct caggggatgccgctttatgggagtcagtcgatggcgcttctatgccaggcttcaggaaca atatagtaatgtccatgatgacctacgggtatatacaaacgacgtccggatcagaacat ctggaaaaggatcactgctagacgccaggtgcataagt	tggatgtcaataactgccactgtgaagtgt ggggtctcaagcctttatgatcagactaagc gcttcgagcgtactgtaccctataatata ttaccgggtgaggttcagcctgtctgcc ctctaaggacagtgttaaacagttctgacg ggttagaacagtgttctgtttaaaccgag ggataacattgtcgaaggctgcctaactcta gcatacaggaggttttaaaacatg
1347-1348	uncultured Erysipelotrichaceae bacterium isolate RUG12334 genome assembly, contig: RUG12334_asm_183, whole genome shotgun sequence GENOME_ACESSION: GCA_902778065.1_Rum en_uncultured_genome_RUG12334_genomic GENOME_ID: 346312 CONTIG_ID: 182 SOURCE: NCBI Prokaryotes DATE: 2020-04-22.	atgcatcaagcgctgcagacattggcgatgggtgctacgtgcggtggacacggacaga acagtgtgtatctgtcgaatacccaacttatagaaaggaggcgtcagtcagtaaggt tttgtcatagggcggaatgtctcgggctgatgccaccaaccacgcttgcctgtatc ctcttccaaaggaaaaggcaatcgtttgaagaagagaccattcaaatcaaacgtttat aagaccggcagtgccacacagcactgtgaggtcgggacgacaccggttctcaacacat cgggtgtgacattgttgggacgatgtgttacttacaagaagagtggtgaacttcgctcc tcaatggagaacgctctcttatggaacgcgcgaaggatgcatcgaagtgggaagcgttat cgtaaaaccagataccgtcaccctaaagtcatgctcacaacaaacgagtcattatgaaa aagctgttaactggcagcgccataagacacactgcatcaagtggaagaatgaattcactt cctacagagaagacggatgggttagcaccagccttaacaaagggtgatcatcacc ggattatagactatatttccgcatgccaagaatgacacatctgaatcgcagctcgca cgggtttgatacacacaagctgtatgctgatgttactgtgagaaggatagatgattgtcc gccttatcaacaggaaaacctgaaggcttatgtttgcaagagacaactacacctgtcag atctcggagctaaagccggccgggtcagaagaagacggcagtcaccgtcaagctgacg ggcatcatatcaattatcgtatccacggcaccacaataatcttcaggatcttctgtgtt gcgacaaatgccatacagcgtgatcatcaaccaggcggaaagctctatcaactgatg atatccaataagaagatcgccgtgggtaagaagatgacaacagtttaattctgtgtc tcaattacgaagaatgtatccttcggcattcttcacatcggcaactataccaaggctaatc gcgatctgatggattgcttaagtcgcacgccaatgacgctgtggctatcgcgaatgcc aggacatcctgtttacgggaaatctgacctgtcgattgtaattggaacgataattacaag cagggtcgtcgcaaaaagcgttcttgcataagcgaacccccgaaaaggccgtaagac tcttaaccgcacagcaaaagcgttattccaagaatacgaagcgtgtgaaaggctatgacct atatgcaaaagtacatgatggaacactgggtacatccggctttaccggtacaagc gccttatccaggattgggaaggcggttatcatcagaatgcagaacaaaacataattgaac atttgcctatcaattacacaaactggataacggccacaattggctgtcccgtatccccc ctccctatag	gtcattaccaccacttagaacctgatgtt cctgaagtgggggctgtgaaaacaagtcg gttgactagcctgagatcttcgagatctacgt ttgcagattatgactctaccgatgtgaccc aagtcggatgagtcggagcaacgaatgaa ggagatgggttccctgaatcgactcgatgca tcaagcgtcgacacattggcgatgggtgc tacgtcgggtggacacggacagaacaggtgt tgtatgtcgcgaatcaccacattatagaaa aggaggcgtcagtcagtaag
1349-1350	Ga0265594_1010503 GENOME_ACESSION: saktaG_20_\$F_33000281_93 GENOME_ID: 37738 CONTIG_ID: 10502 SOURCE: JGI DATE: 2019-01-27.	atggtagttttgtactcaacaaacatggttaaacctttgatgcttgcactcctgcaagca aagaagtgttaagaacaaagagcaaaagtagtaagaagactccattcactattaat tacttggttggtctagtgtgttatgttcaaaaagtaacagtaggaatggactgagcagtaag aaggtagagtagcagcaatctataaacaaggtttgtatcaatcagagataactttgag aaacgataattgtaagaagatgaacaaagagcactctatctgtagaacaagaagagctag aaaacttaggtatcagaaccaaagatttcaacagaggtaaaaaaggttttctgcacctt caatgcttctaagattaattctcatttaaaagagaagaattgttgaagattttaccagta ttgaaatggataatgaacagcaagttttgatattcataaattacaagccaactgtaaaag gtaaaaggtaattgtaaaaggaaatttaaggattatataatgtaaaagcaatattttgcatcga gacaattacaatgtcaaaaatgtaaaaaggtaatttaaaattacatgttcatcatattttt tcgttctgaaggagggtacaaacgaatcaaaaatcttatgtttatgtgatgaatgtcatga taaattacataaagga	gtgaactccccacccttaaaaggtggaggct tctaattatcattaaggagtcgtgttcaccaga acactaacaggagaaaagtaatacaatgta gtaaacgataggaaaagaaattaggtaacctg gaattgcacctcagtttcaagcgtctacggata cagattaaacaactttatagggtataaagtag tgtttgatcgtaaaacctttccatatctgttcg agagggaatcggaattcta
1351-1352	Ga0310786_10003093 GENOME_ACESSION: AusTroGonzalMGv2_3_\$F_3300031998 GENOME_ID: 279931	gtgctgcgcgttaaacctgtataacatcgccgatgtggaccacagagcgttaggctctg cattatccgttaaatcgatcaaaaagacgaaggggagcgagaatcatgacaataattcat catcgcaaaaacgggtgaacgggtgatgcctactgtcagatgcggacgtgtccgccatct gttacaggatggcaagcaatcatcgtaagcatgacccgttaccgttacaattaacgtac ggaagcaatgaatcgtcaggaattatgagctgtgcataagctccggttacgaacacgca	gtcagtaaccaaccactacgtagtgggctt gtatgggactacgagcagtgattagccca agtgcacagcactcgttatagaagaatata atagtcactccgggaggtccacaagtcgg gactctgaggtacagtattaaactatctgag

SEQ ID NO:	Name	IscB sequence	ωRNAs
	CONTIG_ID: 3092 SOURCE: JGI DATE: 2019-06-13.	gggtgagcatgaagtcgacgacgggaatacgtttccgacagctacgatctttcgcacgaaacacgaggcagcagatcaacgtcgtaccgcagaacaagaagaaccgc aaacgctatcgcaaacccgtttcgataaccgcaaacaccgaaaggctggctgcacc ctctatacaacacaagaagaacgcaggtaaggtgatagaacgtttcgtatcagtcgc accgtaacatccataattgttgagggtgggacgttcgacgtccgtactaaaagcgata gagggaagggtacttctatccggagggtacggaattaccagcggggcgaaagatacaggt cgctacgtgagagaagccgtatttcaaagagacggatacaaatgctgttctcggaag aggatcaaaagaaacgcgacgtctgtacgacacatgacacttactgaaaggcaga catgctgacgtgtgagagagctggccgtttgtgtgagagatgcatagccggtacac caccatcgaggcggaagcgtttggggatagagccaaagggttccaaacttgagggtgc atcttttatgaacgcggtacgatgggcaatcatatgaggtgaaagccaaagttttcgtgat acgtacactttacatacggagtaacgactaagctgaaagccagaccctgaggatagaa aagttcacatcaacgacgcttatgctatggggaataaataccagaaaaaggcgaggt ttcgagcattacgcaaaagacgagaacaatcgtttgttgaggtgtctacgaccca aatacattgattgctgcgacgggaaaactaaagcaggcaaatccctcgatgcgaacgt actaacggaagagccgagaacaacgaaaagagccttcgcaagtacagagggtcaa aagatatcaagggaagaagaagtatacgaaaaccattacgatacaaccgggaa cactctgtcgtatgaacacgagcagtatgtagcaaaagggttgcactgaacgggaaaag cgtaagtcttacgaatggcaaaagtgtcactgtttcgaagccagagtataaatacgcac ctgggtggatgcaaaaaaacataacgaaaggatgcgccattctccaccttttacgaag tggccttactggcgagtag	ggtaggagaagtgtgctgcgctttaaacc ttgtataacatcgccgatgtggaccacagag cgtaggctctgcatatccgttaaatcgat caaaaagacgaaaggggacgcagaatcat gaca
1353-1354	human gut metagenome genome assembly, contig: NODE_122_length_8197 1_cov_5.420492, whole genome shotgun sequence GENOME_ACESSION: OMAC01.1 GENOME_ID: 8152 CONTIG_ID: 121 SOURCE: WGS DATE: 2019-01-27.	atgccgaacaaagtgtatgtcatcaataaacacggcgctcgtgattccctgctccca gcaaaagcccgctcatctctggtatgacggaaaagcctaaggcaagaagagaacgctgt taccattcagttgattttacgggacagcggatatactaggatgtcatctcggcgtgat gcagggaagtaaaactattgattatcgcaagatcacagagaagaactcttgacagc aatgtaacccacgcaatgatgtgtagattacttctactgcagagaattccgaagaac cagaagaagccgtaaaacgcggcgtatcgtgaagcgggttcaacaaccgtgttcacgc aagcataaagggtggttagcgccatctgtagaggtcaaaatcaagaacatattacgtc attaagagagctgttaacattctcccatcagcaaaagtgtcgtagagaccgctgagttga tttgaactgtgaaaagcgttgaagaaggtaagcgttcccgcaaggagaggactacc agaaggagaaatgtacggacactacaatgtacggcagtatgtgtcgttgagagacgg ctatactgccagtgctgcgtgctcattctaccaagaagaagatgtcaaaactcacgttc atcatctgagactcgtaaaactggaggagatgctccatcaaccaaattaccctctgtga agactgccaccagaagctgcacaaaggattgattcctgcggcagacttcaaaaaggga agcgtgaagtctaccagagatgctgtttatgggaatcatgagaataaccctgtgacagag gcttcgcagcgaactaccgattccggtgtgtaaacagagggttatattaccaaggcaac acgagaaaagctgctggtacttctaaaagccataacatgacgcactggcaatcgccca aggagacatgcttcaatgtaggatattccagagacatcccaaaaataacaggatttat acgattcgtcctgtaagacaccacaatcgtcaattacacaaagcaaccgttttgaaggcg gcatacgacgggtgaaccaaagcagtaaaagtacgtctacggattccgctatagcaaaa gtcgcgataaacgggttgatgcttctgtcggcaagacaggttatgggctatttccgact gaggcactcttgaggagaaaaatgtcaaatgcgggtgctaattataagaacctcgaactatta gaacgcagtcaaaactattgtatcgctga	gtcaattaaccaccgcttaaacgctgaac gtttgaagcgggggcttgcgttaaaagtgc taagcctgattgactaccctaaacactcggt gtctactgtactcaagaatgtatgtatagca ccggttaggcgttaccgaaactgcccgtct gcggtgtatgattaaaagctctgagggtga gagcgggtgtgcgcgcaaaactgtagat aacattggggagcggatcctgacgtcgaa agactgcgtggtttctattaccacaataatc tacaanaagggaaggcatcatcgccgaa caaa
1355-1356	Ga0302213_1000513 GENOME_ACESSION: II_Fen_N1_2_2_5F_3300 028769 GENOME_ID: 27309 CONTIG_ID: 512 SOURCE: JGI DATE: 2019-01-27.	atgacaaaattattgtatgtatgataaaaaaacaaagccacttgagccaactaattataat aatggttgagattggttagaagaagaagcgacattagttcaatgtatccgtttgtgatt aaattagaacgagaagttgaaacatcggaagtaatttataatgtgtgattgataccgg aagtaaaattactggaattgctatcattctgaatgtaaaaccataaataaagttttgttaagg gaaccatagcccatagacgagcgtgaacaaagttaattgaaattacgaaaggacaaga aaatataatgatatgctcaacattgtatgcaaggcaaaacttatctcagagaaaatttcta aagatatatccaattgaattaacgattgttcagaaactgcgaataaacgcattgattggtg cattgagaatcacatagtaattgacgattgtattgttgcgaataaaaccgtacaagtgtg ataattcaagattgggtataaaacaaatgcgaaggcagagcaaaagccaaaacggacaat gtattgggaattaaacacaagatttgggtttatcacatataaaaatggtgaacacataat ggatatgtgacagcaatttaccagaactaaatgcttgaatctcagtcacatcaaaaacatt gtaaaaaggttaacgctaaaaaggttaatttatgtgagatttaataaaatttattgatgta g	tagataaacgtgtacaatgtgacaacat ctattactcgggacagctgatagggttcac gcctaagtgaatgctaatgaaattgctt gcatgatgagctacgatcgtataaatgccaa agacacaccctggatgtaattctcagctccgt tgcctgtgagtataaaggatgaccactct aatgtcctgaagtcaaacaccgaactacatg tgcattatattgtcattgcaagaagaaata ctgcgaagcgagagtagtcagagatgac aaat
1357-1358	LO132_10002610 GENOME_ACESSION: IMG_3300002465_5F_3 300002465 GENOME_ID: 305139 CONTIG_ID: 2609 SOURCE: JGI DATE: 2019-11-15.	gtggcgcgtgcccgccggcaatccccggcgtctcacgcgcaagggaacgccca gatgtcacaattggtatagttctgcacaaagacggcacaccgctgatgccaccgctcc tgcaaaaggcccgccactgttgacgcgggaaagcgcgctgtggtacagcgcgaccc gttcacatccaattgaagggtctctccggcaagcacgtccggcggtcagcgtgggtgt ggacatgggagcaaaagtggtggagtagcccgcgaggctccggcggggtgctgtg ccagggagagggtatcctgcgggacgacatccgtcggcgagtgacagacgacggg cttaccgtcgcgcgcggcgagtcgcaagtgcgctaccgtccctcgggtcaacaac cgcccgccggcgggagtgctgccaccctccatccgtcgaaggtgacacgacg gtcaaggctcgtgcgcgagtagccttctcctccgtctccctatccgggtggaatcg	gtcaaccacccccgactgaaagtcggggg cttgtaggggagacgttcaagccccggtt gaccggcctcagcccgttccgctcaggcc gatatagtcagggtcactgtcggcaag gctaagacccttcaaacgcttttggccgc gatagggagcgtaccctggatgtctccctc agttccaggctctacaactaccggttcaga caggctgaagggttaggcgcgaaacggg ccggtgatagtcggaccgacaacattggc

SEQ ID NO:	Name	IscB sequence	ωRNAs
		<p>tcaacttcgacacacgagcgtatgcaggccgggggaagcagattccgggatggcctta tcagcgcggggagcagtagcgggtgggaaaacgtcaagatgtacgtgctgcccgcga cgagtacacctgccggtattgctggtgctggtgcccggccgacctggaggtagaccac gtcgtaccgaggagtcgggggtggacagaccggcccaaacctggtggcggcgtgt cacgagtgaaccggcgcaagggaatcagacggcgacggaattcagttacccccag gtggcagaacggagcgggcaatcgtttcgtggtggcgtcacaccagggtggcaag atggctacgctggatgggtgaccgaaatcgccccggtcgagaccacctacggctacgt gacgaaggtagatcggggaagcgtggtttgccgaaggcgactactacgatcggtta gccatcgctcggcggggaagccggtcgaagtgctgggggtctacgagcgctcggg gctggtggcaggggggcataccggcaacgcaaggagaccgttcccacctgtgtcc agcctgccctacgaggtgtttgtttcggcagtgaggacaggggtggcgtaccagacgg aggagtggaattcgtcagggacggcggtcagtggtgatttgcctatccgatctgg aggggcaactgattgcaccgtcggtaattgcaggaggtcgcggtgtgctggccgagcc gtgtgacttctgacgggaaggagggcggtcttccctcgggctgagcctcagggcg aagcctccccgactga</p>	<p>gagggggacctgacgggtggcgtgccc cggccgggcaatccccggcgtctcacgcg caagggaaacgccagatgtcacaaattggt atac</p>
1359-1360	<p>human gut metagenome genome assembly, contig: NODE_9_length_216355 _cov_6.761368, whole genome shotgun sequence GENOME_ACESSION: OWPY01.1 GENOME_ID: 12183 CONTIG_ID: 8 SOURCE: WGS DATE: 2019-01-27.</p>	<p>atggatatgtgttaataataaacggtaacctttaatgccaacggaacggtacggtaaagt cagacgaattttaaagtcggcagagcaaggtatgtaaaaggacaccgtttacaattca gtttgtgtatgaagcaacaactgcaccagcctgtttacattaggcgtggatgcagggtat aaacatacgtgctgtcggctacgacagaagtaagaataatattcaagtgaaattgttg aacgctgtgatatagtcaatttacttaaaaccaaagggaatgtcgtagaacaagacgaaa taacaaacacggtatagaacaaacggttaataacgtgtcgcagtaagcataagggt ttggttagcaccttcagtagaacaataaatacaaacacataatgaagttatataaaatgtgtg cagtatcgttctgtacacaaatcgtattgaacacagctgagttgtgttcataaaatcaa aaacctgaagtacagggaatagggtatcaaacaggcgaaaggtacggctttgttaatac aagaaattatgtttgtggagagataaacacactgcccgtgctgtgtgtaagtcgaaaggc gtactttttgtgtaaatgtctaaaggtataataacagtagcacctgaagattgttatacgggtg gtaaagaatgtctgtataatcatattaaaggtattaaacctttaaagttcaagaaaaaagac atttgcaccacctacacaaatgggtattatgagaacactcttctgacagactgaaaaatt cagtaaatgttctgttgaataacacacggttatgtacaaaaggcagataagggaagagta tggtattaataaatctatacaaacgatgcctattgtattgcaggcaatttacaagcaaaaaag attaacgaatattacctacaaaagaagggtcgtgtcataacaggcaaatatacagaatga acactcctcaaaagggtgaagaaaaacttaaccaagcaccgtatattgtaaaagggtacaa gctaaatgataaagttttatataatcaaatgtttttatccggcagggcgtacttcaggat atttgcataaaaggatataatcggcaaaacttatacagttctgtagctacaaaaaattaaag gctgtgtgaaacagctaaacactatactttagaaggaattgcaaggtaa</p>	<p>gtcaacgaccacgactgaagtacggact tgtaaatttaacagtcactttagttggctgca aaagcagctgtgtgattagcctaagcttaatt gactacgttatatacaagaatatatagacaccta cggattttacgccaagctgtgtgctcgtgtgga gtacgattaaagaggttcgaagggtaggaac agtgctgtcgtatgtaaaacctgtataata cattggcgtggtgtttacagcactattgtg ctgacttacagtaaatgaaaggagtaccgtg atg</p>
1361-1362	<p>NODE_1723_length_100 22_cov_2.7427 GENOME_ACESSION: mgm4790047.3 GENOME_ID: 238341 CONTIG_ID: 1722 SOURCE: MG-RAST DATE: 2019-01-27.</p>	<p>ttgcgtaaaggcaagacacgcgtgctgtgttcgggttaagggaactaaacctttgcggga ttatatccaatgcaaaaacacgaattccggtctatctcgaagtgtaaaccaattaatgccc actactcccgaagggttaggaaattgggtagaaagtgcaaaagccataggcaagtggt ctgacttaaatgtttggcagctcagctagtgttctgagcctctgacatgaaactcagccta ttgtatcgtgtgtagaccctggcaaatcctactcaggaatttcagccagctctaagcacttc accctattcgttccatgtttgtctgcttttaacagagtttagggcaaggatggatcagcgt cgaatccttagacgtggacgtaggggcggaagagtaaacgaaaggtggcatacaacc ttcagcccatcgtcaggttcggttcaaacccgcagaagctccaaagttgcgccatcaa ttaaggctctcgggatactgatttgcgtgttttctgagctgggttagcctttatccaatctc cagattgtttatgagaaagtgaagccgacgtagaccctactcaggtcgaagctgcc cgctctgggaaagggtttagccagtgatggttgggcaacgttaccagctcgtcacttgc ccactatgccccgggtgtcactcgcgagggttggcagaagacggtaattggcacgctct caactatgatctcgttaggtgttttaaggacaaagaaaacaaaggcaaggcttctcgt aaactcacgagtagatggagtcgtttggcctgctcgtcgtttgttcggtacggcaaaaa gtcgctaaccgaatgtgggtgggtggtgtgtttctattactgattcagttcagggtcat tactcgttttgagctgtaaagcgtggaaagcaatattggcttctatcggcgtcagctccatg tagaagtgcctgctaaggctgctgttcgcaagcgcaaaaggaggaaccatcactcgtttg gcgtgagggtcggcgactttgttcaagccaccaaggcaggagaaaccattactggctgg attggcggcttccaatacagaaaagctcaaaagctttctgtctacgactgacttggaa gcgtattgggcaattctcaatctcaaaagtaaaactactcaaaaggccacaggattatgtg ttgtctag</p>	<p>attatgccctttacccgtcgcagcgccgag cttcgtgaccttctaataatgctttctcactc actttcgcgttatcccacagacgatacgccg ctttcgtcttccccaatcctccgtcgtcgtg ggcacactcgcctccggtcgtgagctagg tccttcaactcaaatctccatcaagcgccg attcagccgcgcatctcctaactgtgacttc tcagctcttctgtggcccaattctgcatgattc ctgcctgctcgtggtcgtgctactcttca gaataactcatgcttggatctctcaagtc</p>
1363-1364	<p>Ga0194137_10001287 GENOME_ACESSION: TAR4v2_2_SF_3300018 411 GENOME_ID: 35909 CONTIG_ID: 1286 SOURCE: JGI DATE: 2019-01-27.</p>	<p>atgacgtacaataagcaaggccaggtgtaaggaaagggtgataccgggtgtgtctccac gggtaaagcattgatcctgtacacccgtagtagagcgtgaagctagtagtaagggttaa agcagttaaagagtgctcaaaaggcttcttcatatccagttgctgagctgtaggacggc gacacgcaggagatcgtcgtaggcgtagatccaggaggaagagagaagggttcaca gtgctctcaggcacacgcacttctcaacatcaacgcccacgcagagatggaagggc tgtcaagaaggcgatagaaaaacgcgctaaltcagctaaactagcgtagtaggaata ccccgtaccggtctcctaggttaaacacccgttctgtgaagactgataccccctagcac taaagcccgggtgcagctaaactcaacatcgtgaagcactgtatgaggtttaccgat cacgcgtgtagcgggtggaagacgtggccgagtcagatgaagggtgttatcgttggga</p>	<p>gtcaactacccttcttaaccgaagggtt gtaacagtgaccacatcgtctaacagga ctaacagttgactagcctatgacctagag gtcaactgaacacctcgggtggaaccg agacgcactacggatgcttccctagtctgt agctctgtgtcgtcagtggtgcaaggacc ttaacatgacatcattgatctgtataggact gaacatgacataaagcaaggccagtg taaggaaagggtg</p>

SEQ ID NO:	Name	IscB sequence	ωRNAs
		acaatagcttctgccatccagacgggaaagaactatttctacaaggagttagagggtct agggttaacactcacacagttcaaggcactgaaccgcctctctacgcgctgctagg gttgcgtgaagacctcgtctaagctagctgaagtgtttcagcccactgtgtggactcctggg tgctcgccaaccagggtgttaggtcaacgaatcaccgtagacctaactgaactgtttacctt gatcccgtgtgtctacgccgtagacagctccatgtctcaactgggctaaggagggtgt acggaagcggtgggtgtgtacgtgtcttttaggacttcaaaagggtacgtatgtatcac cctaagtgggtaaatgtacctcgggtgtgtgacgggtgataagagagtgctgcttcatg cttacagtaacaccaagcggttagcacggaacgttaagctggaagacaccaggtagta gcccactcacctggcgactgttagacactcccgcgtagcacacatcacgaaaagag gatcaggagggttaaacacactgtgcattagactgcaacactaggcacgtcagactcca gaggttttcgcgactcaagtcgggcacaccactaccaaaggcgagcatag	
1365-1366	contig_3211414 GENOME_ACESSION: mgm4547164.3 GENOME_ID: 237431 CONTIG_ID: 344745 SOURCE: MG-RAST DATE: 2019-01-27.	gtgttaaacagcagttacaagtaaaagaaacatctacggattcttcacaagttc atagctctataattcctacattaaacaggtctaaagacacagtggttaggaagatactctg cctgttaaacactccgatgtgaatctacctgaggtctaacaggattgagagaatacaaaa tatttgcctcaatgtaaaaggacaaccactaatcctcaactcctgttaagctaaatattt gattaaaaaaggattagctaaagtagttaagcatatccattttgttattaaaatgttagtaa aaacaaggaaattgtagcaaaatgtatattaggagtagatgtattatacaaaatattgga tttagtcagtagaagcaataagaaagattatagaaagaaagaaagaaataaactttggtat aaagatagattaacagaagaagaatgtatagaaagaagaagaataaactttggtat agagaagtaagatttaataatagaagaagaaagaaaggaatggttacctccatcagtagaa aaaagatacaaacattttaggacttattataagattaaaactttgttacctattatgtatgt agtaatagaacagctaactttgatatacagaaaatagaatacctaatttcaaggaaaa ggataccaacaaggagaaacagttaggatttgaataatgtgaacaatatttattacaaga gataaaagaacttgcataacattgtgtaagataatactaaattagaagtcacatattttaa tttagaagtcagggtggaactgataaaccttctattttagaataattgttagtaaatgtcata aagatttacataaaaaatatttagtataaaaagattttaaagattataaaccataatcatt tatgtctatttataaaaaagatttggaaagatataccaacaatgtatgaacatatgttata taactaaagctactagattaaagtttagattagataaaactcattataatgatgcattttgata gctggtgaaatactcaagaaagagtaaaatcaatagaattaaagcaaaaacataggaa aatagagtattacaactaaatagaagaaagattttaaacttctattagaatacaagataaag atacaaccataagattttagttaaataataaataatgtatataacgaacggattacattgta aaggacaagaattatgataaaaaaagtcagtaataataagtaatttagcaaatattttaa atatgggaggtatgtcttttccctcccaaggtaa	gtcagctactaaactaaagcattgggctt gtagtgaactacaactgaagagttgatta ggaggttttagaaggtgtttaaaccagcatt acaagtaaaagaaagaaacatactacgga ttcttcacaaagttcatagctctataattcctaca ttaaacgagctaaagacacagtgtagga aagatactctggcttgaacaactccgagtg aatctacctgagctaacaggattgagagaa atacaaaa
1367-1368	Ga0373632_0083918 GENOME_ACESSION: RO12ediment_MG_6_\$F_3300035181 GENOME_ID: 306458 CONTIG_ID: 83917 SOURCE: JGI DATE: 2019-11-15.	atggcagttttgtatgggacaagcgtaaacaccgctaafgcttgcagaaaaacgg gcaaggtgtgtctgatcgctggacatgccgtagtggtttacacactatccgtttaccattcgt ttgaaaaaccgtattggcggcggaatccagccgttgcgggttaaaacttgacctggcagct cgccatacaggcagtggtgtatgttcgggagataagaaaaactcacggcttgaatttgt tccagctcaatcatcgtgtgttgcgcatcagtgaaaacttaacctacgcccgtgttatgcgta accgtagacgcggtaacttgcgttatcgtccgcagcgcgtttcaacaaccgcactaagccgc aaggctgttagcgcctcgttacaacaccgtgtgtgacataagcataagttgggtgaaca aacttcaaccgttgcctcgtgacagcgtttcccaagaactgggtgcgttgcattatgcaa ctgattacaaccctgaaatcagcggcactgaatatacaaacaggcgagttgcaaggctat gaggtacgcgagtagctgttggaaaaatggggcaggcaatgcgcgtattgtgtgctgga acacataccattggaattgaacatctgtgccaaaatccatctgtggcagtgatcgggtg tctaactgaccttggcttgcgcgttgcacaaatgcaaaagcagccaaccattgaaa ctttcttgctaaaaagcctgtattattgcagaaaatcaaaagcacaagcccaacgccattg agggatgccgcagcagtaataatcaccgcgtgggtctttagctaagcaacttaagcaaca gggttaccgtagaattagccagtggtggcagacgaaattcaaccgcataaccacgg cattcccaaatccatgccttggatgcggtgtgtgttggcgaatagtcgcggttgcctatt ggaaacttctaccttggaaatcaaatgtactggcgccggcagctatcaacgcaccgttt ggacaataatggttcccgctgggtatttactgcgtcaaaaagcggtttatgcttccaga ctggcgaccaagtgaccgcaatgtgaccaaaaggttaaaagacaggcagctataagg acgtgttgcaatagctggcagtggttttcaataatcaccaccgcaacaggatctgtgcaag gcattagccacaacattgccaaactaatcgtctgtaacgacggttatgggtatattttcaac taaaagtagcttaa	gtcaactaccccgccctaaaggcgaggct tgcaagagctagggtgaccagactcagtgct gcaaggcactacgcagataagagctgtta agacttaccaaaggatgcttcttagtccctt gctctaaaagctcgtatcacgctggcgca aggtaaaacggcgaggttggcgccgaca ctgagcgttagctgaagtgtagccgcttctc tgcttgcgaaggagattacgggaaacc gtagcgttactagaccgtaagggttaactatt tgaggcacaacaatggca
1369-1370	Ga0172378_10001289 GENOME_ACESSION: GroweLOW332metaG_2_\$F_3300014203 GENOME_ID: 27072 CONTIG_ID: 1288 SOURCE: JGI DATE: 2019-01-27.	atggatttggatttgatataatgataggcactatgccatgtaaaagaagtaaaagccgt aagctcttatcaaatggcagagctacggttatcataaaaatgccatttactataaaacttttat acaaaacaggttgtgtacacagcccatagaaacttgaatagatactggtacaggcaata tagggattggaattacttctgttaaaaaggttttacacaaaagctgaagtgtctctagaagta aggatatagtcaagttgcttactactcgcgttatatctgtagaagcagacgttctcgttaa acggaatatagacatcacaattccattatcagacaactataaataatgtgtgttgcgtgttaa aacgtcaacataagattgtaaatcaacaagagcgttgggtctaaaggtgtccataaatctt atgtccaacgccctgaaggatggctgccgccgagtagacaatcgaaggtaatgcaca gattaaagtgattgataaatacttcttatttaccacatccgtcattaacataagaagttgtc gtttgatattggcacgaatgaaaaaaccaactgtacataatgaactttatcaatatgggga	atgaattaccgagcatgaaatgcaaaagtatt tattacagctaataacttgatgactagcctc agtcctatattggactaagttatgaatgcaat acatattgttagcagataccctaactctaat gttttagcctgttaaactggtttgcgagacg aaacagtcacactaaagcagaacattctat aaccttggcgagggttacttacgtgaataaa cacacgggcagatcagttttgtatctcggcta aagttataactgtagaaggcgaggtgttat atg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		atgtatgattatgacaatatcaatgtttatgtattgcacgtgataattatcatgcaatgctgt catacaaaagcaacaccgcaaaacaatttaagctgttaatacaccatatactttactca aaaggcggtagtgatagaccgagtaaccgtatcacgatatgtgaaagggtgcatacttc aaaatcatcagccggcggtatcttataatgttgatggttaataaacgtaagggtgctaaa acttatagtgatgcaacacaaatgaataattatagtagacgtatgttcaagcatttccgcaa gcaacttttatacatcggaataatcactaattcctgatcgaagaattgatactttcaaaaagc cattgtaatgatgccatagctattgctcagcaaggacaggaacaaatattcatgatatacct actgtattgtatataaaaacacgtacgtaaaaaagacgctcattgcatgaaggaaftccacg taaaaggcgtaaaaacacctaaccgtgcacagattcgaacagtaagaatacaaaaactgt tggtaccctccatctaaccgactgtgttaagttaacaatacgaatcgggtggataagcggatt ttctggtaaatctgcttatgtaaaagggttacaaggacatatattcaagaagtaggaagg cttataagcttatatacactcaaaagcttagtaagtaggaatgcgggcaattgctatafc aatacaagatagcctaa	
1371-1372	human gut metagenome genome assembly, contig: NODE_90_length_10476_6_cov_9.938421, whole genome shotgun sequence GENOME_ACESSION: UPME01.1 GENOME_ID: 16637 CONTIG_ID: 89 SOURCE: WGS DATE: 2019-01-27.	atgggtatgttttagatattaacggacatccattatgctacagaacggcatggaaaaagt aaaacatctttaaaaaacggaaaagcaaaagtagtcaaaaagatgtccgtttacaatacaa ttattgtataaaagtgttacttatacagcgtctattatattagggtgtgatgcaggaagtaaca tatcggtctttcagcaacaacgaacgataaggagttttatgcttcagatgagaacttagaa ctgatatagtataaaacttatacctaaccgtagagagttaagacggagcagacgggaatagaa agacacgttacaggcaagcacgctttgacaacaggaaacgaaaagacggatggttagct ccgtctataagtcaaaaaatcaaatgccatataaaagtagtgagaacgtacacaagatac ttccgattacaaatattgtttagaacaacgacatcatttataatcaaaaataaaaatccta attaaagatactgaatatcagcaaggcgaacaacttgaattttggaatgtcagagaatatgt cttatttctgtgacgaacatafatgtcagttgttgaaggtaaaactaaagataaaaatcctta gttcacacatagaaaagtcgtaaaaccgggtggtatgctcctaacaatttaacactttatgt gagacttcccaaaagggttatcacaatgggactgttagattacctaataatttaaacgtgg aataatttttagagatgctgctttatggggattatgcgttggcattttataataagctaagag aaaaataataaaaatgtttcaatgacttatgtttacatacaagaatacaaggattgaaaata aattacaaaaaagcattatgttgatgccagatgtattagtggacaccctttagcaaaacca tctgaaatgatttactatcaaaaaagataagatgtcataataggcaggtgttcaaaaataat actttgaagggtggcacacgtaaaagaaccaatctgaacgagtcataaaaggatatata cttttgataaaagttaaatatcagggttaaaagaaatttttcttttgtagcgtataaagtgaatt tttgatataagagaacttaattggcaacaagggttaacaagggaagtataaattgcaaaaac taagctaattgaaagatgcagtggttaattatagaattagaaggagacgggtgctctc tcataga	atgtcaataactcatgactaaagtcacgagct tgtaaaagcccgatattgaccagcctaagtg tttagtacactacgtttaaagagaaatataatg tacctgtgaatgttatccagttcgcagctcta aggctcagtggttaaacaggttctgtggagtag gaacagtgctgctgataataaaccttttattaa ctttggcgaggggaacttaactctttaatttta aaggagcattatgatttaatttcatcatttga gaaaggagtgccgtgaaatg
1373-1374	Ga0395718_011872 GENOME_ACESSION: 713C1X0metaG_FD_5F_3300037107 GENOME_ID: 312125 CONTIG_ID: 11871 SOURCE: JGI DATE: 2020-04-07.	ctgcttcaagtcagatccccggctaccgacagccccgcttcgacaaccgcaggcggat ggaggcgctgctgctcccgctggaagccgggtggccaacctgcagacctggga aagaagactggcctgtgttctccgttggaaagcatctcagtgagctgggtgaagtttga caccagggccctgcagaacccagagatctccgggtggaataccagcagggtgagct gatgggatacagatccgggaatactgtctgaaatggggccggaagtgcgctact gcggcaggacagaagtcccgctggaagtggaaacacattgtccgaaagccggggcg gctcgaaccgggtcagcaacctgacctgctctgc	gtgtactcgtgtctcgtcgtcgcacttcatg cttttctgggtccagtttcatgtatgctggga gaccttttccagtttggccttgcggaagatct cataccagctggcatccagatcttccattca gccaccaggtccacaccttccggttgcac ccacaatgtggaatggcgacgtgcacggtt caggcccatcaggcgcgtctgttttcatcg gcgctaagcgcggcggtatttccgcgggt ttcttccaaataacggcatgtcacacctctct ctatatta
1375-1376	Ga0208461_1004216 GENOME_ACESSION: AD_des_26_5F_3300025 613 GENOME_ID: 21324 CONTIG_ID: 4215 SOURCE: JGI DATE: 2019-01-27.	atgattagtaatagaattcaatcagagtattgtattgaaacaataataaaccactaatg ccttgcacccggcaagagcacgtaaactacttagtaaaaaagaagcagcagtatatcgt agatatacatttacaattatcttcgtaaaagaactactaatttcaactcaaccgattgaattaa agttgatccttgtagtaaaagtatctgtatcgttttagttagtaacaacgtccttttatggggc gctaactctgagcatcgtggacatcaaaftaaatctgatttagattcacgtagaagtattcgt gatgtagacgtataaattacgttatcgtcaaccaagatggtctaatcgtaaacgtgaaaac gggttgcttccgcatccttaattgtctagatgataataataactaatttatgttatagattaag taaacgagtagaccacttaacaatatcgtataagaaactgttaaafttgacatgcatctatgca gaatggatgtcaattataggtactgagtatcaacaagggtactttacaaggatataatgtacg agaatacttacttgagaatttcaagagcttgtgttactgtggcgaactgtataccttta gaagtagaacataattataccgaaatcacgaggaggtactaatcagattagtaacttaactct agctgttaatatgttatacaataaaaggcagtagcactgcaagtgaattggttatccga atgtacataaactagctcaaatgccactaaaagatgcagcagcagtaattgctactcgtat cgtattggttaagatcttaatttaagggttccagtagaataactactggtgctcaact aaaatgaatcgtataaactcgggttaagtaagacatttgattgagctgcatgactgga aatactggaagtaataatgtttcccaagtaataatagtagtactcaatattaaagcaatg ggtagagggaatcgtcaagtaacaattatggacaagatggatttctcgtaaaactaaaa ctagggttaatacaacctaaaactgtaaaagagttcatggatttctactggagatatagtta agttacacacacatcatagatttatgtctggagtttacgttgaagaattactgcaattcg atctagaggtagcgaagtataaattgacgggtatgagattgatttaattggcgaacttt aagttacttcaatacaacgatgggtactcattcaacataataa	gtgtccatcaactgcccttaattatttcatg ctcatccagcttaagtaacttctgtattacta cgttaattacgaatgagaatttatttcaata ggtactccaggatgcactccagtcggga ctctacggtaacgttaftaaacaacagacc gcgacaatgggtagtgtacattacaataaac cgttaataacattagcaggagaccttgacc cttcgagaagactaatcccttacgggattatc ataggaaagtaataatgattagtaatagaattc aatacga

SEQ ID NO:	Name	IscB sequence	ωRNAs
1377-1378	human gut metagenome genome assembly, contig: NODE_33_length_114217_cov_11.444533, whole genome shotgun sequence GENOME_ACESSION: USYB01.1 GENOME_ID: 17033 CONTIG_ID: 32 SOURCE: WGS DATE: 2019-01-27.	atgtccgtttccgttttttttgacaggaaagcaccggcctctgatgccgtaccgaccgc tcggacgcgaagttaattctgacgcgtcaaaagcgggttcacgggtccacgggtgaca tggtgcgggccgaggtccctcggggaaggaggtcggaaacctatccgggtcgcgtgac tgttctaataagcgggtctttcaacattcagacgccgaaggcatccagggtattggt ggcgtcattgtcaacagcgtttcttacacgacggataccgggtatgcgtgcgtccgc acccattctctccccgtttga	gtcaaccacccctccttgaagagagactt gagcaatcaagccggtgaccggcctcagt gatccgaaaggagaactacgttgcgtcgg gttacaagaccgctccgggggggtctcctc agctccgaactccggaacggcagaagca gacaagcttcgggcaagtacgaagcgtct gtcgttcaatgccggacgacacattggc gaagggaattccgttttcggggccctc cgattaaccgtaaggtttaccaaatgtccg tttcc
1379-1380	Ga0137392_10005405 GENOME_ACESSION: 15con2h34AmetaG_2_\$F_3300011269 GENOME_ID: 20764 CONTIG_ID: 5404 SOURCE: JGI DATE: 2019-01-27.	atggagcaatccatcatttatgtctgcacagtcgggcaagcctctgatgccgaccagaa accgcaacaaagtgtgtgtattggctccgagagactggcccgaagtgtcagccgtga gccattcacgattcaactcgtgtttgagaccacgggtacgcccaaccgggtcacgggtgg ggtagataccggctcgcagacgggtgggatcgtgcatcaccagagcagggcagtc taccaggccgaggtccatctgcgcgacgacattcatacaagctcataaacgacgaaa gtcacggcggaaccgcccggagccgcaaaacgcgtaccgcgcgggttcgttcgcaa tcgtcgtcgaagccagcgtgtgtgcgtcgtcgtcgtcccaaggcagagggccacag tgaaggctgtgtctttgtggcagcggctcgtccgtgtgccagatcaacgtcgaagtagg gagtttcgacacgcagaagatgcaagatcccgcagtaacgggtcgtgagatcaacaag gccagatgcaaggttacctgggtgcgagatcttttaccggaagtgaagcgcacgtgtt cgtattgtggtcgaactggccttctctacaggtggaacatctcagtcctaaaccagagg tggttcagatcgcattcgaacctgggtcgtggtcgtgtatcctgttaacaaaaagggga acacagacggctctgtgatttggatttccgacacctcaggcgcagggccgtgtgccactg agagatgctgccacgtgtcttctatcaagactgcccgtgttcatcagctcgttcagcagt cggaccagatcaggttctgtgacctatgggtatgaacaaatacaagcgtcttcagatc ttccaccttccaagtcgcacaccaacgacgcgattgcttattgtgtcagtgaggagag gtgtgtcaacccgctgcaagcattttaccaccgctgtgttctcgcgggaactatcagtt gtacaacgggtgtcggagcgcagcaaaaggtgtgggcgccgaagaagctgcatgggtg gaagctgtatgaactgggtgaggccaaagggtggtgtgtatattggcgttagacgcat ccaaggggcttctgtcatgaaagatgtgcaaacaggcaagggtatctcgaagtaaccc cgcgcaaacctcggcggctgtgctcggcctggcacagctggatcgttggccggcagtc gcttgaagaaaggaggagcgcgttctctccccgcctag	gacctccaggcttgcctggagcagctcgc gaaagacccttgggtttgagagcatgcc gggtgaccagattgtgccttggagacagg caggccgagcaagcgtgctgttgcctgtc gcagaaaaaggtccgatagcagaagtggt tcaagcactaccgagaagctgtgtccccc cttctgctctagaagggctttcttaaccagt gtcgggtcagccaccgtggagagccgaca tcccgcctgtatcagatcgggggaaacg cttactccgaaaggaaagcttatggagc
1381-1382	metagenome genome assembly, contig: NODE_226_length_40329_cov_4.185215, whole genome shotgun sequence GENOME_ACESSION: OJKJ01.1 GENOME_ID: 6455 CONTIG_ID: 225 SOURCE: WGS DATE: 2019-01-27.	atggtgtatgtattgaatagaacggaagtcctttgatgccatgtaaaggcgaaggcta gaagactgttaaagtatcataaagcaaaagtatataaaaggagccatttacgattcaatta ctatttgaatgtgaaaaccatatacaggatattacgcttggattgatgccggtagtaagcat attgattgtcagcaacaacaggcgaaggaattatatacccgctgatgttgaactgagat gtgatgttggaaactgttatcaacccgcagacaatacaggcgaagcccgtagaaacagat tgagataccgggtccgagatttcataacgggacccattctaaaaacaaggctggcttgc tccgtcggtagaacagaaaatcaagactcattacgggttccaaaagagcatattacagat gctcgtctgtatcagggaattccacaggcagaaacgctgggtattattttagcaaaaga agggtgcgtgtcataacaggcagatccataaagcaaatatcaaaaaggcgggagcag gaaacgcaatcagatggagttatttagtaaaaggatttcgattatcgataaagtcagatatg ataatcaggctggcttattctttggcagaagggaagtgatattttgataaaggaaagttg aatggcgaagtaattcatcgtctcgtcatatataaaaaattagaattattggatacaagaaa aacattgttaacagaaaggagacaagcagttatctctatggcttaa	ttccaggctcaataaccaccattaaaaatggc gggctgtgtaaaagcccttattgactagacttg gtgttttaagcactccgttatgtaagaatatat agttaccagttggagctacagcctaattccact gctctaagaattgttaataacagttctgacg ggcaggaacccgtgttgcctgttagtaaacctt atataacattgtctaagcgtcatgacttcag catacaggaggtatttaaatcatg
1383-1384	Ga0131092_10008720 GENOME_ACESSION: IMG_3300009870_\$F_3300009870 GENOME_ID: 312511 CONTIG_ID: 8719 SOURCE: JGI DATE: 2020-04-07.	atggtttatgtaattaaacaaagcagaagcacttatgccaccgaaaggttggtaaagt aagaagactattgaagaatagccttgcctatgttgccttatccattcacatccaatt agattatgacacaactaactttacgcagcccgttaagtttgggcattgatgctggtatgaac atatcggcatttcagcaacgacgagtgagaaagaattgtacgcagcagacgtagaactta gaaacgatattgtagaaaattatctaccgctgagaacaagaagaactcgtagaaca ggttacgctatcgaaggcaggtttcagcaatagagtggcatcgaacataaggaggtggt agcaccatctgttgaacaaagattcaactcactaacctgttagagaaaatatacaagt ttctacctataactaatatcacagtcgaaactgctcttttgatatacaaaaaattagaaccc taccatatctaacgaagagtagcgaaggtgaccacttaatttcttaattgtcgtgagt ttgtcgttttagagataatcacacttgccaacattgtaaaggtaagagtaaggataatc taaatgtacaccactgaagtagaaaagttgtggagattgccaaataatcttataaca ctttgtgaaacctgccacaaggcataccataaaggagaatagaactcaaggtgaaacg agggtcatcatttaggaggtcgtcttttgggaattatgaggtggaatttctacgaaagatt gaaaacactctacgctaattgtgagtgtacttgcgtatatacgaagaatactcgtatta ctaattgtttgccaaaaacattatgtggatgcgaggtgtataagtggtgcatcctacggc taaacctcttgggtatttttctaccagaagaaggtcgttgcgaacaggcaaatacata agggtcaattttctaaaggcggtagaagaactcaatcaagcaccgttcttgggttaagggt tttaggctgtttgatttgggtgagtagcaaaaaggttgaattatcttttggtagaagagatag	aaaaaaatagatatgtaactaccagagg ctaaagacctcttggcttgaagaagcccaag ttgattagactaagtatacccgaagggtt gaactacgttaggagagaatatatagttacc aagtggtgttctcaagcctcttctctaa gggtgtgtgattaaacagttctgtgaggtagg aacagtgtgtcaacgaaaacctctccata acattgtcgtatgagcatttaacggagaatc cgactataataattattaaagtaagaccgaat g

SEQ ID NO:	Name	IscB sequence	ωRNAs
		tggtttctcgatattagaaaacttgacgggacaaaagtaacaaaggctctattagttgtaa acaaatcagattgatagatcagagaagaagtataataacagaagaagggaatagtggttc aatctctcccaaaactaa	
1385-1386	Planktothrix serto PCC 8927 isolate BBR_PRJEB10992 genome assembly, scaffold: BBR_A_PL8927_scaffol d59, whole genome shotgun sequence GENOME_ACESSION: GCA_900010725.1_BBR_PRJEB10992_genomic GENOME_ID: 176102 CONTIG_ID: 58 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	atgtctaattacgtctttgaattgatccgaataaacgcctcaaaaccagttacatcccgcc caagctcgattgttattgaatcaagggtcaagctgctgtttatcgccgttatccgtttactattat cttaaaaggagtcacaaaccagcaccagaaatacaacaataacccctcaaaattgaccccg gttcaaaaaccacaaatttaaccgcactcgattaaacctacctaataaccattgggttagatgca gctgtgttggaaaacttgaactttaaggtgttaaccaataaacctttattgattcaggcga cgggacgtgtgactcgtcaaatgtgtgttacagacaatacggattccaaccttccatg ttttcttgtgattatgctgaaaaacttatgctgagtcagttaggataggatctgttcttcag gtcttatcgattcagataggttaagtcagaagaatcattga	gtcagcaccctgctcttctgtgacggggct tcattgcctcaacttttagatagctcgtacccg cctcagcctaaccggtacgtttctggaatcat aataccaccgaatcgacgctagtttggc tctattgttaacaattaacagttttacgagg gtaagacagttgtttaacgcaaaagtca gaaaacattggcggaagcgaacattacccta gaaataggagttattcatgtctaattac
1387-1388	human gut metagenome genome assembly, contig: NODE_655_length_4111 _cov_3.00098, whole genome shotgun sequence GENOME_ACESSION: UYAI01.1 GENOME_ID: 17510 CONTIG_ID: 654 SOURCE: WGS DATE: 2019-01-27.	atgccgtgccagccgcgaaggcaagaatctgtccgtgacggaaaggcaagggtca tcaagcgcgagcctttcacatccagttgaaattcggtccagtgatataagcagaagat tctcaaggcggaacacgaagctgaatcaggctgcctattccattcatgatttcattctgt tcgacaaagtacggtttgtctggacaagaatgcttcgtgttggcgcgtaccagcggata cttcgatcttcggaagctggacggaacagttgttcaccgttctgttctggaggcagttgc agaaaacctgcaatcgtaaatctatctgattagtaaggaggtgatcgccgtctctccac gactgaagtcgccggtgtccgcgccgaatatcatga	gtcaattaccctcgactgaagtcggagcgtt gcaggaaacgcgcaagcgttaactgacaagc cctgattgactagactaagttctccgagaact atgttgccttggttatctcaccctgggtgttt atcctagcctcgccgtctgaggcggtctgt aaaagtccaatggtagggacggtcaacc gcccggacggccggaattccgggaagcc ttggcaacattgtcgaagggtactttactcgtt ttgagccgtctcgccggcgagaaggagg tagacgtctgttc
1389-1390	Ga0209200_1016766 GENOME_ACESSION: AD_des_33_5F_3300025 784 GENOME_ID: 21332 CONTIG_ID: 16765 SOURCE: JGI DATE: 2019-01-27.	atgaaagcttatgtattgaacaagaataaagaacctttgatccctactactccggctatagc ccgcattctgcttaagaggcgcaagctaaagtggttaagaaaacaccgtttactattca ctactgaacgattctacggggttcaacaacccagtcgccgaggattggatcgtgccc gtgcatttaggggtgtgcggcgtgtctgataaagagggtttgtatgtccgaacacttctt gaagacgattataccgtaaaacaaaagatcaacgaagaagggttccggcggttaag aaggagccgaatcagataccggaagccgccattatcaaggcgatgatgtaagacttg ctctgcaaacagagttgcagaactaaagaatcgaagtcgccgaactgtatccgaatt tctcaaaagaaaagtggttttggcgaatcaagcttctgtcgcaccatctgtacaa acaaaggtcaaccaccacgtgcaggaaatgagaagggttgaaaaaactactacgtatc gaggtgggtaacgaaccgcgaatttcgaccttcacaaaataacgaaccggaggttg ccggggttggttaaatatgcccacaatacaatttctacaatgtcaaatcatatgtcctcg acagagatggctacaagtgccaagtttgcggggcgtccgtatcaaaaagatggaacc gttttgacgttcaaccacatggtccgaggagattatcaacagcagatgacccctctaa tctcatcactctatgtatgctcctccacaagaaggtacacaaatgaagccaccttaaac gctaaagccgatcaagaagggaacgccacgctacgttagctaacaccgttcgagccc gtttgattaaagctcttgactgtaagaacgtttgactacgtttgatatcaacaaaattcaa aaggcaggcaatcttggcgttccgaaagaactactctgcacgcggtctctattgctgtt gacgggaacaaaaaaccaaaattgagcgatcgggtttatctgtattcggagagtacccaaa aacgagtagcgaagagagcgtagcggaaagaaaactgcaacgtccggaaaaacgtat cctgaggttaaaatccgcgcaggtataacgggtataggaaatgggactgtgttcgat ggtgatgtaactgtttgtgaatatttggcaacaaaaggtttagtgattacagacctt gacgggaaacaaactactatcaatcagttctcaacaaaagcctgtaattctcaagcgggca gccgttatccgacggtgataaaaaagcacaagccattgtataccacaacaaattcattg a	tcaagagatcagatagctttactctcggc ttccccacacacggtgtgacagtagctattga aactctttttgaaaaagggaatagtgcttc ccccacacacggtgtgacagtagctactggtt ccaagcgggtttgagggttaactgactgtcc tatgtgtcccccacacgtgtctctgcaaaa gagatgggattccagactctgactggagaa atccagtcctcgttagtaatgtatgatatccg tgtatgccgcctcagtagacggttagatcgtc tagtgtctaaaaaggcctgaagtgagggg cccgccgctaagtgtaaaaagcattattaa cattgtcgagaggacgtcgaattttatctg gtaacaggataaaactacgcacaaccttttt ataaggaggattagcttatgaaa
1391-1392	Oscillatoriales cyanobacterium isolate PH2015_10S_46_199 PH2015_10S_scaffold_8 37, whole genome shotgun sequence GENOME_ACESSION: GCA_004292945.1_AS M429294v1_genomic GENOME_ID: 252669	atattaactcaaggtaagcagcagttacagacgttatccgttcacaataatcctaagtat tctctgtgaaaaaccgcttaccataatcgacgaaattaaagctcgtatccggctcgtgaaaacc acaggattagccgtattggaaggtgacaaagtattctggcgaccgagttagttcatcgg ggaaatcacatcactgtgtgacttagcatccgagccgcaatccgtcgaaccgcaggaa tcgcaagactagataccgtcaacctaggtttctcaatcgtactcgcactcccggatggttgc caccgagtttagcttcaagagatcgaatttgcagttgggttctgttaacagctatatctc aagaattagttaattcgacaccaagcaatgcaatttctgaaatcaggaattgaatat actcaaggcgaactagcaggttatgaaatccgacaatatttttagaaaaatggggcagg aaatgtgcttactcggcatcaagaatcgccttttagaagtcgaacacattcacaaggga ggaagttcgcgggtgtcgaatctgactctggcctgtagaactgcaatctggctaaagggt	gctatccaccaatcaccggacataacagga caaccagtagcttaaaaggtgaagagatagcc agcctcagttacaactacgtttcagggtgat gacacctcagtagctttccagcctgtgct ctgtcgtcaatcgttaaacatctgtatttggta aagaagtgcgggtgtgactcaacaagccgat taactcaaggtaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
	CONTIG_ID: 590 SOURCE: NCBI Prokaryotes DATE: 2019-05-30.	aatagagacatcacagattttctcggaaaaatccgacattttatcaagggttttaaaatttg cttgcttgcattgaaggatgcagcagccgtgaattctactcgacggcgctgtttgagaa actccaacacgggaggggcaaacctcagtaacatgaactcgaatgggattgccaaaaact cactggctgatgctcctgtgtaggtaaaagtagaacaacttgcgggtactaagtaggcaa ccactactgataactgccaaaggatggggcaaccgtcagatgtgtacgcccaataaafat gggtttccagcaaacacagaaccaggtgcaagacttttttgggtccagactggagata tggtaaaagcaggtgatagcagtttggcgtcgggaaattgcaggtactacgttggtga ggctggcagtttagagagatggagcttttgatttgaggacagcttggggaaaaat	
1393-1394	Ga0376086_0019836 GENOME_ACESSION: LAC_S2_MG_FICUS_ FD_SF_3300035686 GENOME_ID: 312530 CONTIG_ID: 19835 SOURCE: JGI DATE: 2020-04-07.	atgattcgtgtcccagtagtaagtcaaaatggaaaacctttaatgcccaaaaagctcaa gagcaagaaagtgggttaaagcaaggagctcgtgtaaatgtcagatgtaggcatt tactatgtccaactatcaaccagcagacgaagaacacaaactgttatagccggagta gatccaggtaaatcgtatgcaggaaatcgggtgacaatctggcaaatccactgtctagatt tcataatctaccctttggcgagtaaaacaacgaatgatcaacgaaggatgcttcgg cggtcaagacgctctagaaggatttaaccgcaagggttgcctttaatcagaacacatgcac agggtcggttcaataaccgaagcaatcgaagggtccaccatcaattaagcgtctgccca actagaactaagagttgtaaggaaatcaagaatatttccgttaacagctattggctacg agcgagtcagagctgacgtatagatagaaccaaacgaaaacagccaagtcgggtaaaa gccttctcctgttaagtgaatgtcaactttgggtatttcccaaatggagactattgcacctgt ctttgttcgtgaagggtggcaaaaaatgggaatggaacgtctcaaatcaggcagcattta ggactgattaaagacaaagaaaaaaagggtgaagccaaacctgaaactcatgctgttgat gctgtggcttagcctgcggttactttgtcaatcacgccgccaccttacctcaaatacca aggatacctctggaagatgaagtccaagtcacggattcagatcagaattatcacccgtt ctggtgctgcaaacgaggttaagaataggtattcttggcgctcaactgatttcgaggtt cccgaataatttgaacaagaaaaagtaaaagggggcaggttacaccctttgatttcga gtgggtgactgtgttaaagccaataaagcccaagaacacattgtgtgtacattggtggatt acttctaccaataaatcctcaaatgtttctatctacgactacacctgggaagcga	ttcactgaccccgcttacttctgtgaaggcg gggattgaggagataacccctcatgaac agtgaatagtggcatgagtagtacgagata aacacctccgaacacttcccaattcggaaa aactgtaagattctggaatcagaaggtg ggtaatgccagcaacatcgtgtaattgcc gtaaggacttttaactcacactcctagga ttatctccatgattc
1395-1396	Chloroflexi bacterium isolate CF_131 14_1009_09_20cm_scaff old_4778, whole genome shotgun sequence GENOME_ACESSION: GCA_005881015.1_AS M588101v1_genomic GENOME_ID: 285819 CONTIG_ID: 485 SOURCE: NCBI Prokaryotes DATE: 2019-07-22.	atgtcacaggtctttgtcttagacacgaccaagcagcgagcgtcaaccgggtgcatccggg gcgtgcccgttctcttgaagcagggccaaagcggcggtgtatccgctatccgttcac gatcatcctgaagcgtgcggctcgagcaaccttcttgaagcgttccgagtcgaagtcga cccaggggagccaaaccacggactggcagtggtcaacgatgccagtgccgaagtcgt gtggggccgtgaactgaggcaccggggcaagcaaatcaagcgcgacctgcagagcc gccgcgcccctcagaaggagtcgacggcaacgcaagaccaggtaccgccagccacgct tcgacaaccggcgcaaaaagcaaggcagcgtccgccatcgttagagagccgagtcgtg caacgtgttcacctgggtgcggcgtctgtcggcggtgtgtccgaltcggccatcagtc aagaactgtccgggtcgatagcgaagcgttagaagaaccagctatcaggggtgtcgaat accagcaaggccagctttcaggctatgaggtcagggaggtatctcctgtcctcaagtggaaac atcgggtgcgctactgtgacgcaagcgtcagtgcccttggaaactgatcacgtgcatccta gagccaagcagcggtccaaccgggtgagtaactctgtggcgccctgcagcccttgaat cggcgcaaggggcaaccaggagattggcgttctcctcggcatgatcctgagcgactgg cgcgcatcctggctcaggtcaagacgcccgtcatggatgcggcgcggtcagtcgac gcgctgggcaattacacgaccgactcgtcgggttagggctcccgtggaatgcggcaggt ggcgggagggacgaagtacaaccgctgagggcaggggtgcccaagtcctcattggtc gatcggcggtgctggtgggggtcagcacgccggagcatcgtgacgtgcatgctggtg ccctgcacatcaggcgcaacaggcgatgctcgcggcagannnnnnnnnnnnnnnnnn ggctccttcaatcagcacagagcagcaggaccgtccaaggatcagccatcgtttgt acgctcatcgcgctagcgatgggtatagctaccaacaaggaaaggagcggggtcctcc tccgctcgtgaaggcagacaggtctccgcaccaggttttgaagccatacggc tggtcgagggttcatcgtccctacagatgcaagaaatccagtgccacgatcgcgcatg acgatgtcgtggtgcggttcgggtcggcatttgcacaccgacgtccactatcgag cgggaaatcgctgtgcgacattgccaggacgctggggcagaggtcgcgggtgt ggttgcaacaggtggcaagcaggtaacatctgcaaggtggcgaccgctcgtcgtg cattacgtctgagctcggcaactgtcttattgagtcggcgcaacgaacattcgtcgt gcggggctctatgttggggcgttatgccgatggcggtatgcgaatactgcggctgc cggaaactaatccctgcacctaccggcggaatccctttgaacatggcgctatcctga tgtgttcgtctccaccgtttccacgcccgtcgcaagtcgagactcaagagtgccgaga ccgttgccatcttggcgtggggggactagccatcggcagttcagcttgcgtatgccttt ggcgcgctcgtatgtctacgtgtatatacaacgagacaagctagggtgcccgaaaa atacggagcgattccggtgaatggcaggtcgaacgatcgggttgcgagatacgcagg ctaccaaaaggcaaaagggtggacgttgcgctcgagatgatcgactccacagacgat gatgcaagccgtgcaatcgttggcggtcagggcgcgcggtcgtggtgggtatcagc gataaaccgcttgagatcgaactaccgggaactgttgcagggaagccgaagtc tggcacgagcgaccatctgttcacaggttgcacctcgtcgaactgacccggcgggg gaaagctcgacctttcagaagccgttacaagaactgtcccttgatgccggtgcgatca accaggtgctggataatcgaacacttccggcgagatgtcgcaccgtgattgtgcatga a	gtcagagaccgtcccgtcgaacgggagc ggcttgcgtgtgagacgcagggccccctc tgaccagctcagccaccgtctcgttccg cgaggctgacgggggtcgttgaagcga atgcatagggaacgtccgggtcttccacc cccgaccgtctcgggagcattgaagcaa gctgacgggtaaagccagtgctgtccac cggaaaccgcttcgaacgttgcagaggtg agcattacctgggcaaccagagggccatgc gggcacatcacataaggagtcctgtatgca cag

SEQ ID NO:	Name	IscB sequence	ωRNAs
1397-1398	Ga0370494_000001 GENOME_ACESSION: Collapse_03_16_2_\$F_3 300034130 GENOME_ID: 280058 CONTIG_ID: 0 SOURCE: JGI DATE: 2019-06-13.	ttgtttgtattgtcgttaatacaatcaatcatgccaacaacaccagtcgccc aaacgctgattcgtgacggcaaaagccaccattctggaagcgtggtcttttgcattc gattaaatttcgatccatctgataaagagtcagaactgtcgcctgtaggtatcgtatcgtgc agcaagaaggaggtatcaccgtcaaatctcaagccafacctatctcaacctacaagca gatgccgtcacttgggtcaaaagatgccgtctcaactgcagagatatgcgcaaaagcag cagatttcgcaaaacacctgcaggaaaaatcgcatgaatcgcaaacgtggtgtcttcc gccttcgacaaaagcccgtggcaatggaaattgcggtatctcgtcatctgtgtaaaatc atccccattgccgattgtgtggtgaagacatcaagccaaactaaaggacagcgcaa atgggatgtatcttctcacctttacaagtgcgcaaacatgtgtctatgaccaattagaaca aattggctttgtcattgtgtggaaggctgcgcaaaccaagaactgcgcgatagtcagggc tatccaagactaaaaacaaatggctgaggtatttgaggcgcattgtgttgatagttgggtg ttggcgaattgtgtgtgcggcgtcacatcaaacggataactgcgatgtatgtattgac gccattacgatttcacggcgtcaattgcatgttttcagccgatgagtggtgaattcgtaa atcctatgagcagtcgcaaaagtcataattgaacgcggtagttgttaagcatatcaaat tggaaattgttatattgtgtgttatgaaggatagaattgtttacatcatgttaagatggg acacgtttgacacaaaccgcaaacacagttgagtgcaagttttaactataacacatggc gaggcttcccccccccgaaggacgggtctccgctcgcaaaaagatga	gtcaattacccccccctgaaggcggggct tgttgaataatgttccacacatcttgaagga caagattaaaagcaatcgactagccgaatat gtgtgaagaccacatgacgaactgaattga ctaccatgttaggtggaatacgtgttaga tgtcattctagtcgacacctctatgtcagtc gtggcgaagaatagatatactcgaagga ggactatctgaactttaaagaactcattt gttt
1399-1400	Gammaproteobacteria bacterium isolate CSSed10_2 CS-Sed10- C100006, whole genome shotgun sequence GENOME_ACESSION: GCA_003560635.1_AS M356063v1_genomic GENOME_ID: 245344 CONTIG_ID: 0 SOURCE: NCBI_Prokaryotes DATE: 2019-05-30.	atggccgtttgtgtattggacaaaagaacacggcgtgatgccctgtgcggagaacag ggcacgggtgtctattggagcggggcggtgtgtgttcacaggatgcgtccgtttacgat ccggctcaaatgctgacgtggaggaatcgaacctgcaaccgtccagataaagatcg atccggcagcaagaccacgggggtcagcgtgatccgcggcagacgctgatccag agcagcagcaggtcgtgatgctgacatcgaatcgcggccaatcgtatcgtgcg gcctgtgtcggaccgtcagcggcgtacccactggagcgtgccaccctgaccgttcg ggccacggggcgggcagttatcagcgcacccggctcaaccggtttggttcccggt ggctacgtgatcgcaaaaacagtgaaagggc	gtcaactacccccgcctaaaggcgggagct tgtgaagcaagcgtgggtgaccagggaga gcggccaatcagccgtacgttggaacag gtcgtcaagaccaccagcgaatgtgcct cagttcgtcgtctggaagacgggacatc gcaggcgaaggtgaagcggcgaaggttc ccgtcggcgcgaagcgggagccggttgc cgacctcccgaggggagccggggcga agccccgcgtcactgagctgtgaagggtg tatgtttggagcattgagatggcc
1401-1402	Trichococcus palustris genome assembly, contig: TpalDRAFT_scaffold-1, whole genome shotgun sequence GENOME_ACESSION: GCA_900067125.1_AS M90006712v1_genomic GENOME_ID: 181408 CONTIG_ID: 4 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgttattcgtgttaagcaaaacaaacaaatggacaactgtacgctgcaaaagca cggatacttctcgtgacggctcggcaaccatacaagcagtaacctttaccatccggt gaaggacaatcgccgcacaccgcccgaacaaatgtatcagtaattggaccgggc gcaaaataacagcgttgccttagtagacagtgagcccatgcggttttcttggcaac tggaaacccgaggagcgcgttgcctttatgcagacgggtaccgtaaaccaagt ggggaatagctttaaaagaaggaattctaaatcaacgcggatacccgagaccggag ggatggctaccgcttccgtgtgtcgtgattgagcaggatatactgacatttcgtcaagaaaa tgcgaaatctcgtcattgattcagttggcggcgtggaatccgtcaaatcgtacatgcagaa gatggagaatgttccattacggacgtgtgtctaccagcaaggcaggttgatgggttacga aatccgcatfaccgtgtgaaaagaagggccatgcctgtcaatattgcggtggccttcc caagacaaacatctggaagtggagcatatgcatcccaagtcgccggcggttccgaccg cctgagcaactggaatgagcgtgtacatgagcaatcaggataaggaatcgcacgct ggccgaatacgtggagcggtaaaaagctccaaacaaactcgaccagacgcgcat aaacggatagagcaaatcctaaggacaaacaaacgttcatcggctgcgtatgcggct tggccaacagcagcgcacacatttgggtggtgacttagaaaagtgttgcacacataa gcaaggcaccggcgacagcgaatacaaccgcagcggcagcaggttggcg aaagacacattatgacgtttatgcgttgagcaatcccttccagcggtatcgtttgta acagacaagggtttgtgcatcaagtcctatgagcagggcagccgggttccggccggac gaacagttgcggcatcaccagcaactgacacgcaaaagcagttttcagcttcag acggcgacgtcgtccgagcaaccgtgccaacggcaagaaaggaagcattcactc ggacgagtggtgttcgcaaaagcgttatttcaacatcgaagcaccggattagttgtc aagggttagttacaacattgcccatcatcgaaggaatgacggatacggctatactt aaaacaaaggagttga	gtcaactacccagcagtaagctgtagctt gtaaaagccttagtgagcagactcagtttcc gcttggcggagctacgttaattagccatca tgttctctgtgtctccagcaggaaacgic atgtcagcagcttaaacagtcgaaggggt aggcgctgtcgtgagtgtaaaaactttt aactttgtcagggagagagaggccgaacg gcctccgttacctgtctgacgttcttcga cagatagagatgttttttattgttaaaat aaatgcaaaaaggagcgtgaacgaattg
1403-1404	Ga0209167_10018693 GENOME_ACESSION: CoaSoiSPAdes_4_\$F_33 00027867 GENOME_ID: 24548 CONTIG_ID: 18692 SOURCE: JGI DATE: 2019-01-27.	atgtcggctctgttctggataggcacaagcggccactcatccgtgttcggagaacga gcacggctgctgctcgtgtaaacgcggcgtgttccatcgttcttccctttaccatccg gctgaaagagcgcacggtccaggagagcctggggcgaccaatcgtgctcaagattgag ccaggctcccgaaccacgggatggcactggcccgatgagcagactgaggaaggc gaagtgcacacgcgtgagctttccgaactcaccaccggcggaagcagtcacatca cgccctttgcaacgagcaggtatcggcgacggaggagatccgcaactgcgtaca ggccagcccgttcgcaactgtaaacgcagagcggcgtggttcccccttccctttga gcagggtcgggaacgtgttactggacgcagcgtatcgcgatgggttactgtatg cggatcagggtgagcgcgtgaagtgtcgtatcgtacgtccttcaaaacccagaggtcgc gggctgcagatcaaaaggcgaactgttgggtggagatcaggagctatctgctgg agaaattccagatcgtgctgtactcggggcgcaaacaccgcttgaactggac catatccggcctgttcgcggcggaatcaatcgtgtcgaacttgcgtcagttgcc	gtcaacgacccacagtaaacgtggggg catgctcccagtttacggttccctgtgt ccagcctgagttcgtatcctagaggcagg gagctgagggaactacgttggcaggtcat ggcaccttgatgccacctcagtcagag cactgcccgtggtgttaaaagccctcggg gtagggtcgggtcgccagccgaacaa cgtgtggcaacattggcgaagggttgc ccgggcaacccgacccagcgcttacta ggccagtaattgggctcctacacgaggcaa aacggaaaggagtgcttccatgtcg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		accactgcaacacatcaaagggatcgacagccagcgaattcgccaccgggaagt ggaagcccgccaggaagccacttcagacgccgctgcgtaacgactcgtttc gccctggctcaggcgctgcgtctctc	
1405-1406	Ga0222637_1000159 GENOME_ACESSION: AntAcemetagen143_2_\$ F_3300023435 GENOME_ID: 23014 CONTIG_ID: 158 SOURCE: JGI DATE: 2019-01-27.	atgcagaaggttagaaaagagaatacatatacacctacgggtgctccacaagccgtag caactgtgacctgttaattaacagagctgtggggtaggctcgggtgtcggggttataaa accttttcaacggctccgtagtggaaccacttcagcatacaggaggcttaagctaagt tgtatgtatatgtaataaacatgatggtttaccgttgatgccatgcaagcctgccaaagctc gtcactattgagagacgggaaagcattggtagtgaagaaaacctttattatcatgttgt tgtgggattgtgaattatagacaaccgggtgtattaggcttagaccctggtatcgtatata ctggattttctcgagggtcaggtaagaagaattgatttcggggacagttgtggggagaa ccgatatccctaaaaagatggagcaaacgaaattatctagaactcgaaggggtaaat tatggtatcgggaaccgagatttgataacagaacaaggaaaagggtgtgtgctcctc gattcggcacaaactgattcgtttgtaagattgtttgggtatatacaaatatttaccgata acgaagataatgttgggtgacagtttgcataacagaagataaagaatcctggtatca gtgtcaagggtatcagcgtgggtgttgaaggattttggaatttcgagaatacgtttg catagagacaatcatttatgccagaagtgaagggaagtcgaagataaaatactcag gtacatcatgtacacggtaaaagggaaggtgctactgtagaccgtgaggagttgttaactg tttgcacaaacatgtcataaaacatcataaaggtaagatttgattcctgataaagagata aagagattaaagctgaacatttatgacgatgggttagatggaagttgtgaataaattgata gatataattgtgacattgtttctacattatgggtatatacgaagtcgaatagaattggatt gggattggaanaagtcattgagcgtgacgtgtcaattagcatgggtagaagtataaa ggaataccattaccgatatattataatggaccgaattatcgaaggtgtaacttttgatg tgaacacgggtcagggaataatcgttcgatacagattaatcgaacgggattcaggccct cgatacggagagcgtatccgttccagcctcatgacctggttagataaacggtaagg aacatcgagtgaaagcgagtcattgttatggagcgagattgtattagataaagaatc agtggatgtgaaccgagtggaattgataacttatgttaaaggattatgttacgcttaa	gtcagataccaccgctaaagcgggtggc ttagcctactgtctgattaggagcattaaac atgcagaaggttagaaaagagaatacatata cacctacgggtgtccacaagccgtagca actgtgacctgttaattaacagagctgtgg ggtaggctcgggtgtcggggttataaaacc ttttctaaggctccgattgtggaccaacttca gcatacaggagggttaagctaattgtgtat g
1407-1408	Ga0070698_100047374 GENOME_ACESSION: KBSK1StandDraft_6_\$F _3300005471 GENOME_ID: 30496 CONTIG_ID: 47373 SOURCE: JGI DATE: 2019-01-27.	atgtcgcacatctttgtccctgtggtgcatcagggaacacaccccgctgatgccaccagg ccatcgcggggcacggcgctgtagcaagagcggcaagcgcaccactcttgaaaggg gggggtttctcgtcgtgcctcaacaggcagccatccgcgcgagcagcaacatcag cgggtggcatcgtatcctggttccaacgcgaggggcctcgtgttgccggagcttccac acctatctgaatatccaggctgaggccagagacggcgctcaaggaaagcggagaggac agcacgcgcatcgcgcgacccgtagaggcgccgaagacccttaccgcaagccgcgc cagaaccgcaagcagagcaagaagaagctgccgccctccaccagggcacgttgcac tggaaactgcgcctgagcagcctggctctgcagactcttcccgtagcgtcttgcgtg gaggacatcaaaagccatgacctggcgaanaagcgtgggatcagtccttttccccgtta gaagtgaggcaagcactggttctatgcggaactctcgaagttgctcctgtccagatcaa cagggtctaccagaccagggaactgcgcgagcgggttggactgaagaagacgcacaa gaagttggcagaggtgtgggaagcgacgtgtgctgatgctggatcgtggccacagc gcagtaggagggcggaagtccccagacaatcgccgggttgcctgcgtcgcgccactga attggcaccatcgcgaactccatcgctcgcagccagagaagagaggcaagcgcaagcc ctacgggggcacacgctcgtgggatcaagcgttggcaccctggtcaagcaccgaa gtggggcaaaacctatgtaggtggcacgatggacgggaactgagcctgcacgacca caaacgggcaagcggctcaccacatggccaagctgcgcgattgtcggctgatcaagct gttacggtggaagcgcgctgtgtgctctcttccccactaaataa	gtcaagcaccatcccttgaagggtatgg gcttatgaagggtatccggtgtcaccggg gaaccaacgagggaactgcgagactagc ttgccaaagctctcgaagaggcggttgaa ggagactgaacgatctggcggtgacagcc caccggatgggtgagcgcctccctaactg caccgcctccgctcagtagcgaaggga ccatagcacaccgaaagggttatcgca cat
1409-1410	Ga0070739_10029292 GENOME_ACESSION: CoaSoiStandDraft_10_\$F _3300005532 GENOME_ID: 24555 CONTIG_ID: 29291 SOURCE: JGI DATE: 2019-01-27.	gtgctcacgtcctacctgcacgtgcgcggccagcgtgctgcttagaactctggtctaa acagaggttaagggttaactcagtgctcagaagatagccgctgagaacctgtcgag gccaactttaccctggcaacaggaggtctgaaaggagcaaaacagcatgtcgaact ctttgtcgttgatccaataggacaccgctagaccgggtgcatccaggctacgcacgcct cttgctacccaaagggaaggcagcgggtgctgcgcgcttcccggttcacacttattctaaa gccgtgtgtgagcaaccacaggcagagcggctgcgcgtcgaactgatccagggaagc aagaccacggggctgctgctgtaacgagaccacgggggaagtcgtcttcgccgca gagctctgccatcggggagcggcgagcaccagaccctcgtgagcggcgtagggta cggcacagtcgaaggagtcgcatatcgctcagaaaagccgcgcttcgcaatcgca agcgaaggcccgctgctgccccgtcttgagagtcgagtagcaacgtgtgtgac ctgggtcaagcattgctgcgcatctgtccatctccagcatttccaggagctcgtccg ttcgatatcagggcgtccaacgccagaatcagcggcattgagtacaaaggggac gaaggtcaagggtttcagaccgggatatgtgcgtgcggtgtgtgaggaagggaacc aaagtcgggacctataccggccgggtggtatccgactcgtggtccttcaatattgcc acagcatgtggagcgtcaagatatcagtcattgttttgcagggtctgcatcactgtga tgggtatagctatcagaaggagagcgggcaatgcctcccgccccctag	gtcaggaaacctcccgtagaacgggatag gcttgtataatcaagctggaactgaccaga ctcagcctggcaacgggctacgttctagg gagtgctacgctcctacctgcacgtgcgc ggccagcgtgctgcttagaactctggtt aaacagaggttaagggttaactcagtgctc agaagatagccgctgagaacctgtcga ggccaactttaccctggcaacaggaggtct gaaaggagcaaaccaagcatgtcgaac
1411-1412	Sutterellaceae bacterium Marseille-P2968 strain Marseille-P2968T genome assembly, scaffold: scaffold00021,	gtgcttcagacaacaccccttccaaacttgccgaaggatcaccacggcgaaggt cgagcagacggagccgtgaggtatccgtcaatggagaacactttgaaagtgtttgtgtg aatatgcgcgggcaaccgctcatcggctgctccccgccaaaggcgcgtgaagctgcttcg ggccgggaaggccgtgctgtgcgccgaaccctctgtgattcaactgacgggtgccg atgggcgaaccaagcagccgatccttgggtgtgagtcgggctacaagcagctcg	gtctactacccctgactgaagtcagagcgtt ggcaacaagctgtgttgactagcctcagtc cgttttcggacggactccgttgggttggaaatc cgttccgtcgaagacgggacgcaaacag gcaccgcgggatgtcgtatctagtcgccg

SEQ ID NO:	Name	IscB sequence	ωRNAs
	whole genome shotgun sequence GENOME_ACESSION: GCA_900128485.1_PRJ EB18048_genomic GENOME_ID: 186107 CONTIG_ID: 20 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	gcatgagtgcaacgaccgccaaggaagattgttggcttccgaggtcgaaactgcggca ggagctgacgggcttgcctcgaatcgcttggcacttccgcccgcacgcaaccgca agacccgctccgctgcccgcgttttgataatcgcttgcgatcaaacgcaagggttggc ttgcgccctccggtggagaacccgattcaggcgacacatcgcgcacgctgcgagggttgc ccgaagacgcaattgcgggacgcttttgcattgctggcgtcctcgacgcaaacgctgc ggggcagtagctgtttcagaacacagcgcggccacaaatgccagatccacaagctg acgactctcaaaaggccgtgtgcgcaagcggcatcagggtccgtacatgtgcacggctt caggctctttgacaaggctgtgcaaaagggggaagtgcgtttttttggcagacgctcct cgggtgcattcgaatgtcggcgtctgacacacaaagatttctgccgtatttctacaa gaaactttcgttcttgagaagcgaaagatgttttaactgaactgagaaaggaggccgc gattcctcccgcgtctga	ctctgcggcttgcgggttaaaagctctgagag gtaggagcgggttctcagacaacaaacccc ttccaacattggcggaagatcaccaccgac cgcaaggtcgagcagacggagcccggtga ggatccgtcaatggagaacacatttgaaa
1413-1414	fermentation metagenome genome assembly, contig: NODE_189_length_9351_1_cov_123.725550, whole genome shotgun sequence GENOME_ACESSION: UOOR01.1 GENOME_ID: 16049 CONTIG_ID: 188 SOURCE: WGS DATE: 2019-01-27.	atgcctacaactcaaaagaaaagcaagactattgtttaaacaaggaaaagctaaaatacac tcataatagctcttttaccatccaacttaataacggcaacagggcaacaaaacaaagacataa ctctaggtatagatactggctctaaactataggaatttcagcaagttaccacaaaagtaga attatattctgctgaattggaattaagaactaatatagtcgaattattatcaactaaaaagcaa tatcgaaggagtaggcgaatatggaagacagatatagaaaagtaagatttttaataagg gcaaaactcaaaacataaaggatggttaactccgctattgaaataaaatcaggggccattt tagaattgtagaaaaagttaatcaactctacccaataaatagaacacatagtagaggtagcaa gctttgataacaaaaataaacaacccagctatcaaggaagaaataatcaggaaggaa atcaattagggttctggaattgtagaataatgtgtatttagagatgttataaatgtcaaggt aagaaaaattgtgaaggcgaagatactaaacgtacatcagatagagtcagaaaaacagg cggtaatgcacccaataaattgataacttttgaagattgccataaagattaccattcagg aaagctaaaaaaggccctttaagcggagggaagagcttcaaggattcgacattcatgggat aatgcgttgactttttacataagattgaaagaaatatatctaatgtaaaaatgacatattg atatataactaaaaatacaagaataataataaactgaaaaagccacatagataatgatgca agatgtataagcgggaatccattggctgaagaatcacatgtttggtatcgtattaaacaggt agaaagaaagaaagaaagccctcatgaaagcaggttctaggaagcgaagaaatgccaa atagagaatctaaaaaagaaatgcaaaaacactaaggaaataatatacaaaaggaaagaa tggtgcctgtatgatgaagttgactaaatggagatcgcgggttatcaggaatttagtggc aacatggtatatacaagataatattggaaatatattcaaatatctccaaaataaagcaa ataagcacaaacaacatagaatttaataaaagaaacaacaaattatatttagaggtatcg cctaa	gtcaactaaccccccttgcagaagaggag acttataggcaattttatgaagcttagttgt ctagcttaagtgtctgtcactacgttggattg catgcagaacctgcaatgatactctagttt gtagctctttctggctctgtaaaagttctgag ggcaaggaaacgtcaaccacattgtgaag gctttccacattggcggaagggttaataact ctgaaggagaacgttaacttgaga
1415-1416	Ga0373625_0031164 GENOME_ACESSION: YL18G_12A_MG_FD_\$ F_3300035701 GENOME_ID: 312749 CONTIG_ID: 31163 SOURCE: JGI DATE: 2020-04-07.	atgttagtatacttctgataaaacaataaaccattatgccttgcgtctcagaatagcc aagtgttattgaagcaggcgaagctgaggtgtacagagaacgccgtttactatfaaac tggattttgtagtagcgggtataaacagatatcaacgctaggcgttgatgctgctctaaac acgattggattatctgctattgatcaaacctccgttgaatcaggatcaaaaacacggacaa gaaagaactattttcggcagaagtgaactcagaactgatatagtatgatttattatctactc gagacagaatcgtagaacaagcgaagccataagacaggtatcgtgaagtaagatttc taaatcgagaatagataaagggttggtagcacctgatataagcacaataatagatactca tttgactatggttgacaagatacataagatttaccgataacaaatatcatagttgaggtagc acagttcgatatacagaagataaagaatcgggatataatcaggtatagaatatcaacaaggt gcgcaactagatttttcaatgaagagagtagctcttatatcgtgacgatcataaagtccaa ctatgcaatggcaaaagtaagacagatatataacgttcatcatatagaagcagaagaa ataggtggtaacgctccaaataatctgattactttgttgaacgctccataaaaagcatca cgatggttaataataattaaaggtaaaagggtcagttgttcaagacgctcttttatggg tattatcggtgggctttttacaatacactgaaagataaatactctaatttagttgacttatg gggtatctaactaaaacacaggataactaacaagctaccgaagaatcatagaacagac ctttatgtatcacaggttaatccaacagttaaaggcttctgatacattgtttctataaaaagg ttagatgtcaaaatagacaaatacacaagacaactattaataaagggtgggtcaagaaagt aaaccaactgcctttatctgatttgatatacgtattttgataaggtaaaatacaacaatcaa gagtggttttttggtaggcgtttaagtggtattttgatatacgttacttgacggaacaac ggtcaacgcaggtattagctataaaaaggtaaaaggtttagagcaacgtaaaacagtttta atcgaagggagaaggcaattcctcccatgcctgaaggcagggggtatccttgcggaaa cgagagatga	gtcaataaccccacgcataaatcggggggc ttgaaaaagcctttgtgactagcctaagtagt gatcgggaatacccgatcttctactacgttgttt tggtcataaacctcggaattatcagcttagt ttagctctgttctctaattaaaaactctga gaggttagggcggtgtgttaggcgtgacaa gccattacaacattggcggaagactgtaacc ctgaagggggtatattttgtta
1417-1418	Ga0114939_10003065 GENOME_ACESSION: AshMeaCrysSpring_2_\$ F_3300009455 GENOME_ID: 23357 CONTIG_ID: 3064 SOURCE: JGI DATE: 2019-01-27.	atgcagacgttagaagaagaataacacacagatagatgctccacaagtcctgc catctgtgatcaagtaataaacaagatttaagtgccaggaaactggttgcggcttagtg tactgtatttaaaactcttctaacttccgaaagtgagacactcagaaatgagagaacag aactcgagagctctctgttaaatctaaaactatgagatcattttatcagctaaagctgaaaaagt gttaggtctactaaagacaaaaatcagaccatcagattttcaacaaatatacacaag gaatgaccttcgctcaaaaattgaatacttaacatgctgtagcagaagaggaagaaat catagaaaaggtatttaattcaagaagaaaggagaaaagcagaagaaaagagcaaa gcaaaagatagaaaacaaaagaacctcaaaagtataagaacaaaaagctatgtattc gtattgaatatgagaggacaacctctgatgcctcatcacaagaaggaagcaagactctat tgaataaataagaagctaaaggtagtcaaaaagatgctatatacagatacagtttaattac	gtcaactaccgccccctgaagggggtgct tgcacagtaatgagcaaggttaaaaagagtt gattaggaggcatttaaaatgagcagcgtta gaagaaaagaataacacacacagatagat gctcccaagctcgtcgcactctgtatcaag tattaaacaaagatttaagtgccaggaaact gggtgcggtcttagtgacttgatttaaaact tcttctaacttccggaagtggacttactcaga aatgagagaacagaactcgagagctctgt

SEQ ID NO:	Name	IscB sequence	ωRNAs
		<p>gaccggagaaaaaacaacacattatagtttaggaatcattcaggatagaacatgtagg ttatctcgcacaacagataaaaaagaattatttctcagatgtcacattaaggactgatatt cctgataagatgactaaaaaggctatgtatcgaagggaagaaataggaaactatag gtatagagaaccgagatggttgaatagaggattccagaaggatggttggccccatcaat ccaacataaattggacagtcattgtagttaataagataataaaacagataattaccatcaa taataatgataaactattgaagtggtctacatttgatacacagaagatgaacaatccggaga taagtggaaatagaataccaacagggaactcttcaggatagaagttagagaatatctatt agaaaaatggggcagaagaattgtcactatgtaagtagagaattgtaccatttgagattga acatatcattcccccttcgagatttgattatgtggaataatagaatccaatctaacgata gctgtcattgattgtaatacaaaagaattgtaatatgacagcagcagaatttggcatccaga agtacaagatttggcacagaatactctaaagatcggcattcatgaatttggatgga gatttagcaattcaattcctgaattgtcatcatatgattgtactaaatataatagaatt agattaggacttgaagaagtcataatgaatgattcgttaacgccaatagaacaatca acttaatacaagaagatgtaaatcgttcgaggtaaacagattagacgcaacaatagatca ttacaataatagaagaaggcttcgaaccatcaataagaagaagaatataaatattctcc tggagatctagctctatcaagaataatcacagggtcacgtctcgttgaagaaga gtttcaattatggtgaatggataagattagtaaatcttattcctggagagaagacattagc acaaatattaagaattgaagaatagtaaatatggaagggaattcgggtctcatatccagat tttctgattaatcctgatattgttaattgaagacaaggaaatagaagaactattattagg acattgagaacatgagacgacgatacaatgatcgtatttggagaagaattataagac aatcaagaataatagaagaagaacagttatcaaaatatacaaaactaatatgatctgtt ggaaagatcattaaacagaacaatcaagaataaactaagaagaagacggagactgt tgaatcaattgattcgacaacacaggtaggaattgataatgcatggaattga</p>	
1419-1420	<p>uncultured Erysipelotrichaceae bacterium isolate RUG813 genome assembly, contig: k87_14161652, whole genome shotgun sequence GENOME_ACESSION: GCA_900313495.1_Rum en_uncultured_genome_ RUG813_genomic GENOME_ID: 192862 CONTIG_ID: 396 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.</p>	<p>atgctgaaaaacggcaaggctaaattatcggccatgaccggttaccattcagcttcagt atgagtcggaacaacacacgcaacctgtagaactcacagaagatagaggctatcagaac atcggactttcaattaagtcggaacacagaatttgaagccgggagtataaactgcttcc ggacgaaaagaacaggcatcaggcacagaagcggtattctgacgcaagacgtaaccg gaagcggtaccgcaaacctatgtcaaaagcggcggaatgttccgcaggaacagaaa gactggctcgcgctgggtcagaacaaagcggatcagacgttgcattgtgaaacg gtttatgaactgttccccataacatcgttggctgctgaaatgggacagttcgataatgcag tgcttccgctgtaaatcagggaattgcggctcccggaaggcggttcttatacaatggacc taaatcggattcgacagcttgagagaagcggtatttggccgcgacattatcgtcgtat gctgtggacgctcgtgatgaaagaccatgtgacgcttgcattcatcacagaggttatcg caaaggtgacagatcaaacgtcttcaaaccttgcagctgtatgtcgagagaccataacc tcagcggaaacacaagaaggcggttaaaccttgaatcttcaaaagacggagggtctct ggcgccatccgcattcatgaatcgggtcaaatgggtacatctgggactgtgtaagcaggctt gggatggagacaagattacatattgcgctgttacaacacgggaacgtctgcagaggaa cattatcaagtcacatgccatgatgcgtattgtattggcgtatttcatcaaaacatcgtact catagcggatattacagcaaacgcagaaggaaatgacgctgctgcaaaagtctatgatg cgaagtataaagatgcagctacgggtgctgcatcaaaaggcgctggaacttggatgtaac gcacgaagagaatattccgcgcaataacgaacgcaatctgcgtccagtcgcgcgga aaaagtaagctgtgatacaacagattctgtaaaagaaaggcattcactcaggcgaggcg atatgttctagtaaaaggacgcaggtatgaagtgaagacgggtacgatttaa</p>	<p>acatctaccgtgataactggtcctttcatagg aacctccctcatgcataaactcgtgcaaga cttccccgaccctgtatcagaccattttctg gattacatagcgatcgcatgtgtgctgt gcgtgctcattcgtataaacgggaatcgat acaagagggaagcagaacgttttaattgact gactcgcggcattgcgggtctaaaaagta acacgtccttacttaccaggccagctgtgac tgtatgcctacagcagaaggcagaaactg ctg</p>
1421-1422	<p>Ga0307376_10002158 GENOME_ACESSION: TR2MG_2_SF_33000315 78 GENOME_ID: 281014 CONTIG_ID: 2157 SOURCE: JGI DATE: 2019-06-13.</p>	<p>atggttatgtactcgacaagcgcaagaacccctaatgccctgttccgagaagcgggccc cggctgctcctggagcgcggctggggcggtgtgtccaaagctggcgccgttcacgatcc ggctcagggaaccgctggttgaggacagtgttttgaacccctcgtatgaagctggatc ccgggtcgaagtaaccggggcgcggtgttccggagcgaaggaaccatcggc tgctgtgagtgtctcaccggacggatataaaggcgaacctagacccccgcgcagcc agcgtcttccctcggaaccgggaagactcggatcgggaagcgcgggttcgataaccga catccggagaagtgcgcggcctgcggcggaacgcaaacacgggagccgggtactg ccggccctgcggcgagaccgggaactcgttaaacacggccatcgggagccttgctg ccggcgtcgtccggcccggttgaggaaactatgtcctgggtggcaaatgagcga agctcctgccattaccggggtcgcgatggagctggctcgggttcgacaccaatggatg gagaacccgaagatctccggggtggagtatcagcagggtatgtcggcgctacagg tccgagatgtctcctgcagaagttcgccacctatgtgcttactgtcggggtgctccgg agaccgggttgaatgtagaacatgtgtgtccaaaggaacccgggagccggggaag ggtacggaccggatcagcaacctggtcatcgcagcaagacgtgcaacgaagcgaagg gcaacctgcagccggaagaatggctggaagagtgcggcgctcgggaaggctctgg accggatccggcggaacacctgcgaagacgtcacaacagtgaagcagcctctgaa ggatgcggcgatgatgaacgccaccaggtggcgctgtatcaccgtttaaagcgctcg gccttccctggaaacagcgacggcggtctgacgaagttaaccggaccaggttatg aagctaccaagaccacttctatagcgggttgcgtagggaagcctccggaaac gggtgatgtccgttcgtcaggtctacacgggaacggcggggtaaccggcagatg gcccggatcgacaagtacggttttccgtaccgggtgcgggagcgggaagagggccac cttgcttccagaccggagacgtggtgcccgtcgacatcccgaaggcacaatacaag</p>	<p>gtcatagccccagcgctaagccggggt tgaggaggccctcctccgacagtctctggtat ttaggcgagaagcccgaaggcgctat gaccagccccagctcggggtccccga cgagcagactacgttattccgggttagggca ccctgggatgcttccagtcggagagctg ccggcgcggttaaacagcgagcaggggt ggatgtaacaatgcacgcagcacattcgtg gtaatccccgtactttatgcagcagtgccgt cggcgcaaacggggataacatcggcgag gagagatgaccacctcgggtgggtgccc cggtgccgaacccggcgcggaacgggtcc cgttaaggatctccgaaggagagaagg gatgaaatcgtg</p>

289

SEQ ID NO:	Name	IscB sequence	ωRNAs
	_3300023210 GENOME_ID: 33123 CONTIG_ID: 62 SOURCE: JGI DATE: 2019-01-27.	gtgtgcagtagtttcaaatgaaaagggtttgtatcagtcagaggtacaactaagacaagac gtttctaaaaagatgaacaaagattatgtatcgaggactagaagatgcagaaagctta gatatcgtaaaatgagatggcagaacagagcttctttaaagaagaaaggagactagca ccaagcattagaagtaaaatattctcatcttagagaaaagagttgtgaatctattttac ctatacaaaaatggatagtagaattagcttctttgatactataaaataactgatcctaactg aaaaggaaactgattatcagttaggcagtcagaaggattttacaatgtaaaagcttatgttt atatagagataattatattgtcaacattgaaaggaaaatacaaaataaaaagcttactgtt catcatgttatatttaggagtcgaggcgacagatctccataaatttaataactttgtgtg aaacgtgtcatgaacaattacatgatggaatftaaattagaagaaaacgttctaaaacc aaacatgctacagaaataggtatagtttaagtcacaattaaaaaagaagtgatagcctttg aagagactttcgggttcgaactaaatatactagagaaaaatgttaggtttacataaaagtc atgctaafgatgcagtggtctattgtgttaaaaatagtagaattgaattgaataatgataatgt ttattgaaaagacatgtttcgtcagggcattatcaacaaccaaaggtaaacgcagcgaa actaggattctacaggaagttatttgattagaagaaattgatcgaatcaaaacgactaa aggcggttggttcgtaaaaggaaaaagatctctggacactttgcttgtttactctagataa aaaaaagttcatctcatcagtttagtgaatacaaatgttagtaaaattcagctagaaaaac tactttgatggaaagaattgcaataatgaatag	agaaagaaattagaacgttggaatgccgtc ccagttccaacctcttgggtgcagattaaac agttctgtgaggcaggaacagtggttgaac gttaaaccttttattatccggcgagggaag tcgaattctataattactccaaaattatagatac gcactactctgtttacggtcagaggtatattt atgta
1431-1432	Gammaproteobacteria bacterium DM2 chromosome, complete genome GENOME_ACESSION: GCA_003072645.1_AS M307264v1_genomic GENOME_ID: 158616 CONTIG_ID: 0 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgcaacgagtttttagttttagtagtaagaagcaaccgcttatgcttgcctccggcaag ggcgccgggagctactcaaaagaaagaaagcgctgtcttccgggtccgtcacga taatcctgaaagagcggggaaggtggggaactacaaaggaccgagttgaagttagacc cgggagcaagacaacggggatggctctgactgttcacggggataacggcatccgcctt gtttgggtggcaacctatctcatcgcggccatgcgattagcgaagctgtcatcgagag cagcccgagcgagagcgagcgcaaccgaaagactcgtatcgccggcgcgatfcg acaatagagcggaagcccaaggctggctgtctccagcttgatgagccgggttcacaac tgcgagacatgggcttctcggctaacccgactcgtccggttaactgacattgccattgaga ccgttcgtttgatagcaactcatgctaacccgaagattgccgtgtcagatcagca aggcagcttgatgctatgaactccgtgagttattgttcagcgggacgggcacacttg ccgatattggcgccgcttccgggtgacctgttctgaacatcgaccatgtgcaaccag agctaaaggcgcgagcgatagcgtaagaacctagtcacctcatgccgaacctgcaatg aggataaaggggccacactgctatccgattgttgaaagctgtcagaaagtcacgctcca aactcaacaaggcaagggaaccagaatacaaaaggtcatagctggccggagcccttc aatgcgagatcgccggcagttaatgctaccgggtatcggtatggcagatgtgatgaagg cgaccggcctaccactactttctgagtgccggacgaaccaagttcaatcggtcgcag caaggttatcaaaaagaccactggttgatgctgttcgtggttgagactggccgcaa gtgtttatcccttctagtgctgtgcccgtggcaattaaagcaaacggtcacggctccggc aaatgacgctgcctgatagtacgggttcccaagaaccaagcgaaaggttgcatcg gttaaaagggttcaggactgggatgtagttaaagctgtcgtccgtgttgtaagtatgcag gcaaatctctagccgaatcctcgtgagaaaaacaggtacgtttagcctgcaagattcag cagggaagagatgtgagtcatagtattgcacgaaggttactcgtcgatgggtatt cgtatgcctga	gtcaactaccggccctaaaggaccgggct tgtaaaagcccatgttgaccagactaagcctt aactggctcatgttaggaaagatccatagcg accgtgggatgaccttcgaagtcaccagct ctcgggtaacagctaaccaggtgtgaagag gttaagccagtgctgttgccatcaaacctttc catacaattgtcgtgctgattttaccgggga aacccgagactgcaatagaggttaacacgat gcaacga
1433-1434	Sutterellaceae bacterium Marseille-P2968 strain Marseille-P2968T genome assembly, scaffold: scaffold00022, whole genome shotgun sequence GENOME_ACESSION: GCA_900128485.1_PRJ EB18048_genomic GENOME_ID: 186107 CONTIG_ID: 21 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	gtgcttcagacaacaaaccccttccaacattggcgaaggatcaccaccgaccgcaaggt cgagcagacggagcccgagaggtatccgtcaatggagaacactttgaaagttttgtgtt gaatatgcgcgggcaaccgctcatccgtgctcccccccaaggcgctgaagctcttc ggcccggaagccgtgctgtgcgcgaagccctttgtatcaactgacgggtccg acggcggaacaaagcagccgatccctgggtgtggatcggggtacaagcagctcg gcctgagcgcaacgaccgcaaggaaagattgttgcttccgaggtcgaaactcgccga ggacgtgacgggttctctcgtgacgcttggaacttcgcccgcggcgacgaatcgca agaccgctaccgcgaccacgcttcgacaaccgcttcgatcaaaagcacaaggctg gcttgccgctgctggtgagaaccgattcaggcgcatatcgcgcattgaagcggtttg ccgagtgcttcgatcacaatactgtgattgaacccgatcttcgacattcagaagatc aagaatcccgaagtcgaagcgacggactcagcaggcgagcagctgtgcttttgaa cgtgcgcaatatgtctgttccggcagggctacgtttgccaagcctgcaaaggcagatc gaaagatctgatttcaacgtgcatcattgagagtcgaaaacgggagggcgacgcgc cgggcaacctcatcagctgtgcgaggcggtgccacaaggcgatcacgaggaaggt gaagcagttcagtcgccggcggtcttcttcaggcgagagatttcatgggcatcat gcgttggaggtgctcaaccgctcgcgagcgccatcccgaattgcctgtcacgaata cctacgggtatctgaccaaacacaagcgattgtcgagggttgcgaagacgcattgc cgggacgctttctgattctgtgcgtctcgcgcaaaacgtcggggcgaatacgtttc agaacagacgcgccccacaatcgccagatccacaagctgacgattctcaagcggt tgtgcgcaagcgccatcaggtctcgtacgtgtgcaggttcaggtctttgacaaggt gctgtgcaaggggaagtcggcttatttttggcagacgctcctcggtgcatcgtg cgccgtctggacggtacaagatttctcggtatttctacaagaaacttctgcttctgag anccgctaccgcgaccacgcttcgacaaccggttcgatcaagcacaagggttgc ttcgccgtccgtggagaaccgattcaggcgacacatcgcgattgaagcggttgc gagtgcttcgatcacaatactgtgattgaacccgatcttcgacattcagaagatcaa	gtcaactaccctgactgaagtcagaggctt ggcaacaagtctgtgtgactagcctcagtc cgttttcggacggactcgttgggtgggaac cgttccgtcgcaagacgggacgcaaacag gcaccgcgggatgtcgtatcctgtcccg ctctgctgtctcggttaaaagctctgagag gtaggagcagtgcttcagacaacaaacccc ttccaacattggcggaagatcaccaccgac cgcaaggtcgagcagacggagcccgga ggtatccgtcaatggagaacactttgaaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
		<p>gaatcccgaagtcgaaggcacggactatcagcagggcgagcagcttgcttttgaac gtgcgcgaatatgtgctgtccgcgacgggtcacgtttgccaagcctgcaaaaggcagatcg aaagatctgattctcaacgtgcatcacattgagagtcggaaaacgggagggcgacgcgcc gggcaacctcatcacgctgtgagggcggtcccaagcggtatcacgggcaagtgtg aagcagttcagtcctccggcgcgctctttcagggcagagactttcatgggcatcatg cgttgagcggtgctcaaccgctgagcgagcgccatcccgaattgectgtcacgaatac ctacgggtatctgacaaacacaagcgcatcgtcggggctgcccgaagacgattgctg cggagcgtttctgattgctggctcctcgacgcaaaacgtcggggcgaaatcctgtttca gaaacagacgcgccccaatccagatccacaagctgacgattctcaaaaggcggt gtgcgcaagcgccatcaggctcgtacctagtgcagggcttcaggctcttgacaagggtg ctgtgcaaaagggaagtcggctttttttggcagacgtcctcgggtgcattcgtatgtgc gccgtctgagcggtacaagaattctgcccgtatttctacaagaactttcgttcttgaga agcgaaagatgttttaactgaactgagaaggaggcgccgattcctcccgctctga</p>	
1435-1436	<p>Candidatus Acidulodesulfobacterium acidiphilum isolate AP4 contig002, whole genome shotgun sequence GENOME_ACESSION: GCA_008534395.1_AS M853439v1_genomic GENOME_ID: 298711 CONTIG_ID: 1 SOURCE: NCBI_Prokaryotes DATE: 2019-11-09.</p>	<p>atggtaaaagtaattagcaaacggaaaacagcttatgccgacaaaagatacggcga ggtaaggcggtatgctaaaagaaggcaaggcggtataataatcgaaaaagccgttactat ccgggtattgttcgatactcccgaattgtccagcagcgaaccgttagcggtgacccggc gatactaccggtacgagtagcggttggaatacggtaaaatagtagaaaaggcggaat cagggttaagaaccgacgttaacttttactgtcggaagaaaagtattgagacgtataga agaacagaaataccggttacccgaagcgcaagatttttaaccgctgcaataaaaaagca ggacggttaccgtcgagcatagaacaaaagcaaggcagataggttaataaaataacgaa ctcgtatagctattatcccgtattacgtttaaataagaaatttaataaattcgaataa cttattaatcccgacataccggcggttgaaatcagcagggctctttatacggctatgaaa cgccaagcagtttttactcgttaagagaagcggaatgccagctttgtcataaagggttac aaagaaaatgacggctggcatattcaccatataatccgagagcgagcgggcgactaa taaacccgataatctgctgttattacataaaagcgtcatataaaaggacataagacagga gctttatacaaaactcaaaaaagcaagcagtttataatccggcgaatgtataacgccgtaa gatataaaacttatggacgaatttaaggctatttatgagataaagggttcatttaccacggct atcttacttctataaacagcgcgagcttgaattagaaaaagagcattataacgacgccata gctattacagggcataaacggaagtaaaagataaacgttattctataactgttattaagcaggc aagaaaaaagaacgttccctgcacgaagctacggcgaggaaaggcagaaaaagacc taactgctcattcggtcaggaacgctaaaaatacgaagaaattaccgtaaacggcga agtttgcatatgggataaagttaagggtatgcggtaaaaaggatatttccgggttacc ggaaacgctgctacgttcaggatattaacgggtgactatataaaaggcgagggcaaatcc tataagcaggttgcgcagataaaatagagcatataaggagggaataacaatttgataacg gaaagaatactggcgtaa</p>	<p>gtcaatcaccggccctttcggacggggtt gaaccgtaagggtcaagagtaattggtgtct aggctaaagtatcgagacataaccgatacta cgttatttgcgtgatataccctgacgtaattg cccaagcgtaagctctatcgaggctctgta aacaagaggaaaactcttagtaaacctcattt aacaagcgcttataaactgcccagtgaggca ataactccaaaaggagactacagtaactt tttagagtaagaagt</p>
1437-1438	<p>Ga0209647_1000722 GENOME_ACESSION: GraSoiAngeSPAdes_43_ \$F_3300026319 GENOME_ID: 26992 CONTIG_ID: 721 SOURCE: JGI DATE: 2019-01-27.</p>	<p>ttgagtaaaagttttgtcttagacacgaacaagcaacaactaatccgggtcatcctggga gggagagattactgtcactcagggcaagcgtccatattcaaacgtatcctttcacaat cattctcaaggttgccattgagcaaccagaggttacacctctcagaatcaagatcgatcca gggagcaaaaccactggcatagccatcgttaaacgacgcgacggcggaagtcgttttcg cgcggaacttgcgcacgagggcaagcgatcaagaagcggttagacgagagacgtg ctgtgagacgggtcacgacgccagagaaaaaccgctatcgcaaggcgcatgagcaga acagcgacggggaaagggttgctgctgcccctctttgagagcagaatagcccatgtt ctgacgtgggttcaacgctacggcgtctgttcttcatcgtggcaatcagtgcaagaactggt caagtttgacttgaagcgatggataatcctgaatcagcggtcgaagtatcgaagcgaag aacgctggtgctgctatgaggtcagggtatctcgttgagaagtgaagcgcatgtgcg cctactgtggcaaaaagaatttccctctgcatagagcacatccagccgcgcgcaaa gacgggaccatcgggtgaagcaacctctgttagcctgtgagcaatgcaatctcgccaaa gggatacaggatatacagggtgtttctcgcgaagaagccagaggacttgaagcgcatctt gctcaggccaaagcaccgctcaaggatgcagcgcggtgaataccacgcgatggcg ctctccgagcgattgaaggaggcggtgctgagcgaatggaagcggtggttga ccaaattcaatcgagcacgcgcaacctccgaagatgcactgctcgtatcgcccaat gtcgggaagagcacgccagagaagctccagattcgggggatcatgccccttgcgcatca cgcccaatggccacgctgcccagatgtgcttgatgatgagacaggttccccgc accaaaccacaagaagcattttacgcatgggtttcgcactggcgataatcgtgcgcgcc atcattctgctcagctcaaccatgcaggggtgcatgtgggaagactgtccgcaaggcc aaaggcggtttaccatcgccaccgcaaaaggcaaaattaccgggtgtgggaagaagt attgccgctgctccaacgagctgacgtgtacgggttacgcgcagcaatatgcgcctacg ggcgtctttctccctcagcctgaaggcatga</p>	<p>gtcaggaacccaatcccccttttaggggatg ggctgtgtgaacaggtcacctgaccaga ctcagctcgaagagctacgttaccggaga aattaggtacgttgcgggtgcaggccagcc gcagcctctacggcaacaattaaacagag gtacaagggttaactcagttgttttgcacta aaccttccagtaacctgtcaggccaccatt acctacgcaagttagagcgttggcaacaagc aaaaaggaaccaagcattttagtaaa</p>
1439-1440	<p>Chroococcidiopsis thermalis PCC 7203, complete genome GENOME_ACESSION: GCA_000317125.1_AS M31712v1_genomic GENOME_ID: 47152</p>	<p>atgcaaaattacgtttttgtgtgacacaaataaacagcctctcaaccaatttctccagca agagctagagaattgtgactaaacaaaagccgtgtatttagaatgtatccctttacaatt attctgaacacgctgttttaaacctgcaccaagccaaftaactatttagctagaccagg cagcaaaattacgggactggcgattttagaaggagaaaattgtatttgggttccgaactt gagcatagaggaggaataatcaaaacgctctctctgtagcggttctttacgccgcagtc gcaggaatcgtaaaactcgtatcgtccagcacgttttgatacagaaaacgtaagaag gatggctacctcatcattgatcagctgggtttgactactagacttgggttaaacggcttt</p>	<p>gtcaggaacccccgacttgaagtcggggg cttcaggcagataacctgaagccctagtftga ccagactactgtttgaggcaagtgttaaga cctacttttaggatcgtgtatgcttgaacct agaaccagacgattaaacagacttattgggt taagtcatgtctgttgatagttaccgacct</p>

SEQ ID NO:	Name	IscB sequence	ωRNAs
	CONTIG_ID: 0 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	gcccgtacgcaccaattgctcaagttgctatgagtggttaattgacactcaaaaaatg caaaatcccgaatagatggtgtgaataccagcaagggaattagtggtatgaggtat cgcgaatactgctcgaataatggggacgcaaatgcgctattgctgatacatcaggcgtg ccactccaattgagcatattaccacgcgccaaaggcggcagtaatcaggtttcaaac ttatgtttgagttgccaacgggtgcaatatcaaaaaaggagtcgaattgattgaagaattctg aaaaaggataactctagactagaaaaaatcaagcaactagccagggaagcctctaaaaaga tgcccttctctgttag	caaacatgtcgaagcaaacattacctgtaa aaaggatgcagcaatgcaaaattac
1441-1442	Ga0310694_10002417 GENOME_ACESSION: 174DNAGHGhip2v2_\$ F_3300031992 GENOME_ID: 279760 CONTIG_ID: 2416 SOURCE: JGI DATE: 2019-06-13.	atggctctatgactgaaccgatacggcaaacctctgatgcctaccaccggtagcgagg gtccgcccgtctgctcaggaaagggtacggtgctgtgctgattaccgtccatttataatcca gctcacttatgacacgcctaaccggtgagggaggtcagctagggctagacgcgggta cgaagcacgtaggcttctccgacgacaaagaagaagggtactttcgaagcagaattat tgctgaggtcggatactggtgagaactctccacacgaaggaggtccgtcggggccaga aggaacaggaaaactgctaccgtaagtaagattcctgaacaggaccggctccaagaa gccgggatggtgctcaccctcggtgaaggcagaagggtgattcccatatccactggatttca aagatatgcaagttctgctatcaagaagaataacgggtggagggccacgcaattgataccc agctgttgaaggctcagggaacagggtctcgttgcggcaggacaccgactaccagaag ggtagcagctcggattctggattgtgctgtagtacgttttattccggatgggtcacaaat gccagtgctgcaagggggaagtcgaaggacagtaagctgcatgtccatcatatcgaagt agaaagacaggcggcgtgacccgaacaacctcatccctgtgctcggaaatgctatgc caagtaccatcggggagaattgatctgctgaagattgcaaggcgtggcaccctgctac gtgatgccacacagatgggtatcatcggcagggctctatttaaccggctgaagaaagaaa tcgggtgagggaattactgctcaagacttacggctacatcaccgaagtagtactgtatgtaa atatggactgccgaaggaaatgtcatagatgcgggtgcatcagcggcaatccgtgtg catgttcagatggcaagtagtataatccgcaagttgctgctgaacaacaggcaactgc acagggcaactataaataagggtgaaagcgggaggaacaaccaggacccccgtgaag taagggggttccggctgatggattcgtcagtagtgcatacagggtgcttctcgtgagcgc acgtagagtgtcgggctcctttccgttgcgataccggaaaagtctatcagattcgc ccagttacaagaactgaccttaaacatcacacaacacttatattatggagggaagcagc gctcctctcaccactaaagatgggtga	atcaactaccacgggctaaataccgtga gtttgggggtgaccggactcatagttgattag actcagtgaaagggaataatccgctgaact acgttaggaaggtatgcacccccgggatg tacgggtcaagtccccgctctgcccgtgtg gttaaacagtcctgatgggtaggacagtg ctcgaggcacgacaagccttctaactgtg cgtgaccacccacagggaatccctgc attacattatagttagggtcaaaaaaacaa ataacggataacgagcacatg
1443-1444	Ga0070697_100040654 GENOME_ACESSION: KBSK1StandDraft_5_\$F _3300005536 GENOME_ID: 30495 CONTIG_ID: 40653 SOURCE: JGI DATE: 2019-01-27.	atgaagggtgttggctatcaaaaaggaaagtgtgtatgccacgacgccacggcgt cgaagggtgtgtggaaggcaaacgtgccgtgtcgtgctgcccacgaccttaccatt cgattgcgtttgccaccgcaacatgtgcaaccggccaagggtggcgttgataccgg ctccaaagacgtgggcatcgtgcaactgccaatggcaagggtgtcttcaggccagg tccatctgctgtatgacatcaatgaagaatgacccaacgacggagctttcggcgcaatc ggcgtgcgcgcaagacgcgctatcgtgaagcgcgtatgacaatcagccgacacctg tgggtgtgtccccccttgcgtgactccaaagcagagggcagcgtcaaggccgtgctg tcattacttctgtttgtcgtgtgtcgggtcacagtgcaggttggccgtttgacccaaa agagggcaaaaccagacatgctcacttggaaatccaacaggcgcaactgcaaggata cttctcgtgtagtattgttagacaagtggcagaggaaatgcgctgactgtgactcac gggtgtcccttggaaatcgagcatattgtccctacgtcaaggggaggaagcaatcgggc cagcaatctcaccctggcctgcatgctgcaacggcgcaaaaggcagcacacagca gcggagtgttggcttccagaggtgcaggcaaaaggctggtgccgtgaaggatgcagc gcatgtgtcttactgaaagcagggtgtgctgacgtgcaagcagttctggagagag ccaggtgagcatcacatatgtttaccagaccaaataaagcatccaggtgctgacct gcccaggtcccacacgaacgacgtatggcgtgctgcaacggcgcaaaaggcagcacacagca aagccattggagatgtccatcagatccgatgtctagcacggggccatcatcagcgtttc aatggcctgcacagtgcacaaagtgttggccccgcgcaaaagtgcgagggtcaagc tctacgaactgggtcaaggcgaagagggtgtgggtacatcgcgggacggcgaggaga agggagcgttctcatcaagacgtcatcagtggaagaacctgttgagggtcacggcg cgcaattggtacgagtagcgcgtccaccagggtggtatgattacgcgacagccgat gctggagagcatcagaagggaaggcggtccttcccccaattga	gtcaatgacctccccactgtgtggggaggt cgtgagggggtgcccccttgcgaagcggat tgaccagacgacacctccgagccagtcgc ggcgagaanaagcagcgtcgtggctggcgc tgaagagggtgagcagatagtaacagggtg cagttcgcacccaggatgctgccccagttc ctggctctgcaacttctgacttaagaactgc gggtccaggaaagggtgtagaggaagta ccgggtgctatcaccgtgcaggggatcgtt ttactccgaaaggagccttacctgat
1445-1446	Bioreactor metagenome contig_111, whole genome shotgun sequence GENOME_ACESSION: AMWB02.1 GENOME_ID: 96 CONTIG_ID: 5442 SOURCE: WGS DATE: 2019-01-27.	gtggctcagtagcgaatggggacgaactataagcggagggttaaccttactgtatgaact agtaatacaggagcatcgtatgtcagaacctcaacctaaagtctaactgaactggacag acttacaaagtaccagctcaacactgtaaacatccaccagtaactacgtacagttgtca gtagtacaggtaaacctctgatgcctgtcatcctcaacgtgcaagagaactgatgaaaa aggtaaaagccaaaccaagatggcactcggtagcatcttcatcaccagttactgaacgt gaagatgtgttaactgctcactaagtgaccttagagtagatccaggtatgaaggtcgaag gctactctatccgatctactaaactgacttacttaacacctcagtcagttgcacactcatgt aaaagatcgagtagctgatacaggaatagcgacgatcagacgttccgtaatactcc atgtcgtgaagacagaagtagtctgtagtaacttaactgaacggtaggatacctccatca actagagctagatgtgagctgaatcaacatgaagctatgaagcactctaaagtataacctatc gttacgatcgaatcgaagacataaaagcttatactaaagaacaaaagtccgatggc ggctcctagtaacaaactgtggaacaaagactctcgcctaccagatcggggaagagtag ttatatactaaattacaagagttatctaaagctcactattgttggaggtctgtaactatca actccgtatgtctactacaacccgtcaagaaagctaaagcgttaaggagcctaagatctcga	gtcaagtaccccccttaaggaaaggggctt gtagtctaggtcacaagggtaccttagtgaact ctaggtagaanaagactagcttaaccaagtaa cctagcctagtaatagtgttctaggggacga ctgaaggtaactgaaccatacaactgat gtacgggtgctcactagctcgtaacctctc agggtgtagtagcgaagtgaggacgaacta taagacgg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gctcactgtctagatgctgggtactagctggtatcggagtaggtcatgaagctattgaag aacctgataatatgagcatgtggtgtgtagaaccgatactactgcatcgtagacagttacat gttaccagcccgtaaggatgggttcagaagatcttatggtgtactagaagcatgggc tggaaagcgtggtgtagtcatcatccttaagtagggtcacacttacgtagggtgtagta gtaataatcggatcagcttacatagtagtccgagatggacaacgactagctcagaacgtcta tccagaagagtgtaagttccgtaccctcaacgccatgaagatcaagttcatccagcaggtt gagatggctcgaagatctgtaaactact	
1447-1448	Ga0209511_1005681 GENOME_ACESSION: SSSSmetaGSPAdes_ \$F_3300025147 GENOME_ID: 35187 CONTIG_ID: 5680 SOURCE: JGI DATE: 2019-01-27.	atgggtattgtattgtcaaaacaaagaaacattagacatgtgctctaatgtcaaaacaa gtattgtctaaaaaagggttatgcagtagtgcataaggatttctttacattagacttaaaa aagacgtgtcaataataaaacaaagaaatataaaaaatagaccaggttctaaata tactggattatctattatagataataatgccaaaggttttttcttgcataatagacatagag gagaaaagggtgtgaagcaatcttataacaagacaacaatcacgtagaaatagacgccaaa gagaaacccgttacagaagctgtaaattataaatcgaagaaacaaaaagatgctaagta taggggtgcaactaatagaccgaaggtatgggtccaccatctgtaataatctatagaacaa aatataatgaacttataaagaaggttaaaagggttgaataaactcaagttcaatagaata tgttaaatgtacactcaactaatggaatcttaaaatcaatggaatacaatcaaacagg tactcttttgggtatgaataaagagaatctttatcataaatggacatacatgtcaattgt tgggtgggtcctaaaaagataatcattagaagtggaacatatgtattcaaaaaaaatagt gggtctaatctaaftagaatftaagtttagcatgtcactatgcaataaagataaagactctt aafttagaccaatgggttaactaaccttaaatcaftaaagactactaacttaaatgatacaag aataaaagaatgaacataatagaatgaaggtactattatagaacaacagatagtg catgggttaaatgggtataaagaaaaatagtcagaactaaaaaactaattccagatata gaactaggtacaggaggcacaacaaagtcataatgaaggtatgaataaactatcaaaaca cactattatgatgcattgtgtgtggagctataccaagtagtttaaatftaagactactgatgt gttaactattaaagcgtatggaagaggttctcattttagaggaagaactaatagctgtggca ttataattcaaaagttaccaagacaaaaaactatctatgctttcaaaactggagataaattag tgcaaccgttaattaaaggttaaaaaataggtagtattttggaagagtggtcacaagaagt agtggatacttaacattcaactaaagaaggtcacaatacaaggtataaacataaaattg taaaatagtgcaacgaaatgatgatactctataacatagaaaaaagggttatttaa	ataaatatgtaattacccacgactgaagtcg agggtcatgttaggcgactaacatattcttaac ctaagttaattgagcagatttaagttaccttag tcatcataactattatgcttctcagtaaatgt cgttatgggtacataatcaaacatggctttgt gggaacgtgtattccgaaaggaatacatct aaaaccaagtgatgtattcataaagctatgt aacagtatcgaggagagacagtcgaaagg ctgcgtacgccgtgaagggtgtttcttaca atgtaggagggttttaattag
1449-1450	Ga0118733_100051634 GENOME_ACESSION: Marsedof8samples_ \$F_3300010430 GENOME_ID: 32450 CONTIG_ID: 51633 SOURCE: JGI DATE: 2019-01-27.	ttgaaagtattagtttaataatcatgaagaagcactaattgccttgcctcgcgagaagggc aagaattctctcaagaataacaaagcgaagtcgttagacgtttgcctttgttattcaatta aatgattattctctactggatctacaggggtttaacactaggaggtgactccggctactcta caataggtttgtctgttacttcagaacaaaaagaaatttattgctttagaactccaattaaagaa tgatatacaagtaagctagaacaaagaatgtatagagaaggaagaagaggttaggtt aaggatagaaaaacctagatttaacataagactagaagaaggaaggtggtcctccatc agttcaacataaagtgatagtcattgttaaaatttattgattatcagagatacttaccata agttaaattataatagaactgaagttttgatattgcttaggattataatcctagcatagag aatagtgattatcagaagggttaacatattgcttttaacaacgttaagcatattcttcaa gggataagtatactgtcaatcaggtaagaagggtctgttcaaaagagctccatgtacatca cttgatttcagatcaaaagggtggaagtgaacatcctaaaaacttaatacgttatgtaaga agcatcacagacaattacatgatggtaagttgaattataactttaaaaagcataaaagtttaa gggtcagcaacgggtgatgaatataataagagtttatataatataaagaggttaccaggagc agagagcgtttgatatattactaaaagtattaggttggaaacaatataagaaaaacgcata gtaattgatccttttgaatattctggagggttaatacagaagaggttgatgtaattgattataa attgagaagaaggaataaccgtatttacaacaaagacagaatggattttctagagatt aggagagaagataattatcaaccacatgacgtatgtgagtagaaggaaagagatgat gtattacaggcatgatgaataaaggttaagatccaattaatgattgatgtgtaaaaga agacaaagcctccagtttaaggttaacgagttatataatgacgaaggagtaggtattgtftaa	gtcaaccacccatgaatgaattcatgggcat gctccgtgaggagcatgataatataatggg tgattagattagcattgagaacatcaagctac gtttgttaagtataatccttagagtgctcc ctagctctaatgcttattgttaattataaag agtttaacgactcgggtgttaagtgtgaaa gctttacaactttatcgaagggaattctacc ttagagaggaggttagtttagcaattaaagctat agctatctttaaactaaaaggaaagaattga aa
1451-1452	Clostridioides difficile strain 7468-NonSp/ST97 LCH7468 contig000007, whole genome shotgun sequence GENOME_ACESSION: GCA_002302755.1_AS M230275v1_genomic GENOME_ID: 130261 CONTIG_ID: 6 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	ttgggtttgttaattaataaaaaaaccattgactcctgtcactttgctgtagcaagaa aattacttaacaaggaaggctgtttgttcataaaacatttccctactattagattaaaaga attaaaagaagctaatcctaagggttaattatataatagaccaggagctaaacaacag gcgtagcccttcacaaacctaattggtgaagttatacttctgcaaccatagaacatagagc atttagaaggacagaagaatagaaaaacagggtatagagcaccaagattttgaaatcg caaacgacttgaaggatggttaccaccatcaattcaaaagtatagtggaataatcaaaaatt ttataaaaaacttaaaagttatgctatataaagacaatctatataagaacagtaaggtttga tactcaattaatgctaaatccttaattgaaggtagtattatcaacaaggcactctttagg gtatgagattcgtgagtatctatataacttatggccatacatgtcaattattgaaaggactc acgaatgaccatcaattagaattgagcataagtagtcttaaatcaagaggaggttcaatt ctattagaatttaaccttagctgttagaactgttaaccaagaaaaaggaaataatacacttg atgaatggttacaatctcaagtaaatcaaggttagataaagaaagagtaaaaaataaac tcaatcttaaaactaataaacagtaaatctaaaagatattgccaaagtaataatagtagtag aatgct	gtcaataaccaccccttaaaaggttagcctt gaagtattctacagtcattatgaccagatta agtttaaacatacattatgaaggtcataacac tacaatgcttctccagttttagcctctgtt catattattaaacagtttgggtacaaacagt gtaatatgctatcaagccttaataatattatcg aggagagcagtttcaactcgtgtacatga agtttaatttcaactcatagagttactcttaa aggagggttttttttgg

SEQ ID NO:	Name	IscB sequence	ωRNAs
1453-1454	Ga0307994_1018066 GENOME_ACESSION: EllFjometagen261_2_\$F_3300031660 GENOME_ID: 280160 CONTIG_ID: 18065 SOURCE: JGI DATE: 2019-06-13.	ttggctgtattgttttagataaacagaaaatccacttatgccatgttctgagaaaagagca agactgctcttatagaaaaagggcagttgtgcacttaattatccgtttaccattcgtttga aagaacgagttgggtgatgtttacagctctgtacatctaagcttgatccaggtgctaaaaca actggagtgctgttttagtagagctacaagaacgttcaaaagtgccgatgtttcgccactt agagcataacgggtttgcaataagtgaaaaactaactcagcggagagcatttcgtcgcag acgtagaatcagctttgtatcgtcaagctagattgacaatagaaccaagccgaaagg ctggcttgcgccatcgattcagcaccgagttgactcaacactttcatgggtaaaaagattt ctaacccttgccttgatgtattgagttgagcgttgaagttgacattcagaagatgca aaatcctgaaattcagggtttgattatcaacaggcgactttatttgggttgagttgaaaga gtatttgcctttacaacataaccaccaatgccttattgctcaggtattagtaagatccaata cttgaggttgacacgttcaaccaaggcttaagggtgttcgaattcagtgaaagtgttgg aattgctgttagagcttgaacgaggaagggcgttgccttttaacccagtggaataaac aattgagcagaagtgttggccataactagaattagggtattgaccgagttttatcagg taaatggcgttgatttagagataccgcagctgtgaatgccacagaacgccttgttagct gatattttaaccttccatcgaataaaaaaacaagagcaaatagcgtctacacaggca caggtgctatgacgaagcttaacgaaagcttcaaggtttacaaaagaccacgccatag atgcagcgggtgtgggtgaaacccaacacttgcagaactggcaagtaccgttctca gcatcaagagtacaggcagaggcgttacaagcgcacaagacttgataaattgtgtttcc tcgctgttacttaattcgtcagaatcgggtacaaggttttcaaacgggcacatggtaaaa gcattgttttaaaaggcgaataaacaggtgaatccttggctgtatcgtgttagagcca gtggaagttaattatcaatcaaatgaaggtggtgatacagggtattgataagtgctgta cgttattgcaacgcaataatggttatggtttgttaacaaagatagctttaaccaacggag aagagagaaaagctgcctaa	gtcaactaccccgccctaaaggacggagct tgtaaaagcttgggtgaccagagggttaa ctttagtttaagcagaagatcgtaacaggtcgt taagacttaccctgggatgcttcttagtccca agctctaaaagggttaagaatgcaggcata aggtaagtccgaattcttaagtcgtgttag aaatacaggagccggttactgtcaatctcga aggagagattgtgttctccaacagcttacta gacctgtaagggttaatttgagttggtttt ggct
1455-1456	Ga0114919_10027729 GENOME_ACESSION: BlataG_2_\$F_330000952 9 GENOME_ID: 23889 CONTIG_ID: 27728 SOURCE: JGI DATE: 2019-01-27.	atgcccacaactcctagaaaagcacgcctactcttgaatacaataaagccaagatcaaa acacacaacccgttcacattcaactaacaacagcaacaggagaacaaaacaacccat tgaactaggcactgacccaggtttcaatttattgggttttagcgtctacacaaaaaaggaa gaactaatcagcgggtgtgttaagataagacaaaatatacaaaaactaatggtatccgct agagggttcagaaaagcacgcccgaatcgaatggcatagaaaatctagatttaataac cgcaaaagaaaagcaggttgggtaccaccaaagtcacacaataaagttcaaacacataa aaacttaacagaacttcagagtattctccaatccaatattctatatagaacaacaa gttcgacactcaaaaatggagacccagatatacaagggtgtagaatacaactggaac tttacaaggctacgaagtcgaaggattttattagagaagttgaaagaaaatgcgtctact gtagaatcaaacacacctttagaagtagaacatgtcataccgagatcaagaggaggtga gtgatcgcatctcgaatctcatatcatgtcatgattgcaatcaaaagaaaggcagtcac acagctgaagaattcggataccgacatccaaaaacaacaaaagatcaattaaaagc agcgacacaaatgaatataatagtcgaagggtattggaagaccagatctgaacactat cccaatcctgctttagaaccagtttaggagaaaacactgaacttagagaatcgcat gtcaacgatgctttgtcatagcaggaggcacaacacaaaagatcacaacattatc attaaacagacaagaagaataatagaaaacttcaacaaaacagaaaagatatggtattt caattagaaaacaaagatatccaataagcgttaagcttaattgagatcgagtgagagg cgagagtaatttagatttttcgagtttaaatcaacacaaaggttgagaggttaattgttca gataataagttgatcaaatcgttttccaacaggttcaaaagttaatttactgttcaggtaa aggcctttgtttattgagta	aaggagggtcagtcataaacctcagggtgggt agagggtgtgacaacactcagtttaagtt gcaacacgcccgggcacaaacttaactgaac agtcaagcagcagttgggtcaggtcatgata cttgggatgctcaatagtccttcgctcctatc gcttagtattaaaagtcctggtggcgaagga cgggtgactaggccaaaaagcctcccaaa caatctcgaagtggaacaaactccaagacga agaggagggtagaacttga
1457-1458	Ga0256404_1000183 GENOME_ACESSION: RumenRJG_01DNA_2_\$F_3300026549 GENOME_ID: 34936 CONTIG_ID: 182 SOURCE: JGI DATE: 2019-01-27.	atgacatacgtatgtgtcgtgagcaggagcggcagccgcttattgcccagcagaagatg cggccatgtgcgcatcctgtgaaggaaaagaaagccagggtgtgtatcggcagggccg ttcacagtgcagcttctgtacgatacagaggaagcagagccatctctacggaggaaca gaccggggcaggacgaacatcggaatgcgctgtgacggaagacggcgagtgctgtt taccgcgacaaggctgagacaagaatgatgaagtgccgaaaggcgtgtccgacagg aagaagaatcgtcaggctcggcgagaggtgaaaggtcgtgaaggaaagcggcg aaaagatgcggcaccctgtcgacaaagctcggaaagcggcagacttatccggggacga aaaagcctacgctgtgaaggatcatcaatcaggagcgcggttccggaaccggaa gaagaggcagctcataacaccgagttgtaaacagctgttcgacacgcatctgaaccagct cgatcagataagaagatacttccggtcaaggctgtgtcctggaagctaacggttcgc gttcatgaagctgaggacggatcgggtcagaggtatcgacttccagaacggcagactga aaggctacgctcggctgatgttctgtgtacgagcgtcagaaggaaggtgttctgtct cgggagcccgatcgaacactaccatgtggaaggagcagaagcgaacggctttga cggaccggagaacaaggtgggtctgtgcaatagctgccacactaagatacatatcgggt agctcagctcgtatgtcagggttcggcaagaagtatcaggcgtgagcgtgtgtaac caggcaataccgtatcttctgtggcctgtcgaacgcttcgatgaagagaatgttctcat ctgcggcgatcagacagaaggagatagcgaagcagccggccttgacaaggacca cgatatcgtatctgtgcatcgcgtcgtatggcgacaggcgtgataccaagcagcctg aggaaacaggccttcagtgtaagcagtagcggccacgacagagcgaagatcaaca atcagcctgaacgcacatacaagcttgacggcaagaccgttgcgaggaaacgaaagcc gcgcatcgaccagaaaggcctcgcattgtcacagttgtacgaaaagcaggtcagacttt cggcgaggaaaagaaagctgaccgctgtcgtcagagctgaaggtaaggaaagctctacag	gtcaactaccccgcccttaactcgtgacgc ttcgttagaagcgggggctgtgaaggcgag agtattgagcaagagcgtgatgtatctcac actcggggttttccagtcgccacatact ggtctgatgcaccaagctcagggaacatc acccgcccacgagcggagactaacagtt ataactcaactagaaggaaaggagacgc acatgacatac

SEQ ID NO:	Name	IscB sequence	ωRNAs
		acgctacaacaacctgtaccgctcatgccgggtgcgacgttgcacatgacggcagga tcgaagtattggagagacagcagaacaacggctattactcaaccttcgatcggggagaa taaaggcttcagaatgcgagatcctgcaccataacgcagggtgtatataatttag	
1459-1460	Ga0272443_10006725 GENOME_ACESSION: SSMAcet12_2_\$F_33000 28883 GENOME_ID: 280921 CONTIG_ID: 6724 SOURCE: JGI DATE: 2019-06-13.	atgccaacaacacaggaaggaaggtgtttaaacaacaaataaagcaaaagtata taaaagatacccttttactatacaataaattatgctactgtgtaagctaaacaaatattaca ctcgggtattgatactggatataacaacataggatttagtctattacagaaaaagaagat atattcaggaactttgaattagattcaaaaactacaaaagactacagggaaaaaagtgt atagaaggacacgaagaacagattgcgctatagaaaaccaagattcaataacagggtta aaaacaaaacataaaggctggttaccacctcaataaacgaaggtatgaacacatac aacttaataagaaaaataagaataattgccaataagtagactaataactggagatcgcaa agtttgacatccaaaagatagaaaaatccaggaatagaaggttctgtttatcaacaaggtaa tatgtatggttatcagaacttaagaattacttaattgtcacgtgaaaaagggaagtccaatt atgtggcaaaagatttagtaaaacagacctgcacatataccattgtattcaaggaaatg atgtgggtacataaataacctaataattggctttattgcataagattgtcatcagagactac ataaaaagaatctaaatcataagctaaaaagaataaagaataaagggttctacattcatg tcaataatcataagaagttttataaagatttaccagctcggaaagtaacttatgtataataa catttgaatagaaatgttttaggttttagaaaaatcacatagtaaatgatccctttgtaatagc gggtggttaataatcacagaaaaagaataaatccaatagtaataaacaacacatcgtaataac aggctcttaggttaaacagaggaaaggatttgcacctcaagtagaaaaaaagatataag attcaaccaatggacttagtaaaagatagctgtaaatgtttaagaactaatggagtagattg taaaagtaaaagttatgtaataaaaaagataaataaataaagtagagtcattct atagtttcggaagttttatatttaactaa	gtcaataaccctcgactaaagtcagggtt gcttggtagaagtgagggttaactgtttgatt agactaagtagctttaggggtactacgtttt aagttataactctgtgggtactccactagcc tgcagctctattacttggattaaaaaggttaa aagactcgggttattaaagtgtaaaaagctttt agaacattgtcgaagtggttaactctgaaa ggaggacgaaacttgagag
1461-1462	Ga0118725_1014359 GENOME_ACESSION: IMG_3300009377_\$F_3 300009377 GENOME_ID: 311715 CONTIG_ID: 14358 SOURCE: JGI DATE: 2020-04-07.	atgagagttccagttagtatcaaaagatggttaaacgcctcatgcctaccaagccagccaaa gcaagaagatgatagaaggaggagtagctaaaaaatctgtgtcaaaaactggcgtatt ctatatacaaatgttaatacctgtgggcaacaagttcagctctggtgactagcaatagacc caggcagcaaatatgatgctacgctgtgtctgtgtagaaggatgtagcgtcaaggct atggcaataatgccagaaaggtacaaaagaaagtaactgagcgcctcaatgaagacg tagtagacgttatcgtatacagacataggaaggcacggttgacaacagaaaaacgca aatctgggtgtagacacatcgcaattggctaaagtacaaatccgtatcaagatagtagc agatttagccaagatattccctcctaactatctgcagtagaggtatgaattcaatcatta caaaaagcgatgggttaagtattctctacagtgagataggaagacgatgtgtacga agaacttggaacgacatgttcaggtaataaagtagcgggttgccaacagcagaagcac gtaagtattgggcattagaagtaaatgttaaggatgctttaacgccccaatctatgct aatgatgctcttgcgatgcttaafgagatgttcggaataaattgttgattctgtgcattttct tgtttggcgtgagctggaattcgccagacgttcattacatcgccagaattataagaaaggt gtgtgactctacgatttggcggaactaccaacggacattatagctaaaggagatcttg tctctgtgagcaaaaagatagacaattcatcgggttgggtctgtggttaccacagacaa gacaagggttattgctgtgctgatgcaacaggttaagcgttttagcacaatgtacagaacg aaaaatcggaactaatgcgccgggttactggcgtaacatgggagagtcagtagataccaaa gccaccaatagcacttatagcgaggagcccaaccctcaactacaattttctaa	tgttctgcaataaccctctctgattctccg aatcagaagggttgggggacacaggat ctccgacgcaagtattgactagagggtccc gaaagtttcgggacagccgcactcgcgatg gtacatagctcatgtgtttcgtactgtctga cctctataactgtctctgtggcagtggtggg ataaagacgtgacatcttctgcagttgcaatc tcgaagcgacctttactttagtcccgaatct gggactccgagaggagaaaatcatgaga
1463-1464	Ga0335001_0013376 GENOME_ACESSION: TYMME1r0054_2_\$F_3 300034064 GENOME_ID: 281041 CONTIG_ID: 13375 SOURCE: JGI DATE: 2019-06-13.	atgcctaggagaagttttatgactaaaaatgttgtgttctctagacaacaccaagaacc actaactccttgcaccagcaagagctagaatgtttgtacgggatggaaaagctgcagtt tggcgaactgtccgttttaataatcataatgcaaaagtactataattgaaccaataatcaaac gcttaaaagtgaagatagacccgggttcaaaactactggcatagtctctgtgaatgaaaaa aatcaggttatatacgcagctgaattagttcatctgcgatagataaaaaagtgatttagaa tcacggaagcataatcgaaaacttagaagatctcgttaaggtagatagataaagctagg tggttaaatcgtggacgcaaaaaagatggttaccacatcaattttacataggggtgattac aaccgtaacgtgggtgataggtttatgaagtggctgtttatgaatggcttatcagtagaaca aataaatttgacacacagaagatgtgcgattcagtagtatatgtgttaaaataccaacaa ggaacgctgtctgatataacgtacgagagatctacttgagaaatggaaggaaatgt gcatattgtgggacatctaactaccattacaggttgagcatgtacatccgaaatccaagg gtggttcaaatgcattagtaatttagttatggcgtgcgccaatgtaattgatgtcaaaaaat cattggatattgctgatttttgagaacaaacaaatacacttaaaactataatgtcacaatt acagaaaccgttgtagatacagcggaatgaattctacacggataaaagttaataacggta ttgggtgaatatgtttaccaatgaacttggttaaggtgcacaaacaaaataatcgtgtt aaactaggatagtaaaagaacattggattgatgcgtcgtgtgtgtgattctgtgtgagtc agttatcattcacccacaacaaaacggttacgataataatctatgggttatgtaatagac agatgactcgtattaatagatttggttttccatgtgctagtgcacaaacgggttaaagttgtatt aggaataaaaactggtcgtttgttttattaaacacaaaagggaataacacgggaat atacacatcgagagtttctgcaattaggtataaagatgattattgcaataacagtttccggt agaaaactatggtttccgctaaactagccagaggttaaacaaatagccgatggtattttagat atgtaa	gacaatgaattatatttttaaacgtgccgga ctttccagttgagcgaccagattaacaatagt tagtattttatcaatgtatttattgataaagctc aaccagcttacttgagaattcaagtagacgat acgaattgaagcgaagcgacacacatag gggtgttcaccagctctatgcaactgtgtgtc gaaatgtatgactgggacatctgtgtgtcca taagcatgtgactgaacgacattctgtctccg agcgaggtaaacattacaggtatgccttag gagaagttttatgactaaaaatgtt
1465-1466	Ga0311301_10014764 GENOME_ACESSION: Sb_comMetaSPAdes_\$F	atgtcacgtgtattgtttgatgccaaaggcgcaaccgctgatgccctgcacgcctgctc gtgcccgtatcctactggcagacgcaaggcagcggcttgcgtcttcttcttcttctt atcttgcgcgagacaaagccacagccactgtggcaccgctgcgcctcaaaattgatcc	gtcagcgacccaccgctaagcgggaggg catgtgcagagcatgtctgtgtgaccag ccccctgtgcttgggcagaggagcgttag

SEQ ID NO:	Name	IscB sequence	ωRNAs
	_3300032160 GENOME_ID: 280981 CONTIG_ID: 14763 SOURCE: JGI DATE: 2019-06-13.	aggcgcgaaaacgactggcatggctgtctgcatctgagacgtctgaggtgctctggg cggcagagctcagcgcacgcggcgagcaggtgctcgcgcagctgatgagcgacgg ggcgctcagacgtctcgcgcgcgcgcacatcccgctacagagccgctcgtccgcaa tcgcccgcgcgtctcgtggctgctccccctctctactttcactgtacacaatgtcag acttggtaatgcgtctcgtccgctttgtcccatagcgcgctctctctatgaactgtcgc cttcgacacgcaactgtccaaaaccagatattgagaacctcgactaccagaaagggc ctctcttggcctcgaactgcggcactatctcgttgcgaagtgggagtagcaggtgtcgt ctgtctgaccacaggggtccgctcgaatgtgacatgttgagccgcgagctcgggga ggcagcgatcgctggccaatgtgtgcctgccacgcctgtaacaggcgaaaag cagacaaactcctggaagaattcctggcagaccgccagaggttctggcagcagtgcaa gccaacgcgcaagcaccgtcgaaggacgcagcgcctgtgaataccaccgacttgcctt gtatcgaagtgtacagcgacagggcttgcggtggagacgggaacaggttgcgtgacc aatggaaacggcaacagcaggggtgccaagaagcgcactgggtcgtatgccgcatgtt gcggcgctcactccggcgacgtgcgactccagacgtgcgcccctggctgactcact gcgacagccgacagaacagggacatcgcaatgttgacaacacgggttccccgttg gcccggccaaggtcgcgagccgggtgcggttgcacgggagatagtcgaag ctgtctgtcctgccatctgaaggcggcggggactcacgttggcggtgtcttagtcgca cacgtggaatttgcgtgcaaacgaggcagcgtcgggtcaagataccctgcacgt actgccagaggttccaagccggagatgggtatcgctatcttgggtcggcgcttcccc tactcccgaatgtggaggtcccgcgccggcgtttag	ttggaagcggcccaagctcggactctgaat ggcctttccagctcagacccccaatcccg atgttaaacagcgagacggggagcagag ccagtcacaggggaagatggcgccagct aactggcgaggaacacattactgtttggt gctcaagagaccagcagcgttccccctt acggggtccacaaggagttgcatgatgc acgt
1467-1468	human gut metagenome genome assembly, contig: NODE_9203_length_484 2_cov_4.000000, whole genome shotgun sequence GENOME_ACESSION: USUO01.1 GENOME_ID: 16942 CONTIG_ID: 9202 SOURCE: WGS DATE: 2019-01-27.	atgaaacaaaaaagatgtcattgttagataaaactggaaaaagattatctctacaga tatcaataatgatgcagatgaattcgcacaaaaaaagcaagatgcattaaataatccaa tgactatgaagtaataaaatcgtaaaagaagaagaattgactcctgttttgaattgg aattgatgatgttctctcattgttgattctcagtgattcaatattgtataaaaaatacaaat tacaagaataaaaagttattcaaaaagcaacatgattcaacggcaagatgtaaaacatctt atggatattagacgtatgtatcagcataatcaccgttctgagaaaaagcagagcatgta gatttaataatcagtaataatcgtaaaagaagaaggacgtgtgtacacacattaaacaaa aagagatgcagttgtaaggtgttaaaagcttataaaaactgtcaatgtgacacaaatg cattagaagatgtaagcattgataattaggttttaactgaaggtgaaaaattagaaggaaaa gattatcaaaaatacaatagacaagatgacaattctgtagagctgtttatctcgcagacaa gggaatttgcacaaatgtgtggagaacataaaacaaaaatggaaattcatcattatcctc aaagattgggtggaccagatagattataatgaaattttattgccatgactgtcatgcac tattaatggaaaagaattgcaatacaagatacttttataaaaaaattatggcaaatatc agaactgattgtgcacacatgttatgcaaggaaaattatttaagaaaggaaacttcccta cttgagaattacatttaacataggtgtgtaaacagcaataaacgaaattgttggaaacat tgaaaaatcacatagtaattgatgcaatctgtattgttgaagatattcatatgaaacctgat acggtgaatgtacaagaatatacaataaaaccaaagacacataaaaaagaaatcaataca acatctatgggattgtgattgggtgattatgtggaactgaaataagaagtaataaattaaaa aagaaaaataagtaaaagggttacattactgcatttataaaatgtcagaatggaaaagata aaggaaaattacattatataaatttaactgcagatgatggaactatttacaagaatattctt aaaaaaatgtaattactagaaaaacaaaaacatttactgtttatggaataa	ttgcttcaagccttagtgagtgtagactcgg ttagcaaatgaactacgttatgccactgct taaaactaaccttagattttctcagtcgaag ctcttagagtatactttgtatacatgtgca acacatagcattggcaagagaaaaacatt aactgttactttcattgagattacatcaattg aaagaagatagaataatcaattttattatgga ttatgatttctttcataatccataactattta tttttaagaaaggagatattcctatgaacaa aaaaagat
1469-1470	human gut metagenome genome assembly, contig: NODE_64_length_52191 _cov_7.228997, whole genome shotgun sequence GENOME_ACESSION: OMDM01.1 GENOME_ID: 8240 CONTIG_ID: 63 SOURCE: WGS DATE: 2019-01-27.	atgggtgtacgtacacagacataaattgtaaacctatgatgccacacaaggcatggaag gttaggagactgcttaagacaagaaggcagtcgttgtgaacctatgtccgtttaccatca aattaatgtacgtacacatctgattacaacaggaattgtttagcggttgatctgtacta aacatgttgggttatcggtctaacacgaaagcaaaagacattacagtagtgaaatcctt agaaatgataattgtatcttctcaccagagggagctacgaagatcaagacgaaata gattgagatagataaaacctctgtttgataatagaataaaagtaagcgtccgggatgggt agcaccttcgggtgaaatacaaaagtagacgcccattctggttattgacaatatatgttctat actaccaatatctcgtattgttattgaagtagctcaatttatactcaaaagattaacaatcct gaaatatcaggtaaagaatcaggaaggtgatcaacttggatttggaaacgttagggaaa gacataatcgtcagatccataaacaagaattccaaaaggaggataaaaaagaccaaat caatctcctttgaaggttttcttccgattgtttgtagagggttattgttgaacaggtattattt tatattcgcaaggcgtaaaacgggtgattttaaattcagagatattggtgtaaaacaaa gagatattacatacaagaattgaaattatcaagggtgaaacgctttatgttcaaaaaggaa atggattga	ataaccaatttattgtattatgcataatagcc aaaagctattccgattattagcctaaggtgtg aacaacactacgttatttaagaatagatagti acctacggatgtttacccaagtcgtagctct aaggtaagtgttaaacagttctggtatttga ggaacaggtgttcttacgaaaactttaaata cattggcgatgggtactacagaggtttcactc tgacttattgtgaataaacattaaaaaggtttg tagatatg
1471-1472	Ga0394881_0018228 GENOME_ACESSION: WSA_S1_MG_FD_SF_3 300037311 GENOME_ID: 312717 CONTIG_ID: 18227 SOURCE: JGI DATE: 2020-04-07.	atgcaaaaagtattcgtattagacacaaacaaagccttagatgtgccaccctggac aagctcgaaggtgttaaaagcaggattggggcgagctctacgactgtttccattcacata attctcaagagagaagttactgctcaagaactacaaaactatttctaaactagatccagg aagttaaaacacaggaagtgcaatagtaaatcaagaacaggagagatgtgtatttgcag cagaaatagaacatcaggagaaactattaaagctagcttagatagtcgaggggtgttaa ggaggcgcaagacgtcacaagaaaaactgtttatcgcaagcaagattcaataacagaacc aggaaaaaagggtgttacctccatcactactaagcagagtagcaaatgtagaacttgg gttaagcgactatatacttattgtccaatagccggaaattagcttagaactagtcgaagttgat actcagcttatgcaaaatgctgaattgaaggtgtcgtgagtaacaaaggcgcaattagctg gatagagttacgcgagttaccttttagaagaagtataagcgaattgtgcctattgcgaaaa	gtcacttaccacggctcaagcgtgggctt gttgggtagataactaaacgtccgagtgac cggactaagttcaagaactacgttattagga gagttaaagaagacacgttgaatacgtgcc agttcaacgactgtacacgaatgaattaaac aattgcgaagggtatagcagatgtcttattgg atgaactgattataacattgttaggcaaac attacttccgaaggaaaggtttaacaacta tgcaaaaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
		acaaaactgcttgcagatagaacacatagtgccaaaaagcagagggtggaagtaatag tgaacaaacctaacattggctgtgaaaagtgaacctaaagaaaagcaacaaacagc ttggagtttggtatcctcaagtgcaagcacaagtaaaatgccataaaagacgctgcta gcgtaaatagtagaaggtgggaaatttcaatagactaaaagttacaggattgccaataga aataggaagtggggattactaagtacaatcgtagaggcaaacctacctaaagcac attggatagatgctgctgctgttggttaactactcccaaaaactagatgttgataaafta acctattgcaataaaagctacaggacatggagcaggcaaatgtgcagtagatgataagt ttgggttccctgctacctcaagaaagctggtaaaaagtttttgggtttcaaacggcgatat tgcaggggctgtgtcacaaggcgcaaaaagttggtacttatattgtaaaagttgccgtta gggctactggctctttgacattttaccctctctggcagagtcacaggattggctataagtat ttactctttgcaattcttgcatggctattctgttttacttaattcaaaacaagaccgcaattc ctcccacgctgaagtcgtgggcttctctgggtgatctgtga	
1473-1474	Ga0070717_10057385 GENOME_ACESSION: LARL11StandDraft_3_\$ F_3300006028 GENOME_ID: 31022 CONTIG_ID: 57384 SOURCE: JGI DATE: 2019-01-27.	ctgaacatccaggccgaagccagggtatggggctcaaaaagcggaggaaacacagcac gcgacagcgccgaacgcgccggggcgcaagacccttgcgcaagccacggcag aatcggcagcagagcaagaagaatttccgcccctgcagaaagcgcgaggggcctcgg aaactcgggctggcagcgttctctgtcagctcttccgggtgcgtgtcttgcgtgcgaag acatcaagcccgccacccagggaagggcgctgggatcagcagtttccggttggga ggtggggaacactgtttatgcagaactcgcgaacgttgcggcggtggtgaccaggc agggctatgacacatcaaccttgcgtgaccaattgggggtcaaaaagatcagcaagaag ctggccgaagtgtggagggcattgcgttgatccctgggtgctgcgtatagc	gtcaagcacccccgcataaatgcggggg ctgtgaaagcgaaaccgacttgcacgggg aaccagcgagggttaactgcgaacgagctt atccaagctcagcctcggggcagtgatcg gaagctgaacgatctgggctggcagccc acagcaacgagcgagacgctccctaactc cgtcccgcttgggtgggaagcctcgaagg gaacagtttactgaaagaggcttatgcac at
1475-1476	Ga0224423_10003602 GENOME_ACESSION: 3300021431 GENOME_ID: 241590 CONTIG_ID: 3272 SOURCE: JGI_Dark_Harvest DATE: 2019-01-27.	atgggaagccggacagggcatacacagcaatgtatatgttctccgggagctgtagagac agtaaacactcaattcagaagggagcgctgacgttaattgacgatttctcgcgaaga cgggtcccggctgatgcgacaaacatcaggaaagccagaagcttcttgataaaggga aagccgtcatttacaacatcacccgttcacgattcagttaacgggtgaatccaggcactg tactcaatcgtatgagttctgtaagatacaggatcagaacatatacgggtgtagatcaaat cagaacgcgatgagtagttcatgcacaatcagatctcagtgatgagaacacagagac atgaggcacaggcatcgtatcgaagaagcagggcggaagaagcgttatcgcaagc cccgcttcgacaacaggagaagaacgataagtggttagctccaacagtagaacacaa gaaggacaacctatagctatcttggatgtatgcagaggatgttccaatcacaagagca gctctcgaagtcggacagttgtatccggctgctatgcaggcacttgaggagacaggaga gtgctccaggggactgattatcagcatggtagaagtttcaacttccaatctgagagaag ctgtattcacaagagaccgctacacctgtcaggtgtgcggtgaagtcagtgaaagacgga gtaatttctatgcgcatcacatttaccgttctaagtcgggaacagaccgtatcaacaac ttgatgactgtctgtgacaagtgtcacactccgaagaatcataagccgggaggcaagctc tatgtctcaaacactgaccgggtatatagagacgtaaacattcaatattgtcagatg gtacatcaaaaatgataatcagaatcgttatccggatattactgtcgaacacacatgtgttc atatacgaagcgctcacgccgtgatctcgggtcagctgcctaagacatgcgaatgatgc atatcgatgggtgagttccaccacaagcacagatgcaaaagaaacgcatattgtgaac gtcgaggaacaatcgtgttctcgaagttctcagatgcgggtgtatatcgacatcgatg ggaactgttaagaagggttcgagattggatgaaccgcacagatcgcaatattccccgc agcaaccaagcaacgagaggatgtatcgccggcataaagtgtaaaagggaagatgat ctgtcagaagacagagatacagcatccagcccggagatacgggtcgtatagagggaagt attccccatgccaaggtgtgactgttaacgttacaagggtgatgctggatacaggcaa gtctgtaaagatcacagatgttgcgttaaaagaacaggaggatggcaattctcccc gcctga	gtcaataaccccacctgattcaaaagaatcag atggggcttgtgaaagaaaagcaatttctta cataagcccgttactagcctaagtgtcttc gagcactacgttatgcagtgacgcataagc gacttacctgtgaatgtccacaagttcacag ctctaagggtatgcattaaacatttcagtggg taaagagaagtgttataccttctaaccgtgc ataacattggcagatgtggacccgtc
1477-1478	Ga0395631_0033223 GENOME_ACESSION: 716E22024metaG_FD_\$ F_3300037196 GENOME_ID: 312229 CONTIG_ID: 33222 SOURCE: JGI DATE: 2020-04-07.	gtgacagtgaaacacgtctctgtctcaatcaaaatcgtgagcctctgcagccgttccccc tgcgggtggcccggatgctcagtgatgcggccaagccgcatctcgcgacactccccct tcacgataatcatgaaatgaccgtgagcggccaacccacacctaccgcatcggagtc gatccgggcagccggaagacgggaattgtgctgtgagcaatgagacgaacgaggtgc gtcttcgcatgagatcaagcatcgggggaagtgagatcaagaagtccttggatcagcg gaggggagtgccgctgtctcgaaggtcccggaagtgcgacatcgggaaccacgattc ctgaatcgagcgagaccgaaagggtggtctccgcccctccacagagtcggatcgaaa atgttgtgacctgattggggcattttccgcatcgttccggttcagatcaatctggaa gatgtgaaattcgacagcgactcctacagaatcctgacattcgtgggttttatataca aggagagttgtggggctacagggtcaagaattcctcctcctcgtgtgagacacgagt gtgcctactgcgggaagaagatgttccctcaggtggagcagatcattcctaagagcc ggaacggatcagatgaaatcctcaacctgacctgcctgctcctgcaatcggaaga agggaaagatgacggctgcggatcggga	tccaagcttccactgaaagtcggaagcctg tggcgaaaccttgcttgggtgaccagactc tgtgaagtgtagaccgagaggcctcgccca actcctgtgcgattgttcacgacaccttggga tgcacgccagtcaccaagctctgtcgtccgg ttcatgaggattgaagagatagctccgaa ggagccggaccgacaagccttctgcacat tgtcaggcggaacgtggtcatgagaatga gcctgacacccgtgagggataacacccaa gcggagacactcaattttcttggagagtg acagtgaatcac
1479-1480	Ga0137383_10023908 GENOME_ACESSION: SagtaG_26_SF_3300012 199 GENOME_ID: 35413 CONTIG_ID:	atgtcacacgtgtgtgctatgatcaggacaacggccacttgatccgatccatccgggct acgcccgaactgctctcgtccgcaagcggtgtgtatgcagcgttccgttctgtct gatcctcaaacgccaggttccggaagctcaccgccctcgcctcctcaaaattgatcc ggggagccaaccacggcctggcgggtgctcaacgatgcgaccggacaagtggtatg ggctgcggagctcatccacggggggagcaagtagacgcgggactcagaacgcgc aggggtgcggctggagccgttcccgcatacactgtaccgaccgctcgtcgtggcgc	gtcagcgaccccacgccaaaggcgggg ccttcttcggcagcccgatcgtgacagt cccagggcatcgcgctgtgtagtcttcgc gaagaatgtataggcactccgggtggcct tccagctccgactctgcggcagtcgttaa acagcgagatgggtttcaatcgagtcgcg

SEQ ID NO:	Name	IscB sequence	ωRNAs
	23907 SOURCE: JGI DATE: 2019-01-27.	aatcgccgtcgcccaaaaggctgcttctctctctgtgctcgcgggtgcagaatgtggagacctggacgagcgggtgatccgttggtgcccgtcgtggtgccatcctacgagggcgtgctctcgacaccaggcactccaaaaccagagatcgaggggtgagcctaccaacacggggacccttgcggcgttgagggtcaagaatactgtgtcctaatggggatatagatgtgtctattgccagaagacgggtcttccctagaattgaacataattgtccgaaaattcgtggaggaagcaaccggatcaccaattgacctggcctgtgaggtgtgtatcagaggaaagccatcaaaactgcccaggaattcgggttctcaggtccaggcgcaggcgcggcagccgctaaaggatccggcggtgcaacgcgacgcgttgggtgctcttcgcgccctgc aagcgacgggattgtccatcgaacaagcagggggggcggacaaatggaatagaaagcaacgcagcattccgaaaaccattgggtgagctgtgtgctgttgggtccatcactcc ggaacgagtgactggcagcagggtgacccgctgctcatcagagccaggggggcagacatgcccggcgtgtgtatgttgacgagttgggttcccaaggagccgtccagagggc cgagcaagggtctacggcttcaaacgggggatatgattcgtgcacgggtcaccaaagg gaaaaagatggggaccatgttggcgagtcgccatcaaaacagatgggtatttcaagatt accggcaggcatggcatggcagggtcattcatgctcgttattgcgcgccattcatcgc gatgatggctatgggtataccaaggcaaaagcggcgcttctctcccaggtgtga	gctcgcaactcaaggaaaccttgcgtgaagt tggacgaggaaaaccttacctgggcaacca gaggccggcaacgggcaaacgacaaaa ggagccaactgctatgctacac
1481-1482	Ga0208478_1004398 GENOME_ACESSION: NGSur0720SPAdes_15_SF_3300025475 GENOME_ID: 32998 CONTIG_ID: 4397 SOURCE: JGI DATE: 2019-01-27.	gtgaatcaggtgtctacgtttccgatagggtgagcggatccaccaggccgtgcttccagc ggcttgcctcgaactcgcgtcagcagacacggcgtgggtgagctccgaaacggggcg cagcgctgaactttctagtctcgagcagcagtcgggaacatgatcaaggaggagaccggc agccgcacccggccggcggcgctcagcaaaacacagccagtcaggtgtgtggcaa cagatgggtggagtgattccaccgccaatcacatggacgtgcagtcggcgcaaac agcagtaacaagacaaccgggcaacccctcgtatgggtgcgtcgcgacaacacgg caagccgctgatccgtgcccacccagccgcggcgagctattgcaaaaggcggc ggcggtagtggcgcatcttcaagccgttcgtatccgctgaaggaccgcacccgtgaac agtccgaagtcgagagcgtgcaggtcgccatcgacccggctccaaacacaccggcat cgccgtattcgggttgcgctcagaccggggagcccgcaagggaattgttcggcc cagttggaccaccggcgtcagaagatcagaaagccatgacctctgctcgagatgc gccggcgtcgccgatccgaacgatgcgggtaccggcgccgcgggtgcgaaccgca ggagggcgaacgggtggtggtgcgctgtgtgcacccgctgaggggtgtcatgtcc tgggcatccaagctgcagcgggtggtaccgatccgggtggtcatatggagctgttcc gttcgacctgcagctgcttgaaccccgagatcagcgggttgatgatacagaacgca ccttggccgggtacgagttgcgggagttacgtgtggaagtggtaccgcacctgcgc atattgcgacgccacagggtgcccctcaatatggacatatccatccgaagtcaagg gcggctcgaaacggcgacgaactcagttggcgtgcatcccggtgcaacaa	gtcaatcaccggcgctggcgtgcccggct tgttcttcgctgtgagcagcggtggcgtgc cccgttcgaggggttctgttaccagatca agacaccaggtatgggattgagtgatca ggtgtctacgtttccgatagggtgagcgaatc caccaggccgtgcttccagcggcgttgcctc tggaaatcgctcagcagacacggcgtgg gtagctccgaaacggggcgcgagcgctga actttctagtctgcagccagtcgggaacat gatcaaggggagaccggcagccgcaccc ggccggcgggcg
1483-1484	human gut metagenome genome assembly, contig: NODE_4377_length_367 4_cov_3.131498, whole genome shotgun sequence GENOME_ACESSION: UPWN01.1 GENOME_ID: 16899 CONTIG_ID: 4376 SOURCE: WGS DATE: 2019-01-27.	gtgtcagaatgacaaatfatgctttgtgtttagatgcgaatggaaacaaatggcaccac aaaaagagcagaagactggttcttattcgtaaaaacgtcgacattggttaataatc caatggtatacaacttaataatacaattccaatgggtggagatacagtaataaacgaatt gattggaattgaaaaatcacatagcaatgctgtattgtattaccgacttgcacacagat gtatatgaaattaaagattggataataaaacaaatgcgtagacaaagtaaaagct	tattaaataataaaaaagtgacatataactt ttgtttgtgcttaatagtaggtttcaagcctg agtgactgctactatcgaaagatatgttgcg atattgaactatgttagatagaaggtaaaaac acacctttagatgtaattctcagctctaaagtc tgtgagtgccaaccaagaacaatgctaatg tctgcatgataacagggaacacatatctct ctatctgacattggcaagaagagaatgctc cgaaagaaaggtgtcagaatgacaaat
1485-1486	Ga0066650_10013185 GENOME_ACESSION: Cry Gey StandDraft_18_SF_3300004239 GENOME_ID: 25052 CONTIG_ID: 13184 SOURCE: JGI DATE: 2019-01-27.	atgaaaaatacacacagaagtaggaagagagaatacacacacctacaataacttca caagttgtagctctgtggtgtgtctttaaacaagagagaaactctcagtgagcatggctta aaaaccttccaaacagcctcgtatgtgaacaaacagctcgcaagactggacaggatttg agaattcctgttttaatatatcgtggaacaccttgcgtccaacacaccagcaaaagcaa ggcaattgttgaacaaagaaaagcaaaagtaatacaaaagaaaacatttgttatacaattg acaatagcaacgggagagactaagcaaaagattacattaggaatagatagcggatatag tcaagtaggattcagcgcaaaagacagaaaaagagaattgatataggagaaactaacttt aagaaaaagatgtatcaaaaagttagaagaaagaaaaggtatagaaaacagaagaga aataagttatggtacagagaagcaagatttgataaccgagtaatttcaaaagaattgttagc accaagcatcacgacaaattggatacacatgttagattgattgaaaaaatcaaaagtttac ttccaatcacaaaaacaataattgaaatagccaacttggatacacagaagatgcagaatcc agaaatatacgaataaaatacaacaagggtgaattacaagggtatcatatcagagaatat ctttagacaagtttggcaggaaatgtgcttattgcaatcaaaagggttctcttagagggt gaacatattattccaaatcaaaaggagggaagtaacagagcatctaaccttacaatctcat gcagtaaatgcaatttagaaaaagggt	gtcaatcaccacaccctaaataatagatttta ggatgtggccttgagccgtgaggtttgagag cgattggttgattagagggtgaaaaagtatg aaaaatacacacagaagttaggaaagaga aatacatcacacacctacaatacttcacaagt tttagctctgtggtgtgtctttaaagagaga ggaaactctcagtgagcatggtttaaacc tttccaaacagcctcgtatgtgaacaaacag tctggcaagactggacaggatttgagaat

SEQ ID NO:	Name	IscB sequence	ωRNAs
1487-1488	Microcoleus chthonoplastes PCC 7420 scf_1103659003791 genomic scaffold, whole genome shotgun sequence GENOME_ACESSION: GCA_00015555.1_AS M15555v1_genomic GENOME_ID: 40746 CONTIG_ID: 53 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgcgcgtttctgttagacaaaaacttacagcctcttgacccttgccatccagcacgggc aagagaattactgaacaaaggagggtctaaagcattcaagcgtatccattactattgtc ctgcaagacagaaccgttgaaagagtcagtcacgacccacatcgggtcaaaatagacc tggcagcaaaacacccggattgctgtgttcaggaaagaacaggggcagtaacaagc gccattgaaatctcacaccgagggaacaataaagattccctcttagctcgcaggctat taagaaggggagccgtaaccgtaaacccgttaccgccagcctcgtttctaaacagaa cccgtaaaaccggatggtaccaccatcgctggagagtcggattgccaacattgaaactt gggtacgacggattaaaaactctgtccgatttcagccatctctcaagagctagtcagggtt gacttgcagcaaatgcaaaacccctgaaattagcggagttgagttaccaagaggtgaact gtttgttttgagggtcaaaagattattactgtccaaatggaacagaaatgcgtactcgc agggtgaaaacgttccgttcgagattgaacacattctgcaaaaagtaagggtgttcaaa ccgagttagtaattcttgccttagttgtcattcttgaaccaagttaaagggaacagacctgtt gaagaattcctaaaaaagaagccaggagtccttaaacgggtattagctcaagccaaagc acctctaaagatgccgcagccgtcaatgctaccggatgggaactgtatcgaaggcttca gtcaactggtttactgttagaggttaggtctggaggtgcactaagttaaccgtaaaacc agagggtgattgaaaagctcacgtttcgtatgcagcctgtgtcggagcaatctactcgtgaa cgattataattcgaggaaatcaaacctctaaaattgccgcaaaaggacgcggaaccaga caacgttctgccttgataagtagcgtttcttaagctcatgtcctaaagccaagtatttc cagggtttcaaaactggcgacattgttaaagctgatttcaaaaaggtaagttcgcagggtc aataattggtcgaattgcgattcgatttagacctagttttgtcttcaggtiaccacacaaa gtttgatgtacatcccaataacttgagaaccattcataaagcagacggctatgaatacaat cctaa	gtcaacaactcaccgctaagtctatcgaact atagcgggagcttgaataacaggctctagt tgaccagactaagttctacgagaactacgtta tttagctatgacacccctaaatgcgtgccca gttttaggctctgtctacaggttttaaacagggt gtagagagttaaagcagtgagctgtacctaa caaggctagataacattgtcagggtcaacttt accagtaagttagaggtgacgtaatgcgc
1489-1490	Ga0302349_1003924 GENOME_ACESSION: PhyAR1v2_SF_3300031 370 GENOME_ID: 280832 CONTIG_ID: 3923 SOURCE: JGI DATE: 2019-06-13.	gtgactacgttcacacacgcgagaagaccaccaggccgtgtctctcagcggcctgc tctggaaatcgactcagcagacacccccggggcaggacgaaacggggcgtcagagc ctgggaccacccaggacacgggtgtgcacccgtggggcaggggagaccgcggctccg gccgcggcgtcacccgggaagggtcgtgtcccgccccggagcggcgggagtgatccc gtcacgcacagcatccggatccgaggggtggacctcgggtgcgcctcgcggcgcc gaggcgcgggtgtcgtctgtcgtggatcggcggggaagccgctcatgccacgactcccc gtcgtgcgcggcagctgtctgttccggcgtgcccagtgatcaggggtgcagcccttc gtcctccggatcgtcagccgaggggtcaggactcggaggtccagccctcgtcctg gcacgtacccgggattccgacataccggcgtgcctctggcgcgcgagcagggtcc ccgacccccgacccggcggggcagcagcgtccgacacggcctcttctcctcgtcgggt ggacatcgcggcgccgtgatccgcgacccgtctcggcagcgtccgctcgcggcg ggccggcgctcgcgaagctccgctaccgtgcggcgccgttcgacaaccgcggcgt gcccggggatggctgcggcggtcgtacccatcgccggcgagaccaccgtgacctg gcacggcgctggcgccgtggcgacccgtgacgcggatcgacctcaggttcccgcc ttcgacggcgtgcccgtgatcggccgatgtcggcggcggcggcggcggcggcggc ccctgcacggcaccgaggtgcgcgagtagtgcctggagcgcgacggcgccggcgtgcg tgtactcgggtgcgagcgccgtggcgccgctcgtccgctcagctggacctgtg agtccccggcgacggcgggcgagctgcccggcgaacctcgtcggcgctcgtc ccctgcaaccggcacaaggcgacccgaggtggaggttccctgcggcgccggccc ggcgtgctcgcgggtgcggcgagcctcgctcgtcatcaggaggacctcgtg tctcc	gtcaggcgccccaccacagacatcgacgg gccccaggggcacatcggttgatgtcaggtc tgaccagctcgagacaccgctcagaggga gggtactacgttcacacacgcgagaagac ccaccaggcgtgtctcctcagcgccgtgct ctggaatcgactcagcagacaaccccg ggcaggggacgaaacggggcgtcagacct gggaccaccaggacacgggtgtgcaccg tggcgaggggagaccgcgctccggcc gcggcgacccggggaagggtcgtgt
1491-1492	Leptolyngbya sp. IPPAS B-1204 Mongol_c4, whole genome shotgun sequence GENOME_ACESSION: GCA_003724315.1_AS M372431v1_genomic GENOME_ID: 202667 CONTIG_ID: 81 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgcgaattccagttgtttacgcgataatcaaccactcatgccaccactccagcgagag tcaggaggtgattgcatcaggtaaagccatcaagaagtgtctgacctgggagctgtt atggattagctctgcatgtgtcgaattcgttaactacgagtcgttccatctcgcgaacactc gcgggtatgcttgactgtgttcagttacgttaacgcctgctatttcagggtgattcgcaga cctccatttctcgcaggaactacacttaattgttccctccattggtggagtacgcaggaa gtatggcggaactacaacacggcatggagtcagaaaagggtgatgtgtcaaaagctgaa atggcggtcgggtctctgtcggctgggtcagtggtgacactcaagacagatttccgtg tctgatagcaattggaaacgcttagggcagttacggcatctaaagtgttgttattgccc aaatactgggtgtgtgttccggggcatcgttggcgagcctgcgcggagcgcataacc ccactcaaccaccgctcctatccctcccatgcctaaaggcagggtatctcggaggaa gtttga	gtcaggaaacccatagctagaagctaggg gcttctccaacaatttgggcaacgcaagt tctgactagccctagagccttcacctgtga cagacttccgaatacttccatgttcggattta ctctaagcctgattgttcagacgttgggttaaf gccaagacatcctgttgtaggtggcggaag ggacttaaacaaagcaatttgggtataccaa tcatcgca
1493-1494	Ga0114918_10006480 GENOME_ACESSION: LantaG_2_SF_33000091 49 GENOME_ID: 31359 CONTIG_ID: 6479 SOURCE: JGI DATE: 2019-01-27.	atgcctttagtctcaagaaagctagactcgtacttataaatggtaagctaaagtagccc gtagggtgctctttactattaagatactatattgctccagtgctataagcaggaagtaaga gcttctttaattccaagtgtctaaagttaggtatagcttcttctcaaatgggacttcttata ctcttctgaagtagagtaagacatgatatttcaagaaaatgaaaggaagatcttctacag aagaactcgtagaatagaaaaactgttatagagaatgcagattcctaataagaaaatca gatagaaagtttactccaactatgatgtctaaatttgaagtcagcaagagaatttcaag attatctaagctacttctattagtaattggatgttgttaaaaattctgttaagaagattatca aggacctaagatttagagtggttaaatctacaagggcaaaatttgaagagatagattta aatgtactattgttaaggtaaatctaaattgtatgaactacatgctaccacctaatatttag	gtgaatcacccatagtgcttcgacaactatg gcttcttctgtaggaggagttaaactgttca ccagactctgattagagaaatctaatctacgt tatttttagttagacacctaagaatccgcctc agttcttagctctgtagaggtctgttaaacag ggactaaagttccagtcacacttcagagacgc cagacatttctgtaagctattataacattgtt gagagggaagttaggattcttaattgttaata

SEQ ID NO:	Name	IscB sequence	ωRNAs
		<p>gagtgaaaggggagaagatgttttagacaatcttattacttatgtaaaaaatgtcatgtatc ttatcataaaggagagatagagcttaaaaaagattaaaagtaaaagaaaaatgatactgag tcagccatcattagaagaacttaaaaattcatgataatccaagaatctatggtttgaa gttaaaagctaagagaaaactattaaacttaaaagccctacacctattacaatgctgtagtttt tagaaatattatcagataatagttattatattaaaaatgtttctaaaggagattatcaaatgaca aaagggtgttagaagccagatgataaattcctaaaggtaaaatatttggaatttaaaagtttgac aaagttaattcaaaaaataattttactctattaaaggtagaatgagtagctgttttttatagg gatggatatttttagaaatacccttaaaaggtaaaactctaaaagctaaagaatgcaaaactaa tttcaaggagatcttcatgccttataacagagacggtagaagaaaatacatgtttacaatgcc atctaa</p>	<p>gattaaagatactcaatactcaggaatagga gatatacttaattgta</p>
1495-1496	<p>NODE_1208_length_400 4_cov_2.52747_ID_7945 90 GENOME_ACESSION: mgm4799983.3 GENOME_ID: 238380 CONTIG_ID: 1207 SOURCE: MG-RAST DATE: 2019-01-27.</p>	<p>atgaacagagttttgtattaaagtaaaacaggtgaagacgttgatgccctgccatccggcaa gggcaagacagctattgcatagcaaaaaagcaagagtttaaacgtctttatcctttactata cagtttaactcagcgttcaaaaggatataacacgctgttgaaactcaagtcgatccgggttc aaaacaaacaggtgtagggtgtgttttactatgggaaaaacaggtgtctgcccataatg cgcagtggttaacgcacggtgggacaggaataaaaaaacaatttagactctggcggtatgat aagaagagccagaagaaccgtataaaccgctatagacagccagcgttctttaatcgtgt agaatggtctaaagcgtttcaccgattggcctcctgtgtgtttatcactgttgtaacagtttaagt ttgatatgcagaagatggaaaaacgtcgcagttcaaggagtggaagtagcagcgtggcagc ctgtttgattatgaagtaaaagaattttgcttgagaataaccattatagctgcgtctattgtg gaacaaagatgtgccatttgaaaaagagcatgttatcccacgtagccgtggcgaggagta acagaataagtaactgtgtttatctgtcatgattgtaacagaaaaagataacttgcctaa tagaagtgttttaaaagataaccggcggtattgaaaaagataaaagcacaattaaaaagt tcgttaaaagatgctgcagcgggtaaacataaccgtaaacaaatcgtaaaaaactgtctg attttaaattgtccctgttctgacaggaacgggagcggaaacgaaatatacagagtaagtc agggatatgaaaaggagcattatagatgcgtgtgtgagcgaacgggagcgaag atatataccctaaaaagttaaagccgtgtgtgataaaaaaggaagaaagataaacaggc agatgtgtctgttgataaatatggttcttagaggaagcgaagagatcgaagatagtg gcatggatttaaacaggtgatatagtgaaagcagttgtgtgaaagggaagaaaaagg tgtatataaaggaaaagtagcggtaaggtaagtggaagcttcaacataaacgtaaaaca aggaagagtagaaggaataggctggaagaactgcgtaatgtcttaccgtttcgtatggta tagctatactactattag</p>	<p>gtcagtaacccccccctgtcggacgaggct tgtactgaaaacagccttatcgttttctaaa caagtcctccgacttggttaactgaccagcct tagtctgtgaaacgctgagttttacagactac gttaaaaaacagatgacacgttagaatgctt ctccagttctatgcagtgtcgaatacattaaa catttcttggggtaaggaaaagtgtgtgcta catgaaatgtttttaaacttggcgaggagata ttactttatgaaaaacgctgttttcataaaag ataaaaggtaacttattaaaaataaaaaag gaaagtatgaacaga</p>
1497-1498	<p>Ga0307376_10007847 GENOME_ACESSION: TR2MG_2_SF_33000315 78 GENOME_ID: 281014 CONTIG_ID: 7846 SOURCE: JGI DATE: 2019-06-13.</p>	<p>atgggtgatgtattggacaagcgcaagagaccatgaacccctgcaccgagaacggggc tcgcctactcttggaaacggccggggcggtgtgtcacaagatggcgccgttcacgatcc gggttaaaagcagtggtgtggccgcccagctgtccggtttacgctgaaactggaccgggt tccaaagtaaccggggggagccgcaatccgggacgggaaagaagtcgttgatgttacg agtgtcaccaccgaactgacatcaagacaaagatggacggccggcgggggcaacggc ggagccggccgggacggaagacacgttataggaacccctgttggggcaaccggcacc cggagaagtgcggccgtcgggtgggaacggaacacggcagccggtagtccgtc cctgcggccgggtaaggcacttctgtgacacggcgtcgggagacctggctgcccgc gtcgtctcgtgcccgtgggaagaaacctatttctgggtgaaaagatgcggcggtctt gcccgtacccgggattgcaatggaactgtgtccgttcgacacccagttgatggagaacc ccgacatccggggttgatgaccagcagggcactctgaccggctacgaggtccggga gtatctcctgagaaagctcgtgacgtgtgtgctactgtcgggtacctcggggatcc gggttctaacgtggagcagctggtaccacggaacccggctcaaggaccgaaagcagc gatcgggttcgaatctggttatcgcctgcaagacgtgcaatgacgcgaaggacaacctg cagccgggaagaatggttgaaacagtgtcagggttccggaaagaagattgatcaggtacg ggcgggaacctgcccgaatgttctaaagcaacttaagcaacc</p>	<p>gtcatagaccacgcctaaaggcggggggc ttgcggagaaaacccaaagtccacgtctt aaggcgagaagcggtatgaagtaccttc gggtacaaccatafcgcgctatgaccacc caagcccgccgggacatcaggcgagcg ggctacgttatctgtgcatgacaccttggga tgccttctagtcccatgccctgtcggcggtg gttaaacagtaagccgggtgacgcagcgag agtgtgtcggcgcaaaaagccgggataa cattgttagtagagatgccggcggaagcc ggcggtcagcaggtccgcaagggtgtc cgaaagggtatgcacaagatg</p>
1499-1500	<p>Ga0209800_10000857 GENOME_ACESSION: HigcelmetaSPAdes_2_SF _3300027800 GENOME_ID: 27174 CONTIG_ID: 856 SOURCE: JGI DATE: 2019-01-27.</p>	<p>atggctgatgccgacaatccggcaaatgccgaattttgctcaagcagggcaggggcaa ggcaaaagtgattcaaaagaaaccccttgcatttactgtttatgagaccaccgagcat attcagccggtgaccgttgggattgatgacggagggatcaatacgggattgcagcgggt tcaaacggcaaggtttatttcaacagaaagttgtttactgttcggagtcgaagtcaaaactg gatacccgagggaataaccgtataatccagaagaaacccgaaaaaaggtatcgaaagt caagatttctgaacagaaagtcattccacatgcaaggtgtgtggcgggaatgcc cggtatctcaggtgatctgtcgtatcctgcctgaacagagcggatgggttcatcaaaaata tgcgaagatcaaaaaaggtttttccgaatcccacatcaatcaaggcaaaaaaatgatgc gattatccgggtgtcaagcagatccactgccatttccagaattgtgtggaagatctct atttgcatttccaggcaatggagaatccggacatttccggtagcagttatcagatggaga tctgctttatcacaagaatttcaaacaggcgtgtctgtgtgacaaagttcagatgccgtg tttgcgtgtgcgcaaaaaactgcaatgccatcatatccgtcaaggggcaaaaggcggg acagacaagctctcaaatctgatgacgtttgtgatgtctgcatgatcagatcataaaga agggtcacaacttccgaaacaaaaggttcttctacatgtcggcagcagatgccagca gggaaagcactatctgcaagctaagttgtcacgaatcgcgccattacggacgacattcgg gtatatacaccagtcattttcgaaacatccgggatagaaaaatcccatgtcaatgatgtc ttatcattgcagataaacaggcaagtcctctggaccggcagatacagacaaaacatgtgc</p>	<p>gtcaactacccccctgaaatcaaaagattcag aaggggctgtgaaaagcccaagttgactag cctcagttctttattaggactacgttccgc aggatgtatgacatttggatgtaatgccca gtccgaagctctgtcgtggtctgtaaaaagc tctgtgaggaagagcggtaaacacgttgt gaagcctgctgaacattggcgaggcgaaac cttaccctcttggagggtgtgcataactga aaggtttttatgaaa</p>

SEQ ID NO:	Name	IscB sequence	ωRNAs
		agtcacggaaaagagttgcatgaagcaactgcaagaaaaggagaaaaccccgaa ccgaaccgagaacggataacaaaacgtattaccctgaagggtttt	
1501-1502	Ga0180438_10006853 GENOME_ACESSION: SS_ome_18_SF_3300017 971 GENOME_ID: 35200 CONTIG_ID: 6852 SOURCE: JGI DATE: 2019-01-27.	ttgagagtacatctatgtgtgaacatcagggggcaaccctcatgtctaccacacaac acaacggcaagaagtgtgtaaaacaaggcaaggcaaaagtgttcagagatgtccttttac cattcagttgaactatgcaactggagaggcgacacaaccgattaaactgggagtgagatat tggatatacgaacattggatttagtccaagaccgataagttgaagtcataagtgccacg ctgaccttgcgaataatgtttctaacaagctagaggaagacgcaggatcgtgacacg cgtagaggcagattaggggtatcgaccaccacgatttgataacagaacacgaccggaagg atgggtcgccttagtattcagcatagggcatgactacatattcattagttgaaacacttga aacactgctaccgattacctacgaaaaggtagaagtgccgaatttcgatgcgcacaaaat gcagcatccagaaattacaggcgctgcataatcagcaggcggaactcagggtctatgaag tcagagaatacctgttagacaaatggggacgaaagtgcgcctactgtgtgaaaagaac gtgccattagaagttagacacatcattcctagaagttagaggtggaacagaccgagtatcg aatttgacgatcgcatgtaggaaatgcaatctaaagaaaggagataagaccgagagga gtttgggtatcctactatacagaagcaggcaaaagcaacctctgaaagcacctgcatgtctg aataatattcgttggaaattagtagaacagtttaggagcagagcatacctatgtatgacgtca ctaagtacgaacgtacaaattagggctggaaaagtctcatgttaatgacgcgtttgtcatt gctggaggacaactcaagagcgctgcagggtcttatgaagtgtgcaagtcaggcgcaa taatcgttcgtacagacgaatcgaaaggattcaaacctctattaggaacacgcatatc cagttgcaacccacgatgtgtaactataatgggaaagtatacgaagtcgaagggttat tcaattatggcactgggtacgattacgtgactcacaggagcggtcgtcaacacaca ttaaaaacgtgaaactgttgaatatggaaaaggcttagtattaccttcaaaaacaactaa	gtcaatcacccttagcaagttagggggct tgtccctgagggagtgagagtaattgtgtgat tagcctaagagaggtcatgttggaaacaac atggcgacatctacgttatcaggaggttca agaacacaccaggggatgtcttctatgtccc ttgctctgtaaatgtgtttaaagagagagg caactctcagtgatcaccgtatagtgactgact gataactttggcggaagagaaccaaccgactt taaaagcaattaaaggagagataggacttg agagtacct
1503-1504	Ga0098074_1007018 GENOME_ACESSION: 32 taG_2_SF_33000067 90 GENOME_ID: 20884 CONTIG_ID: 7017 SOURCE: JGI DATE: 2019-01-27.	atgcagtcacaacagggtcttctgtggtaattagcagttgggggaaccgctgatgccgtgccat ccggcacgagcgcaaaagtgtcctaagttggagcgggcaagattaccgttaccac cgtttgcaatccagttgttgattgtgagaatcaggacgtccaaccgtagaggtcagggt cgatcctggcagcaagacgacggcgatgtctgtgtggtcctatggcaaacgtgggggta gggtcgtgtggcgccgagagctacagcagctggggctgattatcaagatcggtctgag aagcgtagtgtgttcgacggggccgacggccatgaaatgcccgtacagggcccaaac gctttgacaatcgaaaactgaacagttgacgttgggttggcccgcttacagagcag gggtggacaacattgatgtgtgtagggcggtctcgaagtagcaccgttgctgctggg cgggcggtcgagacggccaagtgtgataccagaagatggaaaatcccgaatcggg gattgaataaccagcggggtacgttggccggtgggacgaacgcgaatacctactggag aaatgggggtcggtgtgtgttctgtgtgtgctgagaatgtgccgtccagggtgagcattt gataccaaaatcgagagcgggtacttctgagcgtacatatacaaaaggcagatgcttggg tccttgcatacaaaaggaggccaacaaactgcggccgagttcggtttccagaactgat ggacaaggccaagcagctcttaaatcaagggtatgctgctcttatgaattccacagctat gccataggcgatgcagtgaagcgacacgtgcccactacgttctggacagggtcaagga ccaaaatgaaccgcataaatcagggtatcccaagctcactggatcgacgtgctgtg tgggagaggctgtgtggcgacgtgttcttggcgagggttaccggccttccagataagag ccgtaggtcgggggagccgacagatgtgtgcgatggatcgtatggcttcccgctacc gggtccaaaacgggtcaaacgggttttgggttccagacagggtatgtcaaaaggcagac atcccaaaaggcgaagtatgctgggctacacattggccgtatctctgtccgtgctctg gaggttttagttgaaatgggtaccaagggtcaactggaatatggcagctaatccagcg ggctgatgcatacgaatatggatgattgactaa	ggacacaccgttccgcttaccatgtgg ggcgaaacctaacctttctgtgtgaaatggat agggtcgtctacgaaccagtcacagccag gggataatcgtctgtattatcccaccgggt acgttaccgtgaatgtataggcaccgtggg gtgtcttctagccccagctctcggtgtgt agttaaacagacgtgataaggagggtcggt gctgccagcgagaaaccagtgataacatt gacgaggagacgggcaattgggcatagc cccgatgcctaaaccaccgaaagggtgagaa cccgcgtaagccgggaaacagaagag aagggaatgcagtcacaacagg
1505-1506	Draft_11733483 GENOME_ACESSION: IMG_3300000558_SF_3 300000558 GENOME_ID: 27455 CONTIG_ID: 35644 SOURCE: JGI DATE: 2019-01-27.	atgaagggtatatgtaatacaacgacgggagaccgtttagtgcctttagcccagagaacc gctcgtttacttttacgagataaggccgcgaagtaatacaacgagatcccttactatcaa aattctagtcggagttaaaggctacacacaagatttaactcttggattgatcctgggtcaag atacattgtgttctgcccgttagagatgataaaaatcaagcttattatctatcacaagtagagca cgctactgatgttaaagcgaaatggtatcaacgaagaatgtatcgacgaactagacgttaa tcgtaaaactcgatatacgttaagccaagatttatgaatcgaaaagctagtactaaagacgat agatatccaccaacgctagaatctaataatggagcaatcgtgaagagatcaatttctgtat gtaataatttcccgttaagaattgtatcagagattgcgaagtttgatacaagtgctcttac gaatccaacgtacttgatgatcattgggtgtatcaacgaggtccacagcttggattctaca atactaaagcgtatatactctttagagatacttatactgtcagtttgaagaacaagagaa aagactctcgattacatgttcatcatattcaacacaagtttcaaggcggaacggtatcaacc gaataacttaacacgcttggcaaatcgtcatgatgatcgcacaagaaaaagattatgct taatacaaaagcatttaaacacataaataatcttaaacatgcttcacagatgaatgtttatgt agtatgattaaaagtagattgtttctggatcttatacagacgctgtgtggaattgcaaaa gggtgtcagagaatttcaataatccaaaggagcattattgggacgttcttggagactt gaaaacggaaacactccaaaactattaatagaccgagtagtgatgaagaagtgtgtcgcg aaaggatcttatcaattgactaacggtaagagatctgagaacgggtaccacactggaag atttgtgggtatcgatctgggacaaagtgttctatcaagatcaacagatttctgaagaggt cgtatgtcactgggtatgtaacttctgtgacatttccgggtactcagtttaagattcgtccgat gcctaagtttaataaaaagacaagaattggagcaagagatcattgattatgaccacataa	gtgaactaccattgattaaagatgatggg ctttggatcaaacgggagccggagatcacc agactaagtactaagaattagaactacgtta ttatgatcgatacctttcggtgccctc agccgattgctctatcgtggctctgtaaacat cctcgggaaagggaagtcaaccacatcac gaagcatagataacattgtcagaggaagt cgaatttactcgtgttaacagggtggtaaat acgcattactccttaaggagctcatatgaag

SEQ ID NO:	Name	IscB sequence	ωRNAs
1507-1508	Ga0104854_10039936 GENOME_ACESSION: IMG_3300007643_SF_3 300007643 GENOME_ID: 311670 CONTIG_ID: 4122 SOURCE: JGI DATE: 2020-04-07.	atgaaccagggtattgtgctcagttaccacgagacagccattgatgccgtgctccccagctc gtgctgcgacgattgttaaaagccaagcgagcgccggtctccgacggcaacctttacg attattctcaaggtcgaatcgctggggctacccaaccagtggttgccaaatggatccg ggaagccgcacacccggccttgacatcgtggggcgagttcccgatcagggccggctg ttctgtgggctggcgagctgcacatcagagacacctcattcgccaaaagcttctagccc gtcgggcaatgcgcccgaaccgtcgatctcgaagaccggatcgggctccccgattt ctcaatcggcgccgcagaagactggtctccacctctctcaagtcgcgtgtggataat gtcgtgagttggacaagaaaaatccaacgtctgtgccgattacctcgatcgcatggaa cttgtcgggttgataccaaaaatcatgaatctgagattgaggggtggcataccaaa gggtgtgagctgtttggctatgaagtcggagatctgtcgcgagaagtggaaccgtactt gcgcatattgtcacaagatcaatgtccgttgaaattgagcatttggtccgaggatcgt ggtggcagcgaccgtgtggaacacctcacgcttgctgcacgaaatgtaataacgaaa gggtaatcagacggctacagatttggtatcccgattgatgaagcagggcccagcaac cgtaaaggatatgctggcgccgtaacaataccgctggcgctctatcggaaccgtcag tcactcgtctccagctcagcgtggtcggcggaaccaataacatcgtacccaa caaggctaccccaaaacacattgattgatcgccgatgattggagacacatgctcctg cttgatcctaataatgagaccgctgacctacagctgtgggtcgtggcactcggatgctg ttcggactaacaagtctgttttccgttagcaaggcgaggtcgaattaacgactctgtgt tttcaaacggcgatctgtcgaactaatccagccgctgggcaagtgatgctggtgatggt ttgggcacctcaaaagcgatacgtctgacagggtattgcaactcagaacatcactggga aagtcggcgccctcatcgcgcttactattttgcaacgaatggcggtcacaatttacc caagaaagggttccatctctagtgcgtatcgaatcaatcttgggaataaatcagaatc tatttcgataggagaatcfaatggaagcaaacctctatcgatcagcgatgagtgca	ccgtctcggtcctatctatatgtccaacgtgg tagcaaaaaaggtccgaggggacattgaaaa taattctgacctgtctcaaccaggctccgtc cccatcaattcagcaatgattgtgatggg actccgttctgcggaatacatagggtacct gggtgcttcacagtcggcgccactacgg tcagcgattaatcagcagcgctcgggaca acgggcagtgctgtggcaaaaacccgc gcgagaacattgcgaggtgaccataactc ccatcatgggacgtggctgttaaagccaataa agaaagggaagtcagatgaaccag
1509-1510	Ga0376082_0008098 GENOME_ACESSION: CLC_S1_MG_FICUS_ FD_SF_3300035674 GENOME_ID: 312359 CONTIG_ID: 8097 SOURCE: JGI DATE: 2020-04-07.	atgcgagttccggtattatcaccaaggagaaacctctgatgccagctaaagccagcag ggtgagcgctgctgctaaaagagggttcgctaagggtgtcaagaacaaactcaaaatctt tcaagttcagttagtagaagaacctatcgggacagaagttcaaggctgcgtagccggaac cgatcctgggaagctgtacaccgggacttgcgggtacagaccgcaaaagacacgcttggc tagggcacattgtcctccctatgcagtcacccgcaaacgattagatcaacggcgatgat gcgtagggggcggaagggtcgcgaattgatcgcaaacctaccttatgccagcgtaaatc acagacaaaaacgattcaataaccgccggagcaaaaaactgcctccgtgatccgagc aagtcggcagttagagattagcttgataaagacgctatgctctatcaccataaactgca tagtctgggaaaagggttagtagtaagggtcttcatgctctctccgtaagtcaagcgca gtactgggcaatcaataatctagaaaaagctgcggatgcgagggtcagcagctagaag gctgggaaaccttaacctccgacaataatctgggcttgcacaagcaaaagcatcaaaaa ggagacgctataccagaaccatgctgtcgtatgctgtgctctgctccagttgtttgtt ggtacaaagcaacacaaggtaagcaggtggtgggaagggttcggtcaagttacgcc cgctcctttgtgtcattcgcctccaccgattagccgtagacagctccacctaatgcttcc ctcaaaaggtgggtgctgcgcaagtatggtggaaccgttaactcgtcatgatttgcga gggggattatgttgcgctgaacaagcaggacgcatctactacggctgggtgacggcg acaccgcaagcaagttttagtgaggttctagctggaatcggctagggcaattttctgtt aaaaaagtcctgtctgttcagcgaagcaggggttaatcgtcgtgcttcagccgattgt caaatccccctattgaaagggtcgatttga	ttcaactaccctaggtgctacgctcaagctg agggattgcccggagcaggggtgacg caagttgtaatacagccgctgagactagcc tcggtaagcacttctaggatcttctctagtc ggattatcgcataaacctgtttacggtttc atctaaaaaggtggacatcttaggttattgt caaaagagcgttctgaaggacacatactc tgaaaggagaattatctaaaatgcga
1511-1512	Ga0373620_0000796 GENOME_ACESSION: YL18G_03S_MG_FD_\$ F_3300036760 GENOME_ID: 312746 CONTIG_ID: 795 SOURCE: JGI DATE: 2020-04-07.	atgaagcagaagttaaaacagagaactaaaaagaacacacctacggatgctccacaagt ccgtagcaactgtgctgtcattaaacaagaggaacctcttagtgtgcaggtcctaaaaa acctgttttaacaatcccgaagggtattatctcaacatacagagagcgttaagccattgt gtatgttaattagttatgaagcgcactttaaagccatgcactccagcaaaagcgaaaaaa ttacttaacggttaatcgctgaaggttataatgtatcattactattaaattaaattttag tgcgagaatcaggcgcaagaatacagacttggaggttgattccggatttaacaattgttgaa tgtcagccattacccaacgaaagaagtattgtctcaacaattacgcttgacagcaaaac atcaagtaggctgaccgaaagaagaatgtatcgaggttaagagaacaattatggta tcgaaaatcaagattttgaatcgagggaaatcaaaaagaaggttggcttccacctcaattc aaaggagatatgacgtcatttgacgctaatacaagaatgcataaataatttgcataattca aaagtcacaatcgaacttggcaattttgacatcgctaaaattataaattccagacattatgg cattggtatcagcaaggagattttagtgctaccataatagagacatactgattgctgc gagaacatggtttatgccagttgtgccaacaagaatttacaagaagaaatccatcgataatt catcactgcaaaagcgtcagcaaacagggtctgaccgttcagaaaacttggcgataatt cataagaaggtgcacaagaattgcataataatgattaaaaatttactccaaaagagt aaagcaatacattatgtccatcatcaacataaattcaaacagaatttccagatgtaaat atcacctttgatacaaaaacttctgttgaagacaaaaactgggattgaaaaaatccacg ccactgatgcgttgtattgttgcgggagcaacacaagagcgatacgggtcaattacaat tcaacaaaagcatagaaataacaggcgattcaattaacagaaaaggatttgccttca atccgaaaagcaacgatgtctatcagccaaaagatttgattgattaacggaaaaatattt tcagttggcgggatgcaagacaaggcacacaggtttaaattgaagattcaaaaagggtt tattctataaaaagtgtcgaaaaaatgtatcatttgggtgatttttttaactaaaaaaat gaaacagaaaaaataatgccttatcctctcaactaaatcaacagggttttacaggaaaat	gtcaatcacccaatcaaaaagaaagggtt gagccgtgaggtttaaagggtacaggttgat tagggagcgtaaaattgaaacaatgaagca gaagttaaacagagaactaaaaagaacac acctacggatgctccacaagtcctagcaaa ctgtggtctgtcattaaacaagaggaaact cttagtgtgtcaggtcctaaaaacctgttttaac aatcccgaagtgatttctcaacatacaga gaggttaaaagccatt

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ctcacaacaagcgggtgaagcaggatgagattgatttcacccagcacaataataaagc cagtagattggctatctcccagcttaaataggtctgagggctctatcaacacaatgattgg aaatfaaaaaggaaatccgagactcagactacgggtgcattacctcagcatcatcagca ggacaacgattcgacataaaaaataatagcgacaagccattaaaaataatctcatcatcac ctattgtattagcctaacgaatccactaccattccggctctacacctcactgacagac caagtagagaagcgaagcgaatactggtatcggaaggaatggctgagagcgaagctg acaccgagatgcatatcgtttattcactctccacgaaaatccaactctcaaaatgtcatc aagactcgatagtttccgcaatcattgacatagcgagaaccaaagagcacaattatccag cttfaaacctgcataccctaaagccaaagacggcaaatgcgtatcgagctcacctatcgc ggattaatcaagtcctgacagatagcggtatgataaaagttatggacgctcatatagttat gaagatgactatgagtcgagtatctccagccgagaataaaataccataagccaaaag ttgctaagaccgaagctgaaaataatgcacgccaaatggccgagcttattctgtcgccat actgaatgacggaacaaagcattatcctcatgagatattggaagctcgcgaaaataga acaaatgagtagccggagcgaagcgaactatctataccgaatggaaaacagatagta caagaagtgtgcaatcagaagtcattataatcctgccaaaaggactacgctaccgga atacatccaacgagcaattatgattgagtagaaaatcaagcatcatgagtagcagta aatcgggtgcgggcaaaaaagaggtggaatggaattttcaataatcaaaaattga ctga	
1513-1514	Ga0180008_1007149 GENOME_ACESSION: MM_PW_MetaG_2_\$F_ 3300014613 GENOME_ID: 32274 CONTIG_ID: 7148 SOURCE: JGI DATE: 2019-01-27.	atgagctacgtattttttgacacgaacaagcaaccactcaacccgggttcacaggttg ggcaagaagctattatcatcaggacgtcgccgagtagatataagataatccattaccatta tcttgaaggctccatcccttcagcgggaataacacacattaagaattgaccccg cagtaagacaacgggactgctattgtcaatgacgtttcaggggatgactgtttgcccgtg aacttaaccatagggggcagcagataaagaagtcattggatagcagaaggcctattcgg agaggcagaaggaataggaaaacgagataccgcaaacacgtttcaataaccgagga ggcaagaaggtggttccaccatcgtcttagtagagtagagaacattgagacgtgg gtagcccggtacaaaagttatgtatgattacagcgattagccttgaattggttaagttgat atgcaagcaatggagaatcctgagataatcgggttggaatatacaacaggagagaactcaa gggtatgaggttagagaatatctccttgagaagtcaatcgtaatgctctactcggcgt tgagaatgtaccattacaggtagaacatattgtaccaagagcaaggggtgggacaatag agttagtaatttaactctgtctgcaacatgtaataaaaagaagggaaccaaacagca aaagagtttgggtatccaaagctacatcccttgaagaacattgaaggacgtctg ctgttaatgctacccgggtgggcatatctgctggacggaagcaacgggattgctataga agttaagctactgtggtagaacgaagtacaatcgtagtataagaagctacctaagacaca ctggtctgatgagctgtgttagggcgtcaacgccagaagctcttgatgtcgaaggat acgtcctttggataattactgccactggcggtgtgtcaagacagatgtccgtgtagataagt acgggttcccgccacatcggtcaagaagttcaaacgtgtcatgttttcaaacaggcga tatgtttaaagctgtttaccactgggaagaatacaggagcgtatctggacgtatagct attcgggcttccggcagttttacattaagacaccaaatagtaccatccaaggcattagcta tcgggtactgccaaactgtgcagcgactgtatggttatactacgagaggagacggcattc ctccccgctga	gtcaactacccctcctaattctccgaatcaga aggggcttgcgaagcgagctcaagggtg accagctatagttctttagaactacgttagg aacgaatgtataggtactgtcgtgggtctca ccagcccgacgtctacggtatgtgtttaa cagctaaaggaggataagcagtgacgcata caaaaaccgtccgcaacattagcgaggt gactattactcttactctcgagtaaggagg tgtctt
1515-1516	human gut metagenome genome assembly, contig: NODE_1050_length_109 80_cov_1.51689, whole genome shotgun sequence GENOME_ACESSION: UXMT01.1 GENOME_ID: 17221 CONTIG_ID: 1049 SOURCE: WGS DATE: 2019-01-27.	atggtatattgacttaacaaaaacggacaaccattatgccaaacacagacacggaaag gtaagaaggcttttaagataaacaagccaaaggttattaaagctgtccattacgataca gttgtgtataatacaacaaatgcatacaaaatattacattaggtgtagatgccgttctaag catattggattatctgcaactacaaaagataaagattgtttgaagctgatgagaacttctga acgatatacaaaaacttctgaggcagcaagaaaattctgactcaaggcgaatcgtaa aacccgttacagaaaaagcagtttaacaatcgagtatccagcaacataaaggttggtg gctccaagtattgagcataaaatcagactattttgcaatggtgaaaaagctacacaagat gttgcctattacaaaaattgtatgagaacacgcatgtttgatatgaaattgcttaagctca actggaaaggtgaaccattccaaaaggagcgattaccagaaggtgagcttacagggt ggaaatcagggaatataatttccaccgagataaattacacctgtcaatgggttaaaggtaaa tcgaaagatttaattcttgaacacaccacatgcctactggaagcgaccatacaaaaca aacctcaagctgattacactttgtaacacttgcattgatagtaaatatcaaaaaagaa gctaattaggctttgggttgggagccaaaataaacaactctataaacacgcggccttta tgaatgttatgagatgggtattttataatcggctaaaggaaatttatgcaaatgtttctatgact tatgtttatattacgaaaaacactcgtataaaaaacaattaccaaaagacacattatttagat gcacgttgattatgggaaatccaaaggcctaagtagtgaggagatttttactataaaaa ggtcagatgtcataatcgtcagctatataaagctaatacgtttaaagggtggcatccgaaaa cgtaatcaagctgagtatacagtcgaaggttttaactctttagatagtagaataatcagaat cacgaataattttatgttgaagggcatcagggtttctttagatagaaatttaaatggcca aaagggttaacaaaggctctgtcagtttaagaatttaaaataaagaacaaacaaact atttaattgaagggtgtacggtgatacaagagatgatttagctctctatga	gtatagtaataaccacgactaaagccgca ggcttgcaaaagcttttattgactagcctaag tgtttcaaacactcgttaaaagaaatata gttaccaaaagggtttatgactagccttttgc ctaaggctgggtgtaaacagttctgatgggt aggaacaggtgttccgataaaaacctctat taacatagcgggaaggtatcttacagttgata tgtactggcttatagcatacaacatacactat taagaaaaaggaggtgtttgtatg
1517-1518	Ga0268283_1011674 GENOME_ACESSION: saktaG_17_\$F_33000282 83 GENOME_ID: 37735	atgccatgttccaagaacagcgtcgtattctctacgcgataaatgtgctaaagtaatctc acgagatccatttacaatttaatttaattaggtccaggttatcacacaaccattagttga ggattatgacctggtacaaaatataataggttcagcagctcgaaccaagaaaaacaaagcaa tttattatcagagattcatcaactacagatgttactgataagatgaagcaaaagaaagatgt	gtggactaccaggaggcgttagctccgg acttcagagtaactgtgaagcctaagatcacc agactcagtagtgaagcgtactacgact atctatgagtagatactcgcaggtgccct

SEQ ID NO:	Name	IscB sequence	ωRNAs
	CONTIG_ID: 11673 SOURCE: JGI DATE: 2019-01-27.	atagaagaactagacgcggaagaaaactgatacgaacaaagatttctaatcgaagaatagcacaagactaatcgtttaccacacactgaaacaaagtttcacactattcagaaagaagttgagtttgaactaagatactccaatcaccaaaatgtacatcgagatgggaaagtgtgacatgcacgcaatgatgaacccagctgtatfaaatcactcgtgatgtatcaacaaggtaacgaatttggtttacaacacaaagcttatgtttgaatcgagatgattatcatgtcagtgttgaagggaagagaaaagccctcatttaccatgtcaccacataatctatcgaaagaaatgggtgtagtagagctgataactgttagttgttgaactgtgcacgatgatacagcctaacaagatt	cagctcgtcgtctatcgtggtcttgaacaacctctgaggaaacggaagcacaaccattacgaagcatagatattcgtcgagagggaagttggaaagctactggtataggttagacgctacacattactctgaaaggagacataaatttg
1519-1520	uncultured Methanobrevibacter sp. isolate RUG10470 genome assembly, contig: RUG10470_asm_5, whole genome shotgun sequence GENOME_ACESSION: GCA_902759475.1_Rum en_uncultured_genome_ RUG10470_genomic GENOME_ID: 344641 CONTIG_ID: 4 SOURCE: NCBI Prokaryotes DATE: 2020-04-22.	atgttcgtttatgtactaaatgcattggtgaaccgttaattccatgtaaccaagaaaagc aagaaaattactgaaagaaaagaaagcaaaagtggtgaacaggacaccattaccatac aattgtgtatggatctctgtgtataaacaacgggtaactcgggtgtgacccggaagca aatacattggattatcagcaaccaccaataaaaaagaattattaaagcaacagttgaatta agacaagacatacctaattgttagaaagtgaaaaactcagcgcaagaatcgaggag caaatcaggttatcgtccggcagcgttttaataatcgtgtaaaaaaggaaaactggcacc agttatcagcataaaatgacagccatttaacaataatataacgtatttgcataataatccca tagaaaacatcattgttgaactcgtgaattgacatgcataaacataagaaacccaagtgt tctgtggtcggattatcagcatggtcgtgcagaaaggttttataatgttaaatcagctgttta ccgtgacaaattacacctgcaaatatgcggagcaaaaataaccacacgaattgtatc atatcagattcagaagtgtggtgtccaatcgcattgacaaatctagtaacattatgcaaa gaatgccatgataaagtacatagcggcgaattgaaatatacaaaaaagcaaatcattca aacattccagctcatatgaacatcagagaaaaaattatacaatcttataaaaaagaattca gaaatgttttgagactttcgggtatttaaccaagtatactcgtgaaaaattggcctgttta gtctcattgcaatgatcgtttgttaattatgcacaaattgtgagccgaacatcgaattgtga tattcatataagaagttagacgacataatcgcacaaatcataaattcaaacctagtaagg tggaaaacgtcaacgcaatcagtcaccatattgttaaatggtttagaagatagataaag taattgtacaacggtaaggatgtttataacaagtaagcgtgacagcgggtattttcagttaa aacatttgatggaactgttaattagccaaggagtaagtagtaagcgggtgaaattattgaa cctattaaaggttggttaatagattggaggttgcaattcctccccacctaagagggtcg gggtatccttgccataa	gttgattaaactaccacggctgtagaagcc gttgatttctagagcttttctaggaatccact agtttaatagactcagaccattattgtctac gttactcaagaattattaggtaccgtggaatgt ttaatcccgattccacgctctacgtgaagtga ttaacaggttctgagatggttaggaacagtg ttgctacaattttaaaccttgggataacattgt cgagggataaaataactcaggaataggag gtaaacgtttaatgttc
1521-1522	Ga0401685_0570 GENOME_ACESSION: 1704MEGAvirome_2_F D_SF_3300036182 GENOME_ID: 307114 CONTIG_ID: 569 SOURCE: JGI DATE: 2020-04-07.	ttgaaagttccagttttgagaccgcaataaaaagccgttaatgccttgcacacataaagagc gagaaaattgatgaaagagggtgaagccaaaccttactggtctcgcggagttttctgcatt attttgcaaaaagagccgagcgggtcgagccaagcaaaaatcgtagtaggcgtggtatag cgttccaaaatggaaggctatacgtttaaagtcgaagagacaggttttgaacatcaa tcagaagcggctacttgggttaagcgaaggtttaaactagggcagaaagctagcga ctcggagaactgaaacaatccatcctgggttaataaataccacggcgttgcctgagaa aaaaccgaataccggcgcaacaaggctcgttgggtttgaaactgcgagttgtaaat ggctaagcaaatcttatccgatttcgcacgtcgtgttgaggatgtgaaagcggtaacaaa gtctgcgaacgtgcgtggaataaaagtttttagtccgctacaagcagcgaagcgttgtt ttacacgtcgaagtgtgagaactcggtagcgtcgcacaaaagcaaaagcaaaatggat gtgaactttaatacgcattgcgtgactcgtggttatggcagacaattgctttaaggcag acgggtgttgacaatagcagagttctgtttatcagccgttaaaccttgcctgcagacaact gcataagttcaatgctaaaaagggtgtggtggttaattacggaggacgcgaagtttg ggctgaaccgaggagcgttagtcaaacatgtaaaacacggcttatgtttatgttggcggaa cctcaagggggaaaattcgtctacacgcagccgacacttacgaaagattgtaccgaaatg caaatcggctgactcgaactcgaacaacatttcgttggcgggtccagtggttga	gccaatgggctgtgaaagaaagaaacacac ttctgaagcaattaaaagtagcactaccctac cggcactcgcgaatcgagagcgcaaa cttagcagcagcggcgtgtagccgcaag agtagcgtgaaaatgttctgtatttttacc tctacggcgtgaaagtgagggaagcagtttaa aacttggcgggtgaaccgttaaaacaaaagg gagaagttag
1523-1524	Ga0400245_005261 GENOME_ACESSION: CampLake17_FD_SF_33 00037285 GENOME_ID: 311512 CONTIG_ID: 5260 SOURCE: JGI DATE: 2020-04-07.	atgtcggcttctgtgttgatagacgaagcacccttctggtgcacccgaaaggcgt gcgcggcttttctcggcgtggtcgggtggtggtggtggtggtatccgtttacgatcc ggctgaaagactgtccgttggcgtcacacagccggtcgtatcaagatcgtatccggc agcaaaaccaccgggatcgatggttcgagaaaacggcgaagcaacacgttctgg ccttggcgtatcaagccaccggcgcggcgtatagcgcacacgttaacaaagca cggtttccctcgtgactatctgtgcgcaagaagaaagttcaagggtttcagaccgggtgac atggtgcgagccagatgccgacaggcaggaagccgacatctggccggcggcgggtcg cgttcgcaagaccggcagcttcaatccagacacggcggcggcgcatacaaggcat cttatcgtgtatttcacattaa	gtcaactaccccgccctgaaggacggggct tgtagaaaacacccctgcaagccaggttgac cagggaaagcggacaccaaccgcaccgt ttatcagaggtcgtgaagactaacgccgaa cgcttccctcagttcggcgtctgaaaggctg ggatcagcgttgcgaaaggcaagcaccg aagggtccggtcaccacgcaagtaggagcc gggtatagacattcccaggggaagagagg cttcggcctccgtcacaaggccgtaaggg cattcatgaaagaaaacgatgtcg
1525-1526	Ga0394872_0000490 GENOME_ACESSION: CP_S2_MG_FD_SF_330 0037308 GENOME_ID: 312363 CONTIG_ID: 489 SOURCE: JGI DATE: 2020-04-07.	atgttaagagtaccgtattagacaaaacaacggacctttaatgccacttggccagtc gcgctagacgttggctcaaaagaggcaagccatcggcaagtcaacaattaggacaa tttatgttcaactccaagacatcctagtgaacaacaaactcaacaatcgtattgtatt gatccggtaaacgtttctcagggattgggttcaatcagctcttcatcattagcaggtatt acttacacagggtgtgttctgagatagctcattccaagtatcgtttccgattttaattgga aacggctgggacaatttactgcctcaaaagtggagttgctccaacgctccacaggactaa tttgaacaggaaacagatttctgttggaactgtttaa	Aacgttaattgaaagccattgaaattcatttc aaggcggtttatctggttgggtcaagatcga gtattagcctaattctggaacaaacctccaga tacttctagcttggaatcatttaagtctttt gggagaacgttgcagacaagacatcttag attagttggcggaagagactaaaatttta ctcgtatagattatctatgtttaaga403.g bk

SEQ ID NO:	Name	IscB sequence	ωRNAs
1527-1528	human gut metagenome genome assembly, contig: NODE_8903_length_3490_cov_3.366812, whole genome shotgun sequence GENOME_ACESSION: OLQX01.1 GENOME_ID: 7913 CONTIG_ID: 8902 SOURCE: WGS DATE: 2019-01-27.	ttgagagtattgtttgaacaacgcggaagccgctgatccctgttccggcaaaagcgcgccttctgcttaagagaagaagctattgtgtaagcgacgccttcaccattcagctcagattcgaccggcgagtgcaaacagccgggtgagctgggtgtgatccgggtaaaaatctgctgacatcaatgaaaggtgaactttatgcatccgaatcgaaatgctggaagacatccgatctgctctctgctgctgctgacggaatcccgccgtaacggcaaacgcgctaccgacggcgaggttcgacaaccgcatccgaccaagcgaaggtggcttgaccgtcggtgaaacccaatcaacgcgcatfctgctgcatagaaagcggttctccgactgctgcccacacgaagattaccgtggaacggcgctcttcgacatgcagcggctgaagaatccgacatttcaggaaaagagttaccaagaaggtgaacagctcgcttttggaaacgtccgagatgtgtctttttagagatgggcacgtctgctggcattgtcacggcagatcgaagacccgggtcttaattgttcacactgtagagagcagacgtaccggcagcgaattcggccgacaacctgattacgtttgtgagacgtgccataaggcgcttcaccggcggaatcacgtgaagcccaagcaggagcaatcgctccggcgcaagctttcatgggaattatcgccgggaggtttgaaccgcctaaaggcgctgtaacctgagctgaagtgacagaacactacggctacttaaccaagcagcgcgcgattgcgaacggcattgtcaagtcacatggtgcagatgctttctgattccggcaacctggagccgaaagctcgccgaattcttctccagaagcagactcgccgaacaacccggcagattcacaaactttccatctcaaaagggcgatcgaagcgcaatcaggcaccctttgagccaaaggctccgtttttgataagggtgctgcccaggagaagaaggcttcaatttcggccgtcgtatccgggtctttgatgtacgaagctcgacggcaccgcaatttcagccggcagctgagctgcaagaagctgctgctggaagaagacgaacctatttaacagaaatcgaa	gtcaactacctcgccctaaaggccggggcttgaatagcctctagttagctagcctcaggccgctcatttggcggactacgttggttgagaataataggcaccgggagtgtagatactagctcccgctctgctggccgatggttaaaagctctgagagtagagcggtgctgctgcaaaacccccccaacattggcggaagatcacaaacggctggaagaccgaggagacaaaacttga
1529-1530	LHMISPF_alex1_c86 GENOME_ACESSION: IMG_2049941002_SF_2049941002 GENOME_ID: 27373 CONTIG_ID: 1246 SOURCE: JGI DATE: 2019-01-27.	atgagcaaaatattcggttagacaccaaaaaatctctaaatcctatccatccaggccgagccaggatgctgctcaagaggggcaaacggcgctgctggaacgttaccgctcaccatcctgaaagatgttctacaaaacccccaaatgcacccctgcgattaaaaatagatccaggcgctaaatfctgcttagcactggtgatgatgcaacaggagagaagtagctggcgccgtaactagacatagagcgtttcaaatcgcgatcttaacaaacggccgacgttaagaagaggaaagacgccatcgaaaaactcggatcgtcaaccacgcttgaataccggaaacggccctacagggtgctcccccagctgcaaaagcgagtgagaatatcgtaaactgggttcgtaggttacaaaagctgcttgaattctgataattctcaggaattagcaagtttgatactcaactcatggagaatccagacatcagtggaatcgcatcaacaaggtgagttagctgttacgaagtgcgagaattctgctttcaaatcaatcactgtagtactgtggcgctaaagatactggttggagattgagcatctttcccaagctaaaggtgttcaatagaattagtaatttaggcattgctttaggaattgcaacaaaaaaaggccgtcaagctcagagaaatttttagccaaaaaacctgactgtgtacagcgcaattctgagacaagtaaaacaacagaaagccgacactgctgcccgttaattctaccggttgggcaattatttaacgactcaaaagagactggaattatccgtggaagtgggactggtggcggaacaaagtttaaccgatgctgcaaaaactacataaaactcactggctggatgctcctgtgttggggcatctacacccctcaaaactaaaagtttcaggggtagtgccgctgtcaataaggcgctgttggattggaatcgtcaaatgttcaagttgacaaaatttgctttcctaagcgtggttaaggaaagtaagctaatcaaacgtgagcgacagaaaattactttgggtttcaaacggcgacatggttaagctattgttctactgggaaacaagctgtactacgttggtaagtcaccgtcagaaaatccgggtcatttgattgactgtgctcatgtgcgcttcagagatttctgttgaagcactgcccgcgggttcaccgctttgatggttacagttatgcctcttgtcccaaatgtcagtttttaaacgattag	gctatccggaaccactggacaaacctggacaatcccagataaactctggttagagatagcccggttaagtctaaccacagactacgttaagaagggtcacgacccctgtgagtgctggccagctcctgtgctgctaccagttaaactctttatttcgtaaggagtgctgtagcagacagccctgagcaaaagtcaccccttaacattcaaggcaaaagtcacccgattatcgagggtcaatatgagcaaa
1531-1532	Ga0172362_10014068 GENOME_ACESSION: LakKivs1_kivu2a2_2_SF_3300013133 GENOME_ID: 31130 CONTIG_ID: 14067 SOURCE: JGI DATE: 2019-01-27.	atgcagaagttgaaaaaagagatacacacacactacgaataactctctagtctgtagcaactgtgattctgtgattaaactgtgaattcgaccaaggtgtgcagaatttaaaactttttcaacaactccgaagagaaccaacagctgtggcaagactgaacaggactgtcgagttcctgtttaaactcgtgagacatcctttgatgcctcaaaatcgaagaaagcaagaaagttgttagaacaacaaagctaaagtagtaacaataaaacattttacgatacaactaaattatgctacaggagaacaaaacagaatattatattaggcattgatgcaggatacaaaaagttgggtttctgcaagaaccagcaaaaaggaacttatcttgagaaagttgaattgagaacagataacataaaaactacagaagggcaatgtatcgcagaaaccgaaaaaactgctttgggtataggcaactacgatttaacaatcgtggacgaacaaacggatggctgcaccaagtatacaataagcttgatagtcacttagaattgttgaaaaaattaaacaacttcttctatcacacagattgtagtggaaagttgcaagttttgatatacagaaaatcaaaacccagaaacgcaaggtgagcaatatcaacagggtgaacaactgtgttctggaattgtcgtgaatatgttctcatcgggatgaacattgtgtcagcattgtcatgtaaaaaagatgacccctgtgtcatgtgcatcatatacaacggtaaaaaagaggtgctacagaccgtccagaagaacttttaacagtatgcaagacgtgtcatgatgagcatcataaaggcattgatattatccctaaaaagataatcaagaatttcaaagcctgtttttagtacaaccgttagatggaaaatagtcacgccttaggtgtgagcatacttatgggcatatcacaaaaaacagaataaagcaggtttgcctaaatcacatatcaatgatcctttgtattgtctgggtgactgaccagcaacgttgaaccgattgtttcaaaacagataaggcgtaacaatcgttcaatacagattatcgttaagggttaagccatctatttcagacagaaaatatacaactcaaccaacgataatgtaaatatttgggaaggaatgcaaggtaaaagccatattcaacaaggttaattgattaggtgtgtgattgtaaacattatttaattcgaacattaaaaatgtagagttgttatattgaaaaaggaattattatg	gtcaactaccaccactgaagtaggtggcttgtcccgttaagggtgatgcaagagttgattaaggagcatagaaaaatgacagaagttagaaaaaagagatacacacacctacgaataacttctagtctgtagcaactgtgattcgtgtagcaactgtgattctgtgattaacatggtcaattcgaccaaggtgtgcagaaatttaaaacttttttaacaactccgaagagaaaccaacagctggcgaagactgaacaggacttgcgag

SEQ ID NO:	Name	IscB sequence	ωRNAs
1533-1534	Ga0134404_104661 GENOME_ACESSION: IMG_3300014769_\$F_3300014769 GENOME_ID: 280515 CONTIG_ID: 4660 SOURCE: JGI DATE: 2019-06-13.	atgagcacttgcgttctgttctcagtaagaatggcgaacgcctgatgccactatccgtcttggaaggtacgccattcttctgaagacggaaaagcaaaatcattaagcatcatccatttactattcagttactgtatgacagcgaacgaatttcaaccattgaaatctgtgagacgtcggttacaactacatcggcatcagttgaaagccaattctacagatgtatctgacagtatgatacattgcaagacgagaagcgtgtcagatgcttgcctgataatgcgccgaactcgcagaacacaggctacgttaccgccaaaagcgttggacaaccgcaaacgtgacaaagcctggctgcaccctctctnnnnnnnnnnngcgaagttgtgcttattacgcatgcaaccgttgaagttggttcttcgacacaatgcttgaaaagcaatcgagggaaggtaaagctacaccagaagcgcagattatcaaaaaggccctcgtacaatttagcaaccttgcgggaagcagatttctaccgtgataactacacctgccaaagtttggggcgcgaagccagtgaaaggtgcaatttgcacgtgcatcacatgttttattggaaggtgcgccatgacaatgtctcagcgagcttat aacagtatgcgagaagtgccatcacacagcgaacatcaaaaaggcggcaagctctatg gatcgggtgaagataaaagtttggcaacccttctgtgtgcagcatttgaacactgtacgctggcaaatcgttaataagcaactttacgctgttcttggaaag	gtcaacaacccccgccttaaccagtcgcgtggttatagatggggcttgcagggcaacccgt aagccccgggttattagcctaaagtcgtctc cagcggcaggaactacgttgtgtactaata atataggcaccttactatgcctccacaagtg gtaagcactgcggacggctcgttaaacatctc taagggtaggagaagtgcaacgtcatgtc gaaagggctaaaacggataacaacattggc gatgtggaccacaggcgcgaagccctgac tttatcgatttacaactattatgaaaaggatg aacttgcagcacttgc
1535-1536	Ga0123338_10030355 GENOME_ACESSION: grotaG_2_\$F_3300009686 GENOME_ID: 37706 CONTIG_ID: 30354 SOURCE: JGI DATE: 2019-01-27.	gtgatcctcaaaaggacaatgaacatagaacgatgttaaaaaactgatgcgctaaggcgtggcaataggtcttataaaaggtatcacaaaaggtatagaatttgcaggttgcacatcgcgcacgtctaaaagaaaagggaaggattgcccccaagcattaaacagaaaaagattcgg tcttagggcaatcaggtgagtaaacggattagaatttaagaatttggtagaagatg tcatgatagatattcgtgcatcaacagcggatgtaaacgtattcgtggcaatatacaaaaa tccaacagacttgatgagaacttacgaaaagcgggttactccttagggataaactcaagtcgc cagagtgccggaagacaaacaccgttttagagggttcatatataaaggctaaaagatgta atggctcaaacaccttggcaatttaacgcaattatgttctagtgtcatcaagaacacgca ggcagagaaaaagagtttgaggaaaaactttagctgcatcaatggcaaaaatgtaagg cttgacgatgcacaacatgttatgcaggggaaaaagctatctcagacatcaattatcaatgct ggggcaatttagcattacaatgtggcgactactgcaacaagcgaatagattggagt	tagaaaaatgtataatgtggacacatag gctattgcagctagactgttgagcaggtct caagctttagtactgtacagtgcaagatg tttgtgcagatgaactacgttggatagtaa ggttaagacacacaccttggatgtaactgca gtctaaagctctgtgagtgcccaacgaagaa ctgcgctaatgccatgcgtagataacaggg aaacacaagtcctctatctgacattgacaag acgaaaaactactccgaaaagatggcttaca gaaatgatactagagaagaatattcattcgt catagtgcaggaaggaaaaagtttaactcct accccaataaacaacgcaggttttgcacag aaaaaggcaaggctaaattagtgacgaatatt ccaatggcaattcagttgtaagatccattga aaaagaggatgttgacaagtcagctatttgt gtggaattgacgatgggtcaaacacgttg
1537-1538	NODE_10969_length_3841_cov_6.272065 GENOME_ACESSION: mgm4743570.3 GENOME_ID: 238071 CONTIG_ID: 2352 SOURCE: MG-RAST DATE: 2019-01-27.	atgacacgaagacaaaatcgcaggaaacagaagaatcgtaaaactcgtacagaa aggaagatggttaaatcgtgccaaactcttataaaatggcaggttagtctacgatgcg gagtaaacctcacagccatataaaagaatagatataaataatctattcttcaattacaga aatggtatttgagcagggtcagtttgacactcaccttataaaaaatccaagccttgcatactc caaagttaaacattggggctatcaaaaagggtgtaactcggatttgagaatactaaggca atagttcttaacgagataactatctgtcaatctttaaaggcaacataaagacagtaa gttagaagtacatcatattgtattcgcagtcaggcggttctgatgaagaaggcaatctaa ttactttatgcatacttgcataaaaaacttgcatatgggaaaaaattaccccaattgagtg gcaaaaatgaagggaacccftaaatattgtacgcagatgaactctattcgtgaagcaacttttag actatacca	gtgaactaccaccacactaaaggcagtgagg ctctgttaaatgttccaccagcaatagttgtt agaaatagcaactacgatatttagtcaagat acccttggttgacgcaacagaccatcgcctc atctgataactattaaagttaggtcagatgaaga acagcccttgtgatgtgcttataaaagccttta tatattgtcaggttgagtcggaacaactgc atgtgaacagtgtagcaagtagcgcacaac cattgcataagtaattgcattatgtgtaattg tccagggaaggagaccctgttattgg
1539-1540	Ga0114919_10008950 GENOME_ACESSION: BlataG_2_\$F_3300009529 GENOME_ID: 23889 CONTIG_ID: 8949 SOURCE: JGI DATE: 2019-01-27.	atgcaaatgcaacaacagaagttatgagaagaatacatcacaccattgaatgcttcac aagttcaatgcaactgtgatttggcattaaacaagaggaaaactcttagtgtgcaaatftaa aaactcctaatacaatcccgaagtgaatctaccgcagaatacagcgggtcttaagaca tttatgtattcattaaatcaattgacggtaaaagctttaaagcgtgtaaacagcgaagccca aacatttattaaaggaaggaatgctaagtaattaaacgtgccccatttaccattcaattaa cggttgatgtaaaaataaaatacaagacattacgttaggaattgattccggatatacaaaata ttggatttcagcaggttgcgaattgaaagaattatattctggaacgttgaacttgataacaa aacaaaaagaagactcgacacaaaaagaatgtatagaagaataaaagaataagttgtg ggtatagacaatccagatttttgaatcgttttatacctgaaaggatgttaccaccatctataga gcgtagatataaaactcatttaatttaattgaacgtttaaanaaacattttacattataaag taataattgaagtagtgaagttcgataattcaaaaacttgagaatccagatattaaaggcgfta gttatcaacaaggagaatgtaccaatcgaatagaaattggatatttcttcaagagaa catggagtatgcaatttgaaaaaacaagggtgggggaaactggcgattacatcataatttt gggcaaaagaaaaagatagacctaaagattgggcttattacatgaaaaatgtcatcaactc ttacataaaaaaacaagaaacattttaaagaaacaaaaacttaaaagttataaagactc aacatttgaacataatcgtaaaaaattcaacaagatataaagttgtaattaaacttatgg ttattatctttcaaaatagatgtaatttaatttagacaaaagtcattgaattgacatttgtt	gtcaactaccactgagctaaagactcagtggt cttgcaatgtgaattgcaacgcaaggttgat tagggagttgaaaaatgcaaatgcaacaac agaaagttattgagaagaataacacacacc attgaatgcttcacaagttcaatgcaactgtg atttggcattaaacaagaggaaaactcttagt gtgtcaatttataaactctcaataacaatcc cgaagtgaaactcaccgagaatacagcggg tcttaagacattt
1541-1542	Bioreactor metagenome contig_107124, whole genome shotgun sequence GENOME_ACESSION: AMWB02.1 GENOME_ID: 96 CONTIG_ID: 3559	atgcaaaagatttttagtttaagccaagataaaaagccattaatgcttgcctacgtcccgga gctagaaaacttttaataaatgcaaggcaaaagtatatcgttattatccattttgtattatatt gcaacatcgtatagatggcgatatacaaccaatcatgtgaaattcgatccgggttagtaag acaacaggagttgccctgttggcttattcccaaaaggggaatctgtttatggggagctaa tctaacacatcgtggctacgcaatcacaaaactctgttaccagaaaagcaaatctgctgtg gccgcagaatcgaacacatagatctgcccgccacgtttttgaatcgaatattcttctaaa ggttggcttccgcatcgttactgtctgaatttttaattatctatcgtgtgcaaaaagctca agcaatttgcaccaatccaacaaatcgaatagagagcgtcaatttgatacccaaaaact	ctctacaagcccacctcaaaaaggcgaagtt ttttagagaatctggcgaaattcaggttgac cagcctaagacctcaggtaacgaagagg ctactgttgcgacgaatccaagttaattggtt aataggcagctcaggatgcttactgtgtac gctaagtgcttactcagctcctgacctctgc ggcaatccattaaacaggtgtaattgtttta ggccagtggtgaattgcaaaaaccgtgtttca

SEQ ID NO:	Name	IscB sequence	ωRNAs
	SOURCE: WGS DATE: 2019-01-27.	gcagaatgctgaaatctggtatagatgacacaaaggctactctgcaaggatacgaatt cgcgagatttctagaaaaatcgggcatcagtggtatattggtgcaagctgatgtacc attagaaatagaacatgtagtatgtcgtcataatgtggcagtgatcggaattagcaactag ttatcatgcccgtatgtgaacgaaaagaaggctactcaagacatcagaatattttaaaaa ccaagcccgtctattgcaaaagatacaaaagccaattacaagtcgcactcaagacaca gcccagtaaatgccactgatafgcaatttagacgcactgcaacaattcaactaccaat tcaatatggcactggtggccaaactaaataataatgcacgacacaaggctatgcaaga tcattggattgacgcccgtgtgtggcgaagtggttttaagtaaaaattccccataaat gcacataataatacaaggcagtaggacgcggtactgccagatgtgttaacagataag tacggtttcccggaacagggccgaagcaaaagaaacggatatacggcttcaactgg ggatatagtatatgctaaagtacaaagggcaaaaagccggggaatatatagccgag ttgctattcgtagcaaaagggtatttgcatactggacacacactgcttcataagtagcaattc atacgttatttattctcccaacgaagcgtatgggtatcaatatgattag	agtaacattggcaaggagagcttaacctgc gcaagcagagtggtacgtaagacctctct gaaagggattttatgcaaaaga
1543-1544	Ga0070707_100019090 GENOME_ACESSION: KBSK5StandDraft_9_\$F _3300005468 GENOME_ID: 30508 CONTIG_ID: 19089 SOURCE: JGI DATE: 2019-01-27.	atgatgtttgctatgttgcgaatcaggacagcacaccgatgcttgcacgccagcga aagcccgaagctctgtaggcaggcaagggcgaaggtgatgaatcgtgtcccttcacg attcaactcactggtgtgtgaggacacgttcgaagatcacactggcatcgataag ggcagttctacacaggacttctgtgttggcaataggaaagcgtctgttcagcagagat ccaccatcgccgagatgtaaaagacaaatggatgaccgcccgtgacggcgaagagt cgtcgaggcgcaagtggtaccgcccgtcgcgatcaagaatcgcgcatcaagcaggc ggagtgggcgattgctccctcgatcaaaaagatgtcgaggaggtcattcggtatgtg aagcagatcccgtgcccatttctcgtcgtgcatcgaggacgttcaggtagacattgctc gcttgatgacccacgcttcaaggcagtcgatacgaatcccacgcgactcgtatgag aatctgcgcatggcctgtttgatgcgcagggctacgctgacagtcggtgcaagca aaagggttcggtagaagctcatctgtgtgtacaggggaacaggcggaaaagatacgc tgacgaacttgctcaccctgtgccaagcatgtcataaaaaggttcaccaggggaaaatcc ggcttgaggagtgaggagtcagtgacacctggaccagattgcgcagcgcacgatgca gggcaaaagctactgtatgcgacgctggcgccacatacccccttgagcacgctcttgg ataccagacggccaccctcgccaaaccagagggtggccgaaaacgcacctcgccga tgccctc	gaaaatagttcctaaaaaggagagcgtttga gagcgtggtactcatggagctaccagcgg ggcagcaatggcagcagttcttgagagat cacacctccgatgttcgtccagtcggag ctctgtgggcagatattaaaggtaagcgaaa ggcgaaggtgatctgctgggaaaacctcct ggaaacgtcgccaggagcaatttactcgc aggagggccctaaagccatgatgttt
1545-1546	Ga0187899_10002485 GENOME_ACESSION: Goafecreptomycin_6_\$F _3300019376 GENOME_ID: 26928 CONTIG_ID: 2484 SOURCE: JGI DATE: 2019-01-27.	atgacgaccttctgcatctcggcagagggcgaaaagctgatccgacgacaacatcag gaagataagaagctcctccgctccggaagcgaagagctgtgaagaacgcgcgttc accgtacagcttctgtacgagagcggaagcgtgtgcagcctatcgaatttacagaggat accggctacagttatcggagcctctcctaagtcggagaacacgaatacgtaaagcggc gagtatagcttcttaaaacgagaagcagatcaccgatcaccgcccgcgaggtga gacgtcctcgccgaaacagaaaagcgtatcgaaagcgccgttgataacaggagaag acctgaaaagctgctcgccgctcgtcagaaacaggcgaccgccacgttgacatct tcaggatgtattatgcggtatgcccaataatagcgtctcgtcgagatgggccagttcga tccggcggtccttgacgctatcgagcaggggaagcctgtacctgaggagctgattacc agtatggccctcgtacggctatgacacactgagagaagcagtccttcacaagagacggc tacagatgccttctgctgtaagtcggtatagacgacggagtcatacttcgcatacatc acgtaggcttccgcacggcgacagatcaacaggcttagcaatctcgcatcagtatgc gagaagtggccacagtcggaagaaccataagccggcggaagcgtctgggatctgaag ccgcctaaaggtacggcatccgacgctatgaacaccgttaaagtggcatattctcga gagataaaaggcctcgttatagagaccatcacctacgggtgacgaagcgac gagacttgatctcaatatcggaagctcagcgaacgatgcctattgtatcggaatatga gaccgaacaccgcacaaggacgagtgctgttgagaagcgccgcccgaacaccgc atacttgagaagtctacgatgcgaagtgtgtgtatgtagcagacaagcggaagag gcggcggagctcggatgtaaccgcacgtcgagatccataccggaacaaccctcaga acgagcgacacttccgggggaagaagatccaaggcagcgagatccgcatgca gagataattgtacggcaggcgatattgtcatcttgggtggagaagcgcatggtcaag ggcact	atcaaccaccaccgcttaggctaaccgctt gaagcgggggctgtgaaacacaaagta gttgattagcctaagtgtctcgagcactacgt tatccgcaatgctcgacacaaataggtac gtcagatactccacaagttctgacctctacg gatatgctgtaaacatcggtgagggtaggc gaagtgcttatatcatcaaacggggaataa cattggcagtgatggacaccgcccacgggg cgagatccgcatacgcctctacgaaa aacggtagtcgggaagccgtaaggcaaatg acgttagagcaaaagtaagaagagaaaagg aggagcgtatgacg
1547-1548	Ga0394882_0012572 GENOME_ACESSION: WSA_S4_MG_FD_\$F_3 300037422 GENOME_ID: 312718 CONTIG_ID: 12571 SOURCE: JGI DATE: 2020-04-07.	atgattaaaaaaccacatgtgttagtttagatgctaacaaaaccattaattccgtgta aacatcgctagctaaaaaatttaaaagctcaaaaagcgtccgtttaccgaagatttccctt ttacgcttattctgaaaaaagaatgtactgtaaatcagcaaaaactagaatctcgtcgtcagt atcgtcgttaagacgaagtagattgaggtatagaaaacgaagatttgacaatcgccaaa gagctaaagattggttagctcctagtttagacataggggttttaactacgatgaccattgtga gacgactaattaaatcgctcatattagttcgtatgtgagtggtgtaaaagtgtgataccca aaagctagaaaaaccagaattttaggaattgaatatcaacaaggagaattattaggtat gaaaattagagaatatctattagaaaatgggagagaaaatgtactattgtttcctaaaagg tattcctttgcaaatagacatattgttccaaagtcgagaacgggttcaaacagacttagtaa ttatgctgactgttcgaaaaatgcaatcgaaagaaaggagtaaatcgtacgagcaatttt tagccaaaaccctagcttactaagtctatttttagtcaagctaaaaaacattacacgatg ctcgtcgattaatgccactagatggaattgtttaactctgaaagaaactggtttaacaa tttaactggagcgggagcaaaaactaaatttaacgtttcctcattgttagaaaaaac attttatgatgctgctgtgttggaaaaatggataaattagaatttttggcaaacacctttg	ctaaaaaacctatacaaacctaattaaggttg gtattttgttgaggttcagactaagttatca cgtaaagtaactacgatagcaagtcagat acctacaaatgaacggcaggtttgtagctctat cgtagaaaattaagaaagggtcaaaaggt ttataccaagcgtgtttctagcttaaaagct tgcatatcattgtcaggtcaacttaacctat taactaggaaactaatcaaacattgattaaaa aaccacaaat

SEQ ID NO:	Name	IscB sequence	ωRNAs
		attattatttgcacaaaggacaagggggaagacagaagcagcattataataaacggatat cccattagatacaacatctttggacaaaagcaaaagcattcacgtaataaataagac gagtcacataaaaggatgatagattatgcacataaa	
1549-1550	human gut metagenome genome assembly, contig: NODE_12009_length_3547_cov_2.304410, whole genome shotgun sequence GENOME_ACESSION: OURT01.1 GENOME_ID: 11174 CONTIG_ID: 12008 SOURCE: WGS DATE: 2019-01-27.	atgtatgtattcgtcatagggctggtatggaacaggcttatgacctgtataaccaaggaaag cccggaaattaatagaagctcacaagctgaaatctataagaagcagccattacaataa gattgtgtataagacaggctgtgtctacacacccgtaacaataggtatagatacagggttc ccagcatatcggaaattgcggtatcaatcaggaacaaagcttctaccaggcagagatag aattgcgcagcactatggacaagcgttcgtcatggaaactagcaggtaccgcaga agcagaagataccgcaaaaccagggtacagaagctctaaattcagattcatacaaaacgg acatactctgaacactgtttaaacgttaagactacaggtgttatgacctactgggttaaca tactaactcaatgagtacaacagaccggacggatggctgaccatccatgcagtcgaa agcagatcatcatatccgatggattaacagatttttagatgtactccaccagatagaaact tcgggtggaaattgccagattcgtatagcaaggatgaaaaatccggaagtgcagatgtga attataccagtatgggtccagatgatgtacgaaatttaaggcttatgtatttgacagaga ccattacaatgtgtgtctgtaagaggaaattggggagtaaacagccggatggacattct ttaaaaggaaatgatgcatacataaattccgggtctaaaggagctacggacaatcctgatc agctgtttacagctcgcgaatactgccatactcctcaggcgcataaaaaaggcggagtact ctgggactgc	ttaaataatcgagcatggttaatgattgtcg cccgttaaggcagagtagctagccaaagtt ccaagagaactacatcagaagtacaatccc agttccagacagatccctaatctgagcgtcg ggcgggcaacatagatccagagagaaag aatgatgtggccgggtttcaagtactctg ataattgcgaagggttaattgctgactttggcc agcatagac
1551-1552	human gut metagenome genome assembly, contig: NODE_3708_length_6101_cov_2.219618, whole genome shotgun sequence GENOME_ACESSION: OQNV01.1 GENOME_ID: 9976 CONTIG_ID: 3707 SOURCE: WGS DATE: 2019-01-27.	gtggggacgacgacctgggtctcaagcgcaccggacacagcgaggggcgaacta ccccgtaagggtcccggggctttcagtcgccgggaagaggaaataatgaaggtttt tgtcaattgcatgactgggaaagggctctgcccaacttcccgtctaaagctaaaaactg ctgaaaacaggggaagcacatgtttacagcggtgtgcccgttcacgatccggcttgattata agaccggaggtccgcaccggcgtgaaactggggatcgatacaggggagcagcata tcgggtacagcagatgtaacgatgccagctcttctgtaaatcagaaattgaactataaaat ccatggaaaagaagaaagggctggaaaacgcataatataccgcaggtcccgcaggtga cagggaagaccgggtacagcgcacgaaatccgtttccggacaaagcggattatgtgga aggcgggtacaggaacgcgaagggcgggaagatgggtctccggcattgggaagaaatg cccaaacacggcatgacgaaccggcatgaaggtggctcccggcgtctatcgagtcga aggtaagcaccatcatgactggtacacccgggtatgtgatgtctccgtcaggggacca tgctcggcatagaagtagccgggttcgataattgccagatgaagaccgtctatccgcg gagaactttaccagttcggggcgatgtatggccgggaaatagcgaaggcgtatgtactg gcaaaagtgtgattatcctgcccgtctgtgaaagaaagtttgacaggacaggaagccg cgtatgcaccatgtcaccatgcgggaagacgcagccactgacaaccggatgagtagc accagtatgtgtctcgtccatfcccggggaggagcatctccagggggtgtgtcgataa gctggcaaggaatgcagacgcagggaataccgtgaacctacgttcagtaacatcttc gccggcggtgtgtgagacttaccgggaagcggaaatcacataggggaatatacacaatg ctgaccgggaagatgtggggctggaaaaaacccacgcgaatgatcgggtctccattgcc atgcatggggcgaaaaaagataagactgcgaagatacgggtctattaccgccaggttcac aggaagaacgggtccctcataaagcaaccccgcaaaagggaagaaagagccgaac cataccgcaaaagggaacgtgaagaacgtcccttatgtaggcaagttccacatcaatgat aaagtgaaggtccagacggaaatcctggattcataaccggtttgacagggaagtgcggc ctatatcagcggattcatcggggaatacatctatccggaagggaaggattacaacaacat acgtttccagcctgcgggtacgtccatcattgtgggaattggatcaccagctccgtatcag gtcgcgctgtatctcacgactaa	gtcaataactcatgactgaagtcacgagctt ggaaacagcttcaaggcttctctgttagccgt acatgaacggcaggttgactagcctgtctctt ttccctgcaccgaggggggaaggcgagtc ctgatgcgttaataccgccgacagacagaaa aagccctgcaaaaggcggtcccatctgca gcgcggttcaggaacggatgtcccttga gtggggacgacgacctgggtctcaagcgc caccggacacagcgcgaaggcggaactac cccgttaagggttccggggctttcagtcgcc cgggaaggagggaataatgaaggttttgt cattg
1553-1554	Bacillus sp. FJAT-29814 scaffold2, whole genome shotgun sequence GENOME_ACESSION: GCA_001510715.1_AS M151071v1_genomic GENOME_ID: 96811 CONTIG_ID: 11 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atggctcttgtgttaaatgcgaataaacaacgggttaaaccttgccatctgcaaaaggcaag aagggtactcaagaaagggaagccgtagttcataagaagtttccgtttactattcgattaa aggataggaagatatttcttgaatcaacaacatcatcgaatcaaaatggatgttgggtcc aaaataacaggaatcagcatcatgaagaagaatgaagtggtgttcttagcggaaactccac cacaacacagacatcaacaagaatgaagccgtctgtcttatagaaaatcaagacgg aatcgtgaagaccggtaccgtcaagctagatttttgaaccgttaaacgggtcgaagggtggt tgcccccttccctacagtcgaggtgggacaatgttaattcatggactaatcggttgaaaaa actaatctttgacagatacagcttagaattagtcataaattgatactcaagatgatgaa ccctgaaatcagcggcgttgatfatcaaaaaggaccttacaagatacgaagtcggag aatacttttagaaggtcgggttgcaatgtgctattgctggaaccaaagatgctccattgg aaatgaaacagctttatccgaaatcaaggggaggttccgatcgagtcagcaactgactct tgctgtcgagagtttaacgaggaagaaatgataccttaacctggatgaatgggcaatccaa ttaatcaagaaaaaggataagcggagcaaacgaattttatcagtttgacctatcaaaaa acagttacagaaacctttgaaagatactgcagttgtcaattcaaccgttggaactttatca aatgcttttgaccactgatttaagggtggagtgggaacggggcgaagaacaaaatgca acggatacaacatgaatttccaaaggaactactacgatgcggtttgatcggagagtcata ctcttccaccccaatacatttaaaactaactatgtgttacaatacaaaagcgaagcaga ggctcaaggatcgcaggggaacagataaatacgggtttctatccgtcagttaccgagag tcaaatgatccatggctttatgtcgggagatatgtcaaaagcattgtgcagagaggttaa atatcgggaactgtgtcgggcaaatgtatgcgttccagtggttatgtggatataaag atatgaccggaaagcgaatagcacaagggttcaggttaaaatgctgtcagctgtacagc	gtcaactaccacgccccaaaaaaggcggg ggctgtgaaagtgaagcccgagttgtccag ttcagtttcttttagaaactacgttggttgtg catcataccaaggatacttccctagtccttg gctctatggaaatctcattaaagcaagcaag ggtgttagcgagcgggtgtggattcacgac aagcctgtcaacccctaacgaagggaacag ccttacgggtgtgttaccgtgcaaacggag aaagattcacaatctttaaaggagccgag tcggatg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gttttgatggatattgctattttataacaacgaaaggagtcggcaattccgcacatgtgtga	
1555-1556	k141_28014 flag=0 multi=166.3446 len=3617 GENOME_ACESSION: mgm4773634.3 GENOME_ID: 238274 CONTIG_ID: 349 SOURCE: MG-RAST DATE: 2019-01-27.	atgggtgctgtattatcaagtagccggtatccgtctgatgccgacatccaactacaaagcacg gaagcttcttgcgaagcgccgtgcagtgcaagagaataacacccaatcttcacgattcgc gcttttagatcggagtaacggagaatacagccaattgaatcagattctgacacaggagct atccatgtaggatctcgaatcaaatcacaatacgaatgtgagcgaacagagagat ctcctctctgatgaaacagagcggcatgatatcgaggaaataaccgcagacagcgaag gagccgggttgcgtaccggaagcccggtttgataaccggagagcctcgaataaagaa gggttgctccctcaggcatccagaaccgtatggaccagcaagtcagattttgagaat acgctgaaacgttccctgttacacaggcaaccttcgagatgggacccctcgaacccagc ttttacaggcggcgaatccgagaacccggagccgaaaggacagattatcagcaagg cccaagataccggacagaacccctgcgacaggcagatttttaccgtgacggatagctg ctgcttctcggagcgtggtatcaaaagaccacgaaagctccatccacatctggggttc cggaatggagatcatagacacggatgagcaatctggcaacagtttgcgaaatgccca tacgccgaagaaccataaactgggtgcatccttatgatgcgaaacccaagctgaaacc atttaagggggcggcgtttatgacagcggtagcgtgaagatgtgggacatgctaagaa catccagaaccagatagattccatctacatcgaacccagacaaactgacccgc cagaagctgcatctggagaatccacgcaaacgatgctatgctatcgggaactccat ccgaagcaccggacacaaacgggtttatctgcagaacgccgaggaataaccggtgtct ggagaagttttacgatgcaaaatcatagacagccgtaccggcaaaaggctcgggtgac aggaactgttttccggaagtcgaagcgggaacccggaccctcggagagaacctacat cggtaccggggaaggaagatctctcagggagcggggaatccggaccagacattatct aatccagccggaacattctgaaattgatagaagcggatctgtgcaaaagggttcat tgtaatggaacccgtgtaatctcgagaataaagaatcggatcagttaaaaaagtaaaa tttaataataacgggtgggtggttccctgcgagaagggaacgggagagtttatcag ggggcgactga	gtcaactactcgggtcattgaataaccgagcat gcaaaagcaagatctgtcgttcgattatc ttgggtgcctgtttcgaagcgggtgattag ccttagctgtccggtagggaaaccgacga ggcggctacgttgcgggaagatacaggc acctggggattctccacgtccccagctct gcggccgtgtattaaacatccctgagagtaa gggaagtgtgtacgttactataacccggc caacattggcgtggaacaccccgaaacc agggttttttgcgtggaagtaggataccc gttcagattcatcagtaacgggaaggcgtaa ggcatcaatggaaggaggcaccggaatg
1557-1558	human gut metagenome genome assembly, contig: NODE_6_length_339981 _cov_10.284268, whole genome shotgun sequence GENOME_ACESSION: UEYJ01.1 GENOME_ID: 15178 CONTIG_ID: 5 SOURCE: WGS DATE: 2019-01-27.	atgacaggcagctatataccaattttaccgtcttatttctgtcgaatagttgacgtacccac cgaatgacaggcagctatatacaggctacgggaggaagaaatggcagtagtgtttt aagtgcaaacgggtaacggctgatgccaacgaaatattcggcaagtgccatctgc tgaaggaggggccgctgctgattgcaaaagcgaacccgttccatccagctgacctac gagaccgtacattatctacgtccatcgaatgtgctgcggacccggaatgagcatac ggcgtcagcattagacaaggcgaagaggtgctgcgagcaaatatgatttgcacg aacgagcgtcccgccatgacgactgccgagtcaccggcggacgagcgaagaaccac ctgcggtatcgggtgcacggttcataaccgtgatccatcagaagccgggttggttgc ctccgtccctggataataaggtggagcggcaccgtgataattatctccgctatctgtctgc atgcctgtcagggatgctttatcaaggccgcaacatagacacagacacgtccttgcggcgc tggaggcaggggagcctgtcccgaggcgaaggactatcagcatggccgcagtagc gccacgatacgtcggggaagcggctttgaacgagaccatactatctgtgtattgca agagaggcctgaaagacggcgtatcttcatgtccatcagcctactactggaagggtc tgcatgggaacagcataggggaacttgcacgtgctgcgagaagtgaacacacgtc caatcacaaggaggcggaaagcgttggggcttcgataagccttgcggaatataccg gcgaagcgtttatgaatagcgtgctggttcttaccagcgtgcgatgctcacttcca gggctgctggacatacacatgacctagcgggttatctccaagcgtgttcgaccgacct cgccctccgtattctcgtcgtacgagctcctacgtgagggcagctgctccgaag ccagatgcgaacagaggtcttcagaataacagcggaaacaacagagtcctatccaaa ttctatgacgcaaatactacgacacacgggaaaaggaggttaccgctctggaacga gctgtcctccggttaggacgaatcgaatcataacctgacggtgaaaatctgcgccgggt ccgtggtcgaagaagtcgaaggccggacctcgaccagaaagcagagatagccttg cagccaggagatgtcgtgtctttggcaatcgcaaatcgtttccaaggcgtgctcagct atggcaggggaattaaagctactcagcgtggaagccgcttgggtcagcaggaagaaa atccagctgtcagtcacaaggcggcgtgggtgcggctccccaggcggcagccgggg caaaagaataa	gtgcgtaccatcaaacccagtgatctttggc gatgtggaccacagtaacgccgtgcaagc ggtaagtgtcgtatttaccaaattttgttttct tgaagcctgtatcgaagactaccaactaca ttgagttgatacccccgcaagtagctatacc ctatggcgaatcagaatcaactctcgtatg ttgacgtacccaccgaatgacaggcagct ataccaaattttaccgtcttatttctgtcgaat agttgacgtacccaccgaatgacaggcagc ctataccaggctacggggaagaaagaaa
1559-1560	JGI1357J11328_1001262 3 GENOME_ACESSION: Subgrocontami54m_3_\$F _3300000574 GENOME_ID: 35777 CONTIG_ID: 12622 SOURCE: JGI DATE: 2019-01-27.	atgttgaccagggttaagcaggattaggcaggcaaccgcttgagttctgtcctgttgcg gatagtgccaaagaccgacggcagatgttctccagccctgcccctcgaactcacgc cagcagacacgccgggtcaggcagcaaacggggcgcgagagcagaccggtccgca acattcccaggagagccttcgagacggatgccaatctcagcgcgtcagggccaagg gtacgaccgaagccgtctcgtcgcgaggggttaacccccggtttctgttcgaccgg catggacgcccactatgtcgtgctgattctcggggcaaaaggaattgctcgataagggt caggccgttctcagatcagcgcctttcaggttctcgttgcgtggtgggtggcgc gggtccaaaccggtccgccaagatcagccggctcgaagatgaccggcatcgtggtt ttcgcgatgttgacacgtttgatgcaacgagcaggtgaagtcacccgcgaagccgtggtg ctgcacctgcttgaattagcagcgcggcgagcagattcgaagcgccttcagcagcg tcgaggttccgtcgcgtcggcggcgaatctccgatctcgcgccccgcgggtcg acaaccgaacaagaccgatgctgctgctcccgagcctacaacaccgggtcgtcac aacaatgacgtgggtctcgcgctcgagacgagcgtgtccatttccgtctctcgttggag	ttggcaatcaagacaagcacaacttttgaagt ccaagaccgccgaaggtttgacgtcatctc gaacgcaatcatatggcagcacattagcgggt ggtcatcccgaatgagatgcctcgtat gatacttccagggaatcaagaccccccaatt catccaaaaggacccttctgatcctgatagc cgacgtattattgttgaacaaacttcaatt gggaacactccaggtgctatgtactgtgtg gttttagggatagacaagagtcggttaaac gggaaagtgcgactgctcatctggttgaaa aaccgaagaagggaaca

310

SEQ ID NO:	Name	IscB sequence	ωRNAs
		<p>tgtgggtgagtcgggagtcgatgtctgtatctacccggccataagccctcagtatctca</p> <p>tctatgggacgagggagcagcaggtagtcgaaccgaccgtatggattccggcga</p> <p>gtaaggcagggagaataaaaagggtctttgtttccagaccggggatctgtcaactgg</p> <p>ttcaacccaacgggaagtatgccgggggtacacaaagggaaggctggcgggtatcagagt</p> <p>gaccggacaatttgatatcaaggcttcaattggaaaaataacagcg</p>	
1567-1568	Ga0123353_10014374 GENOME_ACESSION: TergutLab28metaG_6_\$F_3300010167 GENOME_ID: 36036 CONTIG_ID: 14373 SOURCE: JGI DATE: 2019-01-27.	<p>atggctctatgtcgtatcacaggaaggaacccctttaatgccgacgtatcggcacaataagg</p> <p>tatctaagttgcttgatgagagaaaagctgtgtgataagaacaacccgttcacaattcaa</p> <p>ctgctttatcaaacggaaaccgaagtaactcaaccaatcacattaggcatcgacgagc</p> <p>tacaaaaggctgggtacagcgcaataactgaccaaggaggaactcatcgcgcggaact</p> <p>cgaactattacagggcatgaagaagcgtatagcggagcgtgcaatgtatcgcggtcaga</p> <p>gacgcgcgagcgtccgccaccgtgaaccgctgtttgataacagggcaagagacgagg</p> <p>gctggctggcaccgagtatacagcataaacttgacagccatatacggctcatcgaaaag</p> <p>gtacttgggctaagccagtgacaagtatcataatgaagtggcaagctttgataccaccg</p> <p>tattgaagaatcccggtattcgcggcaggtatcagcaagcgcaactaatggacttttg</p> <p>gaacacccgtgaatatgtacttcacagggaacgacaaatgtcagaaccccgattgtac</p> <p>gaacaaatcagagcagccaacacttgaacacatcatatataccgcaggacaatg</p> <p>gctctgacgcgcatgggaacctgtgaacgctctgtagtaagtgtacacacctcaagcaca</p> <p>tgatacatggctcaaggattgggaagcctaatacaaaagggtttaaggctgaacgttcatgt</p> <p>caatggtagcgtggaactgatagaactgctctcggagtaaatgtgacacatactacg</p> <p>gttacgagacgaagggttaacgggttagcattaggctctgtataagacacatgccaacgacg</p> <p>cgtttgttatagccggtgggaacgacacagggccagatgaagcagattaattttgcacaaa</p> <p>gaagacgcaataaccgctgcttggagaatgtgacgacgctaataacagataatcgta</p> <p>ccggtaaacctgaaaaggcgcaatctgtaaacggcagacgtacacgtacacaaaac</p> <p>aaaaccaaaccggagacaatctaaagggttccgtgtgtgagaagttaagcaaaaggccg</p> <p>ccgctcgataagaaaacagcgttatccgtatcagccgggtgataaagttcttttgagggtg</p> <p>agagatatagcgtttacaggtgctcactgtaagggaagaaagcgtataacttaagagaccg</p> <p>gtaaatcggttagcaggaagcaaaactaacacatctatcatagcgaagggtatgtagctg</p> <p>tgtaa</p>	<p>gtcaactaccaagtctaagagcgttgggct</p> <p>tgtagtaaaacgaatacagagcccggttgaa</p> <p>tagcctaagaccttcgaggtctacgttaccca</p> <p>agaataattcaataggcaccataggggtacttc</p> <p>acaagctctatgcaactgcgggtacatgggtta</p> <p>aatatccctgtgggttaggggaagtgtgtgt</p> <p>acagtcacaacttaagcgagaacccgtaac</p> <p>gggacctatgctgtgtgcttgaaccttggga</p> <p>taccttggcaatgtgaaccgatacgggtcag</p> <p>atagtaccagcttaccgtaaaactacaaaa</p> <p>aatgggaaggaggtttatcgcaatg</p>
1569-1570	Ga0209124_10001910 GENOME_ACESSION: BarGra022ASPAdes_\$F_3300025852 GENOME_ID: 23741 CONTIG_ID: 1909 SOURCE: JGI DATE: 2019-01-27.	<p>atgacgtatgtattgatcaaacggcaacccccatgccaaaccgagcgtggcgggaa</p> <p>agtcggaattctgttaaaacaaaaggcaaaagtggtgtgtcgtacccgtttacgatc</p> <p>caactgctttatgatagaaggcgtttacgcaacccgttaacgctgggtattgacagcggat</p> <p>atgccgaagtgggctattcggcgtgtcactgaaacaaaagaactcatgggtggcactttaa</p> <p>aaactgtctacggccagaaaaaccgtctcaaaagacgaaggatatacgtccgtacaag</p> <p>gcgtaacccgttgaggtaccgacgtacgggttttcaaacctgtgcaagaagaaggga</p> <p>ggctggcaccctcattcaacacaagcttgacagccagttggcgtttattgagcgcatcaa</p> <p>aaagcttctgccatcacaaaacgattcattgaggtgcctcctttgatacacagaaaat</p> <p>taaaaccagatattgaaggccttgatfatcaggaaggcgcgcaaaaagacttctggaatg</p> <p>taagaggtacgttctacatcgggataaccacaatgccagaacctgattgtaccaaca</p> <p>aagccagtcagccgatcctcgctgtccatcactgtatttagagacaacgggtggcacag</p> <p>ataaaccttcaacttgattaccctgtgtaccaggtgtcatgtttctaaaatcatcagcgcg</p> <p>gacagttctgtcatgactggatgcaaaacggcaagaagcccaagctcctaaaggggc</p> <p>gacctttatgaccatggtgctgctggcgttggtgacattctgaaagacgtaggcgccac</p> <p>tgacatcgctatgatatagacaagcttaactgcataactttaggtttggagaaaagcc</p> <p>atcacacagatgccttctgcatgtcacatggcggaagaacagaacgattgtcagcgattg</p> <p>actatctgcaacggcgacgcaacaaccgctccctgaaaagttttacgacgcatcctatct</p> <p>ggacatacgggacgggatcaaaaaagcgggtaaaagactttcaagtggtcgcaaggt</p> <p>cgtaacctcactgagaacggcgaaaacctgccttaccgtgtgtaaaaagtcgggt</p> <p>cggccgacgcagtatccgcaccaggggtacccttttcgacaggaacatcgtttcttat</p> <p>gagggtattgttataccagtgccgggaagccacaacaaggcgcaacagtgtaaacgttg</p> <p>gaactgaacgttcgacacaccaggaagctaaaattaatacagtaacatgcaggcatc</p> <p>gcctgtgttgcttaa</p>	<p>gtcaattaccaccacttaattctaagtattt</p> <p>tgaagtggggcgtgtataagctcagttgatt</p> <p>agccttagtccttgaaagaggcctcatttag</p> <p>cttttaattgataggtaccatgggatgccccca</p> <p>caagttctatgctctacggtgctgtattaaac</p> <p>cgtttcgtagggtgaaaccgtgttcgcgat</p> <p>attaacaaaagcataatttggcgatgtgg</p> <p>acaaaccgctgttaacgcaacttttgcggag</p> <p>acggtgcggattaactttacaaggagtttca</p> <p>aatg</p>
1571-1572	Ga0216255_10040300 GENOME_ACESSION: IMG_3300021179_\$F_3300021179 GENOME_ID: 311738 CONTIG_ID: 40299 SOURCE: JGI DATE: 2020-04-07.	<p>atgccatgtggacctcaaaaagcaaggagattatfaaaagaagaaaagctgaagtatttt</p> <p>tgttggtaccatttactattctgttaaccattgttaactggagaactgtacaatcctgtatagct</p> <p>gggattgatcctgatttagcaatgttggaatatctattgtattcttggaataaaagaagtttt</p> <p>gagtagtggaatttctattaaggacagatattgttaactcaattccgaaagaagtcaatatg</p> <p>cagaggggagcgttatagaagacctggtatagatcatcaagatttttaataagaagaa</p> <p>aggtgatggctgctgaccatcgtattcaacataagaagaacagatgttttaaaataatt</p> <p>gatttcataagcaagctgtgccgataaaaaattattatagaagttgcagcttttgacattc</p> <p>aaaagattaaaaatccagcgattcaagaaaacaatcaagaggagatcaaaaggatt</p> <p>tttggaattatctgtagtatgttctgcatagattctcataatgtcaacagtgtaagggtaa</p> <p>gtcaaaagacaagataattgcaggtacacataattgaatcacaaaaaacaggtggtagccg</p> <p>tcctgacaatctcataactctattgttaagatgccatgcaaaagctgcatcaaaacaaaatc</p> <p>agttaaaagttaagaaaagaaaggttttaagccgaacccctcatgtctacggttaggtg</p> <p>gaaaaatgttgaagatttaagaaagcgttgatcagtgtaacatcatatggctacatta</p> <p>caaaaaataatcgtatcaaaatagggttagaaaaaacctcattgcaacgatgcatttgaatt</p>	<p>gtcaacaacccccattgaaatgggagcgtt</p> <p>gcaagagccctatgttgactagggaacat</p> <p>acaataggagatatgttttatggctgtgcag</p> <p>aagttgtttagggtcaagacaccatgggtac</p> <p>tcctctcgtccatgcaaacgggcaactgtt</p> <p>ttttacagaagtcctatggtactgtcgttagtg</p> <p>gttaaaagctctgtgagggtaggagcgggtg</p> <p>ctgctagcagaacagcccttacaacaatcc</p> <p>cgaagaggccttactcccccttttaggga</p> <p>ggacagaacttgaga</p>

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gctggcgggaacaacggaggcacagcctctcgcagaagaaagaagttgatccaatata ctgttaaacaggttacgagcgaataagaaaattattcaagaagtaagaagccatttaag aaatactgcgtcaaggcatgttcattgctctcaacgtttgatcaagttattttgaanaagaa agatgtttttttttgaagaagaagctctggtctatttgcatacgaanaaactcaatgggcaa gttattcatcgttctgtaaaagccaatagcctgaacattttcaatcgtttaacactttactatg ggagcaatctcgaataaggcattcgtctcctccctgattga	
1573-1574	Activated carbon metagenome, whole genome shotgun sequence GENOME_ACESSION: LNFM01.2 GENOME_ID: 1525 CONTIG_ID: 23713 SOURCE: WGS DATE: 2019-01-27.	atggctgttttgcgtgagcaagcaagcagcgtgatgccctgtagcgagaacacg cgcgaggctcgtcttctgctgcggcgccgcgcgtcatcgtgtgttccctttgtcatt cgattgggtgatcgccctccagtcggaaagtgcatggcagccgggtggcaatcaagctgat ccggcgagcaaaacaacaggcgttgcattgttgcctatcaggaagaacggacgaag acaacagcggagattcgcacatcgcctgcgccatttctactgatggaattgggtcatcgcg gcaagcagatcagcaagtcgttacggcgctgcgccttcaggcgccggcgcgagaa gccaacacctgcggcatcgccaagcgctgttgacaaccgaacaaagccggaagatg ggttgcggccttcgtgcaacaccgtgttgacggccaccgtgtcttgggtaaacgcctgcg tcgctggcgccggcgcaaccgggaatcacagaagagctgtgtgcctgcacacccagcc atgcagaatcccgacattgccggcatcgaatcagcaagggcgaactggcggtttacga agtgccgggaatatccttgaagcatggagccgtcaatgcgcctattgcgacgcgaag agggtcccttggaatcgagcatatcattggcaaaagccaggcggaaccgaccgtgttt caaacctgaccttggcttgcgggtgctgcaatcagaagaagagccctcccttctcagc agtttggtaaaagaccctgcgcgctggcgcgcatcctcggccacgccaagggccattg aaagacggcgcgtgccgtcaatgctgcgcgctggcgcgctgttttctcgcg	gtcaactacccccgactaaagtcgggggctt gaagggaacaatcccgaagccgagattg accagggaagcggttaaccaaccgctac gtgtactacaggtcggttaagcgtaccggcg aatgctctctcagttcggcgctctacaaggct tggatcatgctcgggaaggttaagccgcg aagggttcggatcgaccgcacaggggagcc ggcgctacacattcccaggggagacggcc ccgtcgaaggcgcgctcacaagcccgtaa ggcgaatttggagaaagaagatggct
1575-1576	wastewater metagenome genome assembly, contig: NODE_129_length_4304_1_cov_48.731052, whole genome shotgun sequence GENOME_ACESSION: OVQT01.1 GENOME_ID: 11629 CONTIG_ID: 128 SOURCE: WGS DATE: 2019-01-27.	atgaacaacaacaaacaaaaattgtacctgtgttggctgtagaccggtacaccgttaatgc ataacacacatggttacgcaagagtgtaattgctgtgcgaagcgtagctgttaagca aaagatttctactattcaattattacacgaacacaagacaacaaagaagaataaagtgtg aacttaaggttgattgtgtgttggaaatattggttacagcgtgttagcgaaatgaagaatt aattgggttggaagtgaattattaaactggcatctcagcacgaattacaacaaaggctggg tataagctacacgtgtaattgctgcaaacgttaccgtaagccaagatttgacaatcgcaaa cgtgaagtgatggttgccaccatcaaatgaacataaaaacagatcacattttaaataatt gatttaattgttcttataatccagtgaaatcatttaaaattagaagctgtaatttggatgcaa aaattcaagaatccagatattaaaggcaagcaatcacaagaagaatgaatttggataagaat taaatcctacactgcctgtgctatttgcctgattgattacaaatgtcaattgtgtggtgat tctttaagtaagaataaaaacattaaactagaagattcatcatatagtctaccgttcaaaaagt ggctcggattctgaagctaattgtttacgtgtgtgtacaaaatgtcacacagctaagaatcat aaagatggtggcatattgcacaagtggaatgcttcaaaagaaaagcatgggtgatttaaaa gaagctacgtacatgaatatattagcatctagatttaaaacagcgctcatcacaaagctgaga tatgtttggatcgtacggcgtgtaaaaacgaagtcgttaggttctgaaaaaactcacat aatgatgcttttgttaggtgtgtgtgtgatgagacaaccatgagagtggaatcatctact aatttcaagcaaaaacgttaggcatgattaggtctctacaaaattctatgacgcacataatcat tgatattctgtagtggcaaaaagaaaagggcgagaaattaaagtgtgtgtagaactaaacg caataaaaaatcttaattggtgaaaatgagcgaataatttccgggtgagaaggttttaaaaga cgtgttaactcattgttaaaaaaatccgattgaatcgtgtgataaatttctatcaaggaca aatatatgtctcagcgggaataaaaaacaaaggtgcctatattaagtagtgtagaaggt aaagaacaaatgtttaaaaaacctcgaaagttaaaagtcttaaacgttaactgtggaacttaccat acaacgaacacctcagcgggtgctgtaagccattggaatgtaaa	gttaagtaccacagaccacaaagggtcgggg cttattagctctgtataagctcgttaattag cttaagtgtgctgtcgaagataagctgaa ctacgttaccgtatgaatatataggtactcgg ggttgcctcgaagctccgaactacggcttg taattaacatctctgacgggaaggagaagt gttgcaagcgtttaaacattaggataacatt agcgatgggaatttgcataagtttaataccac ttataaaaaggattttatcccatatgaacaa acacaacaaaaatt
1577-1578	human gut metagenome genome assembly, contig: NODE_207_length_1326_6_cov_7.084551, whole genome shotgun sequence GENOME_ACESSION: OXJ01.1 GENOME_ID: 12718 CONTIG_ID: 206 SOURCE: WGS DATE: 2019-01-27.	atggtgtatatacaggacatagatgtaaacggatgatgcctactacaaggcatgggaag gttaggagactgctaaaagcagcaagcgctgtgtgtaaacatgtccctttaccatca aattgatgtacaagacatcagggttacaaaagagattgtgttaggtgcatgccggaaac caaacattgtgtctatcagcaacgcaagaagcaagaagatttaccagcgaatgattt cttagaagtgtatgttagaacttttcttacaagaagagatcaagaagaacgagacgaa ataggttgagatacaggaaacctgttttaacaaacaggaataaaaagcaaacgtacaggat gggtgacaccttcggtagacataggattgatgctatccgtgttaccgacaacatctgt tctacttccgggtatccctgtcatcgtcagggttccggagttgtatgccacgaataaa gaatcccgacatctccgggtgatgaatcaggaaggaggaactcattggttttggaaatgca gggaatatacttggcaagggtgggcataaatgtcaacattgtaaggaagtaaaagc accgattttgaattgtcatcacatcgaaatctgaaaaacaggaggtgatttcttccatcaatc ttattactttatgtgagactgtcataaggagatcataaaggtaacatcgatttgaaggttaa acaggcgaagtcgtctgcgacgcagccgtgatgggcatcatgaaatggaattgtacg atgagtgtgaatccagatgcgacaacgtttcgtatggcgttcggatattacgaaatacaa tcgattaaatcagggattgaaaaatcccatatccgatgcgtgttgtatttctaggaacttc aatgcgaacgaattgaactcaatttttaaacgttaa	atataatttcttataatgttttaatatatccata aggatcggattattagcctaatgtcttgaaca gagctacgttattgagaatcatatgttaca aggaattgttcccaagttccctgtctcaagg taagtgttaaacaggagtagtatttgcga aacgattgtcttatataaaacctcaaaata acattggcgatgggtactaacagggtttttac cttgattgttgataaacattgaattgttt gtaaaatg
1579-1580	human oral metagenome genome assembly, contig: NODE_4_length_74797_cov_34.078711, whole	atgcggggcgaagttccaactctgtgaaaaacctgtcgacgagaagagagatgcgaagag ccagaagaagtcgcaaaaacagacataggaaagcgcgcttgcacaacagaagaagac cggaaggctggctcgcgccgtctgtgagaaaaagatcgataccacccttaagggtgata cgcgacctgcacaagaattatcccaatcgcgaagaatgatcatcgacagactcgccttgcac gcaagcagatcagcaagtcgttacggcgctgcgccttcaggcgccggcgcgagaa gccaacacctgcggcatcgccaagcgctgttgacaaccgaacaaagccggaagatg ggttgcggccttcgtgcaacaccgtgttgacggccaccgtgtcttgggtaaacgcctgcg tcgctggcgccggcgcaaccgggaatcacagaagagctgtgtgcctgcacacccagcc atgcagaatcccgacattgccggcatcgaatcagcaagggcgaactggcggtttacga agtgccgggaatatccttgaagcatggagccgtcaatgcgcctattgcgacgcgaag agggtcccttggaatcgagcatatcattggcaaaagccaggcggaaccgaccgtgttt caaacctgaccttggcttgcgggtgctgcaatcagaagaagagccctcccttctcagc agtttggtaaaagaccctgcgcgctggcgcgcatcctcggccacgccaagggccattg aaagacggcgcgtgccgtcaatgctgcgcgctggcgcgctgttttctcgcg	gtcaactaccaccggctaaagcatatgggct tgccgaagcccgctcaggggagatgtaaaa gccccgggtgactagcctaaggttttgcag cactacettaccctgaagaatatataagccacct

SEQ ID NO:	Name	IscB sequence	ωRNAs
	genome shotgun sequence GENOME_ACESSION: OQOW01.1 GENOME_ID: 10003 CONTIG_ID: 3 SOURCE: WGS DATE: 2019-01-27.	acgcagaaactcaaaaggaccccaagatctgcccgccgaaggcggaagaatc agcaaggcgaaatgacgggcttttgaacgcgagaagatgcgctttcagagatggg caccgctgccagcactataaaggcaaatccaaagaccgggactggaagtgcacacat cgaagcaggaggactggaggcgatgcgccaacaacatgacacactctgtaaacg tgtcacgacgcgtaccacaagggaagatcgaactggatgtgaacgatgcgcacgtt cagagacgcggccttcagtgacatcgtgcgctggcggtgtacggcgctactcaaggag gaacaccagatgtatcgtacgtctggctacagaagaccaagaatgcacgaatagagaa tgggttagaaaaatgcacatagttgacgcgcgtgtatatacaggaaacctcttggccg ccggcagaatatacttctgtcagaagaaagtgagaaggcataaccggcagatgcaca ggaggactatcggtaaggcgctatcgtaaagagaaccaggcgctgcacacctcttcg gttcagattgttcgacaaggtaa	ggatgtactacctagttccacgctctgcgg catgtgcttaaacagtctccgggaggtaggcc cagtgctgcatgtgaaacctgcggataata ttggcgaaggtagcattacgaccgcacggc ccggcttacagcgtaaacgtgcgaatatctt atgaaaggagcgatgccctatg
1581-1582	Ga0222708_1003174 GENOME_ACESSION: AntAcemetage1576_2_\$ F_3300023242 GENOME_ID: 23008 CONTIG_ID: 3173 SOURCE: JGI DATE: 2019-01-27.	atgcagagagtcctccgtacaaaacctgatggcagtcacgcatgccacaaagcgca gtcgcgtgaagcggtgggtgaggcaaggcaaacggaatgggtcaagaccgatctca ggattaaagcagttcgaactgtaagagaacccctcgggcacaataccagggcattgtg gtgggagtagatccggcaagcctactacaggcattggcgttcagtcggccaaagccac gctgttcaggctcactggccctaccccttgagcgggttagggaaacggatggacaatcg ccgcatgttgcgtcttctcggcgagtcggcggaatcaaccggatgttccctttgacact cgcaaccaccggcagaagcggttcgataaccggcggtgcagcaaaactcgccctc tccgcgccaaccgcaacttgaagtcgctgtagtgcggagttgtctacctgtttccgatt accgccattggctatgagaaggtaaggccgatgtggactgacctcgggcccgaagg gtgcccgtctgggaaagggtctctctgtatgtgtggccaggcctacgccattgagc aaatgagccagattcgccggctctacaccgctatgggtggcagaaggacggcaatgg gacatctcaactcgcacggctctagggtgtgttaagtcacaaacaaagcagagcagac ccccagagccatgccgtcgtatgtgtggcgctggcctgcgggtacttcattgagtacca gtcctttaccagtgccgcaaccatggccatgattgttggccaggtcaatatcacgacg gcaccatttccgtgattaaagcggccacctatctcgcggcagctccacctgatgttc cgcccaaggcgcgctgcgcggaaatattggcgaccatcactcggcatgggttca gaaaagggtatctagtaaacgtgagatggcgggcggtgtcggctggctatgtcag cggggacacggagcgacaggtttcgtctccgacatcaattggcggcgattggccagt tcaccgcatcaaaagttcagctctttatcgcgcaacgggcatcctggccacctgcccga g	ttcaatgacccccacttagcaaacgtgaagt gggggagtgccggagataaccggctcgt aactgcgaatagagccattgagcgtttac gagcaaacacttccaaattcttctgatttgg attagatgtgagggggcagaatctagctcca ctgggtattccaggacacggttaaccgt ggcgtaagagacagtagttttggaattatct tccacatgcaga
1583-1584	Ga0207145_1003 GENOME_ACESSION: Alteio_BWMinHeat_169 _SF_3300020828 GENOME_ID: 22696 CONTIG_ID: 2 SOURCE: JGI DATE: 2019-01-27.	atgcttgcaggttccaagctctgttaacatagaagctaaataagactgagggtataagtcct gctctatgatatgccgtgatttaacattagcaggcacacggttaccagaaatggcgga aggaaaaacatggtactagtattcagacaacaataaaagccggtaacccctgcccgtcc tgcagatgccgactactactaaaggaggcgaagctgctcttttagacgggttccattta ctattattctaatgaagactctctcaaaagggttagaccattacggttaaaaaattgactctgga agtactaccactggcatggccctagtggatgacaaaacgggaaagggttattagtctcttg aaattagccatcgagtaaaacaatttcacgatctatgatatcggcggaagctctagaaaa tcgagaagatacaggaatacccgctatcgtgaagcaagggtataaacagaaaaactgc gagttgatgttgcacacctcttttggcagcagaattgcgaacgatgaaacctggataaaa cgggttcagcgatataccccatctccacataagcgtgaaattgctaattcgataccca attgatgcagaatcgaggatgttcggggattgaatccaacaaggtgcgtccaaggata tgaggtgagagatgacctcttggaaaaatcggtcggaatgttgcactgtgggggttaa gcatcgtcccttgaatatagaccatattgtcctaaatcaagaggaggaagtaataagatc tcaaacctagctatcttgttaagcttgttaacaaatgaaaagtaatacaaatgcgatga atttggctatcctgaagtgcacaaatattgctcaatcatcaataaagatgctgcagcaatga atattacaagaataattcttcttaaacgaatgaaacgaatggcttgcgctgataaacagg caccggagctctcacaacacttcatagaatggagcaaggcttggagaagcacattggat tgatgcagcgtgtgtaggatctgcacaccaaagaaattatataaatgatctacaacctt tgctagtaaaaggctacaggtcacggctcacgccaatgtgcaaaagttgataaatatggttt ccgcgcaactcaatcaaaagaaaaataaaaaaaggttcgcttcaactggcgatgatc aaagctatagtaaaagatgggaaaaaaaagggtaccatcttggtaaaagttcggtgaaga tcaagttggatacttcgacatacttagacctcagggaatagtaacaggcaftaaacatacag catgtagaatcattcaaatctgatgggtatagttatagctacttgcctatcattagagaaa agaatagattgtggagttgatgggactgaacccatcacctcatggatgccatcatgc gctctaccagatga	tggtggagttgatgggagtcgaacccatgg cctcgtcaattaccgcggccattattggcgca gctcaacagctcgcgttgaccagcttcagcc aggactagtgtatcagcgggctacgttaatt aggagcgtcctaaaaatcacaccttgggatgct tgccagtccaagctctgttaacatagaagcta ataagactgagggtataagtcggtctctat gatatgccgtgattaaacattagcaggcac acgttaccaccaagaaatggcggaaggaaaa aacatggtactagtatcgacaacataaaaa agccc
1585-1586	Ga0376510_00006 GENOME_ACESSION: B34_SRodWell667_2_\$F _3300035036 GENOME_ID: 304288 CONTIG_ID: 5 SOURCE: JGI DATE: 2019-11-15.	ttgagcatattgttctgatacaaaaagccgcagaatccggtacatccggcaag gcaagactgctttgacagaagggaagccgcagttttcagacagttcccccttcacgatca ttttgaaagagaaagcttggacgtggcatcaacccattgcggatcaagattgatccggg cagtagagagaccggaattgccgtgattaacgacgattccggagagattgttttccatg gaactgagacatcgaggccaacagatcaaaaacaattctggaagcaagaagcgcaatca gaagatcccgagaaaaccggcaaacgcgatacaggaacccgcgtttgagaacagga ccaggccggaaggatgctgcgccttcaataaaagccgggttataatcgaacact gggtcaaccgctgtgcgcatgttcaatattcagcgaattcaatggagttggttcggttg acatgcagaaaataccgattggctgatgccgctgtgtaggcaaaagcacacctgaa	gtcaactaccctcggtgaagaccgaggg gctgaaaaagccccgaagtggaccagtcta agtcttcgagcactacgttagatcggaat aggtaacctggggtgctcgccagccccag gtctacggcgaagtgttaaacaggtctaa gggttaagccggtgctacttgcgcaaac cgatcataacattgacgcggcaacattacc ttgaaacaagagatttttaaaattgagc

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aagattttcagatcgacaaaaccgtgttgattgtaaaggcagacggctatggctcaagac agatgtgcagggttaacaggttcggattccccggacaacagcaaaagtaactgagaaa aaagtcaaaaggctttcagacggcgacatcgtcaaggcgggtttactccggcaaaaag gttggaacgtacattgggctgtcgtgtcagaaaagcggatcgtttaacattaaaacag tggacaaaacgtacagggcattagctggaaatactgcagactgcttcacgcatctgacg gctattcctacaatacagcgtctaa	
1587-1588	Ga0172381_10003644 GENOME_ACESSION: Leawel6488metaG_2_\$F _3300014204 GENOME_ID: 31499 CONTIG_ID: 3643 SOURCE: JGI DATE: 2019-01-27.	atggtttttagtagataaaaaacaaagccctcgtggtttatgccattctgcaaaagcgag aaaacttttagaccgtcaagaagcagtaataaattatcattatcctttgtaacagcgtaaaa aaagaacagagggaagtaaaaaagacttactgcataaaaatagatcccggtgcaaa gtttacaggtcttgcattctcaatacaaatgaggagattgtatctgtcaggttattcagcat aaggcatttgagataaaaaaaacttactccagggcttcattaagaagaggagaagg gcaagaataaccaggtatagaagccccgttttcaaacagagtaagtaataaaaaggct ggatggttaccaccttcttaagatcaagaatagacaatacaaaactgggtaaaaagc ttatggctatatgtcccataggagaataatttactttgagaatgtaaagttgacactcagttaat ggaaaatcctgatgtaaatggagttgaatcagagagggaattatatgtgttgaattc aggaaatcttagagaaaagaccggctttaaatgtgcttattgtggcaagaaggcacga aagagaacttgaataagacatattattccaaaaagtaggggtggaaagtaacagagtaa gtaatttaacttttagcttccataagtgcaatcagaaaaagggaataaaacggctaaaga attcggatattcacaagttgaagtcattgcaaaaaagccattaaaaagacactgccaatg aattcttcccgaaaagcaatgttattgaattaaagaaaataggactccaataaagacagg tacagggtgaaggactaaatgaatagagtaaatcaaaaacttctaaaacttactttg acgcttctgtgttgattataccggaataatggaaagtgaacaagacagcaagcttaaat ataaaggcagtaggaagaggaaaatataaaaggactgatactgataaatcaggatttcca agagcttacaggcgaagattaggaatttccagggattttaaagtgagatgtgtgtaaaaa gtatatacggataaaaagtagatagtcagtgtaagagctaaaggcagtttactattagagaat aagaaaaatgtttaccgaaaaaatgtcaataattcaacgatgtaatgtctataaaa agtgaataaatgaggagaggttaa	ttcgaactactttctgcaaatccgaaaaagt gaaaacttttcaattatgaaccagccaaaagt cgagtactactgttagaaaaattatgacacc cggagggtgcttccctagcttccggtctgtc gtgtagcattaagcaggtcgtgactcggg ctacagtgtgttacacgtaaaaagcttacta acattggcaggggagactttttaaagcgt tacctgtctgtagtaatacaggtgagagat ggccgaaaggctaaattaaaaagaaagga aagcgggttaaatg
1589-1590	Desulfitobacterium chlororespirans DSM 11544 genome assembly, contig: EJ42DRAFT_scaffold00 015.15, whole genome shotgun sequence GENOME_ACESSION: GCA_900143285.1_IMG - taxon_2582580732_annot ated_assembly_genomic GENOME_ID: 187554 CONTIG_ID: 19 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	gtgcatacaaaagtaccagtcgtccatcaagacggcactccactgatgccatgctcaccg gtcaaaagcccggagctattgcaaaaagggtggggctgttaagaagtggacgggaagcag gaattcttcatatccagctcactacgtcaacacagtaaacatcgcacacctctggtattggga tatgaccccggagcgaatacagtgattttgcatagccgtataaaagcaaatgcaaac tccggcatgatcgtcgaaacagaaatcaagaagaagcgtggaacagcggcgtatatt gcgtcgagccagacgggtccgcaagaccagacggcgacggcgcggttttaaacaccg aaagaacaggagaaactgctgcccgccagatcaaaagccaggtggaatgcggat agcgttttcaagcaactcctggccatttccaatcagtcgaatcgttgggaagatgtca aaatcgacggcaataagctcaaaaggacaaaaggccgtcaatactggactggactatg gtaggtaaaacaaagctctaccgatggttagagcgaagaacagagttgagcttatgtgaa ccggaggataccgccagggtgctgcaaaaagaatcgggctaaccaagatagggcaaaag aaagcccatgttttgaatcccaagccgtagatgatttgccttatgcaattgccacattggt actcaagataaaaagcgtaacagcttcagcgttttggagaaggccagagaatccacgaag gcaattgcaccggctagagccgaaaaaaggcgggtatcagaccgccttatgggggcagc gttacactgggattttaaagaatcaggtagtggaatacaaaaggtaaaacttaccgtacag gcgggacaacgaaggggaggtcagctgcacagttttgactacgataacaggcgaatc acgcagaataccaaagccggaagagtgctgtaaaagtgttggcagagctggtttcaca aaagtagtttaa	gtcaataacccccgttaagaagcggggg cttgccctggcggggcgaaggaaacaaag ttgactagactcagccctgagaggggctac cggaggtgggctataagaacctcctgggtg ccagtcaccagcctggggaattcggcagaa ccgctaaatgccatgatagtttttaacctgtct aagcccgctgttctgtaggggacactca actctgaaaaggaggcacgattaccgtgcat aacaaa
1591-1592	Archaeon SCG- AAA382B04 isolate SCG-AAA382B04 AAA382B04_8, whole genome shotgun sequence GENOME_ACESSION: GCA_003034855.1_AS M303485v1_genomic GENOME_ID: 236541 CONTIG_ID: 110 SOURCE: NCBI_Archaea DATE: 2019-01-27.	ttgacaggaaaaaggaggttttaccacaaagacacaactttaggaatcgacgctggt tactcaaaaatagggtttcagcagtcacagatacaaaagaacttatctcaggaaaaatgg aattaaagaacgacatctcaaaaaacttgatgagagaaaaagtatcgagacaaaagg agacatagaaacaccagatcacagagaacccagattcgataacagaaacgaagaagaag gaggggtggttagctcttcaatcaaacacaaaaaaacacccacataaggttagtga agaaaatcaagaattattgccattggaacagtagtagaggttgcaatttcgacca acaaaagatgaaaaaccgaaatctcaggagtaaaatcaacaaggaaacccctacaa ggatacaacgtcaagaactatttattagagaagttgattacgaatgtgcttactgtagtaaa agtgtgttctcttgaagtagcacataactccaaatcaagaggagggaagtacaga gtttcaacttaactataagttgctgtagactgcacaaagaaaaaggcacaacaaactgc aaaagagtttgggtatcaaaatgatacaagaagagcaggagaatccttgaaggaaacag cctttatgaacaaatgcggtgaaaaatagtaaacgaactaggatgcaaacatacactgc gacacataaccaagaagaaaagactggaagaagaatagaaaaatctcacatcaacgat gctttcgtgattgctggcagtgctaa	gtcaactaccctccctaactcattttgtcgt gaggaaggggctgtgtaggagattagccat atcgatgtcactctagggaagtgcgaatg aaactagctcgggactgagtgctgacaata caccttgcgggtgcttcccaagctccacgt ctggaaaacatagtgacgacctcacagcg tctctttccaggctgtgcaatcactacg gtctacgcctgtccaccattggcgatggaa acctgaccgcactacgcggagatgacct ttacgggtcgtaaaaatggttggagaat atgaagaaa
1593-1594	Ga0376085_0000839 GENOME_ACESSION: LAC_S1_MG_FICUS_ FD_\$F_3300035685 GENOME_ID: 312529 CONTIG_ID: 838	atgaagaaagaaaacaaagagttccagtaatttcgccatcgggtgaaccgttgatgccg accaaagcctcgggctagacgctggataaaacacagatagctaaccattttggaa cgatttaggtatctgtcattcaggttattggttagccatcaggcagacaaaacccaagat atcgtttgttgtagcagccaggtaaaaggtactcaggtattgctgtcagttcttaattc acctgtttatgtcgtatcttctgtgaggttacttctcaagcagcagcaatgca ggcgtcaagagaaggtgctctatcttccctactcaggcgtgttcgaagaggaagacg	ttgaactaccgttcttaacctcggggtag aaacggattccaaggagataaccttggaa cgaagcgttcaattacctcaggtttagtag cgtcttaacacttccgggtgttactagctc ggatttcatgtagcttccggaatcagggg gttggtaatgccaaagacgcgctactcaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
	SOURCE: JGI DATE: 2020-04-07.	tattgccgtggtataaaccttcaagctcagaaatcatcgccaaaagcgggtttccaatcgta agaaaactaagctagcgcctcaatcgctcaaacgcaactcgaaatcgggtatttct gaattatcccaatataatcaatgcagcaattcggtgtaaaaagctcgagctgatgtga taagacatcaggagagaaaagggttaagtcgggtaaaaggcttctcccgtgatgtag gtcaggcttgatgattgaacagttatccaaaatagcggcagttgtcacagttgaaggttg gctcaagacgggaacggtacatctcaagttcgtgaacgcttaagcttaccaaaacaa atctgataaagcctcgccagttccagaaactcatgctgtgatgctgtgccatagctgtag tcaattttgccatacaagcccttcatacagctaatactcgtggctgtacttgataggtgat gtgacgggtacacgtccatctcaaaagtcacatctaaacctcggaattacgcccgaagatt gcatgatgccgtgccagccaaaggcgggaattaggagcgttacggcggctcaacaaca ccctcaaaagcgaggaagggtgacttaattgagtactcgactcatcaaaaggagaaatc aaaaaggtgattgataatttgcgggttacacggggaaaaatctgtcttaagtgtgcta tggttacgtctcgagattcgtaattcaaaatgctggattctgccagaatacaggttg gtaatttcggggggtgaccccccggaaattaccctctatcccctacgaccaacctta acgggtatggtgcactaccccgagatctga	tgagggaagtgacctcaaaactatcttcgag gattgtatcatgaagaagaaaacacaa
1595-1596	Ga0137370_10002449 GENOME_ACESSION: SagtaG_47_SF_3300012 285 GENOME_ID: 35434 CONTIG_ID: 2448 SOURCE: JGI DATE: 2019-01-27.	atgtccaaagtattctgtcgtggatgccaacagcaccactgaatcctgtccatccgggt atgctcgtctctgtgtaaacaggcgaaggcagcagttctgcggcactaccccttgcct cattctgcagtcggaagtggagatgcctcaactgaaccacttcgggtcaagctcgtatcct ggcagcaagacgactggcctagcgtggtcaaggatgcaaccggcgaggtgatctcg ccgagagatcacccatcggtgggtgagatcaaaatgcccctgtgatgtcgacgggg cgtacgtcgggggagaaaggcagcgcgatacccgataccgcaaaccaaggtttcctaat cgccgacgaccaaaagggtggtcctccgcccctgttgaaagtcgggtgtgcaatgtcgt tacctgggtgaagcggctgaggcgtcttggccctctgtgccaattctcaggagttatgac gctttgatctgcaagcgtatgagaaccagaaattgcgggaatagcttatcagcaaggca ccctcgacggctacgaggcaagggaatcctcttagagaagtgggtcggaagtgtgtc tactgtggagcaactcgctcccccttacaatcagcatalccagtgccgagctaagggt gggtcagatcgcaatttctaatgtgtctcgcctgtggaccctgcaatctgaaaaaggaac agtggatattcaggtgtttctggcaaaagcagcactcctcaagcgcactctgccccg ggccaggcgaccgcttgcggacgcggcgagtggaattcaacgaggtggcggttatat gagcgggtgcaagcgggtggactgccgttgatgctggaagcggggggtacaaagt caatcgggccagcggggtgagagaagacc	gcatgaaccccggttggaagcggggg cttgacagagggtaccactcgtcggtag gccacgattttccagctctacatgagaga ctcagctctcggggtcagttcggaacgaa ataggtaccagtggtgctcgccgacgcc actgctacgatgagcatataacaggctta gagggttaagcagtggtgtgatgtgaaac cgttctgaacctgtgcaaggccacctttacc cgcgcaagcggaggtctgaaaggagcac ctatcgatgtccaaa
1597-1598	TPA_asm: Prevotellaceae bacterium UBA3839 UBA3839_contig_75, whole genome shotgun sequence GENOME_ACESSION: GCA_002392805.1_AS M239280v1_genomic GENOME_ID: 134605 CONTIG_ID: 205 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atggtatatgtacttaataagaatggtagcccttgatgccgacagaacgtcattgtaaggt tcgtcgtatgctcgtgatgttaaagctcagctgtacagactgaaccttttaccattcaact ggactatcatcggaagaagtcgttcaggaggtttcgctaggagttgacgtgtgtagcaa gcatattggtatttctgccacaacggaggaaatgtgagttgttagcaatgcaagtggaaagag cgtgatgataattgtagaactatagctactcgttagagaatcaaggagacagcgagaagc cgcaagttacgatataagaccatctcatttgacaaccgtcgcccgaaaggaggttggtga caccagttcagaaaaacagggttgcggcacatctgcgtatcattcatttgcataattccata cttctatagcaaaagacaactatagaagttgcacaatttgattcacaataaatacaagaatga caggattcgaggaatagaataacacaggggtgaaacaacttggattctggaacgtccgag aatatgtactttcaaggagcggacacatttgcagactgcgaagggaagaaagtaagaca agattttgaatgtgcatcacattgaaagtagaaagactggtgggaatgcacccaacaatct tataacactttgtgagacatgccacaagcaataaccataatggcaacattcaactaaaggtta gacgaggaaccaatctcgtgatgtgctgcaatgctcatttctgctggaggtctatata cgtgccaaagcagaatttaacaatgtacattgacttttgggtatctaaacagcacagtcg cattcaaaatggcattgcaaaaactcattgtgcagatgctttctgtatagcaggcaacattaa agcagaacgattaccttggtctatattggtggatgcaacgtcgccacaatcgggtcttg cacttagtgaagccggctaaagggggttgcgtcaacgcaacatagcccacattgagat acctaaagacaaaactacaacaatgatgttgggaatggatgtagaaaatcttttatat ctggaagcgttaacgctcgtctgtatctaaagataattcaagggaatatgtatgctacat gaatctggagcgaagaagcgataaatgcaaaagcagctgagattcatcagacgcaaac gtggcagcatgataattgagcaattgacaaaagtgtgtatag	gtttaccacaaaaatgctggcactaacagc ctatagtgagtaagccgattagcctcagtc ttaacatcatatcagaaatagaaattgaa gaaactcgttagaagtgaaatgtagccacc ttgaaatagctgtcaagtttcaagctcgtg gtcagtgattaaaaggagcgaagcaacg gtgttactgataatgaaaccatttataacatt ggcgtatggcgcataacgggcattagccc gacttacagatttttaattgaattgaataatcc ccaatg
1599-1600	Ga0302321_100092658 GENOME_ACESSION: Fen_T0_1_2_SF_330003 1726 GENOME_ID: 280172 CONTIG_ID: 92657 SOURCE: JGI DATE: 2019-06-13.	atgcccctgcacaccagcaaggcgaagggtgtgtgctgacaggggaaaaagctgcagta cagacagatcccttcagcatcatcttaaaagaccgtgtggatggggatttacagccgac gaattcaaggctgaccccgagcagtaagaccacaggagcaacccctgtcgcggactatgc aaaacgcgggtacgactgttctggcaggagagatcaatcacaggagcagggctatca aagacgctctgtgattcaaggcgagcaatcccgagaagcagggcgcccgtaaaactcg gtaccgtgctcctcgttcgacaatcgcaaacgctcgaggatggttgccaccttctc atgtcccagttcataacgtcacgacatgggcactgagactgcagcgggttctccgctg acatcaatagcgggttgagactgtcagattcgatagcagaagatgtaacccggagatt caggagtcgagtaaccagcagggcacttggctgggtggaagtaaggagatctctgg agaaatgggttagacaatgcggccactgcaacaaaaagaatctgccactacaggctga acatattatccctacgacaaggggaggtacagatcaggttgcgaatcaccctgtcctgc gaagactgcaatcagaagaagacaatcttactccacggagtttgggcatccagaagt ccaggagcatgcccctgtgcccgtcaaaagacgcaacggctgttaatttaccgctatgct	gtcaaccactcgggattgaaatcccgagcat gtgaaccaacagttaggtgttttagcgaagc atgttctagtgtaccaggctcagttataac aactacgttaccagagaatacatagccact tgggatactcggcagtcgccaggtcttggc agacgggttaaataggtgaagggttaagcca gtgctgtctgcataccagaagtgagcgtt aagcgtaaaccgtcgataacattgccgag gcacacattaccgcgcaagcgagattaa aggagtagtgcagcaactg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		atcggcaatgcccttaagccttaggtggcgggtgtctttctggtcaggaggcgggacta agttaaccgcacccagcagggttatccaaaa	
1601-1602	Pelotomaculum thermopropionicum SI DNA, complete genome GENOME_ACESSION: GCA_000010565.1_AS M1056v1_genomic GENOME_ID: 39479 CONTIG_ID: 0 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	gtgtatcggcaaatgattaccagagttcctgtgtggcggtggacggaacccctcatg cccaccacgccaagaagagcccgcctgttaatacgggacggcctggccaacagccgg agaaacaagctggggctttttacgtccagatgctccggcagtcggaaccgaaaccca accctatggccctggcgtggacccggcggaagtacgagcgcatatcaatagcttcac acaacaaatcgaactaaaggcgatggtgttctcgcctgttggcgtgccggagaagatg gaaacccggaggaatctgcggcgccaggagatcaggaacactccccgcaggcca gcccgtttcgacaatcgaggagggaagggttactggttagctccacgcagcgttccaaa gtggaatccggctgaaagtcgtccgggaactctgcaaaagtattccccgtctggtctatc gtcaccgaagacgtgcgttcaaccacgccaggagtcgcaacggtaatacttctccac cgtagagatcggaaggccctcacttctttagtacaacaaagctggccgaactcaagct cgtgaaggtatccgataccgacgctggcgtgagaggttcgggttgtaaaagtatacgg aacgcaagtgggaacaaacccggagaccacccaatgacgctggcgtgtgtat gggtataaccggctgcgaaaaagccgtcacgcttctatgtctggcgaaaaaccagta cgcccgccgtagcctgcaccggcagcatttccagaagggcagcaaaaggccctgttc gggtggcgccagcaacggcactttttccgaaggcgactgggtggaggcagaaaaag ccggagtaaaatcgcgggttgggtctgtgactgccgacggaaccacgaaactgggtg gggtgtgcggatgcagacgcaagcggtataggccagttcagcccgaagaaagtcagg ctgctggccaggtcaacagggttttctgataaacgcggcatga	gtcaaccacccccgctttagaagcgggg gcttgcctggcggggacaggggcaactg gctgactaggcggcaaccttacagttgcag aagtcagggggtacctaatacttccgg gttttgcagcccgactatgtgaggca gaaccgctaagtccacgtacaccaagac cgtattaagccccgtggacagcgccgagg cgacaatacctccgaaaggaggccgtgtat cgg
1603-1604	Ga0163153_10017205 GENOME_ACESSION: OliLakLV19MP6IB1_\$F_3300020186 GENOME_ID: 33722 CONTIG_ID: 17204 SOURCE: JGI DATE: 2019-01-27.	atgaaagtgtttgttgaatcagaatggggaaccgctaattgccaacacaccgcgaatg cccgtttactctgcgcgatggacggggcaagaatctttagtcgaagccgtttgcaattcaa ttgtaaaactcaactagtggctatgtgcgaagataccattttagcatcgacgcgggctaca agacaatcggcttttagcatgttaccgaaaaagaagaattggtggcggtgaagtcgaagc tgttggaaagaatgtcggaaaggctaagggaacgagcgaggtatcgccgcaatcgacg ggggagcaagacgcgctatcgcgcaccacagtcgacaatcgccgtcgcgataagag ctggttagctccttcgattcagcacaagttggatacacaattcacgattaatcgagatgtaca agcgagttttccgatccagacgatcatcattgagacggcaaaccttgatattcaggcgct caagaatgtggcagtgaaaggatgtgatacaacagggtgatcaggcaggatttggga attacgcgaataatattctacacagacaatcacgcgtgccaaagtcagatgcgaaaa gagacgcagggaagatttgcctgttcgacgcgagattttagatgtcatcatatcgatttt ggaagcaggatcgaaacagatcgccctagcaacttgattacatttgcgacaaatgtcatgt tagcaagaacatcaactaaagggtgctgtggggatgggaaccaaagctcaaggga ttcaagcccgaacgtttatgtcaacgggttcgttggcggtcaatcaatcaatgatccat caacacttttgggtatattacaaagtcggcggtatcacgcttgggcttgcgaagtcgcac cataatgatgcctttgttattacttgcggaacaaatcaggagcgctctgaaacaatgttati cagcaacgcagacacaacaatcgagctcttgagacctttaaggatgctaagtatatagata cccgcgatggtgaaatcgcgggggaaaggaaattgtcaagcggtagacgaaaacgca acaagaatttgcagtgaaaacttgcgtcaataccgcggggaaaaatgtcagaccaggg acgcgatccataagaagaagcgctatcgcgttgcgacctggggatattgtctgtataaag gacagaacggattgttaaggggacgcatagcgaagaaaaagcgctttctctacgcg aatgcaaaaccggaataatgcacagggtcaaaaatctgttctatattcatatttctgtttg atggcgaaatttcgcctga	gttatagacccccatccctaaaggaggggc ttgaaccgtaaggaaacagccccctatgatt agactcaggtcgcgagacctacgttatcgc gaatatatagacaccaggggatgcttcacaa gtcctttgctctgtggaacaatggttaaacatc gctgatggtagcggaagtgtgttgc aaaccgcgaataacattgtcgtgtgaacta caaccgcaagggtgatcaaaagaaaggttt tgatgaaa
1605-1606	uncultured Roseburia sp. isolate UMGS2025 genome assembly, contig: NODE_85_length_10901_5_cov_32.350597, whole genome shotgun sequence GENOME_ACESSION: GCA_900557275.1_UMGS2025_genomic GENOME_ID: 270068 CONTIG_ID: 5 SOURCE: NCBI_Prokaryotes DATE: 2019-05-30.	gtggcacttttcgatgcttctagtcggaacccctgcaagcagacaccgaatctaagaaa accgtatttcccgcgatgcacaggcttatccccgatccgcccaccgtttggcagtgatga cactgcatggaatcatccaaggagaaattctcattgcaggaacaggcgtatattctatgt gatcggtcgtgacgagcagaccacaatgcccccagaagacatgacatgtgaaaaaa cttctaataaccgcaagcccgtatcgtgggaacacgttccctatagatcagcttttata gaaaacagccctgtattacagccattgcaactgacgaaaaatcccgtcggacgaacat cggcatcgctgtattaaccgaacagggaaaacttgtattctcagcatccgcagaaacgag aaacaaggaaatcataaattatggaaaaaagaaaaacagcaccgtcagaattccagag ccggagagcggaaagcccgcagcgccttgcagaaacatggtacgacactgaaag aaggcctcgtggatcggaaagcttccatattataaggaaagaaaaaccggcgcacatgcaaa gacatccggaacacggaagccagatactgtaaccgcaaacgaaaaagcggtggctg acaccgacagtcgaacaattggttcgcacgcatctgaacctgtccataaagtacagcgt tttctcccatcacagacatcgactggaaatcaaccgatttgcgtttgcttactgcatgat ccttctgtaaccgcatgattttcagaccggctccgtgaaaggatacgtatgatccacg cagccgtatcagatcaacagcatgggtgtgcttttgtgtgaaaagagatagaagcatt tacctatatcgttccaaagcaaaaggcgatcagaacactggcgaaatgtagcaggcg tctgtactgtctgtgaaaaactgcataccgatcagaacgtccgtgaggaaactgaaac agcagaagaggggacagctttaaaaaatagtgctactaagtgccatcaatcaggcaatcc catatatctgcgaacaactgatccgggaataggaagctcacatgtgcaattttgactgg aggcgaacctatcatatgaggggaactgctgggtatatacaaaagaaacggcgatgaca gcaccagcaatagatgcttatgtgattggctgtgatacccttggaaatcattccggaggaaaa accggattttacatttgcatacagattcgtcagtttcgaggcgaagaccgcaggattatca	gccagaagaactggaagaaagcactgaaaa cgcttgcagaagacctgaagatcaagaccg atgcacgctttattccggtaacacatacagat gaagcttctccagcaattatagcggga acagacatttgaacatgtcaggggtgtgat tcaaaacggcaaaacgatcatcagggaagg tcccttgcagtggtcgtgaaagaaaaatcgtc aaaatcgaccgtcataaacgaatcgatttct gaaaagcagtgaccagactaaaagaagcatt actgtttaaagtcgggtctcgagattacagaaa aaacaactgcataacgcatatgggtatttaca aagtcagctgctccggcagaaaaacagca ggggcagttgagcagagatacgtatgc

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aaagccagcgggaacgcacctataatatcatggcacaaccgtggcaagaaccgcaa ttcacgattccaacagaccagtcctcgttgccagactggtatcaggaaatgtgtgacc tacggaaccacacaggcagaaaaaatgcgatccagcttactgtagttaaaagcaccgc aagatacaataacccaaaccgctgtatccgggaaccgttatttactatcgcggcgaaac gcatgtgatgagcggccagctcagtaattggactgtatctgcgagcagtgaggagacgta agacgaattataagcgaagaatgttgggtggcgaagcgggaatgcggggttgggtt gtggagtag	
1607-1608	Ga0307377_10081489 GENOME_ACESSION: TR3MG_2_SF_33000316 73 GENOME_ID: 281015 CONTIG_ID: 81488 SOURCE: JGI DATE: 2019-06-13.	atgcagcagtttaaaggaaagtaaaagcgtacaaaggatacttactagtcctttgctc tacaatctggaattaaacagagatcaaatgtctcagtggtccagatagccttagcaaacg ccttaaacactcccgaagtgaatcaaccgtcaacgaacaacgaggacttaccatcaaaa gtgttcgtattaagatcaaaaggtacaccacttatgcctactaaatgtccagggtgtagaa aatgttgataagtggaagcaagagtagttaaacgattcccttttacaatacaattaaat gaatgtgaaaatgtgtgcgaagaattagtttaggagtgatcactggatttggaaatatagg attttcagcaataacagagaagaggaaattaatgtgtgacttaattctcgtatggaagaa ctaaagaaagacttgatgaaaaaagaatgtatcgtagaatagaagaacaagtattgtgt atcgagaatcgagatggcgaatagggtgaagaaaagaggctgttaccaccttctactg agagaagatatacaacacatttaactattatgaagcaattgaaaaagctactaccatata aatataactttagaagtagcaaaagctgataitgtctaaatgaaaaatccttaataagagc caagattatcaacaaggtacactttatgattatcaaaatatgaaaagttattgagtaga gaaaaaggtaaatgccagttttgtgtgaagattttaaggacaacccacatataacca tattaaatcaaacgagaaggtggcaatgatagaccagacaatctgtttgttataaaga aatgtcatgaggagatgcagctaaacatttagaatacaaaattaaatcaaaatctgaagatt ataagcaaacactattatgaattatcaacaaaagatttaacaagatataccagagttga agattacgtatggaatattacattgtcgtatagaataatttaggattagaaaaaacgcaat ataacgacgcaattgtatatacaggtggaactcagcaatcaagaacaaagcaataaaaa ttaagcaagtagatagaataatagggtactacaattgaatcgtaaaaggtttaaacatca atcaagaagggagaatctaaagtgtatccggagattgtttgggtagcaggttaacaat atactgtacaggaaatgtcaacaaggaataatattgtttgggagtagtcttaaaaaag aatatttcaatttctcaaaagtagtaaaaacattcaaacaggcgagtttttggaaatatata a	gtcaactaccatgagctaaagactcatggg cttgaatcgtgaggttcaacgtaagagttgat tagggagcatattatgcagcagtttaagga aagttaaagaacgtaccaaggatactcac tagtctttgtcttacaatctggaattaaaca gagatcaaaagctcagtggtccagatagccct agcaaacccgcttttaaacactcccgaagtga atcaaccgtcaacgaacaacgaggacttac catcaaaa
1609-1610	Ga0310147_000164 GENOME_ACESSION: K73_2_SF_3300034683 GENOME_ID: 292989 CONTIG_ID: 163 SOURCE: JGI DATE: 2019-08-09.	ttgttagtctgtacttaacaagcatggaaaccccttatccatgtaaacatcaaaagc aagaaaactgttgaacagggggaagctaaagtagtcaaaaagaacctttcacaattca acttctgtatggaaagtagcggatacaagcagaaggtgtatgtgggaattgatgcaggaag taaaaatcgtgtatagcagtaacaactgaagatggcagggttaattacaaggcacaggt agagttgagcagagatataaaagaaaaatagaacacgcaggaacttcgaagagca agaagaatagaagacacgctaccgcgaacccagattcaacaaccgcgaaaaaccaa aaggatggttgcctccatctataagcagaatagaagcacttataatctcatcaaaag acttactcaataatfaccagatcaaaagattagtagaagtagcgaattttgacgtgcagg ctatttctaatcctacatcatggtcaaaagaatatacaaaaggcaggatgcaggtgtttgac agtgttaaggaatattgaaaatcagagatggttttaagtccattatgccaaagttaagacc tgatataccatgttcagggaactgactgtagaccatctgtaccagaagcaagggtga aacagacaatccaaacacctgtgtgtgtgtgaagcacataatagcaaaaggga tatgagttataacaatttactgaaaacaaccacaaaaattgaggattttctgccacag tattatgaatgttttgagagttatctctgtccgatgttcag	gtcaactaccacccgcttagaagcggag gcttgcgaagcccttagtgactaccctcagc cagggtcaaaacgccatcgggctacgttag gcaagtcagacaccatgggataatgccca agtcctcatgtctgtcagtgacttaaacag tcctgaggggttagggacagtggtcatgca tagcaagcttgcctcaattggggatgggca aaataactccaaaggagggaacacttctgtt a
1611-1612	Ga0073582_115410 GENOME_ACESSION: IMG_3300005663_SF_3 300005663 GENOME_ID: 276644 CONTIG_ID: 15409 SOURCE: JGI DATE: 2019-06-01.	gtgttgagcagaagttattggagagaatacatcacaccgaagagtgctcctccagct cttggcactgtgaccctgtgtttaaacagttctaaagcagaaggaacagtgatgcagagtt aaaaaccttaataacactcccaggaggctctcaccctaaagcagagggggacag gacttgagagttcctgttattatgtcctcaatagagaggtgatcctctcatgcctacttcac cacgtaaaagctagaattactcaaaaaggagaagcagctgtcacaaaaagaacacct ctaactattcaactaacaacccacaggaggaggtcaaacagagcatcacattaggagtg gatgctgttatagtcagctgcgtctctgtgacacacaaaaacaaagaattgctatgctc agaagtaactctagaatagacatggttaacttaattcagatcgacagacgtaccgcaga actagaagaacagaaaaacgtggtacagacaaccttagttcttaaacagagtttaagag caaacataaagggtggctccctcctctgtacacataagttagatactcatgttaattagt caaaaatgtacttaacatactcctgtgtccaaatcattgtagaagtaggttaacttggatatt caaaagattcagaacccagagatagagggcaaggaataccaaaaaggttctcaattagg gttctataatgtcagagaatattgacttcacagagacaatcacactgtcagcactgtaaaa gtaagaattgacctctgtgtatcatattgagagccgacagactgggtgagatagacct aataacttaatcactttgtgcaagaattgtacagtgaccatcatgaaggacttattaaagtt aaagttatgaagagcaaaccttttagagcagcgacttcatgtctactatt	gtcaacgactctggggttaaaaccccgag cttgtacaaaaagctgctgtatgaccagg agctttaacaagttgtgaagcagaagttatt ggagagaatacatcacaccaagagtgctc tcctccagctcttgacagtgacacctgtgtt aaacagttctaaagcagaaggaacagtgat gcagagtttaaaaacctcttaataacactccc gaggaggctctca
1613-1614	Anaerobic digester metagenome, whole genome shotgun sequence GENOME_ACESSION:	atgaatctgtgtgttattagatgcaataaaacacactgttaccatgccatccaatgagg gcacgcaagtgtgtgagtcaggttaaggcactgattatagaagatataccatttacaattatt cttaaacgagtggttgaataactctgtatgacttaactagaataaaagtagaccctgggtc taaaacaacaggtatagcagttgttaatccacacaaaaagagtagtatttgcgggttaatt tacaccatagaggagaacggattgtaagtgttacttaaacgtagccaagttcgtagggg	actttataccctcatggcaacagtttagta gagggctatagcccaagtcaaccagttta agctctatgagagctacgttatccacgaatat atagacaccttgggtgtcttccagctcca agctctgtgtgaagtattaaacagtagaata

SEQ ID NO:	Name	IscB sequence	ωRNAs
	LSQX01.3 GENOME_ID: 1539 CONTIG_ID: 89099 SOURCE: WGS DATE: 2019-01-27.	tagaaggaaatagaaaaacaagatatagaagccaaggttgataatcgtaaaaagaag agggttggtgcccaccttctgtttatctagggttaacaatgtggtatgatgacgcaaaaga ttgatgaaatattgtcctattggtgttatcatatgtagaaacagctaagtttgatacgcaacttat gcaaaatcctgaaattagtggtgagtgatgacaaacaggacattacaagggttacgaggt caaaagagtattattagagaatttgactataaatgtgcttattgtgggatacagaacgtacc attagaggttagagcatgtttggctaaaaagcaaaggtggttctgataagtaagcaacctt gtaatatctgtgttaaatgtaacgatgaaaaaataatgcttattgaagaatttctaaaag atagacctgagctactaaagaaaatcagctcacaatgaaagcatctttaaagattcggct gttatgaacccatcagatatgaatcgggtgatggttaagaaggttgggttgccagtttg ttttggacaggtgggagaaccaagtataacagacataaacagggtctccaaaagagc actggattgatctgtgcataggagaaggtggtgatgatttctctatagcctaataatg caaatfttaataattgaggtattgttagaggttaataagacagatgtgttaattgataagtai ggctttcccagaactaaacctaaacagcttaagcaggttcatggttttcaaacagggtgacat gggtacgcttagtgcaacctagtgggaagtatgctggcacatatgtagggaagtattgtt agagcacgtggagactttgacattattacaaaagaagacaaaaataacagcaacatg gaagcgatttactttattacaacgcttggatgggtatagctatacttttagccccgcttag	gcttagaaggtgaggtgtggttgccttacaataaac cgtggaataacattgactaggagaccgttac ttttaaaaggaggtattcttagatgaatctt
1615-1616	Ga0376687_0067215 GENOME_ACESSION: AusPinMetagenome_11_ FD_SF_3300036817 GENOME_ID: 312296 CONTIG_ID: 67214 SOURCE: JGI DATE: 2020-04-07.	gtgccagagccggaagggaaggcgctgatgggcaccagcggtgctctgcatgacac ccagtcgtgcttccgaagcttctggagctgtcgcgagagcgtttaaagcaccctgctggg gtacgggtcggtgctgagcccaacagcgagagcaccagcgcgaggggagcgt gcctccgatacctgctcgggaccgtcagacagcccggtaacggggccctcatccgag ggcacagaagagtgaccccatggcttctgtcgcacaaacaaaaagccgcttatg ccgtgtacccccagaagagcagcgctcgtcgtacgcgaagcggcggtgtgctc ggctcagtcctgtacgatccggcgatccggagcagcatccagcagagtaccgtac agccggtgtgctcaagatcgtacccggctcgtacgaaacaccggcctggccctggcgcg gggtcagggagagcgaagaggagaggtgacccacgcgctgcacctggccgagctcac ccatcggggagagagatccgcgagcgtcgcgaacggcggtctatcgcggcg caggcgttcggccaacctgcgtaccggccagcccgcttctcaaccgtcgcggcg cccggtggtgcttctccgtccctgcgtcacggatcgacaacgtcgtgcttggccctcc cgctaccggcgctgggtcccgtggtgctgctggaggtcagcgctcaagttcgaca cccagaggtcgcgagaataggagatcacaccggcgagggttcggccatcccgaagt ggccgccccggccagccgtccctgcgggatgcccggcgccgtcaacgccacgccttt gcgctctgtgacgaactgcgcgctgggctcgcgtcaccagccttgtgtggaggac gtacgaggtggaaccgtgcccgttctcgtccccaaagcgcacgcgtggacgcttg tgcgtggggaaactgctggtgattgatgctggcgccagcaccaccccttccatcaaac gatgggtcgggtgctacagccggaccaatgtcagagtcgggggttcccgtggg tacttgatgcgcgcagcgcgaggttctgggatcaaaacgggggaccgggtgcgg cggtgttcccgggggtttgcggcgaggggagcgtacaccggcgcatcgggtgc gagccaaccggcagtttcgcatgggcccgtgtcaggggataccgg	gtcaacgatccccacgatgaatgcgggg gcttgcctctgtgttgatgccgttttcgtgtc cagccgaggtgcccagagccgggaaggag ggcgctgatgggcaccagcgttgcctcgt catgacaccagctgcttccgaagctcct ggagctgtcgcgagagcgtttaaagcacc cgtcgggtacgggtggtgctgtgagccc aaaaagcgagagcaccagcgcgagggg agcgtgctccggatagctgctgggaccg tc
1617-1618	human gut metagenome genome assembly, contig: NODE_8362_length_255 8_cov_1.603676, whole genome shotgun sequence GENOME_ACESSION: UZUR01.1 GENOME_ID: 279697 CONTIG_ID: 8361 SOURCE: WGS DATE: 2019-06-06.	gtgaaggatcacaaccggctgcaagcgccaggagacaaaacttgagagtattgtttta aacaacgaaggaaagccgtgatccgtgttctcggctaaggcgcgcatctgctcaaa gagaagaaggctgtgtgtagggcgtacgccattcactattagctgacgacgctacc ggcgaaacgaagcagccgtaagtctcggcgctgatccggatataagcatgtcggcc ttctgtctcaacgaaaaggccgaactctatgcatccgaagtgaactgcgtcaggacat aaccgacctgtctctacacgacttgactgcggcgctcgcgtcgcgaaccgaaagacgc gctaccgcgcacctgggttcgacaatcgcttcgagcaagcacaagggttggctggca ccgtctgttgagaatcgcatcaatgcacatcttcacgcattgctgccgttctgaatactg cctcttacgaagatcactgtggaacagctgccttgatagcagcgttctgaagaatccag atgtttcggcggaagcgtaccagcaggcggaacagcttggcttttgaacgttcgcgaat atgtcttattcagagacgggcacgtctgcccagcagtgctacggcaatcgaagatccgg tgctcaacgtccatcacttggaagcagacgtaccggcggttctcctagcaatctact cacgctttgcgaacactgtcacaagcgctgcaccgtggcgagatcaagctgaaactc aacagggaatcgttcagagcagaacacctcatggcgtatgcgctggacgggttctga atcgactgaagtctgcatacccgaaactggatgtcgcgaacacctacgctatctcacga agcacgcacgcatcacatcagggttgcgaatctcattgtcggatgcgttctgcatcgc cgggaatttcggagcaaacgactcgggtactacttccaaaagcagaccgtcggc acaaccggcagattcacaagcttcaatcctgaaggcgcgcatcgcgaagcgaatcaa gcgcttacgaagtcaagggttccgtctctttgacaaggttgatgcaaggagaaggt gccttcatctcggccggctcgcgggaagcttcgatgttcgaaagcttgatggaacgc gaatttcagcaggtcagcttcaagagcttctgctcgtcgaagcgcgaacactctt aactgaaattcgaaggagggcgccgttccaccctgcctaaccgacgggggtctccgc gtcgattttga	gtcaactaccccgccctgaaggccgaggt tgaatgagccctagtgtgactagccctagtc gtcatttgacgggtactggttgggaatgt acaggcaccggggtgctgacatctagct cgcgctctcggccgatgtttaaagctctt agaggttagagcgggtgctgctgcgaagaa acccttccatcattgtgaagatcacaaac cggctgcaaggccgaggagaacaaactt gaga
1619-1620	Draft_c0007806 GENOME_ACESSION: IMG_3300000032_SF_3 300000032	atgccgacggaaaggcagcggaaggtgcggcaatgtttaaagcggcagggcgag agtaatacaagccagggcggttacgatccaactaacttacgaaactacaagttacaccaa cacattaccctgggcatagacggcggttaccagaatcgccctatctgcggatctccc gggaaggagttgtatccggcgatgcaccccttcaaaaggccaggtcgagcgcaacaa	gtcaacaaccnccgcctataganccgggg gcttggggnagcgtatagnggaaaagctat acgtcccaagcccggttgattagccctcagc caccggcggttaaagcgtcagggcgctgcg

SEQ ID NO:	Name	IscB sequence	ωRNAs
	GENOME_ID: 27423 CONTIG_ID: 3439 SOURCE: JGI DATE: 2019-01-27.	agaaagattaatgtaccgcagacagagaagaagccgctgcgttaccggaaccacgc ttcaataaccgtaaaaaaccgaaggtggtgcccaggtattcagcacaactgga cagccacctgcgctgtgttgcctgggcaaaaaagctgtttcatcactgacatagtata gaagtagcaagtttcgacatccaggaatataaaatccggcatccagggaagagta ccagcaaggagacagtagcggcttctggaaccttcgcgaatcacctgcaccgggacg gccaccagtgccagaaccttaaatgccgaacaaaaaccaggtgtctcaaacaccac ataggtttttgaaagacgatatgcggacagggcaggaacatcacaccctctgcgac aatgccaccggccgggaaccacagaaggaagcaagttctgtgggatgggagccg aaagtaaaagccgttcgggcagagacatttatgacaccgtaagatggaagatagtaaa cgctctaaaatgccggcatatagcccatataaacgaatcccgaggatagggtctaa agctggaaaaaagccatgccaacgatgcctttgtaataagccggggcagttgccaaca agggcaataacttttagtatagagcaggtccggcgcaataaccggtccttggggaagttt acgattccgatatagatatccgcacaggggagaaagtcagtgccaggagttgaact gtggccggcgaccgggaacaagaatctgaacagtcggaacctgcgggtatagccggg gcaaaaaagtatctaaaggcaggggtgagcataaagaaagcagaggtaccgtatcagcc caaagacattgtgtaattgtgtggctggcggtacacgggttaaggtatgcatagctacggc gcctatatcaatctggcgggaattgagcaagccgggtgaaaaaggaactagtcagactgt gcggtacagcaaaagttttgtgtgctgtagggaaaaactgtcgtctgcggcgacc ggcatctatcttcgcctaaagagcgggagctcttcgcgggattatgataaatatattca ggaattaatactacacatttaattgattgaaagtgatgatgaattgcatactacgcaattca aagaacattagattgtctaaaagatactccaattgggcaacaaataacacgttttag gtacaccaatttgactttgtaattgagtgattttacataaaagcaattgtaactttattac atatagatggcaaatctgtgacataaacagagcatatgggtgagaaattgaaagtctca aagagtataagcatgaattcaattcaagagagtgcttggaaaaaatatcctacctgaagt aaagtggcgatcgatccaaaggccatcatgcaacaaatatgcaaatgtctagtgt ttgcagcagtaaaaggatactgttattacagagaaaaatccgtcttcttgcgttgggagcc catcgaaaaaataatgatggattgtggactgatgtaa	ttgcaatggaatatataggcactcgggata ctccactcgtcccgacactgcggccaatg gttaaatcgtctgaagggtaggcggaagcg ctgttggcgcaaaaccgtgtgaactgtggc gaagtggaccactgcgaaagcagcctgtg ttccgaaaggagatcaactg
1621-1622	Ga0376513_002019 GENOME_ACESSION: B37_SRodWell667_2_\$F_3300035039 GENOME_ID: 304291 CONTIG_ID: 2018 SOURCE: JGI DATE: 2019-11-15.	atgatttatttgtgaaagataagaatggttaaacgtggtcatccgacaagacgtgcagatta cataactaaacatttaaacgtggtactgctaagataaattagtcgactaaagatagcttaac agtaaaactattggaatattgaattaaagatgaagatactattgacgtgaatttagtagtagg tattgatcctgtgacaaatcatattgattgtctctatataaaattataacaaaattattacttt gtaataagtggtgaagcagatattagaagtactgatataactaagaatcttataatcggaa gatgtatcgtataactagacggagatctgtgaaaaaacgttcttcgtaaatatggaaaa gctaagtttcgtaaaccaatttggaaaaacagaaaacgtaaacattttcaactactttcaag catttaattaacacacataattatgaatttaaatatttcttaaatgatgttcaattaataagta cataattgaattgctaatttgatactcaaaaaatggttaatacaaacattaaatctatttgta tcaacatggaccacaatcatgtttgaaatactaaatcttatatccgtatgcgagataattat actgtcagatatgtaaaaaaattgtctgattataatgaagcatcatattatattggcaaga atattggtggatctgataaacagataatcttatattactttgccaatttgcataagaaagtg cataaaaaatctgttaaatgtccttctattccaactcagaatcttaacaaagcaggcttcttaa ttcatgtatgaaatataattatgataatttgaaaaaggtgaccgactcaggatactcttggaa agtattacaaaaatagtaagaattatctggaatagagaaaactcatgaaatgatgctaa aattatagcgttatgtgatagtcttgatttacaagattgaaattataatataatcgacttaa ataactcatgtaacagtttaacaataatcgtagacatgatcgagcttgggtaaaaaagatagaa gatcgtataactatattactggacaaggtaaaaaagtattgtctataatcgtagtaaatca actgtcaaaaaactcgtggattagatgagttaaaacaagaattaaaaaacgtggtataat taataagattcaaatattgctaagcctgtgtgccaatatattgtcgggcaatataataaa cgatttattcctgtgcaatttaattataatggagattgatactgcggaggttggcgct ctacacaatgcaaaagtaacttgaaaataatggatatgttaacaaaaattatgcaaaagta attcgtataatagtgattagttgttttag	gtcaactaccgggtgcttagcaccggcata cagatgtggcagatgcccataaagtaagtc tagttaatactagggttgaacagacaattta gctaaagtttactcaggggtgccactccag ctctgaacacatttaaaactagaggtaaaa agaataatagtgtaaatcctatcgtgactatt aggtaacgaagatcagtgctatttagctaa agtcgagggagacgtaatgcaagtttctcag atcagaggaaaattttaaattgatt
1623-1624	Ga0071116_1000899 GENOME_ACESSION: IMG_3300005077_\$F_3300005077 GENOME_ID: 305312 CONTIG_ID: 898 SOURCE: JGI DATE: 2019-11-15.	atgttcgtaccaacaataaacaaccaagcattccggcaatgccacaacgccccaaag agcaaggaggttatatgctctaatactgctactcactcgtgagtcacggcaatttctgtatt cgactaaacgtagatgtgcaagacaataaacaagatatagccatcggtatagacacaggt agtaaaaagggaagccttttagcgtgtgctctaaaaagcacacgtacctaataatcaggca gatgcattgacctacgtaaaagctaaaagtagaaacccgtagaatgcttcgccggagtcgc aggagccgaatactccatacgcgaatgcccgtggaatgaaagataggaagactacc gccctcgaccaaaagctcgatgggattggaactcgcgattgttaattctaaagcgagta atacctcttactactcttgcgtcgaagacattaaagccaaggccaagaaaaaagaaagg aagtggacatatcgttttagtctattgaagtgggaagaaactctataatgcacttgatt cgacactaaagcctaaagtaattgtggaactaacgctgaagctacgaaagggaagcc ggactcaagaaaaacaagaaaaactgagcgaaagttcgagactcactcgttgatgc gtttgtgttagcgtgtcatatcatttgcgacataatccggacaataaagtcaccttgc gtttcgccaatagaattaaaacgcagaatgctacatgccaatggtgccttcaaaaggagga actcgaagaaaaatcggaggaacaagatcattagactaaaacgtgggagtttagtcaa acataaaaaacacggaatatgctactgttggcgccacatgaaaaatcgaatctcactaca	atcagtaaccacgactaaagtattgggtct gtcgaaaggaaactaacctctcggaggat ggattaaaccactgattagacgcgatttca cgggtgaaatcaagctagccaactggacgtg gtagtccgaagaatgtaggatgctctca agtctgtctccttctgttgcca

320

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gataaaaaaagaatggcgaagaaaagtagctaaaggaaacattacatgctcagtagctg gccccgtagatgataatgatggttactctgactggaaaacagaagatgtaagtggaac gtatggtaaaaggattacagttcaactaa	
1633-1634	Ga0137415_10000002 GENOME_ACESSION: CZOHybAssembly_2_\$F_3300028536 GENOME_ID: 280031 CONTIG_ID: 1 SOURCE: JGI DATE: 2019-06-13.	gtgataagtagtagtcaacttaagataacattggcgcagagtgactaacaagtccttacgg gattattggaattttgatgaaaaaacaagaataatggaagacgtatcgttatctcaaaat caggctcagtaaaagttaaaccattttttgtagtgaattggaagtgatggttaaccac taatgccatgtagtggcaaaagtgcaggaaacttttagataaaggctgtgcgaagtaaat ctgtagagaaccgttcaattattcaattactgtatcgaacacaaatcaattgtaattacaacct acacaaataaaattgatcctggtagtaaaacaaccggattagctatcgtgcgtgaagata gaccaaacgaaattacagttattcgttaattggaattagaacatcgtggccgtgttattaatt aaaactttataatcagctatgcatcgtcgttaacggcgttcgtcgttaaacacgttatcgtg aagctcgtttcttaaatctgactagaccacaaggatgtttacctccttcatccaacatcgtat tgatagcacgtagtcttgggttaacgacttagacgatgtttccattacataaaattgtttat gaggatgttaagttcgatacacactggctcaaaatctcattagaaggcataatgtatca acaaggtagcattattagatacgaacttcgtgaattatttactttaattcgtatcataaatgtg cttactcgcgaatgggtgttctaaagggttttcaaatagatcacgtaattcctcgttagtagag gtggtatcattcgtttgatttaattgttctgtcgtcgaagagtgtaatacaactaaaggtaatt aaaagttatcaagagttcttaattttatcaaaattagatagataaatttaagaacaaatatt aaaacctcttcgagatatggcgggttagaacgtatatacgaatgaacttagattagcactta ttaaaactgggtttatttaattgaaactggaagtggtgggttaaacgaaatttaataagattcgttt aaaatacccaaaactcatgctctagatgctatttgggtattgttataatgtaaaatcatggat ggattcataaaacattacataaaaatgtaattgtagaggtagatatacaagagcaatacct gaccaatattggtttatttaaaagcttatatacctagacagaaatatttcttgattcagaacag gtgatctgttcgtgtaagatcttctatcaaatctggtattaaacccaccatttgggtcaaggt cagtcgtgtagtatgccgagcttctggttaagttgttttaaaataaatgaaagatttatcaa ccttctatcgatactgtcaacgtgttcaacatcgtatgatatgttacataactgttatata a	gtcactgtcgaataccttagtaacctcgga aacattaaaaatggaaaaagttcaccactc cactttaacatcgtagttgtaagactagatga atcagcctcagtagtaaaactactggttttt gatgcatacttggatgccactcagttcca agctctatggtcttataaacaactctgtg agataggagtagtgtgataagtagtagtcaa cttaagataacattggcgcagagtgactaaca agtccttacgggattattggaattttgatgg aaaagaacagaataatggaagacgtatcgt taattctcaaaatcaggtcagtaaaagttaa ccaattacttt
1635-1636	Ga0256842_1002984 GENOME_ACESSION: 4559240metaG_2_\$F_3300026512 GENOME_ID: 20948 CONTIG_ID: 2983 SOURCE: JGI DATE: 2019-01-27.	atgccaaftaaacagataaaaatttttagtggatgcagaagagaacccgtgctaccta ctacacctgcaaggcgaagaatacttctaaagaagggtgaagctaaagttacagaatga ttccctttaccattcagttaaagcaggtttatcaacacctcgtcggaaaatttactgcctctgtt gatgatggagctaaattggtaggaattgctgtaaaagggaagatgaaattgtttttgtag ctaattgtgagactcaggcaagatgtaggcagaagaaagcaaaagagcaatgtatcgt agaacacagaaggagcaggttaaaagtataagacctgcgaagtttttgaacaggaagagagc taaaactgctaccaccgtctatcatagatagaaaagaagttgttttgagaatcctcaatg acttaaaagaatttctcaacattacggaagttgttattgaacaggctcagatttgcatttctct ctgtagctggaaagaaaactgttaggcagaaattccaacagaaaagatatgaagggaag aagctttagagaaaaagttcttaaaagagacaattacacctgtcagatttgggaatcaaaag agaatcttgaagctcatcacataattccaagaagcaagaggaacgaatctttagaga atggaattaccctctgaagagctgtcacagagcagttcatgaaggaaaagataaaaataa ctgctaataattcttttctaaagctccatcagcaattcagcaaggaaagtggtggttataca ggaaagtaaaagggaacagtttaaaagtagaataacttttgggtatctaacaaagaagaa gagagaaaaactcaatcttctaagaccattatgcagatgctgtgcaatgcttaactgca atcatgtatttctcctgtttatctcctcctccacgaagaaggagacctgagataaacaact ccacaagaagcatgaagaatataaggattcagacactttgatctgtagtgtccttatcat agaactaaagggaaggtgatagggtgtgtgaggagtttaaaagaaagagggttagcttta agaactaagttttcaataaactttgtagtggatacacgaaaagcagactactctggagacc aaaagggttagttagtattgtga	agttttcattacggtttggatgggtctgaggtg gaggtagctcctgatgaaggataagctgaac ctgccgaacagggaaggttaggaggtgaaga ggagagggaattccttcccgctaccctacgtt ctgtagttgtccataaggacagacacctgtg agtgaggacaccagctccaggtctgtgtgtg gggtggagacgaaactcagcctgagatggt gttgggtaacactcctcatcgtttagagact acgggactcaccgaggtgcagtttaaccccc tttatgggaggttagcttaaatgccaaftaaac agataaaa
1637-1638	Ga0209777_10021516 GENOME_ACESSION: FrelakHBHBSPAdes_\$F_3300027896 GENOME_ID: 26210 CONTIG_ID: 21515 SOURCE: JGI DATE: 2019-01-27.	atgcagaagttacaaggaaagtttaagaacacacaaaggatgcttcccaagtccttctgct ctgtaaattttcaattaaacagagacgaaagctcagtcgtgagaataaagtactgactgt aacaatcccgaagggaatctagctccagcatcacaggacgtcattaaaagcaattgtgtat gtatttaacaaaggagggaaccccttaattgcctttagtgtttgcaaaagctaaaagattgtg aaaaaaaggagcgcaaatgttgaataaagatttcttcaaatcaaaatattttgaatg taaggaaagagattcagtcggtatttttaggaattgatcctggatatacagaatgtaggattttct gctaggactaccaagaaagattgataagagtagaagtagtaagaacagatatcaat aaaaaacttacagaaaaaagaattgacagaaagaatgaagaacaaactttgttacag acaaccaagggtgtcaaatagattgctcaaaaaaacagggtgtgtgtccttctgtg cagcacaaattggatttcatataagattagtaaaagaaatctgtaaattgcttctgtatca aaaataattgtagaattggcaaatttgatattcagaaaattataatccagataatttctggag tagaataccaacaaggctccttctgtaggatatgaaacagcaagcttacttaattgctaga gaaggaaaggtcagttgtgtaataaagaatccactaaaggaaatccttttaaaatcatca tataatacaaaagaaagatgggggcactaataagccttctaactagcttactgcataaaa aatgtcatgataaattacatgagcaaggattaaaactaaaagcaagtaagcaatttaaggc agagactttttagtccacagtaagatggaataatgtagaagaattaaagaatctccacta gtattactttttagatattacaaaagtaaaaggcaggaagaatattgaaaagctcatt ccaatgatgcctttataattctgagggaagaccaaagaaagcgtgtgtgcctcaatatt	gtcaactaccacgactaaagtcgtggcctt gtcaggtgactggtaacgtgaagattgatta gggagcttaaaaggaaactttatgcagaagtt acaaggaaagtttaagaacacaccaaaggga tgcctccaagtccttctgctgttaattttcaa ttaaacagagacgaaagctcagctgtgaga ataaagtagtactgttaaatcccgagg gaatctagtcacagatcagacgtcattaa aaggcaatt

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aaacaaaaagaaagaataatagatgctacaattaaatagaaaaggctccagccttcta ttagaaaaaagataccccatacaacctaagcattagtaaaaaaatgggaaaattgc tcttcaggaggagtagactgttaaaggtaaaagaattattgtaaacggaaaaagtgcaat ataagttattagataaatggaaatttaatttggaggtttatggggcgagcaattcctccca catctaaagatgtgggttcctgctgcgaattatga	
1639-1640	human oral metagenome genome assembly, contig: NODE_138_length_6003_8_cov_8.947085, whole genome shotgun sequence GENOME_ACESSION: UPOV01.1 GENOME_ID: 16706 CONTIG_ID: 137 SOURCE: WGS DATE: 2019-01-27.	atgacaaaaacataaccaaaaaaacacgcaagctgctgctgtagatcaacaaggtaa accgttaatgccaacgacacgtcttggcaaaagttatcgcttttaaaaaacaaaaagctc ataattgtcatatgaaccgttcacaaatccaattggattatgaacctgacacacacattattca accgatgacactcggcgttgatagtggtgctattcattcagggttattctgtaactaataaaaa ccgtgaataactatagtgctgaagtcacgcacgtaataatattctaaacgtcttttagataga cgtatgtaccgcaaaaatcgtcggctcagaaaaaacacgttaccgcaaacacgattfaaa aatcgtaaaaacaaaaaaaagggttggtaccaccatcacttgaacaaaaagttgcagtac aagtaaacgaaattgaccattacatcactatttcttattgaacaatattgtgaagtcgc tgaattgtatattcaaaaaatcaaaaatcctgataattgcaggcattgaataccaacaaggaa ccttacaagggtataacatccgtaactatcttctgaaaagcatggacgcaaatgtttctattg tggaaaaagcgttctaaagtcgaagttgcaacacatgataccaaaaaaagaggtggctct gatcgtattgataatcaactatcttctgcatgactgtaataaaaaagatacattgaccg cagaagaatttattcgacaaaccttaccacctaataaagcagctaaagggtfaaaacaatt gcctaatgaaaaacgtctgtttaaatacatggcacacatgaacgctaccagatgggctttat acaatgctattgacgaaaaatccaatgtcaaaataacttatggctacattacaaaatac aaccgtattcaagcaggtcttcaaaaagctcaccacattgacgtaaatgtattactgatttt caaccgttccatctataaaacaaactgttataaaaaaaatgcttagacacaatcgtca ttacatcaattgactttcagcaaaaggacatatacgaaggcagcaagtttaccacaattat cttcgggtttcagttatagacctcgtgttattgataatcattattattataaaaaagccgtcga agttcagatcatttgaccttactctatagaaggttfaaaaaatacaaatcgaagttacaaa aaattacattattagcacatacaaatgcatacttaactacagatacactaatgactaa	agctagatcaataaccaccactgaaatgg gcggtttaattactctttacaatgagttatc aaaccgtgattgattagccttagtttttaaaa ctccgttactgttaatatataggtaccgtgg gatatctgcctagtcceccagctctacgcgtta ttattaaatgttctgattgtggaacgggtg gataacgataaaaaactacaagataacattg cgaaggcaacaagggtttgtttataccgg cttaccgcatacaataaacaccaatacaact agaaaggactttgtcaaacatagacaaaa acactaaccaacaaacacgcaa
1641-1642	Ga0167622_1004112 GENOME_ACESSION: IMG_3300015158_SF_3 300015158 GENOME_ID: 28719 CONTIG_ID: 4111 SOURCE: JGI DATE: 2019-01-27.	atggtttctgctcgaaaaaataaacacccgctcatccgtgctcgaacacgcgcg cgctcctcctcaacgtggggggcagcattcacacgctctaccattcacgacccgtc tcatgctgcactaccggcgttctgcaaccaatcgctcaaaaatgatcccggtccca aaacaaccggatcgcctcttgcgagagcagtcgagggacgccaatgtactgttcttg cagacctcgcgcaccgcggaagcgtatttccaagaagatccacgagcagcagggtta tcggcgtcggcgccgctcgcgaaccccgctaccgcgcgccacgattaaataaccgc acgcgaaaggcgggttggtcgcgcctcgcctccagcaccgaattgacagtacgttctc gtggggccagaaactccaagcagcgcgcgccatcccttggcgcgaagagctgttga gggttgataccaggtatacaaaaatgtgaaatctccggaactgcatacagcgcggga ccctctacgaatacgaagcagcgaataatcttgcgaaatcggagcccaatgcagct actgcgacacgaaacgggaccattgaacctcgtacatgcctcaacccaaaagccggg cgatcgaaccgctcgcgaacctcgtaccagcatgcatccgtgcaatacatccaagg ggagccagcctatagaacaattccttagccacgacccgcttcgactcgcacgcatccgta accagtaaaagacgccccttaaggacgcccggctgtcaatgtactcgtggcgcttt tccaagcactgtcaagcctcgcgtaccactccaagcgggaacaggcgggcaaaaaa atttaactgtaaacggtatgacctgcaaaaacgatacgtttgatccgtctgtgttgcat gatggacacgctcgtacgacgaacagcaacagcaattcgttaactactgtat gggacgcggagctataacgcacgcgcgtaccgcaatggtttccacagggatatt cttatgcgggcaagcggcgcggttgctacagggcatcgtttatagcatcggtta ccaaaaggcaagcacaccggcaccatacgggacgctgtccgtacgggcaacagg gtcattcaatatacagaagataatgacgtcttcaaggcattagctacgacacgtccgc cttctcagcgcacagatggataccgataattgtcttccccctccagcgaagacgccc ctcctcctcggcctgaacgcccagggtctccgcagcgtaa	gtaaactccccggcctaaaggccggagct ttctgaaaggacgctcatattaccagggtt ggcgtgatgtagatcagcctccgttagagac agggttaagaccgacgttgggatgttctc agttccaacctctcgaagttgtagtagaga cacgcccggtaaggcagcaaacggattgc aacaacacccgcatctcaatctccgagg ggagcagccccgaaggctccgttacaag gctcttagcgaacacaaatagaaggga gtacgtcatg
1643-1644	Draft_10008528 GENOME_ACESSION: IMG_3300001594_SF_3 300001594 GENOME_ID: 27524 CONTIG_ID: 8527 SOURCE: JGI DATE: 2019-01-27.	atgaaaaatcgtgtttcgtgtggtgacgccaccagcatctttaatgccgtgctcgtcaaaa acgggcacggctactgcttcaacgcggcggtgcccgcgtccatcgtttcaaacgtttac tattcgttggctgacgctccaggcgtcgtcggcgctacagcctgttaccgtcaagatt gatcctggtagtcggtatacaggattgttagtggtcgtggtcgccgataggcgtttcacg catctccatgggtgtttctgatggagctggagatcgcgggctcgggacgcagcagc cctcgtcaacgagccgactgctggaatcggcgagctgtaaaacacgctaccgg cctgtacggttctgaaccgactcggcaggggctgttagccccatcgttgcctc gagtcagtcgacattgacgtgggtgacgacgtcaccgctgggtacctgtgaccgca ttggcgaggagctgttgccttgacacgcagcgcacatgcagaatcctgaaatctctgtg gttgagtagcagcaaggacgctactagggtacgagattcgtgagtagctgatggccaa ataccggggctgttgcgttattgctgggtgtaatcggaaggacgtagagatacagattga ccacgttcgaccacgggctcgtggcggtgcgaatcgcgtgagcaatctcgtactggctt gttggcgctgtaatcagtcgaagggtgctcgacctgtagaggagtttctcagggcgac ccgaggcgctgaaaagggtctagcgggactgcgtcaaccgctacgggatgctgccg cggtaacgccacgcgtgggtcctgacctgacttgggaacggggtccccgtc caaacgggagcgtggtcccaaacgaatgggaatcggactcgggttgattaccgaaaa	gtcattcatcgtaaataaaaaggagattcaca atgatcaaatgttgatcagcttcaaaccta agcacttaaggagccagactcagctccc gtaaaacggaggctacgatccccagaat aggtagcggcgatgctcctcagtcgcc gctcagcgttagtcatfaaacagctcag ggttaaggcagtggtgtagcgaacact ggggatacttctcagagaggagcccgct ctagaagccggcgtcacatggtccgtaag gaaacccacatgaaatcgt

SEQ ID NO:	Name	IscB sequence	ωRNAs
		cccatgcgttgatgcgttggtgttgagaaagtggtatctgtggcagatacaccgggccc atcgcatgccgggttaaatgcgcgggaaggggactttaccagagaactataacggaccgt ttccgggtcccgtcgcggcaccggatgcgccagaagcttgttcaggcttcaagggtgtg gacatttgcctactgtcatttactacgtagacaaaagcgatgttagacacatcacct gcgcctaaggggcggtttcagttttatgacatcagtagtaagcttcctacgacgtcaattt aggttttgcgtagagtgcgatcagtgatggatttctgtttttaccaagaaggagatctg tga	
1645-1646	Ga0315284_10052402 GENOME_ACESSION: YL17G09_16_MG_2_\$F _3300032053 GENOME_ID: 281635 CONTIG_ID: 52401 SOURCE: JGI DATE: 2019-06-13.	gtggttctcggccctaaggcgacaaaaccaaggagattttgaagttggtatgtagtta acaaaaaatggaaaccccttgatgcctgttaaggaggcaaaagctaggcatctgctaaaaa gcggtaaggctagatgttttacttaggtgccctttacaaattgtctcaattgggaatgcgaag aggtgtgacaagaggtttagcttactgatacaagtagtagtgaggtaggtagtctgt taagcgcgatagtgggagtgcccttactgctgtaaatagaactcagacaaaacgtaac aagagctaccgttcaattgtaaaaggtaaaagaagaagtttcacgaagaatgcagta taggagaccacgtaggaatagaaaaactagggtatagacagtgtagaaataggaaatagaa ctggttaactatgtcaaaagtattcccaacacttaggagtaaaattgaaggcatgagag agaagtacgtagggtagaaaaagctgttaccagtaaacgggtgttagtggttagagggtc aaagggtgaaggccattttaaggatggcagcttggaagaacagtggttaaatgttcagcg acagggtgttgagagagatggttccggtgtcgtcactgttaaaaaaggtaagagagaactt catgtccacatttgaagcacgttaaggatggcgggtcttgatactctagaaaaacttagtaac gctgtgtaaggagtgcatggggattatcatcagggtcttatactttaagataggtaaaaca tacttataaaggtaaaagttagacagaagtagctataataagaagaatttagtgggtgaaa agtcggaagacgtttacggctttcaagtaaaagctaaaaagaatgcactagaactatcttat tcaccacttaagtgtctgtgcggcttcaatgttaaacctcagactaatgtttatgggtattag gtgtatccctagaggggattaccaaaagactcgtgtcgtcactctcaacaagtagttcct aaggggaaaataatgggttttaaccgttttgataaggtagatatttgggaaaggcacttttt attaaaaatgcgaatgagtagcgggtattttaaattacagatataataataaagataacca aaagtaatactaggaagaaaactgaagcttttaggaaggagaagatcctgtcttattgtcttt gaccaccccaacctccccctaa	gtgaacctccccctccctaaggaggagggtc tctgaagtgatttagagggtaacagggttca ccagactcgggtctgaagaatttagaactccgt tatttagcatgtagaacccctagaatgccgcc tcagttcttaggctcttcgtggcactgtaaac agacttggtttctatgaaaaaagaacgtga gtcagtcacccacaagggtgaaggctaaata acattgtcagaggaagtcctgtgttttagg tggttatagctaaaaacaggcataacaggt ggttctcggccctaaggcagcaaaaacaa ggagattttgaagttg
1647-1648	Ga0207193_1000837 GENOME_ACESSION: IMG_3300020048_\$F_3 300020048 GENOME_ID: 30012 CONTIG_ID: 836 SOURCE: JGI DATE: 2019-01-27.	ttgcatagaaaaacagtaaaacatcaaacagtggtgtttgttgaataagacaaacagc ctctttctccctgtcacctctgtagagctagggtgtgttagacaagaagaagcatcagt ctttagactttatccattcacaattatactcaaaactcaaaagaagaacccaaagtttcaaga tactagggttaagatagaccagggtctagaactacaggcatagctttaatcatagaagga gagaagaagggtgtgttctaatttgggtcgaacacttagagcataggggtcatgctata gttagaagaactcatcttagaagacaaagaagaatcagaagaacagaaaaactag gtacagacaaccaagattctcaaacagggaacactcaaaacccaagggttggcttccctcc ttcttactttctagagttaaacattgttaaacactgtttaaggaattcagaagttctgtacg ttgaatcctgtactgtagaactgtgaagttgtacacacaaaagatgcaaaattctgagatt aaagggaattgagtagtaacaaggagaactacaagggttacgaagttaaagagtagtctcta gagaagtacaagaagactgttatttctgtgaaaaactgatgtaccactagaagtagaagac acattgtacctaatacacaaggcttaacaaaaatacaagaagaatcactggatagatcgag ctgtgttaggaaagtattctggaattgaagttttatttccaaaagggttaaacctataataa aaggctaattggaagaggccatcgtagattctgctctatgataaatagggttccctaagtc taaccttagacagagaagaacagatagaggagtagaattcaacgatacagtagagg gctactcacaagaaggtagtctttgtaggcagagtaggattgtctaacgaccaatttacaat gaagtagaacaagggtacatttagatttaagctagcgactgtaaaattctacacaaggag gatggatattttatagttttaa	attgattataatgatccaaaaatcacttacctt taaaaagggtcacttaccaggggcaagcc ccttgcttttgaccagactcagtgcatctag atgcactacgttagagagaatataaggtac tagtgggtgtcttccagcccatctgctctac gcttgattttaacagatattgggagatata gtttagcaacgttaactcttagcagtgagttc aagatataaaacctctcatttaacattgtcgag gagaacttaacagctcgtgtagatttagg aggaatcctacctgcatagaaaaacagtaa aacctcaaacagt
1649-1650	Ga0180732_1000047 GENOME_ACESSION: KR11_01MetaG_2_\$F_3 300020171 GENOME_ID: 30715 CONTIG_ID: 46 SOURCE: JGI DATE: 2019-01-27.	atgcagaagtctgtaccagtgtacgcgcgacagggaacattaatgccgactaataat tatcgcgcgaacgaactgataaaaaagaacagagcaaaacgccgtttcaaggctgggga tttttatcaagtttaattccagcgtactgatgtgtaacgcaggacattgtggtaggaaatcga tcctggcagcaagagagaagcgttcagggtgaaatcagaattacatacttacttaaacattt taagtaattctgtttcatgggtaaaagaagcggtagaacaaggagaaacgcgagaaaa gtacggagatataggacaactccatgcaggagaataaaaaacatagatctcaacaac ctttttagcaccagtagcaaaaagcacgttggcaggcggaagttaaggattgcagtcatttc gcgaagttgttcttataagacattttagtagaagatataaaagcacctacatttaagaac gctaaaagattggaatcacattttttccactcgaagtcggtaagaagtggttcataaaga acttgaaaaaattgcaaaagtagaataaaatctgtttacgatacataccaataagacag gatttaggacttataaagacagggtataaataacgccaatttttccgcgcacaattgtg actcttggtcattggcttgggttcagtcgaggtggtactgttccgaaaaacaaagaatt acaagatttaattcaattcagtttcacagggaacacttcatcttttaatttctcaaaaggc ggcataaagaagaattatggcagcaaatgtcacttggattaaaacgcggcagccttga acatcctaaatcaggtatttgtatgtcgggtggcagtagtaagatagattaaagtgtcat	gtcaataacccctatttaaaaataggggcttg cgaaagagaagaccatcatcttagcagggtt aaacaattgactacacagctagtaactgaatt ctctgagcgtgtagctcgaagaatgtaaa gaatgcttcccaagttctttctctctgagag tcagttggatgggagaaattgtagtaaggt taggtaacttaactataataaaggagtaat aaatgcagaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aatttagagacaggagaacgattagctcaaaacatagataagaaggatttagtattcttcttataaattactggaggacgacaattcctctgcacacataa	
1651-1652	Ga0256404_1003517 GENOME_ACESSION: RumenR/JG_01DNA_2_ \$F_3300026549 GENOME_ID: 34936 CONTIG_ID: 3516 SOURCE: JGI DATE: 2019-01-27.	atgccgaaccagctattctgtaataagcttctgatggcagcagcattgatgccaccttagctgtaaaagctagaaaactactaaaagccaaaaagctattatgtaagctacatgccgttcgccattcagcttacatataaagtggaacacacacagccgggtggagattgtatcgacacaggtctgggaacatacgggtttccgtataaatacagagaagcatgaatacttcatcgcaatttgacaacctcaaggatgagaacaacgacacgaagcaagatgatcgtcgacacagacgaacagattgcgttatcgaaaccagattcaacaatcgaggaaatccagaaggttggttagcaccactgtggagcataagaaaaaatacatccagattatcagcagcgctgttaggactcatccagtgaaagacatatactgttaggtcggacaattcgatacacatcttatgcaagctcttgagcaaggaataaagcttgaaggagaagattatcaacgtggattgaagtaattgtcttgcacacagagagaagcagattatcgtagagacaactatacttccaatgtgtgggaagagtgtaaaagattggagtaattctccgcgtccaccatatagtccatcgaagcaacgtgtgttagcgtattctgtagcaatctgtaacaatgtataaattgtcacactcgaagaaacacaaacccggcgggaattagatgtgtctaaagcccaaaattctctgaaagatgcaagcttcatgaatgtcgtgctgtgtatattgtcaatgagcttaagaaaaattcccaggaataggatccacacaaactttaggaattacacaaagcttcaagaagagctttaggtcagatagaaaattctacgcccaacgatgcataatgcaatgggcaactttcaccatcacacagaacacacgaatcagttatcaagaacatagaagaacacagagcttcaagtattctacgatgctaagtatatagattcacgcaccggagaagttaaaagcggaaagcgaattagcgtgtacaagaacaaatcgaagcattccaagaacaaatcgaacaacgagcgaatctcagaggagataaggtttctaaaggagaaggtttccataagaagaagcattacgaatacaaccaggggatataaacttacagaggaagaatacagcagtaggagttcattgtaacggcacaagaataaatttagagacaagcaatccattagataacggatttacaataaaagaaaagagggtggcaattcctctccgcctatag	gtcaaccaccaccacttaaatcctaaggaatcctgaagtggtgggttataagccaagccaggttgattagcctcagtggttcgagaactacgttatgcagtgatgcttaagcaacttaccaggaatgtccacaaagttcctggtctaaaggtatgtattaaacatctctgatggtaggagaaatgtgtcatacttcaaaatgcataacattggcgtatgtggaccacctgcatgtggcaggaatcagaaaagggcgtaagctctcctgggaaggtgtagaccagggcgaaagcctgaacataaagaaaggagcatcaaatgccgaaca
1653-1654	human gut metagenome genome assembly, contig: NODE_911_length_1802_3_cov_19.4281, whole genome shotgun sequence GENOME_ACESSION: OWAV01.1 GENOME_ID: 11837 CONTIG_ID: 910 SOURCE: WGS DATE: 2019-01-27.	atggaaaaacgccggcagcaccggcaggcatccagaagaggagaacaccttgcgaggaaacggctgtcaggagactggaaaccagatgggaataatcctggaaacggctgctcctggctgtgaaaaaccagttatggtaaggacatcatcaatcggagtcgccgattcaacaacccggcggcgaaggagattggctgacgccaacagcgaacacagctgttaagaacacatctgaacctgtgtgaaaaggtatgtcggatcttcccgtagcgggattgcatggaggccaacgggttgcattatggaactggaaagcggcggcattcgaatcggagtggaaccagtgtagacctttacggatactgcagattctgtgaagccctggaagagcttcaggaatgaagatgtctctgtcggcgaacgtgacattgaacatgaccatcacctggttctcaggtcataaggggtggaagcaacactctgccaatattgcccgttattgcgaaaatgtcatagctgtgtacatacgaaccaagaagctgcggagaactggaaagccatcaaggcaggagagcttctcattggaaaagcttctgcatag	gtcaactatccatccgctaagcgaatgggctgtgtaaaagcccagagactttacaggtgagcagttgagcagaacggtagcagcagcaggttcgcttatactccgggatgacgccaagtcggacactgatgtggctgcgccaggttatgacaacactcagggacattgttccctgttatagcaaggagatacaatgaactat
1655-1656	Ga0255345_1013787 GENOME_ACESSION: wastewater_treat_10_ \$F_3300028568 GENOME_ID: 281791 CONTIG_ID: 13786 SOURCE: JGI DATE: 2019-06-13.	atggtatattgataaaacgaacggacaacctctttagcctactaatagccacggtaaaagtaaaacacttactaaaagtggaacacacaggaagattaaaggtttaaaaaatcgtataaggttctgttggaaaggaatatattcaaaaggtcgtatgagcagcggatagctattcttatggatgctgatgtaacaagatagactttcagatatccgaaaggctataagacaccgaaattaagtaattcaagagaatttctgctctagaggatgtgataacaacaatcgcttag	atggaatagccctaacttatacgcctgtgacattgtgtaagacattgagttgcgtatcatcgtagccaatgcagtagtggttgaagcaggaagctccgacttcaaatcgtagcattaaagtcgagtagttcacgttaggtcagagtaattatagcctgtgatgtatgtgtaaaagtcctatcatcatgtcagagcaagaacgattctgtatatgtgaacgtatagatagtcgaaccactacataagtagtggcattatgtattaagttacatagaaggagtgccagttatg
1657-1658	Ga0376497_007558 GENOME_ACESSION: B210_RoWell1209_2_ \$F_3300035535 GENOME_ID: 304274 CONTIG_ID: 7557 SOURCE: JGI DATE: 2019-11-15.	ttgcatagagaactaaagaccaacggcgggatgcttctcagtcgccgtctcttgaagcccggttgtagaagccctgtgtcatccattgggacaacctacaggggtgtggagcaaacgggtccggcagacagccgtgtgcaacaatcccaggggagagttcaaccgatgccagggtgaaccgtcacccgcttgcgggtgaacggcgtgaagccgtgttttgcactcgacaagaacggtagccgttgatgccgtgccatccggccagagcaagaaagtctggcacgggcccgtgcccgggtccaccagatgtttccattaccatcagactcgttgaccggatccgggagaacagtgatgtgcagccggttaattgttaaaattgatccggagccaaaacaaccggcatggccgtgtgtgagacaggtgtggccatgcccaaatcctgcacctgtcagaactgacacaccggggcgctgtcatccggaagaaatggatcaagaatccaattaccggagaagcgagaacccgcaaacctctgtgtatcggaagaaaagattcgacaacagaacctggcccgaagggtggcttccgccaagtctcagatcacgggtggataacactttatcctgggtgaaaagtatcaaggatctgcccgtttacagcattgtcctggaacgggtccggttcgatactcagaagctgcagaacccggataattcagggatgagatgacgaaggcacactgttcgggtatgaggttaaggaataactgcttgaaaaattcaaccggacatgtgcttactgcaacggcttgaagcaaggatccgggttggagattgagcaatttcatccctgcaatccctcaaaaggggataaggatcaaacggatcagtaacctggccattgctgttaaaacctgtatcaggaatcaaaaagaacctgcagccggcagattgtagcagctttatccagggtcacagaaaagatagatcaggtc	gtcaactaccggcctgaaggccggagctgtgtaaaagcctccaggttgaccagggaatgttcaaaagttttatggcaagacaaaaaagagagaacattagaagttgcatagagaactaaagaccaacggcggatgcttctcagtcgccgtctcttgaagccggttcgttgaagcccggttgtagaagccttgtgtatgtgtgtagcagagattctgtatatgtgaacgtatagatagtcgaaccactacataagtagtggcattatgtattaagttacatagaaggagtgccagttatg

SEQ ID NO:	Name	IscB sequence	ωRNAs
1659-1660	Ga0247608_10001764 GENOME_ACESSION: 176DNAGHGlwgp2_2_\$F_3300028805 GENOME_ID: 279762 CONTIG_ID: 1763 SOURCE: JGI DATE: 2019-06-13.	atgctgaaggacggacgggcagtgatgtcttatgagccgttcaccatcaggctcaca tatgagcctgaacacaagtatgtcgagaagtaattctcggagtagacactggaagcaaa tatgtcggactgagtgctcgaagtgatgagaatgagttattctgtcgagaagcacagatac gttctgtcgaaatcaaggaaactccttgatgaaaagagaggttgcagaagacaaggaga acaagcaagctcgtttacagaaaacgggttcaataacagggtttcttcaagaaaccc ggatggctccaccgtcagtgactgcaaaagatgacgagccatgtacagtgatataaacttg ttgaaaggattctctatatcagagattgttagagaagggaaagttcgacacacacaaa ctgaagaacccggatgtatccggttctgactatcagcacggaagaagaatggtttgag aattcaaggcattcgtcaaatacagagacaattacaatgtgtatgctgcaagaagaat ccgcaaaagacagaaaacttgaagtgcatactatcttataaaatgatgggggttctgat gacccggacaacctggtgacattatgtaccagtgccatgagaaaatccacaagaataa actgaaacttccgagggaagttaatgcacgttcaagacagtcacaaatctcagggatgc agcatttatgaatgcaatgtcggaacaaactgtcgaacttctcgtgcggaattccacac gtcagtgtaaaaccacatacgggttactacgcaaaaacaaaagagaagaagtgagggtt gttgaacgacattcaaatgatgttggcaatcactggaactgtcagccaagccgtata gcccgttctcagaataaaaagagtacgcagacacacccgcaagataattcaatgtcaga atattcaagaaaagagtcaggaagaacacagagtgatgtcaagttcatcggagggttc ctccggtgggacaagggttcttaaaagaacccgtgttaccgttctataacaggacgg atgaagaaagaaaacatgcagtagtctgtaatatggacggagagaattacatgagaa aagcactattgctgtggaagattaaactgttaagcatgcaaacggtatgctttatgaag aaacaacatataa	gtcaccacccggcggtctgaagaacccg cggctgtgttccactcgggaaccgtagccg agtgaacagcctgagtgaaaggactgatgt ccttgaaactcgttaccagaatataataggc accgtcgggtgttctccagccttattgcact gcgtccggctgttaacaaaaactcttgcgg aagagtgtgcagccggaatataaaacacttg ggataacattggcgagggacaaccatagg attcaaatctacattgcaattatttctctaa attaaatcaagttgaaatgtac
1661-1662	human gut metagenome genome assembly, contig: scaffold82971_2, whole genome shotgun sequence GENOME_ACESSION: ULMK01.1 GENOME_ID: 15310 CONTIG_ID: 26480 SOURCE: WGS DATE: 2019-01-27.	atgtgccggcttaccgtttaaactatatacagaaagggaagtccttaccatggtatgt gttagatatcaatggacagccctcttatgcccagctgcaaggcatggaagtaagaaggct tttgaaacttcttgcacaaagggtggaacgctgttcatttaccagacagcttttctatcca aatacaaaagaaacacagccgttatcttcttggcgtagatgcaggaagtaaacatattggtc tggccgccacaaccgagaagaagtcatttaccagaggagctaatctcgcacagat gttgaaggttgggtgcagacgacgaactttacgcccgttctgaaagaagtcgaaagacga gatacagaagccaagattcaacaatcgtgtgcttccaaacataaaggtggttggctc catctgtagaagtcaaaattcaggagcatatcacagctatcaaaacgtctgcgaagattctt cctgtgtctgagattcacgtagaacagccgaattcgaatccagcgtctgaaagcaatgg aagaaggtgaagcctttgcccagttgtgacggttaccagcttggagaacagctacgacttca caacacgagacagtagtacttcatcgcagcggataccctgtcagtgctgtgtagcacaat gataaagatgtcaaaactgcatgtacatatacgaagccgtcagacagagaggaatgct ccgaacaatctggttaccatttgtgaacactgtcaataaggcattcatgaaggaagattga gcttccccaaaataaagaaaggcaaatcccatcgagatgccctttatgggtatcat gcgttaacacttctggaacgtctgaaaaagaagtagatgtccagcaacgatgacatat ggatatatcacaagtagctggagagaagggcagccttagagaaaagccatatcaatga tgccatctgcatcagtaagcatccatagccaacctttgatgctactatctgcagaaag ccgtacgacatcaaatgccagattcataaagccaacttcagtaaggtgtgatccgttaa acgaaatcaagcaccatacttctgtaaaaggatttgcctgtttgataaggtctcgtatcaaa gtaaaagatttctcttggaaagcagctacagccttcttgaataaagacgtggat ggaactaaagtaaacaaagggttctgtcagttataagaattgaggttcaggacacagca aaagcatatgtaaaagggtgcgagcaattcctcatgaatgaattcacgtgtctccttg catga	ctgtcaactaccgccataaattggcaggc ttgtaaaagcctaagttgactagcctaagtc ttcgagcactccgttaagagagaatataatg taccctgtggatgtaacacactagctacggct ctaaggagacacattaacccgttctgtgagg taggaaccgtgtgtgtctattgcaaacctct cattaacattggcgaaggtgtcatttaccgct cgatatgtccggcttaccgtttaaactat catacagaaggaaggtgccttaccatg
1663-1664	Ga0315282_10018675 GENOME_ACESSION: YL17G07_20_MG_2_\$F_3300032069 GENOME_ID: 281633 CONTIG_ID: 18674 SOURCE: JGI DATE: 2019-06-13.	atgcagaagttgaaggagaataacacacacacacaggttattccgaagatcgtgac aactgtggcttgcattaaacagagaggaaactctcagtggtggcaggtcctaaaacacttct ttaacaactccgatgcggacctaaactcgaagaaggcagaaacttgagagtatctgaca aagtttatgttttaacagaaagggaacagccgttaagtagtaagagaacgcctttactaccagtg attattgtgaaggaggaaaggcaaggttagtaagagaacgcctttactaccagtg gtctatgtcagagtgaaagcaaaacacgtgttactcttgatgtgtagcgggtataagca catcggcttatcagcagtaacagataaaaaagaactatttagttcagaagtgagttgaaga accgatattgttaattgtcgaaggtcgaagccagtcacgttagacacagacgaacagaga ttgtgtatagagaaccaagattccttaatcgtgaaggcctgaaggctgtgtagctccttc tatccagaacaaactggatgctacattaaaggttatcaaccaggttaagcaatactgcct gtatcgaagattatattgaagtagctacctttgatacgcgaagatggtcaatccagaatac agtggcgttgatgatcagcagggttaactcgaaggctatgaagtgagagaataatcttctg agaaatgggtaggaatgtgcgtatttgcgcaaaaagaatgttctcttgagatagagc atattattcctaatacaagggtggaactgatagtagataatttggcattagcttccatg agtgaaccagaagaaggacaacatgactgcagaggagtttgggcaccagaagttca aatgaaagcgttagagacgttaagccacagcctttatgaacatagttcgtggaactg gtgataagctgaggagtcaggaacatagttatgaacttatgatacattacaaga gtaacaggattgcattaaaaatcctaagctcacataaacgatgttctgtaattgacaggt ggaagttaatacaacagatcaacttcaatatttcatcaagcaggtaagggaagtgaaca ggagttattcaaggcaaacctgtttaaagggaagcaaaagaaaagtaatacagataaga gaagcgttggcttcatagatttgacaaggtagctgtataacgggatagaatgtctcatat	gtcaactaccctcacttaagaatgaggggtt gcatcgtgagatgcaacacaagagttgataa ggaggcatgacatgcagaagttgaaggaa gagaataacatacacacctcaggttattccg caagatcgtgaactgtggctgtcattaaa cagagaggaaactctcagttggcaggctc aaaaaccttcttaacactccgatgcggac ctaactccgaagaaggcagaacttgaga gtatctgacaaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ggattaagaagcaagggttatgttattgaagaaactgtggaactaagggtcatagttctgcaaaggcaaaagatgcactctgattgagtgacacacggttttaacagaaaggagaatggctctctcccccaaccttaaaagtgagggtctccgagccataa	
1665-1666	Ga0315288_10049882 GENOME_ACESSION: YL17G11_20_MG_2_\$F_3300031772 GENOME_ID: 281639 CONTIG_ID: 49881 SOURCE: JGI DATE: 2019-06-13.	gtggaaatgtaccagtagtagataaacagcaaaagccactatgacctgctccgaaaagagagcacggcaaatgttatatctagaaaggcgactccttttggaaaaaggggtctctgtattcgtctaaatgtggaaccttcagccagaatitgcaacagtagttgttggaaatgacctggatcaaaaaaagaggctgacggtaaatcgaagcacatacttatcttaattgttcaa gctgatgccattacgtgggttaaagaacatgtgaagctaagacgtataatgcgaaaagctcgggcgcatagaaaaaaccttatagagagtgctgaaataatgcactatgaattgtcttcaccttcaacaaaagcgcgttggaatggaaagtaagacttattacttggtttcagagatgtacccaattactgatttatagttgaggacgctaatacattacaaaaaagaattgctctaaatggaatggtaacttttctctataaaaaataggtaaaaattggtttacagtcataaaaaaatttggaaatctgagctctgtaagggttgggatacaaaaacttgcgagataaagggtgacattggaaagagttcaataaagatgagagactcctttgactctcattgtgttgattcatggttttggcaaatgaaattgcacaaggacacacaaaacctgaaaacaagagagttctttgtgtgttcctttgcgatttcacgtagacaactacatagtagtcccgtctaaagggggttaaaaggcc aagttatgggtgggacagaagattgagtttaaaacggtgtctttgtacgaacccctaaat atggcctgtcttatgttggggtagtgcgaaaaacagaattagttctcaattcttaagggtatgtaaacgtttaagtagtctcgtaaaggtaagtgattgcaagggtttatctttgggtgtactag agcctattgggctgaggtgtga	taaaggacacgcataacatctgaactcttta agtaagaagctagcccttctgaaaaaggca gagaaggtaaactttcaggattagccaag attaaaaatcaactagccatctagattggca actcaagaacacatgaagaatgctaccctag ttttcttccctgtgatggctagggcgaagg gttgtataaacaaggctgtgaaagacaagat gtaataaagagaaagtagtgaaatg
1667-1668	Ga0180435_10000153 GENOME_ACESSION: SS_ome_21_\$F_3300017992 GENOME_ID: 35203 CONTIG_ID: 152 SOURCE: JGI DATE: 2019-01-27.	atgcaaaagatttctgataatctatcgataaaaagccgttaatgcccgtgcgctcctgcgcgcgccagaaagctgctgctgctgctgctgctgcgcgcgtgtgtacgactatccattaccat tcaactagaacaggcgaccaccaacaatcgaacctgttcgactgaagctcgaccgccgctccaagaccacggggatttcatttagtgcatgttgcgaaatggcttccgcgtgtgttgg gcggcgaaattgtcgcacgcagggcactactgttaaaaagaactctgatagtcggccggtt atcgccgtgctgctgtagtcgcaacctcgcctaccgaaagtcacgcttttaaatcgcgg ttggcgataaatcaggatggctaccgcatcgtatgtcgcgtgtacataacgttgaaac ttgtcaaaagccctcaaaagctctcgtcactatcgcgggtgatgttgaaacggtacgttt gatacgcagctaatgcaaaacccggatattcagggtgttgagtagtgcagcgcggcgaact ggtcgattgggagttgcgccagctactgtttaccgacacaaagcactatgcgcttattgc ggccggttgaagcaacgacccaattctcagcgcgagcacattcatccagctgaaagg ttgaaagcaaccgactatccaatcaagtgattgtctgtcacacctgtaacgatacgaagg caatttactccccgagcagtggtgtctgttactcaatacatcaagcaaaaagattgacgtg ttgcgctgacaaactttaaagagattgttgcgggacagagaccaagcctgcgcgataccg gcgcagtaaacgacactgttacaatacggcgacgtgttgaatcgcattttgaccaggt cgaattctgttccggtggcgcaacaaagaaacgctctgatcagggttatcggaag atcactggatcgatgccgatgtgtcggcaccagcgggtggtccgtatttatccctgaaag cctgacacctctgctgattagggcgaggggcagggctacgccagttaccaagccga atgctcgggttttccaaagacatcgccaaggcgatcaccttttttgcggcttgaag acgggtgatctgttgaagcctcagtcggcactgtttaaaggctggcattcacgttgggc ggggtgctgtaaaaaaacggctcttttaattttcaacaaacaaagaccgtgcaggg aatcagtcacagatactgcgaaaagattcattgtcggatgtgtataactacaacttggag gcgcgcttctccccacggataa	gtcatctaccccacgggttaaaacgggggct tgaaaaagcccggatgaccagcgaaggc cagaacatttttcgaagatgttttgactactgt tggaaaagtcacgacacctggaatgcttgc cagttccaggctctgtcgtataagattaaaaa gcccgttaggggtattggcggtgtctttag cacgacaagcctttccaacattggcagagc aaacaaactcatctcggatggagaaccgg cgggtaaccgcctaaccttaaccaatggataa tgatgcaaaaga
1669-1670	Ga0070738_10016471 GENOME_ACESSION: CoaSoiStandDraft_9_\$F_3300005531 GENOME_ID: 24565 CONTIG_ID: 16470 SOURCE: JGI DATE: 2019-01-27.	gtgtcaaaagtcttggtagatgccaaacagcagccactttccccggtgcatccaggg cgagcacagtgctgtcttacgcggggaaaagcggtagtgttccgctaccgtttacc atccgcttgaaaaacgcgcgctttctccagagggtgctcctctgcgcttgaatcgcac aggcagcaagaccacgggaatcgccctgctcgcagagacgacaggagaagtggtctt cgcggcagaactgcgcacgcgggggagtgccattaaacaagccctgatcagcgtcgg gccgttctcgcaggacgcgcagacgcgcaaacgcgctaccgaaagccgcgctttgcc aatcggcgacggcggaagggaagtgctccctctctcgaagccgagtgagcaacgt cctgacctgggtccagcgtctcatgctctgtgtccatcacgcgcgtgagcctggaact ggtcaagttcgtatcagcagatggagatcccgagatcagcgggtgtgagtatcagca agggacgcttttgcgtatgaagtaagcaataatctgtggaataatggatcgggcctg tagctactgtggaagaaaggtgtccgctgcaagtggagcataatcaggccaaagtca atggaggcaccgatcgggtcagcaatctgtctgtgctgtgatggctgtaataaggcca aaggaaacgagataccgagcgtttcggccgaaaagccggagctgctgtgcgctctt ctggctcagccagggtctctcaaggatgctgccgtgtcaatacagccgatgggc gttgtacgaatgcttactggctctgggtgtgccattgagtggtcaggtgggtgacg cagttcaaccgaaggcagcgggattgccaagcaacactggattgacgcggcctgtgt gggcaagagtagtcccgttcaactcaagctggctgcagctcgtctgctcatcagcgc gaccgggcatggcagcagcagaagtgcaatgtcaacaagatcggttccctgttcca aagccaaaggagccagaagggtcaagggtaccagacaggggataccgttcgggtgt tctgacacagcagcacaaggtatctatgtcggacagtgcttgtgcgagccagtg ggtcctttgacatccgcagcgcagcgggtcgatccaggacatcagttatcgtctgtac	gtcatgaacccccgataaattgcgggggct ttgcagacagattgtctgttgcaagccc tgaacatgaccagactcagcagagttctcc ctltgggtgggaaggctaccgggtcctgttag gagcgaagcataggtacgtccgggtgcg agggccagcccggaccgtacggcaagca gttaaacaggtctagttgggttaagcagtgct gcttgcattgtgaaaccgtccataacgctgt cgaggccacctttacctgggaaaccagagg cccgcaagggtagaagcaaaaggaagcaat cgtgtcaaaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gcccattcaccgctcggatgctatagctatcaacaaggagtagcgtatgcacagcataa tcccgcccaataa	
1671-1672	Ga0114918_10012039 GENOME_ACESSION: LantaG_2_\$F_33000091 49 GENOME_ID: 31359 CONTIG_ID: 12038 SOURCE: JGI DATE: 2019-01-27.	atgcagaagttagggaagaaagcctaagaacgtacattgatacttcacacgtccatagc tctataagttcttcttaaacagagttcaagactcagtgagaagaacatgtgctgtccat aacaatcccgaagtgaatcaacctcaatgtaaaaggggcagaacctgagagttatcc aaagttatcatattgaatgctaggttaaccttcaatgcatgtatccgagaaaagcaaaa aaattattaaaagagggaagtagtattgtcaaaaagtttccattcacaatccaattgta attccaacaggggaaacaaaaaataacattaggaattgattctgggtatatgaata ttgggttagtgacgtatcagaaaagggaattatgtatctgtacagttaattatgatgaa agaacatcagaagaattacagaacgtaggatgtatagaaaaattagaaggagcaacct tggtagcggaaatcaagatttttaaatcgtccaaacaaaagactggcttccatctat acaaagaaaatgatgcacattggaatcttattaacagaattaagaagattatcctgttct gaaacaattattgaagtagcaaatctgatactcaaaaattatgaacctgatatttcagga acaggttatcaacaaggttcttatatgaatcagaatcagaagtattttaatgaccaga gaaaaaggaaaatgtcaactttgtgtaaggattttaaggacagtcagtcatacatca ctgtaaacaaagggaatgagaatggaaagtaagaccttaaaatctgtattttacataagg ctgtcataaaaaactgcacaagaaggattgaaactatcaacaaaaagctataaac aatcacatttatgtctatcatataaaagattttggcaggatgtatgatttacagataact atggatatgaacatttatcaaaagaaatgaattaaatttgaaaaaacctatagttactgat ctttgttaatagcgaagggtactattcaaacagatgcaatgaacaaattatacaaaaaa catagaataatagattttacaaatcaacagaatggattcaaaccttattaaaaagaaa aaatctaaagttaatccagaagattttgttgggttaatgaaaaacagtatttgaaggta tgtttaataaaggacagtatactctatggaagcacaaaaagaaagaataactttaatttt ctaaagtataaaaatatacatcaagggaattttgcatggaatatataa	gtcaaataccacgcactaaagtagtggctt gagtcgtgagactcaacgtgaaggttgatta gggggcttaaggaaatttatgcagaagttag ggaaaaagctaaagAACgtacatggata cttcacacgtccatagctctataagttctt aaacagagttcaaaagactcagtgagaagaa catgtgctgttccataacaatcccgaagtga atcaacctcaatgtaaaggagggcagaact gagagtatctacaaa
1673-1674	Ga0137391_10019969 GENOME_ACESSION: 15con2h24BmetaG_2_\$F_3300011270 GENOME_ID: 20763 CONTIG_ID: 19968 SOURCE: JGI DATE: 2019-01-27.	gtgtggtattggctgcgcgcaggtctgtgcagaagctgcagacgcgagccattcacgat ccgtttgaaatgtcaaccgccgagtagaccaaccgccgtcagcgtcggggtagatactg gttcacagacagctggtagctgcctaacgcaatgaaagaggtggctaccagatgcatt gtgtcgtcgcggcaactaccagttgtatacggggaagcgcagcgcagcataaagctcgg gcacccaagaagctgcacggttggagctgtacgaactggttggagcgaagggagtggtg ttggctatatgttggcaggcagtgaaaggcagcttggatgaaggatgtgaccaccg gaaaaacactgtgcgaggtgaccccgccgaacatttattcggctgtccggttcccgga gggtggatcatcacgaacaagaatgtcagagagaggaggcgcgttctctcccg tttaa	gtcaatgaccacccattgagggggtggt ggtgagaattcgtttctgcacgccgaattg accagactgtgctctcgagatcgaggagcc gagcaagtaggcgtctgctgatcaaggaa ggagctccgatagcgggaagtggtaagc accaccttcgggtctgccccagcctgaa gccctggaaactgtttcttaaccagtgtcgcg gtcagccaccgtgaaagcagacatcccg cccgctatcctgtcaggggactacttgac ccggaaggagacttatggga
1675-1676	JGI11643J12802_101273 70 GENOME_ACESSION: Grami_3300000033_\$F_3300000890 GENOME_ID: 27064 CONTIG_ID: 1825 SOURCE: JGI DATE: 2019-01-27.	gtgagagtgccgtgttttagacacacaaaaaacgctagcggcgacaacgccgcgtc cgctcgattgtctttgaaatcgggcaaggcggcggtgttctgctgttaccattcacgatca ttcttaagcgagagatcgaaaatccgacgaccccgatctaaagatcaaaatcgatccg gatcgaacacgacaggcgtcgaatcatcaatcagaaaacccgggaagttgccttggcc gccgagatagagcatcgccggcaggcgatcaaacctctctggactctgctcgtcgtc tcgtcggggcgccgtcagcgaaaaaacccggttacgcgaagcctcgttccgagaaccga actcggccgaaaggatggctcccgtctcgtcgaagccggatgagaacgtctacac ctggcgcgacagtacaacgtgtatattccgctgaaagcgtcgcgtatgagcttcttcg attcgatagcaattatgcagaacctacaatcgaagacgtcgaatcatcgaaggcgctc cttaaggcttgaactcgtgaatattgtgtgatcaattcaaccatagatgtgtttacgcc gacgcgaaaaagccctgcatgaggtcttgaacgtcgaccacctgatccctgttcacgc ggcggtatcgaatcagcgagcaatctcgtttgcgtgtcgaacacacacgaagagaa aggcgccctgtcactgaagaattcaaaattgctgtgctgcatgttctcaccgattaaag ctcaggcgaaggcgccctgaaagacgcccgctgtcaatgcgacaaggtggcgct cttcaatcgttgaagactcgtatctgcccgaattgaacgggatcggtgggttgaccaaa ttcaatcgaacaaagctggagggccggaaggcgattgggtcgacgccgatgcgttg gaaagagcagcctgagaactgaagatctgaacgttcagccgcttcggatcagagcc atggggcatggtcgcgtcagatagccgaaccgataagtcagcgttccaaaaacacat cgacacgcaaggcaattatcagggatttcaaacgggacatcgtgaaggctgacatt ccgaagggggaagtttgcggggccgcatgtcggcagattgtcggctgtacgcaacgac catcattcacattgaaggcttgcgttccaccgaagcatttaagcgaatccatcacag cgtatgggttaactataaaatgaaaactgtgggtga	atctgctatgccacagtttctgagaaaaagaca aaggcattgattcagcccgactaagtgattg actcggcagattatgaactacgatcttggc aatataacaatgtcaccttgggattgcgatg ccagtcgccaggtcgtgaggtcgtgtttaa cagttgatccgaacgcgaagacagtgacgt gagcaaaaaaacgcaaaagatcattgtcc aggcgcaattacgtgagcgatcacgtcaa actcgaagagtttaataacgtgagagtg cctgt
1677-1678	human gut metagenome genome assembly, contig: NODE_539 length_5461 2_cov_5.601316, whole genome shotgun sequence GENOME_ACESSION: ORVF01.1 GENOME_ID: 10838	atgagtgattcgtggcggcgtgaacggatgccgctgatgcctacatctgaaagagag gcccgtttattactgaacatgggaagctctgtttatcggaaggtccatttaccataaaa ctaaattataagacaggcagtagcacacagcgtgttatttggggatcgataccgatcgc agcatatcgaggtgtccgttctccggaagatgtaccgtgttataaaggaaagatcg gtctcagggtatccatgagtaaaagaaaactgtatggatcaaggctcttataaagagag gaagacgtatcgaaagaccagatcggccatccaaatggagactaaagccaaactgt gtttattacgaaactccagaccgaaaaggaaagacactggaagaagcaaaagatcacgtt cgcgtccaaacgaccgcagggtcgtctctctatcgtcgtcagtcgaagacgatcatc atatcagtgatcaagaagttacaggtatcttctcagatggatagctgttctgtatgaac	ttgataaccgagcatgcgaaaaagtagtgt gaagtcattcatgtgtcccgaaaggcagtg ctgactagactcagtgcaacggaacccatt ttgtgaggaataaccattagttactgtccca gatatcaatcccgaccggtagtatcccaa gccgggaacagggttcaggaacggatgt ctttcaggagaagaggacactgacgtggag cacgaagtatctggaactttgtcgaagggtg atcactccatccggaggggcagacttttg

SEQ ID NO:	Name	IscB sequence	ωRNAs
	CONTIG_ID: 538 SOURCE: WGS DATE: 2019-01-27.	ttggctgttccgatccggcaagaatgaaagatacagagatccatggagatctataccagaa aggaccacagtagtactatgaaatgttctgtctatgtcctcgtacgtgacagatatacttg tcaggtagtggaaagagggtggaaattacatataccatatacctgtacagaagtc gggtgcaccgataatccacagtagtggctacgagatgacgaggtgacacagag aacatcaaccgggaggcatcctttatcagtgatgcaggagcagaagaatttaccag aggactaagagatgctacattcatgaacatcttaaggaaacgtctgatgaaagcatttcca aaagcaacttttacctatgaaatatacacaagcagacagagagaactgaagattcct aaaagccatgaaatgatgccatgcgacgcgaatcgtaaaaactgggacatgacgggt aaaggataacgaacctgtgatctataccaacaggtccgaagaaaaaacgttctctgcat gaagaaactccgagaaaaggacgaaagaaccgaaccgttaaagcttcacgatacaata agaatacgaagcagtcaccgtgacaaaagaagaataaagaaagatacgtacact ggctgtctgttcgacgtgtgtagaactgaatggaaagaaaggtgacgtctgtgatttactga taagtcctgctatataaaggatgagaatgacaaatattcggatataccgaagtacaac aggtagtctgtcaaaactcaagatcctgcatcactgtggaactgggtgataggagcaa gaagatccctgtgaaggatga	tgtacctgccataacaatatataaggaggac cagccatga
1679-1680	marine metagenome genome assembly, contig: NODE_629_length_2182 2_cov_14.725040, whole genome shotgun sequence GENOME_ACESSION: ORDN01.1 GENOME_ID: 10380 CONTIG_ID: 628 SOURCE: WGS DATE: 2019-01-27.	atgccttgcgttcccgcaagagctaggcaggttgcfttaaaaaaggtaaggcatcggtatc cgtcgctttctttacgatcatctaaacgaacgggtagggtggaagaactcaacctaccga actaaagtatgaccaggtagcaaaaccactggcatcagctctgtgtgcatgtgtaaaac cgtagtctgttagtattggcgaggagaactaacatcgtatgtggagtattgttgggca ctaaatactgcagaacattaggagatttagcggtaccgttaaaacccgttatcgccaac cacgcttcttaaacccgactagaaggatggatggctaccgccatccattcagtcgcgtgt ggataatctgaccataacgcggaaattacaagtatttgcgccattatcatcattgtgtg cgaagtgtgtaagttcgatacccaaaaactcgcgaatcctgtaaatcgtgggtcagagta ccagcgaggtacgtgtgtgttcgagctatgggagttatttattgtaaatggaatcattca tgtgtctactgtatgctaaaaatcgcctagtagaaagagcatgtgttagctaaaagtaa aggcggttcaaacagtgtactaacctatgctgttcttgcgacctgcaacgaaagaa ggcaccataatcgatgacgaattttgaaagaaacactgcgttttaagaaataaaatc acagcttaaaagccactaaaagatgcggcgcaatgaatgcgaatcgctatcgcatag ccagagaattagaacattcaacttacctgttaattacggaacagggtgcaacaaaat aacgcacaaatcaaggctatcgcaaaacactggatagatgcgtctgtgcaaccag cacaggtgagaatatacaatactggatccatgacgccactcatcattaaagccgttggc cgtggctcaagaaaaatgtgccaacagatgggtatgttccctagctctaagccaaaag ctgtaaagaggattaatggaatccagactaatgatactgttaagctattgttacgaaagta aaaagaaagggtttatgttggtaaagtatcgattagaagcactgatacttaattataaaa caaaacaaggtaagatacaggccataaatcagatactgcaacctattacaagagatag atggctatgaataccacaccta	gttaactaccccgccctgaaggctgagctt gtaaaagctacgttttcgggattgttgacc agcctaaagttcttagaagactgttagatag gtcacgataccctagaatgttccacagttct aggctctatcgttaagcagttaaacaggacg gtgcgggaacaacgacagtgctgttgaatg taaaaagccttttaaccttggcgagtgagac cgttaccgcagaatcggtgagtcgcgtaa gcgacataaatattaacaaatgaggtat
1681-1682	Ga0063293_10084471 GENOME_ACESSION: IMG_3300003886_F3 300003886 GENOME_ID: 311638 CONTIG_ID: 84470 SOURCE: JGI DATE: 2020-04-07.	ttgagagtcctctgtgtaaatatgagtgaaagcctctgatgcctacaagtcgagaaagg caagaatacttttgaagcaaggaaagcaacagtagtgcagagaatgccctttaccatata gttaaaataccatcaggcgaaacaaagcaagatattacattggacataataatagatatt caacaattggcttcagtgacgaacgaagaaagcgaattattcagtgagttacccct gagaaaaaggcatctcaaaccccttgagcagaaaaaagaaattacagacgagttcgaagg aataaattattgcatagaaagccaagactcaataatctagcaaacccaaagctggcttg cgtcaggtattccataaactcaatacatctgagattgattgaaaaattgaaagaaagta ctaccataacgaagtgatagtagaattatcagacttgcatactcaaaagctacagaatc cagagattaaagatgagaataccagcaggaaggttgacggagacatgaagtaataagaa tacctctggagaagtggaagcataagtgcttactgtggaagataatcttctctgg agattgaacacatcatacaaaatcacgaggaggttctgataagatttctaatctgacttta gcgtgtcataaatgtaatcagaataaaggagataagacggcagcagagtttggacatcca gaaatacagaagggtggaccatacagaaaaaaacaggagtattctgttaa	gtcaattaccccggttaaaagggggtgtgt ccgtgaggagataagcaattgttgattag gagatattaaaaaatgtggaagtattgtgt agagttaaagaacacatcaaggcctcctct caagtctctctctataatctctctgacag tacttgccaaataaactctgctcaagacata cctctgagcaggttggaagaaactgaga g
1683-1684	Ga0074255_1053809 GENOME_ACESSION: LumHo_2088090016_F3 3300005361 GENOME_ID: 31619 CONTIG_ID: 2696 SOURCE: JGI DATE: 2019-01-27.	atgaataatggccgccgatagcggcctacgtgcaggccatcgacggtcaccgctgat gcccaccgcgaagaaaggcaagggtgcgccacctgctgaaagacggcctggccacgggt cgtgtgcagggtgcccgttcacgatcagggtgacctacgagaccacgaggtacgtccaga aggtgtcgtcgggtatgcacacgggtcgtcgggagcgtgcggcctgtcggccaccacgc gcgacggcgacggcgagcggcgaggcgaggcgagggtctctacgaggaggagct gacctcagggacgaggtcaagggaacatctccaccaggccgagctccgagggg gcgcaggtccaggaataccgctaccgcaggcgagggttcgacaacaggcgacgcc cgagggtggtatcccccttccatgtcggagaaggttgacagccacgtctccatgtgtt agagggtctccgccatactccgggtgtccagcataacgggtggaggtggggcagttcga catgcagctcataaggcgccaggaggagggccttcccccccgaggggaggggccta ccagtcaggccccatggcgggtgactggaacgttcgcgagtagctgtgtacaggagac ggccacaggtccggaaactgcggcgagggtcagggcagacaggatactccaggtgcac cacctggagttcccgcaagaccggagggaactccccgggcaacctgtcacctgtgca ggacctgccacgacgggtaccacggcgagggtggaggtgaagataaggcgcgca aggcccatggcccgccgggaccggcacgaacatcataggccagcgctatagacatc ctgaggagaggtacggggtgagaaggtccgcctcaccgcgggtacgtaccaag	ctcaagtttctaagtaactgtctcgcagaga gtaaagcggagccgcaagcgcgaaatttag gagcgggttagtccaagccccctggccct tgggggctacgttggatggaataggtac ccccggatgtgaaggagttccaaagtcgg ggcactacggcgagcgttaaaagccccg aaaggggcggtgccccctggcccaaaccc atccgcaacatcgacgtgggaccacaact cctcctccgggagaggttacaccgaaa gggtgtgtcgaccggacatccaccaggt ccaagaaaaaagaaagagatg aaaggacagaaaggagaaacatgaataat ggccggcgatagcg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		atgcccgcagcgcagcagcgcctggagaagaccactgcatcgaccccgtgcatat caggcaacccccggcgggccccgtagaggggacgatgtaccggcgcaggaaggtc gcgtgccacaccaggtcccttcacgtgatgaagaccggcaagaaagggcgcagggcgc agcaaggtgcacccacaagataggcagggcgccttcagaggtacacatggtga tgcacgaagggaaggagtgcttcctccggaagcacgggcggaaagggcgggtcga aagacatcgactggaaacagggtaccgaaaagcagtcctggaacgtaaaggaaatcaa gttcgtagaaggatgcgtaacgggataataacagaaggttaaatag	
1685-1686	Ga0209800_10005712 GENOME_ACESSION: HigcelmetaSPAdes_2_\$F _3300027800 GENOME_ID: 27174 CONTIG_ID: 5711 SOURCE: JGI DATE: 2019-01-27.	atggcggtctacgtgctggacaagaccggaaagccgtgatgcctgcagcgaaaagc gcgctcgctgctgtttggagcgcagtcgagcccgggtgcatcgctgctgctgctg atccggctcaccgaccggcgcccaagccgattgtaccctgcagccattggcgctcaagct cgatccgggcagccgggtcaccagatgcgctggtgctgcccggcagggcggcggggt ttcggtgctgctgctgctggagctgacccagggcgcgcgcatccgagccctgc tgcaacggggcgcgttccggcgccggcgccgagtcgcaacctgcgtaccgggcac cgcgcttcgacaaccgacccggcctgcaggtgctgctgcccagcctgcgcccacc ggctggacacagtgctggcctgggtgcacgcctcgggcgctgggcacccatcgcca gccttgccgtcgagcgggtcaagttcgacctgcaggctcgcagaaccagggaattgc gggatcgagatcagcagggcacgttcagggctacgaggtgcgcgagtaacctgctg agaaatggggcgcgtgctgctgctgctggaaccgagcaggtgcgctggaggtcga gcacatcgctccgcgcagcggcgccggcagcgaccgggtgctgaacctgacgctg ctgcccggacctgcacacagaacaagggaatcgccactcgaggtgttctggtgcacgg caacccctgcgcttggccagaatcatggcgagcccaaggcgccgctacgcgatgcg cgcgcggtcaatgccacccgcaacgccctgttgcctcgtcctgctgacggcctgct gggtggagggccggcagcgcggggcaaacgaagtcaaccgtgcgcccgtggcctgccc gaagagccatgccctcgatgctgctgctgctggcgaggtcgaggggtggcgggctgg cagcgccgacccctgctgcatcagggcaccggccggcgaggtaccagcgatcgcg cctgacgcgcgatggcttccgaggggctacctgaccggcgcaaacgccatttcggctt ccagaccggcgatcggttccatgcctgctgcccaccggcaagagggcggtgtccat cgtggccgcgtggcggtgctgcccagaccggcagtttcaacatccagacagcggcggt gtgtgtcaggggtatagggcatcgccatggccacctgtgcccagcgagccgatggctacgg ctattccctgcaatcgctgatagtgcccaactcgagaaggagggcagcgagacagggg cacgctga	gtcaactacccccggcctaaggccggagct tgtgaaagcaggtctgggtgaccagggaaa gcggtaatcaacccgctccgtttgcaacagg tcgcatcgaccaccccggaatgtctctca gttccggcgtctggaaggtcgggatcatgc gggcgaaaggtaaagcgccgaaggttccg atcgccgcaacttcgcccggagccggttggcg acattcccaggggagcgagccgtaaggc tcgctcaccagggccgtaaggcgatgtttg aaaggaggtatcgccatggcg
1687-1688	Ga0068305_10111221 GENOME_ACESSION: IMG_3300005083_\$F_3 300005083 GENOME_ID: 311641 CONTIG_ID: 12741 SOURCE: JGI DATE: 2020-04-07.	atgatttatgttttgaacaatacgggaaaccccttatgccaacacgaatgagggttacat aaggaaagctgttacacagaggcgatgccgtgctggtgcaacacgattgtttacggttaa gtgtttgaacgatactactaccatacggataaggttactcggcatcgaccccggaat tcacacatcggtatttcggcagtttgggtgataaagaacttatttcgggcacacttgaaca ggaagcgcgaggttaagaataatgcacacccatctgatgagcaaacgacttgcgtataagctga tgtatcgtcgtacagcggtcgcagactgtggcatagaaggcgagatggcaaacgct gtttcaacgaagcagaagggtgtgtgtagcaccatctctatgcaagaaagatacact tatctcgttaaacagattgaagaaggatttccgattgaccacatagtattgagcgcaac aagtttgacattgcgcgattgaaacccagatattcagggtgtacagtatcagcgtgta cgtgtacgattacgagaacaagaaggcgtacttgcgtgcgaagcaagaggcgattgtc cgatatgcggcgaaattctgaagtcgtacgttcacgtgcatcagctcaagccacgcgcaa aagggtggctggataattgttgcaaatgtgtgctcgtacataagcgtccataagcggttg catcgtgaaaagttaagctgtcaacgagcgggttcaagaagaagcacaagcagacac gtttatgaacatgctcagggcagggctgtgtgacgaattacagcgagcgtaacgtttggc agttacacgaaggttcgcggattgagaacggcatcgaaaagcgcactataacgacg ccttcgtgattgcgggtggaacgacgcaagaacgatccgccctgtgtgttctatcagaa gcgaagaacaacaggctcttgcagaggaaacacttcacacaaaaggcgaggtgca atccgtcgtcagcggttcaatttgcagacagggcgattgattggacagcgagcgggatgc acaggtgtggcggtatgacgggtggccgattgttctcaagatgaatacaaggaggga tgtaaatcccaattagaataagtaactaaataatacaaaagcatttgaagtcataaaagcatt tgactcttttcggataa	ctttccaaagaagaattgataaggtttgaa tttgattatccgaatgagtcacatgacttgg gcggcagtcgaattagagggcaactcgga ggagttgacgaggttgaagattgagatcat acacaccaaggggtgttcccaagccctct gctctgtgtgtataaggacgtaaactgaaa gagcgacaccgaaataacttaaaacatctt aaacaacctcgatgggaatcaaccgcttc ggcggtggcgctgaaaacaggaacaa
1689-1690	human oral metagenome genome assembly, contig: NODE_7457_length_375 0_cov_2.000000, whole genome shotgun sequence GENOME_ACESSION: UPKJ01.1 GENOME_ID: 16590 CONTIG_ID: 7456 SOURCE: WGS DATE: 2019-01-27.	atgttcgtatattgtttggacaagaatgggcaaccactatgccaacatcacgttttgcaaa gttcgtaggttactaagagataaaaaagtaaaagtagtaaaagcgtgttcatttacgattag gttgtgtacagcctgaaactaaaatgtacaagatgtgtgctggtgtagatactggttc taaacacgtgggtgtagcatgtgtaggttaataaggtattgtatcaatcaaaagttgaac taagggtgacattaaagaataatggattctcgagaatgtatcgaagaagtcgtagaa atagaagactagatacagaaaagcaagattttaaacggggcaattctattaagaagg taggcattgtctacgattgttagtaatactacggcgatgagcgagagattgattctgta agaagattatactgtcaagatacagtccttgaacaggtaaatttgacactcaattatg gaaaagcctgttgtaacaataatagtgggcctatcaagaggtgtgaactatgggtatg cgaatcgaggggaacatgcttagttagtagatagtatctgtcaatgtgtgtggcaaaaa gaactgtagagtagaggcacatcacatagttttagaagtaaaatggcagtaaatgatttag agaattatattacattatggaagattgcccaaaagcagttcatttaggggagattgattga aactgaatgtgtaaacgtaaaagtgatttacgacatgcaactcaaatgtctaccatcagaag	gtaaacatgcactaggttaaatcactagtacc ttctaggtacggtgtttacaagaccgtggt tagaaatagccactacgatagttggatgata cacatacacaccctgtgtgaatgctcaaga ctgaggttcgtgattgctgattaaagttggct aagagtgtgtaaatgctgtgttagtaattt caaaactctaatattcttgcgatgagaatgc cgacactattttgttaataagaataagtttag cattacactcagtggtggtactgtcttacaag taaaactacttaaaaaatcaatcagaaggag ttaagcgtatgttc

SEQ ID NO:	Name	IscB sequence	ωRNAs
		catgtttgtaaagaataccatgatgctattgagacfatgttttggacaaaagcaaatag ggaaaacttgaatatcgacaaaagaacattacttagatgct	
1691-1692	Ga0070770_10195732 GENOME_ACESSION: IMG_3300005078_SF_3 300005078 GENOME_ID: 27773 CONTIG_ID: 4481 SOURCE: JGI DATE: 2019-01-27.	gtgaacagcgtatttctactcgacaatacacaacacacactctccccgtttcacctgcac gggcacgtatgtctgttcgtaaggaagcagcgggtgtatcgccgctatccgttcacca ttatcctgaaacaccgtaccgggtgggaacgtgcatccccgtagaactaaaactgatccgg gcagcaagaccacagggcgttgcctgtgtgcaacagaaggcactggtgcagcgggtat cttcgcagccgaactcaacatcgaggcgcggccatcaagaaagcactcagcacccgt cgcccccagcgtatagccagaagaagccgcaataccgcgtaccgtgcgccgcgttca acaaccgcaccccccagcactatccaaagctcactgggtagatgcagcctgtgtcggg gattccggaaaaacagtaagactcgaccctaagcgaacccctgacaatcaaagccac gggtcgtggcagtcgtcagatgcgtgggtggataaattggctttccgcgg	gtcaacgaccaccaccacaaacctaccgggt tctgtgggggtgtgtgaggaagtcgga agttgaccagggcacagccaccagcgaactt tcgctgacggggctacgttagtgccgaata cataggcaccttgggtatgccaccagttcc aagctctcggttggcgattaaacaggctta gaggggttaagccagtgctgccgcataaa aaaccgccgataacattgccgagggtaac attaccgccgaagggcagaaagccgggtg tagaccggcactgattcaaaaggagtgac gtgaacagc
1693-1694	Thermogemmatispora aurantia A1-2 DNA, sequence09, whole genome shotgun sequence GENOME_ACESSION: GCA_008974285.1_AS M897428v1_genomic GENOME_ID: 312957 CONTIG_ID: 8 SOURCE: NCBI Prokaryotes DATE: 2020-04-22.	atgtttttgtgctggacaagagaagaagccgtcatgcttgcctggagaagcgagca cgattgtactggaaacgggacgcccgtgtccaccggctccatccgttcacgattagg ctcaagaccgggtactctgaggagagtgctgtgcagcccgtcgggtgaagctgatcc aggtagcaaggtcactgggtgtgcgggtgtgcggcggaagagagggcggtgtgtcc ttggatcatgcttgcacctggcggtgatcttaccgtggagatgccattcgggaacg gatgcagaacgagcggcacaccgcccgggaggaggaacgcaatctgcgtacc gcgccccaaagttcaacaatcgccggaggggcaaacggctgtgtccgccttcgttcga agctgtgttgggaacgtactctgttggggcagccgttaccggcgcttggcaccattgc ttttatcgagatcgagaccgtccggtttgacacgagggcgctccaaaaccggagatctc cgccgtgagtagcagcgtgtgcgagcttttcggctatgaagtcgagaaatctgctgga gaaatggggggccgctgtgtgtactgtgtgtgcggagaacgtgccactggagatcgag cacatcgttccgagatcgagaggcgggaactgatcgggttaagcaatctcacattggctgt cgagcctgcaaccaggccaagggggaccggagggcgaggaggttgggtcaccggga agtgcagcccaaggcaagggtgcccgtctcaaggagcccgccgctcaacacgacgc gctacgcgatccggaaaggctgtgtgcctatgggcagagaggttcgctatggacggg cgggcggacgaagtggaaacgggaacgggttggactcccgaagacgcatgctgttga cgcactgtgctgtgggcatctgcccgtgtgtccagctggcagctgtgttctgaaatc aaggcgtctgggacgcccggcagcgggtgcccggacgaacgtggacgccacggtttccc gcgcgatacaggatgcttccaagacgggtgcgcggatttgcacgggggagctgtgtg cgtgcggaaagtactgaagggaagcgcacggagtcatttggccctgtggccgtgtgc gcccctctgcatgttccgggtgggaaagcggatggatttctgtggagggtgtgtgcgc tgcctcagcgggcagatggatacgggtatatgaaggaggggcgccgcccctctcccc aggcgtaaacgcccgggcttccggcgcccatga	gtcatagacccacgccaaggcggggg cttgggggatgacccgaccgcccgtcta tgaccagcccaagcgggtgagccgaagga caaggcgaaggaagccgtacgtcggcgcc aggccaagaccactcgggatgcttccc cagtcgccgacttggagctcgggtgca gacaaccttgggggtgtgggcgaacggg ccggagcagatggcagccggcgacatg ggcgagggggagcgttccgggagagtcca ccggggcccgctcaccggcccgttaagggg gtcccgaaggaggaagagatcatg
1695-1696	Ga0207421_10015550 GENOME_ACESSION: SodLak8KL_SPAdes_SF 3300027784 GENOME_ID: 35608 CONTIG_ID: 15549 SOURCE: JGI DATE: 2019-01-27.	atggaaagagaatttatgttcaagacaaaatgttaaacctttagacccacaaatccag ccagagccagaaaattattagacaaagcgagagctgaagtagtacaaagagaaccttta ctatacaaatctgatagataaaaaaagaaattcttataactaaagatgaactttagggtgag atgcagggtataaagaagtaggggttttcagcaatgaatgaagagtaattagtggtg agtgttaaaactacgcaatgatatactaaaaagttagaacaaaagctaattatcgagaa atagaaggcatagaataactaggtatcgcaaacctcgtttgataatcgcaaaaaagaag aagggttggtagccctagttatgacataaagttaaatagccataaatttggtaatatagatt aaaagataattaccattacaagaatgaatgttgaagttgcaattttgacactcaaaaaat gcaaaatcccgaataagaagggttgtagtatacaaaagcgaatttcaggggtattagtt tcggaaattatctattagaaaaatttaattatcaatgtgcataattgtggaagaaataactctt tagaagttgagcatataaattcctaattctagaggaggttctaatagatgaagtaatttaactat tagttgccataagtgtatcaaaagaaaggtatgcaaacagctaaagaatttgggtatccta aagtgcacaacaagctaaaaagagtttaaaatcaacagcttttatgaatatagttagatgg gaaatagttataaatttgattgtgattatacttttggcatattacaaaaatgaatagaattta atttagatttagagaaaacacattataatgatcttttgaattgcaggaggtactaatcaaaa tcgagttaatcgaatagaagttacattatcgagaaaacatcttcattacagactaatc gcaaaacttatggtagaagtgttagaactcaaaaatattcattatcgcaaatgatttagtta aatatgaagggtgaagtttgaaggttaaggaacgtttaactatgttacttgggttagaatga aagatagtaaggttaacgttaattcaaatgttaagatgtagaattagttaaatatagcc aaggttatcattataagtttaa	tcaactactccacctgaatcaagctaaagatt cagatggagcttgaanaagctcataagttg agtaggagattagctaaagcgtatcttctct aaggtgaaggcaaatgagctacgtcctggg ctgagttgtgacataccctggagtgcttc tcagctcagggctctggaactatgtgtgga cgacctataccgtttatctgatgtatgcaa catcaccatagcttgcctgctccagaggtt gacgaggagaaccaccctcatttgagag atgacctttacaggtctgaanaattatgttta aggaaggttagtatggaagagaaa
1697-1698	Ga0180732_1000469 GENOME_ACESSION: KR11_01MetaG_2_SF_3 300020171 GENOME_ID: 30715 CONTIG_ID: 468 SOURCE: JGI DATE: 2019-01-27.	atgcagaagttattaaagagaatacatcacacctacagatacttctctgtctgtagctct gtggccgtgtctttaaacagagaggaaactctagtgagcattggcttaaaaacctttgata acaccctcgaagagaaatcgacctcagcatagagagggcgcaaacattatctccaaggtc tatgttaattagtgtaggtgtgtgagagctctaattccgacaaacccaagaagcgaggattt tgtaaatcgaggtaaggcgaagtgattagaaagccccgtttacgggtcaatttaacgttt caaaactacaggtttattcagccaataacacttggaaattgacagtggttcaaaaaacattgg gttttagacaataacggataataaggagttaatgagtgagaaagtttaactgatacatga tgagcaagagattgggtgatagggaattgtatagaaggcaaaagagaacaaactttgg tataggaaaccaaggtttataacaggggtgcagtaagaagaaggaggtgttacctctt	gtcaactcccaaacctaaaggattgggctt gttccgtgaggttcaacgcaagagctgatta gagggcgaagaacaaatggaataaaaaat gcagaagttattaaagagaatacatacaca cctacagatacttctctgtctgtagctctgtg gccgtgtctttaaacagagaggaactctca gtgagcatggcttaaaacctttgataacacc ctcgaagagaaatcgacctcagcatagagag gggcaaacattatctccaag

SEQ ID NO:	Name	IscB sequence	ωRNAs
		cgattcaagaaggtatgacacacattfaaatttagttaaaaaaattaaagggttggcctat tactaagggtatcatagaggtgggcaattttgatatacagaaattgaataatccgaaattga gagcaatggatataagaaggctcattgtaccaatcagaatgttagatctttattatttcg agagaaaagggttaattgcaattatgtggaaaggataagggaagtgatctggcgctct catcatattataacacaattaaaagggtgactgataacccaataattttgctttactacacc ttaatgccataaaaaattgcacaacaaggattagagaatcaattcaaaaagaataaaca ataaaagcatcgacatttatgaattattcaaaaataagtttcagcaagattggattgtgat atcacttttgggtataaaacctatgtggacaggtgtgaactgttttagaaaaatcacacgg caatgatgctttgtatagcgggtggaaacggacaggagagaattgacccgtttaaagtt atgcacaaaacgcaaaaacaatagggtttacaaaagaatcgacggggtttgctccagcg attcgtaaacaagatatcaattcaaccgaaggacttagttcaaatagccgggtgaatgtg cagagacaacaggtactcactgtaaaggctcaagaataatggttaataaaaagtcfaat atatcagattagtgaaaggtgtttcacacaggaactttaatatggagacaggaatttctc ccaaactcgaagggttttctgctga	
1699-1700	Ga0394881_0000594 GENOME_ACESSION: WSA_S1_MG_FD_SF_3 300037311 GENOME_ID: 312717 CONTIG_ID: 593 SOURCE: JGI DATE: 2020-04-07.	atgcacaaaagtattgtgtggacacaaacacagacgcccttatgccctgccatcccgtc ggcgcgagagattattgacattggggaaggcatcggtgttcaaacgatttccatttacgatt atttcaagaatcgtgaagggtggcgaactgcacatctgctcgaatgaccccgcc agttaagaacacagcattgtaattgtgggggattcaacgtggcaaacgggtgattgg gcgggtgtgctagagcctcgtggacacaaagtaagaaagcgtatagaaatccgtcgac tttaagacgtggcgtagggcagtaaatgccactgttggcgattgtacaattgtctgaaa aattcgggattgcctatcgaaatcggtacagggcgacgaagttaaccgcactag gcagggtaccacaaaagccattggaattgacgaattgtatcggtgagagtgccatctg tgtttatctgagccgaatatgtcttactgtcacatcaaacgaacgggtcatggttcgctca aatgtgtggcaccgacaagtttgctttccaatcgccatcggttacgcaaaaacgccatt ttgatttcaaacaggtgatattgttaagccgtgtcatcaaggcaagtattgtaggtgttc atgttggacgggttcgtgtcgggcaacaggtattttgatacacgaccaagcaggta aagtaactgttaattacaaaattgtatcatctgcatcgcgagatggctatagctacgct tatcaagaagaggtggcattcctctgtagctaa	gtaagcaatccatcgccaaagcggaagacc ctttaataagccctcatgcttaccagattcagcc cattcgggctacgttaagggtgaatacatag gtacttgggatacttaccagctccgaactc tacggtacagattaaacaggcaaaatgga atatgcccgtgctgtacataaaccactcac taacattatcgaggtgaacattacatccgcaa ggacaggaaaaatgaggttaactcatgcaa aaa
1701-1702	Ga0209591_10074768 GENOME_ACESSION: MAT01SPAdes_SF_3300 027850 GENOME_ID: 31815 CONTIG_ID: 74767 SOURCE: JGI DATE: 2019-01-27.	atgcttaaaccaattacgttttctgtagatcccgcaaaatcccactcacgccatgctg ccctctgtagcatctcgttattggatgcagggaaggcagcggtttcagaatgtttccgttc accatcattctcaaaaaggaggtgaattgttacactgaacctattgaactaaaaatagacc caggctcaaaaacaacaggactagcaaatcctgcttgggtccaaagtcattcttgcggctg aactaacccatcgggggagagcaatcagcgcgttttgaatcaaggcgagcatcaag aggtagccgtcgggcacgcatcccgatcgccagccgcgatttccgaaccggacg cgccctaaaggctggttggctcctagcctccaaatcggttcttaactacctgagctgg gtggagaaattccgcgtctatcgccattgtgctaacatcactcaggagttatgtagcgtttg atttgaacagattgagaacctgaaattgtcggggtcgaatatcagcaaggtgaactcg ctggtacgaagtcagagaataattgtctcaaaaatgggataagaaatgtgcttactgtga cgcccaaaacctacccttcagggttgagcatattcatcctaataaaggaggaggagcaa ccggatcttaattctgctctggttgcgagaagtgcaacctcaaaaagggaactaaga tattaaaggatttctcaaaaagaacctgaactactgaaagcatttctgctcaagcacaac ggccatta	gtcagataccagccctaaaggactgagctt gtagactgcccactgagcgagcggtat aaaccaagctaatctgaccagcctaagctt cacgggttcgactggctcaccgttaactggc tacgtttttgagtcacgacaccttgaatcaa gctagtccctgctctgcatctgcaattaaac agtcttaagggtcactgagacaggtgttcagg tctaacaagctcttaaaacattggcggaagca aacattatccctacagaaactaagcgtcat aatgcttaaaccaattac
1703-1704	Lactobacillus mucosae strain CRL573 Contig_36, whole genome shotgun sequence GENOME_ACESSION: GCA_000766905.1_AS M76690v1_genomic GENOME_ID: 67659 CONTIG_ID: 35 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgatgcacaaaacaaaagaattatgtctacgtagttgacccaacggcacaccc gtcatgccgacttcacgttaggtatgtgcggcgtggttaagtcgggccaagcggctc tggtagcgaataatgtctgactgactacagtttgcctcgtcgtttagatagcagaccaac cgctgacgctggcgctcaacaccgggttccacttaggtatggcagtcgttggcaatgatc gcgaataactatgcttcagaagaattcgtgaatgcggaaaaggatcgatcacggccgc cgcgagtatcgacggacacgacgcaaccgctgcttgcgtgaagccacgatttaacat cggcgcgcaaaagcaggttgctcgcggctcgttgcgttcagcatcggtcgtgatttgcgtt tcgagagattcagcgcttgcacgggttgcgcatcagtcagtcgtgctgaggttagtc ccttcgacaacaaaacttgcgaatcccaacatcaagccgtgggaatacacgcaagga aaaatgcacggctatcaacagtttaaggtattctgctggccgcgatcataaccgtgatg ccctggatgtcaagctacccggcctcgacgtgcgggtgcaccacctgttcagcgc aaagacgggtggcacgaatcagcccgaacacctggttttggtagtgacgtgcatcacaat caagccaatcacgttaacgggacgctggccaagctggcagccaaccgacaacgaac atcgactatcggggagcttactttatgacgctgtgccagccgtttgaatgcttatttctg actacgtgcagacgcagggtctatcaacggctaactcgtaccgttacgggattgtga	agcagtcaacagaaccactactaaagtag cgggcttgcgtgcaaacgtaagcatggt tgattagccctagtcatttggatggctacgt taccgctgaatagataggtactctggaatgc cacccatgttccagactctacggtagccatt aaacagagccaaggttccggctcagtggtg ctgatatacaaaaccagcgataaccttggc gaagggtactcagtcgtgcttttgaccag atgcgacttattgattcaattgattctaga aaggagtggtcatgatgcacaaaacaaaaca agtttaattat
1705-1706	Ga0256829_1004322 GENOME_ACESSION: 131447metaG_2_SF_330 0026518 GENOME_ID: 20587 CONTIG_ID: 4321 SOURCE: JGI DATE: 2019-01-27.	atgaaccgagtaatcgtactcgatacgaataagcaacaactgacaccttgcattccagca agagctagagagttacttaggaatacaacagcgtcagtcctatcgtaacaacccgttcacg atcatactaaacatactgtatgtatgtacacacttagatagacgaactcaaatagata ccccgtagtaaaactacaggagtcggctgtacaacacaccaccaatggttctaaagt agttatgggtatcaactacaacagaggacacatacaacacaaactactagctag aagacagttacgtattctagacgtatgagaacttaagacatagaccagctagggtataat aaccggagaagatgctcagactgttaccctctctatcaatctagactagataaacgtac gtacagtagtcaatcgttactaagaaccatacctataactgctataactatggaagatac	gtcaactaccactacgtgaaggagtggtt gtaataatgctacgacctaaccataggtgaag acttgaatatagcttaagtttccagctgaag ctttagaggtactacgttaggatagaatatata tgactcccccggtggcgttccagccag ggctctacgggtatagtttaaacaggctgtg tgggaaggacaggtctatatacttacacctta tccataaccttgacgaggagcacataacagt

SEQ ID NO:	Name	IscB sequence	ωRNAs
		aagttcgtaccagctaatgactaacaataacatatacggaaaagagtagaccagaacggtaccttagtaggaacagagatcaaaagagatctcttagccgctaataaaggtagctgtagtactgtcacggtataaataatgataatcggttagaggtagaacacgtacacccataatcaagagggtgtagtaactccgttactaacctaacgttagcgtgtaggacgtgtaacatagctaaaagtaacatgctgctccagagtggaagcacactggggctctctggtaaaccgttagatataatcgtatctgttagatataaactaagactcaacaagactgttaaactcaaatgctacagttatgaatgtgctacgtaagcatcttaaacatctcttagagtcgtacgggtaccagtagaatagccttcggttacttactaaataacatagaactaatcagaacacagaaaagatcactggatagatgctgctgctgtggttactactggtactaacgtataatccataagaacaacaaatgctaacatgctgctgctatgcaaaataacaacagacagatgtgcttagtgaataatgggttccctagaactagacctaaggctcctagtaattgctgtggttactgtaggagatctagtaacgctgtggtacttagaggtatagtagtatacaggaagaataatccgttagaacacggttacttcggtatataaaatagattgtaactataatataatgaagttactacaagctaatgataactataaatactcttatggcacactacactacataatcatacatagagagaacaaaacctcttaaacatgttcattatagtaggcgaacacactacactatccaaagatggttaagcatgggtcaacaaggtaacactgtactgtatcttaaacgggttactatagcggcatag	tatcctagtggttaactagaacatagtaaggaa
1707-1708	Archaeon IPdc08 DNA, NODE_37_36, whole genome shotgun sequence GENOME_ACESSION: GCA_006538205.1_AS M653820v1_genomic GENOME_ID: 289565 CONTIG_ID: 35 SOURCE: JGI NCBI_Prokaryotes DATE: 2019-07-22.	atgcgtggaaaaccattaatcctacacgacctggaaggcaagaatacttttcagcagaaggaaaagcaatagtaattcagcgagcgccattacttactcagcacaacttgagacacatcagagattgattgaaaaactgaagaaatactacgtgaacaaagataggagtagaagtagcagattttgatatacaaaaaaccagaatccagagattaaaggaatagagtagaccagcaggagaggttacagggttacgaagtaaggaaataatgctggagaagtggaagcagaatgctgctgactgcggcaagattatcttctctggagattgaacatattatatacctaaattacgaggtggaacagatagagtttcaactcgtgttagcgtgtcctataaattgaatcagaagaagggaatcaaacagcagcagagtttgacatcctgaaatacatcagaagaagcaagcaagcaattaaagaaactgcgtttatgaatattgtagatggagactggtgaatactctaaagtgggttggacctgtgggtatattacaagcatgctagaattaaattag	accaatcaccacttattaaatgatgctgctgctgtagagataagggaacaggctgatgatggggcggttaaagatgcagaagtattgtagagtttcagaaacacaccagtagtagatactcctcaagtcctcgtctgtaagttagacataaaacagagataaaaagctcaggtgtacccctgtatagtagtggccaataacagctccgaaggacttacactctggcaagagtgagacaggcgttgagcg
1709-1710	Ga0209204_1004127 GENOME_ACESSION: AD_des_37_SF_3300025 631 GENOME_ID: 21336 CONTIG_ID: 4126 SOURCE: JGI DATE: 2019-01-27.	atggtttatgttttattatctattgaaggcaaacctttaaagcccaactaaaaggtagtgaaggtagacatcttctaaacaagaactgtgtaaaaggttaaaggttaaaccctttacttactcagttactttatgaactactaactatacacaagacattacacttggtagatagatggctataattataggttttctgctgaacagagaaagaagaactaatctctggtgaaggttaagctaaagaatgatatacctgaacttatcaaaagaaaagcaaatgtatcgtagaattagacgaagtaaaactaagatagaacacctaatttaatacagaacatcatctaaanaaggaaggtggttagcacctctatataacataagtttagattctacattagattatagaaataccttaaaaagatactccctataaataatatagtatagagtagtactaattttgatacacataaactaaaaatcctaactgtagaaggtgaaggttatcagcagggaagaacaagggctttataatctctgtagtatattctttatagagatgactatactgtcaattatgtggtaaaacaaatgttccacttgaaagtagaccatatagcttttggaaaggagatagacagatagaccaagcaacctaataacactatgactaaatgccaatactcctaaaaatcataaagaagtggaacactatatgtatgaacactatacaaaaactcactgaagggaagcagatttatgtcaactgttaaggtgaaattagttatatacattaatgtgtaactatacttatgttattacaaggtctaaaagaataagccttaacttagaaaaactcattataatgatgcgttttgcatagcaggtggaactgttcaacaagagtagaacctatatattttgaacaagtgagacgaataatcgttctttgaaaagtttcatgctaaatatatagatataagagacaactctatataaacagggaagaactatttaacggtagacgaacacgcaacaaaaactacaatatagaaaacttaagactgtttagggggcaagagatcttaaaagtagaagaaatattcgtacacaaaggtatttttatcaacctaagatagttatattgaaggtaacaagtatattgctaagggaatacaaaaagggttagttatattaaacttgaaaactgcctaaacctgttaaacagaattagttaaaccttatatgtttgaaaaggattatgtattctgtag	gtcaactaccaccgcctatagaggcgggggctttagaaaatacaagctcgggtgattagccacgccaccagcgatagctgacggggctacgttatatgggaacatatagtcacataggtatgctcccaagctctatgctgagggtagtggttaaacatctctgaggggtaggagaagtgctgctactatacaaaaacccatataacattggcgtgtggacctacaggcttttgcctgaattacccttaaggagggttaatatg
1711-1712	Ga0102924_1014628 GENOME_ACESSION: PaiPotIIAssembly_SF_33 00007982 GENOME_ID: 34130 CONTIG_ID: 14627 SOURCE: JGI DATE: 2019-01-27.	atgaatgtatgtatgtactttaccagacaaggtgccactcatgcttgcctgctgcgcgatgctcgtctgttgcctcaacaaggaaaagcaaaaggccgccgaaccccttttaccatcaactgcatactcagccagagagcacctacaccaacccctgacgcttggatggacacaggcagtgccggtgatgggttcggcagtggtgatgaacagggaacatattgtatctctgaagtagaagtgccgaacgatacgcaccacgctcaaaagacgggcttctatcgcgcaatcgtcgtcaacggaaaaccgctaccgcaagcccgatggctaatcgcaagaattccatcaagacaggcacttttccccacgatgacgagcaaaatcgacgcgcacctgcga gaaattcgtttgtccagtcgtgttaccgattcacctaatgctggaacagggacatttgacctcatgcttgaacatcccaggtcttgcagaacaaatggatgaccagaaggcatcaactacggttttccaataccaagcctatgtgtcacacagagatggctacacctgcagcactgcaagggaaggaagcaacagatcggcggttagaggtccatcacatattttccgcagcgaacggggaagcgatgaggaagcgaattttgtcaccctctgaagacctgccagatgccctccatgcaaggagcatcagctcaacaaaaaggcgagagaaaaggaaatcttctacacgacgagatgaacagtatccgcatccagttgctccggcgcttgcggcagaggaaaacctggggctttgtgaccaaaagaaacccgcttggcaaaacttcccaaagcc	gggagccaccgaggtcttcagccctgtagcttcttcggagcctgagttaccagactcctttcagaaaatgggaagaccgttcgagtggttatgatactcgcggtgacgcacagactgccgctctatcgcctgtagttaagtagaactgagggaaggttcgggtgctacaggtgcaaaaagctgcttgaacactgtcgagatgaagtcggattcttctgctacaggggaggtacgcaccacccggcgcttgcggagcatttttccgaaggaaatgcgcgtatgaatga

SEQ ID NO:	Name	IscB sequence	ωRNAs
		catgtgttgatgcagcgggtatcgccacacggggaacgatcccaacctttcacacgaca atcatgcttgcgaagagattcaaaaacctgctggtttctgattctctctgttgaatgg tga	
1713-1714	Draft_10001674 GENOME_ACESSION: IMG_3300001592_\$F_3 300001592 GENOME_ID: 27523 CONTIG_ID: 1673 SOURCE: JGI DATE: 2019-01-27.	gtgcaagataaagctacgtagtttcaatgaacaaaaagccactaatgccaactacaaact ttgggtacgttagaattttatgaacggaaaatggctaaagtcatactatcaaacggcttta ccgtgcagttacttttgaaccggaagaaaatgtgtaccgactgtgttaggacttgatccc ggctgggtgaatctagctattactgtattagaaaaaagatctggtaatttttaattctcgg agtgtgtaaccggtataaaaaagttgttccaatctaataaggaagcgaacggcaccgaa tgattcataggggctgccgaagagccaaaaaaatagtagagctaacaattctaagacg gtctttaaaaatgccaaaacatcatacaatccggcgtttgaaccatcaaggtaacgt atattaagccaaaaccagtaggttttaaataggactagaaaaaaaatggcttactcca acggccaatcatcttttgaaacgcaccttcattatgtaaatcttatatgcagaatcatacc gtaagtgaatagtcattgagtagcagaagttgatacagaaaatgaacaatctgcaat tcaaggaaacgagtagtcaagaaggactgttaagggtatgtagtgcattgattatattt atcaagcacaagacaagcattgttgatgtgcaaaaacaagattaaacatgatcatatatt gttctcagcatgaagcggtagtgaccatccggataacatagcgggtcttttgaagaaat gtcatactaaagctcatcacagataagacgtttgtgtgtaaaatgaagcgaaggaagg atcaaaaaatcctatgtagtgcggaattcttaattccatcatgccaaagtagtactcaaa caattcaagaattgtatgggacagaaaatgtaccataacgacgggtgatcgaacaagtc aggcgaaggaagcgtatggcctcaagaaaatgactgtaacgatactgctgatttggccc tgtcaaatgtgatcatgtagtctatagcgcgataatttgccttatcattgcatgcaataccg gaggcataatcggaaagtgtgtagctgtgaagagaacgtcattataaaaaatagacaag aagttagtgcacccaacagaacgaacgatacagcgaaggaacaaattctttatttgagt ataaaaagcaactcaagaagcgggactgactaaaaagaaaatcaagcaatctatfca aatftaacgtgtggccatctaaagaaacggcttaaacgcccaacaaaaaacctgtcatta atgcaggatgcaccgttcttataggtctaaacgatatgtggttaatgtagctttaataaag gcataacattgaaattccacgggtttgataaagaacgttctctattgaaaaatgcaaatat tacaaggaaatcagggttagtatgtagttaa	gtcaataacccctcattaaaatgagaggcat gttagattagcactttccgcaaggaaagaac tttacataatgttatcaagtgctgtgtagt tattgaacagggaagagagatttacaagaca ccttaaaaataccactccagtttaagccctgt gatactgtatatactgtcccaaggagactca ccgacttaactgtttgccaacaacaggaa gggtcgtgtttacaaaattacgaaaggaatt catataagtccaagataaa
1715-1716	Ga0306921_10158369 GENOME_ACESSION: sta12C44000080v2_\$F_3 300031912 GENOME_ID: 281678 CONTIG_ID: 158368 SOURCE: JGI DATE: 2019-06-13.	atgccacgcggccagagctcagagaattggtaggaaaggcaagcagtcgttgcgtcgt cgctacccgtttgttattcgcctaaaaaaggatcattcggggcggcagccaatagccatca aaatcgatcccgagcgaacacaaactgggagtcgcgcgtgctcgattagcttcatctactca ggctcgttttgcacgttccgagttgacgcacgggggggagtcataaggagagatgctcaata aacggcgtgcatccaggaaaagtgcgcgtgctcgcaaaaacccgttatcgtgctcgag gtttaataaccgcatcaggcacgaagcgtgtgtgacaccaagtcgcatcgcgtgtaga caatgtgatttcagggcagcgcgtaccagctgtggtgacggattagcgggatcgtcata gaggccgttctgttcgacatcaaaaactccttaatccggggatattcaacgcggaatata gacaaggcacattgtttcgtacgagctaagagagatctcctcgaaaaattccagcggac ttgcgttactgtggaagcgagaacgtccgcgttgaattgatcatgtgcagccgcgtc caaaaggcgaaccatgaaccttaaacctcgtactggctggccacgggttgaaccggg ctaaatggagccagccgattgaagactccttgcaataatcctgaacgcctaaagcggga tcaagagccagcttcagacctcgttagagcgactgcggcgataaacgccactcgcaca aaaattctgctggaactttcaggatcatctgcgggttgaggtatcgacgggaggtgaaa ctaagtttaacgaactcggcttagcatccgaaaagccacgattagatgctgcatgcac ggcgaacacaggaattcgtcgtgaggaatgtctgtgctcgtgtaaggcctgtgg caggggctcgtatcagagaactcctcagacaagtagtgggttccgagaggtttttgatcc ggaagaaaaaaccaaggcttcag	gtcagctaccacgactaacatcgtgggctt gtgaaatcagagccgtgtgtgacagacc aagaagcgaaaaaatgcgactacgttataaa tacaagtcagaagaccactccggagtgcttc ctcagctccggaccctggaaagtcgggttg acgttacgggaaggttaagccggcagc accagacgggtcccgagatcgagatca gactgttttcaacatggtcgaggggagcgt ccaaaaggactgtcacttctcagcaacga gaagagaaaagcgggtaaccgc
1717-1718	Ga0103869_10001850 GENOME_ACESSION: 3300009257 GENOME_ID: 238854 CONTIG_ID: 2212 SOURCE: JGI Dark Harvest DATE: 2019-01-27.	atgcaagagcgttgcgtgtagacaagaaaaaacacgttgatgccctgccatccagc aaggggcgcgagaactattacgtgaaggggaaagcgcgggtcttcaagcaataaccttttac gattattctcagaacgggaagcaggcaggtacagaatgtcgtttcaagatagatcc gggtagtaagaaaacagggaattgttagcggcagatttcaaacgtgtaagcgtggtatt tgggcagcgggtactgaacatcgttgacagcaaataaagccgacgttgaacccgtc taacctacgagtagccgccgcgtcgacataccggttatcgtccagcagcgtttttgaac cgacatacgggcaagggcatttaccgccgtcgtcgaatcgcgttattgagaatgtctgg acgtggctatgtctatccatcgtgtatgccgattagtagcatctcacaggaactcgtcg cttgatacccaactcatgcaaaacactgagatttccgggtgtgagtatcagcaaggaacg ctgctcggtaggaagtcgcgaataactattcgaaaatggggacgcaaatgtgcctact gtgtgaacaagatgtgccactggagatcgaacacatcacaccaagtctcgtgcgcgc tcgaaccgtgtcagtaacctaataagcgtgtcacaattgcaatcaagaaaaggctcc caaacgcagtcgaatttgcgtatcgtgaattcaagcaaatccaacaaccg	gtaagcaacccgcacctacaggtggggg cttttgaagcccaatgctaccagactaagcg aaagctacgttaggggcgaatcatagagca cttcggaatgctcaccagttccgaattctgc gggtgaggtgttaaacgggacgttggaggt aagccagtgccattcacacaaacccgtcc ataacattgtcaggtgaaattatctcccga agggaagattttgaggtactaatgcaaa ga
1719-1720	Ga0114939_10008276 GENOME_ACESSION: AshMeaCrysSpring_2_\$ F_3300009455 GENOME_ID: 23357 CONTIG_ID: 8275	atgaaaagtaaacgagacagagacaaacactcagaagtacttccctgaccttgccaa gtctattgaagatttctggctgatgctatacctcttgagggttagggccgaaagcaatttc gaacgacacagcttctctttttatctcaatggcggaaggagcgtctgaaaaagaaagagaat gagaagaggaggagaaatcgaagatctatctcaatctggacctgaatctggaataaa ggacaacacaaatgagattatatacaaatcaattattcttcgagacagaatccgaagaa acagtgctgtatttggaccagataagaaaaagattaaaaatacagaatatacgaagcga gataatcgaagcga	gtcaactaccccgccctgaaggcgggctt gttgtgagactcaggactgcgaagataacc aaaaaggtgaataagttagactagcctacta gaggtgaaataagtcaatgaaaagtaaacg agacaggagacaacactcagaagtacttc cctagcctctggcaagctattgaagattct

SEQ ID NO:	Name	IscB sequence	ωRNAs
	SOURCE: JGI DATE: 2019-01-27.	agaaatggataaaacaaaaggaagcaagcaggtcaggaataatctagacttgaga atftcgatagacttattgaacaaacctgttgtaagataagagttccggtatgataaag atggaaagccattaatccaacgtgtgctggcgaggaagagatggattgagaagg gaaagctgtcggtaagagaactaagacaggaatatttattgctcatttggtagaacctt ctggaagaataagcaagatattgtagcaacattagaccaggaagcagatatactgga gtagcagtatgttcaagaacagggtctatattggtatctagaactatagctgatgaa aaagaaaatagatttgcattatcaaaaatagaatggataagagaagaattgagacga ggaagaagatafataaattgtagaaggaaggaagccagatttgataatagatcgaaac gggtaagatggctccatcaataagagcaaaaaaactggaaactgaaagtatcatcag aattatgtgagatatttctatatacagttattgggataagaagatgtactttcaatcataca aagagatggggtagaacttcagtcaagtagaggtaggaaagaattggtatatgaaga actgaaagaaataccctggaattagagaagttagactaataaaagggtatgataataatc agaagacaacaaataggactagagaaggggctaaagaaagaaagagaagtcaga gcccatgtcaatgattgtatagcgtgggactataatattggggattggaatagagacaa agactcggtaagacttggtaaaatttgataatcattacaaggcctaaatattcaagaagga aattgcatgatgagaaggcatcaaaaggcggaattaggaaaagacattgggtgacaca atagattggacaataatcagattagagattatgggggctgctagaactgaaaaaggta ttgaaagatgattatagaggttggcgtgcggaatcagtagtgaagaatgataataat tcgatatctttatattgatttggaaaagaatagggtcaatttggagttaataacgtgaagatt attgaaatagaaggacaggactaatggctcaaggtatggagatttcagaaaatagataat atctgtaaatatgggactgaacagggtgagaattgaagatccttggaaagtag	gggctgatgctatacctcttgagggtatggg tccgaaagcaatttcgaacgacaca
1721-1722	Ga0311341_10020266 GENOME_ACESSION: II_bly_9_\$F_330002990 8 GENOME_ID: 280489 CONTIG_ID: 20265 SOURCE: JGI DATE: 2019-06-13.	atgcaagtaccagtaataagacaaaacccaagatgctctgatccgactctccggcaaga gcagcacagtggaattagtcacaagaggtaccacatttggaaatttgggtatattctgtgt gcgggtgaatcaacccacgggaataatcaagaaaataatttctgctgggggtggacagcg gaagcaaacgggaagcgggttggtaagctctcaaacacactatgtcaacattttagc ggatgcggtagacatgggttaaggaagcgggtggaacaaaacggaatgctcgtagaaca agacgaaatagaacgactccatgtcgggaagacaataatagaactcgtggagatt gcctctctccaccaaggaagatggaattcaaaactccgattatcaataaactaaggaaa atcattcttataactcatatgtggttgaagatttgcgtctccaccaaaaagggaagaaa tggaaatgtacgttttagtccgcttcaatgtggaaaagaatatttcatatggaattggaga acttggaaaactcagactaagcaagatgatgaccaaagagatgcgggataaattgg gattgaaaaatcatccagtaaaatggctgagattttagggcacataacgtatagcttgg gtgttggctcattcagatgtagggtgagatttagacaacaaggaaactaacgagagtagttc cattgagattccatagacggcaactccatatttcaaacatctgtgggtgagattgcgaag acctatgtgtggaacgatgtcactgggattaaagcgtgggtcttgggttcgcatcaaaat atgatttgtgttaggtggaacttgggagaccgaatttctctgcatatctaaagatg gaaagcgattatgccagaacgccaacacatcagatattaagtcttgcatacaactatttt cgtgtctctccctactctaaaggatag	gtcaataaccacgacctaaagggtcggg gcttgaagaataataaaccatattgtttg acaagttaaacgtgactacacaatcaaca actgaacgaatggatgtgtaatccaaat atattgaggatgctcccaagctctcttct atgttcgtca
1723-1724	Ga0265294_10058290 GENOME_ACESSION: MunlanPumphouse3_2_\$ F_3300028602 GENOME_ID: 280699 CONTIG_ID: 58289 SOURCE: JGI DATE: 2019-06-13.	atgagggttccggttctgattctcagagacgcccgtgatccgaccacaccgtacgg gcccgcctgctcatcaactggggcggaacaaacgtactggaacaagctcggcatctt ctgcatatctgacctatggcgtacagcctgacaaccaggatctgtgctcgggacga ccccgcgagctcctttaggggtgtgctggtcggcaccacaagagactgttgaacg gcatgtcggagaccccccaagcagctcaagaagacgtggagactagaagatgtcgtc gtcgggcccgcgagggccggaagtgtcgtcgtcggcggcgggttcaacaacggc tccgtggtcggcagttcctgcccgcgagcaggttgcggcggtggaacgcaaggtacg gatcctcaaccaactcagaagatcctgcccgtacccgctgtgctgtggaggacgttgc cagcagcgacgaagaagcactgcaaacggtggaacacgaacttctcggcttagagggg gggaaaacagtgttttaccagacgatccggatcacggccttgcctgcacccgcgtg ccgggtacgaacaaagcactccgggacgggttcggttgaagaagacaagccaga agagcaagccggcttccgctcatgcatcgtggttgggtgatggcggcgagcgtc tcgggtgcggagcagccgaccgaacgggggttcttactggacgccgatccggttgc accggcgagcgttgcaccggttgcagccggataccggcggtatccgaaaaccctacg gaggaaacccggtcgtcgggttaccgggggagcaggtccggcagctcaagtacg gattagcctatctcgaggaacgttgaaggaagatgtcattgcataacgcggtcacc	gtcaactacccctggftaaacgggggcct gtgaacggctacgcaggaccaaacagtag actaggagggaagaattgcagcagcgtta cgggctacaagaacggcggtatgttccct agtcctgcccttccgggcacgccgaacc cgacgtacctctcggggtactaagcctgct cgcaactccgaagggaatgaaccccg ctgcaccgggagaaggtaatcattaggg gttccgggttctgattctgcaggacgggt gatgccgaccacaccctacggggccgct tgcctatcaaacggggcggaacacccgt actggaacaagctcggcatcttctcatcatc ctga
1725-1726	human gut metagenome genome assembly, contig: NODE_1363_length_780 9_cov_11.459247, whole genome shotgun sequence GENOME_ACESSION: UAOH01.1 GENOME_ID: 14968 CONTIG_ID: 1362	ttggttggatgtacagcaccgtgggatgtttatctatgccacgctctgctgtgtgt ttaaaggtctgagaggtaggaaagggtgatgcagacaagaacccattccaactgacg aaggatgagaatcgccctcgggcggaacaaacggccttcggcgagcaagcgga gcctgcgggtatccgcaagagagatacttgaagattttttgaacatcgcgacgaag ccgctgatccgtgttcgacgaagaaggcgacacatctgaaggcggcgcaagggc gtcgtcgtgctgcgaacgccgttcacgatcaagctgacatccgacagggcgagacga agcaggacgtgacgttggcgtcgtgacggcgcaagcgacgttggcatttccgccac gacggaaaaggaggaggtctcgtcgtccgaagtcgctcgtcagaggacatcacggga cttctgccgatcgtctgcatctccgacgtgcaaggcgcaatcgaagacgcgtaccg cgctccgcgttcaacaatcgctcgtatcaagacagaaggatgcttgcggctccgt ggagaaccgtatcaagcgcacatttcgcgacgaggcggtctgcagactgttccctgt	gtcaataacccccgcctaaaggcgaggct tcaagagcctttattgactgtctagcaaa cctcctctgggagcgagctcgttgggtt gaattacaggcaccgtgggtattatctt agtcacacgtctcgtctgtttttaaagt tctgagaggtaggaaagggtgatgcagaca gaaacccattccaacattgacgaaggatga gaatcggccctcggggcgaacacggcc ttcggcgagcaagcgagcgtcgggt atccgcaagg

SEQ ID NO:	Name	IscB sequence	ωRNAs
	SOURCE: WGS DATE: 2019-01-27.	caccaagatcgtgattgaaaccgcatccttcgacattcaaaagatcaggaatcccgaagt cgaaggagcgggctatcagcagggcgaccagcttggattctggaacgtgcgcgagtac gttcttttcagagacgggtcatactgtcagcactgtcagggccgatccaaggatccgatcct caacgtgcatcatcttgagagtcgtaaacggcgagatgcgccaacaacctgatca cgtgtgcgagacatgccacaaggcttatcacgcaggaagatcaagttgaaggtcaag cgggtcaatcgttcaggcggaagccttcag	
1727-1728	sediment metagenome genome assembly, contig: NODE_138_length_1178 39_cov_9.207605, whole genome shotgun sequence GENOME_ACESSION: OVX001.1 GENOME_ID: 11753 CONTIG_ID: 137 SOURCE: WGS DATE: 2019-01-27.	atgccatgcaccgagaagcgggcgcgcctcttctcagaaaaggcgtgcggtgtcc accgcatctcccccttaccatccgcctcaaggaccggctcgcgaagagagcaacctc cagccctgcgcctcaagctggaccgggattccggcagaccggcgtcgcggtcgtgc gagagagcagggaccctcagcgtcccgccaccaccgctccactgcgcagtgtacca ccaccgccaggagggttcggtccgacatctcgcgcggcgtcgtcgcgcgcgc aggcggagctccctgcggtaccggcctccccgattcgacaacaggcgcagggcccaa agatggctccccctgcgcgcgcgcgtcgtcaacgtctttcgtatcgcccggtcgt tacacgaggtttccctcatcgcgccctgcacgtggagctcgcgaagtctgacacca caagctccagaacccggaggtctcgggggtgtctaccagcaaggaggagctgtgggg ctacgaagtcgggagctgtcgtgagaatggggcgcaggtgcgcctactcgcggc gccacaacgtcccccttgagatcgaccacatctcctctgttcaggggtggagcga ccgggtcagcaacctcagcgtgcctgtcggcctgcaatcaggccaagggaaccga acggctgaagaatcgggcaccccgaggtgcagagaggacaagaacccctggtga gggtgcgggcgaggaacgccaccgggtacgcgagctcagtcgctcctcggctcttgg gctgtccgtgtactccacacggcagcgcggagcaagtggaaccggagcagctcgg gctatctaagtcgcacactgtggacgcctctgtgtggggagatggccgagtcgccc gatggagatgccggtgtgtggatccgcgcgattggcaggggtcgcgggtgcggagc gaacttcgacgcgcaggggtgccagaggatctcgcgcgcacgaagcgggtgttc gggttctgcacggggatctgtgtgtccgaggtgccaagtgggagtgatgcaggcc ggcacatgggaacggctctgatccgtccagcgggaggtttgacatcaaggacctgcg cgggcggcgggtgtgcgcaggggatcaggtggcagcactccgctctgtacagcgggt ctgcgggtactcttacgaaaggaggaagcggcgctctctccacgcctgaaggcg ggggccttcggcgccgcagaaggatgtga	gtcatggacccacggctaagccggggg cttgcgatgagcaagcctgccatgaccag ggcagcgttgcacggacaagggg atgacggcgctacgttgcggcagtgctcg cagaccactccgggtgttccccagtc cggaccctggaagccccggctgcagaca ccttcgggtgtgagaaacggggcgggg cacgtgcggccggcaacatggccgagg ggagcgcggagaggagtcacccccgcg gtcaccaggccgtgaaggcgccccctcg ggggcagaaggaggtagggagccttg
1729-1730	Ga0376086_0041273 GENOME_ACESSION: LAC_S2_MG_FICUS_ FD_\$F_3300035686 GENOME_ID: 312530 CONTIG_ID: 41272 SOURCE: JGI DATE: 2020-04-07.	ttgtcactgtttcctaanaaggcgaatcatcctgatagaggtgtgcgaagggaacttcaa acttgactcaaggattatctcaatgcgaggtaccaggttttatcgccagaaggacacccatc atgccgactaaaccaagtcgtgcgaaggcgtggcttaaaaaggcgaagccaaaattgt ttataacgatttggcatttttacccaactgatgttggatctcgtgcactaacactcaa ccaattagtctgctgtcgtatccgggtaaaaatatacgggaattggcgttaacagcta aatctacttggtagtgcatttacaattaccctttgaaactgtcaaaaggcggatggagc aacgacggatgatgcgtgtgtgcgcagggggaagcaggttaaccgcaaggttccctat agtcaagagcgcacgcccaagcaagatttgacaatcgacgtccaaaaaggttccacc atcgctccgagcgtcgcagcagctagctgcggatattaaagggaactgactaaaacta tccttaactacagctgtttgggaagtgtttactcgtgcgaaggaattttcacccgt aatgtgtcctaatttggcaattgaccaactcaagcaactgctccagtagtcacccaa gaagggtggcaaacatcaatcagacagtgctggaattgaaaaacagaaaacca aaaaggtgatgcaattccggaaccatgcagtagatggagtggtactggtcttactca gtttgtccgatacgcacaactcaagggttaacaaggatgtgtggaa	gtcaactacccgcactccgctcagctgc gtgacggggcgtgcaagagacaagggttgc aacgcaagagtgtactagcacgttgagcct ctatctgcacgcactctcaggtatttcccta gacctgattatctgcaaacagctgtcactg ttcgtctaaaaggcgaatcatcctgataga gggtgtgcgaaggcatttcaaaactgactca aggattatctccatgcga
1731-1732	MIS_10021274 GENOME_ACESSION: IMG_3300002026_\$F_3 300002026 GENOME_ID: 305123 CONTIG_ID: 21273 SOURCE: JGI DATE: 2019-11-15.	atggctgtattgtattagtaaaaggttaattaccctttatgcctgttctgaaaaaaggcga gaaaaattattagagcctaactgtcgtgtagtgtgtaggggtatattcgtttactattagattaa agatagatcagatgtattatgcagccctaaatgttaagcttgacccggtagcaaaagtta caggtattgccttaacaagagatgtagaacggtaactattgatgtgaaatagtaaaaac ggtaaacgttttaactatttgaacttactcatagagggaactaattagtcttcttataaaa ctagaaaagcttttagagctagctgtagagctagaatactagataccgatctcctagattt taaatcggttagacctaaggtgttgcctctagtctaaacctagatgacttctattt taaatttgtaattaaatataaaatagtagcatttaactaataacacaagaattagttaa gtttgatatgcaaaagatggtaaatcctgagataagttgtattgaatcagcaaggtacgt tgcagggttatgaagtacgcgaattttgttagagaataataagaacttgcgtatattgtg gtgtgaaaaatgtaccttacaatagacataccaagctaagtcaaaagggtgtaccaa tagaattagtaattgacattagcatgtgaatcgtgtaataaaaagaaagataatttagatt agtgtgttttgaaaaataagcctgagtttaataatatcttaaaacagtaagtcctcc taaggatgctgctgtgtaaatgctactagatgggtcattatttaataacttaaaaaatag gtttaagtaggagtaggagcggagggttaacaagtttaatagagtgaacttaggacta cctaaactcatgttttagatcgtgaagtgttgaataatgttatcaataattggaatat cccatccttaattagaatcaacgggtagaggtagatagtagaacctataataattctttt gggtttcaataggtactaattgaaaaacaaaagtattaaagggaattcagacaggagatac tgttaaggcactaattacaaaaggtaaaaagataggggaatttttgtaggttaacaataa aagccacaggatatttactataaaacacgtagtaccactgtaaatagcttattctttaagt actttacgtgattcaacgtgcagatggctatagctatacttttaa	ataagtacccccgcagtaatgcggtatttt gaaaggaaataggtatttaacagactaagt gaatttttaactacgttgaatagagttaag acctacatgaatgcttccctagttatagct ctagaagtttaactgcagacacgcttaggg taagcacgaacgggtttagacataatgctg aattacaactgtcagggggagattcacag aaatgtgagcgttaacaaggcgcgtaagcgt aaaggatttttaattggct

SEQ ID NO:	Name	IscB sequence	ωRNAs
1733-1734	JGI24712J26585_10005370 GENOME_ACESSION: BioPla3DNA2_2_SF_330 0002168 GENOME_ID: 23823 CONTIG_ID: 5369 SOURCE: JGI DATE: 2019-01-27.	ctgtcgtggtggtcgccgaccagagagactgtcgcaaacggcatgttggaggctccaaaa acgtcaagaaagcaatgaaacccgaagaacgtcgtcgtcgtccgtcgtcaccggaa gtgctggcgtcgtcgtcccggtttgacaacggctcagcggccggcgttctgtacc gagcacgttgcgggtggaacgccagatccgaattctggaccagctccagaagatcc tccgatcaccgatgtgtgtgtagaggatgttttgcgggtgacaagaagaattgtcggcg gtggaaacgaaacctctcgtccgttggaggctggggaacaatgtttatcaggcaacca gggacgcggactcgtatcacctcgtcgtcgggtgtatgagagaaagaactccggga gcgggtcggcttgaagaagactcagcagaagagtaaacgggtgttgcgtcccacgcgg ttgatcgtggtgtgatgagcagctgacgtcccgggtgcggag	gtcaactaccccccctgaatgttggggctt gcgaacggcctaagcgagaccaaacggtag accaggaggcgaccaatcgattgcagcag cgtttcgggtacatgaacgggtgggtgtt ccctggcctgtccctcttccgggcacgccaa aacccgatctgctctctggagcactaagccc gtctcgaactccgaagggaatgcacccc cgattcgaccgggaggagtaacctgcc atcggtgtaccgtatcagaccacaaatcaaa caccgttgatgccgacgcccgacggcgtg cccgctctctctcaagcagggaagggcga agccgtactggaacaaactcggcatctcag tatcatctca
1735-1736	metagenome genome assembly, contig: NODE_590_length_2905 5_cov_3.058414, whole genome shotgun sequence GENOME_ACESSION: OGEZ01.1 GENOME_ID: 4324 CONTIG_ID: 589 SOURCE: WGS DATE: 2019-01-27.	atgaaagtatacgtaatggatggatgggttcgggttcggccaacgacgccacggaa agcacgcattctgctgaatcaggcagagcaaggtgtgctgttcgtccaccattcaccatc cgtcttaaatataagaccgggtgtgctgcacccggctggaatgagcatcggcgtggacac cgggtcccagcatatcgggatcagcgtcatcagcagagagcaggaaattgtatcagca aagagggaatcagctcgtccaccatggaaaagcgggtcttctggagacaaggaa gaacctgcgaagaagtagacgggtcacgaaaacacgttaccgtctatccgaaatttagac cggataccaggcgcacctatcaggaaacccagattagcgaataaaaaatgatcacac tggaaagaggagacgaatacatatgagtcgtcgtgagaccagaaggtgtccgccatc catgacgtcaaaactggaccatcatatccgtatcttccagatatacaggaaaccttgcctc ctgatacagaccttaaggatcgaagtgggaggttgcgtccagcatatggagaatccgat gattcacgggtgaaatgtatcagaaggagcacagatgaatcagagaatctgaaagcata tatcttctcaagagatgtgtacacctgccagtcgtgtaagcaaaaagccggaacgaaacg gaaggacggaacgaccgtccgctgcatgctcatatcgaattccggagcaaaagggtg cttctgataacccgaaacactgtatcacagatgtgacaaatgccatcgtccgagccca taaacggggcggtatctgtacgactggatgtcaaaaaacaaacagtagcaagaggat accgggatacggcggtcatgaacgtcttaagacgtcgtcgtcgtggaagcggtccgaat gccgaatttatacaggaataatcacggcagcggacagaaaacttccggtcttgagaaa gtcactgtaacgacgcaacagcgtatcggcgcacatggactgaacagatccgggaca ttccggatacaacgtattatcggcatccggaacagaaacgttctctgcatgaggcaa caccacggaaggacggaagggaacggaaccgacgcgcaaaagcgggaatccaaagaa actagccagtcgggaggtctgtacctgaatgataaagtgtgggtattggacagtagcggat ggatcagcgggttcagcggaacaagcagtgcttatataagaacgggagtggaataac attcgggttcccggaaaagccatagcgtgatccggcctaagggtgattgcaaggact cattgtataactgggctgtttatagcgcataatagattttatagacagcggggaaggaa tattaagaccacagaacataa	ttgaataaccgagcatcggaagtagaaaat gggtgtctgtgagggttcgtcgtcgcagacta gttgcactagcctaaggacaaggctccataca gccagtaagggaagccttagtactacgtctt atggagtgtaataaccctaccggacaaggttgc ccatcggtatgggttcgggtacgtatgtgaa caagagttcaacgaccgggtgatagaagct cccgacattggcgaggacaatacacatcg aaagatgaaccgg
1737-1738	fermentation metagenome genome assembly, contig: NODE_421_length_4970 0_cov_20.654930, whole genome shotgun sequence GENOME_ACESSION: UOY01.1 GENOME_ID: 16056 CONTIG_ID: 420 SOURCE: WGS DATE: 2019-01-27.	atgaaagaataaagtatttctgttagtagatagtaggggaaaagccttaagctactaa atctgagaagacatgtattttataaggaaaggcaaaagcaaaattagtaacacagtaccca ttaacaattaaagtgtacttatgaggtagacaatacggatattagtaaacacacatgggttta gacgtaggacaacaacaggaatagccttagtacaagaatgtaaaacaagaataaagt aatatttaagggtgagatagtcaccgttaaatgtatgaatgacttaatacgaactcgcaaa gggtatagaaaaaacagacgtagtgaagagggtatagacctgttaggtataataaagt catcttcaagtgaaggggaagggttagcaccttcgataaaaaacagacaagatgagatac taagggttagtaaaaagggttacaaaagtattgtggtatagacaaagttagtaataagggatgt gagttttgataaagggttttaacagatggctacaacccgtatgggtgggaatataaaagg ggtaaatggttagatgagaacataagaagaagcgaactaattaggggataacgttactgtc aggaaatgtgtgtaaaagacacaatattggaggcacatcatatagatcccaaaagggttaa aaggtagtgacacaataagcaatctataaacgttatgcaatagttgtcatagtagtgaaca ggcaaaagagggaattatataagataaatttcaagctcaactgtgtggaagcaattagggtt acgctacgcgattcagtaatgcaaggaaagacatatttatataacagtataaagtaaaattgg tgtcagatatagcaaaaacagatgggggtacaacaagtaataagaagaattgattggggtta tagttaagagtcataatgatgcaatagcaataaaggtttaaaaccagatacagtagg cgtatagcagtagataatatacaacattaagaagaagcgaatgtaagttagataaaagt cagtaatagtgcaaggggacagagtaataatactccaaggggttaagttctcaataaactg ttatgtaacgcgaataaataagagtggaaggttaaaaggatattacaagtttaactgggttatt agatgttaaacgttacgggtcgtgtaagtgtaagaagcttaaggaaattgtaacagaccg aggtttacgcataagtttag	tataaaaatgtataaataactattgttttata gacataactgttgggttgcaagcctaagt agtacagcattggaacagtggaatgacaatg aactacgttagtaggaagaagttaagacttac ctttagatgtaacatcagcttaaaagctctaa gagttattgaggaagactgttaactaagtaa tattagcataagttataaagaacaaacaa ccgaaataacatgaacttctattgaccttgg caagaagggaataactttgaaagaaaggta ggagaatcctgatgaagaataa
1739-1740	Ga0137379_10004000 GENOME_ACESSION: SagtaG_36_SF_3300012 209 GENOME_ID: 35423 CONTIG_ID:	atgtcacagggtgtttgtcttagatagcaataagcaaccgtcctaatccgggtgcacccgggtg tgcccgcctgtcttgaacaggggcaggcgtcgtgtaccggcgtcaccgttcacgt gatccctcaagcgtacgtgtggagcagccggagcttcagccgtcgtcgtcagatcagtc caggcagccagaccacggcctcgccttggtcaacgatgcagcggggaggtcgtgt ggggcgcggaagtgcgggaatattctctcctcaagtggcatcaccagtcgcctactgc gacgcgcgcgtgtccgctggaactcaccacgtgcagccgagagccaaaggcggc	gtcaggaaaccccccctagaagggggca ggctgtcgtcacaagaagaggaaagcctact cttgaccagctcagccaccagcttctctctc ggaaaggcgtcagcgggctcattcagg gcgaatcatagggcacgtccgggtgtc ccagcccgaccgctgcggagcagcatta

SEQ ID NO:	Name	IscB sequence	ωRNAs
	3999 SOURCE: JGI DATE: 2019-01-27.	tccaaccgctcagcaacttaacactggcctgcacttcctgtaaccaacgcaaggccaat caagacgtccgcgacttctcaag	aacagggtggacggggtaagccagtgtgct gctcatctgaaaccgctcaggaacgttgacg aggtagcatcaccgcgcaagcggaggc tcgcaagagcaaggaacgcaaggaatgg ctcgtatgtcacag
1741- 1742	human metagenome genome assembly, contig: NODE_7302_length_427 7_cov_2.90502, whole genome shotgun sequence GENOME_ACESSION: OEBY01.1 GENOME_ID: 3487 CONTIG_ID: 7301 SOURCE: WGS DATE: 2019-01-27.	gtgggtgtaccgcagcaatgcgagattttatgaaggatataatggttttgtattagaca aatctaaaaaacctttagatagatcacaaacgctgaagctagaataactaattagaaaaa gctagctgtaatccatagatatatcttccactattaggcttagagacaatagtgtataagt aatgatagaacttataccgttaagcgttgcacaaagttcaagacatactggagtggtctattat gatgataaagatcaagtagtcatgttagcagagcgttagcatagaggctcatgttataaa agaatatggatagcagaagagctattagacgcaaaagacgcagtagaagattagatat agacctgcgaagatttttaatagaaccaagcctaaggttggttagcacctagcattaaatc aagggtgcgaacgtgattaatttcaataagaataaagaagctctgaatatcaataagg ttatgattgagaatgaagtttaagtagtctcaaatgaagcttaacactaagcttgggttaa taattaccaaacaggacctttgtaccaacaaaagcttagaagttttatattggccgctcaa acggcaaatgtgtactcgcggagctgaggcaaccgagatagatcacattgtgcctaga gctaaaggtggaaactaatgacacataataattagtagcatctgttagagctgttaatacaatg aagcttaataagactttgaaagaa	gtcaattactcgacaaaaaacaatcagctc gattgaccagactcagttatctaacttaactg ttagattaaactacgatagataagttatcacac ccatggatgcttccagctctgtgctctgtgt tatctatcattgtcaggagagactagcttag aaccttaattggacattagctctgtgtgata gctgtgaagtagcgtgcctactgtaaaggt gggtgtaccgcagcaatgcgagatttttatg aaaggatatatatg
1743- 1744	Chloroflexi bacterium isolate CF_117 14 0929 09 20cm scaff old_4585, whole genome shotgun sequence GENOME_ACESSION: GCA_005880275.1_AS M588027v1_genomic GENOME_ID: 285783 CONTIG_ID: 892 SOURCE: NCBI_Prokaryotes DATE: 2019-07-22.	atggtcttcgctcttagacacaaagaaaagccgctgatgccttcacaccaacgaggcgca cgacagctgcttccagagcgagcagctgcctccacgggtgcaccccttcgtcattc gctcaaatgcgcaggtggagagctcggctcgaacccgcttccctgaaatcgtac cgggcagccagaccacggcatgacgctggcccggtgtggaagacaggagcgaggga gccatccaccatgcgggtgtgctgctgaagtcgagcaccgagccacgaggtgcgag cgcgcaaatgcacgaacgccatgctgcagacgcccgcagtgccaatctgcgcca ccgggctgcgcgcgacgaaccgcgcgcctcgtgggtgctgcccgcgtcgtc gcttccgagatcggaacgtggtgagtgaggaagcgctgctgctgttctccgctc acacgcgtggatgtgagtgctgccttcgatacgaactactcagaaccccgagatc actggcgtccaataccagcatggggagcttttgggtgggaatacagacctatctcctgc tcaaatatgctgaccagtcgcttactcgggaaaacgcgacacccctttgagctcgtac tatccggcccgtagccgaggggggtccagtcgcatcacgaacctggcattggcctgtc acgactcgaatcggaagggggcgcaaacggcgccggaatgggggacaccggga agtggaggtgcaggcaaaaacaccgctcagggaatgcggcggtggaatgccacag ccacaaactcgtccaggagctagggcgtgttcgggtaccctgctggcactggagggc ggggcgacgcgtggaatcgccccgcttggcatcgcgaagacgcatgccaagac cgcctcgtcgttggagatctcgcaggcgtgcgagaggcaggctgaagacctggcg atcacggccaccggagcgggggagcattgtcgaccaattggagcaaacaggcgctcc cgcgcgataccgctgctgccacaagcagggtcaaggggttcaagaccggcgaccggg tgcgggcccgtcgtccggacaagctcaagaccaggggcatcattgtaggacgcgtgca gggtcgcaaaagcggctccttggatccagacgcacgagcgggaggtcgaaggagtc aatcgacactactgttgcctgttacagcgaggcgatggctatgcctatgctgtcatag	gtcacgaacccccgatgacatccggggg cttgcagccagacgctcgtcgtcgggtg caagctccatagttccagccttagcgtcca atggcgagggccccaatacggagcgtc cgttcgaggggtcgtgaccccgagagc cctctcaatccctggccctgcgtccaggg ttaaagggtcgcaggggctgagccggtgc tctggaccatcaagcccgtgacattggc gagaggagcatcttgggaatgcctgccag gatcgcagacagccccttaggggtctct ccaaagcgcgagaagaaaggaaaggaga tg
1745- 1746	Ga0072941_1026019 GENOME_ACESSION: IMG_3300005201_SF_3 300005201 GENOME_ID: 311642 CONTIG_ID: 12710 SOURCE: JGI DATE: 2020-04-07.	atgtttgtaggagtaattggaataacaggaaaacaataatgccgtgtatgccacgaagg ctagactcttttgaagctggcaaggcggagatcgtgcggtatgaaccgtttacgattaaa ttgaattacgacacagaaattacgaagaattactgtcggagttgacctggtagta aacatcttgcagtttcgcaacaacagaaagagggtctatatacaacgcaagtagaat taagaacagatattgttattgttatcagggaagaagagctttacgagaacaagacgta cgtaaaacaaggatcgtcaggctagtggtgcaaaataggaaacctaatttggctt ccgccaaagtgttcagcacaaggttaacgctacgatcaaaatattggatacaattgacagt gtttccccaataaccaagtttatcgttgggtgtcctcgtttgacattcaaaagattacaatcc tgacatcaggggagacaataaccagcaagcgatcagtggtgttggctaatgtcgcaga gtatgtattgtacagagatgggcataaatgcagggtttgcggaaagtcggtaaatcac gttcaccatagagagtcgcaaaacaggaggaaacgcccacaaacttagttactttat gtgaaaattgtcataaggacatcaacagtaacaaactgataaagtttaagcaggagac aagtttttaagatgcaattcaggtgacaattattcaagagtgatgtgactggctagtac atcaagggtattaggttgagaaaacatacggatcattcaaacataaacgcaatcaggt tggattacaaaatctcatttgaacgatgctttgtgatcagtgagggaatttaattgtcaggcg gtctgatgaattgtgcttcaaaagagagtttagagcaaatatcga	ggatagtgtaattttacgaaacggaaagaa aggatgattgtaagcaactgttccctata gggatggaacttttagtcttcttactagatt gagcaacacaggcaactgtgtcactgttatt tgagaatacataggcactttgggatactgtc ctagtcccaaacctgcgggtttgtattaaac agtcctgattgttagggcaggtgataaaac atacaaaacctcaaaaatacattatcgaagg acaccaactcccgtaggagagacaacattta gttatattttattatgttt
1747- 1748	Ga0374549_000851 GENOME_ACESSION: MayFmetaT_FD_SF_330 0035593 GENOME_ID: 311788 CONTIG_ID: 850 SOURCE: JGI DATE: 2020-04-07.	atgccatgtcatcctgcccgccagcgaactactcaaacgaggaaaagcggcggtgtt cgccgctatccattctccatcatttgaagagcgcgatgggggtgatcgcaatccgtg agcttaaaactgatcctgtgtagccagacaacggcgatgctcattgcgaatgaacggg tgatattggcggtgaactacacatcgctgacgcggatgaatgacgcgtcgaccag cgtcgcgctgtgcgcacagcaggagcgagcgttaaaacgcggtatcgaacacggctt tctgaaccgtactcgtccaaaggttggctaccggcgagctctgaaagccgctggcga atgtcgagacgtgggtgtcgtggttcgtctgattgtcctccatcacggagatagccttaga agtagtgcggtttgatcccaatctatgcagaatcagatgtgagcgggattgagatcaa caggggacgctggcggtcagaggtgcgggaatatgtacttgagaaatggaacgaa	gtcaatgaccaagcctttcagggttgggt tgtgagagcaagccgacattgaccagactc agcccgtgtgggctacgataggcgggtcat gacaccgtagatgcaagccagcttctacgt ctgtcgttggtagttaaacaagcagacggct tatgtcgtgctgctccagcagacaagcctcc ctatcatgtcgaggcaccattaccgcgcta aggcgagaagcaacggttaacattgcac aatcac

338

339

SEQ ID NO:	Name	IscB sequence	ωRNAs
		cgcacaagttggccgaccggacgtgaggagcactggctaccaggaaggatcagat atggcttctggaacgtgaggaatacgtgctgtggcgtgaccggcacacctgtcaggcg tgcaagggaagttggcgatccgatcctgaccgtgcatcacctccggcaacgcaaaag agggtggcagcgaccggccggcaaacctggtaacatgtgctgagacgtgccacaagg cgaccacgccggcgagccgctcaagctggaagcgccggcaccctacgggatcg acacattcaacgtatcaaggcgctacgtgagtcgggtcactgtccatctgaaccggctg atcacgttcggctacatcaccaaggcgccggcgagtgccgaggggctcccaaatccc acgtcaacgatgcttctgtatcgccggaggccaggagcagggccgatcacatgtct acctggcggggtctaccggccgcaagcgaaggtgaccaaaaggagcgagatcgc acattcgcaacaccatccctcaggcattcggttcaggcgcgagatcggtcgagaatg ccagatggacgacaagggtttatctatggcttgcgcacgtccggctattgacgttcggc agttggctgctgcgaggtgttgagccactcggtcagttggaaagctaagccgttgaga aagcgctacacacgaatcgaaggaggagggttgaaacggcctgcctga	gagcacgcccctgatggcgactacagccg caaggctgacttatccgagaggattgagcat gtta
1763-1764	Ga0209985_10000066 GENOME_ACESSION: NOAmetSPAdes_SF_330 0027806 GENOME_ID: 33087 CONTIG_ID: 65 SOURCE: JGI DATE: 2019-01-27.	atgcaaaagagctttgttctatcttaccacaaagcctcttatgcttgcctacacgcagaa gcaagagagctgctacgtgatcgaagcagcagctacagaagaatctttcatgatc atccataaagagcgtcagatggggatcttcagtcattcgctaaagcttgaccagggg agtccgacgacagcattgcatgtgctggggaatttaaaaggaaagcgcttattcttc gcagctaacctttctataggggcgaagcgtgtgaaagatgctgctgaatcacggcgatcc ctaaggcgctcgtataggcgaagcaagacaggtatcgggtcctcaaggtttttaaactg gactcggccataggctggcttcaggaaacgccttaagagagtttctactaccatgag cttttcacggcggaagaacaaagttaacagatcaaaagcgtatcccaagatcat tgggtcgtatgcagcttgcgttgaataacggcgaaagggtttatatcgccctttcataa ggcttcaattatgaggcgagaatcaagagcctcaaggcagaatgcttccagataagta tgggttccaaagacgcgccaaggcccaaaagcgctcttggcttcagactggaga tctcgttcagccaaagtcacgaaaggcaagaagcggggaagctatcagggtgctgt gctgtcttcaactggcaactttaacataaagcacagggggcgaggttgccaaggcat aagctataa	gtcaatcaccggccctaaaggcgaggct tgtaagaaaggagactgatttcatgctagt gaccggactcagtcacaaaggggctccgt tacgagcgaatcacatgattaccccgatg cttcctcagctctgggtctaaaggttgcggt taaacagatgaagaggttaagtcagtgcc gcaagcaagtaaacgcgtgataacattgct taggtgaccattacctgcgaaagcgagaa ccgaaggtagctttaaagcgaaga
1765-1766	Ga0307379_10050423 GENOME_ACESSION: UN2MG_2_SF_3300031 565 GENOME_ID: 281080 CONTIG_ID: 50422 SOURCE: JGI DATE: 2019-06-13.	atgttagtttttacactcaacaacaaggaaagcctttgatgtcttcagtcctcgcaaaagca aaaaggctgctcagagctaaaaaagcaaaagctgtaagaactacaccttttacgatcaaa ctgctttatgttcaacagggtataagcagggaattacagctgggatggatacaggtagta agaaaattgtgtagctgctgtatctgctgatagaattctctatcagtcagaagtgaactg agaaacgatataacaagaagatgaaacaaagactatcttaccgcagaacaagaagg gtagaaaaaccaggtatcgaaaaccgctgttcctcaacagaggaaaaagcaggtttctgg caccgagtataaagagtaaaattgaattctcatttgagagaaaagagattacagaagaatt ttaccagatcaaaaggatacttgaactcgaagtttcgataatcttataaccaatccta gtgtgaaaaaaaaggctatcaagaagggtgaacaaaaagggttttacaacacaaaagcgt atatctgcatagacggatataaattgctagaatgtaaaaagagtaaaagttaagttacag gttcatcatatttcaaggctcacaagaaggaaactaatgaacaaaaaacctataaacta tgtgtagactgtcatgaaaaacttataatggtgacttcaccattatgggaagtaaatcaaa aacaacacatcgacagaatgggtataaattaaatcagttaaagaagaggttggaga attcaaggaaagttttgttatgagacaaaactctggagagagcaagtttgggtttacca aaacccattacaacgatgcaatcagcatcgtttgtaacggagaacaagaagagattaaaa tattgaataatgtatttacaacaaatgtttccagaggagattatcaacaaccaaaggc aatcattctgaaaaaaagataccaacagtaaacgttttgggttaaaaaattcgataaagt aaagtataacaatgtttatactttatcaaaaggcaggtatgagtagtggtatgctatattgca aatgtatttggagagaaggttaaaftaaagccaatacctaagtttagtaaaatggagagaat aaatgcaagaacaacacgaggttgcaattcatccccacctaagggaagggttttct tgccacaacaggataa	gtgaactccccctgctaaaggcagaggct ccagaagcctaagttaccagatcactaata ggagaagaagcagtaaaatattagtaacc gataagaagaatttaggcaccctgggtgc cacctcagctccaggcactgcggagattat taaacagctttaaagggttatagacaggttta atcttaaaacattcatatctatgcagaga ggaaggcgtgaagctgagttgatctcaca ctcagccacgactcctaaggagagggg aaaaactaaatgtta
1767-1768	Ga0113559_100101 GENOME_ACESSION: IMG_3300008081_SF_3 300008081 GENOME_ID: 28361 CONTIG_ID: 100 SOURCE: JGI DATE: 2019-01-27.	atgaccaattttaattgagattatattataaaggaaataaaaatatgaatttatggattttgt attaattgtgataaaaaataaataaacatgcataccaataaaaatgtgaagcaggttattt attaagagaacataaagcagaaattataaacacgagccattagttatcaaaagaactgat gattataattcggatttagaaaatagagatatattgaattaaaagtagatagtggaatttaa atatcggatttccgtaagcgataatgagcacgaatatattgcaggtcaagttgaaatgcta aacgggaatgagtaatagattttagaagaataatctatgcgtagttcagaagaatcgttt acgttatagaaaaataagaacattgattataaacagtcataatcctacttataaaaatgg aatgaagatggatggttgcaccgtctattgtacataaaatggaacgcatacagaatca ttgagcaattaaaaaatgggttccatcgataaagttattgtagaggtggcaaatgttga ttgctgcaatggatgcatattaaaggacggaactattttaaatggaaaagattatcaaatg gtgaaatgaaggatagagaattgttcttatgtaagagccagagataatttctgtta ctttgcaataagaaaaaagaaagattgttacataaaaggaacaaaaagaatg aagtcaccataaaatacgaatcctgggtggtactaataatccaggaaattgtattgc gtatgtcaaggatgtcatgaaaaatcattgaataataatacaacaaatattttaaagaa ttattagaacaagctttacagaataatcatttaagattctacatatgaattgttagatg ggaattattaaacagacttactgaaaaatccagagcgtgatattgaagctgaatatgttta taacacaaaaaataagaaaaagcaggtttgagaaaattccattataatgatgcagttt	atgaccaattttaattgagattatattataaag gaaataaaaatatgaattatattgattttgtatt aattgtgataaaaaataaataaacatgcatac caataaaaaatgtaaaagcaggttatttata agagaacataaagcagaataataaaatcag agccattagttatcaaaagaactgatgattat aattcggatttagaaaatagagatatattgaa ttaaagtagatagtggaattttaaatacgg attttccgtaagcgataatgagcacgaatata ttgcaggctcaagttgaaatgctaaccggaat gagtaatagattgttagaagaaaatctatgc gtatgttcagaagaatcgtttacgttataga aaaaataagaacattgattataaacagataca taatcctcttataaaaaatggaatgaagatg gatgtttgcaccgtctattgtacataaaatg gaaacgcatacagaatcattgagcaattaa aacaatgggttccaatcgataaagattttagta

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gtataaaagaatttaagaatataactttatctaaaaagtattttgtcgaacaaaaacgttg taatgatagaaaaatggaaagtttgcagatgcaaaatatatagattcaagagatggtaaaa agaaaaaggaatgatctaaaagtaattaggcatagactaaatcaaacgctctacaaa taagaacatattgataatgaaagaatttgaagaagagaagatcgaaggttaaaata caatttgagtgctatttataattgcgtaaaaccaggagatttaataacataaaagaaggtaaa cataaagggaattgcagaagtttctacaatcacagattgtaggcgaataaattcctaatc cgataattgataaattgagataaataaaaaaattgatttcaacagagagttgaaaaaa agaaaaacaacttctaatatgactgattatcaaaaacgttttgcataaatacaaatagattta catataaagaatctgatgcagatggaccatctataacgtttacgcaaaaagaatatgaaaa gcttaagagaataaatacagatagagtaaaaattattcgtaccagacgtggattgtttgga gagagtacgacagacttcatatgaagccgaaaatatggtcaggaagaaaaaagaafta gaagtaaaaaataagaacaagaattaaaagctgcataa	gaggtggcaattttgacattgctgaatgg atgcatatttaaggacggaactattttaaatg gaaaagattatcaaaaagggtgaaatgaaagg atatgagaattgtgtatcttatgtaagacca gagataaattattctgttacttttgaataagaa aaagaaaaaagggtgacactaaaagagaa accaaaaagaatagaagtgccaccataaaat accaagatcctgggggtgtactaataatcca ggaaatttgatttgcgtatgtcaaggatgtcat cagaaaaatacattcgaataataatacaaca aataattttaaagaatttagaacaagctttac aagaaaaatacatttaagatttcatatataga atatgttagatgggaattttaaacagactta ctgaaaaataccagagcttgataattgaagct gaatatggtataacacaaaaattaatagaaa agaagcaggtttgagaaaattcattataatg atgcagtttgataaaaagaatttaagaatataa ctttatctaaaaagttatttgcgaacaaa aacgttgaatgatagaaaaatggaaagtttt gcagatgcaaaaatatatagattcaagagatg gtaaaaagaaaaaggaaatgatctaaaag taattaggcatagtactaatcaaacgctct acaaataagaacatattgataatgaagaa tccttagaaaagagaaagtatcaagggttaa aatacaatttgatgtcatttataattgcgtaaa accaggagatttaataatataaaaagaggt aaacataaagggaatttgcagaagtttcta caatacagattgtaggcgaataaattccta ccgataattgatattatgagataaataataa aaaattgatttcaacagagagttgaaaaaaa gaaaaacaacttcaatatgactgattataca aaacgttttgcataaatacaaatgatttaca tataaagaatctgatgcagatggaccatctat aacgttaccgcaaaaagaatatgaaagctt aaagagaataaatacagatagagtaaaafta ttcgtaccagacgtggattgtttggagaga gtacgacagacttcatatgaagccgaaat atggtacaggaagaaaaaagaaftagaagta aaaaataagaacaagaattaaaagctgcata aa
1769-1770	human gut metagenome genome assembly, contig: NODE_2190_length_16758_cov_18.186733, whole genome shotgun sequence GENOME_ACESSION: UPLV01.1 GENOME_ID: 16628 CONTIG_ID: 2189 SOURCE: WGS DATE: 2019-01-27.	atgccaaacaaagtgtatgtactcaaacgacggcgctccgtgatgccctgctccccg gcaacagccccgtcatctgttagatgacgggaaagccaaagtgaagaagagagacgcct ttaccatccagctgcttttggcagcaccggctatcgcaggagattattctcgggtggat gccgggaagcaaaaccatcgcccttccgcttctaccgaaacggaggaactgttctccgt gaggtaaaagccctcgaacgatgtgtgtagacttactgtccaccgctgcagagctctcgtga gccagaagaaaccgcaaaactcgctatcgtaaagccaggttcgacaaccgctgcgga gcaagcataagggttggttgctccctcgtgtagggcgaataatccaagagcatattactg ctatcaaaagagcctgcccgaatttgcctgtcagcaaaagtgtcgtagaaccgcccaggtt tgaccgcagcttatcaaaagctgtcagggaagggaacccatcccgaaggagaggac taccagaaggcgagatgtatgtgctactacaattgtgcgccagtgatcttgtgctgat ggctatgtctccaatgctgcggagcgtatgcgacccataagaagaggtgcggtcca tgtcatcatctggagagccggaaacggagggcgacgtccagtaattctcattaccct atgtgataactgccacgagaattacacaagggcattatcacggcagtggaacctcaagaa gcgtgaagcggcttctcacgggatgccacattcatgggcatcatgcggaagactttgct gcaaaaggctgtgtgctgaactaacattctgttgcggagaccagaggctacatcacaaa ggcaaccagagaaaaactattgttctgcccagaagccataactaacgactctggcaat cgcgagggtccacagggttccgactggatacttccaagcatccgacaagcaacc cgcaattataccatccgacccgtgcggcatcataaccgacagctgcacaaagcaaccatc ctcaagggtgtgtgcgaaaagctaatacaggcggaagatgtgtctacggcctccgct ctacgacaaggtgcttacaacggattgagtgcttgggtgggagggcgagcagcgg ctcttctgctacgaaggctgtgtggagcaaaaggtcaaggatggcgtgaagctacaagcat ttaaattattagagcgagtcagaattatctgatagcgtaa	gtcaataacccccaggctaaagccgggact tgtggaacacagagccctgattgactaccct aagtgcttcgagcactactgttactcaagaat gtatgtataggcaccggtggacgtttgccca aatctccgctcgtcgggtgtgtgattaaaca gctctgagggttaaggagcaggtttgcacgc aaaaacctgagataacattggggatgggta ccttaccggccaaaaggctgagcggcttaatt tttagcccaactctgaaaaggaggaaggc atcatgccaaacaaa
1771-1772	Ga0137365_10001047 GENOME_ACESSION: SagtaG 29 \$F 3300012	atgaaggtgtttgtactatcgaagaaggaagcccttgatgcccaacgcccacggcg cgcgccgggtgtgcttaaggcaaacgtgccctgtcgtgcggcgagaccctttacca ttcggttgcgtttgccaccagcagcatcacacaccggcgaggtggcggttgcacac	gtcaatgacctccctataaaacggggagact tgtgaggggttgcgcccttgcgaagccggatt gaccagacggcaccctctaagccagtcgcg

SEQ ID NO:	Name	IscB sequence	ωRNAs
	201 GENOME_ID: 35416 CONTIG_ID: 1046 SOURCE: JGI DATE: 2019-01-27.	ggctcgaaacacgtcgccatcgtccatcaccaatggcggaggtgtattccaggccga tgccatctcgtgatgacatcgtgaaagctgaagcaacggcaccgggtatcggcgca atcgccggggcgcaagacacgctatcgtgaagcgcgttatgacaatagacgccgacc tgatgggtgctgcccccttcgtgcaatccaaatcacaggcgaccgtcagggtgtgc gctcatcgcctgtgtcgtggtggccagatcacagtgcgaagtcggtagttttgacacc caaaagatgcaaaagccggacatcgtgggattgatacagcaggcgagcttgagg gctacctcgtcgtgaataatctcctggagagtttcagagaacgtgcgtgtatttgaggc ccaaggggttcctttagaatacggaccatctgattcccaagctcgtggagggtcgaacag agcatcgaatttgacgtgtcgtccgtccttgcacacagcgcaaaaggtcagcaaacagc ggcgagggttggcttccggcggtgcaggcaaaagctcgtgtccgttgaagcatgcag cgcatgtgtcttactgaaaagcagggtgatccatgacctgcaagcgatcttcggagaga gccaagtgagcatcacataggctacgagacgaaatacaaacgcgtccagggtatttct gccccagtcacacgacgcccgtggcgattgcctgcgagatcggcgagggtggc aagccattggaatagtcacatgacccgtgtctaggacgggccaatatacagctttca atggtctgcacagcgagcacaagtgttggggccgcgcactttgcgggttcaagctct acgaattgttaagaccaaaggaggcgtgggttacctgcgggacggcgaggagaag gagcgttcgtcatcaaaagcgtatcagcggaagaaagtggtcagggtcagccacg caaatggtagcggtacgcgtccaccagggtcgataattactgactgccgtcct ggagagcatcagaagggcggcggtcctcctcccaagttgaacggggggcacc ggccggcatga	ccgagaaagcgagcgtcgtggctggcgct gatgaggtgagcgatagcatcaacggttgc agttcgcaccaggagatgctccccagttcc tggctctgcaacctctgacttaagaatcgc ggtcgaggaacgggtgagaggaagta ccgggtctatcaccgtcagggggtatcgt ttcactatctcacaagagaggagccttacct gat
1773-1774	Ga0070717_10000790 GENOME_ACESSION: LARL11StandDraft_3_5 F_3300006028 GENOME_ID: 31022 CONTIG_ID: 789 SOURCE: JGI DATE: 2019-01-27.	atggctaccaactgtttgtcatcatctgaacggcgatgccctgacctcctgc taaggcacgcaagctctgcgagatggcaaaagcagagttattgtcgtgcccccttacc atcaaaactcttgggactgtgaagagcaggtccagggaagtggtgctggccctcgataag ggcagccactacacgggggtctgtgactgcaaaaggcgagattcttctatcgtgcgtg attcagcaccggcaggatgtgaagacaaaatggagaatcgttcagcaaccggagaa atcgtcggagaggaagtggatcgaccggctagattcagcaatcgtggtatagcaag cggagcggcggtaccgcttcgcatcaagtcgaatgtggagaggtgatccgtgtcgt ccggcagattccacttcgcatcgaagatcacaatagaggacgtccgggtagatcgc ctcgcttaaacgattccacgcttgcagaaagtcgtatcaggacccacacgggtcgtg aaaatctctgtatggcctgcctcatgcgagacgggtacacgtcagcagtgccgcaaac aaaagattcgtctgaagcgacacacactatcttcagagagcgaggaggaagatac gctgaataacctgctgactctgtgcagcctgccataaaaagcttcatcagaacaagatc cagcttaaggcgacgggggtgagtggtgcgtcgtatcagatcgcccaacacatgca aggcaagagctatctacgccgtccttggcacacaggcagcactcttaggttttggc taccagacggcaactctgcgtaaaagcagggtatcctccaaacacatgacgtgatgc cctctgtctgccacctcagatacggcgagacggtag	aatagtgcttaaaaatctggaacgtctctac cggattcatggcactaccagcggtgtgtt gtcagcagcaggtgagatggccatgacacc accagatgttctccagtcgggtgcaactgt cggcatcctaagggtgagcggaacacgga aatgtggatgctgtaaaagccacttca caccgcgagggtcatcatctcctaatac aggagggttaacgcatgggtaccaac
1775-1776	Ga0211664_10024503 GENOME_ACESSION: IMG_3300020455_F3 300020455 GENOME_ID: 278939 CONTIG_ID: 24502 SOURCE: JGI DATE: 2019-06-01.	gtgagagtcgttggctcgcacatcagaacaacgcgtgatgccttgcgtcgtgcgga gcacgccagctcgtgctgatggcaaaagcagcggtgttctgcgttatccgttactatcgt tattgaaaggacgacggcggtgagatacaactgttttacttaattgctgatccgggca gcaaaaacacaggtattgacttgttgcgttattccaaggtggccacgctgcgtctacgg tcagcatattgtccatcgtggacaacaatcagggtgacacgaaaccgtcgcgtac ccgtcgttcacggcgaacagaaaaacacgctatcggaacacgggttttttaaacggaac ccgccatttggatggctaccaccatccctaatgtcagcagtgcaaatgtggaacgtg ggcgaaacgtctaaagcaggttgcgcgtgtcagctcagcgaatgtcagacgtgcgt ttgacatcagcgttatcgaaaaacctgacatcgtggaatggactaccaacaaggctcact gtttggctggaggttgcgtgaatacttgcgtatcgccaccagcatcctgtgcgtactgtc acggcctaacaaggtatcggttcttgagaaggacatatcatcccaaaacattagggtg gctcaaacgactggcgaaacatgcatcattggccgacatgcaacgaggac	gtcaacgaccacgacttattcgttagcga attaagtcggagcatggaagcctgagttgac cagaccccatcgggttaactgatggagcgt tatcgagaatataaggcaccgtgggagta cttccagttccacgctcgggaaacgat taaacagcgtaatgggttaagccagtgtc tttttataaacctcgaataacagggtcag gcaaacataaccacttgcgtggagaacgt ctgttaacagacattaatttggagtttgata acgtgaga
1777-1778	Ga0181555_1021736 GENOME_ACESSION: 011ome 28 \$F_3300020 051 GENOME_ID: 20407 CONTIG_ID: 21735 SOURCE: JGI DATE: 2019-01-27.	atgttagttttgttctgaacaacatggcgaaaccttaatgccatgtagtaataaaaaagct agacttctttaaagaaggtgaaggtgaaattcattacataaacattcacaatccaattg aaatttggtagtgatgtacaggatgacgttgaaattgtgtggatactggattcac aaacgtaggaaattgtcgtactactgaacaaaaatatttcatgtaattgagattaaatttcg taattgacgttagcgaatttaaacactcgttaattcttgcgaagggtcgtgataatcgca aacaagatacagacctaaagtctttgcaagaagtctaaacggattttctcatgaacaga aaagatgattaaggttaagattaaataccgtggatgatgattgttgccttagttattt agctaaggaaagtaacttaatttgggttgagaatacaaaaaacgtgtaccatgtagta agctaattctgaagttggcatttggatgttgcataaataaataatccagagattagtgcaa agattatcaagggggtcacaaaaaggattgagaattgtaaggcatatgtcgtcatcgc gatgaatacagctgccaacattgcaaggcgcaagaaaaaggacgtgcgtttagaagtgc ccacactatttcaaaagtaattgtgttcagataagccagataactgttattgtctttcaaa actgtcatgataattgcatcaagggttaattgaagcctaacttaagattacaaaaatttaca aagaagcaactttcatgaatacttagcatcagattgagagaatttaccagaggctga gttaacttttgggttagactaagttgcaccgcatggaatttaggttgcacaaatctcatgc gaatgatcgatagcggtaacaacataaagttctatccgcaagaaaaagattgaggtttg tacgttaagcagtgccgtaagaaaaacgtatgcttcatgaagctaaaccacgtaaaagt	gtcagataccacgccccaggcgcggg cttgtgaaagccctacttactagattcagtg agtagatgttcttaggataatcgaactacg ttatgatagttatctgttggatgatacaa agtccaacactatcaggctctgtaaacattc ctcttgggtggggaagtcactcgaatc cttgcggatatttggagttctgtaactca cagctatcataacattatcgtggtaataaa cttctttagaaaggaaaaactttatgta

SEQ ID NO:	Name	IscB sequence	ωRNAs
		agaaaagaacctaacactacgtcatcaagaaatgctaaaaacacaaaagcagttggtaa catttgcctattcgacaaagtcctgattgataatcaagttggatgggttagtgggtttacgggt gggaattgtatgtgtgtgttttagtgggaatacttaagttttaaatactaaaggtaca gattcaatcaggaattaacatggggcgactaaagtaatccacagaacaacaattggattt ctgagattagataa	
1779-1780	Chloroflexi bacterium isolate CF_154 14_0903_05_20cm_scaff old_174, whole genome shotgun sequence GENOME_ACESSION: GCA_005879655.1_AS M587965v1_genomic GENOME_ID: 285751 CONTIG_ID: 151 SOURCE: NCBI_Prokaryotes DATE: 2019-07-22.	atgtctgtctatgtcaaaaatcaagatggggcaaacactcatgccctgcacagaggccaaa gcacgcaagttgctagaagcaaaagaaagcgaagttgtagactatcgtctttaccatcc aactgacttggcagtgccgagggcgacccaagaggtgacctgtgggacgataagg cagcagtgccacgggtctcgcctgtgtgagccatggcagagtgctgtgcccgcagag atccatcatcgtagggtatgcacggacaaaatgcagatcggcgggacgacgtaaaag ccggcgggcacgtcgtgtgtatggcctgtcgttttcaatagggttcaagccgacg cagcggacgctgtcctccttgattaaagcaacgtcggagagatcattcgtgtcgtccg gcacattccttgaccacgtctttggctaccacacccctccctacaggaatataggaaact tgccaaaacgcataatcgtatgcctgtgtgcacgaacccatgccgacggcgacgtg gttctgtgtgagccgaagaatgtgtaccacattcgtattcgtgcccaggcaaacacgcaag cactatcacagtcagccgcaaaaaggcaaaaggccgggttaataccaagtcaatgagg agttacaaggctctcgcaaggagatttggcttggctcaaaaggcaaatatgcaagctgatt cactctatctattccgacggctatttggcctttccaggtgtcaaggagagaacccaataaagc actcccgcgagattgtcgcattctagaacgagaaggacgattttatggaaaatggtgg ataa	gatctacaatagttcctaaaagttaggttta cggacatcatattcaaggagctagcagcgg gctagcgtatgcttagcagcagtttgcagg tgatcacaccttgagatgttgcctccagtc aagctctgtgtgctccagtttaagggttagccg aaagtgaaagtattggagcagcacaag cctgatgaacactcgcgaaggagcaactcac tcttgacagaggagggttacctagcagatgt ct
1781-1782	Streptomyces sp. HG99 scaffold3, whole genome shotgun sequence GENOME_ACESSION: GCA_002742045.1_AS M274204v1_genomic GENOME_ID: 145796 CONTIG_ID: 22 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	gtggctacgttccgtacagacagacaccccgactgtgttccctcagcagtcggct ctggaacccggctcagcagacaacctcaggagcatggacgaaacgggagccggggc ccgcgacgcggcaccgggtacgggacatgtgcgaggggagacgcccgacgtttac ctgccgtgtggcgcgtacccctggacgtcgggtcgggaggggtcgtgaggccc acccatgtgtattgtactcagacagcagcggatgccgttgcagccgtgcagtcgggcac gcgcgcgcaagctgtcgtggcgaagcgggacgtgtacacggcgcacgcgcgttctgt gatccagacgtgtcaccgttcaagaccggggacttcttgcggcagtggttccactgg caagaaggcgggaacccatagcggcggggtcgggtccgcaccaaggcgagcttcaa catcaccacccggcagcgattgtgcaggggagcattaccggcagctccgctcctcc agcgagctgacggatagcctacacagagaaggaggcagggggcctccccaggc atagcctctgggaatccggccctctagcgcgcatga	gtcacaggcttctccactcgaagagtgga ggcgtcggcccatggcggcgacgcctgc tgtgaccagcgccagccatcaccaaatagg aggtgaccttgggtgctactgtccgtacagg acagcagaccaccccgactgtcttccca gcagtcggctcgtgaacccggctcagcag acaacctcaggagcatggacgaaacggga ggcgggcggcgacgcgaccccggtta cgggacatgtgcagggggagacggcg cagttcacttccgctgggcgcgtcacctc ggaccg
1783-1784	Ga0370532_0143 GENOME_ACESSION: 20181204_8_FD_330_0035014 GENOME_ID: 307140 CONTIG_ID: 142 SOURCE: JGI DATE: 2020-04-07.	atgtccaaagtgtttgtaattgattcagaaaaaagaccttttagctccaatcatcggcgca agcaagacaactgttaagaaataagaaagcagccgttttaaacgctttccatttcttgat tcttaaaaaatccgtattgactgttctgtcgtcagcgtctaaataagacctgtgtc taagcatacaggaattgctttgttgatgactcaaggggtgacgattgggacaaaaaaa tggggcaagccctaataacacacgataaatcgtgtgtcttccagatcaacgcaagagg gcttgcctccattttccagtcattccccgggtgaagtgtcttctgctccgaactaaagcata gagggtttgcaatcagagagcttctgtgtccagggagacagataagacgcagcagacga aacgtaagaccagatcgtcaaccaaggttttaaacagaaacgcgtctaaagggttgg tagcgcttagcttcaaaagtcagtcgaaatatacaaacctgggttaacaaactgcgca aggtagcccggaattgaagcgattagccaagagctgtacgtttgatagaagaaagtgg tgtgtacctcggaaggattgcaaacgtgcaaacaggtattcaatattcaaccaagac ggattagtcgaaggaattagccaaggtttgaacacattcacaanaagagatggttactc gtatgcctaa	gtaagcaacccctgcctaaaggctaggggc ttcgacaaatgttgaagcccagcttaccaga ctaagtactctgtactacgttcaagaatagt gttaaaagtcttactcaaaatgcgtagctagt ttgtgtctctagaacaaaagtaaacatct gtatgggttaaggaagtgtcttttgatagta ccccctcgggggtatgcgcaaaagcgacgg cttgaggccgaacggcagtcgctcctgggg gaaaccaaagccagacacctacggaggga aacctcctgcatggctagcgtccccaaag accgcgtcctcagcgtattcgaacattgt caaaagtaactttaccgtaaggaggacat ttatgtccaaa
1785-1786	Oxynema sp. AP17 chromosome, complete genome GENOME_ACESSION: GCA_012295525.1_AS M1229552v1_genomic GENOME_ID: 331850 CONTIG_ID: 0 SOURCE: NCBI_Prokaryotes DATE: 2020-04-22.	atgtcaaatgtttgttcttggacagaaaagcggcaacttgaccagtgcatccggg aaaagctcggcgactcctctgtacagaaagcgtgccgtatttcgccgttaccgttcacc atcattctcaaaagcctaccagacataccgtgcgggacgtggaacttaactcgtat ccgggttctcaggtcacaggaattgcgacgtacagggcacgaaggtaatcttcggggct gagatcgaacaccgaggtcaacaatcaaaagatgcagcagccgtcaattcaaccgatg ggcgttgtcaatgctctgaacaaaacaggcttgcctgtggcgacaggtacggggggac aaaccaagtcaatcgacgcgattagacctacccaaaacgcactggctagatgctgca agggttgggaaaattgattcactaaagcttgaccaccaaacgggttgaatttcagcaaa aggacatggctacgtatagatgtgcgggacgataaatatgggttctactcgtcatccc cccttttagtcccccttaataagggggttaggggggattcataaaagatttcagactgg cgatctgttactgccaggtcacaanaaggcaagaaatagggttattgttggcgaggt ctctgtcgcgcatcagccacttggatattgacgtgattgggtcgggtatctggaattag ccacaggtattgccaaacattcacaanaaggatggttatgcttatgctgtcgtcatacca tagcgcgctaa	gtcagccccccggctgaagcagcgggc ttcatgttcagctcagtagggggcttctg cgaaacggccctaccagcctaagtcttaac tgactacgttttaaggctatggccactaca atgctttccagtttgaacccctgcgttaacg gttaaacagccctagaggggtaaggcag ggcgttagcccaacagccctaaaaacattg gcgaggaaaacattaccgtttacggagac cttcaatgtcaaat
1787-1788	marine metagenome genome assembly, contig: NODE_793_length_1894_7_cov_12.879862, whole genome shotgun sequence GENOME_ACESSION:	atgaatcattacgtttatgtacaagatcacacaggacaaccactgatgcctaccaagcgtc atggctgggtgcgtcgtgattgcggaagcgaacggttgaagacgtcgccttt tacgattaaactcatttaaggtgcaatgaattgacgagtgctcattcatctgtatgc aggatacgtcaatcgggtttctgcacaaagcagaatcagaagaattactcggcggtgtt tagagttatfaaaaggtatgtcagaacgcataaaagacgcgctatgtatcgtacacagc tcgttctgttgcgttaccgcaaacctaaattcaaaaagggtgtgattttgttcaaccgt cccgtggttggtagccctagattcagcataaactggatgctattgaatttaattcgtga	gtcaactaccccgccctgaaggcggggc ttgtgaaaacagccagttgaatagctccct tatttataaggagacgtttagagggaatag atagggttctgtatgctccacaagttagcag gtgacacggaacggattaaaaaccactga agggttaagggtgcgtgttctgttgcagcaaa ccttcgcgaaccgagcgtatggtctacag

SEQ ID NO:	Name	IscB sequence	ωRNAs
	ORDN01.1 GENOME_ID: 10380 CONTIG_ID: 792 SOURCE: WGS DATE: 2019-01-27.	aaatccagcagatfttaccattacctctattcaaatgaaaccgcgacgttgatattcaga aaatcaacaatcctgacatcggaaggcgtggctatcagcaaggcgaacgattagggttct ggaatcttctgtaatatatcctcacagagataaccatacttgccaaattctgactgtaaaa acaaaagcaaaagaaatccttcaagtgcatacatcggtcttggaaagcagatagaa gcgaccgtccagcaaacctgattacctatgcaccaaatgccatagcactaaaaccacc agaaaaatggcttctgtatgggtgggaaccacgctggaagagcttcaagcctgaaacct catgactacgggtcgtcgtctttaatcatccaagcaagagaaatattgacttacctgttac tgaaccttgggttacttaacaaaagcaagccgtattcaaaaagaggttgcctaaaagccatc actatgatgcgttggcgtgggtaacgcaattaca	ccgttaggctgacttactgtaatacgaattgt tattactaaaaggtaaaatgaatcattac
1789-1790	Ga0101770_1058525 GENOME_ACESSION: IMG_3300006674_SF_3 300006674 GENOME_ID: 27981 CONTIG_ID: 24152 SOURCE: JGI DATE: 2019-01-27.	atgtatcgtagaattagcgaatagacttagatagagaaacctagattgataacagaat ttcatctaagaagaagggttgtagcgcctctattaaacacaagttgattctcatgttag gtttattaattatcttaaaaagataattgcctataactaacatagtcataagtgcaaatftg acacacataaacttaaaacccaacgtagttgggtgaaggttatcaacaagggtgaacaac aagatttttgaatgaatgaagaatacgttctgtatagagataaattatccctgcaactatgtg gcaagaaaaacaccatactgaggtacatcatataggctattgnaacaagatagaacag atagaccaggcaatctaataacattatgcaccaaatgccaatcctaataatcataaaga ggaaggtaaattgtatggaatgaagcctgttcaaaacctaagaagcgaacattatg tctactataagtgaaactgttcaatgcgtgatgtgactacattacgggtatattaca aaatccaaaagagttaagtcattgttggaaagacacattataacgatgcaatttgcatagc agggtgaattatcaacagagaatagaacctatctattttgagcaattaggagaacaat cgttcactcgaagattctatgatgcaaaatattgtatataagagataagtcattataaaca ggacaagagctttctgtgtagaaggacacggaacaaaaacttaaatgaagaaaactct cataagtatcgtggagctaaaaaatcaaaaggcagaagaataatctgtaagcaagatac gcttatcagcctaaagataattgtt	gtcaattaccaccgcccttagaggcgggg gcttgtagaataacagctcgggttaggac tatgtcattagctttgctaagattagttag cagaatataatgacccgtggatgctccac aagtcctcatgntctgagggtaattggttaaac atctctgagggtaggagaagtgtgttacc attaaacctgctataacattggcgtatggac ct
1791-1792	Eubacterium sp. An3 An3_contig_51, whole genome shotgun sequence GENOME_ACESSION: GCA_002159465.1_AS M215946v1_genomic GENOME_ID: 123888 CONTIG_ID: 61 SOURCE: NCBI Prokaryotes DATE: 2019-01-27.	atgacggggcattcgtccgtccatgggaacgcgtaagcgaaggagggaacctat ggattttgttatctcgaacgggaacgccctgatccgaccagtcctaccgtgcaagg aagctgttaagaaggggcgccggaagatctaccagttaccctcgttaccatccagga tacagaccgggaagacgggggaagtacagaagatcgaatactgcacggacgtgggat ctccatgtaggatctccgtgaatcagaanaacatgatctgtgggctgcaggcagac ctgctgccgatgaacggagcggcacaatgccaggcgggaagtaccgccgtaccgg cggaaccggaaatgctacatgaagccccgggttttaataaccgttaaaagagagaaggctg gctggcaccatccatccgccataaggctcgaaggcacagctgctgtgtatcgaagatcct cgctgtctgccccgtctgtatcgtgtggctggagatggggaaagtgcacacgagctgct aaaagccccctcgtcggggggaagccgcccgaaggagcggagttaccagcagga gagaggtatgggatcgaacactccgggaagcggctcttaccaggaccgccatacctg tgtattctgtggcgggggatcaaggaccatgccatcctcgtatgccacatgctgggtat tggagaagaccgcagcaaccggcttgcacacttgcattctgtgagcagtgcca tacgccagcaaacaccagccggagcggcatcttatatggaaggaaaccaaactttccgt ctctagcaccggcaacttatatgtccatgatccgggtggcagatgctcgcgatgtgaagg agatgcacgataagtcctatactattgtgtacaggacaagatcgagcgtgaagaac gccatttgaaaaagacgcataatgatgcgtactgtattggggagttccgccgaaca ccggacgccggacgtctatattgaaagaaacgccggaatgaccggatcctgcagaat tttatgatgcagctatatacaccggcgagatgggaaggaagttccgggaaggagcttt ttaatgggagaatcagccgggaaccacaagaaggattccgagaatcgtcatcttatcgc agaagaagatcaaaaaaggacatcttccatacggaaaagcaggacgcagctcaacct ggaaagtgtgtctataaaaacgagaggaagatcgtaacgggacacataccaggtat tacaacccgaaaagatcaccgggtccgggtgaagcgtatcaatgtgcaatttaccgaaca gcctctgatggcgagaaaagtgacagaccttacaattgacactgtgatcgacaggaattc atccgtggctggaagcaggttatagataa	gtcaataaccccgctgatttgccttcggcaa attcaggcggagcgttgcggtcgccgtgag cctgattgattagcctcagctgtcggctcatt cgctgcagctacgttaccaccagaatccatc atcaggtaagacagcgcaccagtgatgctc ctcaagtcactgctcgtgggtgaggatta aacagctctcagggaaggagaagtgtccc gcataaaaactggggataacattggcgat gaggaccacccccgaaagggaagacccc atgtatgacggggcattcgtccgtccatg ggacgcgtaagcgaaggagggaaactc tatg
1793	human gut metagenome genome assembly, contig: NODE_3304_length_853 9_cov_2.986681, whole genome shotgun sequence GENOME_ACESSION: OUOT01.1 GENOME_ID: 11097 CONTIG_ID: 3303 SOURCE: WGS DATE: 2019-01-27.	atcgagataccgggtatctccgggatactgcatggcacattataaaaggagcattc atgcctgatacagattacatatgtactcggcgttgacggttaaacccgagatgccgaca aacgaaagagacatgtcaaaagattacttaataccggcaaggcaggaattgcatgccatg tcccttatacaatagatttattgaaacacacctgttctccagcctgtaattgctggctga agaccgggacgcacaaatcggcatggcgattttatcataaaagggtgaactactcat gtctgtgtcgtggaacacgcaacagggaatcgtgaaggctcatgaatcaccgaggg gatacgcagggttcaagaattgggaacgaaggccagacagcgtcttgcaaacgtt tcggcacaatgcttaaggcaggtctgattatgcgtaaaacttccaaagtttgcgctgaca gtttatcacatgccactgcataaaaatcggagtcggtttcttaacagaaaacatccg gatggatggcttacggcacttcaaggcagcttgttgaacacatacaacttgatgcaca agatacagaataacctgccgttacagacatcgccatggaagtaaacgctttgcaattct gtcgttggaacacccgtctgttctgggagtagacttccgaatggaccgtttaaaggctt gacagcctgcatgatacagtaaatgaattacagcatggcaaatgctgtgtgacaaa aagatagaacattatcatattgttccaagatccataaacggctcaatactattggcaat atagccgggctgtgcatggaatgcacgacaaggtcataaatgtgcagcatatcagaaa aagcttaacgatgcaaaaaggactggataaaaagtattgctgattgagtgccctgaat	

SEQ ID NO:	Name	IscB sequence	ωRNAs
		caggcagtgccatttatctgtaaatgtctgaagctgaattggcagagagcatgtatcttac tgtaccggccgggacactgcaaaaatgcgcagttctttggattccaaaagcaaaaggat aatcagatgcatgaaactgacgcatggtgtataggaattctcgcattaaacaaaattccga atgtggttccggattttggacagacatactgcatcagacagttcagacgtcaggacagaa gcctgtgttgcgccagacagaaagagctacaaactggacgggtgacgattgccaaa accgtaaaaaacgtaccgaacagaaaacagattctctcgaactgtgacacagacag gtgactttataccggcaaaaagagctgattcatgcccagccgattgactgtcattaaaga gcaaaagacgctataatgacttaacacgctcatgcccgggagcagattcatgtacaacg gagagtgctacgtcttaagcggacgtctcacagaaggacaatactccgcccagtggtg gacacaaaaacaaattatctgcttcaagtgtaagataatcagacacaatgaaggactg gtgtttgtatgctaa	
1794-1795		ttgctcgaagaacataaggctaaagtggtagaaaatgtccatttacaattcagcttaattat gaatgtgtaaacattattcaacctgttagtcttgggtgagacatcggttgaataattggaa tttctgcaacaactgaaaaagatgttcttggaggctaaagtcataatagaacagatatag ttaaacttcttcaagaatagagataagacgtacaagacgacacgaaagacaagata tcgtaacactcgggttctaaaccgaaaatatacgaacctaaggatggaaagcaccttca ttcagagctagactagattcacatgtagaacttattgcaaaagttcagatttctgcctatc ggcaagatcattgtggagattgcaagttttgacatacagaaaatcaagaatcctgagatca aaggtgctgagtatcagcaaggtgaaatgcttggattcaacaatttaaaagcttatacttat gcagagatgggcatctatgccgatgttgcgaaggtgaagagtggtgataagatttaagagt tcatcatcttgaatcaagattaactgttggcaatgctcctaacaatctcataactttgtgatt cgtgtcacacgaaatatcacaacgggttgatagtttaaaagatttaaacgtggaaattgt ataaaactgaatcatgtatgacatctatgaaaaatcaattgattagagatctacgatcaagt attctgaagtatatgtaacgtttggtagaaaaaccaaatttacaatafaa	atcaggaaaaataagtatctgtcataaaata acttgattttctgacttagtcttagcttag caaaagttactttcaaaagtaaaatgagctacg ttatgaaagatatatagttacgtcagaatgtt ctacgttagtctgacctctaaggttatgatta aacagttcacacgtttgaacaggtgttagac aacaacacctttcaataacattgacgaaggt agttttacagttttgagtagtctgcttatagca ataaactgtacttttaaaaggttttaaatga aattt
1796-1797	Ectothiorhodospira magna strain B7-7 genome assembly, contig: Ga0116939_119, whole genome shotgun sequence GENOME_ACESSION: GCA_900110965.1_IMG - taxon_2675903141_annot ated_assembly_genomic GENOME_ID: 184488 CONTIG_ID: 10 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atggcgattctggtactggacaagaggaaaaagccctgatgccctgctcgagaagcg ggcaccggttgcgtgctgagcgggaccgggcccggcgggtgacaccagtag cgtggtgcccacccaatacaatcgaacggcgttggcattcccaagaccacgccctg gatgcagctgtgtcgggccagtcactgccaccaccctggaatgtgcccctctga	gtcaactacccccccctgaaggccggagct tgtgaaagcaagccgggttgaccagggtca gcggtgaaaaccccgctacgtttgcaacagg tcgtcaagaccaccggcggtatgctgcctc agtctgccgctctggaaggttaggatcatgc tggcgaaaggtaaagcggcgaaggttcca gccgcccgtcagcgggagccgggttgacag acattcccaggggagatgagcgcaagc cctgcgtcactggaccctgaagggttctgtt taggagattgaaatggcg
1798-1799	TPA_asm: Euryarchaeota archaeon isolate HyVt- 292 HyVt- 292_k295_371144, whole genome shotgun sequence GENOME_ACESSION: GCA_011040935.1_AS M1104093v1_genomic GENOME_ID: 326302 CONTIG_ID: 122 SOURCE: NCBI_Prokaryotes DATE: 2020-04-22.	atgcagaagttattgtagatttaagaacacatcaggggatgctcctcaagctcccctgct ctgtaaagtgggcattaaacagagaggaaactctcagtgctccctgcatagtagtggcca ataacaactccgatgaggacctacactctggcaagagtagacaggacttgagagttcctg tgtaaatatgcgtggaaaacctttaatgcctatgagacctagaaggaagaggtttcctg aaacaggaaaaaggcagcagtagttcagcgaagctctttatcatagtaaaatccctt caggtgaaacaaaacagctctaaaattaggtatagatgctggatattcaacaataggctt cagtgctataaatgacaagagtgtaattactctcagggtgagtttaactttgcgaaaaagaatct ctaaactccttgagcagaaaaacattacagcggagctcgaaggaataaattatggtaca gaaagccaaaatcaataatcgtagcaaacccaaaggttgggttgcacctctatagacat aaacttgagacacatctgagattgattgaaaactgacgaaaatctgctgtaacaaaaa taacagtgaagatcagcttggaccagagaagatgcagaatccagagattaatggag tagaatacatcaaggtgaattacagagctacgaagtaagaataatctgctggagaat ggaagcataagtgcttattgtgggaagagtaataaccccttagagattgacatattata cctaaaattcagggtggaacagatagagttcaaatctgactttagcgtgccataatgtaa tcagaagaaaggagataaaacagcagcagaggttgatctcctgaaatacagaaaaaa gcaaatcaaacattaaaaacactgcattatgaatactgttagatggagactgtgaatac tctgaagtgtagctggacgtatgggtatattacaagcatgctcgaattaaattagatatgg agaaaaatcatataaattgatctttgtgattgctggtggaacacacaaagtaggagtaaa ccttataaagtgacacaaacagacgaataatagaagatttcaacaaacagaaaaaggt tttaaaccttctatagaaaaacaaagataataaattacagcccaatgatttagtaaatattata aatctctatgtaaaagtgaaagggtgtccataattatgggaataatgtaataagtaagggaa aatagggaaaatcttgatattaatgttaagaagtggaatttaaaatattggaagaaggaat acaattctaa	gtcaatcaccaccattaaaatgagtgccct gtctcgtgagagacaggggttaacagggtga ttaggaggcattgaagaatgcagaagttatt ggtagaggttaagaacacatcaggggatgct cctcaagtcccctgctctgtaagtggggcat aaacagagaggaaactctcagtggtccctg catagtagtggcaataaacctccgatgag gacctacactctggcaagatgacaggac ttgagag

SEQ ID NO:	Name	IscB sequence	ωRNAs
1800-1801	Ga0074646_100108 GENOME_ACESSION: Etolagbewater25m_FD_ \$ F_3300005470 GENOME_ID: 311557 CONTIG_ID: 107 SOURCE: JGI DATE: 2020-04-07.	atgaaatccaaaagtattgttttgcagaaagatggacaccattgatgcctacaactgcac aagggtctagacgaaatfagggtatggtaagctaaagtgtgaaaagagaaccttcac aatcaattgctttacaatcaggatggatgattcaaccaatgaattggggattgattcgg gctatcaaaatagatttttccgcaacacagaaaaagatgaattaatatccggfagaatt aagttggataatggaatgcaaaagagaatccaagacaaagccatatactgtaggaatcgc aggaatagacttagatacagacaatcaagttgataatagacacaaaaagaaaattggt tgccaccatcaattcaagaagatttgataataattagcttattaacaaattgagatcaat acttctatttctgagattattgttgaggcaggtatttgatattcaaaagtacaaaatcctg aaattgaagtaacaataatcaacaaggtgaattgtatggttatgcaaatctaaatcttattc gtttgacaagagaaaaagtttatgtcaattgtgtgtaagttcacaanaagtgccagatg catcatattattccaaggtcctaaaggtggaacaaatagaccnaaagattttgattgtggg cgataatgtcatgataaattgcataaacaatttatatcacaaattaaaaagaaatagac aattaaaggatctactttcatgctatcattagaagagattctatgatttggatataacgta gtttatggttatcagacctgtggatcgcaataaattgagttatcaagagtcagcaaatg atgctttgtaatacaggtgtgataaatcaaaatagatgaatattgtcatagtaactcaaa aaagaaaaataatagattgtctgcaaatatagaagtcgggaattttgattcgtagaag agatatagtattcgacaaaagatttagtaaatattcaggtgaagattgttgaagtgatagg attattagtagggattaaagtgtgtctgacggatggaattaaaaagatttataatccaca tcaaaattggatgttgatcttcatagaagacattgattggagaactgtataggctaa taa	gtcagttgtcaaggcttaataacctgtgtgtt aggtgactaacaacgattaaaagagctgatt agtctaagtcttaattgactacgttatcttga atatataggtacttcagaatgcctcccaagtt ctgaacactacggatgatagttaaacagagt tcaaaagactcgggtctattattaaaaacctaa gaataacattgacgatgggaatcaactccga aaggagatttataatgaattcaaaa
1802-1803	Ga0318466_10004665 GENOME_ACESSION: GoaFecpsamplesv2_ \$F_3 300031555 GENOME_ID: 280328 CONTIG_ID: 4664 SOURCE: JGI DATE: 2019-06-13.	atgatttttagttatgttttgaataaagatagtaaacctttaatgccttcaaaagctcaaaagc aagaaaaattattaaaatcaataaagcgaaagtcataaatagaactccttttactatcaatta ttatatggaaagttcagggttataaacaacctatgataattaggagtgatagacaggaagtaaaaa tatagggttattctgtaacaacaataaacaagaatttttccgctgaagttttattagagat gatattcgtagaaaaattatcgacagagcagagaataagacaaaatagaagaatagaaa aacaaggtatagacagccaagatttaacaatagaattaaatcaagaaaaaagggttggtg ccaccttctcattgacagataaattagtcataattcaataataaataaagtataataaatttt acctgttacagatttaataagaagtggaagtttgacattcaaaaaatcaaaaaatcctaa tattcagggaaatagaataccaacagggagacaattgggattttggaatcacagagaatatt gttttatttagataaccatactgtcaatgctgtaaaagggaatcaaaagataaagttttaa atgttcacatataagagatcgcaaaagttgtgagatgcaccaataaatttattactcttt gtgaaactgtcatacaaaatataaaggagaattaaatattttttaaagaaggtcttc atctttacgagatgtgctttatgaattcgataaaatataatattatataataaactaaaaaaa atttatccaaattgtaaaattctatgggtatatacaaaaatataaaagaataataaattt atctaaaactcatgcagttgatgctttttgtatacaaaataatctgttagcaaatagattaaatt attttatcaatataaattgtatgagaaatcataatagacaaattcataaattgaaaattataaa ggaaattgtagaacaaatagtttaggttaaatatgtacaaggatttcaagcttttgataaa gtaaaaattcttaattcaacataatagatttataaaagcaagaagaaaaacaggtagtttt gtaatatctgatattttggaataatcatataaaataaacttataaaaaattaaaattgttaga aagtagacgaagttatttagataattactaataaaggagaaataa	gaaagttaagattattgtcacccatcatattg tgataacgagacaaaatttcaatttaacat aattgttgattgaagtttagcctcagtaaga attttctttactactgttacttgaatataatggt acctacaatgttaagtcaggtttgtagctcta aggttagtgattaaacatgttttttagacaagt gttgcataatataaaaacctaaagataaattg gcgagggtctacctaactttaaagtaagg attactgttttagtaaatcaatttttttagga aaggagtttatctatgatttta
1804-1805	Ga0182240_101142 GENOME_ACESSION: hydrates4_PW_90_2_ \$F _3300020032 GENOME_ID: 37710 CONTIG_ID: 1141 SOURCE: JGI DATE: 2019-01-27.	atggtatagctactgaattgtgaaggcaaccgccttatgccaaacacggttcggcaagg tacgccgtatgttaaaagataaccgtccagggtgtgaagagctaaacctttaccatcca gttaacttacgagaccaccaactcgtacaaccggtgactcttggcatagacgccggcta cgaacacagtagcgttttgcagtcaccgaaaaaagagagctgattgccgggtgagtgcc agcttctgaccggccaagcagaactgaacaaagagcgcgaatgtaccgccgcgaac gccgcaatagattaagataccgcaagccccggtttgacaaccgcaagaagccccgagg ctggctttcgctagcattcaacacaaagctggacagccatattccgctggtgaacctgtg aaatccatactgtcggtagccagggtattgttgaagttgcccagtttcgacatccaggccat caagaacccgggtatccaggcgaagagtagcagcagggcgagcagtagcgtttctgg aacctcggggaatatactctgaccgggacaaccaccagtgccagaatccagactgca gatccaaaacccccgaacttgagggtacatcatatcgggttactgaaaggtgatagacag accggccgggcaacctgattaccctttgattaaatgccaccgccccgaaaaccacaaa aaaggcgggctactgtgggctgggaacccaagtcacactttttaaggcagaacatttc atgactactgtacgttgaaaatggtcaacatctgggcagcgattatacttaccggttacatt accaagaaaaacgcatggaattaaacctgtccaagagccatattaacgatgcattgttat cgccgggtgtactgctcaaacccgggtacaaaccgctgactataatccaagtgcgccga acaaccgtagcttgcagaattttacgatgcaaaatcattgacatccgcactgccaagaa agcaaccggccaggacttgaactgtggaagcgtaccagaacccgtaaccttaacggctc cgaacctgcgcatttatctgtgccaaaaattatcgaaggcagggtagcaggtacgtagaa aacggcatcgttccagcctgtgtgatacagtaattttcagggcaagaagtacactgtcaa aggcaccagaaccgcggagattactgtcggctggcggagctaccacccaagccgggtcaa ggctggtttactcaatcatctgtattatggcaaggactcgggtgatttag	gtcaactatccccacttaacggctaacgccg tttgaagtgggggcttcgagaagtatagcaa agatatgtccatccctgcaagcccggttgat tagcctcagccaccagcgaagagctgacag ggctactgtaccgtgaatatatagccacttc cgggtactccacagctccgaacctgcgg ccagcgggttaaccatcgctacgggtaggc gaagtgtcgtcggcgtaaaaccaggggata acattggcgaagtggaacacagccacaaa ggctgatttatctcgtcaaggagattcaaaa atg
1806-1807	Ga0310695_10004545 GENOME_ACESSION: RumRJGDNAv2_ \$F_330	atgcctacagaagaccatcgtaaagtaagggtgcttctaaatcaggtcaggcaagggttg tccagagaacacattacaattcagttacttcaactactcaggtttataaacagatattaa tttaggcgtggatagctgttagtaagtaattgtttatcagcaactacttctaagaaagatt	ggaacatcaataacctataaacttttagtttaa gaagaatttcttattgattacccttaagctgg tttaccaggaaactactgttatgaaagtatga

SEQ ID NO:	Name	IscB sequence	ωRNAs
	0032007 GENOME_ID: 280889 CONTIG_ID: 4544 SOURCE: JGI DATE: 2019-06-13.	gtttgctgaggaaagtactattagaaatgatattacagaattgttatctcttagaaaaatgtatc gtaaaaatagacgatggagaactactagatagaaaaaacgggttttaaatagagttaa agtgatgaaaaaagggtgggtagcgccatctattagagctaaacttgaatacatttaaatat aaitaagaaagtatacaagatattaccaataactaaattaatagtagaagttgcttcatttgat atgcaaaaaatcacaaatccgaaatfgaaggatagaaatcacgaagggtcacagtttg gcttttggaaatgtgagagagtattgtattgcatagagataatcacgaatgtcaatattgtcatg gtaaaagtaagacaagggtgttaaatgtacatcatattgtacaagaaaaactggaggga attctccaagcaatttaattctttatgtagaactgtgcatcaaaagtatcatagcggagaat taaatfaaggtaactaaaccaaactcttfaaaagatgcagcatttatgaatataatgagatg gaagttatataatagctttaaagaaatataataatgtgtatagacatttggatataattaca aaaacattcgtattgaaaaatagcttccaaagaacattatataatagctagatgtatttctgg aaatccaaaatctcttcattaggattttatattgggcaagctagttcgtagacataatagg caactcataaagcaacaatcgggaaaaaggaatcgttaagagtaaatcaatcgccaaaa tatgttttgggtaccattgtatgataaagtataattgaaaggtcaagttgtttatatttcta gaagaacaaagggtactttgatataagacatattgtagggactagagtaactgcaagtat aacatataaaaaaatgaagctatttgaaaaaagaaaaacgttgttcttgagattgttaaag tttaa	taggcacctacggatgttcctcctagtctgta gctctcgcgctaatgattaaacagctcgtag aggttaggggaggtgtgttagcatataaac ttctaataacattgggttaagggaactaacgt ttactttgtaaacgagttactcctcgttagt gactacttttaagtaagaaggagtaacttta cttg
1808-1809	Ga0310138_000236 GENOME_ACESSION: M273_2_\$F_3300033986 GENOME_ID: 280650 CONTIG_ID: 235 SOURCE: JGI DATE: 2019-06-13.	atgggtatgtatttcaaaaggtgtaaaccttaattgccaacaaaaaggcatggtaaaagt tagaagacttttaaaacaaaggtcttgcataaagttgttagaagagaacctttacaattcaact attgtatgatactacaacatatacacaaacctgtgttagttggcgtagataatcgttcactgcat aaagcaaatccaattaaaggtggaagaaagaccagtaaatactgtaaaggaagttagagg gtttgaagattcgacaagtaagatataaaaaatcaaatgtgtataatctacgggttaagaa tctctggatactttgatagcgttcatgaagggtgagaaaaaattcacatcagttaaagtggtc aaacttaaaacttttagaaaaagcaaaaacactaattatagaaaggagggaacagcgcat tctctccatcttaaagaagatggagtcctcgtcgtgcttattga	gtcaactaccaccacctatagagggtggag gcttgcaaaagccttagtgactaccctcagc cagggggaagttaattctcctatcgggctacg ttagactggcatgacacctgggatgctgc tcaagttccaggctctgtcgtatgtacttaa cagtcctgaggggtagggacagtggtgcat gcataacaagcagcttaacattggggatga gcgcctaactcctaaccaaggagggttac cgcatatg
1810-1811	TB_LI09_3DRAFT_1004762 GENOME_ACESSION: UncGr_3300000229_\$F_3300000229 GENOME_ID: 36368 CONTIG_ID: 4761 SOURCE: JGI DATE: 2019-01-27.	atggtttatgttttcaaacgatggaaaacctttaattgcctactcaaggcatagaaaagtc agattgtggttaaaagagggaagccagagtcataagacgcttccgtttactatccaac tattgttgaccttgttgcacaaaaaacacaaaatttaacgttaggttagacgtcggatttaa aacgttagcggtcagcgtgtttctgacagggtggaagcgttttagcggcgaattcaactc cgaaacgacgtatctgccaacatgacagaacgccgtatgtatagacgtataatcgccgtaac agggttagggcacagaaaacaaagatttcttaaccgttaacaaaaagcgggttttagcgctg gcgttaaaacaaaaatcgattctatttcaatttaattgccttattaaatcatattgccatta caaaagtcacgttgaaacctgctcgtttgatccccataagctaaaaaacctcatgttcaa ggaaaggattatcagcaaggagaacaatatggctatgaaatgtcaaaagcttatgtgttag ccagagacgggtatcaatgtcaggcaaatcaaaaagggcataagtcctatcctgaacgttc atcatattcaatcgagagacaagggggaagcgataatcgggataatttaataacctttg taaaaagcatcatgaacaattgcacgatggtaaaatcaggctacatgttaaaagggttaa aatcttaaaagccgcgacagcgatgaatatcgttagaagccaattgttaaaaaaatgcct gaagcgattgaaacgtttgttcttaccaaagcaaaacgtcaggaacaaaagcttcag aagagtcacgaactgacgcgtttatcattgccggcgttaattggacaaccgagactcaac ctgttggaattgtttgaaagcgttaaaataatcgatcattgcaaaagaacagaaaagggttt agtccctctattcaggttcaacgatatgcaatcacgccttatgatttagtatttcaaggca agcgatataagcgataggaattcagaacaaggggcttacttaaaagatgacaacaggt attcataattagtgaaaaatgtaaaacaataagattatctatcatcaaaagactttgttt gtgtgtga	gtcaatcacccttaagctaaagacttagggg cttgtaaggcgactacaggatagctcaaaa aagagtaaatgttgattagggggcgaaa ggcagaagttacggagtgatagatcgcac accgaagatgctcctcaagttttggctctgt gattgtctttaaagagagagaaactcag tgaggcagatttaaaacctccggttaacaac cccgaaaggaatcagagtttgacgttct cggcttacagcgctaaaaactgtcgtttctct taaaaaaaggattttgatcataatg
1812-1813	Ga0180007_10021296 GENOME_ACESSION: MM_PC_MetaG_2_\$F_3300014656 GENOME_ID: 32273 CONTIG_ID: 21295 SOURCE: JGI DATE: 2019-01-27.	atgacagattctcagaacttgagtattttaaataaataatgtagaaaaggttaattatctta aatctcaaaatgtgtcacaacatcaagaacaaagaaagctaaaaaggatataaatcagaaa aaatatctgtcaaaaataatcaaaacaggacaggtgtgtgtccatgtattgaacatgagagg taaacctttgatgccaacgtctccaagaaaagcaagacttctattaaagaataaaaaagca aaagtgtgtcagagaagtcatttaccattcagttgaaatattccaactggagaaaacaaac aaaccaatcaaatagggttagatagcaggttataaattcgtcggatttccggcgtgacaaat aaaagggaattgatacaggagaagtgattattaggactgatattcctgaaaagatgactg aaaagaagatgtatagaagggaagcggaaatagaataactagatagagaacctaga tttgaataatagaggtatcccagaaggatggcttccatcaacacacataaagttagata cctatgttagattgtagtaagtttgacaaaatattacctataacgcatatcaatgtcgaat agctccattcgatactcaaaagatgcagaatccagagatactgggattgaaatcaacat ggagagttgcaaggatagaggttaaggaataatctacttgaaaaatggggaaggaaatgt gcatattgtaaagcagagaatgtatctttgagactgaacatataattccaacatcaagagg tggaacaaatagagtcctaaatctgacgatagcatgtcatgaatgtaatcaaaagaaggt aatatgacagcagcagaatttggccatccagagattcaaatcctagcaagacaatcattga aagctgcgacatttatgaatgttggaggcgaagattagttgataatataaagaaatcgttc cagatctatattgtataatacatatggatataattacaaatatacagagaataaaattagatt agataaatctcatgctaatgtatgtcgtatgacacattccaatagtaggatgattttgg tcagataagatcaaaacctatcagggtcaaacaaatagaagaacaaataggtctctacag	gtcaactaccccgccctgaaggcgggggc ctgtcagtgatgacaagagcaatagttgatta ggaggtatacctcaaaaatgacagacgtta gaagttagaataacatacactcggccgatg ctccactagtcgggggaatcgtgatcctgt attaaataaagattcaagaaggcagccaaa gctgtcttcagggtctatgtgtaggattttaa aacttcttaactcctcgaagtggaactactc agcaatgagaagacaggactcagaggtctc tgtagggtgatta

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ctaaatagaaaaggattcaaacctgctgataaagaagaagagatataaatattctcctggag atttgataaagagaagatctgataatcaaacactggatggggcgagaaaaagataaaa aaagatagatcgatgatgtacactgtcaagggaatttcaattatgttaaatggattagattg gctaactctatttcaggcgaaaaagatatattatctctattgaagatgtgaacattgaaat atgggaagcgggtatttatttcaattagcaaaatcaaatgaacataaaatagagaagaagga gaagggtatgaaattgaataagaagaacagaggattatagatatgaaggagcagaaatc aatagatgatacatggagttaa	
1814-1815	Ga0310133_003358 GENOME_ACESSION: M171_2_\$F_3300033998 GENOME_ID: 280645 CONTIG_ID: 3357 SOURCE: JGI DATE: 2019-06-13.	atgtcgtgcaaaccttcaaaagcaaggaaactgcttaagcaaggcaaaagcaagatagt caaatatgaaccattcaaatccaaactgtttatcggcagtagtggtataacaatgtgtgac tgctggtattgacgtggcagtaaaaacataggcatagcagtaacaacagatgtggcag aataatttataaagcacaagtaataatgaacagacatcaagaaatattgaacaaaa cgtagacttagaagaagcagaagaacagaaaaacacgttacagaaaaccaagatttct aacggcaaaagaaaaatgggtggctccaccatcctaacagcaagaattgatgcacat tacaattatataaaaaagttaataaaattataccatttacaacacattgtagaagtaggac aatttgatacacaggcattataaatccaaatatacaaggtaagaatatcaaatggagat atgaaagggtttgacagtggttaaggaaatcgtaaaaataagagacaattatcaatgtcattat gctaaattgaagcctgatataaccatgttcaggcaaaatgacagtagatcataatcctaa aagtaaggcggaacagataatccaaactatctgtatgtgtcgtcggaagaacacaacag acagaaagacaatctattgtacaagaatttactggcaaaatccgccaagcaattaggaat tttaaggctactgctttatgaatgtcttaagggaattatctgttctaaattacaggaaatagc accaacagataatacttttgactatatacccgagaaaaacgcaaaagaatggaatttagaa aaatcacacataaatgatgctattgtctattgtgggataaaaccaagacaggaggtatcag tcagtattatatacgaagcttcttaaaagaacgcagtttcatgaagaattccacgt aaagggaagtaaacctaatagagatgctaaaaagaacgaaaaaacataaaagaaaa tcataactaataaactgctggtgtttgtgggataaagtatatatacctcaatagataaaat aggctatatacagggtttacaggaaaaatgggtatatgtacaggacatagaaggtaattatt acgaatatcagaaaaatataaacaatcaatccgaaagaattgcagctaattgtagaat aacaattatattagtcagcaattcatctccacctatag	gtcaactaccccccttataagcggagg cttgaaaaagccttgggtgactagcctcagc caccagcagtagtctgacagggtacgtta gacaggttatgacacctgaaatgatgtca agtttcaagctctgtcgtacatgacctaaaca gttctgtggtaggaacagtggctgtacgt gtaagctgtctaacattggcgatgggcaaa taactctgaaggagggaacactttatgtta
1816-1817	human gut metagenome genome assembly, contig: NODE_78_length_10418 4_cov_6.349394, whole genome shotgun sequence GENOME_ACESSION: OWCF01.1 GENOME_ID: 11873 CONTIG_ID: 77 SOURCE: WGS DATE: 2019-01-27.	atgcaatatttagattttacttttagtagtagataaaaaataaaacatgtgtaccgattttaa cggtagagctgggtatttcttagaataataaagcaaaaattattatcatgatccattagta ataaaacgaatagacgattataagagtgattttgaaaaataggagataatttgcagttaaaaat gatagcggatatttgaaatagattttctgtcagtgataattatcatgagtatttgctggac aagtcgaattattaaaaggaaatgtcggatagattaaacaaatgaaatggatcgaagaac acgaagatccgaattaggtacaggaataaataaaatgttgattacaaaaccgtacataat ccaacataaaaaatgaaatgaagaaggatgttgcctccatcaatacaacataaaatg attcacataatcgtttatagataaaatgctcatggttccagtagataaaggtaattgtaga agttgctaaattgataccagatgataaaagctttgactgcagggaagaataatcaggga aaagattatcaaaatggagaatgaaaggatgaaatgcaagcagcgttatgttagagata gagataaacataatgtcgggtgtgtggcgcaaaataaaatgttggattgaagttcatat atacaaccggcttcaaggagggaacccgataaaacgaagtaataatcttattgtcata gttgcacgcgaagggtgcatccaataaacgacaataaatattttgagaaagtaagagt atgaagttatcggtacatacaaaagacagttatataatgattgttgggaactttttg aaaggctttctggcaaatatgacgtcaaaatgggtatggatatacaaaaaaataataga aggaatggcgggtttaaagaaatttattatcggatgctgtttgtattatgattacaaggatg tgacactaacggagaatatctattgtatgataaaaaacgatgcaatgacaggagtagtga gacatttagtgatcaaaatacatatgtgtacgtgatggaaaaagaaaaagtggaataac attatataaggaaggcttccaaatgctccgtctaacagagtcacgcaaaaagaaatata aacaatatgagacaatttctgtgtaagaaaattaaactggtaaacgcacttttgtttgaatt cataattgttgaattgtggagatttaattacataaataagtgaaagcagtagaggaaatatcg cagaagtagaatccatgcaaaaactacctaattgtaatttcaaaatcagattatataaaa gcacaacagtcgaatacccttctataagataaaagccagaagaatatgaattattaaaga ataatttattagacaaagtaaaaaattgaagaacaggcgtggaatgatttggagaaaaat aatcgtctagaataacgaagcgacccatgcagatcaagaaggaaatggctgtataa	gcaaatcctcaaaaataaaactctataaaa ttaagcgtttaataaaaaatagacttagtta atctacgttattacaagaaatattgatatgt aactctagatcaaaattatctgttagatttaa acatgcgttacgtcctaatttctgtcgaagt gatttaagcataaaaactgtataacattgtc aaagagtattaccgactataaagtcaggttatt aatttaaaaggagaaaattatgaattattaga tttt
1818-1819	Ga0082212_10029072 GENOME_ACESSION: TergutNt197P3_\$F_3300 006045 GENOME_ID: 36056 CONTIG_ID: 29071 SOURCE: JGI DATE: 2019-01-27.	ttgagtaaaagtaaaagacacacctgaggatgcttccccagtcaggcgtctagaagg aagagtagcagacaagctagtgtgaagcacgaaacggatttccacaacgcgaagc gagcagctgatactcaacattggcgagaggagactggccgaaaggctagcgtcacaga gttctgactctgcccagtgagggtgacctcactgcccgaagggaacaaacacccctcc aaaactgcgtagcagattggacaaaaaaggagccactacagccaacctcaccgcg ccgagcacgcctcctgctcaaaaaggagagcagcagtcaccgctctacccttca ccatcagactagtggaccgctcccaagaagaatcaacattgacggagtggaactcaaa gttgaccaggcgcaaaaaccggcgttagcaatcaccgtgaccctaattgacgggtga aacaagaagccattctctatcagaattagttcacacaggatggctcattaaaaagaatctgc agaaaagggtgcttctgctggcgaggccgcagaagcagaaaactaagataccgggtc cacgttggcataatcgagcagggcgtcctgtcgtggattagacgtcgtgctggcccaa gtattcggagtcgtctcttgcctactttaaattgacaagaagtgggccaaactcttccca	gtcaactaccacacctaaggaggtggcc tgcaagtgaatcagcttctgattcaccaaaagc caccgctaagtaagtgaattgcgaaagca atcacgggtgaccagcctaagacacccctg aaggaggtgactacgttgagtaaaagttaa gagacctaccctgaggatgcttccccagctcc caggctctagaaggagagtagcagaca

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ttacttgcgtgtgattgaaatgtaagtttgacatgcagagactacgtaactcctgcaatttc tgggaaagagtatcagcaggcactcttgcgtgttacgagttgaaagagtacttactgga aaagtttaatcgtagttgcttattgtgacaagagtggataaccgtaaatgtagatcatg tggtaaccagggtctcgtgggggtactgatgctgttagcaatctgttttggcttgtgttaagt gcaatcagaagaagagcgctaaaggttagatgaattttggcggagaataaacaagat tggctaaagattaagcggcagttgaaagtaccgctgcgcgatgctgcagcggttaattgtg gctcgaaggtttacatcgcattgggtgagcgtgttttattgaaagcttggtagtgggtg ctcaaaactaagttgaatcgaagcagtttggattcctaagtcgcatgcttggacgcttgt gtgttggcgtgacattaaagtgaagtgacgagatctcagttcaatgttggcattactgtc aaggtcgagggtgctgcagcgcgagttgtagataagtttggtttctcgtggaaaacc aaagcctagggtcgaagcaggttttgggttctactgggtgacttgggtcagcggtagttcc taagggtaaaaaacagggtgagtttccggcagggttgcgtgctagcgggtgtaatttt aatataaagactttgaacgggtactgtcagggtattaaagatttcattgtcagctcgtgcaa cgagctgatgggtatggtatcagtttagtaaaatataa	
1820-1821	Ga0247608_10014894 GENOME_ACESSION: 176DNAGHglowgp2_2_-\$F_3300028805 GENOME_ID: 279762 CONTIG_ID: 14891 SOURCE: JGI DATE: 2019-06-13.	atgaaagttatgtattagttgaaaggaaaaccaattatgcctacctctagaggaggtag agtttaggtggcttttaaaacaggaaaagccaaagtcgtgagaactattccccttactattc aattatttgaaagggaaggaaatactatacgaagaataatctaggtataaaacccgg atctaaagaattgggatactcgtacttcagaaaaggagagctgttctgtgcatctgtaa aattgagaacagatagataaattcctactactagaagaactaggagaactagacg aagtagacttagatagagagctaggttcttaaatagagtttaaaacaaaaaacctgggt ggatagctccctctatcagaataaaatcaatttcatgcaagatagtagaatttgaataaa gattttacctatactaatatgatacttagatagctactttgatttcagaaataaaaaatc ctgattataatggagtacagatcagaatggaccacagaaggttctggaaatgtacgtga atatgtcttatttcgagataatcatgtctgtcaaatgttcttggaaaactcggagatgaatcc tcaatgtacatcatattgtatcaagaaaattgtgtgaatttccaggaaaacctgattacttt atgtaagacttgcataaagactatcatgaaggtaagtaagttatccattaaaaagagta ggctcttaaaagatttaacatgctcttaataataaagagatagattataaagagttattcaag agatatggatttatatactattgtgttaccaaactaaatgtgacgtatatactaaagggtta ttgaaaagtagtgatacagatgcttattgtattgttcaggtgatacattcctttattgagtg taccagatattgtattatcagggttcgacgacataatagacaatacataagtttaaaattctt aagggtggaagctcaaaaagaatcaagcttaggtataaagatttggatagctgtcttaag atatagttaaagtacaaagggaatcgtattataggtgtcggcgtgagagaggatcttc aatattaaatcattagaggagataaagaatttagatattataaagaatttaaaattcttat atgaacctagacgaatatttatgtataatcaagaagaaggttactccaattcctccacg gctaaagcagtggtgacccctggaggtgaa	aaaaatcagtaaccacgactgaacgatat tataagaaaatattctgtcgaggcatgtaa atgccactatgctgattagctaaagtgaagtc tctatgtattatagaattatagaagaagatga actacgttatgtgtgaatgaatctcgagatct cgtatttctcggattatacaataggctacacg gatatttctcgaactctgtagctctacggttatt gattaaacagctctgttgggtagagcagtg ttgatgacaacaaaccatacaataaactga cgatgggagtagtactttaaagtattatgac ataataatgattagtaagtataaaggtgtctt attgttaaacatttttaaaataaaatgaaa
1822-1823	Ga0373631_0015577 GENOME_ACESSION: RO123_strea_MG_FD_\$F_3300035180 GENOME_ID: 312623 CONTIG_ID: 15576 SOURCE: JGI DATE: 2020-04-07.	atgcataatagaatattgttctgtcgtcaataggagagacctaatgcttccatccagca agagcaagggtctcctgagagacagataaagcaaaagtgtgaaggtttatcctttacaa ttaagcttacagagagaaccacaggtgaattgcaacctgttcaattgaaaattgaccacgg ttcaagacacagtggttagctctgttgaacttggggaacagcacctgaagggtatcttcg gtgcagttctacatcaaaagggtcattgtattgaacaaatcttctgagctgtcgtgagcctgc gtcgtggagtaggcagcgtgaagacaagatacagacagcgtcgtgggctaataaggaa aagggtgatgggtgttaccacctcagctatgagtagaagtaataatctgaaggttggga cacagaattcagtaggttaaacacgggtgattcaatttctgttgaagacgaaattgtat actcacttattgttaaccctgaagtgtcgtgttccagatcaacagggaacattagaag gttataccgtaagggaatatctctggagaacacataagaacttgtgtctattgtgggct aaaaatgtcccacttcagatagagcatatccatccacgttctcgtgtgttcaaatgctata tctaacttactctatcttgtgtccatgcaatcagagggaagggtgactagacattggaggga gttcttctcgtgaagcctgagcttcttctgaagataaagggtcagacaagaagagcctttg cggatgctgcacaagtgaagcaatccgcaataagctcttgaagttctatgtgactttggt ttacctgttgaaagtacacaggagcagaacaaagtttaacagaactagactaggttatg gtaaaagaacattggaattgagcagcctgtataggctatcaggacaacttgtttcaatccac agacctgatagaattatgttttagacataaaggctatgggtcgtggatgtagaatgttctt gcatgtgataaattggttcttctcaagaagaacctaagactcggaaacgagtccttg gttttgaaactggagattacattaaagactgtattaaaggtaaaagccttcaaggtcagaatg agcttgaaggtttcattatgtagagcagattttgatggaacaggcaaaagccttaaaagact gtaggctaattcagaagaatgatggatacagttacatcacttacttaaa	atgaactacgatttataagaatcttctagcaac ttgtgctattggaatagttaccgtgggtgtg ctctccagccccacgctctaaggttaggtgt taaacatttgtgaacactgcaggaatgcat ctaacgacaacaccttataaatcattgtcagg agactgtccgtaaggacaataaccgacata ggatttccattgttgagaattcaaaaggt acttt
1824-1825	Ga0194138_10000007 GENOME_ACESSION: TAR4v2_3_\$F_3300018 405 GENOME_ID: 35910 CONTIG_ID: 6 SOURCE: JGI DATE: 2019-01-27.	atgcctgtcatccggctagagctagagaacttctaggttaaaggtaaaagcagcggtatttc gacgcttccctttaccattatcttaaaagaccgggtgggtccatccctcaacctacggca atcaagatcgtatccggcgacgaagaccggggtcgcagtcactgtcaaaaggttaaac gcggcgacggtgtatcatgggttccatctggagaccgcgggtatcaacaatcaagaag aatctgtcagtcgtgctacgtccgacgtgaacgtcgtaatcgtgaagacgcgttaccgtga aaccccgatttctcaaccgaactcgaaccggcggtgggtcccccagctctaatgcac cgggtgctgacgactctgacctgggtgaacgactctgtctagcagatcctgtctccaag atcgagatcgaacatgtgtcctttgacaccagaagatgctcaacgggtcgaatcaaaagc aaagggtatcagcaaggctacactcgaaggctacaagatccgtgagttacttctatcgg	gtaaccgccccatcctttaaccgaagata atctgctggtatagaggtttagttgaccagac ccagggtttgtaatagaacctacgttaccgaag aatacatagtagtactccggatggccgagcc agtcccggaactctacgccgagctgttaaca agcctgtgaggcagggttagtgcagtcagg atacaaaaccttgggataacatctgtaggc tcagttacaccgacttagatcgatagacta

SEQ ID NO:	Name	IscB sequence	ωRNAs
		tacaactacatctgtcagttatgttcgggtatctccaaggataaacgtctggagactgaaca tgtactcccaaggctcaagggtgatcgaactcgtacacacacgtgacctgtctgccat acctgtaacgcgcacaaggccaatcgactcctgtccagtgagggaatggccctcaagg caaaggggatgtgtcagtcagcgcagacgcaaaggctgtcgcggagttgcccaagg aagaaccacctctcagagatgcagcagccgtcaacagcacagctaaacggtgatcca agaagtcctgtgaaatcggactcctcgtgtagaacgtccaagctacatgacgaatacaa ccgtcatcgtcaagggtatcaaaggatcactttatcgtacagcggtagctgggtgtgcc tgtgacattgtatacattccgaagggtgagaccgtgactgcaaaagcattgggacatg gttcgagacagatgtgtcgggtatcagtagtggttccctagaaccttagctaaaggctct tctatgtctcgcgataaccggaccggcgatagtcgaagccaaagtaccagaggttaa atatacgggtgtgtatgtagacgtgtgctgtgaagaagtagtgcccaattgtcattaca ctgacgttggttaaaatcaccaccgtcaacgtcacgtcagtcggttacagcacaatgacg gttaccgctacgcgcatcaaccacataa	aggggaaacctaatgaaggtaacctac taactat
1826-1827	Ga0102947_1000378 GENOME_ACESSION: SalPon2A_A_D2_MG_2 _SF_3300007619 GENOME_ID: 35439 CONTIG_ID: 377 SOURCE: JGI DATE: 2019-01-27.	atgataccagtaataagataaacaccaattacctttaatgcctacaacagaaaaaggcga gaaagtaattgagacaagtaagcaaccccttttggaaaaagggaattttgtattagat taattttgagcctctagcagaacaacacagcaaatgtgttggaattgactcaggctca aaaagagaagcttatacagataagaagtaaaaacatcgttataaacattcaaacagaaa cgctctgttgggtaaaaagaaagttgaaacaaggcgaaatgccagaaaagcaagacg gttccgaaaaacacatgtagaaaacctagggtcaatacaaaatcaaaacctaaattac ccaagcacaaaatcaagatggcaatfaaaatgaagtttagattttcttccaaggcttatc caataactgatgtgtttttgaagacatcaagcaacacaaaaagggaagaaaagttg gaatttaatttctccattgcaagttggaaaaaatggttttgtgaacaagtaaaaacaaa ttataaattactattaaagaaggttgggaacgctgtgtagaaaagaactaggtttag aaaaaattaaagaaaagttgtccgatcgtttgagggcacattgtgttgacagctttacatt caagttttctatttttagataaaaaagcaaccagaaaaacaaaagtatttatactaaaac tcaatttaacagaagacaacttcattgttttaattttcaaaaggaggaagaaaacattac ggaggaacaatgtcgttagattcaagcgtgggaagtattgtaaaacatcccaatacaac attgtttatgttgggggaagttcaaaaggaagataagctacataatttagaaactggcaa aagactttgtcaaaatcctaaattgaagatttgaatttttgagctataatgtttccagataa ggagggaagccattcctccacagacgagcctgtgggtatcctggcttgagattaa	gtcaactacccccatgacaagcatggggc ttaaaaggagtacaaaagcaattggaac aattagactaaacagttgactagactacg aatgtacgtttgacaagtcacaaatgaatca actctgtgtggcacagaaaaatgtctcgg atgcttcccaagtcggaccttftgtgtatca gtgtcgaagggaatatacagaagcccttta cgggcatftttataagaacatgatacca
1828-1829	Ga0210402_10005866 GENOME_ACESSION: IncBWC7MmetaG_2_SF _3300021478 GENOME_ID: 30130 CONTIG_ID: 5865 SOURCE: JGI DATE: 2019-01-27.	atgccgtgcaatgaacgtcgacacggctgtgtttagaacgcggctgggcacaggtaca tttgcggtatcattacgatacgggttagtgcgtgcggctgactcgcactacaac cgattaccgcgaagctcgaccaggttcgaaaaagaccgggttggttagtgcggcca tcactcgtactcaatcggtaaacgttttaaatctaatagaatagagcatcgtgttcgcaa attccgacaaccttacgaaacgcctggccaccgtgaaacatcggcgtagtaaaactcaatt atcgtgccccccgttttaacaaccgaaccgcgtcgaaaggatggctagcaccctcactac aacatcgggtagatggcatcattattgtctagccaaattacgtaaactggccctattacg gatttagctcaagaactcgtcgtttcgatcgaattaatgcaaaactcgtatattaaaggt atcgaatatacaaacggcacgttacaggggttacgaaatcagagatgtgtctagaaaaa tggggacggcagtgctgtgtactgtaataagcgcagcgcatactaaatttggatcatafcg tgcctaagtcgccaggcggctctaatcgcctgccaacctagaccagcgtgtattaaatg taatccatcaagggtatcgttcgatagaggattttctgcccattgaccagcacgattaa aacatatacttaattttgctaagacaccgctcaagacgccgctgtgttaacactacacga tgggcattatggcagcgtttgacggcgactagcttatttggtagtaggcaccggagg cgaccaaagtacaaccgacaccagtagcggcgtggcctaatacatgtaaatgatgcgggt gtgcgtgggaccaatggataatgtcaccaaaatcaccggattaaatcggcccttgttagct atcaaatgtatggggcgtgtgttcctatcgtcgcactcgttcgatgcacacgggtttccgt gggataattgcatcgggacgaagcgaagtatatggatttggtagcgcgatatagtagtgc ggtagtacctaaaggccgttaccaggaacgtatataaggacgagcagctataagagcaa atggcttattgataattaaacgttcaatggattaaaatttgatacatcttataaaaattgcact tgcctcaagaaatgatgataataatgatgataataa	atcaacgaccccgctttaaaccggctta ccttaataagtaagcgtgtgaacaggatt agctcgaagggtcagctgtgacggaggc aaaaactcaaccagtcagaatgcttccag ttctgaccactgaaatctcaatagtagaca gctacaggggtatgcaggaacggattgaga tttagcggggaaccaagcgaagcgggt gaaccgtaaaaccgctcgacaactttccga ggggaggcttcgaagaggccgtaactag gcccgtgaaggcaccagaaggcaaaaaa atg
1830-1831	Ga0136175_10101454 GENOME_ACESSION: IMG_3300010163_SF_3 300010163 GENOME_ID: 311728 CONTIG_ID: 101453 SOURCE: JGI DATE: 2020-04-07.	atgggtgtattgtataataaaaaacggaaaggcattaatgcccctgttcgctagaaaagc aagaattttactaaagaaaaagaaagctcaattttatgctacaagccatttacaataaaat gctttacggaagctacgggttacactcaacctacaagaataggggttgaattaggtgctaaa tgtgttggaaatagctatactcaagatgaaaacataatagcaaaagggtgaattacgctaag aagtgtatgtaaaacaaatattcaaacacgcaagctataccgtcgcagccgaagaaca gaaaaacgaggtacagagaagagagattctcaacaggattaaagtcagaaaaggacgg ctggttacctcaagatataagaagcaggattgagaacactttatgttgatagacaaattca aaaagcgttccgaatccagtgcttagattgagcttggcaaaaacagaatgaatttttag aaaacaggtactactttatgcattataaggaaacgcataataaggagataccctcat tctgaatttatacacaaggatttgactacgctcagaagaatagaattagacctagagaaga cttattatacaacgccatagctataagtggtgtctgagaaaaataagaagaataaaaggaa tatattcaagattgtgcagttcagaaagaaaaaaagagctgtgcaggaacaaacctag aaaggggaagaaagacaagaacgtattgtcaagagaagcgaagaaacaaaaagca aattaaaaactgggtccctaatgatactgaagagtattcggcaaggttgatttatatgcgg	gttgattttccgtagtccctagcagtagacg ccttaaaaggcttaaccaactagcctaagttaa ttattatgctaactcgttttaattagctacgtt attttcatacagaacccgtggaatgataccct agtttgcagctctttcgggctctgtaaacag ctctgagggaaaggagcagtcacccgtgttt aaaggttaataacattggcgaagggttaaat aacctaacaggagggaatacattatggtt

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ttttccggtggagtgctgttatgtcaaaagacataagcgggaaatatatgttcctgaag gcaaaaatacaagcagataaattgaccagtttaagctaattgtaagaacaacaactg gcaatagcgttgctaccgcttagtcaataa	
1832-1833	human oral metagenome genome assembly, contig: NODE_46_length_12485_5_cov_3.86146, whole genome shotgun sequence GENOME_ACESSION: UPCW01.1 GENOME_ID: 16407 CONTIG_ID: 45 SOURCE: WGS DATE: 2019-01-27.	atgggttttagtaattgataagcgtgaagaacctgttaatactataagtgagcttacgctcga atattactatttaacaacaagcggtaattcataagagattccgttcactataagattaaaaa acgtagatgctgtattgaaagatagacctaactgttaaggttgacctgttcaagaac cactggaatagctattgtgataacaaggattcagtggttatgctagctgaaattgaacata gaggtcatattattaaaaaagaacatggatagtagaaggactcttaggcgtcatcgtagaca aagaaagactagatataggcctgcaagattctcaatagaccacaaactaaaggatgggtt agctccgagtggtgaagtctagagctgataatgtgataaattttattagaataataaagaatt ttgtaattataaagaattatgattgagatgttgatctcgatgactcaaatgactgcaaac aataacctgttaggaactgcttatcagcaaggctcttataatcaaaacaacattagaagcttc atatttaacctgtcaaatggttaattgtttactgcggagccaaagcaaccgagatagacc atactgaccaagagctaaaggcggagctaacagcgtttacaatctagtagcatcttcgag aagctgcaatgagaaaaaacaactatcactaaaagcctttgttaaatatgggttaaag acttttagtaagttagaacctaaaaaattgcctaaaggatgcaagtattgttcaagcagctaga aacttatattgttcaagaatatgctaaaaactgttcctagcaccgtagggttatgactgtgctta ctaaatcaatagaatgagctcgtttacctaagaacattattacgatgccttatctgttg gtgaaattccatctaaattcaatttccactactgataaggtattacaatatctgctcaagctag aggctctagacagatgtgtagggtggacaagtacggattccacgaacatctgctaaagc ttcaaaagtcagtttatgtttccagactggagatattggttaagccatagttccaaatggatt aaaaaaggagaaataacctgtgtagagttgctgccgctcaatggttaactttgatataaaag caaaacaaaaaactatccaagtataagacataaatattttcacataattcaaaatgtgatg gataattatagttataatgagcgtatctag	gtcaataacttgtaactagttttactaagctt ggttgaccagacttaggttagataacagata atgttctttaaactcagatagatagttatca caccatagctagctctccagctctgtggctc tgttaggcctctgtaagttgggttaaagccct gtcaaccttagaacggctctagcaagccat ctatcattgtcggaggagactaacttagaa ctttaattgggagcattagctctgtgtaaatag ctgtaagttagcgtaacggctgcaagactag attaactaaaaaggaaaaattatg
1834-1835	human gut metagenome genome assembly, contig: NODE_1068_length_23742_cov_9.115844, whole genome shotgun sequence GENOME_ACESSION: UPOJ01.1 GENOME_ID: 16694 CONTIG_ID: 1067 SOURCE: WGS DATE: 2019-01-27.	atgggtatattgtttaaatatcaatgtgttccacttatgccaaactgataggcagcgttaagta agaaaactactaaaagctaaaaaggtcaagtagtcaaaagaacaccatttacaattcaat tactatatgtttcaacggaaatatacacaacacatttaactttggcgtggatgcaggctcga catattggaatatctgttacatcaagtaacaagaggtttattgtggtcaggttgaaacctgt caggatgtatctaaacttttagatgcacgaagaaaatcaggcgttcaagacgaacacatt taaggatagacagccgaggtcctaataaccgagtggtcgtatcaacacaaagggttggttag caccgtctgtcgaatgtaaaatcaacacacataatfagggttaagataagataattacaataa ctgcctattactaataatagctgttgaaattgcagaattgtatacacaataaattgcttaactctg atatacaaggtgttgaaatcaacgaggtacgtgttatcaatcaaatfagaacaatatgtgt tagcaagagataattactgtcaatgtgtgtaaaaggtaaaatcaaatfataaatactatc tacaccattggaactattggaataatgacatagtaataaacctgcaagtttaattactcttg tgaattgtgtaacatgtaaaacacataaggaaaacgtttcttattgtgttggaaccta aattacaataattcagagatgccgtttaatgaactgtatgcgttgggtttatacaacag attaaaggaaatataccaatgtcaaatgactacggatataatcacaagcgtatgcgtt ctgaacacataatacaaaaatcacatgtaaatgatgcctttatgtattaccggcaatcctgatg taacacgaatgaatacaatgtatctttcaaaaagaacagatcacaaatcgacaatacat aaagccaacaactccaagagcagattgaaacttaatacagcacttatattgttaaaag gttttcgattaaatgataaggtttgtataacggcgctatctgtttataacaggcagacgaaa aacgggtattttgtgctaaaaaattgtaagacaaacctatcaaatcggccaaatgga aagatttaacattgcttgaaaaagcaaatattactctatagaaggaattcaaggagttag	gtcaactaccaccactgaagtgcggcctt gtaagataattacagcttttgcaagtgga gaatttgatttaacaaattccgttgattatc taagtgttagcagcactacgtttattagaat taggcaccagtggaatagttccaagccac tgctctgcgttatgcaattaaacagtgactga gggttaggaatacgtgtgtgcatcagttaa ccttaaaataacattgacgatggagctgaca gtattattactgcatcaacaggaaactgtttcct gcattattattaaatattatagaaggaa tgcatatcaatg
1836-1837	mouse gut metagenome genome assembly, contig: NODE_4542_length_6004_cov_2.376702, whole genome shotgun sequence GENOME_ACESSION: OFGO01.1 GENOME_ID: 3940 CONTIG_ID: 4541 SOURCE: WGS DATE: 2019-01-27.	atgggtattactctgataagcgcaaaaagcctttgggctttgctctgagaacgcgccc cactctgctgaaaagcggcggcctgcgtttacaagtatttccattcaccattattatcaa ggacgccgatgctagaccattgaacctgttgctcttaccgggtgaaaattgaccccg cgaaaagcacaccggcctggccatcgccgcaattcggacaactgcgtggtgtattacct gcaggttgagcacctgtcagaacgcacgtcaagaattctgaaacccgcaaaagctgcc aggcgaacccggcgagcgtgagacccgataccgccatgcaagtggaattaacact acctgccgcagggcagcaaatatacaaacagagctctccccccgggaagggtggttcc gccgtctgtaaaatccattgtgataacattatcaactgggtcatccgattcggaaggttag taaacattaccacgtgtcttttgaggcggtccggtttgataccagctgctggacaatccg gataatcagcggcgtgctgaccaaaggggacactctttggttacgaaatcaagaatac ctactgacaataatggccaccagtgtcaactcgtcgcggtgctgcccaggacccgtgtt ttggagttgggaacacattgttccaaatccagggggcgagcgcagacataaaaaaac ccacccttgctgccatactgcaaccaggcaaaagggaattcttttagaagagtggtct cgccaaagaagcggctgccggagcggcaaggcccaaggcaaaacaggagtttag caaaagcccgggtatcaggtatcggccatgctctgaagggaagccccagaaaaag caatcgtactgtgctgggtgtgttctcccgccgtatgttgagactgggtgttcagca tctttggcaacgtgggaatgctcctctggcggcagactaagttcaaccggcagatgctga agcttgcccaaggaccaccactatgacgctctgtgctgtaggggaaattccagatgtgtgc tataggacgtgacgcatgggtactgctgtatcaaggtctattgtgtggtgacccgggt tccgggggaaaatcaacaagtgcgggttatccaaaagctgcgaacaaacacaa gcgccattcgggtccagaatggggtatgttcttgccaaatgccccagccggcaata	gtcaactactcggtcacaagtagccgagctt gacaaagagtaactctctttgtgagcctact gctttagattagttgagcagacatcagcaa gtgctacgttacaatcagcggccagggtg tcttctccagccctggcgcggtatcgcc acgtgcggcgctacgccgatccggaggc atgcagctcaaacaggcaatccgagtacaat agccagtgctgcatgcgctaacggttgata acattgtcgaggagacgtgttctctctc cattggggaggggagaacaccattacaag cccttaacaggagactatttctcagaaggga

SEQ ID NO:	Name	IscB sequence	ωRNAs
		caaaggacgccacattggcggtgtatgaccagaaagctggtgtttgatataaggaccacagggtgataacctggtaactgtaaacacaagtattgcaaacctctacagagagacaacggctaccaatatcggtag	
1838-1839	human gut metagenome genome assembly, contig: NODE_2011_length_119_52_cov_14.972463, whole genome shotgun sequence GENOME_ACESSION: UESP01.1 GENOME_ID: 15163 CONTIG_ID: 2010 SOURCE: WGS DATE: 2019-01-27.	atgagatacgtatatgtactggatgtggacgggaaaccgtcatgccacgtgccggttcggcaaggtgcgcggatgcttaaatccggacaggcaaaagcagtgataccttgccgtttaccatccagcttacctatagaacaaggaccgcctcctcaaccggttacttttaggccagatccgggacgaaccaatattcggtatggctgtctgcgttcggacggaaaggaaactgggcccgtttcactgcataaccgggaataaggaaatcccgaagctgatggcagaccgcatggcagccagaaaagctcccgaactgtggagaacgtctggcaaggaaacggctcggcagaaagctgcatacaacggcaagcatctgaatgggaggtattctccaggatgcagcgaaccgatggcggtcaaggatcatcaacacagaatcacggtttaaacacgctctgagaccggagggatggctgacaccgacggccacgcatcttactgagaacacacatcaacctgttcgtaatctggcaaaagattctccggcaccggacgtggccgtggaaactcaacaagtgccttcatgcagctggataatccggagatgaagaacgggagattgattctgccgcgtccgttaccggaaacggagggcgtggaagcggcggttaaggaaacaacaggcgccgcaaatgtctgtatgcgggaaagaacaaatcgggtcattatcacatctgtgccggttccagaagaggcaatcaatcgttcagaacatcggcggtcttgcctcaagtgccatgaacgggtacataaggatgtgataccgagagagactgacggagatgaagccgggctcacaagaagtacgggtggtacatcggtattaaaccagatcatcccgaactcgtcacacagctgtcggaattatttccggccattctatgtgacaaacggatggaatacgaaggaattccgtgagaacatgacctgggaaaagaccatgatgtgatgcctactgcatcgctgcagccacctggaatcgggaaggggcactgatggaaccgaaccgtttgagatcctgcagttccggaagcatagccggggcaggattaaacatcaaacagaacgtatttataagttgggtggaacgattgtgcccagaaaacggcggaacgggatggaacagaagcggattcactggaagactggtataaggatatgtaagccaatcagggaaagtcaaacgggatacattgtgttccaaattggctgttatcaaaaatccccgggtactacaacacccgggaagaatcatgccgggagctgtgttctgtacgaaaggaaacgggtacgtgacccgccaagatcacgaacggaaagtactatcgtcttacggccaggagaaacgggaattccggcgcaaacgtacggataccgcaaaaataccggattgtatttgcagcatga	gtcaataaccatgactaaagtcacgggcctgcagatgcaggtcctgatggaaagggttattgagcagagagagacatgtcaatcgccgctgggggtgattcaagccccgtgctccggttagacatgccccatgctatgaaactttaacggtatgtgcggtttccgtacataccggcttacaaataaaggagctatcatgagatac
1840-1841	Ga0136851_10002495 GENOME_ACESSION: IMG_3300010413_SF_3300010413 GENOME_ID: 312516 CONTIG_ID: 2494 SOURCE: JGI DATE: 2020-04-07.	atgaaggtagacgtaatacaagaatggagagccattaatgccttgcaccacgtaagccaagattattactgaagcaaaaggacgtagtaccacaaaaggtaaaacagggtacttcacataaagcttctgcatggtgtctgtgttacaacaacctattacagttgggatagtatctgggtgctaaatattgtctatagcagcaacatctggttaagaatgtcttatacgttaaggagaagattccgaacagacgtgaaagggaactagaggaagagcagctgctcgaagacgtagacgcaatgatacaggttatagaccggaggtgtcaacaataggacaagcagaagtgtaacgcagtgtggagagaacaacttaccataaagtattgtcaagacaaaagaactaatgacgttcttaagaatgaagctaaaggaggggtctgctctgcagaagatgtgcagctgaaggtgaaaaaggtgtagggatgaagacgtatattatgccttcagtgaaaaatagagcagagagattattatgataagacaactgttattctttaccaataagtgtggtgtgtttgagagacgtatctttgatactcaaaagatggctaacctgacataaaaggattagatataacatgataaggaaggatgggattgagacaatacatcttactataaataggcacaatgtgtttattgtgttaaggatataagtgagagaaagaaagtaaatatagaacatattccattttcaagaggtgtgttcttttagaaaaatcctacatgcgcttcgaaggatgcaatagatcaagaacgcagcactcctaaggagtggttagatttctactatagagaagaataaggagccaaactaaacgaacagaagttacttggataaagaatcttctaagttatcttattaacaaagttagtaagacttttacacagtgcatattctcaatcttacaataattacttattagacagctgagagaaagatggaacacatctaccactacaggatagaaacgaaatgggctagatctcaactacacttagctaagctcagattatagatgcaatagtgattgcgtctaaagtgaaagaagtagagatccctaatacatctataaaagagaacagatcaagaagagataccacacgattacataggaccattaaagaagaatgtaaaagagacatattctcctagagaagatgaggtatattgggtcaggtgtggatcgagtaattgccaatcatgctaaaaaaggtagaattggaaggctatgtaacttctcgaaggaaagtggaaagcttctattagtaacttagatggtgagctcctaataaggagaatcagttataaaaattagaacttataagaccttactcttaactatgtaagaggtggataaaagcttaa	atcgactatgctagtgtgagaagttactgttagatttacaggagtagaaagttataacttagattacgacagaaatattattacataagctctccatactataaaggagagattcaatatctatgagatcattaaagatctttagagacagcggacaatctaagatgtaatacatattgattgtacagaattggttcacagaatgtgggtcatcaaacgctaaccataatctaaagcagttatatactaaaaaaggaggggaagatgaag
1842-1843	metagenome genome assembly, contig: NODE_68_length_10156_3_cov_30.349657, whole genome shotgun sequence GENOME_ACESSION: OGUD01.1 GENOME_ID: 4712 CONTIG_ID: 67	atgacaaaattatgctttgtattagacgctaattggaaaacaattagcacctacaaaagaaca aaaagcatggtttctaattcgtaaaacgtgcgacattggttaataaatccaatggttaatacaacttaataaacaattcctaattgtggagatacagctaataaacgaattgattggaata ttgaaaatcacatagcaatgctgctattgtattaccgacttgcacacagatgtatatgaaa ttaaagagtgagataaaaacaaatcgctagacaagaagtaaaagcgaacagataatgtttt aggaattaacatagagattgttgaatatacattaaaaatggagaacgcataaagggtt atgtaaccgcgttatccaaatcaaatgcttttaaaatgccaaagtcacgaagcagttgta aaaaagtaaatgtcagaagaatgtaaatcttggaaatataataaaatttggcctgaatgtgtatcttaa	tattaaacataaaaaaaggtagacataaactttttgttggcttaagtaggtttcaagcctg agtgactgctactatcgaaagatattgttag atatgaactatgttagatagtaaggtaaaaacacaccttagatgtatcttcagctcaaaagctc tgtgagtgccaaccaagaacaatgctaatgtcttcgattgataacagggaacacatattctctatctgacattggcaagaagagaatgtctcgaagaaaagggtgcagaatgacaat

353

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ctccacaagcggctggcgctggcgctggcgctgaaaccggctctggcgctca ccaaatggaacaggcagagcaggagctccgaagggcactggatcgatcagcct gctggcgacctcaactcctccggctgctgctctacaacagtcgctccctggctgattg aggccaaggccgacaagctcgtcagatggtaacgttgaccatctgggattccctcga gcaaaccaaaaggccaaagccgggtacgaggatttcgacaggcgaatctggcgaagc cgtggctccaccgcatctcgggtgcgaagcgctgcatacggcgagctcctggcgca cgcgcgccactttgacatacagaccaggcatggacgagtcgaagatattcctgcacgct attgacagattgcatcaaacgatgggtatgctatcagcttggggcgcgctcctccg catgcctga	
1852-1853	Okeania sp. SIO3B3 3B3_NODE_558, whole genome shotgun sequence GENOME_ACESSION: GCA_010692535.1_AS M1069253v1_genomic GENOME_ID: 325543 CONTIG_ID: 2076 SOURCE: NCBI Prokaryotes DATE: 2020-04-22.	atgaataaaaattccgtctttgactggatacaataaataatctaatccagtacatcca gcacgagccagaagaattataaagcaagtaaaagcagcagtggttcgtagatatacattta ccattattttgaaagaggaaatcagcaaatgaagtaaaagaattaaactaaactagatcc aggaaagtcgcttcacgggattagctttgtatcagacataacattgtttgtgtgcacaatt agaacatcgagggtttcaaatcgcgaagctttaactaaacgtagaactttgagaagttctc gtagaaccgcgaactcgttatcgtcagcaagggtttctgaatcgtaaagatcctaaagg ttggttacctctagcttaattccagagatatttaacatagaatcttgggttaagaacaagg ttatcagcctttgtaattaagctatggggcatggaagtagacaatgtgcaattcagataa atatggctttctcgtggccagccaaaactcagacaaaactcattctatagttttatgactgg cgatattgttaaagctgatactcctaaaggtaaatatgcaggtactatactggcagaattgc tactagaagaactggaagctttaaacctaaaacatcaactcaaacctttgatgttaaccata aatattgtcgtcatatccataagctgatggttttagctacagttttgtgaattagtaagtac aaagtcaaaactgataactaa	gtcagcaccctggctctaaagagacggagct tcgtgcatcgttgatttagtgatgctgacccg gctaagcttaactgactacgttattagtaaga gttaaaagctcctaccttgaatacctcacaag ttcaagctctagaatcagatagttaaacagt cctacgaggggtgaagacagtcgtattgaaa agtaccgactaataacatagcttatgtacttt taccctagcaataggagttttcaatgaataa aatccc
1854-1855	Scytonema sp. RU_4_4 NODE_236_length_4553 4_cov_15.648975, whole genome shotgun sequence GENOME_ACESSION: GCA_012031965.1_AS M1203196v1_genomic GENOME_ID: 331080 CONTIG_ID: 176 SOURCE: NCBI Prokaryotes DATE: 2020-04-22.	atgcaacgagtagaccagtggttagataaagacggcaagccgcttagccaaccaaacccag tcgggctagacgttggctcagaagggcaagccaagattgtacgcaatgacctgaatgt ttttgtattcagctattagtagaacatcgtgatgtggcactcaaccaatagctttgggatta gaccaggtgaaagattactggtgttgggtccaaatcgtccaagttcattttgtcatggca catctaattctccttttctgatgtgacaaaagagatgtcaggaaagcgtgatttgcagctgtg ccagacgaggttagacgcatcaaccgtaaggtgttcattcaacaaaagagcgcatcatcaa aaacgggttgataaccgtaagcagaacaattgccacatagttatcgggctaataaagaac tggaaattacgagttaccaagaattgtaaaactgtttctgttactcaaatcattatgaata tgtcaaaagccaaaggtgataaaagatttagccagtgatgggtgtgcaaaaagtgatgttc caatgggtggaaaagattgcaccaaccaaaatccaagaaggttggcagacttcaatactc aggcagcaactaggttttagctaaagacaaaagataaatctaggcaagtcctgaaact catgctcatgatggagtgctggttggctgcaagcaactcatgaagtttgagaatttcaga ctgcaaatagctgtggtcatcgtctggagagggaatcacagtcacatctgcgccatttc gagtgattgtctgtcctaactgtttctgtcgcgaactcattttgagaatcctgtgaaaagatgc accaggtatagaaaacgtaaaagggtggagacagtaacaccttttgggtttcgtctggcgatt tagtcagggtgtaaaaagcgggtaagctgtatttggatgggttgggtgataactcaaac acccaaaaccagaataatttccgtctataaccacaattggcacaggttggacaggttcagc ccgtcaaaagggtgcaattaatcaaacggagtagcagattatgcgtagcactcctaa	gtcaactaccacactgatctgagtacagat acagtggtgggttgaagaaccagctttcaa cgcaagagatgactagccccaccagacttg ttttgtacagactccgaatgtttccttagtgc ggattatctctaagcctactggtttagggcgc tttagaaaagacatgtagacaagttgggc taaggacttattcttctgtaaggatt
1856-1857	human gut metagenome genome assembly, contig: NODE_59_length_61911 _cov_6.971902, whole genome shotgun sequence GENOME_ACESSION: UZNL01.1 GENOME_ID: 18041 CONTIG_ID: 58 SOURCE: WGS DATE: 2019-01-27.	atgatttatgaaggagcaaggatggaagcattgatgccaaagtgcgctggaggag gataggctatcttctgccatggcaaggtcatgtagtcagccgtgttctgttctgca gttgattatgagacacatctacacaggaagtgagccttggcattgatgctggctca aagcacattggcggtttcggctagttccgagaaggaaggcgtgctgcagcgaggtcga gttaagaagtgtgttgaacttctatctactcgaaggagttgagaagacaagcg aaaccgcaagacacgttaccgaaggttcttttgataaccgcaagaagaagatgggtg gctagcacctagtgttgagcaaaagggtgagagtcactgaagggttatccgttgggtcgt agttacttccaattacgaagaccactatagaggtgtcctcgtttgatgcgcaaaagattaag aatcccgacatcaagggtgatgagatcagcaaggcgagcagatgggcttttggaaactg gagggagtagcttttggctagggtgggcacaagtgtgttcatgcaaggcgcaagagca gagacctatctgaacgttcacatttggagagccataagactggtgttaattccctagt aatctgtaagcgtttgcgagacctgccacaaggcttaccatcgtggggagttcactga aatcaagcgtggcacaactttcgtgatgctgcgggtgatgaacattatcgttggctggt gtatgaacgagccaaggctgatttgggaatgtgtacttgacctatggttaccattacaag cacactgcatagagaatgataattgagaaaactcatgcagccgatgcttctgattgcca agaacgtacacgcaaggcggtcgagaactttcttatgtctgttgtgtacctcgcatacag agagcattgcacgttgcgaaccggaagaaagggtgtattcgtaggtcttgcattgcctc ataagataggcaagtcgtcttcagcgttgcacatggtatggttggaaaggcaaggat gctttatcttttggagcagcgacggaagaccaatattgcgtgatgttgaaggaaagcaaat tgcaggacaaccgagtgatgaatacaaacgataaagtttttaagagattaagaataaac attttagtggaagaaaggacttccgaaggttggatagaataaaggtaagttag	ttcaagtttcggaagcctttgattaggctaca gcgattatccattcaatcgtccggagcggatt agcctcagccccgaatggaattaggagct acgttaggggtggaatgcataaggcagctcag gatgtccgtccaagtctgacctcgggttc gtggttaaaagtggcgaaagctcggtgct gcgggcaagaaccatctataacattggc gatggcgccacaaccaccttcgaggtgag atttataatttgattaatgattgattatg
1858-1859	Ga0302251_1000232 GENOME_ACESSION: AAG_RA_Gln_MG_2_\$ F_3300028625	atgcagttggtgtctgctggaccgggaagaaagcaaccattgatccgtgccatcccgcg cgggcgcggaactgctgcggcgaggcgggcgctgctccatcgccctgcgccctttta ccatccggctgaaggagcgcgtggcgagatggtacaaccatacgcctgaagctcga ccccggcagcaagaccaccgggattgctatcgtgcgcgagtcaggagcaagcgggt	gtcaatcacctcagcgtggagccggggg gcttggggggcgaactccaatgacctgattga ccagggctagccgggaagcctgacgaacg gctgtaccggctacgttggcagatagata

SEQ ID NO:	Name	IscB sequence	ωRNAs
	GENOME_ID: 279887 CONTIG_ID: 231 SOURCE: JGI DATE: 2019-06-13.	cgtgctgcacctggccgagctgcaccacaagcggacatcaagaagaattggatcaa aggcgggcctatcggcgcaacggccgacccggagagctccgctaccgcgctccgcgc ttcctgaaccgacccgcccggaggggtggctgcccgcctccctgcccggcgagggtg gataacatcctctcttggctgctgacgctatcggcggtgcccgttacggccctctcc ctggagctggcccgttcgacacgcagcgctccagaacccggaatcagcggggta caataccagcaaggcgagctggcggttacgagtgccgagtgatctctggccaagt cggccacgctgctactcggcgcgaccgacgtgcccctggaggtcgagcacgtc gtgcccaggtccaggggcgccgacgaccgggtgagcaatctcgcttatctgtcacgt ctgcaatcagcggaaaggcgaccgcaccgcccggagttcggacacccggaggtgca ggcgccaggcccgcagccgctccgacgcccggcgatcaacagcacgcgctggg cgctgtggcgggcgctggcgacgggctggcggtcggaggtgggacggcgggg cgaccaagtacaaccgtactcgctgaacctgcccacacccacgcccctcgatgctt gtgctgggctgtccacggcgagcgctgacgactgaatggcgcggggctactacc atccggcgagggcgccggcgagtgatcggcgacccctcgctgctgtagcggttccc gcgtgggtactgatcgccacaagatggtgcgaggcttgcactggcgacgtggtgc aggcgaccatcccaaggcgagtgatcgggctggcgacggcgacggtgctggtgc gcggcggggggtatttgacctgaagcgcggggctgctggtggccaggggattgc cgcccgtattgcagagagcgcaacggcgacgggtatgggtatgcttgcgccat cccccttag	gtagaccaccggcgggcgcttctcaacc cgccgctctgggaaggactggcgagaca ggctcaggggtaagcggaacgcgcca gttcggggggccaccacccccacgggtc gccaacaggtccgaggggagccgcggg gaaagtcggccgcccgttaccgggt aagcgggggtggagtaactcaaatgcagt tg
1860-1861	Hymenobacter sp. CCM 8763 NODE 27 length 66586 _cov 28.8038 ID 53, whole genome shotgun sequence GENOME_ACESSION: GCA_003417065.1_AS M341706v1_genomic GENOME_ID: 170640 CONTIG_ID: 44 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgtctaataaattgtttgtattagattatgatgtaaacattatcaccaactaaagaaa caagggtgtgttctaatatgaagataagcaacattagaanaaataccaatgacc atacgtttaataaacgtgtagaagataaagactgataaatacaaaatgcattgattg gatgatgttctaatacatgtaggggttacttattgccaagagagagactaaaataaagt agtttttaatactacaatagaattaagacaaaatgtaagaaatgaatgatacagaagag gtttaagaagatacaaaagatacataaagatacagacccttcacgttttaataatgtaaat cttctaggaatacaggtagattagctcctagttaataaaaaaagaacgtctattatacga gtagttaagaggttaaacacacataagatttaacgttattcattagaagatgtagctattg atactagagctatgacagatggttacaacctataaatgcaataatactaatcgtt tagatgaaatatacgtaaagctgtttatctacagataataataacttgccaatgagtgta ataaaaatgaagaatggaagttcatattatttctaactgttaaatggttctaactata aataacttaattactttgtgtacagaatgcataggggaagtaaacaggtaagaagaacaata cattgattattacaatctattgtaggaagaaacacacaccttttaaatcatgcttcacatg taatgataggtaaaacttacttaagacaagagtatttataaattgcctccataagattaaca caggtgggacacagcaataaacgtattgattggaatatagaaaaaacacattctaag atgcaatcgttaattacaatttaattcctagttaaggatttaataatagaagaatgtagtta accaataagaagaaaagtaaatctattataataatgtaaggaattgaagcatagagatat tgttaataatactatagaatggtgacacgcatataggctatgttacaggtcttatcctagt agaaaggctatttaatttccaatcgtcaacaagcattgcaaaagcagtaactgtaagaag ttaagttattggagatttacaataatgtattggttaacataa	tagaagattatttagtaagaatagaagtt aaaataaaggacttacttattggttca agcctcagtgactaggggttggaaacaac gtatcctagatagtaactcgttagaggataa tggtaaagacttactcagatgcacgttta gtcgggaagctctaagagtgtaacaaagaa acacttcaagtcattgatgaagtataac atggaagcacatattacccttgcatttgcc aaaaagcaaaataccacgaagtgagggtta gacagaaatgtctaataaatattgt
1862-1863	Ga0070707_100063183 GENOME_ACESSION: KBSK5StandDraft_9_\$F _3300005468 GENOME_ID: 30508 CONTIG_ID: 63182 SOURCE: JGI DATE: 2019-01-27.	atgtcgtgtgtcttctgtgataccgagcgccgcccgtggaccgggtgcatccggc gcccggcgccgctgctctcccagcgggagggcgagtggtggcgccgctccgttca ccctcatcttgaaagcagggcggtgcccggagcgcccccgcagccagtgccgctcaagct cgatccggcgagccgacacactggcctggcgctggtcaaggagcgacggcgacggccc ggcagcggaagatacggcatcgccgcccgtggcgagacggcgggcggtggtggtgg caggggagctgacccacggggggcagggcagctccagagaactggcgcgccg agcctgctgcccgtcccggcgagcgccacacccgttaccgtccccgcccgttcgag aacggcgccggcgccggcggtggtgcccgtcgctggaagccggctcgccaa cgtggagacctgggtggcccggctgctgcccgtgacacacgtgacggcgatcagccag gagttggtaagttcgacacccagcggttcagaaatccggagatcagcgggcggaat accagcaggggacgtgctggcggtatgagctgggggaatactgctggagaagtggtg ggcgcggttgctgctactgtcacgccacggcggtgcccgtgcaaatgagcacatctg cccaaaacgcggccggcgtcgtgacgtgcccagcaatctcacgtggcctgcccgc cgtgcaaacgagcgcaaggcgacgggacggccggaggaattcggcaccagagggtg cagggcg	gtcagatccccggcgtaagccggggggc ttgcacggcggaacgcgggggagggccc cacgtctgaccagccccgctctcgaggg gagcgagcggtgctggagcgaatgcatag gcaccggcggtgcttaccagccccggc ctctcgcgctagcggttaaacaggtctacg gggtgtcaccagtgccgctggcaagcaag gaaacgctcacgaacacggcgaggtga gcatcaccggcgcaagcgaggcccgcg aggcgaccagaactgaaaggtcttctgt atgtcgtgt
1864-1865	Ga0307376_10008654 GENOME_ACESSION: TR2MG_2_\$F_33000315 78 GENOME_ID: 281014 CONTIG_ID: 8653 SOURCE: JGI DATE: 2019-06-13.	gtgaagcaaatgtagagaaggccgcatcagaacggattttgtgatatcaaatcggg aaggccgttggcccaacagctggccggggcggtgctgctactcaaaagcggtt gtgcgcggatgtagctatgaccccttggcgtccaaactgctgtacgactgctctgagttc gtgcagtgcgaggtaccgtcgcaatcaagaggatcgaaagatacacaattgttgc gcagaacacattacaagatggacactgtccatagctacgtaagagatcttgcagc gggctggtgcttgcgtcatgtgaacgcgctacggagccggcgagccgagggaa ccgcaagaccgggtacagaaaaaacgggttcgacaacggcccaaaacttattgttccat ttggcgctgaaccacaccccgaaacgtgggcaaaagtcgaacgcaagaccggcac aagtctaaagaaggtaggagtcggtatgctggttgcgaagtgcgaacatcagg ggcttggcgagcatttggcgctgtgcccgaacacatctacccagcgctacggaac agagttgatgtacgtacggaagtaggaattgactgccaatcatgccagtgacgaa	tctacctaccagcgccctgagaccgtggtt ggacggaggattcttgaactgtctctgtct taaggcgagcggaagctccatagagagac agcctaagtcttaagtactactgtctccgg ttaacagacagtcagatgaacaagaccctt ctccgggttagcagctggtcataacttctg gaggtctgttaaaagcccttgggggatgc ggcagctcaacctggagcgcccggtt ggcggaagccggttagaacttggcgggg aagtaaacggcggaagccccgattatct cccttccgggagaacacaaagatgctgt

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gataatgatggagcctacggcaccttatgcccaataatggcctattgaaatgggcaactt aataaacctgcatcgagcaccatgctgggcatactgtattgttagggagtagcttctgg aaaacacgggcatcagtgctgttactgcaaggtcacaagcggcgaccgttcgctgtcaa ggaacacgtgattccaaaatcacggggggaagcgatgccctatataacctgggtatc gtgcaagacctgcaatgatgcgaaggggggcgaagacggcccggaattcgatacc ggagattaacgagatggctgcgaagtcttctgctgctggaatggggcgccattatcaaag gtaccagcggatgctctggcaggagttcgagaatcgggcacagcggtcgaggtaccg tttggcagttgcacaaagcataacaggcttaggacaccgttgccgaaagtatatacaca atggtegttagccccaacggcttgaagttgccccccgaaggaatacatggtcgaaaa gcgattaaaaataaggagtccttttcatcgattaccaatgagaacagaaaagggtggcca ttaggaagaccctttccatgagagaagtcgggtgggttcaattgcacgacgaggtctcgt ttatcgatggtaacgatcaaaaacctgctgttacttactgcttgcgtaaggacgatcg gcagaagtgtctgacttgaaggaaatcttattcaaaaagagcttgagaagctcacat tggaacataacgtgtatcgaaaaagggttctcgaagcgtcgtgacgtttgagggggaataatg a	gagtgaaagcaagtgagagaaggccggat cagaaaccg
1866-1867	Ga0209096_1014402 GENOME_ACESSION: AD_des_43_\$F_3300025 859 GENOME_ID: 21343 CONTIG_ID: 14401 SOURCE: JGI DATE: 2019-01-27.	ttgcagcgcgtttctgagtgagaccacaagaacacattgagccctgcacacggcc aggcgaggatgctcctcggaagggtgaaggcgcgctatcgcgggaaccttca cgtatcctcaagtacggattgaagaacccgtgcccgccgtagaggttgcgaattgati ccggcagcagaaccaggggagtcgcccgttgggggaagtgaagaagggggaccaa gtggtatgggcctgcgagttggaacatcgggggtggccgttagggatgctttaactcc cgcagggcaatccgggggggacgcaggaacaggcatccgctaccgccagcctcg gttcaataaccgtaccggccgctgctgaggtgcgccttcccttattgtacgggtaaat aacgtcgttactgtgctggggaagatcctcgccctggcgcctgtgacccatccatcgcctc gagaccgtcgggttgacaccagcgctcgaatccggagatcagcggatcgtgatta ccagcgggggacgttctggttactgaggtccggaggtacgtgctgagaagtggggg cggaaatgcgcctactgcggggcggaagacgttcaattggagatcagcatattgtccc caagaaacccatcagggccacacgtgtacggacaggataagcaacctaccattgctt gtgtccgtgcaatgaggacaagggaacaaagcctatcaggaggttctcgcggcaa gccggaggtgctcctcggaagatactgccacggccaagaagccccgtatcgacggcg ggcgatcaatgctacaaggtagccattggcaacgcgtcaagtcctgaacgtgctggt gtccttctggtccggaggcgaccaaagttcaaccgctccaagcaggattaccgaagg cccactggatagactcgtgctgtggggaggccggcgcaagtgttctcgaagcc ggatgcaagccctgaaggtgaaagcctcgccgacggggcacgagacaggtggtgaa atccgacaagtacgcttctcctgactgcccgggaagaatcaaaagggtacacggctt ccaaacgggggacatggtactctcaaccaaacctggcaagtacgcc	gtcgtactgctccggcggtatggggcaac tcgctcaaccagactcggccgatagtcat aacgtgtgtcgggctacgttaccggcgaat atataggcactccggatgcttcgcaagtcc cggacccttgcggcaggtgtgcaaacagcg cagaccagggaatgagcagtgctgctg cacacaaccggcgataacattgtcagg cgacctttaccgggggaacccgagtaaccg gacgggaacgtccaccatgttgaggga gagattttgcagcgc
1868-1869	Tissierella creatinini strain BN11 NODE_17_length_55240 _cov_1267.366911, whole genome shotgun sequence GENOME_ACESSION: GCA_005046945.1_AS M504694v1_genomic GENOME_ID: 263665 CONTIG_ID: 28 SOURCE: NCBI_Prokaryotes DATE: 2019-05-30.	attatgctctatgtgctaaataaagattggcacaacattatcctactaataaggcatgtaaa gtaaagcacctgtttaaagaagggtaaagctaaagtagttaaagaacacatttaccatac aactgcttatagtactactgaattttgtacaacctatcattaggtgttgcaggtagtaa gattattgtttaaagtgctacacagaagatgaaggtgttttcatccaggtagttaaag aaatgacattgtagagctattatctacaagaaggcaaaatcgaggacaagaagaacc gtctacgctacagaaagccaaggtttcaataagagtaaaagctaaaaataagggtggt agctccatctattaggcataagatagatagctatcaaaaggtatagctaatatataagat attaccaataagtaacctattatagaacaggcttctttgacatccaaaagattaaaaatcca caaataagcggcaccgattatcaagaaggtgagcaggttaggcttttgaatgttagagat atgtaattgtggcgtgatggtcatgaatccaacattgccatggttaaatcaagataatgtt taaatgttcatcataagataagtcgtaaaactggtggtgactcccaggtaatttagtaaca ctttgtgagactgtcatattacacatcaaggcaaacctacattaaagataactaaaat aagtcatacagggtgctgcttcttattgtgattatgagatggcgtgttataataagcttaaa gagatatatccaaatgtcactatgacatatgtttatataactaaaaataatagaattactaatg gcatcgagaaaactcatactcagatgcctttgtatagctggttaattaaaagcaaaaaga agtagcgtatctaccttcaagcattcaaaagaggcataatagacaaatccacaagcta acttcttaaaagggtgtaagaaaaaactaaatcaatgccttcatggttaaaaggcttttagat tgtttgataaagtactgttcaagaacaagaatgtttatcttggtagaagagctactggttat ttcgatttaagaaaattggatggtacagtaattcataagagtgcaaggtttaaagattataaa ttattgaaaagaaaagcacactattatagaaggcaggtggcataaa	gtcaactaacctcgcactaaaatgaaggct tgaagtgtacagcttttgggtctattttaaagc ctggttgaatagcctaaggtgcttgggcacta cgtttttttgtcataacacctaggcggtttaa ccctaacctgtggtcgtgttctgtgattaaa agttctcaggggttaattgggaacggtgtgc agggttataaaagcaatttacaacattagctaa gggtatttaaccagttattactggcattatgt atttaactacattaaaaggagaagtattatg
1870-1871	Ga0208687_1000090 GENOME_ACESSION: Jundes_13_\$F_33000254 69 GENOME_ID: 30316 CONTIG_ID: 89 SOURCE: JGI DATE: 2019-01-27.	atgaagaacgagacagaagaacacccatccgcgctccctaacggctggctggg ctgatgctgtacctgagagtagcggaggtaaacctgacacagctcaatgctgcgctcagc ttgcgctgattacagcttctgtttatctcgtatggcgaaggagctctagaatttgaat cgaaagaagacaaatcaaggatactgctaggattatgccaacaattatctgagacgaa gaaaaagaattaaaagaagaagaagaaggaccaaaaaacaattatagatgaccagct tgaatccagatggaaactcagctatgctcctacgaagtaggagagcgaataaattggatta aagaaaagaagaagaatagtcacaaataaattaggaataatttcagattcaattactttc gaaccaagtggaaagaagaacacccgatagtaattgacagttgatcctgacatcattt acaggaaataggatcatatcaaaaagctgtgttatgtgtgtacattgaattgctg atacaagcaagggttcaaccaagattgagaagaataaatttggtaagaaggtgagaa	gtcaactacatccctagcgatcagcgcta gctgagagcttatgccggcttccggcaca ctaaggggcggtgctgtatctcagaataca acatgagagtttaactagcctatcgagaat acagaatctatgaagaacgagacagaaga caaacaccatccgcgctccctaacggct ggctgggctgatgctgtactgagagtagc gaggttaaacctgacacagctcaatgctgc gctcagcttgcgctgattaca

SEQ ID NO:	Name	IscB sequence	ωRNAs
		atatgcgaatgctatagttgaagggaatgataagaggagagaattaaagaggtagaa gacatagaattgtcgaagaagagaagagatggtaaatagaagtagatctaagatcc ctccctcaatacttgcgagaaaacaattagaactgaaagttgcaacaattatgcaaaaata tatccaattataatgattggtttgaagatgtgacattcgaatcttttaagacacgaaagggtg taaaagggtcaattcttccatgttgaaagttgtaagaattggatattgagagaactgagaa agatagccaccaagaggagttaaaataataaaagatatgagactgctagaatgagaagtt tgttattgatgaaaaaggaaggagataagactgttagatctgttgaaagccatgtcaatgat tttatagcgtaggagctatgtttaatggcggagttgagactaagaataaattcaata tgatactattacaagacctaataattcgagaagaattctacattagaacaacctacaaaag gtgggggtgagaagaagatatggaggaacgacttgcattgataaattggacaataattagga aagggtgactatgttgaaagacagagagagaatgtgtatagatcttatgtctggattt actaccatgaatggtagaattatatactgtctgatttgcactggaagagattgtgtaa gataaacacagacgcgacgtaccgtctaatattaagatatgtagtagaatagtggat tattggtaagaagtgtgtaaaattatgacaaaagtaattattttgacataaaactggg acagtacagatgaattatgaagatgcattgggggttaa	
1872-1873	Petrotoga mobilis SJ95, complete genome GENOME_ACESSION: GCA_000018605.1_AS M1860v1_genomic GENOME_ID: 39841 CONTIG_ID: 0 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atggaaaagggaatctgtacacgagttttgtattagataaaaaacaacacacttatgcc atgtcatcctgtctaggcgaagagattatgaagaaaggtaagcagcagatttctgatac catcattcacattatcttaaaagatcgtgaaggcgcgcgacacgcaacctatacagggtta agatcgtatcctggaagtaagatcactggagttacttttagtagcgaattttaagaacggtaa gaaagtaatttgggtgctgagattcaccagaggtcgaagatcaaaaaagcattagat actcgtagaggtgtaagacgttcacgagaagaaacgaaagataaaggatctgtatgcaag gtttgacaacagaaaagcgttcaaaagggtgttaccactagctttataagccgtgtagaaa atatactacatggtattaaacgtattcgcgctttaccacattacaggtatttcaattggaatta gtacgtttgtatctcaaaaactacaagatccagaaatgaatggaatagagtatcaacgag gaacgtgtatgtagacgaaataaagagtagctcttctgagaatggggtaggaatgtgt atattgtgtaagagaatgtaccgttagaataagaacatattctacacaaatcaaaagggtg gaagtgcaggatagaatgaaccttaccattagcttccacgaatgcaatcagaagaaggta atcaatcaattgaagaattttaacaataatccagaaggcgtgaacagatcaaatcaga atcaaaagagaccactcaagatagacgacgacacatgctacacgttgtatataatcaat caattacgaggaacacgctcacggcaggtgaaggagaattacattgaagttggaac agggagacgcaacaagataatcgtgagacacaaaactatcctaaaaagcattggatag atgcagctgtgttggagaaggtggcgaagcgttcaaatgaacgtgatatgcaagtttta gagatcaaaagcaatgggtcatggaatgcgtagaatgtgtttgtatgataaatacggattcc caaaaaatattctcctcaaaagacgaacatataatgggatacaaaacaggcgatatagtgt tggcagttataccaaaaggaaaaatattgggtatccatttggacgattgcaatcagaca tagggccaagcgttttattaaatgtgtgtgtgtatcatccaaaatacttactactacaaa aaatgacgggttatggatatacaatattcttaa	gtcaactacccccgtctaaagcggaggcctt gaagcctcggaggtgaccagcctaagttcg aagacagcctcacggcaaatftgaagaact acgttattctgtgcatgacaccttggaaactgct ggagccaggttccaaagcctgtcgttcaacat taacaggtatagcgggttgaaccaggtgtg ttgaatgtaaaagccagaataacattggcg aggcctactttaaccgaaaggagagaggaa ggtaacttcatggaaggaatctgtacaa cga
1874-1875	Chlamydiae bacterium isolate K940_chlam_3 K940chlam3_32, whole genome shotgun sequence GENOME_ACESSION: GCA_011065065.1_AS M1106506v1_genomic GENOME_ID: 327255 CONTIG_ID: 147 SOURCE: NCBI_Prokaryotes DATE: 2020-04-22.	atgcgaggtgttagttatagacaaaaacaaaaacgctaattgccttgcaaacatcaaga gcaagagagccttttaagcaagggaagcggcgtgtctattagacgataccattcacgatt atacttttctcgtgagggaggtgtccatgcaagaagaacggaactaaaagtgatcctgta gcagaatttcaggaaatgcattagtgtctaatttaaacaggagcgaagggtagtttgggc atcaaacctacatcatagaggacttgcagtgagaatgctctgtattctagaaggcgattg agaagagggaagcagattcgtacacacgttatcgaagcctcgtttgataaccgaaca agggcgaaggttggctcccccttcttgcagttctcgtagtagtaacgttcatcagtggg caaaaaactcaacgttttgcgatttcatctatcgtgtagaactgtccgttttgataca caaaaaatgcaaaatcctgaaatattcggcattcagtatcaacaggagattacttggata cgaattcgaagaatattttagaaaagtggggggaacacctgtgcttattgcatgctaaa gatgtccctgagaatagatcattgttccaaaaagcgtggaggagtgagcgtctgta tccaacttaacaatctgttccggccttgcgaatgagaaaaatccaatcaatctgtccaa gtttctgaagagtaaacaggagttttatctaatttcagaaaaagaaacgattatcttctgc gacacggcagcgggttaatgcaacacgtattgccaatgggaatgctttaaaggacttagaa ttgccgataacattctggagcggaggaatgactaaataatcgtttctccaaagatataca aaaaatcacttgatagatgcagcttgtgtggagaagcggagaagcgttacatttcc cgaaatattctcgttttagaaatctcagctacaggaagagggtcacggcaaatgttaga gttgaccgctacggcttccaagaacattagcaaaaagcaaaaagcgcgttaattgatttcc agacaggagacagagtgcatgcttattgttccaaaggaaagaaagccggaacacatatt ggaaagggtgtctgttgcgtctgtgaaattttaacatcaagacatcttgcgaacgggtcca aggtgtacatgcaagatactgtaagcgactattccaagcagacgggtacgaatatgtacaa ttataatfaaaggaggcggcgttttccctcgtctcaa	gtcaactacccctcccctaaagggaaggcctt gtagagccttattgtgaccagccttagtttcc acgaggaactacgttaggagcgaatacat aggcaccgtggagcttctccagctccac gctctcggtgagtggttaaacaggtgtaag aggttaagccagtgctgctatataaac gctctataacattggcgaggaacattacc cgggaacccggagattggcgggtaaccgc aaaaaggaaaaaagaacatgcga
1876-1877	human gut metagenome genome assembly, contig: NODE_192_length_3834 5_cov_13.286210, whole genome shotgun sequence	atgcgcggtcagccctgatgcctgtcgcctgcaaaaggcaggcatctgctcaaggc tggaaaagcgtgtggaagcgtcgaacgcattcagattcaacttcgaatcgccaccg gtgaaacgaagcagaacgtgacgctggcggttgatgcaggcgcaagcatgtcggcct ttccgtacgacggaaggaagaggtcttgcgtccgaagtcgaacttcgacaggacat cacgggcttctgctgcccggctcgtcgtcgtcgtgagcggccgacgcgaagacg cgtaccgtgcggcgcgttttgaaccgtgtccgatcgaacacaaggagtgcttgcga	gtcaataacccctgtctgaagacagaggcctt gaaagagcttatttactagcctcagcgcc cctcttctgagagggcgtcagttggttggga atgtataggcaccgtggatgtacatcctag ttccagctctcggcgttgaattaaagctc tgagagtaggagcggctgttgcaggcacc

SEQ ID NO:	Name	IscB sequence	ωRNAs
	GENOME_ACESSION: OWJA01.1 GENOME_ID: 12050 CONTIG_ID: 191 SOURCE: WGS DATE: 2019-01-27.	ccgtccgttgaaccgcattcaggcgcacatgtcgcgcacgatgcggctcgcagactg cttcccgtcaccagatcgtgattgagacggcatcgttcacattcagaagatcaaggatc cgtctgttgaaggcagcgtatcagcaggcgcgatcagctcggcttttgaacgtcgtg aatacgttctcttagagacggatcgtttccaacactgtcagggcatgcgaaggacaa gatctcaacgtacaccacctcgaagccgccaaaccggcggcgatgcgccacaac ctgattacgtctcgcgagaggtgccacaaggcctgtcagcagggaagatcaagttgaa ggccaagcgcggccaatcgttcaggagggaagcctcatagcgcacatcgttggacatt gctcgtatcgtcgcgaaggcgcacccgcgttgcctgttgaacacctcgggtatct gacgaagaacacgcgcacacctgggactcccaagcgattgcgtcgtatgcctact gcattgcgggaacctcaaggccgtacgaagaggcgtccacctctatcagcgacaggt gcgcaagcacaaccgtcagattcacaagtgtacgggtcctcgaagacgttaaggacg gcacgaagatcggataccgaagctcaatcaaacaccgcatcgttgcgaagaactcagg ctcttcgacaaggtgagatgtcttgggcaaacggcttcatctcggcgaagatcgtccg gttgcctcgtatcgcggaggctggacggcgtaaagcttctccgacatcagctacaggaa gctcacgcttctcgaagagaggcaccatttaaccgaacttatcaaggaggacggcg cttctcctctgtcagaagacagagggttctcgcggcagtttctatgaacaagccgtgact tatcttcttccgctctgatggcgtccgctgtcgtcgcctgtcgcgcttcggccggcg acaagtcggagctcaggtcccatggaggcggcgagaaggctcgcggaccgtc cctactacgagcgtcgtcgcggaggccgctgaagacggttggagatcgaaggc cgcaatgagcgggttcggccaggcagaaggagatcgcgggtcgcgtcgcgtcat gaaggactgggtccacacgcgttctcgttaacgcctgcatcaaggacgcgaggac ctgcacacgagcgtgagaaggagatcgtacgtacgcctcaaggccgacgagatga tccgctcgcaccgctcgcagcgcaggccgtcaggaagccagaagcagaac gtgaagcagcccatgaagctcggcgtcaccgagctcggcgaagccgcagcaggt cgcgcggcggagggtgaagcaagcagagccgcggcggaggcgcagggccct tgaggaggcgaacgtcgtcgtacagggaagcaggcggggccgccaagaagg cggaggagcggcacgatcccatcagggtgaagagccgctgaaggctcgtcaagtc ccattgaaggacatctcggcgtacggcgtcgtggtgaggcgggagggccgagg cgccgcgagcgtatcgtcagaggacgccaattgcggccttcaatgccaagcagg ccgaggcgaagaagcgccttgaggaggccgaggccaccgcccgaacgcaaggc gagccgcgaggcgcggcagcgaactcaacaagacgctcaggagcgtcgcggccg cacagaagcgtatgaggagtcagggaacaaggactccggcctcaagaagtattt ctga	aaacccctccaacattggcgaaggatgca accggccttcgggcccgtgaagcggaaacct gcgggtatccgcaaggagatactttgaaa
1878-1879	Ga0326513_10003181 GENOME_ACESSION: 2_2465_0518_D_2_\$F_3 300031760 GENOME_ID: 279785 CONTIG_ID: 3180 SOURCE: JGI DATE: 2019-06-13.	ctgagaacagatttattgaaatttatcaaccgaagggaacgagacgaaccgacgt tcaaggaaattacgatatagaagcaagatttgcgaatcggaacgtacggaataatgg ctgcctccgtcaataaaatcaaaactggaatgcaatttaacactatagcaaaactgcaca gggttttgcgataggaaaaatcgtgtcgtgaggtgcgaggttcgatattcagaaaaataaa aatcctgagataaaaaggagctgaattatcagcagggtgttcacatcagatttggatgtac gtgaatacgtttgtatagataatcataatgcagatattgcaaaagtaaatcaaaagac agggttttaaatgttcacatctgtgaatcagcaaaaacggcggaactgcctgacaattt aattacctatgcgaacctgccataacatgtatcaccgggttttaataagcttgataatgt aactaaaggcagttctatagaccgaaacatttatgacatactgaggaaatattcgtattg aaaacctcaggaaaacatacaaatgtaagtatactatggttatgaactaaaagttaa agaattgcaaacaggctggagaaa	taattatggctttgcttagtctaagtatttcg ctgaaaagtgaatgaactacgttatgagaga atgaatttatgcaataaattcaaatagtacc ttcggatgttcatctagtctgaagctctaagg gtatgattaaacagtcggtaagtctgacagt gtgtacacggcaacctctcaataacattga cgaaggtgaatttaccgttttgcagagttccg gctgttaacagcataaactgcagctcataaa gggtttttgtatg
1880-1881	Ga0326511_10001255 GENOME_ACESSION: 0_2496_0518_D_2_\$F_3 300031867 GENOME_ID: 279726 CONTIG_ID: 1254 SOURCE: JGI DATE: 2019-06-13.	atgcgcggtaaggcatttaagccaactagtaatagaagaacacgtatctgttaaaaca ggtaaaagctaaagtgtattgtataatccatttacaattcagcttaaatatccaacaggtgaa acgacacagccgactgaattagggatagaaccaattataatgatttaggcatactgttgc tttcacaggaaaaagtacttcacaaatgtgagatacattttctccaggaaatgcatactaatct aaaatcaagagggtctttgagatcagcagcgtgcacgtaaaacacgttatcgtlaagca tcgttttcaaatcgtcgaagctgttataacgtatgtcaccgaagtagtcggaataagctaa gacattatataaactggataacaagtttaagcattgttaccgaactgcacagtgatcattg tttgcgtgatccagcagaacagataatctgagcttaattgatgaacaacacaaaactaac aaaaaacatttctaatatagcaatgcgtgttttgaacgtgacgggttacagatgtcagatt tgtgtagaactgggttaaaagctagcttacatcacatcagggttagaagtaagaatgtag aatgatttagataatctgacacagtggtgattgattgtcattctcatgaaaatcattaccg ggagggtattttagacaatttaccgacaacagaagcgtgtcagcaagcttaaatcggcaatta ccgactatccgttatctcgaagaatgacatgcacagcttaaacgagcatttctgatatgttaa gtgtagctaataatgatgtactaataatagtcacacattaggattaacacctactgtgtata ataaagcaattgcagcaacaggatataaccagcttaaacagaacacaaatgacatttac aatcagacagatgcgaacaaagatagaagcatacataatttcttccaataaaaggtag aaaagaaccaaacagaatgcaacacgtcgtatacaaacatactgtacataccgaggat gggttaagaatgataaagtcgtgtgtcgggtactaagggtgttttataacaggatttgta attttgtgctgtgttaaggatatcaatggtaattacataacagaccgaatagaagaact actcagtcatttaagtagtttaaaattatgtgtcacaataataattggcagtatcagacagt acaattgtgagtaa	gtcacggctacttataataataatagtacacac aaaatcgtatgcactcattgaatatataattca aactgtgaataacagttacatcgattactctaa gttctctgagaactcgttattatgttatcata catagcgtatgttctctagtctgtattcttatgt aggctctgtaaacagttctgttgataggaa cagtcacactaagggtgtcaggtacgacaa gcatttataacatggagggaagggaacaaa ctttcacagaagggtataccacttgcgtagta ggtatttaattagaatcagaaaacaga

SEQ ID NO:	Name	IscB sequence	ωRNAs
1882-1883	Ga0265293_10014478 GENOME_ACESSION: Munlanlewell1138R_2_ \$F_3300028603 GENOME_ID: 280700 CONTIG_ID: 14477 SOURCE: JGI DATE: 2019-06-13.	atggtttatgtaatttcaaaagaagggaagcctttaatgccaactaaaaggcatggaaaaat tagaattttacttaaaaaatagtaagcaagggtgtttaaagaaccccttactaitaagcctt ttgtatgataccacaattatactcaacctgttaactttggcggttgatagtggtactatctatat tgggtttcagcaacaactgaaaaagaagaactgatttcaggagaatgtactttatataaag gtcaatccgagagactgaaagaaaatctatgtatagaaaaaaggcgaagtagacta aggatctgtgtccacgctttgataatagagcaaatcccaagggttggttagcaccatccat tcaacataagtatgaatctcatcttgatgttgcctatttacaatctatattgccaatttcaaa aatcatcatagaagtagcgcaattttgatgttcaaaaaatataaaatcctaaattgagggaa aagaatatcaagaagggtgaacaaaagacttttgaacttgaaagaatacatttgcacgc agataatcacaaatgtcagaatccagattgcaataacaagtcaaaagataagatattgga gtacaccacattggttttggaaaaagatagaaccaatagaccaagtaatttaattactcttt gcgataagtgccacaatccaaagaatcacaaagggttaactttcttgggttggaagc ctaaataaagattttaaagaagctacctttatgtcagtagtctgttggaagccttgtaag cttaattgcgaacatgcttatggttcgacacaaaatcaaaagaattccttgggtttaga aaaaactcatttaattgatgcttttgcatacaaatggtctctactcaaaaagagtaaaacc aatctatttgaacaaatccgagaacaataggtctttagaagaattttatgatgccaaatat atagattcaagaacaaacaaaaagtaagggtggcaagaccttttagtggctgtagagca cgaacaaaaatctaaattcagagaatcttcaataatcgaagcaaaaactatcaaaag gtcaagaagaataagaactcaagatacttttataaccaaaagatttggttaaatatgaa aacaataatctataggttaagggaattcaaaataaggagcctatataaattagaaaattta tctaagccaattaaaactgaattagtagttccatagattgtaggaaggcatttgcgtgatt aa	gtcaacgacccaccacttagcaccgtaggt gcttgaagtggtggccttgagaaggctctg ccctttgaaaagcctagttgaatagcctaagt cttcaagaactacattatgtgggaatatatag taccctagaatgtccctactgttctaggctct aagggttagattaaacatctctaaagggtga ggagaagtgcttacatttgaacccccatata atattggcgaagtgacttaccactactttta ttctaaagtagtgtgaattactctataggag ttgaatcaatg
1884-1885	Ga0265298_10029446 GENOME_ACESSION: AusTroGonzalo MG_8_ \$F_3300028832 GENOME_ID: 279935 CONTIG_ID: 29445 SOURCE: JGI DATE: 2019-06-13.	atggttagtgatgtgctagacagagcggaagccgctgatgccaccgataggacgg cgtgggtggcctatgccttgaagcacggagaagcggaaggttgtagcgagagcccttc accattcagttgtccgcgatagcacgaactatctccaaagggtgaacgcttggtgttgatgt gggaagcaagcacatcggtgctttcgaacaaccgaaagaaggaaactgtattcgcac aggtggagttacgggacgatgtgactaaactctgtggctcgcaggaggttttaggagag gaaggcgaggggcgaacacaaactggtagacacgtgcgaggtgggctaacagagcg aacgaaaagcgcaatgcggcattgcccgaagcataataacaaggcgagcgacacact tccgcgccgtgcagttcgtgtacaactacttccatcagaagacacgggttgagttgg gcaagttcagcttgcagaaataaagaaatccgagcatcggaaggagagaagtaaccaaca agggtgtgcttggcggtgggagaacctgaaagcctatgcaagtagcgtgatgtgtacaa gtgccgtgtgttgcaagagcaagcacaagacggagtaaaagtgggaagtcaccac ataatccgcagggtgacgtggggactgatgttctgagaatgtggtgacattgtgccac gagtgacagagaagaaccaccaaggtgagaaaaagttgagattgaggcgccaccac agcacaaaggcggaagcgaatggaatgctatgcgatactattttagatagcgttccat acgaacaggaagatagagaggacctatggctacaagacggcgacggctcgtcgcgaa cacggcatagagaagaccatgtgtcagatgcgtactgcatagcggggaacttcaaggc ggagcgaagcaacgataatgtgtatctgcacaagttgttgcgcccaacaataggcagtt gcataagacgacctctgaaagggtggctgtcgcaaggcgaatcaagccccgaagtatg tgtttgggttcgctgtttgatgggttcgtattgctggcctaaattgtctgtgtcggtag gaggagcagcgggttcgtttgacatcgaacactggaagggaagaagatttggcgggg gtgtcgtacaagagactcaagccattgactaaatcgaacaatattaactgaaggagg atgtgcgatttctcccaaccctaa	gtcaactccaacccctgaatggttggcctt gtgaggggaacccagcaagcccggttga ctaggggactgaaacggaaaggaaactag aagtggaagtagcagtttaggagagaatgca taggcacccctcgatgtgagtcgaagtcct ggctctcgggtgtcgggttaacagtcctac gaggtaggggcagtgccgcacacgataaa cctctcccaacgcccgatgggcacctaac acaggggaggagaatctaactgtgacttac cgagaagtcggttaaggaaactaagatgga ta
1886-1887	Ga0335394_10082827 GENOME_ACESSION: MAT03spassembly_ \$F_3300032456 GENOME_ID: 280657 CONTIG_ID: 82826 SOURCE: JGI DATE: 2019-06-13.	gtgtggggcgagagttgaccatcgtgggtcagagtattcgggacgcccctgttgagtcgg cgtcagcagaggcattccgtcgtcagcgccagacgcgctatcgtgcggcggttcaa caaccggcggtgtgttgcgggatggttggctccgtcgttgcaaaagtcgttgagaacat ctggacatgggtgtgtgctgtggcgtgcctcctacatcgaatctcaggaactt gtgaggttcgatgtcagttgatggagaacgtgaaatctctggtgtgagtagcagcaa ggcgcggttcagggctacgaagtgcgcgagtagtctcttgagaagtggtggcgccaggt gtgcctattgtggcgaaaagacctgccgtagagattgagcacatcattcccaagacac gcggtgttccaaccgtgtgaccaatctgtctcgtgcgtgccacgattgcaaccagaaga aaggcacacagaccggcgagagttcggtacattcagaaagcaagccaaagc gcccttg	gtaagcatccccacacctaagggtggggc ttttagccccaatgcttacagacttagcctg agaaggctcggttagagacgaatcacagag gcacttcggaatgcctcaccagttccgaact ctgcgtgtaacgattaaacagggggcttgg gagccaacaggtgttgttcacacaaaccgtc tataatttgcgaggtgaactttatccgcg caagcggaaatatgaaaggtaactttcaatg caacgt
1888-1889	Ga0209720_1016953 GENOME_ACESSION: AD_des_11_ \$F_3300025605 GENOME_ID: 21309 CONTIG_ID: 16952 SOURCE: JGI DATE: 2019-01-27.	atgttagtatacgttttaataagcatggtgaagcctttaatgcttgcacacatcaaaagcc agaaaaacttttaagaagcgaagcaaaataacacaaagagaaccattcacaattcaa cttatttatgtagtagtgatacaaaacatttgcattaggtagatgctggaagcaaa tttatcggagtgtctgtacacagaaacgcaagaactatttgcgcagaagtagaactaa gaaacgatatagtacaattattatcagaacgtaggcaaatatcgtataagtaggagatata aaaacaagatatagaaaaccacgatttttaatatagtagcaagtaaaaaataaagggttgg tagcaccttctgttgaataaataacacaaacattttaaataatagaaaaagttcataag ataattgctataactaaaaataatfatagaagtagcttcttcgataatgcaaaagattaaaaatc ctaataatagaaggcgtagataatcaacaagggtgaacagtttaggtatttggaaatgtaagaga atatgttcttggagagacaatcattcttcaaaatttgaaggtaagagcaaaagacaaca	gtcaactccccaacctatagaggttgagc ttgtaaaagctcaagttgactacctaagtc ctcaggactacgttatatgtcataaacct acggcggtttctcctaactgtagctctgtt cttaactttaaaactcgtgtaggtagggat agtgttaagtgtaaaaagcctatataacat tggggaaggagacttactcaaaaggag gtatactttatgta

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ggftaaatgttcatcatatagaaagtaggcaaacaggtggaaatgcacctaacaacctcat aactttgtgtgaggtctgccac	
1890-1891	JGI12150J12686_102082 GENOME_ACESSION: ANMFosRound1_\$F_33 00000857 GENOME_ID: 21413 CONTIG_ID: 325 SOURCE: JGI DATE: 2019-01-27.	atgcagaagttacgcccgaagaatacatcacccccacagtgctgctcaagcttgcgg ctctgtgagaggtggttaaacagttctgagagggtaggaacagtgccgtccagcttaaa aactccgcataacaactccgatgagcacttactccaaccagcgggagggcataaagccg atacgtctgtatatgtgattaataagaacgggaagaccattaatgccgtgtaaacctgc agcgagacatctgctggaagcaggtaaagcagaagttgttcagcggacacatttaccat acgcttcttgggattgtgaagaaatcacacagaacatccctcggcattgacgcagg gtacacaacaatcggtttcagtgctgttaaccacagataaagaactaatgcggggaact cgaactccgcaacgatgtcaaaagactgcttgagaagcggaaagcatacagcgcacg cgaagaagccgtaatgtgtacagaaaaccgcgattcaacaaccgtggaagaaagggt ggcttgcaccgagcatcaaacacaactcaatagccatatcaagcttatcgaaaaactga agaaaatcctaccgatcacgcgaatcatcgtagaggttgccttcttcacacagaaatg cgacaacccggaaatcttaacatcgaatcagcaaggtgaactgcagggtatgagg tcagagaatacctgctcgaagaagggttaggaaatgcgttactgcagcaaaaagtac gttccactggagatagagcagatcgtaccgctgcgagaggtggttcggatcgagctcc aatctcacgatagatgtacagtgcaacagacagcagagggttcggacacccgc agatacagggtcaaacgagagaagtccttgaagcaactgcgtttatgaatatcgtgaa atagaatcgttatatactgggttcgatcagacttatgctatatcacgaaatgaccgc atagaactctatcgaagaaatcacgcgaatgatgcttttgcagctgacggaatg atcaggaaacgcgtacacacatcagctgtaagcaggtgaagacgacagaaccgatca ttgtttaaagctaactcttcaagaggaagtaaaaaaagaacactgtgaaagaaagttaa aggattcaggaggtttgataaagtgaatacagcagcaaaagaatgtttatcacgagact cggagttccgggtacttcgatcttcggacaatcactggagatcgaatcggaaactgtgga acagcaagaagctaacgcttctggaacgagcgagaggaattatcagagagatgtcgc aatctccccgcgcttgcggacgggtctccttgcacgggtga	ccccttgaccctaacgggaccgaggaatgg gcttttaagcccccttagttgatcaggagga tagaaatagcagaagttacgcggaagaaat acatacaccacccagagtgctgctcaagctt gcggtctgtggcaggatggttaaacagttc tgagagggttaggaacagtgccgtccagctt aaaaactccgcataacaactccgatgagca cttactccaaccgcgggagggcataaagc cgatacgtg
1892-1893	Ga0137378_10002753 GENOME_ACESSION: SagtaG_37_\$F_3300012 210 GENOME_ID: 35424 CONTIG_ID: 2752 SOURCE: JGI DATE: 2019-01-27.	atgtcacacgtgttctgtgagcggccacaagcggccacttgcctcgatcatccgggg cgagcacggctctgtcgaaggcggcgagggcgccggtctcaagcgaattcccttcac cctcctgctcaagagcccggagcgaagccgcccggcgagccgctccgttgaagatc gacccgggaagtcgaccacgggtctggcgtggtaggagagacgagtgccgaggtg gtgtggggcggggagctcaccatcaaggagagacgatcgtggagcgttgcgcaaa cggcgagccgtgcgccgaggacgcaggcagcggcagcagcgttaccgacaagcgcg cttcgcgaaccgatgcaggtcgggaaggctggctcccgccctccagaggagccgcgtg caaaacgtggtagcttggctcgaacggctgcgcagcgtgtgctccatcacggcgctctc tgtggagttgtgcttctgatacgaagccatgcaggcaccgggcatcgaggcattcc agtaccagcaggggacgtcgtggcgtacgagacaggggaatatgtgctggagaagt gggggcacagatgcgctattgtgatcgcagcgggctgctcctctggagggtgaacacatt cgaccacagagccgtggggggagcagccgcgagagtaattgacgctctctgtgtgc cgtgttaaccaggccaaggggacgcaggacatccgcgttcttggacatgatccccag cgctcgtgcacattctgcacagggcgaaggcctcgtcgggatgtggcagtggtc tgcgacccgctggcgctctatgaacgggtggcagccttcgggctcccgtagagggg gggagtgccggcgacacaaatacaaccgagccgccaggcattccgaagaccat tggacggatgccgctgctggcgccagcaccgccgagggttgcggaactggcag accgtgcaccctgtgctgattactgccaccggcagcgaatgccgtcagatgtgcaacatg aataagcagggcttccgctgtggaagccaaaggggccgagccgatcacacgggttcc gcagcgggatatgtgtgcggccgtatgcagaaaggcgtgcacctgggacacacgt gggacgtgtgccaatcaagtcgatgctatttaaacacgaccggggccagggtgg tggaaaggcatccatgccgctattgcacccactccatcgtggcgatggctatgggtacg catgtgggagtcggcagcgcttctccccaggcctga	gtcagcgacccacggggaacccgggg gcatgttctgagaagggtgtgccctgatg ctgaccagccccctgagtgcttggagga gagagcggccagggaaccgttcgtggga gcgcgaagagtcggaccctgagatggcctt tcagtcgccagcgcttcaatcccgtgttaa acaggtgcacggggaagcagagccggtg cagcgggaagcggccccaacgaacca ggggcagggaacatgaccgcgaagc ggagggcccaagcggggcagacaacgaa aggaagcagtaggtatgtcacac
1894-1895	SMTZ23_10055571 GENOME_ACESSION: IMG_3300002053_\$F_3 300002053 GENOME_ID: 27600 CONTIG_ID: 55570 SOURCE: JGI DATE: 2019-01-27.	atggttttgttctgagcaaaacaaaaaacactaatgccctgccatccagcaaggccac gagaactgctcaagaaggtcgagcgggtgttcacaaaggtctatcccttcacgatccgtct caaggatagggtcgagggggacgtacaaccgatcaggagagaaggtagaccccgccg ctaagacaacgggtattgtctcgaatcagaagggaagaggttcttttcttgcagagctt aaccacaaaactacgattcatgacaacatgagaanaacgagcaattaccgggtagacg aagaagtgcacaacctcaggtaccgagcaccgaatcaacaatgcacacggccaaat ggatgtgtgccaccgagctccaatcaagggttgatgctattacgcatggactgtgcggt caagaaactatgcccgggtcaccgattgctcgggtgagaccgtacgattgatacgcagaa gctccaaaatcctgagatcaatgtgtgctgagtaaacagggaactatcggatga agtgcgggaataactgttggagaatgggtcgaatgtgcctactcggggaagaaa atgtacccctggagatagagcatatccgaagtcgaggggtggtacaagtcgagtat cgaacctaacccttgcgtgttacacgtgtaacgagaagaaagtaataagaccgcagcg gaatttgggtatccagaagtcgaagaaaggaagcgaaggttccctgaacagggcagctatt gtcaatgcaacacgaatgctctctacagggaactgagtaactgttcataaaatgggttg aggtatctacaggagccgaacgaagtacaacaggcaaggtcggattgcccaagac	gtcaatgacccccacttaaaagtggggctt gctctacggggcagccctcaattgaccag cgctcagcaccggagccaacgacacggct acgatgtgtcagctttagtataaggtacaag accactccgggatgcttctcagtcgccgga cactggaaagttcagcagcagcaaaagttag gggtcgtacacgaacgggctgaacgat acgccgtatactaacattgccgaggggaga tttcaggaatgtcgtatcctgaacgttaccg gcgaagccgagaacaggaggttaaca atg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ccactacttcgatgctctatgtgttgtaaatcaacaccaacgcagcccggtggtttcagga acatagattcaacgccggtgctatcaatcacctgtaaggcaaggacaatacgaagaa cgaatgtgaataaatatggtttccgagactatctcatcgccagagaagatggtcaacgg attccagacgggggatattggtgaaacagctacacaaaggaaaatattctggcaggtg tcgaggagcttggccggttaggaaaaccgggtacttcggtttgatcgtgaataagaaattgg ttgctgaggagtcgaagctcccaatccctattggttcaagattcgatgggtacacata ctctttacagagtcgggtctctccccacttaa	
1896-1897	human gut metagenome genome assembly, contig: NODE_117_length_59318_cov_12.435854, whole genome shotgun sequence GENOME_ACESSION: OLON01.1 GENOME_ID: 7851 CONTIG_ID: 116 SOURCE: WGS DATE: 2019-01-27.	atgtcagttgcagtcattagtaaaacaggcgcaagattgatccaacaagtgaatacaga gcacgtaagctgctcaaatcaggtaaaagcaataaataatagttatcatccgtttaccataca gctcactgaaagaaaaactggaaatattcagccaatagagctttgtatggatacgggat atacatattggaatatctgtaaaagtcagagaagcgaatatctgcagacaagattgaca cattaacagacgaaagaagcaggcataatgcacgccgtatgtataggagccagagacg aaacagaaaaagataccggcagccgcttcaacaatagaaaagataaaggctgg attccccgtctttggaacacaaaaagaatacagcttcagggcaatttcacgcatcaaca atgtcatgcctcttccagataattactatggaatgggaaatttgacactcagggtattaaaag ctaaagaggaaaggacacattaccgcaaggtgtgagattatcaaacgcgtgaacgctac agtatcgcaacgcttcgtgaagcagttttgcccgtgatggttataaatgtcaatgctgcgg cagaacaataaagatggagcagctacacgttcacatcatatataacagaaagccagggg cggaaactaatgtatatacaatctggctacagatgtgataaattgcatacaccgaagaac cataaaccagggtgggaaactatagcttggaaagccaaactccatcattaaaagggtgca actttcatgaccacgatacagtgagcagctttacaacgaggtcaaggctttatccctgataat gatattcacattacatacggagcagcaaaaaagaacggcgccgagagctgaatatcga taaatcacatacatagatgctgtttgtatggggcagtttaccgagacaccgtataaaaag cggttcttataagaaaaaacgaaggataacagatgcttggagaattctatgatcgga gtataattgacagccgcgacggaaagaaacgaagcggtaggaattattaaacggtagga ttaatcgtatcataagaaggattcgtgaaatctgcacagatcgggtacaaaaagtaacc cgaggaaaaaagactgttaggaacaacattacagtgctcagcctcagcatatcatat acgaaagcagaagcgtgaaactacagatgccactgtaattggagcacgggtgatgct actccctgataaaaaatctgtatctattaaaaaatcaaaatatacagatagctggcggtta tttaaatcaacatttaattaa	gtcaataacccgcgccgaggtaaacctcg gacggggcgtgcaaaagaataagtttcttat tccactttgtttagccccgttacagtttacgg aaggctgttggtgtaaaagccttattgatagc gccggggcaccgctggttaccacataagg cttacagccttattgattagcctgtgtagtaa gaagccggccaaaccaggtcaggtataatct ttaactgtcagttaccggtataataggcacc gtgggatgctcctcaagctcagcgtctcg gtatgtgttaaacatcttaggttaggaga tggcgatagagacaaccgattctgactagg cgccgcttgcggagccgaaaagataga gaacctgtatggactgttttaataatcttcg atgtttgaaagcgggttgtagaagacgta agtcaccccgcttgcgggaaaaatataaga ataaaaggaggcagcttatgtca
1898-1899	human gut metagenome genome assembly, contig: NODE_3680_length_4513_cov_6.120682, whole genome shotgun sequence GENOME_ACESSION: UAPB01.1 GENOME_ID: 14988 CONTIG_ID: 3679 SOURCE: WGS DATE: 2019-01-27.	gtgccggcgcaacaaccccttcaacattggcgaaggatcttaaccggcgcaaaagac cgaggtaacaacttgagagattgttttaacaacacgggaaagccgctgatccgtgt tcaccggcaaaaggcccgacttctgctcaaaagaaaaagccggtgagacaaggcgca tgcccttctactatccgggtgacaatcgccaccggcgaaacccaagcagccgggtgacgctc ggcggtgatccggatacaaacacgttggcgttgcagctgtactgaaaggctgagctgt atgcacttgaaagtgaactccgacaggacatcaggagctgctcttgcgcgttggcgcc tgcgtcgcgtcgcgaaaccggcagacggttatcgcgcccccacgctttgacaatcgc gttcgcagcaaacacaaagggttggcttgccttcggtcgagaaccgtatcaacgcgcat ctttcacgattgacgcagctacttcgattttgccgatttcaaaaatagtcacgaaacggca gcctttgatacgcagcttttaaaaggatacaaacattgaaggcgcaacctatcaacaaggc gaacagtttgggtttggaattgacgcgaatacgtgctcttcgagacgggcaggtttgca gcattgtcacggcaagtcgaagaccctgtactcaatgttcatcttgagagtagacgt acaggaggcgatgcgcgaacaatctgacacgctatgcgaacgtgccacaaagcgc gtcatcaaggcaaacctacgctcaagtaaaagcagggttctactataaggccgagacct ttatgggaattatcgatggacgggtcagagatcggtcgaagacagcctatccgacattg agattcgcaacacctacggctatctaccaagcattcgctatttcacatgggattgccaa gtccccattggccgatgctattgcatcgccggctcatcttaaggcaaaacgtctggccggt tattactatcaaaagcaaacccgtcgacacaatcgccagatccataagctgacgatttga agggtggcttgagaagagaaatcaggcagcttacgaaatcaaaaggctttcggttttga caaatgaaaggttcttgagaagaaggctcatattcggtcgaagggttcttactttgatg ttcgcaagcttgacggaacacgtatttcagcggcgatcagtagcaaaaagctgactctgct ggaaaaacgcgaacatttaacagatacagaagaaggcgccgacattcctctctgta atgaattcagagggttccggccgctgttctatga	gtcaactagccatacccgctctaaagacg gaggctggataaaagccttggtagtagcc ttggtagcttcttttgagaggtaactccgt tgggtgggaatgtataggaccgcgaatgt cggttctagtgcgcgtctcggttgggtg ttaaagttctgagagtaggagccgtgccc ggcgcaacaaccccttcaaacattggcg aaggatcttaaccggtcgaaagaccgaggt aacaacattgaga
1900-1901	human gut metagenome genome assembly, contig: NODE_128_length_79994_cov_7.882986, whole genome shotgun sequence GENOME_ACESSION: UPL001.1 GENOME_ID: 16621 CONTIG_ID: 127 SOURCE: WGS DATE: 2019-01-27.	atgatggtatgtcttagatgtaaacggccagccactgatccaacagaaagatgtggct atgtacgcaaacgcgtgaagagcaataaagctcaagtcgaagtggcaatccgtttacaa tccaatgtttatgagacggaaacaatacacaagacattgatcgaatacgtatgctgt tatcagcatattggcatttctgcttgcgtgatacaaaaaggtttccggtgaattaaagct tttagaaaaccaatccaacgattagacgaccggcgagcatatcgctgcacgcgtcgca atcgtttgagatatacgaagccaagatttgataatcgagaagaccggatcaatggttgc accagcattcagcataagctggacagccacaaaaaggtcattcgaaagttgatgtgctt tcttcgattacaaaaatctatgtggaacacgcaaaatttcgattcaagcattaaagaaccc aagcatatccgggtgagaataatcaaaagggaataatgtacgacttccgaaatttgcgtga gtacacttctacagagacggatacacttccagatctgtgtgtaaaatgcttccgagatg gtgcagtgcttcgcatcatatcggatattggaaaacgatcattcaaatacgcctgca aatacattaacgctatgtagcaaggtcatatcgtacgtaatcatcaaaaggcgagatttt	gtcaattccccagactgacgtgtaaacagt cagtcggggcgttgtagtaattacaagcca tagttgattagctaaagcattcgagtctac gttatgttttaagcagcttcgctgctccaaat aggtaaccaagggtgctccacaagccctt gcactacggtgtattgtaaaaatctttagtg ggtcgaagacgtgcagtacacaacaacaa ttcataacattgacagtggtgttaccacag accgggcttggctcgtctcgcgctggt cattccagaccaaaattattgaaaggagccat atcttatgatg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		<p>gtatggtctgcaaccaagcaaaagcttcaaacggagacatttatgtctacgggtcaga cgaatgctaatactactgctgctgatgaatacacaattccgggtgtagagacgttggcta tttgacaaaacccaacggatagattgacgtggacaaaacacattacaacgatgcgtat tgtatcggcgacaaacagccaaagcaccgggtgtaaacgggtgttctggcaagagaagc gtaagaataatcgtgtctgtaaaagttctacgatgccaaatattggatgcaagacgggt caaaaggctaccggtaagaattatttaacggctgaacaactcgcaacaaaacgtgaat gggtgagaactgcatccatacagacagcaaaagatttcaaaaggcaggatgtctgtcaga agacaacgatattccgtaaacccgatgatacgggtgttggcgttaacagaacatttgaa gtgtgtggcgcacaaaaccgtggcgttacgtatctatcaaaaataacgatttcaaaaag ttgtaagtaaccaagcgattacagccattcaaatcgcgaaacaggttataatcacgcatag</p>	
1902-1903	human gut metagenome genome assembly, contig: NODE_1355_length_19542_cov_13.053583, whole genome shotgun sequence GENOME_ACESSION: UEPG01.1 GENOME_ID: 15076 CONTIG_ID: 1354 SOURCE: WGS DATE: 2019-01-27.	<p>atgccctgttaccggcaaaagcgcgccttctcttaagagaagaaagctattgtgaag aggcgaacgcctttactattcagctgacgattgcaacgggtgagtcacaaacagccgggtg actctgggtgttgatccgggtacaaacatgtgccccttccgcatcaacggaaaaggct gagctttatgcatcagaagtcgaactcgtcaggacgctctctgactgtctctgtctgtc tgcgttacggcagtcgccgtaaccgcaaaacgcgtaccgcgcgcgcgagattcaac aacctgtctccgacaaagcgaaagcgtggttctcgcgtggttgaacccgaatcaa cgcgcaactgtcgcgcgatagaagcgggttctcgcgtggttgaacccgaatcaa ggaaacggcgtcttgcgactgacgtgctgaagaaacccgacatttgaaggaaagata ccaggaggagaaacagctcgaactctggaacacccgcgagatgttcttgcagagacg ggcatgttgcagcattgttacggcagatcaaaagacccgggtcttaattgtcatcatttg gaaagcagacgtacggcgagattcaccggcaacctgattacgtctgtgagacgtg ccataagcccttcatcgcagtgaaatcacgtgaaggcaaacgcggacaatcgttcc gcgttggccttcatgggaattatgcgtggaaaggtgctga</p>	<p>gtcaactacctcggcctaagggcgcggcctt gaaaagccctctagttgactagcctcagcc gtcatttggcggactatgttggcgggatttta taggcaccgcgggagtcagatcctagtccc gcgcctcgtcggccgaggtttaaagctctga gaggtaggagcgggtgtcgtcggcttgaac cccttccaacattggcgaggatcacaaac ggctgaagaccggagagataaaacttga ga</p>
1904-1905	Marinitoga sp. 38H-ov contig_19, whole genome shotgun sequence GENOME_ACESSION: GCA_011057715.1_AS M1105771v1_genomic GENOME_ID: 327122 CONTIG_ID: 10 SOURCE: NCBI Prokaryotes DATE: 2020-04-22.	<p>atgaaaagaaagaggtacagaagaacaagaataccgcaaaagaagatatagaaaa ccacgtgttctaaatagacgcaggcgtgaagattggcttcccaaatatacaatgaaag ttaatgcacatataaaaattatttaactatagcaaaagatactaccagtcaaaaaagttgtg ttgaaatagcacctgttgaacacaaaatagtttaactctgattgaaggttaagaatatac aagaaggtcagcaaaaaggttttgggatgttagagagactgttggcgtgcgagagata caaatctgagacatcaggtaaaaaggtatattggaagtacatcacgtcataccaagaag ccaaggcgggaactgataatcctccaatttaattttaaactgcgaagaccacaaagcta ttcatgaagataaaacttaaaattcaagttacagcttaagaaaatttaaaacttaaaagatg cgagtcattgtttcaacaatgctggttattgtgaaccacttaaaacaaaatattgattg gtataacatatgttagtgcataaaatcaaaagaatagaaataggattggagaaacaca taggaacgatgcttttgcatttctggaggaaataattttaaaaaatttagaccgttga</p>	<p>gtcaattaccgcaacctatagaggtgggtg ctgttaaagccatagttgactacccctcagcc aggggaggttaactctctatcgggctacgtt agctctggctataacacctggagatctgctc aagttccagctctgttctgtcattaaacaa tcctgagtgtagggatagtgccgagaca tggcaagccagtcataacttgggatgagc acctactccaataaataggcttaccatcat atg</p>
1906-1907	JGI25616J43925_10007359 GENOME_ACESSION: GraSoi2013_100cm_2_\$F_3300002917 GENOME_ID: 26954 CONTIG_ID: 7358 SOURCE: JGI DATE: 2019-01-27.	<p>atgatgtgtgtatgtacgtaacacgatggggcggctctatgccttgcacaccagcca aagcacgaaattgttacgagcagcagcgaagctgtagcctaccgaccccttacc attcaactacctggcagtcgaggggaaggtccaggagatcacctgtggattgataag ggcagcagcattacagcgctgctgttaggcaatggcaccgtgttctcgcagcaga gatccatcatcgtcgcgatgcaaaagacaaaatggaggatcggcgtgaccgacgcaag agtcgtcagcagcgcctctgtatcgcgcggcccttctcaatcgtcgtcgtcagtaag cgagtgggcgttgcgcgcctgatcaaaacacgctcgaagaggtgatccgggtcgt ccggcagctcccttaccatcagcgcgcctgtgattgaggtgtccaggtggacatcgc aagactcaacaacccccacttacaggggccacgatcaagaccgacgcggtggtgga gagaatctgcgcattgcctgtctgatcgcgacgggtgtatgcctgccagcactgcgggaa acgggggacgcgtctggaggccaccactgatcttctgcaaaacggaggcaaggat acgcttgcgaattgtgacctgtgtgagggtgtcaccaccggctgcacgaaggaaaa attaccctgaaggcagcgggagtgggcgggacacctggatcagatagcccaacgcacta tgcaaggaaaaacccatctatgcagcactgagcgcgttcgcacctctgacaaccgtgt ttggctatgaaacatctgcataccggaaatctgtaatttgcgaaaacccatattgattg cactgtgtattgccacgcttacgaccggcgaagtggttgcctcgcagagcgcataatct accacatcagcttgcggcaggcaaacgcgcaagcggatcacagcttgcgcgagaaa gggaaaggccgctgaaatatacgttcaatggcgaactcgaagcgttgcgaaggcg atttggctgtgtcaaggagaagtatatgaagctgtcaatttcaatttgcactggctatc tggccttccgcgcgtcaagggggaacccaatgctgcccgtcccaagactgtcgaatct tagaacgagagggaactgtgttgggaaaaatgtggaataa</p>	<p>gatagatatagttgctaaaaataggcgaa aatcttccggcgttattcagactaccagcg ggcctctcggaggcagcagcttgcagagc gatcacaccttgggatgttgcctcagcttca agctctgtgggtatgattaaaggtagcgga aacgtgaacgtgtgctaccgtaaaaacgtc gtcgaacagtcgcgagagcaattcactcc gacaggagggtcttatagccatgtgtg</p>
1908-1909	Bacillus sp. AFS014408 AFS014408_98_D11_Co ntig119_140514, whole genome shotgun sequence GENOME_ACESSION: GCA_002557915.1_AS M255791v1_genomic GENOME_ID: 141756 CONTIG_ID: 20	<p>atgcgtgtattgtcaagaatgtaagaggagaaccgctatgccttgcagtaatcgaagg cacggctcttctcaacaaggaaaagcaaaattgtgaagtacacaccatttacgattca actcctatatgccaccgggtaaacgggtgcaaccgttacaattggcgttgcagcggagc aaagcatacggattgctgattactactgcagataaagtgttagcaaaaggaaaccca gctgcgtcaagacgtcaaaagaaatcttatacgaagctacattacgcagaggtagaag acaacgaaaaacaagatcgaagaatgctgtttctcaaccgaaaaaagaagaaggat gggtaccgccatgattcaaaagcagagtggaatacaaaatcattggattgaacattcgc tctgtattaccatctcaaaagtgattgtggaagtaggggaatttgatgcacaaaagctaaa aaaccccaatatacagggaacagaatatacaacagagagatacttttggcttttgaatcag agatactacgtatttgcgagagacaactacacctgtcaaatgtgaagaaaaagggtggtat</p>	<p>gtcaactaccaccgcttaaacgtaacacg ttttgaagtggggcttgaaaaagctcgtgt tgtctagcctcagcttctgtgactcgttgc taggttgcatacccaagatgattccctagt cttggctctatgggtcgtctgaacgttctga ttgggaagggaacggtaaccacatgccttct tgcatgagaagttgctacacacataacatt ggcgaagggaacaaactcttagaggagga caaacatgcgt</p>

[illegible]

SEQ ID NO:	Name	IscB sequence	ωRNAs
	NCBI_Prokaryotes DATE: 2019-01-27.		
1916-1917	Ga0121719_100007 GENOME_ACESSION: 3300011832 GENOME_ID: 239423 CONTIG_ID: 6 SOURCE: JGI_Dark_Harvest DATE: 2019-01-27.	atgccccgtaaagtggatgtaaacacgggtatattttagtggtgaccagcatggtcagcc gttgatgccgtgtttgtcgtcacgagctaaaaattgctgaccagaaacgcgctcgaatt acagcatgcaaccgttcaccattcagatattggaccgatcggtaatagcagtgcttaca gagagcgaaftgaaattggatccgggcagcaaaaccaccgggatggccttgggtcaca aggggatccgaggggtgttcgtggccatggccatccattgaacaccgaagtcacat tatccgtgatgctgtaattcgcgaagtcgaatccgacgttcgcgcacgcaaccgcaaac ccgttaccgcccaccacgggtcaaaatcgttcaagaaagccagaacagcacaaaggtt acaatgtatggttccccgccttgggtgcacgaattgaaaacattgtaactctgtaatac agtgtggcaaatattccccgtaacacagattggcggagagtttaacaagttgacacgca caaaatgcataatcctgaggtcaatggtattgaatataacgcggcacattgtatgggttg aagtgcgagaatactgttgcaaaagtacaaatacctgtgtgtactgtcagaagtgtgca tttcaaaataaataatccacgggtttaaattggaattgatcagtggtgccccgttccaaa ggcggcagtgatcaagtgaacaattgggtgtggtgtgcacagttgcaaccaaaca ggcaatcaatccctcaagaatttctcaaaagaaaaccgtctgtgtcgaagctattcaag agagtttggggaaccattgaacttaacagcgcagctcagatgaacaccatgcgcgaa gagttgtatcaggtgttgaaaaaaggaccgggttgcctgtgaaaacctatagcgtctc aaaccaagtcaaccgcgtgcagcagaacaccacaaagaccattggattgatgcagcc tgtgtaggtgagagcggaaaccaggttaattatcagcacatctctcatattattttgaaa gcagtggggcgcgcgaatcgtcagatgtgctaatggtatgctatgggttccccgaacc aaagcaaaacagtcgaagcgtacatcacgtccatgggtcaaggttccaaaccgggtgat atagttcgaactaagcagccttcggggaataccaaggaacacatacgtactgtttcga tccgtgccagaggtcaacttgatataccacacaaaacggcgaaaaatcacatccatt ggaaaaatttgcgaataatctccggctcaacgggtatcacgttaaccacaaacgagtatct tatggttga	ctcaaccagcagatattgttaaaaaaagg agggaattaaaaacacatacagctgata catgaatatataggctctgtgtgttacc agcaagcagctctacggccacagttaaac agaatttctcatgtggaacagttgtatg tcgacaatgctttgtattggtaaaccaag aacgcgt
1918-1919	Ga0307928_10003230 GENOME_ACESSION: AceLakmetagen232_2_\$ F_3300031227 GENOME_ID: 279923 CONTIG_ID: 3229 SOURCE: JGI DATE: 2019-06-13.	atgcaaaaagtatatgtacaacaaggatggagaatctttaatgcctactgaaggtatg gtgcagtaaggagatggcttagagacaacaagcaactgttgaacctaactccttttac aatcaaataaactttgattgtgagaacaacacagaagtagttgttaggtttagatactg gtgcagtttaattgtgttcagcagtaaacacaatgggtgtttgtatgcttccgaaacaa aactaagaacagacatacataagaaaatgcagagaagagttatgtacaggaaaaactaga agagttgaaaaattgagatatagacagccaagatttcataataggacaagacctaaagc tggctgccgccagcctaaaaagtaagttggactcaactgttaaggtataaaggcaattag cagagatctaccaataagtaagtgagagtgagattgctaaagttgacacgcagaaatt gcagaaccagagattcaaggtaaaagggtatcagaaggggttactgaaggatataata atgttcgggcttatgtgtttgagagggataagtatcctgccaaatattgaagaaacgaga aggaaatttgcacactaccatacatacaaaagaaagatgaaggaaatagacctga taattctgccactgtccacaagattgttcattctgattttcacaaaggatttaattcaacaca tttagaaaaccaaaggcatataagatgcaatccaggtaacaatctaaaagatttcattg tagtgaaactaaagaaaactcaaaagtgaagttacttctggtatatacaaaagagaata ggatgagattaaacttccaaaatctattgtttgtatgcagttgcaatatgtaactcaaaaa gattgaaaggctaaactatattcaaaaggaaattgataactcaaaagggtatcagatg actaaaggcataagagtgagaagaagttgcctaaaggagaattgtttgggttcagcga tgggataagggttaagattaataatcgggttgatttatcaaggttaaaagaagttcaggatt tttgacgtttgtatattgatggaacaacatacattcaattaaatacacaacactcaa agattatgtgggaataataatggaggttaagcgtttctctccacgactaaagtcaggg tatccgaaatgcaaaagtattatga	gtaacacattcccacctctaaaggggg gcttttagtaactaagcctgtgttaccagac ttagcaaaaggaaattgtctagttattgagatg tagaaaccaacaattgtcagttgttggct ctttctgtgacactgtaaaagtcttggaggatt cgagacagtcacaccattgcgaagctcg ataacattgtgagagcaacataacagtttga aggaaactggttacgcaaaaactctcata aacaattgcacaaa
1920-1921	[Bacillus] selenitireducens MLS10 chromosome, complete genome GENOME_ACESSION: GCA_000093085.1_AS M9308v1_genomic GENOME_ID: 40342 CONTIG_ID: 0 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgttcgtgtacgtattaaacaaaacggcggaacctttaatgccgtgcaaacgagaaaa gcacgattattgtgaacgaaaagaaagcaaaagttgtgaagcaaaccttttacgatcc aactcgtttacggttcaagtggtcacaacacgcctgtttcattgggtgtggatgccggaac caagcacatcggatcagcgcaacgacggagaaaagacgtgttatftgaaggtgacgttca actgagaacggacattcaggataaactggcaacaagcggcagtagccgacgatcaaga agaagccgaaaaacaagggtacagacaacggagattcgacaatcgaagaaagaaaaa ggatggcttcccccttcgtgcaacataaagtgattcgatatacctgtttgtgaaaaagt ccatgcgttggccgatcacaacaataacagttgaagtgccacagtttgatattcagaaaa tcaaaaagcctgaaatcgaagggaagggtatcaacaaggcgatcagtttgggttttga acgtcagggagtagctgtttttcagagacaaccatcgttgcagcactgcaaaagggaat cgaaagacaagatactgaattgtcatcacatgaaagccgtaagacaggcgggtgacagc cctgacaaccttctaacgttatgcgaaacgtgccacaagaagattcaccgtgaaggactt gaacatttatttcagcgaacacgtacaccttgcgtgacgcttcacagatggcggtcatgc gttggtttatctaacgggctgaaagcccgctatcctgatgcgaagctgacgtatggttat ctgacgaagaatacccgatagaacacagcttagagaaatctcatgctcgtatgccgt tgtatcagtggtatccgacggcaaaaccaagcgaagccacgtatcagtaacagcaggt tcgcaagaacaacccgcaattacacagacgaacgattttaaaggcgggattcgcaagg caaacaaagcgggaacgtttcgtgaaagggttccaattatcagataaggtgctttgcgaagg	gtcaacaaccaacgactaaagtctggggc ttgtaaaagccctgttgactagcctaagctt tcgaggactacgttgatcggtcaagacac cttacgggtcggattctagctgaagccctgt cgcttgtgattaaaagtcgaacgggtgtg ggcgggtgtgcaagcgttaaaagccttccaa cattggcgaagaatacctgactcctgaaagg aggtagacgttatgttc

SEQ ID NO:	Name	IscB sequence	ωRNAs
		<p>gcaacctgcttcttttggcgacgaaaaacagggtcgtttgattgctgattgctgatg gcacagttatcagcaggggaaaagtataaaaaatagcgttaaaagaaaagccacg tcgttgcttgttgaaggagtgaaaccgtccacaggaaacccccgaattcattcgggggt ttcctgtggacgtatcttatgaagaatcagacagaaggcctgatactggccattgaaacga gctgtgatgagaccgagcccccgtcattaaaaacggcaatgcgattctgagtaacgtg gtggcttcgcatcgcagatcagatcataaacgcttcggcggcgttgcgggagatcgcattc agacatcatgtggatcagatgacactgacgatcgtatgaagcgtccatcagggcggggt gacctcaggatctcgtatgccattgcccgtcacggaaaggaccggggctcatcgggtgcac ttctcgtaggcgtgcatcggctaaagcgtatcgtttacgctcgggatcccgcttattgt gtacaccatacgcagggcatactacgcgaacctctgacagaacctttgacataccctc tcgtctactcgtcgtttccggcgggcatacggaaactgatctacatggaagcggacggga cgtatgaactcagggggagaccgtgacgatccgtaggtgagggcgtatgacaaagtc gcgcggacactcggctcccgtatccggaggtccgcagattgaccgcttgcgaaag aaggggagggccaccatcgcctgccgagagcctggctcagggaagggtcttttgatttc agcttcagcggattgaagtcgcgggtgatcaacaccttcacaatgccaaacaaaagg acagacgataccgatggcgatctggtcgcgagctttcaggcaagtgatgatgtgctt gtggataagacgatccggcggtgaaggagcgggggtggacagactgggtcattgca gggggaggtgcagcgaacagcggccttcgggagggccttgatggcgcccttgaaag gcccggcatccgtctgaccttccccgatcctctgtaccgataacgccgccatgatc gtgcagcggcgtttgtcagaccggcgcggtgtattcagtgatgaccggatgaacggga atccgggtcttgagcttaagaccgttaactaaagcggagtaa</p>	
1922-1923	Ga0393278_0005903 GENOME_ACESSION: LZ_38metaG_FD_\$F_33 00036985 GENOME_ID: 312543 CONTIG_ID: 5902 SOURCE: JGI DATE: 2020-04-07.	<p>atgttagtctttgtctgaataagatggaagccattatgcctgtggaaccaagcaggc aagaaggcttctcagatggtaaaagcaaaagttaagagccggataccattacgattaaa ttactttatgctctctgtttacaacaacaatttcattaaaagttagattcgtgttctaaaggt gtaggatgtgcagcagtaagaagtgtatggggaaacattatatacaagtgaggtgaaac tagaaaagatattcataaaaggatgacgcgcagcttcttataagaactagaagaggt aggaaaaactgttatcgtcaagcaagatttagtaactgtttagatcaaaaagggttgcttac cccaacattaagatctaaaaactcaactcattgaaagaaattaattatgtaaaatcaattttac caataagtgaattgatttgaagcggcatctttgtatattcataaaattataatccgaaat aaatggaattgggtatcaacaagggtcaacaaaaaattattataatgtaaaagcatttgtatt tcataagacaaatattgtttgcaaaaatgtaaaaacaaaaaataatgtaaaattaaatgt tcatacatattatttgcctctgataatggaacaaattctccggataatttaattacattatgta aacctgtcatgataaaatacattctcattatgacaaaaaaagaaatctttaaattacaaaa gaaaaagaaaaatatacaaccgatgctacacaagtttataaattgttcttatttaaaaaag catctgcaattcaagaagttttgggttatgaaactaaatttaataagaaacaaattggattac ccaaagcgcatttttagatgctatgtatttggattgactgatggtaggttgcaaaatgc caacttatgttttaaaaaaattagatttaacaaaggagattatcaagaaccaaaggatca gatcagaattggaatcccaatgggaaaaaataatgggatttaaaaaactggatagggttaa gtattttaaactacagcttttgaaggtagaatgtcaacagggttatgcaattttaaagac atttatggaaatgaattaaatttggacatactcacaactgtattctatgaaaagatttggag ctaggaaaacatgtcttacaatcaaaatttcatagaaagttcacatcaaatatcatattat tctgttgaatatatgaaaaaacatcttcaagaaaaaaaattatag</p>	<p>tgaagcctcccttagactaaagatctagggg cttccaaaagaaaggagaagacaaggga gctacttcaccagaacaccaagagagga tctaaaatgttgtaaacgatagggtgaata tataggtacttcagaatgcctactcagttctga actctacggatgtaatttaaaagtgcagaga ggagttcgggtgattacatttataaacctac ccatatccgttcgagagtaagtcgtgtgaaa tta</p>
1924-1925	Ga0209719_1005600 GENOME_ACESSION: AD_des_12_\$F_3300025 677 GENOME_ID: 21310 CONTIG_ID: 5599 SOURCE: JGI DATE: 2019-01-27.	<p>atgcagaagttatcagagagacaatttaaacacacctacgagtgcttactagctcgtg gctctgtgttccgtctttaaacagagaggaactctcagtgagcggagcttctcgacctct gataacaatcccgaagtgaacaacgccacatacaggcggtctttaaagacaaatgtgtat gtattagcagtagatggcaaacattatgcatgttcaagagcgaagctagaaggttat taagagcaacaaagctacaattgttaacataatccattcactataaaactaactttgaat gtgaaaaccaaattcagctattacattaggagtagatactggatatacaatataaggattat ctgctaaatcagaaaaggcagagtattgtcttcagaaagttgtattgcgaacatttctcctt tattaacagaaaaaaaatgtatcgtcgcggaagacgcaataaactgtgtatagaaac caagattcttgaatagaaaacgtaaaaaaggatggtttacctcaagttatgattaccgcata aattctcatataaaactcattgagaagatgttcttgttaccatttccattatgttgaa gtagcaaaactttgacatacagaaactcaagaatcttgaataaaagggttaggttatcaac aggagatctataggatagaaaatataaaatcgtatttaatcgaacgcgaacacgccag atgtcaactatgccatgaaagatcaacacgtactaattcattcagagtgaccatataatac aaaaaagttaaagggtgcacggacaacacctgacaattctgttgccttcacgaaaaatgtca tactaaattacacaaagagaatttaggacaccttaataccaagaacaaatacaaaagca gaaacctttatgtcaattatgaggaatacgttgaactgaactaagaaatcgcatacgggt aatggaaacctttgtcatattacaagatgagaagaaacacattgaacattgaaaagctc catattaatgatgttttgtatagccaaggcgagtaaccaagttcgtctgcgcctctgact ataattcagaaacggcataacaatagggtcttcaattgaacaggaagggttcaacctc cagttagacaccagagatatccctaccaacccaagacgtagttatgattcaagggtctta ttatgatgttactgggacttttaacaaaggttcttggatagggtaacaaaagcaggaacag tgttcaacttttctacaaaagaggtgaaagacattatgtaactaacgggttggcgaattc cacacctgaaggatgtgtcttcttgcctccacaggataa</p>	<p>gtcaactaccacgtcctaaggcctaggtt gaacgggtgacgttcgacgcaagagttgatta gggagccttaggtgaaaatctatgcagaagtt atcagagagacaatttaataacacctacg agtgtctactagctcgtagctctgtgttcc gtctttaaacagagagaaactcagtgaga cggagcttctcgacctctgatacaatcccg aagtgaacaacgccacatacaggcggtct taagacaaat</p>

SEQ ID NO:	Name	IscB sequence	ωRNAs
1926-1927	Ga0394882_0003342 GENOME_ACESSION: WSA_S4_MG_FD_\$F_3 300037422 GENOME_ID: 312718 CONTIG_ID: 3341 SOURCE: JGI DATE: 2020-04-07.	ttgggttacgacacccctgggatgctagctagctccctgcccctgcatctgcaatfaaaca gttctaaggctacgtggagcagtggtgcaggctgaacaagccagataacaaggtcgaag ctaacgttacgagcaatcagaaggacgcaacaatgtctaatttcgtttgtactgacac caatcgcaaacccgctctcaccctgcaaacacggggtagcgagatcttgcataatgctgg caaggcgccgggtattccgtcgatttccgttcacgaattatttgcacaaggaggtagaggca acgacctgaacctatcgaaacttaagttagatcccggttctaaacaactggaatcgcttaaa acaaggcaacaaatcattgttggcgcagagctaaccatcgggggcacgccatcaaa gccagictgtattctgctggtcattgcgcccggggagagcttctagacatactgctatc gtcaagcacggttcgaatcgacccgctcccgatgggttgggttagcccaagctgaagc atcggttagaacaacggttgacctggtttgagcgattctgcaaggttgaccgattggctc aattgttcaagaactggtaggggtcgacctacaacagaggtcgaacctgaaatctcagg cgttgaatcagcaggcggaattgcaagggtacaggtgcgcgagctacgtctgaaca aatgggacgggaagtgtgcttactgtgctaaagcaacggggcatggtactcgccaatgt gtcgaccgataagtattgggttcccttctcggtatcgccctcgctcaagtttgcgaaggc tttcagacaggcgacattgtgaaagcagttgtaccacggcgaagaggtcgggtgaatac gtcggacgagttgcagtcgatcagtcggttcgtcaacatcacaatctcaaggattagt gcaaggcatcaatcacaagtactgcaaatcaattcaacgaaaggtggtatagctatgg cttttga	gtcactaacccccgactaaagttcgggggct tgcagactccccgacgctacgcgcggga cttgcagccgtagtgaccagaccaccga aagggtagccgttatttgggttacgacccc tgggtagcgtagctagctccctgcccgtca tctgcaataaacagttctaaggctactggag cagttgttcaggtcgaacaagccagataa caaggtcgaagctaacgttacgagcaatca gaaaggacgcaacaatgtctaatttc
1928-1929	Ga0334896_1003582 GENOME_ACESSION: GraSluR3_MG_18_\$F_3 300033428 GENOME_ID: 280364 CONTIG_ID: 3581 SOURCE: JGI DATE: 2019-06-13.	ttgggtattgtgctaagtaaaacaaaaaccacttaactcttgaataatgcagtagccaga aagctactaaacaagggaagcagttattcataagaataatcttttactattaggcttaaa tatataattaatgagcctaagtaaaaggaatacactttaaaatagatccaggagctaaagt actggcgtagccattgtagaaaaaagccataaccatgctaaagtagtttttctagctaattta gagcatagacaaaatattaaatctatattagatgataagagcatttagaaggaactaggc gaaatagaaaaactagatatagaagccaagattttcaaatcgaacaaggactaggggtt ggctaccaccttctatcagtcataatgctgtaataattgagcttggactaaaaagataaag aaattatgcaatcacacgcatagctgttgaacacgtacgcttgacactcaattaatgga caatcctaattatgagggtgtgaaatacaacaaggtacgcttttaggggtacgagcttcgtg agtacctaataataatgaggcatgagtgcaatattgcaaaaggcgaatcagaagacag cgttcttaatatggaacataatgatttcaaggctaattgaggtagtaataagatctatctaat actttgtctctgctgaccttgaatgaggataaaggctcacttaattatctaatgtttagat tttaaaaacacaatcagaacaaagtttaataagagagaaatlaaaacatagaagttatcc ttaaaaaaggacttcttaagctttaaagatgctgctaagggttaattcatcaagaaaagcta cctatagagttgaatgaattatactcagatttagaaggtatcatcagggggcaggactaaat ttaatagaactacttcaaatctctaaacacactactatttgcgtgctatgcgtagggaaga atacttctgatagttttaccctccaaagcattgaaagtgctaaatattaaagctatagcca gagggtcaagaagctgactaacftaaataaataatggcttccctagaagctacttaccaga caaaagtatttctcgtttccaaactggagatttagtaaaagctgaaatcccaaatggcaa gtatcaaggcgctgtattatggctctgtagctgtgctagtaaaaggtagctttgacattaaac tactgatgggaagagaataagcactaactacaaatacttttcaataacaaagacttgatg gttacaactatgggttgaggatattgcttttaa	gtcaacttctccagctttaaagcgggagct ttttgtaaaagagctcatagttgaccagctta agctttaaagactatgttattaggttatgaca ccatagggtacttctctagctctatgcaactgt cgattagctttaaacaactcctggcagaga agagtatgtgctaatcttaaaaagcctttata accttgacgaagagagactgtctttgagaca gcgttacctacattttgtagaggaaggagttt taccttgg
1930-1931	Ga0101770_1016828 GENOME_ACESSION: IMG_3300006674_\$F_3 300006674 GENOME_ID: 27981 CONTIG_ID: 7458 SOURCE: JGI DATE: 2019-01-27.	ttgggtattgtgctggacaaacacaagaacccctgatgcgctgcaccgagaagcgtgc gcggtgctgcttcaaaagaggccgcggtggtgcacagatagcgcctttaccattc ggftaaaggaccgggtgagagaagaagcgctgcagccgttgcggttgaaggtcaa gcctggcgcaaaaagagaccgacatggcagctcctgcgggaagagaccacagcgcaatc cgaagctgcatgctgcttgaactcggcacaagccggcataaaacagagccttgaca gccgcaggggccatgcggcgagcaggagaacccgaagaccggtaccgcaggccc cgggttcgacaaccgcaatcctgcgaagtgcggcgctgcggcaggaacgcgagacat ggcaaccgttactgcgctccttgcgaganancnanaaaantttgctgataacgggttacgt gagaagcggctgccggcttccctggaagcccgctgcgaaccagatcatgaacacagtca agagactgaccaggtgctcctctatagctgcgatacatacgaaacgccaaattcgaca cccaactgctgcagaaccccgagattgccggggtgagtagaccagcaggcgagcttttt ggctatgaagtcagggaatacctgttagagaagtgggggccgagtcgctacttgcgg caaaaacggcggtgccgttagaatacagacatcgtgcccaatccaggcgccgaccc gacaggggtgccaaccttaccctggcctgtaaaactgcacaccagagaaggggacnc tgacggcngaagaattngccancccgaggtgcaggccgaagccagcgcccaatg aaagagggaagcgatagtaacgaacccgctggaagctgttcaaccagttaaagaat cgcccttctctggaatgcggcacggcgccctcacgaagaagcagcgcatgagca gggttgcgaagagcaaatattacgaagcctgcgtatggggaagtagctcccggca aactggctatcaggcagaatacgttcaggtctggacagcagttggccggcgagccg gaagatgtgcaacacagacaagtagcttttccggtcagccaccggggccagcagaag atcgcttgggtttatgaccgggactgtgtgtgtgcccagatacccaaggaaatac gctggtcgtgggttggcagggtggccgtgaagctgcagcggtacttctgacgtaaaaga cggcgcnaggcaggcggtgtgcccagggtgtccacaggcatttccgctgttacag cgtgccaacgggttgcgaatacgaagaagaaaattagcggcgctga	gtcatagaccccacgctaaagagggggg cttgcgtgggaataactccggcaggtctccg ttttaagcggaagctgctcttctggcggc gctatgacccgccccagcgtggagccga tggaccaggcgatgaaccacgctacgttaa ggaggctcatcacttctggaatgcttcttag ttccttgccttgggtgcggcattaaacagc gagcgggggtgtcagcagcaggtgtgccg cacgcaaaaagccttcaacttggccag gagagacgccccggcgagtcggcccg gagcgtcacaagccccgttaagggtgtcc ggaaggagatgctcaaatg

SEQ ID NO:	Name	IscB sequence	ωRNAs
1932-1933	Streptomyces sp. NBS 14/10 NBS-14-10_1_paired_trimmed_paired_contig_41, whole genome shotgun sequence GENOME_ACESSION: GCA_002224125.1 AS M222412v1_genomic GENOME_ID: 126671 CONTIG_ID: 63 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	gtgctcctcagcagcttgcctcgaattcagtcagcagacaacccgggaacaggga cgaacgggactcgaacaccgcgtcgggtaccgggcgcggaacctgtgcgagagg agaccgcaccgcttcacctggcgtatgcgggcatcaccccaaccaccgggtgggga gaagccgcgtgagcggcaccattcgtgttcctcgaagcacggcacaccgttc aacctgcactgcggccgtgcccgcgaatgctcaagaaggacgtgcgggtgtcca ccgtcacaccattcgtatccctcaagaccgcaccgcgcccagtcgagaca gacgggttgagatcggtatcgaccgggcagcaagcacaccggtatcgccctgtca ccgccgcgcggggagcgccgggcccgtacagcgtgcagctcgaccaccgggg cgcgacgatccgaagaagctggagcagcgtgccctaccggcgccgacggcgca cccggaaacctgcttaccgcgcaccgaggttcccaaccgcgccggcgcgagggt ggctgccgcctgctgcggcaccgggtggacaccactgtgtcgtgggtatcccgtc gcccgtggggcaccggtgagggccgtgcacgtggagcgggtggcgttcgacactac gcactcgcggccggctgccctggaagcgccgagttaccggggcaccctgca cggctacgaagtcgcgaatacctgctgcgaagtggaaaccggacgtgcctactgc ggcgcgaccggcggtccgtgaacatcgaccacatcccccgtcagcgccggc tccgaccggtatcgaaacctggccctggcctgcccctgtacgagaccaagtga ccggcttgggacgagttcctggcccacaagccgaacctctggtcgatcaaggcac aggccaaagcaccgttgcgggacggcgacgggtacagtcaccgcacggccctat ggcgccgcttgacggccgtggcccacgcacaccgctccggggggcgaccaag tggaaaccgacccgctgccacctggccaagaccacactctcgacggcctcgccgtag gcacgctcgacacaatcgacaaaacctgaccacggcttgcctcgccggtatgcacgg gcgcggtaccataccgcaccgcacgaacaagcacggctccccggttgcgcatg ccccgcgcaagcggttcttgggttccgaccggcgacactgacacccacgtacc caccggcaagaaggccgggacacacaccggacgggtagcagtcgccaccaccgga gcttcaacatcaccaccggccacggcaccgtccaaggcatccggcacaagcagctccg tctctccaagaagcagacggatacgctacaccaccttccgagcattccacagcgat accccgacgcaccgcggcgaggactga	gtaaccgggtcccccgctgaggcgggg gcttcagccgaactggctgagggccgtgt gttaccagcaccagccaacacctttggagg aggtagccgaagtggctacgttccgcga cgacagaagaccaccaggctgtgcctct cagcagctgtcttggaaftcgatcagcag acaacccgggaacagggacgaacggg actgaacaccgcgtcggttaccggcg cggaacctgtcgagaggagaccggcacc gcttcacctggcgatcggggcatcaccccc aaccaccgggtggggagaagccgcgtga gcggcaccattc
1934-1935	Ga0334813_001619 GENOME_ACESSION: 713E1X1metaG_FD_5F_3300036884 GENOME_ID: 312156 CONTIG_ID: 1618 SOURCE: JGI DATE: 2020-04-07.	ttgacaaggggcctcgaaaaggcgggcagaagttgtttggcgcgacaccatgggata ttccgcacgtcccatgcaactgtgcgcgggtgtttaaagtcctgaggcgaaggacgggt gctgccagcagcagaacctttataacaacccgatcggtccttaacctctcgtcggag agagaaaggaaactttatgttagttccagttatatacaaaagcggacgggtccctgatgcct ggcaccgcgctaaagcaaggcggatgctcaggggaagcagggccaccgtagtcaaac gaacgcccgtttgtgatcaagcttgcatacggcaggttccgggtataggcagccgatgacgt ggggacggcagacgggttattgtatgcggcatttccggggtaacggacggaaaagag gtctatgcggcggatgtgggttccggaaaagacatggaataactcaacgggagagaa accactatcgagaggccgaagacagcgccatcgtgtgatcgaaccccggttcga caaccgggaagaacccgaaggctggttgcgccttattcagaacaagctggacacc aggtaaaactatcgagaaaacagcgcgatcttgcgcatcacacgcgtagctgtggaa gtcggcgatcgatatacgaagatacaaaacccggatcagggacactgatacca gaacggagcgcagaagggaattctggaacgtccgtgagtatgtacgacagggagac catactgtcagcagtcgaagggttaaatcgaagaccctgttttcacgggtccaccacatc gaaacccgtcagaccggtggaacccggccggacaatctcgttacagtggttaaacctg tcacgggaagatacaaggagagatcgccccgaatttaag	gtacaagactcgggaaggcgttatcctgg gaggacagggcaggtcgttggccacatg acctcggcgatcgtatggttggcgccca gtcgggtgttcatgacgacgtccccgaaa ccaggtgtgttcgggttaccatagggccc cacagacaatgtcgtcgccgagaggcatg cgtttcacaacttcggagatcggaataac gtagcctccatgttgaaggcgtcgagaaac tggaaaggacaaaaataagaccac agcaccaaatcctgagtcgtcactgcgtca gaatcaaatatagaa
1936-1937	Ga0209617_10003712 GENOME_ACESSION: LakEri2011SPAdes_5F_3300027720 GENOME_ID: 31100 CONTIG_ID: 3711 SOURCE: JGI DATE: 2019-01-27.	atgcgagtagctgaatttcagtagacaatttccattatgcccagccaaacctagtcgcgc tagacgttgattaaagaaagtaagcaattgttaattcaacaagttagttattctatgt ccagttaatagctgaacctatggcatgaaaactcaagaatagtaattgattagacca ggcaaaatgtttctggttagcgggtcaatctcaaaaatcaccttacaatgttgacatt gttttaccattcaaaactgttaagacagaatggaacaacggctctatgttgcgacgcgta gacgtggttagacgagttaccgaaatgttttcaataaaccgtatgacaccgtcaagctagg tttgataatcgtcgggggttaacttctccaagtattcagcacaacaagatttagaataat cggacaatcacactctttgtgataattaccgggttaaaacaaattgtcgtgaagagtaga agcaagaggaaataagatttttagtcagtaattgttaggtcaagatatcaaatcaaccga ttatctgaactagcaaaagtagaattgaagaaaggatgggaacatcgaaatcttctgaac acttgggtttacacaaagcaagctctgataagctcttcaaaatcccagactcacgcagt tgacgctgtgacactagcgtgtccgaatttgaataacaagctgtgggaagggtgctaaa actcatgtgtgattgtgggtgggaatgcaattattacaaattcaatttacgatttacgcc gtcctcaatcagtcgcagacagctacacttaattgttttcaaaaggtggaatagacg aaaatacggagggtactgtgaccagacatggctatagtaagggtgattatgtcgaagctact caaggcaataaaacatataaggctgggtgaagtgagacacaaaactcaagttcagtt agtattttaactggaacgggttagggcaattcagtaaaaacaaagtcgactaattagac gctctattggactaattatcactcgggttaaaaccgtcgcgtcgttccctctcagactaa agtga	gtcaataactcagggttgaagccactgagct tgtaagaaacagttctacgtgtttagtagc tcattgagatgcagtttggtagcaactccga atacttcccagttcggattatctcaaatat cttgaataatgttgcgttgggaagacatct taactcgttgaagcaggggacttaaacctt actccaaggattatctttagtcga
1938-1939	Ga0133913_10206209 GENOME_ACESSION: norCanCoassembly_5F_3	atgaacagagttttttttagatacaaataaacaccactagctccctgcacaccagcaag agctagacagttgttagatacaaatagcgtgctccctatcgatataatccattcagattat attgcatagagaagtagaattagtcacagaataactattctattaaacttagaccagggtg	gatgcaataataagtttttagtaggttttta aactatcttaattgttcaaccagactaagttatt agatttatctaataactcgttatgaagaaaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
	300010885 GENOME_ID: 37731 CONTIG_ID: 206208 SOURCE: JGI DATE: 2019-01-27.	cagtactaccggtatagctattatcggtcattccctaacaaggtgatgtgtgttttgg tgcagaaataccataaatctaaagctattaaagctaaattatgcccgtagcagtttctg tctaatagacgtagccgtttgcgtcacagagaacaaagatttgataagaaactagaaaag atggatggttaccaccatccctcgaatcacgggtgaattgcattactcatttctttaataaatt taaacatttaatacttaagtctaattatgtgcaatattgaattaccaaaatcgatacacaataa atgaataatcctaataatcaaaaattatcaataaacaagggggtatgactaatttgataata caaaagattatttaattcctagagatggggaattctgttctattgtgtgctacagatattaa attatttaaaagaactgttgcgcctcgcgcaacaggttgtaattcagtttaataatctggtgt atcatgcctagcttgcatacaaaaaatcctaataaccagttgatgatttc	gttaaagacttacatggtgtcttccagc ccgtagctctaagagtacagaactaaacag gggctaatacgatggcaaccgtgggtgtac atgaacttcttataactttgtcaggagacc gttactagggtggagaccacatttacacagg attcacatgaacaga
1940-1941	Ga0265319_1000103 GENOME_ACESSION: 4824metaG_2_SF_33000 28563 GENOME_ID: 279827 CONTIG_ID: 102 SOURCE: JGI DATE: 2019-06-13.	atgcagcagttaaaaacagagaactaaaaagaacacacacctacgaatgcttcacaagtctg agcaactgtgttactattaaacaagaggaaactctagtgtagtagactcaaaacctg tttaacaatcccgatgtgaatcaagttcaacatacagaacggctattaaaaccaactgtgt atgtaataattataatggcaaccactaatgcctgtgtatgctaaatctaaagattagt gaaaaaaatgctgcaagtgttaattgaattataccggtttacaattaaactaaattgaattg aaaaacaaggtaaccagttacgttaggaattgatgtgtgtataacaacattggttctct tgtactagtgaataatctgaattatctctggaactctgttttagatgataaaacatctaaaa gattgacagagagaagaatgtatctgtgaataagaagaataaactatggtatagaaaac caaaattttaataagaagaagaagagggtgtgttagcaccatcaatccaagaagata tgatactcattatcttgaataaaaaataaaactatattacattatagcagaggtgattatgg agaccgctaatttcgatatacaaaagatcgaatacctgaaataacagggaattgattacca acaaggtaatatgtataattatcaaaatgttagaagtatttaattgctagagaaaagggt gtgtcagttatgtaataaagaatttacaagaagtaactcaagtcatacatcactgtaaac aagaagcaaaaatggatcaaataggcgcaaaagaatctgtcattattacatgaaaaatgcat accaaacttcataaacaggattaaaactaaaacctgctaaaattataaatcaaatatggtt atgtcaatcatccgtaagagggttttggaatgatatactgtattgaaagtacttatggtacat aaccttttttaaaaggcaagaatttggaataataaaagtcataataacgatgcttttatact gcaaatggatcaatacagaagaataaaaatcaattataaatacaaaaacatagaaaata atagagcaatcaattaaacagaagggtttaaaccctctatagcgaacaagaatgc aatacaaccaaaggacttaatttgatgtatacaaaaagatgtgttattgtatatacaaaa tagcggagcttatataagggtgaaattgcaagaaaatttacctgtaagccagattataaa gattataactttggaagtttaacatacaataattaa	gtcaatcgccctttagctaaggaaggg gctgaaccgtaaggttaagggttaacgggt tgattagggggcaaaaaatfaaaaaatg cagcagttaaaaacagagaactaaaaagaac acacacaggaatgcttcacaagttcgtagca actgtgttactattaaacaagaggaaact cttagttagtagactcaaaacctgttttaac aatcccgatgtgaatcaagttcaacatacag aacggtcattaaaaccaactgt
1942-1943	Ga0207156_10850 GENOME_ACESSION: Alteio_BWMinHeat_179_SF_3300020942 GENOME_ID: 22706 CONTIG_ID: 849 SOURCE: JGI DATE: 2019-01-27.	atggcagttattgtattagacaaaaaacatcaaccttfaatgccctgcactgagaaacggg cccggctattattaaaaaaagggaaggtcggatccattctatgtatcttttaccgattcgac tgattaaaccgaactgtcaggaaggtgtctacacacctttgctgtgaatgatcgaacagga agcaaaagtacaggtattgtctatgtacgagaagactcgaacaagaatgtattctctct tgtagaattactcatcgttgatcagtcacgtcgaagcattgcaagcacgtgcgggtgaag cgaagacgaagacgttctcactaagataccgctccccacgttgcgaatagaactaaa cccaacggatggtcctccgctcttgcgcacatcgatagagactgtgatgagctgggtt catcgattaaaaagccagagcttgcattacagccattacatcgagagggttagttttgatac acaaaaaatgacctctccgaataacagaatgttgaatcacagccaggggactctttggg atacgaatctgggagttattgctgaaagatggggacggaagtgtactattgcaatgttc aaaatgtgcccttgcgaatcgaccacatagttcctaaaagttgtggaggtccaatcgatt gataatctcaccattgacctgcgcctctgtatcagaaaaaaagaatctacctctcacaat atttaatccaaactagcgcaaaagttaacctctataacacctctgcgagatgcagctgcag taaatagcacacgtatgcccctatggcagcaactgatcaagcttttactccctgtgaggcg ggtagtgaggggcaacaataataatcgcaatcggttaggaatcccgaacacatgct ttagatgctgctgcacaggcaccactttcagatccaaaattggcaaatcccactctaca gatccgtgtcaccggaaagggttcttaccaaagaactcgtacagatcgctttgctttctc gaggatttcttctccgc	gtcaactacctctccctaaaggagaaagctt gaaagggaactaacaagcttgggtgaccag ggaaagcggttaaccaaccgctacgtata acaagggttaagcccactccggagtgctc ctcagctccggactctggaagcagcggttg cagacaaccgattgggtaaggacgaaacg gatcgtgcaaatcgccgtgtataacattcc cagggggagcaagccgaaaggctccgtta caaggcccgtaaggcgatttttagagaatta tgga
1944-1945	Ga0187869_10014458 GENOME_ACESSION: June2016WEW_13_1_4_SF_3300018030 GENOME_ID: 30392 CONTIG_ID: 14457 SOURCE: JGI DATE: 2019-01-27.	atgccatgccacctcgaagagcaagggaactgttctgtgcaggaaaggcaagcgtc ggttcaacaagggtctgttctacatcctcctcactaaaaggctggacggattcaaacagcc agtggcacttggaattgatacaggttcgaattgggaagggttcacgggtcaaatcagaagc ccatacttttgagaatctgaacgctgatgcagtcacgtgggtcaaaagataaaatagagact cgcaggaatatgctcgtgcacggcgatcgaatacctcatalctcaatgccgttgg aatcgttcaacaaagcgaactacgacggattcagacttgcaccttcagccatgcacgtt gggattggaactacgtctcgaaggtggtcagaaagttgatccatctcgtgttcgtc gtagaagacattaaagccgaaactaagaagggtgtgaagaagtggaatacctccttagt ccgatagagaatggcaagaagtggtttatcaggaaacggaagttaggcggtgtgaa accaagcaaggctatgaaactttcgatgcgaatcagatgagtcgaagaactaaag gacaagaagccaaggtttctatgctcactgcgtagactcgtgactcttccaactggtt gtcgggtgacatacaaaagcctgagtttgagcatgtcacitgctcactccgtttcagga agcgacgacaattgcacaaactccaaaagtcgtggtgttctgaagaggaggagggc tctaccaatagtctcggattcaaacgaggcagtttgcgaagcatgtcaaggtaggcttgg cttattgtgggggtttatcaaaaacagaatcagtcgtcactccgttcagacagcgcaacg	gtcaatccccacgcctaaaggcgggggct tgaaggtcaaaactaacaggttctggcaat caagttaaactgtctagacaagtctcgaa tcgaaagattggagcaacaaggtgtctgtac gtgttagtagcgggttatgtcaggatgttc gctagtcttcgaccttaagacagcctgtgtc gaagcgggtattagcgtcctcgggacga gaaaaacgtagggttaactttatg

SEQ ID NO:	Name	IscB sequence	ωRNAs
		tttagacaaaactcaaacagcagattgcaagtttcttcccacagcagttggagggtt atgcaccagcataa	
1946-1947	Ga0207639_10055049 GENOME_ACESSION: KBSCorC32SPAdes_\$F_3300026041 GENOME_ID: 30472 CONTIG_ID: 55048 SOURCE: JGI DATE: 2019-01-27.	ttgcgttacgcatttgttttagacaacatcaacagcctctcatgccctgccatccagccag ggcaagacaattgtgtcccaaggcaagcgctgtatttcgttgcattccctttaccattat tctcttggaaagagattctggggaattgtctggccttgatttgaatacagatccaggcagca aaacgacagggcgtgttttaaccgttgatcggaaaaaaacagggtccaacggtacttggg cggctgagttacatcatcgtgtcaagcgattatgaaaaatttgcgtctcgttggccatt cgtlaggggaagacgcatcgaaaaactcgtatcgagcgcctgtttccttaacgttgtc gggcaacgggtaaatagcgccttccctgcaatctagagttgacaattgtgaacaactaac gctgaaattagacgaatcgctcccatcgtatttgggttgaacgagtcgaattgata cccaaaaactgcaatcctccggaagttcattgcaatcaacagggaaacctgtttgtgt tatgaactcgcagaatatttggaaaaatgcaacgaaagtcgcctattgcgatgtgac tcgaaattaccttagcctgtcgagctgcaacaaaaaagggtgcccaatactgcact tttttctggctaaagaaccggcacgtctgaagaagctcaaaagcacagctcaagctcctc acgagatgctgctgccgtcaacaccacccgaaaagccttacaaaactgcctatctgcctt acccataccggcttgcacgtaccgggtgctgaaaccaagtataaccgttgcacaacagg ctatcccaaaagcacattgattgatgccgctgctggggcagccaagggtattcggttttt taccgcctcaaatctgctctctgcaaatccagccatggggcgaggttcaaggcaaatgt gccgctggaataaagcaggcttcccagaacgagcgcaaaatcagccaaacgcgttca aggttttcaaacaggagattgggtagttgctatggttgcctacaggaaaaaaacaggat gcacagagccgagttgccgttcgggctactggaatttcaatatcaagaccttttaggg gttattcaagggtatccggcgtattgtcaggtccagcatcgactggatggatactatta tcaatataccgctatagacgcgttctccccctgctgaagcagggggtctccgcgcgag gtcctatgaacgagcgagctagctggcttggctggactccggcgtctacatcctcagca acagaaagccgacggcggaatgagatcctgtcaggatttgaacttgcattggcgttgaat atccaggcaatagggcctatcaaaagcgtctctctccagatggcgattggttgcggctg a	gtcaactacccctccgttaaagcgggaggctt gcagggaaccttggtaaccgtaagcctct gaggtagccagactcagctccactctctttg attcttctgggtcaggcggttagggattacct taaggcagaatctgctctgttcagagcagcg agataggtaccctgggatgcttccagctcc cagggttctacggttagagttaaacaggcata agaggtggaaagccagctgcttaaaaaaaa cctgccattaagttgtcaggagaccgttac ctgcgaagcagagcgctgtgaacaggc actttttagaaggagttaaagccattgcttta c
1948-1949	Ga0315282_10012226 GENOME_ACESSION: YL17G07_20_MG_2_\$F_3300032069 GENOME_ID: 281633 CONTIG_ID: 12225 SOURCE: JGI DATE: 2019-06-13.	atgcactaccgcaggaagaggttcaaacccgaagcgatctaaaaactgcttgcctc aacgctaaggagcaaaaaagatgctatagtaaatgctgtaaacgggatagcaaaagatact tccgttaaagaaagtaacgattgaacacgcttctttgacactcagggcaataaaaaacagga cataagataccgaactggttataccagaaggctccgtatagatgaagagaacataaaa gcatactgaaggaagaggatgattatctgccaatatttggccaagaccttcacggc aaacgttggtaagtagaccacataaagccgaagtcagagcgccgacagatgacctga caatatggtggcatctgtgaagactgcaataaaagtaaggataatttgaactctggacgaa tgggtaaaagtattggaggtctatcccaacgggaatcaacaaaaagcgattaaggcggtta ccgaagataagagagcaggttaaaaaatctgctggttaggtaccgccatgtcacaagtat gaagaatgcattgggttaaggaataagcgcaacatttccagttcaagagaccaacgggtgt actactaaactgatgcgggaaagcataaacctacctaataacacactgcaatgatgcgata gcgatagcattggacacgtctaaagaataagtagaccgttgcactaatgtatcgaataaagc aggttaggaagaaaaatcgactcttcatgaagcaataccaggaaaggagggggcaa gcctaataagagagggcaagcgcaacaggaagaacataaaagagattgtgataaacaac aagaaatggttgcctatgggcaaaagtatatatacacagaaaggttaagattggatataaa tcggctttacagccaagtggttatataccaggacatagcggtatagatacctcaattatc atccaatacaaacagatcaatccgaagaggttaaaactgatatgcaggaacaacaatta ttttatagaaaatatgtcgtttcatcttccaatga	gtcaactacccgctctaaagaggcaggctt gaaaaagccctcaagttgactagccctagcc agggcagaatgccatcggtctacgttactc gctgtatgataccttttagatgataccctagtc tgaagctctatcacctataatccaacaggttc caggggtagggagcaggtatgggcatgta agcatgataaactggcggaaggcgaataa actccgaaaggagatacactggatgtta
1950-1951	Ga0395714_000473 GENOME_ACESSION: 712E3XmetaG_FD_\$F_3300037225 GENOME_ID: 312105 CONTIG_ID: 472 SOURCE: JGI DATE: 2020-04-07.	atgcagaagtttaagtaagttaaagacgtacactacagatactcccaagctgttagct ctacaattctgtattaaacaaagaccgaagctttagtgcaggaataaagtactgactttaa acaacctcgatgtggaatctactccaacataaaggaggagcagggttgcagagttcctgttat tgccttatgtacttaatacaacggggaattctgttgatgccttgcctgcaggaagcaagatt acttctaaaaagagagatgcccatgtctgtaaaaactaatccatttttgttattcagttaaat cgcgcaagtggaagaacaagttccaagaatgttcatttggatcgattcaggttctaagaata tcggattttcgttattacaacaaaaaagaattgacagtggtacattgattttagataataa aacatccgacagactcactgaacgcgcaatgtaccgaagaggacgtagaataaattgtt ggtatcgtaaaccaagatttaataatagaagaagatagagaaggttggctaccaccgtctgc acagagaaagttgatactatattacatttaataataaacttaaaagttattaccaatcaaa aatattacaattgaagttggtaattttgatattcaaaagatagaaaatccagatattgcaggta tacaatatcaacaaggttcaatgtttgaatatcaaaatagagaagtttcttggtagcaagag aacaaggtaaatgtcaattgtgcaataaagagtttagtaaaagtaattcttcgcacattcacc atataatacaaaaaagtaagcggttactgataggaaaagaatttagcattacttcacga atcttgcataaaaaaattacataaaaccgatcttttagtttgcgtaaaaaagaataaacaat aaagatacaacatttatgaatataatcgtatggagatttagaaaaatattcctgattgcaaat taacttatggttaatgaacattcgtttaaagaataattttgagattagaagaagactattaca atgatcccttgttatcgtgtgtggaatagtcacaaactaaagttcaacctattttttaaaca aaaaacataaaaaataacagagttcttcaacttaatagaaaaaggctttaaacatcaataaga	gtcaactacacttctaaagaagtaggcttg aatcgtgaggttcaacgtaagagttgattag agggcaaaaggttaaatcttgcagaagttta aagtaaaagttaaagacgtactacagatac tccacaagctgtgtagcttacaattctgtatt aaacaagaccaaagctttagtgcagagat aaagtagtactttaaacacctcgatgtgga tctactccaacataaggaggacagggtt gagagttcctgttatt

SEQ ID NO:	Name	IscB sequence	ωRNAs
		ggcagagggtattctattcaaccggagatataattggataaaagataaaaaatattgtta agggtgtcattgttatggaaaatggatcctttgtacaaaataattaaaaattttagtta ttaaaaagtagaaaatgttttcatacacatcaatctatfaa	
1952-1953	Ga0137369_10007808 GENOME_ACESSION: SagtaG_48_SF_3300012 355 GENOME_ID: 35435 CONTIG_ID: 7807 SOURCE: JGI DATE: 2019-01-27.	gtggttttcgttctcaagtgcagccagccgttgaccgtgccacgagggcgcgggc gaggaagctgctggcagcgggagaagccgtcagatcgcctgtatccctataccattc ggctgctccaccgcagagcagccgaatcgcagcagcccgaccggctcaagctcg atccgggcagcaaacgaccggcctcgcctcgtgcagaggatcagccccgcgtggt gtgggctgcccgaactcggccatcgcggccagcagatagcgcgtgctcttggccgg cgggcgctccggcagccggagacacgcaaacgcgttaccggcctgcacggtttc tcaaccggcgcgcaaacgagcgtggtcctccatcgtcttgcacgcagtgctgacta cgctgacctgggcaagaaggctgcggcggtgtgtccgatcggtgccatcagatgga gttgggtgcgttcgacaccagcgctccaaaatcctgaaatcgcaggagtgccatacc agcggcgacattatcggctggaggtctggaattacctgctgatgaagtggggccatc cgtgcctactcggcgcgactaactgcgctggagcaggagcacctgatcccccg caccgggggtggcagcgaccgcatcacgaactcaccatcgctgccggccctgcaac gagcgcaagggcaacctgactgccatcaggttcggttccgcagcttatggccaggc gaagcgccgctcaaatgctcgtcgtggaacagcgccgggtggcactctatcag cgctactagcgacgggctgcccgtcaggtgggcaactggcgagcaggaagtg aatcgccccggctcggctgggacaagaccattggcgtgatcgggcagtggtgga cgagtagccggacacgctccacgtggcggttaggaagcgtgctgctatcgccgta caggccatggccggccagatgtgcccggacaacaagatggttccccacgcgcca tgtaccgcggcgaacgccacctggcttcaaacggcgatcgtggtgctgctggtac tcccgagcggaaaagcagtaggaacgcacgtcgggcggttcgcatcgcagtagg gcaagtcaattgtctacgctgctgattgggtacaaaggcatccaccacgctattgtca gatcatccagcgtggcgaggggtacgcttatggctatcgaagggagggcggttaa	attgtgtcaggaaccgcagcgtctggcg gcggctgctgacccgtcgtgaagccctgctg accagccccagctcgggcaaccgagctac ggccgtcagtcacgacacctccggctgcg cgggcccagggcgagccctgtcgcctgtc agaacgttcgctggggcaaaaggcgcat ctggcgggcgcgacaagctcagcgcgcat ggggcagggccacttaaccggcgcaagg cgagatgcacccgtaaggggcaaggat cattgtg
1954-1955	human gut metagenome genome assembly, contig: NODE_4586_length_312 1_cov_1.805284, whole genome shotgun sequence GENOME_ACESSION: OHFR01.1 GENOME_ID: 4996 CONTIG_ID: 4585 SOURCE: WGS DATE: 2019-01-27.	gtgtgtctattgcaaacctcttattacattggcgaaggtgtcatataggctcgatattgc cgcttaccgctttaagcatatcacaaaaaggaggtgccttatgttatgtattagata tcgatggacagccacttatgcaactcaaggcacggaaaagtaagaaggctgttgaata gtcatcttgctaaagtagtaaacggtgttcatttaccgatacagctttatatacaagtaaaa agaacacagcctgtatctctgtgtgtagacgaggaagtatagatgttggctgtctgct acaacagagcgggaaggtgtctatcaggaagaactgttccctcgaatgataattgtaaac tgttatcagcaagacgagcattgctgcggctcgaaggaatcgaagacgagatacaga agggttagattcaacaatcgtgtgcatcttaaacacaaggaatggctgctccatctgtag aagtcagatagcagagcagatcacagccacaagaaatctgtcaaatctgctatatt tgatattcatgtagaacagcagaatcgtatcagcgtctaaaagcgaatggaagaagga aggctacttctgctggtacagattatcagcttggggaacagtagattctataacacaag acagtagtgccttcacgagatgaatatactgtcagtggtgtgtacgcatgataacaatgt aaaactacatgtgcatcatgtcgaagccgtcagacaggtggagaggcacctaaacaatct gattacactttgtgagcactgtcataatcgtacatcggggaagatgatgtccaaaa ggcaagaacagagcgaagtctatcgtgatgctgctttatgggtatcatcgtaatactct gctagaacgtttaaaaaaagaatgaattgtccagtgacgatgacatat	gtcaactaccggccaatgaattggcaggctt gtaaaagccctaagtgtactagcctaagtgttt tgagcactagtaagagagaatatagtta cccgtgatgtaacacatagctcggcctc aagaagatacattaaaaagtctgtaggga ggactgtgtgtgtcttattgcaaacctcttatt aacattggcggaaggtgtcatatacggctcga tatgtccggcttaccgcttaagcatatcac aaaaaggaggtgctcttatag
1956-1957	human gut metagenome genome assembly, contig: NODE_15547_length_43 08_cov_2.449800, whole genome shotgun sequence GENOME_ACESSION: OIXA01.1 GENOME_ID: 6111 CONTIG_ID: 15546 SOURCE: WGS DATE: 2019-01-27.	atgtgacggctgtctgacgggagggcgaagccggaaggaacttatgttagcagttgt cacgaacactgggaagccccgatgcccagcagatatacgggcacgggaagctctta aaggcgagcgtgcccgtgatctataataccgtccttactttaccatcatgctcgtggacc gggcagaagggggaagtacaggatagcagtagaagtcgataccggctccaacatgc cgggatcagtgctgtacagagaacaggaacttttaagcgaacagcgtgacctccgg ataaggagccggaacaccacaagaacagcagcgttaccggcgggagaaggaac aggaaacggtaccggaagccgaagtgcataaccggaagaagaaccgcaggaagg ccatgaaaaatggctgacacgcacaacccccaagcgtggagatccaggtaaacctgt tcaaaacttctgccagggtatgctggtaacctctgcttactttgaaatgggaaagtgtgata cccagggtactgaaagcagtcctggaagggaagacaaataccacaggggtgaagattacca gagaggggagcagtagcgggacggatacgttactgctgcttctcagggatgact acacctgaggaatcgtcgggcggaaccattaaggacggtgccatcccatgtacacat gtggggtactggatgcaggaccggaccaaccggcctgcaaacctgccacggatgtga gcagtgccatacgcctgcgaacctgagcaggaacgggacgtgtatggaaggaaacct gagcatgggactttaaaggatgcctcctatatacgcgttgcctggtgatgctgaagg aatfaaaggatgccggccgggaagtgaagtacatgttacctatgggtcaccacaaga ggaaagccaggcctgcacctccaagagccatgtcaacgatgccttttccatgggg cggttccatccaagaagggcggtactgcttactggaagaagaccatccggcatga ccgttccctgcagaagtctatgatgcggtatacctggatacaggggacggcagagaga aaaggcgagtgaaacttccaacgggaagatcagccggaaccacaagaaggacagtga gaacctgcataagtagccgggcacaaaggtgtcaaaaggccatgtttccatccgagg acgggaacaaactgaagcctggaagcgtgtcctctataacggggaaggtgacggta cacggcactcatatctacaggaaaataaaaggcggaagggtcaagacgtca	gtcaattaccacgggtcatatcctgacggaat ggccggggcttctgatggaagtaactttt tggcgatattctccagcccggttgattactc taagccgggtccggaagtctttttgatactgc cgggtaacggctacgttactggaatatac aggcaccaggggatcttctcaagtcctgt ctctgcggcgtgtgttaaacatcccttttt ggtagggggaagtgcgcagccggcaaaa ccctttgtaacattgaggttgagaacacttg ctatgtaagaagcccgaggacacagga gctgggaggtgctcccggtacatggtacg gctgtctgacgggagggcgtaagccggaag ggaa

SEQ ID NO:	Name	IscB sequence	ωRNAs
		atgtccagtttacaagaccggcatctgcagcgaagagctgtcttactgaaaaatgtaa gatcgtacaaggaactataacacgggctggaaacgataccggcctgcataa	
1958-1959	Succinivibrio dextrinosolvens strain Z6 chromosome, complete genome GENOME_ACESSION: GCA_011065405.1_AS M1106540v1_genomic GENOME_ID: 327272 CONTIG_ID: 0 SOURCE: NCBI_Prokaryotes DATE: 2020-04-22	atgggttatgttttaaaftgcgaaggtgtaccaatcatgccgactactcggcacggtatggta agacgtttgtctcaaagaacataaagcaaaagtagttaaagatgtccatttacaattcagct tagatatgagtgctgcaaacattaccaatctacacacttggtagacacgggatafgaaa atgttggtatttctgcagtactgagaaaaagttcttttgaagctaaagctaaatcagaa ctgatatggtaaagctctgtctacaagaagagaacaagactactcgtagatcaagaa aaacaagatattcgaagctgttttttaaccgaagtattgaaaaaggttggtactcccat ctatcagagctaaagtgaatctcatttaagattgttgcgaaggttcacgaattttgcctat agacaaaattgtcgtagagattgcaagttttgacattcagaaaatcaaaaatcctgataca aaagtgtcgtagatcagcaaggtgatcaattaggattcaagaatatgaaagcttatgtcttat gtagagatagatgtactgttagatgttgcgaaggttaagctaaagacaaaattttgcgagtt caccatcttgaatcaagactataggtgtgtgactcctaacaatctcgtgactttgtgcaatt cgtgtcacacgaattttcaaaaggtttgatttcttgaattcagattaaagaggaaatcttt caaagctgaaactttcattgggtgatgcgaatttcttttcaaagaactttgcgataaatat ccatgtgttaaatgcacattatggttataccaagcatttaagagagaatattcatttaag aaagatcatcatgccgactcgtcgttcatttgcgtacaccgcaagctttaccaagtgatg tctatctgtagaaaaaggtaagatgtcataatcagacagactataagtttactaccttaaaag gaggtcgaagaaagttgaatcaagcaccgtatacggtaaaaggtttcaggcttttcgaca aggttaaagttggcgatcaaatggatttttttgaagaagaacaagtgacagtttgata tacgcaccgtatcaggtgaacgtatcaatggtatcaataagttataaagacttcaattgatt gaaactagaaaaagttgggttaattgaaaaattgttaa	ctctgcagacagattgatgatattttcattata gatatacatgagagagaattataaaagctct gatgtagatgagattgatatacggaaatctgc aactaaagcttagtctcagcaaaatttactttc aaaagtagattaaagctacgttatgaaagaata gataacaatttaaatgttacgtcagaattgtt ctacctagttctgacctctaaggtctatgatta aacagttttcattgaaaacagagttgttagaca acaacacttcaataacattgacgaaggttag cattttacagttttgcaggatctggcttacagc aaaaaactgcaacttttaaggaattaaatt atg
1960-1961	Ga0311329_10041539 GENOME_ACESSION: I_Bbly_SF_3300029907 GENOME_ID: 280532 CONTIG_ID: 41538 SOURCE: JGI DATE: 2019-06-13.	atgccgtgtactccgaagcgagcgcgaagctgctcgaagcgggtcgtgcacgcgtac atcagctgattccgtttgctattcgaactactgactgctgttgcacaaatctaaactttcagc ccttaaaaaataagctggaccaggttagtaaggatcaggtatgacgtttgtcgtgtgtc agaaaaaacggatggcgacgggtgttgaacacgacgatcatatcctgtttttatgaa ctggtgcacccgagcgcagcaatfaaaaaatcattgcaggctcgcagcagaatcgcgtgg acgtcgtagggtaactcgtcgtatcgcgcgcgcgggttcaataaccgcacgaatcga aaggatggctgcccttccctcaacatcgggttgacaccttattcttgggttaaccga ataggcaggttggcggccattactcatctagcccaagaacttgggtgcgattcgtatgcaac tgatgcaaaacgcgaataatccggagtgccctaccacaaggcgcagctggcaggcta cgaagtacgcgaatcctgctagctaaatcaaccgaacttgcctatttgcgatgcaacg gaggtaccgtcgcagatagagcataatcgaagcaaaagcaaacgcagggtctaatcggg tcagtaacttaacgctcgcctgcaagaagtgtatcagcgaaggggtccttggatattcgc gacctttttgccaagaacccggcacgtctggcgcgaataactaaagcaggttaaaagcacc gttacgagacgccgcggcgttaaatgcaacgcgtgggcctctttaacatgctaaaaac aaaccggctctccgggtcgaaccggcagcggcgggtgcaaaactaaataacatgcactcgc ttaggcgtgattaaaacgcattggttgatgcggctgtgttgggacaa	gtcaactaccccgccctaaaggatgagctt gaaaagcgaagccgattaaagctcagggtg accgagaaaggttagtaaacggaagttta ctgagtactactgttgaagtagtacaacaa cgacccgggatgcttttccagttccggata gaaattcgaagatttctcgaagttgcgg tagcagacaagctacaggggtatgcagaaa cggaccgcaacaaaacagaaaggtatcgct gttgaagccgggttacaacattctcaaggaa agcgggacccgggaaggtctccgtcacaagg cgcgtaagcgcactgttttaccataaagaag tgattttacttaacgaagaagaacgcaatt acgtgaaa
1962-1963	Ga0210051_1023303 GENOME_ACESSION: CryGey14_aSPAdes_2_\$ F_3300025868 GENOME_ID: 280081 CONTIG_ID: 23302 SOURCE: JGI DATE: 2019-06-13.	atgaaaaatacacacaagaagttaggaagagaataacatacacactacaataacttca caagttgtagctctgtgtgtgtctttaaacagagaggaaactctcagtgagcatggccta aaaacctttcaaacagcctcgtatgtgaacaaacagcttggcaagactggacaggattg agaattcctgttttaatatatcgtggaacacctttgatgccacaacaccagcaaaagcaa ggcatttgtgaacaaagaaagcaaaagtatcaaaagaaaacatttgtttatacaattg acaatagcaacgggagagactaagcaaaagattacattaggaatagatagcggatagatg tcaagtaggattcagcgaagacagaagaaagaaagtgtatattagagaaactaacttt aagaaaagatgtatccaaaagttagaagaaagaaaggtatagaaaacagaagaga aataagttatggtacagagaagcaagattgataaccgagtaatttcaaaagaaattgttag accaagcatacagcacaattggatacacatgttagattgattgaaaaaatcaaaagtttac ttccaatcacaaaacaataattgaaatagccaactttgatacacagaagatgcagaatcc agaaatatacaggaataaaatatacaaaaggtgaattacaagggtatcatatagagaatat cttttagacaagtttggcaggaaattgctttatgcaatcaaaagggtgttcccttagagggtt gaacatattatttcaaaatcaaaaggaggaagtaacagagcatctaaccttacaatctcat gcagtaaatgcaattttagaaaaaggtatatttaacagtagaagatttgggttccataacatt caaaaattagcaaaacatctttaaagcaacagcatttatgaattgtgaaggaagagac tgactaggcaggtaaatgcagatgaacatttggctatatacacaataatagaaggatttag ataacttaacaaaagatcatgtcaacgatctttgtgattgtcgtgagggaagtaacagg caagatgtcaaccatcattgttaacacaataagaaggaataatagatgtttacaattgaat agggtatgatacaaacatcaataagaatacagagatacaaaattacaacctaatgatttag taagatataatggtcaatcacaaaaatcaaaaggagattatcaattatgga	gtcaatcaccacccctaaataatagatttta ggatgtggccttgagcgtgaggtttgagag cgattgggtgattagagggttgaagaaatg aaaaatacacacaagaagttaggaagagaga aatacacacacactacaataacttacaagt tttagctctgtgtgtgtctttaaagagaga ggaaactctcagtgagcatggtttaaacc tttcaaacagcctcgtatgtgaacaacagat ctggcaagactggacaggatttgagaat
1964-1965	Ga0307377_10002004 GENOME_ACESSION: TR3MG_2_SF_33000316 73 GENOME_ID: 281015 CONTIG_ID:	atggcgacacagcaagaaggcgaaggttacttttaaaagaaggcaaggcaacagtaat caaccgtacacatttattatctgttacttattgctcaggagaaacaaagcaaccaataat attaggtgtagatgcaggtagcaagacaatttgattatcagctactatagaaaacgaagaa ctgttctcgcagaggttcagtaaggactgatattgtagaattattatctacaagaatggaa aacaggtgtaccagaaggcgtagattgagcagtaggggaaccaagatttttaaatcgaatt aagggtaaaaataagggttgcttgcaccgagtatagagaacaggattcagacacacct	gtcaattaccactacccctaaaggatagtg ctgttaggaaactacaacaaattgaattgacta gactaagcgaagactcgttatataagccat cataccctggatgataaacttagtctcgccg tctatgttttaacattaaacatctgtgtgtg aggggtcgtgtgttgagcatacaaaacttat

SEQ ID NO:	Name	IscB sequence	ωRNAs
	2003 SOURCE: JGI DATE: 2019-06-13.	aaaaatggtgatagatgatgtattctccaatatcaagattattgtgaacagcatc ctttgatatacagaagataaagaatccagttatcgaaggtactggaaccagcaaggagag caactagacttttgaatgttcgtgagtagttctgtttaggatagacaccaacgtaagggt aaaaaggatgcaaggtgtaaaattctcaatgtcatcattgagaccgtaaaactgggtg gagatgcaccaataaacttgattacctatgtgaggattgccaatgacgattatcacatggc aagttgaaatcaacatcaagcgtgtaattcatttaggatgttgccttcatgggtattatgc gttggcattttataacagactaaagagaataattccaatgtacaaaacacatttggatata taacgaaaaacacacgcataactaacaatttaccaaaagagcatagatagatgctttgtg tataacgagtaatacctttgtaataagggtgtagcattgtattgttaaaacaagtacggaa acacaatcgccagatcacaaagtcataatctcaggggcgagtaagaactcaatca agcaccatacctagtttaagggttttagattattgacaagggtgaattacaaaatcaagaat gctttatctttgtgtagaagagctacgggctatttcgacttgcgtaagttagacggtacgtta ttcacagaagtgcaagctataaagatttagaattattagaataaagaaaagctattataca gaaagggaggaagcaggagcaaggaatgagtgagcatctccacgaacctaaaagtat cgtgtttctgccaaagttttatga	ataacattgtctaaggattataactctgtaagg aggacgttaactgaga
1966-1967	Ga0182238_1001861 GENOME_ACESSION: hydfrates6_PW_90_2_\$F _3300017922 GENOME_ID: 37711 CONTIG_ID: 1860 SOURCE: JGI DATE: 2019-01-27.	atgcaacatcaaaagatgctttgttcggacacgaaccgtacccgtcatc cggccaggcgacggcgctgctaaagacaggccgtgcatcgggtttcgtaaaacacc cttcacatcatcattcatgacaggagcggaacccaggacacagaactgcgg atagatcccggttccaaggtaccgggtgtgtgttgggtgacggccacaacggggac cgggctgtggccgggacgagattgggacagaaagccagcaataaagcagaacctg gaaaaacggcggtgtgctgtagaagtagggcgaaccgcaaaactcggtatcgtgccc cacgctttaacaaccggcttcgcccagaactgttttcgcttagccttcaagccggg tggagcatgtgcggacctggcggaatcgttctgcaaatatctccggtatcgagcattgc cgtggagacggtgcttggatcacagaagatgcaaaaccggaaataatccgggggtg gagatcagcaaggggagcttttggctacgaggtccgggagctatttggagaagtgg ggacgaaaatcgccctatttgcggcaaggaataacaccttggaggtagagcacattgt gcccaggtccaggggcgaggtaaccgggttccaacctcaccctgtctgcacggac tgcaaccaagacaaaagggaacaggactgcggctgagtaggtttcccaagatccagg cccaggccaaagctccctgcgggacgctgcggcggttaaacgtaaccggtacgccat cgggaatatgctaaaaaggttgggctccgggtatcctttgtgttggaggtaggactaag aagaatcgatcggcgagggaattgagaagcccatgtagatagccgttgcgttagg gaagaaagggggaacggctatattccaaaggataattccttctgtggccaaagctatggg gcatgggtaggagcttttgcgggtggaaggcatatcctggaagtattgtctattttgc agcaagcagacgggtatatttcttcccgccctga	gtcatctaccccgccctaaaggcgagcgtt gtaagctctgaagatgaccgcttagtcac attaaaagtgtgactacgttgggagcgaata tataggcactccgggatgcccgctagtccc ggacactcgggcaaacggttaaacaggcat acggggttgaagccagtgccgtttgcaataa accgctctacaacttggcgaagcggagct aacggcctttgagccgagaaaaactaaggt aactcatgcaacatcaaga
1968-1969	human gut metagenome genome assembly, contig: NODE_3471_length_592 4_cov_5.665020, whole genome shotgun sequence GENOME_ACESSION: UAOQ01.1 GENOME_ID: 14977 CONTIG_ID: 3470 SOURCE: WGS DATE: 2019-01-27.	atgtcaacaacagtttctgattggacaagaagcaccggccctgatccgtgcccgcg gcaaggcgagacggcttctcggcgccggcgccgacgggtggtgaagcgctcccc ttcgtcatccggctcgtggaccgactgtggagcagtcgagggttcaaccgctgctc aagctcgaatccggctcaaggagacggcgcccgccgtcgtcgggagtaggggaa gaagcggcatcatcgctcgcgttctcgtcatcaagcatcggggcgccgcatccgcg acgcgctcaaggcgcggaagcgatttctgcgcccgaggctcgagaatttccgctgt cggctcggcgcttctgaaccgctgaagcccaagggtggtccccgcttctcgcg gcaccgctggaagacgaccttctgtcgtgaggaggtatgcccgtacgtccggtaa gcggcatcgcgagggttggtaggttcgattcgcaaaagctccagaaccggaggt ctcgggggtcaagtatcagcaggcgacgcttctttagttaggtgcgcgaataccttct gagaagctggccggaagtgcgtctatttgcggcgccgaaggcgtgcccgtcaacatcg accatgtgttccccgcgcaaggcgcgctccaaccgggtctccaactgtgtgctc	gtcaactaccccgccctaaaggccgagcgt tgtgaagcaggtcgtgttaccagcctaag ttttcagaaactcgttgcagacagggtcaa gaccacacctcggaatgcttctcagttccga gctctggaagcctcagaagcagacaagcat tgggtacgcacgaacgggtcagggcaaa atcgccgtctgttcaacttggcgagggga attgctccgcccgcaaggcgagcatgtc acgggcatcagcccgaagatcgtaaaggtt cattttcatgcaacaaca
1970-1971	TPA_asm: Richelia sp. UBA3308 UBA3308_contig_795, whole genome shotgun sequence GENOME_ACESSION: GCA_002361335.1_AS M236133v1_genomic GENOME_ID: 133139 CONTIG_ID: 369 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	atgcgtgtacctgttaattactcagacagcgccgttaattgcttaccacaaagcgaatcgc gctcgacgttgataaaaagaaggtataaggtctaaactaggcattttct acattcaatttaaacaccgctagtgggataaaagcccaagatattgctgttgaatcgac ccaggcaaaacatttaccgggacgcagtttattctacaaaagcactttactactattca cgttgatttaccctttaaataataactaaaagaatggctcaacgtgccatgtagcagctt gtagaagaggtagaagaattacaaaaataaccctttaaacacgtagtcatcgtcaaa ctagattcgataaccgtagacagaaaaaattaccccaagcataaacgcgaataaacaac tagaaagacgaatactaggattaataaagttatttccagtagtcaaaaggttgttacgaagt tgtaaaagcgtcggggaacaaagggttttagtccggtcaggttaggtcaaatagaccaatg caaaaaattagtaagcttgaattatttattgtttcaaaactcttcaaggttacgaaactatc aatccgcgaacacttaagcttgaaaaagagaaaagtgaagtctttaaataaactccta aacccatgctgtagatggaattgccttagctgcttactagtgatgaattacggaattgtga caataattcgatggatgagcaggagagataaccttaactgattcaataatttctggttattct agacctctatcagtcgtagacagttacttaattgttccaacaaaggaggtaaaagac gtaaatattggtgtagtggtaacagcagcagatcgtaaaaggtgataatgtcgaagctgt aaagctaataaaacttaccgtggatgggttaagtgagataactaataccaaagtttctgttag taatgcaactggaaacggcttggacagttctcaaaaataaagtgcgattatttcgacgtt ctactgttattatcatcatctcgaatatatctagtaaaaggcacttaa	gtcagtaactcccgctaaaagccgggagc ttgcaagaacaaatttgcgtgtcgtactag cccgtgagccatcaatcggtaaagacttccg gatgcttcccagctccggattatctctaaaac tgtttgtcagttgctttagaaaggacattt attgatgttggcgaggggacttaaaacttta cacgaaggattatcttattgctgt

373

SEQ ID NO:	Name	IscB sequence	ωRNAs
	5215 SOURCE: JGI DATE: 2019-01-27.	atcgtcctgaaggctggtgccaccgagttgcatcaccgtgtcgaaccacattgacat gggtaaatcgtttcagacgttggcgccgttgagaaaattctgtcagcgggttaattt gacatgcaattgatgcgtaaccctgaaatatcaggcgtcgaatatcaacaaggcagcta gcccgtatagttgagagcgtatttctgctgaaaaatggcagcgaatgtgctattgctg gtgctgaacacgttcccttgaggttgagcatatccattctaagccaaaggcggcgga atgcgggttttaacctggcgttggcttgcgatccttgcaacaagaaaaaggcacactgg atataaccgtttcctgaaaaataaacctgaacaacttaagaaaatcctgtctcaagcga cgatcgtaagcgtgacggcgacggttaatgcaacacgtaataagtgtttgctgaactgc tgaacaccggcttaccgttgaagccggaaccggcgcaaacccaattcaaccgaac ccggcgaggatgccgaaggcgcaattggatcgacgcgctgtgtggtgtaaatccgga gcaactgtttacgttgcgtatgatgaagccgctgttattaaagccaccggacacggaa cacggcaggtgtgctgaactgataaataatggcttccctcgaataaggccggacgagtc aacggatcatggcttcaaacaggtgactgtgtaagctcaagcaaccgaaaggcaaa tacgggtgtagcatattggacgcttagcgggcatcgcgctacggggcaattcgatattg ctgctacgggttgtaagataacagcaaaattcagcaattcacactcattcaagaggtgat ggctatgcttacgcacactga	tcgttaaccattaaacagcacagaccggaac aatgggcaggtgtgtaaacatgacaagccctt ataactttgccgaggtgacaattacgggtga aagccaggagaaagttttgactactaattc a
1980-1981	Ga0334826_004099 GENOME_ACESSION: 713E3X2metaG_FD_\$F_ 3300036886 GENOME_ID: 311349 CONTIG_ID: 4098 SOURCE: JGI DATE: 2020-04-07.	atgatgatgttagtgggtgtgtatataatggaagtgttgccttgttgcctgcaagg gcaagaagactcattgctccagaaaaggcaacacctttcatccataaagggtgtgtctgtat caggttgaaatcagtcaccttctgataataaaatacaacctatataatcgggttagacccc ggatcaaaagaaaggagataatgaactgaaagaaaactgtaattcaattcagtc tggatgctgctacattgggttaaggatgtgtgataatcaggaagaaatgctcgcagagcga gaagatttcgtaaaacccctgttaggaagccaatcatagataattatactgtgataattgg ttgcctcctagcacaagcgtcgggtggcaatccaaattgaaattgattcgttcttcatcag atatctctatgtccatgtgtgtagtcgaggatattgtgctgtcacaagaaaggacaacct aaatggaataaattcttcagtcgaataagattgtgtaaatcatgctttatacacaactcaa accgaaggttcatgctgtgttgaagaaaggattgaaacatccgagttgcgtaaacgac tcggactgcataagtcactgataagttatctgataaatgggatgctcattgtgtcgaatgt gggtactagctgcatctgcgatgcaaacactcccgttaattgacaagagcatgatattaac aaaccaatgcaatttcacgtagacaattgcatgcttccaaacccgcaaaaggcgggtta gaagactatattgatctactgatcaatgggttctgtagagatcaatcgtatcacatcct aaatcggaatcgccaccataggtggcaatccaaaataaattaaagctcacactataaa aacagggtaaaaggtaactcaaaatgtaacatcaagaatgttaattcaaatcattcaaca attacaacgttttctacccccgaagctaa	gtcaatcaccctgaactgaagattcagggt gcttgcgactgggaagattgctatctttgta aaagaagcgagataattgaaactgattgaac tagcctatgcgagttttgtccgcaaggaca aaatagcaaacgttaattatctatccgtgac agggtgaagacatctactgggtcctgtcagc tggtagctctgtgaggtcagtgccgaagat gtatatatacccgatgaatgggagtgccctta cgggcaaatfttagaagggaatgatgatgt a
1982-1983	Ga0310786_10002951 GENOME_ACESSION: AusTroGonzalMGv2_3_ \$F_3300031998 GENOME_ID: 279931 CONTIG_ID: 2950 SOURCE: JGI DATE: 2019-06-13.	atggctctatgactgaacagatacggtaaacctctgatgcctaccaccgggtacggcagg gtccgctgtctcaggaaaggactcgggctgttgcgattaccgtccattcacaatcc agcttactacgacacgccgaacggcgtgaggaaggtcagctctagcgtagacgcggg aaccaagcaggtgtgtctccgccattaccaagaagaagggtgctatcgaagcagaatt acttttaaggctcggatattcgtgataagatctccacgcgaaggaggttccgtcggaccag aaggaaacagggaagactcgtaccgcaagtcagggttccgtgaacaggggcccgtccaag aagccgggagtggttgcacctcgtacgaggaaggttgattctcacatctactggaattc gaagatacgcgaattcttccctatcaagaagataaccgttgaacaccgcgcagtttgatac ccagctgatgaaggctcaggaacaggggcttccgttccacagggcaccgactaccag aagggtgaacagctcggattctggaatgtcgtgaatacgtgctgttccgggacggccat aatgccagtgtcgaagggaagtcgaagacagtaaatgcatgtccaccatcga gagtcggaagacaggcggcgatgcgccgaacaacctatcacctgtgtcgaatgc catgagaagtaccatcttgagaaatcgaaactccgaagaccgcaaacggtgtacctc gctgctgacgcggcacagatgggtatcatgcggaagtccttattcatccggctgaagg aagaaattggaacgaaattccttctacgagacgtacggatacatcaccaagaatattcg caccacggcaggcctaccgaaggacatgtgaatgatgccactgcatcagcggcaat cccgggtgatgttctgatggcaagtacttgatcatccgcaagttgctgcaacaacagg caactgcaccgtgcaaccattatgccgggcggaaggcggagaaacaaccagtcacccc gtgaagtaaggggcttccggctgatgattccgtcaggtatgcatacaggagtgctttct gagcggacgtaggatgtcaggctgttttccgttgcggataccggaaaagtctggtgc agattcgtcagttacaagaagctgacattaagacatcacacaacacttacattatggag gaagcagcgctcctctcaccactaaagatgggtga	atcaactaccacggggttaaagaccggtga gtttgggttaactcaggctcatagttgattag actcgtgatgcaagtagaaatcgtctga actacgttaggaaggtcagcggcaccctggg atgtacgggtcaagtcccccgctgccgag cgtcattaaacagtcctgatgggtagggaca gtgtggcgttcagacaagccctccataact gtcgtgaccacccacaggagtaatccct gcattacattaaagtaggtttaaacaataaa ggataacgactagatg
1984-1985	Ga0393278_0000470 GENOME_ACESSION: LZ_38metaG_FD_\$F_33 00036985 GENOME_ID: 312543 CONTIG_ID: DATE: 2020-04-07.	atgcagaagttagaaaaaagaatacatcacaccacaatacttcttagtttgggca actgtgatttttattaaacagagaccaaaatctcagtgtaaaaaattcaaaaacttttcta aatctcgaagagactaatcccaatagagggaaacaggggcttgagagttctgtgtt gtctatgtacttaatatgaggggagaacctctaatgccaaacccctagaatgcaaga aatttctaaaaagaagaaagcaaggtgtttaaagcatgtcctttacaattcaattgctcat tcaaacaggagaatcaaaacagaaatattcattgtggaatgattctgttataataagata gggtatgtatttttaacaagaaaaaagtttaataatgggtgaattaaaattagatcagggt atgaagaaaagactgtatgaagaaaaggtatcgcaaatataaaaagtaaatgtgg tatcgtaaacctagatttttaaacagagcaaaacaaaagggttggcttcacatctatcag agaaaattgataccacattcttattgaagattaaagaaaatacttctatcacaaga	gtcagataccatgtgtaaaacctgggcttt gtccgagagggttaaggcgaactagatgat tagagagcttggaaggagaaacttaccatg cagaagttagaaaaaagaatacatcacaca cccacaataacttcttagtttggcaactgt gattttttataaacagagacccaaatctcag gtaaaaaattcaaaacttttttaacaatctc gaagagactcaatcccaacatagaggga cagggttgagag

SEQ ID NO:	Name	IscB sequence	ωRNAs
		gtggttgaggagtaggacagtttgatattcagaaaatagataatcctgatatccgagcaat agaatatcaacaagggaatctataatacatacaaatgtaaaagggtatattctttcaagaga aaatgggacatgccaaactgtgtataaaagaaagacggtaagtttcaattacatcataaact ccaagaagcaaaagggtggaacagacagccctagcaatttggcttacttcataaatctggc atgataaattacatagggaacatttagaacattcttaagtaatagtaggaacataagacc agttcgtttatgaatattatttgcaaaaatttaaagatttaatttcgaattacataggat attacatttgaatagaataatttaaattgaaaaatgcacatgcaaatgatcttttggat tgctgtgggaacaactcaagagttcactaggccatatctaatttcaaaaaagaaaaata atcgtatgttgcacaaactaagaaggttttaagccagcatcaggagaagtcattatatt atcaacctcatgatctgttatgttttaaggagagatttgaagttgttgcacgcatagtta tgaaaaagcgatattgttataaaataatttagggcaaaaatttggcgttaggattaaaaaa tagtttcttatttcacacgaatggtttaatatggagatagagtaa	
1986-1987	Ga0207870_100010 GENOME_ACESSION: CeladaP60BSPAdes_SF_3300027319 GENOME_ID: 24472 CONTIG_ID: 9 SOURCE: JGI DATE: 2019-01-27.	atgcagcagagagtggttcgtaggaaggacagaagtcctcctatgccgtgcagtc cgcccgagcccgaaatgcttctgtagagcgctcgccggcgatcctgcggcggtatcctt tcacgatcatgttgaggatagagaagcggtgaggttcacgccggcggtggaagtgc gatcccggttcggagaccggcgatggccttgggtgctgagtttggcagcgggggc caaccggctcggcgagcggagatccaaccggggcaaacaggctccggaaggca ctcatgaaaaggcggttctccggcgagcgaggaatcgaagacgcgtacccgg aagccggcggttctgaaccggatccgaaaaatgcagcggatcggaagaacgcta agcagggcagccgtactgtcgcgcctcgccggcgaggagacggcgatgggttccgg gacaaacgcttcccccttcttgaagagccgggtcgaagacgctgacttgggtgaac cgtctcgcgatatgcggcggttcttataacgggtgaacacgtcaaatgcacacgc agctgctccagaacccgatactcggcggtggaataccaacgaggcaccctttcgggt atgagttgcgggagtagctcgtgagaagtttggccgccagtgtgctactgtgggga gcatcggggagacgaagtgtgaacatcgaccacgtggtcccgaggtcgcgaggcggc agtgcaggggtgagcaacctggcggtgtgtcgcggacgtgaacgaggcgaaaggga aacgggatgccggaagaatgctggagaggtacatgcatcggtccggcgctgtatct cattcggcgagcgttttccgaagcgctgaaaaggctcaaacgcctctccgggatg cagcgcggtggaacacggtccgatgggtgatcgttgagcgttgaagagactcgggtt ccgttggagctgggatcgggcggtcagacgaagcgcaaccgttcgagccgaggttatc cgaagaacactggatcgacggcggttgcgtgggaatcaggggagtggtgtcggctt gacccggacatgcgaatttgcggatcgaggcaaggggcatggaacgcggcgccgg tgcgggacggacaagtacgggtttccatcggcagcgccgtgctgcgaagtctgacat gggtttcgaacgggggaccttgcgggctcctatccgcgaggggaagtatgcgggc acgcacgttggacgcaattgctatccggcaccggcgaggttcaggctcaacgggtcga cgtgcacccgaacacttgaacatttgcagcgagggtgacgggtatgcgtacaacatggg ataa	gtcaaccaccccatggctaaggccagggg cttgtgaagaacaagcccggttgaccag ccgaagtcggagggccattgccttgcggg gctacgttatccggtcatgacaccttgggt gctgcgagccagctcaagccctgtcgcgg gctgttaaacaggcatcgggtcgaagc cgggtccgttcgcatgacaagccgggataa catcggcgaggctcactgaccgtccaggg gctgaagcccgaggacgaggaggagcg taagtcctatgcagcagaga
1988-1989	Ga0224415_10011456 GENOME_ACESSION: 3300021399 GENOME_ID: 241587 CONTIG_ID: 10807 SOURCE: JGI Dark Harvest DATE: 2019-01-27.	atgccgacctttaattcaggaagtgcgcaggtccttaagagcgggcgggcgatatac ttctgccatgagccctttacgatccgggtctgtacccgaaaagcttgacacacagccag tagaactgtgactgacggcgaggacggcacatcgccatattcgtaaagagcgaaaa gcacaggttctgacccagtcacgatccttaagggtatgagaagggaaaagcatgatg acagcgcatgtaccgaaggacgagacgaagccgcagacggcacagaagttccgc ttcgacaacaggaagaagcggaagggtggtcgtccgacgggtgagcacaagaag gacctgcacatcagcattacgagatgtacatcagggtctgccccgttaaccagcatcacg gtagaacggcatcctttgacacacagggccctggaagccatagaagaggactgccgct gccgaagggaagactacaaaagggtcctcggtagacgatgaacaccttgcggga cgccgtgtttaccgtgacggccacagatcgccgttgcgggaataatggcgtgctcag gggtacatcacatcggtactggaaggcgacgggttccaaccgatggggaaccttgc cggtctgcaggggatgcaatacgccgcaaacacaggaaggcggaaggctctacg gctgggagccggagatcaaaccccttaaggcgcgcccttataaacgcggtagagaaa gaggtcggcggaagaactgagagaagaacttcttccgtgacagagacaaacggc agcgccacaagaagcgccagaagcgcttctgcatagacaagaccacgcgaacgac gcttttgcattgggaaaatttaccggagggcagagcgggcaggagcaggtctgcgagaa gcgcaggagaataaccgatcctgtcaaaagtctatgacgccaggtatgtggacatccg cgacgaaaaggttaagtcggacagagctgtcttaccggagggtgaaaaggagcgga atccagaattctgaaaagaacatgcgtcgttccgagggcgcaaggtgaggaaaggc cgtgtcagtaaaaggagcgcgctatgccataaggcccggaaccatagtcgtatcga tggacagaacagtcgtcaagggtatccactgtggcggaacaaggcggttttagaaa acggaaagagcggtatccgtcaaaaaatcaggggactgaggttccccggcgcatgga cgcatatcgtaa	gtcaactaccacccgcttaggttctgatgag ccttgaagcgggggctgtgaaaagggg aattgtttcttttattgcaagccgggtgatt agccgtgcggatgcaggttcaaaaaggga aacccgatccgctgcgcggggaacccgc acaagtaccggcatatccgggaggg aactacgttccgggagaaggatgtaagtga ggcagctcttccatcaccgcacacgagg caccatgggatactccacacgtccatgca gctgcggcgagcgttaaacatccctgagg gtaagggggaagtgcgtcgggttgaaaa cctccatgaacatcgccgaagtggatcac ccactaaaaggagataccgcagcgacaca aactgtaccacgtacagtaaaaggctcgg aaggcgtaaggcatacaagaacattaac actatcgctcggggcatggccggaagca cgaatcggctatggcccgagag
1990-1991	Nocardiosis sp. JB363 genome assembly, contig: Scaffold2, whole genome shotgun sequence GENOME_ACESSION:	atggctactgtccgacaggcgagaagaccacaggccgtgcttctcagcgccgtg ctctggaatcgagtcggtagacacgccccggatggcgacgaacgggacttgcgca ccacccaggggtgatccgggtcggaacctgtgcgagggagaccacctcatctca cctgatgtcgggtgcacacccggccttggctgggagagggcggtgagcgcc acccatccgtcttgcgtgcagaagaaccagattccgtccagccctgtcaccggcca	gtaaaggatctcgggtcgtgaaggaccggg ctttacgcccccttgggtgagagccgtt accagcaccagcatcgctacggggagg tgacctgatggctacgttccgacagggc agaagaccacagggcgttctcctcagc

SEQ ID NO:	Name	IscB sequence	ωRNAs
	GCA_900143625.1_AS M90014362v1_genomic GENOME_ID: 187582 CONTIG_ID: 161 SOURCE: NCBI_Prokaryotes DATE: 2019-01-27.	ggggccgcgaactgttgaacaagggtcggggcgtggtggccggcacacccctttcac catccgggtcaaggaccgcaccatcgcgagtcgaggtggacggtgtccaggtcggg atcgacccgggcagcaaacacaccggtatcgcgggtgttaccagcatgccggggaac gacgcgccgggtacgcatcagctcgaccaccgtggtgcgcgaatccgaagaaaat cggccaacgctccgctataggcgggggtcgtcgttcgggaacctcgccaccgcga ccccggttcgccaaccgacccgcccccaagggtgctcgcgcgctccctgcgacacc gggtggagaacaccaccggtgggtggaccgggtcaccgatgggaccggtcaaag cgggtgcacgtggaacgggtgccttcgacaccacgcccgtccgcccgaagccct agaggggcgtggagtataacggggcacccctgcacggctatgaggtacgcatgtacgt ctggccaagttcgaccgtcgtgtgtgtactgcggggccacgaacacaccactcaact ggaccacgtccaccccgctccgaggcgggtccgacggtgtccaacctggtgttg gcgtgcgtgacgtcaaccaggccaagaacaaccggcccgtcgaaggttcatacca acaagcaggtcttggctcggatttggctcgggccaaggcacccttgcgagacgggg ggcggtgcagtcacccgatggcgctctggcgctctgaacgaacggttgcctactc atgtcgggtcgggtggcgaacggtgaaccggaccgcaaccacctaccgaaac ccacactcgtgacgcccctatctgtgggaagatcgaaccatcaccagaccaccaacc gggtcctggtcggcggatgtgcggcgggggtcctacgcccgaacccgccccgaca agcacggattccgagggtcgcctgccccgaaccaagggtttctcgggttctcacc ggcgatctgtccggccgtgggtccccagggaagaagaccggcacctataccggc cgggtcgcgggtacgcccgtggcgagcttcaacatcaccaccaccaaggcaccgtcc aaagcatccgttaacagacatgctgtactccaacgagcggacggctatggctacact tgaaggaggagggcgtttctcccggtcctga	ggcctgctctggaatcggagtcggttagaca cgccccgatcgggcacgaacgggactt cggcaccaccagggtggtatccggtgc ggacatgtgcgaggggagaccacctca tctacatgatgtcgggtgcgtcaccacccg
1992-1993	Ga0118657_10064171 GENOME_ACESSION: IMG_3300009506_SF_3 300009506 GENOME_ID: 312509 CONTIG_ID: 64170 SOURCE: JGI DATE: 2020-04-07.	gtggcaatctttatttagatgacacagaagaaggagttaatgtgaaagtattgtgcgaac aggcagggacaaccgtttagtccactacaccccgcaaggcgcgtttgttctgaaggc gggcaaggccgagattgtcggcggggaacccttaccatccggttctgtacggctcaa cgggatacaggcaggacgtaacgctgggcatgtacgggtacaagatgataggctat agcggcgtcaccggccagcggaggtgtatcggcggtgaggtgtgtatcgaagggg atgtccgagcggctcaagggaacgggcgatgtaccgcccgggcagacggaaccggaa gcgggtatcggcagccacgatttgaacacggcaaggcggcgtccgggtggtctgcggc tcattcagcacaactggatagcaccggcggtgtatagcgccgattgagtcgttttgc cgtgacaagaacgatcagcaggtggcggaactttgacattcaggccatcagggaaccc agtattgagggcaaggcgtatcagcaagggaacaggcgggcttctggaactgaggg agtacattctgcatcgggacaagcacaatgccaaaatccgggtgcggaaacaaggcc aaggataagggtgctggaagtgcacacatccgggtactggaagaagaccggcggaca ggccggcgaacctgagcacgctttgtacgaagtgccacaccccgagggaacctcaacc cggcggtttctgtacggctggcaacctgcactcaaatctgtcagaccggaacgtttatg agtaccgtccgggtggcgtggttggggcggggaacaagccatacctacggctatc gcaccaagtcggggcgcatgtcgtggggttggaggtacacaccacaacgacgcttt gtgattcggggcggcgagagcagggaacgaactgaagcaattgactttgaacatga ggcgaatgaccgcagcttgcagaagttctatgacgcaaatatcagacattcgcacg ggcggggaagtgaaggggcgaactctatcaggccgaaggacgcgcaaccgcaa cctgaacagcgagaacctacggcgataccgaggtccgaagaacctggccgtctcaa atcagacgcaggcgggtatccctatcagtggaatgacctgattgtaggaggcgagaa gtaccacagcaagggtacgaagaactacggcgatacgtgtgtgacgggttgagaa aggcggtcagggtcgtcgtggtgacccgggttggtggcgaagggttgcacggc aatctga	gtcaactaccatgctttgaaagtgtggcctt gcggtagcgaagcccggttgaattagcctc agccaccaccggcaagggtgacgggcta cgttaaggatgaaacataggtacttcggga tactctacacgtcccgaacactacgggtcag gtctaacatcgtgagggtagcggaagt gcttgaactgtatacaacatcttaacattggc gatgtagaccacccggaacgggtgcaat ctttatttagattgcacagaaaggagttaa tgtgaaa
1994-1995	Ga0120401_1023611 GENOME_ACESSION: IMG_3300009431_SF_3 300009431 GENOME_ID: 305927 CONTIG_ID: 23610 SOURCE: JGI DATE: 2019-11-15.	atgagagttccagtagtatctaaagatgataagccactatgccaaactaagccagctaagg cacgtaaaatgatagaaggtagagtgctaaagaagttgttgcgaagactggcgtattct acatacaaatgttaatactgtggcgcaaaaagttcaggatgtgttcgctcgtcaccgac gacttctgttacggcactggcaatagacctggcgagtaaatcagtgatgatccgtgt ctggtgagaaaagatgtggcactcaaacgatggcaataatgccacagaaggtacaaaag aaagtaactgagcggcgtcaattaagacgagtagacgttatcgtaatcagacgcag gaaggcacgggttgacaacaggaacgcaaaagctgggtggatagccggtcacaattg gctaaagtgcagttccgcaataaggtatgataagatttagccaagatttccccctcaacta tatcgtgtgtgaggatgttaggttcaaccattacaagaagcagtggtgtaagcatttctta cagtgagataggaaggcgtgtgtacgaggaattggaacgacatggcaagtgata aagtacgctggttggcaaacgcagaagcagcgaagtaattgggcaataagaatcaag cgtaaaagatgctctgataccttctccatgccaatgctcttgcgatgctcaatgaagt attcggggataatgtggataattca	tcagtcaatgacccctcctgaatcatagattc agaaggccttgagatgcaggactctaagg gtaactcttgactagaggcgttgaattcct aagcaggcgcactgcgatgtacatcgtc agagtatttcgcaagctcgtacacctataagc tgcctcttgcggtagcggggaaggacc gacatcttgcaggttgcaatcgtatcgac ctttactttatgtccgaatatcgggactccga aaggagacagcaatgaga
1996-1997	human gut metagenome genome assembly, contig: NODE_564_length_3781 4_cov_32.665907, whole genome shotgun sequence	atgagagttattgtattcaatatcgttggcagaccattaatgccatgctcacaaagaaaagc cagattactcctaagggaataaggtctgatttataaatcattacccatttaccgattcagct gactttataactggagaacaagcaggagctgcatataggtatagacacagggttctaa atatataggagctgctcagatcagaggataaggtttttgaaaggcgaatcgagctt cgacaggatcaggtcaaatctgatacgaacgtatttatcgagaagcagacgaaat cgtaaaacaagataccgaaaaccaagggtttttaaactgtaagagaagatgaatgctt	gtcaataaccccgacctacagtaataactgtt gaggtcggagattgtaaaagctcatattgac tagcctaagttcttcgagaactcgtgtttat gttatcacactgcgaatgataccatgtttg cagcctgtgtaggtctgttaaaagtctgtg aggtaggaaacggtcaacctagatgttcgat

SEQ ID NO:	Name	IscB sequence	ωRNAs
	GENOME_ACESSION: OHAC01.1 GENOME_ID: 4852 CONTIG_ID: 563 SOURCE: WGS DATE: 2019-01-27.	ccacctagcttgcaagcaggataaatcatcgtttcattggaacacattgagtagtttggttccaaacccattcttcacatagaagtcggtgaagttgatgtagcaaatgataaatcctgaaatccatggagttgactatcaacatggtcagacatattggtttctcgtgaaaggtattttgtttttgcaagagataactacactgtcagtgctgtggaagtcacaaataagatttgaaacacacatcatcatctaccgagtaattggcggaacaaacagagttgataatctttacagttgacacagattgtcatatcgagaaacacaggaaggtggaattctatcaatggcagagcagcataaaaagftaaaacaatacaagaaccgccgttcattgaataccttacgtaagaggataattgtagcgtatccagatgctgagatcacatattgattcgaacaacacacaaacgtaagcgcatgaaattggataagacgcactataatgatgcgattgtcattagtgtatcaatgaaatcaatgaaatcaagaaatcttgaaagtggtactgataaaacaattccgcaaaaagaaacgctctttacacgaggctaccgccgtaagggaagaaaaagccaaacagaaaacagacgcgtaacagtaagaatcgccttattataaaggattttatctaacgataaagggttcagttttggaaagagtgatatattacagatttacgagcaatgcggcatatataaagatgagaacaatgactataattactctgccaacaaaacctataaacaggctcagcatcaataaaatgaagttggaattgtcataacaacaattggcagatcataataaaaaatgctgtgtaa	cacgacaagcatttacaacattggcggaagggtacaacattcaaaaagaaggagacagcacttgagagtagctgtcaaaaggtaaaaactatgaga
1998-1999	Ga0392328_0075420 GENOME_ACESSION: OC_3500m_C2_20cm_F D_SF_3300037245 GENOME_ID: 312580 CONTIG_ID: 75419 SOURCE: JGI DATE: 2020-04-07.	atgagtgatctgtctatgtgttgaacatgagggggaaacctctcatgccacacaccaccaaagcaagaaatattataaaatgaaagtaaaagtaagcaagattaccattattatccagttataaatatgttactggagaacaaaagcaacattacccttaggaatagactctggttatcaattgttggcttttcagtaatacagagaagaagaatataatcaggagaactagaaacttagaaaagataattcaagaatcttactcaaaaaagacaatcagggaaccagaagaatagattatggtaccgttaaccacgcttttaataatcgggggtctagaaaaaagctgggttgcaccaagtattagcaagaattaacctctcacaagaacctgattgaaaaacttgagcaaatcttccaataaccagaataattattgaagttgctacttttgatgccataagatgaagaaagccgggagtgacgggtgttgaatatcaacaggggtgatcttcaaggttatgagatcagagaataattgcttgaataatggaagagaaaatgtcctattgtacagaacagtgaaattggagatcgagcatattatcttaagtcaagaggaggaagtaatagtagtagataaccttaactattcatgcagaaaaatgcaatcttaaaaaaaggaaataa	gtcaatcaccactatcacctctcgggggtgagacggggctgttgagtgatctacaagagtaattggttgattagcctaagaggagatgaaatataaacttaagttatcagtaaaagtfaaagaacacacaaaggatgcttctcaagtcctttgctctgaagcggttgattaaacagagaggaaactctcagttgttcaccgcaagtgactggctgataaccttgccgaagagaaacttactttaaatttttaagtttaaggaggacagaacatgagtgtatct
2000-2001	Ga0207193_1000100 GENOME_ACESSION: IMG_3300020048_SF_3 300020048 GENOME_ID: 30012 CONTIG_ID: 99 SOURCE: JGI DATE: 2019-01-27.	atgggtatatgtattataaagatgggaaccactcatcctagcattagacatggtaaaatcagaagaatgttgaaagagaagaagctgtagtagttataaactactcctttcacagttaaaactttatcacactcctaaagacaatgtataacagaatctagttctaggaatagaccctgggaagtagaactctgtcaactgtagtttagagttgagaaaaactcctaaactagttctatgtttctgaagttaaagttagaactgatattactagtaagttgaaaaagagaagttcctacagaaaaactaggaagaaatcgtaaaactagatacagagaatcacgtttcttaaacaggaaggaagagaagaatggctaccacctagtggttttctaaacttaattctcacaagaagagataaaagttaattcttaaactttacctgtctctagagtaaccttagagagaataactttgatactcataagtttaagaatcctagatcagtggttaagatacaaacatggtacactctactcttatgagaatcacagacagttacgttttaactagagataagttactgtctgtagaagttgtaaaaagaagaatgtagtctcaacactcatcattacgttttaaaagtaagggagaaaagatcactactctaaccttagttaccctttgtgaacatgccacaagaaggttcacagttgaaagttaaaactacataagaacttctagtagtctaggaacaagttgaaacacattagatgctactcagctacataatcattcttaagaactagaagaatttctttaaacaacagaagagaagaattataaatttttaactacttttgggtatgagacagtagtaaaaagaagaattctaaagtttaagaaaactcactacaatgatgcaattttatctatctcttataaaaccttactcttaagactagagcagaatttatgtttagaagaatcgggtgagtaagggagactaccagcagactaaaggtgtaagaagtcaggttaagaatcctacttgcaagattcatgggttaggaagtttgacattgtacagtatcttggaaagtattacattataaaggtagaatgagttctgggtacgcaattcttatgaattcaagttaaaaccagtaaatgttgaaactaatggttaagagtaaacctactgtgaacttttaaacgtttgaaagaattagtttagaagaagtagtcttataatagaagttagaacacactag	gtgagctacgtcgggctaaagcactggcgcttcttaagagatgagtagttaccagactcagggttagagaacttaactacgttgaagaagttaccacacctaagaatgcccgctcagttcctagctctgtgtaggctctgtaagaagggaagcaaaacccggcaacctaaagttcagatgccaaggaacagactttttcaacattgtcagagaggaaagttggaagattcaatgggtatagctgaaattagtagacgttcatcgaagaatggcttatagcaaaccttaaaaggttaattatg

NUCLEIC ACID-GUIDED NUCLEASES

[0243] In one aspect, the present disclosure provides nucleic acid-guided nucleases. In an embodiment, the nucleic acid-guided nuclease is a CRISPR-associated IscB nucleic acid-guided nuclease. The nuclease may form a complex with one or more guide molecules. The complex may bind and target (e.g., cleaving, nicking, or otherwise modify) a target sequence

of a target polynucleotide. The nucleic acid-guided nucleases may generate a double-strand and/or single-strand break on a target polynucleotide. A target sequence may be a portion of, equal to, or overspanning a target polynucleotide.

[0244] In some examples, the target polynucleotide is DNA. In some example, the target polynucleotide is RNA. In some examples, the target polynucleotide is DNA-RNA hybrids or derivatives thereof. In an embodiment, the nucleic acid-guided nucleases and related compositions may specifically target double-strand DNA. In an embodiment, the nucleic acid-guided nucleases or nuclease/guide complexes may bind and cleave double-strand DNA. In an embodiment, the nucleic acid-guided nucleases or nuclease/guide complexes may bind to double-strand DNA without introducing a break to either or the strands. In an embodiment, the nucleic acid-guided nucleases or nuclease/guide complexes may open, disrupting the continuity of one of the two DNA strands, hereby introducing a nick of the double stranded DNA.

[0245] In one embodiment, the nucleic acid-guided nucleases (e.g., IscB) are CRISPR-associated proteins, e.g., the loci of the nucleases are associated with an CRISPR array. In one embodiment the IscBs may be referred to as Cas IscBs.

[0246] In one embodiment, the nucleic acid-guided nucleases, e.g. Cas IscBs, may have a small size. For example, the nucleic acid-guided nucleases may be no more than 50, no more than 100, no more than 150, no more than 200, no more than 250, no more than 300, no more than 350, no more than 400, no more than 450, no more than 500, no more than 550, no more than 600, no more than 650, no more than 700, no more than 750, no more than 800, no more than 850, no more than 900, no more than 950, or no more than 1000 amino acids in length.

[0247] The Cas IscB nucleic acid-guided nuclease may comprise one or more domains, e.g., one or more of a PLMP domain (e.g., at N-terminus), a RuvC domain, a Bridge Helix domain, and a Y domain (e.g., at C-terminus).

Table 2. Examples of nucleic acid-guided nucleases include Sequence Nos. 1-7 below.

No.	Proteins	Sequences
1 SEQ ID NO: 2041	IscB(-HNH) EFH81386	1 mstdatlirt tpshacadat dtlvatplmp prrvispwpg pgegqslmri pvvdirgmal 61 mpctpakarh llksgnarpk mklglfyvq lsycqepdnq slvagvdpgs kfeqlsvgt 121 kdtvlnlmve apdhvkgavq trrtmrrarr qrkwrpkrf hurlnmqri ppstrsrwea 181 karivahlrt ilpftdvve dvqavtrkgk ggtwnsfsp vqvqkehlyr llramglth 241 lregwqtkel reqhglkktk skskqsfesh avdswvlaas isgaehptct rlwymvpail 301 hrrqlhrlqa skggvrkpyg gtrslgvkrg tlvehkkygr ctvggvdrkr ntislheyrt 361 ntrltqaakv etcrvltwls wrswllrgkr tsskgkgshs s

2 SEQ ID NO: 2042	IscB(+HNNH) TAE54104.1	1 mqpakqqnwv fqingdkqpl dminpgrcre lqnrkglasf rrfpyvviqq qtienpqtk 61 yilkidpgsq wtgfaiqcg n dilfraelnh rgeaikfdlv krawfrgr r srnlryrkr 121 lnrakpegwl apsirhrvlt vetwikrfmr ycpiawieie qvrfdtqkla npeidgveyq 181 qgelqgyevr eyllqkwgrk caycgtenvp levehiqsks kggssrignl tlachvcnvk 241 kgnldvrdfi akspdilnqv lenstkplkd aaavnstrya ivkmaksice nvkcsggart 301 kmnrvrqgle kthslsdaacv gesgasirvl tdrpllitck ghgsrqsriv nasgfpavkn 361 aktvftiaa gdvvrftigk drkkaqagty tarvktptpk gfevlidgar islstmsnv 421 fvhrsdgygy el
3 SEQ ID NO: 2043	IscB(+HNNH) WP_0380936 40.1	1 mavfvidkhk rplmpcsekr arlllergra vhrqvpfvi rldrtvqhs avqplrvld 61 pgsratgmal vrekntvdtg tgevyeria lnlfelvhrg hrireqldqr nfrfrrrga 121 nlryraprfd nrrppgwla pslqhrvdt mawvrlcrw apasagiet vrfdtqrlq 181 peisgveyqq galagcevre yllekwgrkc aycaenvpl eiehivpksr ggsdrvsnl 241 lacracnqak gnrdrvafila dqpelarlil aqakaplka aavnatrwal yralvdtgpl 301 veagtgggrk wnrtlrplk thaldalcvg qvdqvrhwr pvlgrcagr gsyrrtrlr 361 hgfpgryltr nksafgftg dlaravvtkg kkagtylgri airasgsfni qtpmgvvqgi 421 hhrfctllqr adgygyfvqp kpteaalssp rikagvssag n
4 SEQ ID NO: 2044	IscB(+HNNH) WP_0524903 48.1	1 mtnvfvfvid tnqkplqps aavarkllr gkaamfryp aviilkkev d svgkpkie 61 idpgskytgf alvdsdnad fiiwgtelch rgaackelt krsairrrr nrktryrkr 121 ferrkpegwl apslqhrvdt tlwvkrick fvpimsisve qvkdldkle nsdiqgieyq 181 qgtlagtylr callehgrk caycdvenvf leichyips kgsdskfsl tlachkcnin 241 kgnksidefl lsdhkrleqi klhqkktldk aaavnatrkk lvtlqekt f lnvlvsdgas 301 tkmtrlsstl akrhwidagc vnttlivilk tlqplqvkn ghgnkqfvtm daygfprksy 361 epkkvrkdwk agdiirvttk dgtmlmrvk kaakklyip fggeasfss enakaihrsd 421 gyrysfaaid sellqkmat
5 SEQ ID NO: 2045	IscB(+HNNH) WP_0153258 18.1	1 mpnkyaflvd skgkldptk skkawylirk gkasveeyp liilkrevp kdqvnslkli 61 lgiddgtkv gfalvqkcq knkvlfkavm eqrqdvskkm eerrgyrrr rshkrypar 121 fdnrssskrk grippsilqk qkailrvnk lkkyiridki vledvsidir kltegrelyn 181 weyqesnrld enlrkatlyr ddctcqlcgt tetmlhahhi mprddgads iynlitcka 241 chkdvdne yqykdqlai idskelsdlk sashvmqgkt wlrdklksia qleitsggnt 301 ankrityeie kshsndaict tglpvdnid dikeyyikpl rkksakake lkcfqrldv 361 kytkmgety tgyitslrik nnkynskvcn fstlkgkifr gygfmltl nrpklmiv
6 SEQ ID NO: 2046	sp G3ECRI C AS9_STRTR	1 mlfnkciis inldfsnkek cmtkpysigl digtnsvgwa vitdnykps kkmkvlgnst 61 kkyikknlly vlfdsgita egrrlkrar rrytrrrri lylqefste matlddaffq 121 rlddsflvpd dkrdskypif gnlveekvyh defptiyhlr kyladstka dlrlylala 181 hmikyrglfl iegefnsknn diqknfdql dtynaifsd lslenksqle eivdkdiskl 241 ekkdrillkf pgeknsqifs eflklivngq adfrkcfnd ekashfske syedtlell 301 gyigddysdv flkakklyda illsgflvt dneteaplss amikrynekh edlallkeyi 361 nislktyne vfkddtkngy agyidgktnq edfyvylkn laefegadyf lekidredfl 421 rkqrtdngs ipyqihlqem raildkqakf yplaknker iekiltfrip yvvgplargn 481 sdfawsirkr nekitpwnfe dvidkessae afnrmstsd lylpeekvlp khslyetfn 541 vyneltkvrf iaesmrdyqf ldsqkqkdiv rlyfkdrkrv tdkdieylh aiyygdgiel 601 kgiekqfnss lstyhdlni indkeflds sneaieiei hltifedre mikrqlskfe 661 nifdksvllk lsrhytgwg klsaklingi rdeksntil dyldddgns rfnlqlhdd 721 alsfkkkiqk aqiigdedkg nkevvkslp gspaiikkil qsikivdelv kvmggrkpes 781 ivvemarenq ytnqgksnsq qrlrlekl kelgskilke nipaklskid nnalqndrly 841 lyylqngkdm ytgddldidr lsnydidhii pqaflkdnsi dnkvlyssas nrgksddfps 901 levvkrktf wyqlkskli sqrkfdntk aerggllped kagfiqrqlv etrqtkhva 961 rildekfnk kdenravrt vkiitlksl vsqfrkdfel ykvreindfh hahdaylnav 1021 iasallkyp klepefygd ypkynsfrer ksatekvyfy snimnifkks isladgrvie 1081 rplievneet gesvwnkesd latvrvlsy pqvrvvkkve eqnhglrdgk pkglfnanls 1141 skpkpnsnen lvgakeyldp kkygyagis nsfavlvkt iekgakkkit nvlefqgisi 1201 ldrinyrkd lnllekgyk dielielpk yslfelsdgs rmlasilst nnkrgeihkg 1261 nqiflsqkf klyhakris ntinenhrky venhkkfee lfyyilefne nyvgakkngk 1321 llnsafqswq nhsidelcss figptgserk glfeltsrgs aadfevlgvk ipryrdytps 1381 sllkdatlih qsvtglyetr idlaklgeg
7 SEQ ID NO: 2047	sp J7RUA5 C AS9_STAAU	1 mkmyilgld igitvgygi idyetrdiv agvrlfkean vennegrrsk rgarrlkrr 61 rhiqrvvkl lfdynltdh selsginpye arvkglsqkl seeefsaall hlakrrgvhn 121 vnevedtgn elstkeqisr nskaleekyv aelqlerlkk dgevrgsinr fktstvkkea 181 kqlkvqkay hqldqsfidt yidletrrt yyegpgegsf fgwkdkewy emlmghtyf 241 peelrsvkya ynadlynaln dlnnlvitrd enekleyyek fqienvfqk kkkptlkqia 301 keilvneedi kgvrvstgk pefnlkvvh dikditarke iienaellq iakiltiyqs 361 sediqeelt lnseleqei eqisnlkgyt gthnlskai nlldelwht ndnqiaifnr 421 lkvpkvdvl sqkqieptl vddfilspvv krsfiqsikv inaiikkygl pndiielar

		<p>481 eknskdaqkm inemqkmrq tnerieeir ttgkenakyl iekiklhdmq egkclyslea 541 ipledllnnp fnyevdhiip rsvsfdsfn nkvlvkqeen skkgmrtpfq ylsdsdskis 601 yetfkkhiin lakggrisk tkkeylleer dinrfsvqkd finrnldtr yatrglmnll 661 rsyfrvnld vkvksinggf tsflrrkwkf kknrnkykh haedaliian adfifkewkk 721 ldkakvmen qmfeekqaes mpeieteqey keifitphqi khikdfkdyk yshrvdkkpn 781 relindtlys trkddkgnti ivnnlnglyd kdndklkkl nkspekllmy hhdppqyqkl 841 klimeqygde knplykyee tgnyltkysk kdngpvikki kyygnknah lditddypns 901 mkvvlslk pyrfdvldn gvykftvkn ldvikenyy evnscyee kklkksinqa 961 efiasfynnd likingelyr vigvndlln rievnmidit yreylenmnd krppriikti 1021 asktqsikky stdilgnlye vkskhhppqii kkg</p>
8 SEQ ID NO: 2048	Streptococcus _pyogenes_S F370	<p>1 kysigldigt nsvgwavitd eykvpskkfk vlgntdrhsi kknligallf dsgetaeatr 61 lkrtarryt rknricylq eifsnemakv ddsffhrlee sflveedkhh erhpifgniv 121 devayhekyt tiyhlrkkiv dstdkadrl iylalahmik frghflieg lnpdsdvdk 181 lfiqlvtyn qlfeenpina sgvdakails arlsksrle nliqlpgek knlgfnlia 241 lslgltpnfk snfdlaedak lqlskdydd dldnllaqig dqyadflaa knlsdailis 301 dilrvnteit kaplsasmik rydehhdlt llkalvrqql pekykeiffd qskngyagvi 361 dggasqeefy kfikpilekm dgteellvkl nredllrkqr tfdngsiphq ihlgelhaul 421 rqedfypfl kdnrekieki ltfripyyv plargnsrfa wmrksceti tawnfeevvd 481 kgsaqsfie rmtndfknlp nekvlpkhsl lyeyftvne ltkvyvteg mrkpaflsge 541 qkkaivdlf ktnrkvtvkq lkedyfkkie cfdsvaisgv edrfnaslgt yhdlkiikd 601 kdfldneene diledivlt tlfdremie erlkyahlf ddkvmkqlkr rrytgwgrls 661 rklingirdk qsgktildf ksdgfanmf mqlihdsl fkediqaqv sgqgdsilheh 721 ianlagspai kkgilqtkv vdelvkmgr hkpeniviem arenqtkqg qknsrermkr 781 iecikelgs qilkehpven tqlqnekyl yylqngrdmy vdqeldiml sydvdhivp 841 qsfkddsid nkvltrsdn rgksdnpse evvkkmknyw rqlnaklit qrkfdnlta 901 erglseldk agfkrqlve trqtkhvaq ildsrmtky dendklirev kvitlksklv 961 sdfkrdfqy kvreinnyhh ahdaylnav gtalikkypk lesefygydy kvdyvrkmia 1021 kseqeigkat akyyfysnim nffkteitla ngeirkrpli etngetgeiv wdkgrdfatv 1081 rkvlsmppqn ivkktvqtg gfskesilpk msdkliark kdwdpkkygg fdsptvaysv 1141 lvvakvegk skklsvkel lgitimerss feknpidfle akgykevkd liiklpkysl 1201 felengrkm lasagelqkg nelalpskyv nflylashye klkgspeдне qkqlfveqhk 1261 hyldeieqi sefskrvlla danldkvlsa ynkhdrkpir eqaenihlf tltnlgapaa 1321 fkyfddidr krytstkevl datlihqsit glyetrids qlggd</p>

Table 2 (cont'd)

No.	Proteins	Domains and amino acid positions
1	IscB(-HNH) EFH81386	X domain: 51-97 RuvC-I: 104-118 Bridge Helix: 140-160 RuvC-II: 169-212 RuvC-III: 226-278
2	IscB(+HNH) TAE54104.1	X domain: 11-56 RuvC-I: 63-77 Bridge Helix: 100-121 RuvC-II: 129-172 HNH: 211-243 RuvC-III: 279-321
3	IscB(+HNH) WP_0380936 40.1	X domain: 4-50 RuvC-I: 57-71 Bridge Helix: 108-129 RuvC-II: 138-181 HNH: 220-252 RuvC-III: 288-330
4	IscB(+HNH) WP_0524903 48.1	X domain: 7-52 RuvC-I: 59-73 Bridge Helix: 100-121 RuvC-II: 129-172 HNH: 211-243 RuvC-III: 279-322

5	IscB(+HNN) WP_0153258 18.1	X domain: 7-52 RuvC-I: 61-75 Bridge Helix: 101-121 RuvC-II: 132-175 HNN: 215-247 RuvC-III: 284-327
6	sp G3ECR1 C AS9_STRTR	RuvC-I: 28-42 Bridge Helix: 85-108 Rec: 118-736 RuvC-II: 750-799 HNN: 864-896 RuvC-III: 957-1019 PAM Interaction (PI): 1119-1409
7	sp J7RUA5 C AS9_STAAU	RuvC-I: 7-21 Bridge Helix: 49-72 Rec: 80-433 RuvC-II: 445-493 HNN: 553-585 RuvC-III: 654-709 PAM Interaction (PI): 789-1053
8	Streptococcus _pyogenes_S F370	RuvC-I: 4-18 Bridge Helix: 61-84 Rec: 94-718 RuvC-II: 725-774 HNN: 833-865 RuvC-III: 926-988 PAM Interaction (PI): 1099-1365

Table 3: Some other examples of the nucleic acid-guided nucleases are provided.

ID	Sequences
0181581_1004 2345_organize d_organized_- _ >_IscB(0.prot SEQ ID NO: 2049	MSRVLVVDADRRVLAPCTARRARLLSGGKAAVLRYPFTIILKQSYPTASPRPVRLKLDPGSKTT GIAVVTEATGEVVWAAELQHRGQLIKDALESRRSLRSSRRNRKTRYRPPRWRNRKRTGPPVLSSA GEVNQLGKWLAPSLQHRIVMTWVHRLRRYLPITAIHQEIVRFDMQKMQNPEISGVEYQQGTFLG FEVREYLLDKWHRRCGYCGAQNTRLEVDHIVPRSHGGSDRVSNLTLSCEPCNKKKSNRPAALFLA KKPEVLQKLQRQAKAPLKDAAVNSTRYALLERLKATGLPVEIASGGRTKFNRFERQIPKTHWLD AACVGASTPEVLQWEAVRPLAIKAMGHGKRQVTNVDAYGFPKGKPKGIPVHPFRTGDVIRAEVP VGKFAGNYVDRIVAIRTDQTRVSLPLRSQEKGKKKVPFLFQTKYITAKLFSADGYDYGFLQPPEPR TQRTES*
0181587_1005 2842_organize d_organized_- _ >_IscB(0.prot SEQ ID NO: 2050	MSRVLVVDADRRVLAPCTARRARLLSGGKAAVLRYPFTIILKQSYPTASPRPVRLKLDPGSKTT GIAVVTEATGEVVWAAELQHRGQLIKDALESRRSLRSSRRNRKTRYRPPRWRNRKRTGPPVLSSA GEVNQLGKWLAPSLQHRIVMTWVHRLRRYLPITAIHQEIVRFDMQKMQNPEISGVEYQQGTFLG FEVREYLLDKWHRRCGYCGAQNTRLEVDHIVPRSHGGSDRVSNLTLSCEPCNKKKSNRPAALFLA KKPEVLQKLQRQAKAPLKDAAVNSTRYALLERLKATGLPVEIASGGRTKFNRFERQIPKTHWLD AACVGASTPEVLQWEAVRPLAIKAMGHGKRQVTNVDAYGFPKGKPKGIPVHPFRTGDVIRAEVP VGKFAGNYVDRIVAIRTDQTRVSLPLRSQEKGKKKVPFLFQTKYITAKLFSADGYDYGFLQPPEPR TQRTES*
0211577_1000 7484_organize d_organized_-	MSKVFVVDKERRPLAPCTPRRARLLSECKASVLRQYPFTIILKESHATATPRPLRLKIYPASKTTGL AVINESTAEVVWAAELKHRGHLIKDALESRRSLRRGRRSRKTRYRPARWLNVRKPPVFTNTEGV VVTGKWLPPSLQHRIVVMTWVERLQHYLQITAIHQEIVMRFDQKLQNPESGVEYQQGTFLHGYE VREYLLKWSRKCAYCGARDTRLEISHLIARSRGSDQVSNLTACKACRDQKGDSNLEKFLATK

<p>-</p> <p>>_IscB(0.prot</p> <p>SEQ ID</p> <p>NO: 2051</p>	<p>PKILKKLQSQARVSLKDVAAINSTRLLALLERLKATGLPVEVSSGGETKYNRNQQQIPKSHWLDAV</p> <p>CVGASTPENLEWQQVKPLAIKAMGHGKRQMVNVDAFGFPRGKPKGIPVHPFRTGDIVRVITIPKGGK</p> <p>YAGEYEERISSIKTSETRVGIPNKKKEGTIYLQTKYITAKIFSSDGYDYDYL*</p>
<p>0211577_1005</p> <p>5383_organize</p> <p>d_organized_-</p> <p>-</p> <p>>_IscB(0.prot</p> <p>SEQ ID</p> <p>NO: 2052</p>	<p>MSQVFVVDKERRPLAPCTPRRARLLLSECKASVFRRYPFTIILKESHATATPRPLRLKIYPASKTTGL</p> <p>AVINESTAEVWAAELKHRSQLIKKALESRRSLRSGRRSRKTRYRPARWLNVRNPNVFTNTEGA</p> <p>VITGKWLPSSLQHRVEVVMWVERLQRYLPITALSQEIMRFDTQKLQNPEISGVEYQQGTLHGYY</p> <p>VREYLLEKWSRKCA YCGARDTRLEINHIVARSRGSDRVSNLTLACRSCREQRGASNLEEFLATRP</p> <p>ALLMKLQSQAQVSLRDVAAINSTRFVLLERLKARGLPVEVSSGGETKFNRNQQQIPRSHWLNAV</p> <p>IGPNTPENLKWQVQPLAIKAMGHGKRQMVNVDAFGFPRGKPKGTPVHPFRTGDVVRAAIPKGGK</p> <p>YVGEYEERISSIKTSETRVGIPNKKGGGTIYLQTKYITTKIFSSDGFDEFLTSES*</p>
<p>0211577_1006</p> <p>4440_organize</p> <p>d_organized_-</p> <p>-</p> <p>>_IscB(0.prot</p> <p>SEQ ID</p> <p>NO: 2053</p>	<p>MSQVFVVDKERRPLAPCTPRRARLLLSECKASVFRRYPFTIILKESHATATPRPLRLKIYPASKTTGL</p> <p>AVINESTAEVWAAELKHRSQLIKKALESRRSLRSGRRSRKTRYRPARWLNVRNPNVFTNTEGA</p> <p>VITGKWLPSSLQHRVEVVMWVERLQRYLPITALSQEIMRFDTQKLQNPEISGVEYQQGTLHGYY</p> <p>VREYLLEKWSRKCA YCGARDTRLEINHIVARSRGSDRVSNLTLACRSCREQRGASNLEEFLATRP</p> <p>ALLMKLQSQAQVSLRDVAAINSTRFVLLERLKARGLPVEVSSGGETKFNRNQQQIPRSHWLNAV</p> <p>IGPNTPENLKWQVQPLAIKAMGHGKRQMVNVDAFGFPRGKPKGTPVHPFRTGDVVRAAIPKGGK</p> <p>YVGEYEERISSIKTSETRVGIPNKKGGGTIYLQTKYITTKIFSSDGFDEFLTSES*</p>
<p>a0099850_100</p> <p>2913_organize</p> <p>d_organized_-</p> <p>-</p> <p>>_cas9(209.pro</p> <p>t</p> <p>SEQ ID</p> <p>NO: 2054</p>	<p>MSRVLVVDANRCPLAPCTPRRARLLLNSGKAAVLRYPFTIILKQSYPTANPRPVRLKLDPGSKTT</p> <p>GIAVVTEATGEVWAAELQHRGQLIKNALESRRSLRRGRRNRKTRYRPARWLNKRRTGPPVLSSA</p> <p>DTVSTLGKWLAPSLQHRIEVIMTWVHRLRRYLPITAISQEI VRFD MQM QNPEISGVEYQQGTLFG</p> <p>YEVREYLLDKWQRQCGYCGAKDKRLEVDHIVPRSHGGSDRVSNLTLSCPCNKRNQRPAAVFL</p> <p>AKKPEVLQKLQRQAKAPLKDAAAVNSTRYALLERLKATGLPVEIASGGRTKFNRSERQIPKTHWL</p> <p>DAACVGASTPEVLQWKAVKPLAIKAMGHGKRQVVNVDA YGFPRGKAKGIPVHPFRTGDIVRAEI</p> <p>PKGKYVGTYVSRIAETTTSKPLAGFKSKTGKRIQCHTKHMTKLFNSDGYGYGFLKAPEPRQT VISE</p> <p>S*</p>
<p>a0208542_100</p> <p>2724_organize</p> <p>d_organized_-</p> <p>-</p> <p>>_IscB(0.prot</p> <p>SEQ ID</p> <p>NO: 2055</p>	<p>MSRVLVVDADRCPLAPCTPRRARLLLNSGKAAVLRYPFTIILKQSYPTASPRPVRLKLDPGSKTT</p> <p>GIAVVTEATGEVWAAELQHRGQLIKNALESRRSLRRGRRNRKTRYRPARWLNKRRTGPPVLSSA</p> <p>DTVSTLGKWLAPSLQHRIEVIMTWVHRLRRYLPITAISQEI VRFD MQM QNPEISGVEYQQGTLFG</p> <p>YEVREYLLDKWRRQCGYCGAKDKRLEVDHIVPRSHGGSDRVSNLTLSCPCNKRNQRPAAVFL</p> <p>AKKPEVLQKLQRQAKAPLKDAAAVNSTRYALLERLKATGLPVEVASGGRTKFNRSERQIPKTHW</p> <p>LDAACVGASTPEVLQWEAVRPLAIKAMGHGKRQVVNVDA YGFPRGKAKGIPVHPFRTGDIVRAEI</p> <p>PKGKYVGTYVSRIAETTTSKPLAGFKSKTGKRIQCHTKHMTKLFNSDGYGYGFLKAPEPRQT VISE</p> <p>S*</p>
<p>AGA010000</p> <p>46.1_organized</p> <p>_organized_-</p> <p>>_IscB(26.prot</p>	<p>MSKVFVVDKERRPLAPCTPRRARLLLSECKASVLRQYPFTIILKESHATATPRPLRLKIYPASKTTGL</p> <p>AVINESTAEVWAAELKHGHLIKKALESRRSLRRGRRSRKTRYRPARWLNVRKPPVLNTEGV</p> <p>VVTGKWLPSSLQHRIKVVMTWVERLQHYLQITALSQEV MRFD TQKLQNPEISGVWYQQGTLHGYY</p> <p>EVREYLLEKWSRKCA YCGARDTRLEISHLIARSRGSDQVSNLTLACKACRDQKGDSNLEKFLAT</p> <p>KPKILKKLQSQARVSLKDVAAINSTRLLALLERLKATGLPVEVSSGGETKYNRNQQQIPKSHWLDA</p>

In an embodiment, the nucleic acid-guided nucleases and related compositions do not have a PAM requirement, in that there is no sequence requirement outside of the target sequence which defines target specificity. In some cases, the target specificity of the nucleic acid-guided nucleases and related compositions may be determined by the sequence of the guide molecule only, not any sequence within the target polynucleotide. In alternative embodiments, the nucleic acid-guided nucleases and related compositions has a target specificity, more particularly the binding of the nucleic acid-guided nucleases-guide complex is PAM-dependent. The nucleic acid-guided nucleases and related systems may be modified to include PAM specificity (as described in Kleinstiver et al. 2015; Hirano et al. Mol. Cell 2016).

[0252] In an embodiment, the nucleic acid-guided nucleases correspond to a naturally occurring protein, a modified naturally occurring protein, functional fragment or truncated version thereof, or a non-naturally occurring protein. In an embodiment, the nucleic acid-guided nucleases comprises one or more domains originating from other nucleic acid-guided nucleases, more particularly originating from different organisms. In an embodiment, the nucleic acid-guided nucleases may be designed by *in silico* approaches. Examples of *in silico* protein design have been described in the art and are therefore known to a skilled person.

[0253] In embodiments, the nucleic acid-guided nucleases also encompasses a homologs or an orthologs of nucleic acid-guided nucleases whose sequences are specifically described herein. The terms “ortholog” and “homolog” are well known in the art. By means of further guidance, a “homolog” of a protein as used herein is a protein of the same species which performs the same or a similar function as the protein it is a homolog of. Homologous proteins may but need not be structurally related, or are only partially structurally related. An “ortholog” of a protein as used herein is a protein of a different species which performs the same or a similar function as the protein it is an orthologue of. Orthologous proteins may but need not be structurally related, or are only partially structurally related. In one embodiment, the homolog or ortholog of a nucleic acid-guided nucleases such as referred to herein has a sequence homology or identity of at least 80%, at least 85%, at least 90%, at least 95% with a nucleic acid-guided nuclease. In further embodiments, the homolog or ortholog of a nucleic acid-guided nuclease has a sequence identity of at least 80%, at least 85%, at least 90%, or at least 95% with a wildtype nucleic acid-guided nuclease.

[0254] Further orthologs of known nucleic acid-guided nuclease may be identified. Some methods of identifying orthologs of nucleic acid-guided nucleases may involve identifying

tracr sequences in genomes of interest. Identification of tracr sequences may relate to the following steps: Search for the direct repeats or tracr mate sequences in a database to identify a region comprising a nucleic acid-guided nuclease. Search for homologous sequences in the region flanking the nucleic acid-guided nuclease in both the sense and antisense directions. Look for transcriptional terminators and secondary structures. Identify any sequence that is not a direct repeat or a tracr mate sequence but has more than 50% identity to the direct repeat or tracr mate sequence as a potential tracr sequence. Take the potential tracr sequence and analyze for transcriptional terminator sequences associated therewith.

[0255] A chimeric enzyme can comprise a first fragment and a second fragment, and the fragments can be of nucleic acid-guided nuclease orthologs of organisms of a genus or of species, e.g., the fragments are from nucleic acid-guided nuclease orthologs of different species.

Domains

[0256] In some examples, the nucleic-acid guided nuclease, e.g. Cas IscB, comprises an N-terminal X domain, a RuvC domain (e.g., including a RuvC-I, RuvC-II, and RuvC-III subdomains), a Bridge Helix domain, and a C-terminal Y domain. In some examples, the nucleic-acid guided nuclease comprises In some examples, the nucleic-acid guided nuclease comprises an N-terminal X domain, a RuvC domain (e.g., including a RuvC-I, RuvC-II, and RuvC-III subdomains), a Bridge Helix domain, an HNH domain, and a C-terminal Y domain.

X domain

[0257] The Cas IscB nucleic-acid guided nuclease comprises an X domain, e.g., at its N-terminal.

[0258] In an embodiment, the X domain include the X domains in Table 2. Examples of the X domains also include any polypeptides a structural similarity and/or sequence similarity to a X domain described in the art. In some examples, the X domain may have an amino acid sequence that share at least 50%, at least 55%, at least 60%, at least 5%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with X domains in Table 2.

[0259] In some examples, the X domain may be no more than 10, no more than 20, no more than 30, no more than 40, no more than 50, no more than 60, no more than 70, no more than 80, no more than 90, or no more than 100 amino acids in length. For example, the X domain may be no more than 50 amino acids in length, such as comprising 2 3, 4, 5, 6, 7, 8, 9,

10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acids in length.

Y domain

[0260] The Cas IscB nucleic-acid guided nuclease comprises an Y domain, e.g., at its C-terminal.

[0261] In an embodiment, the X domain include Y domains in Table 2. Examples of the Y domain also include any polypeptides a structural similarity and/or sequence similarity to a Y domain described in the art. In some examples, the Y domain may have an amino acid sequence that share at least 50%, at least 55%, at least 60%, at least 5%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with Y domains in Table 2.

RuvC domain

[0262] In one embodiment, the nucleic acid-guided nuclease comprises at least one nuclease domain. In an embodiment, the nucleic acid-guided nuclease protein comprises at least two nuclease domains. In an embodiment, the one or more nuclease domains are only active upon presence of a cofactor. In an embodiment, the cofactor is Magnesium (Mg). In embodiments where more than one nuclease domain is present and the substrate is a double-strand polynucleotide, the nuclease domains each cleave a different strand of the double-strand polynucleotide. In an embodiment, the nuclease domain is a RuvC domain.

[0263] The nucleic-acid guided nuclease comprises a RuvC domain. The RuvC domain may comprise multiple subdomains, e.g., RuvC-I, RuvC-II and RuvC-III. The subdomains may be separated by interval sequences on the amino acid sequence of the protein.

[0264] In an embodiment, Examples of the RuvC domain include those in Table 2. Examples of the RuvC domain also include any polypeptides a structural similarity and/or sequence similarity to a RuvC domain described in the art. For example, the RuvC domain may share a structural similarity and/or sequence similarity to a RuvC of Cas9. In some examples, the RuvC domain may have an amino acid sequence that share at least 50%, at least 55%, at least 60%, at least 5%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with RuvC domains in Table 2.

[0265] In some examples, the RuvC domain comprise RuvC-I polypeptide, RuvC-II polypeptide, and RuvC-III polypeptide. Examples of the RuvC-I domain also include any polypeptides a structural similarity and/or sequence similarity to a RuvC-I domain described

in the art. For example, the RuvC-I domain may share a structural similarity and/or sequence similarity to a RuvC-I of Cas9. In some examples, the RuvC domain may have an amino acid sequence that share at least 50%, at least 55%, at least 60%, at least 5%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with RuvC-I domain in Table 3. The RuvC-II domain also include any polypeptides a structural similarity and/or sequence similarity to a RuvC-II domain described in the art. For example, the RuvC-II domain may share a structural similarity and/or sequence similarity to a RuvC-II of Cas9. In some examples, the RuvC domain may have an amino acid sequence that share at least 50%, at least 55%, at least 60%, at least 5%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with RuvC-II domains in Table 2. The RuvC-III domain also include any polypeptides a structural similarity and/or sequence similarity to a RuvC-III domain described in the art. For example, the RuvC-III domains may share a structural similarity and/or sequence similarity to a RuvC-III of Cas9. In some examples, the RuvC domain may have an amino acid sequence that share at least 50%, at least 55%, at least 60%, at least 5%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with RuvC-III domains in Table 2.

[0266] For example, and as described in the art (e.g. Crystal structure of Cas9 in complex with guide RNA and target DNA, Nishimasu et al. Cell, 2014) the RuvC domain of Cas9 consists of a six-stranded mixed β -sheet (β 1, β 2, β 5, β 11, β 14 and β 17) flanked by α -helices (α 33, α 34 and α 39– α 45) and two additional two-stranded antiparallel β -sheets (β 3/ β 4 and β 15/ β 16). It has been described that the RuvC domain of Cas9 shares structural similarity with the retroviral integrase superfamily members characterized by an RNase H fold, such as *Escherichia coli* RuvC (PDB code 1HJR, 14% identity, root-mean-square deviation (rmsd) of 3.6 Å for 126 equivalent C α atoms) and *Thermus thermophilus* RuvC (PDB code 4LD0, 12% identity, rmsd of 3.4 Å for 131 equivalent C α atoms). RuvC nucleases have four catalytic residues (e.g., Asp7, Glu70, His143 and Asp146 in *T. thermophilus* RuvC), and cleave Holliday junctions through a two-metal mechanism. Asp10 (Ala), Glu762, His983 and Asp986 of the Cas9 RuvC domain are located at positions similar to those of the catalytic residues of *T. thermophilus* RuvC. There are key structural discrepancies between the Cas9 RuvC domain and the RuvC nucleases, which explain their functional differences. Unlike the Cas9 RuvC domain, the RuvC nucleases form dimers and recognize Holliday junctions. In addition to the

conserved RNase H fold, the Cas9 RuvC domain has other structural elements involved in interactions with the guide:target heteroduplex (an end-capping loop between $\alpha 42$ and $\alpha 43$) and the PI domain/stem loop 3 (β -hairpin formed by $\beta 3$ and $\beta 4$).

Bridge helix

[0267] The nucleic-acid guided nuclease comprises a bridge helix (BH) domain. The bridge helix domain refers to a helix and arginine rich polypeptide. The bridge helix domain may be located next to anyone of the amino acid domains in the nucleic-acid guided nuclease. In one embodiment, the bridge helix domain is next to a RuvC domain, e.g., next to RuvC-I, RuvC-II, or RuvC-III subdomain. In one example, the bridge helix domain is between a RuvC-1 and RuvC2 subdomains.

[0268] The bridge helix domain may be from 10 to 100, from 20 to 60, from 30 to 50, e.g., 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46 or 47, 48, 49, or 50 amino acids in length. Examples of bridge helix includes the polypeptide of amino acids 60-93 of the sequence of *S. pyogenes* Cas9.

[0269] In an embodiment, examples of the BH domain include those in Table 2. Examples of the BH domain also include any polypeptides a structural similarity and/or sequence similarity to a BH domain described in the art. For example, the BH domain may share a structural similarity and/or sequence similarity to a BH domain of Cas9. In some examples, the BH domain may have an amino acid sequence that share at least 50%, at least 55%, at least 60%, at least 5%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with BH domains in Table 2.

HNH domain

[0270] The nucleic-acid guided nuclease comprises a HNH domain. In an embodiment, at least one nuclease domain shares a substantial structural similarity or sequence similarity to a HNH domain described in the art.

[0271] In some examples, the nucleic acid-guided nuclease comprises a HNH domain and a RuvC domain. In the cases where the RuvC domain comprises RuvC-I, RuvC-II, and RuvC-III domain, the HNH domain may be located between the Ruv C II and RuvC III subdomains of the RuvC domain.

[0272] In an embodiment, examples of the HNH domain include those in Table 2. Examples of the HNH domain also include any polypeptides a structural similarity and/or sequence similarity to a HNH domain described in the art. For example, the HNH domain may

share a structural similarity and/or sequence similarity to a HNH domain of Cas9. In some examples, the HNH domain may have an amino acid sequence that share at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with HNH domains in Table 2.

[0273] For example, the HNH domain of Cas9 as described in the art (e.g. Crystal structure of Cas9 in complex with guide RNA and target DNA, Nishimasu et al. Cell, 2014) comprises a two-stranded antiparallel β -sheet ($\beta 12$ and $\beta 13$) flanked by four α -helices ($\alpha 35$ – $\alpha 38$). It shares structural similarity with the HNH endonucleases characterized by a $\beta\beta\alpha$ -metal fold, such as phage T4 endonuclease VII (Endo VII) (PDB code 2QNC, 20% identity, rmsd of 2.7 Å for 61 equivalent Ca atoms) and *Vibrio vulnificus* nuclease (PDB code 1OUP, 8% identity, rmsd of 2.7 Å for 77 equivalent Ca atoms). HNH nucleases have three catalytic residues (e.g., Asp40, His41, and Asn62 in Endo VII), and cleave nucleic acid substrates through a single-metal mechanism. In the structure of the Endo VII N62D mutant in complex with a Holliday junction, a Mg²⁺ ion is coordinated by Asp40, Asp62, and the oxygen atoms of the scissile phosphate group of the substrate, while His41 acts as a general base to activate a water molecule for catalysis. Asp839, His840, and Asn863 of the Cas9 HNH domain correspond to Asp40, His41, and Asn62 of Endo VII, respectively, consistent with the observation that His840 is critical for the cleavage of the complementary DNA strand. The N863A mutant functions as a nickase, indicating that Asn863 participates in catalysis. The Cas9 HNH domain may cleave the complementary strand of the target DNA through a single-metal mechanism, as observed for other HNH superfamily nucleases. Although the Cas9 HNH domain shares a $\beta\beta\alpha$ -metal fold with other HNH endonucleases, their overall structures are distinct, consistent with the differences in their substrate specificities.

[0274] In an embodiment, the nucleic-acid guided nuclease comprises at least a HNH or RuvC nuclease domain. In an embodiment, the nucleic-acid guided nuclease comprises at least one reduced or minimal HNH or RuvC nuclease domain. In one embodiment, the nucleic-acid guided nuclease comprises two nuclease domains. In an embodiment, the two nuclease domains are a HNH and a RuvC domain. In an embodiment, the nucleic-acid guided nuclease comprises at least one nuclease domain substantially similar to a HNH or RuvC domain by sequence similarity. In an embodiment, the nucleic-acid guided nuclease comprises at least one nuclease domain substantially similar to a HNH or RuvC domain by structural similarity.

[0275] In one embodiment, the nucleic acid-guided nucleases are in part characterizable by the nature of the guide molecule that ensures formation of the nucleic acid-guided nuclease complex and binding to the target sequence. The guide molecule envisaged for use with a nucleic acid-guided nucleases capable of specifically hybridizing to a target sequence, directing binding of the complex formed by said nucleic acid-guided nucleases and guide sequence to said target sequence. In an embodiment, the target sequence is a coding sequence. In an embodiment, the target sequence is a noncoding sequence. By means of example, noncoding sequences include noncoding functional RNA, cis-and trans-regulatory elements, introns, pseudogenes, repeat sequences, transposons, viral elements, and telomeres. Examples of noncoding functional RNA include ribosomal RNA, transfer RNA, piwi-interacting RNA and microRNA. In an embodiment, the target sequence may be a regulatory DNA sequence. Non-limiting examples of regulatory DNA sequences are transcription factors, operators, enhancers, silencers, promoters, and insulators.

[0276] In one embodiment, where the nucleic acid-guided nucleases is a reduced version of a nucleic acid-guided nuclease, the guide molecule envisaged for use can be the guide RNA which is known to function with the corresponding full length nucleic acid-guided nucleases. Features of the guide molecules are detailed herein below.

[0277] In one embodiment, the compositions and systems are characterized by elements that promote the formation of a complex at the site of a target sequence (also referred to as a protospacer in the context of an endogenous system). In the context of formation of a complex, “target sequence” refers to a sequence to which a guide sequence is designed to target, e.g., have complementarity, where hybridization between a target sequence and a guide sequence promotes the formation of a complex. The section of the guide sequence through which complementarity to the target sequence is important for cleavage activity is referred to herein as the seed sequence. A target sequence may comprise any polynucleotide, such as DNA or RNA polynucleotides and is comprised within a target locus of interest. In one embodiment, a target sequence is located in the nucleus or cytoplasm of a cell.

PAM specificity

[0278] In one embodiment, the nucleic acid-guided nucleases and related compositions do not contain a PAM specificity. In certain examples, the nucleic acid-guided nucleases lack or substantially lack a PAM interacting (PI) domain. In an embodiment, the nucleic acid-guided nucleases may have a PI domain or a functional fragment of a PI domain. In an embodiment,

the nucleic acid-guided nucleases may achieve a target specificity by a non-protein domain. In an embodiment, the nucleic acid-guided nucleases may have helicase activity. In an embodiment, the nucleic acid-guided nucleases may have reduced helicase activity compared to Cas proteins known in the art. In an embodiment, the nucleic acid-guided nucleases may comprise additional components that contribute in mediating target recognition. In an embodiment, targeting specificity is obtained by a central hairpin structure in a guide molecule.

[0279] Examples of PAM sequences for the nucleic acid-guided nucleases herein include NGG and NAC. For example, the nucleic acid-guided nucleases may recognize PAM sequence NAC.

[0280] The PAM interaction domain or PI domain as referred to herein is reported to be responsible for determining PAM specificity of nucleic acid-guided nucleases (e.g., IscB proteins). By means of example, the PI domain is contained in the NUC lobe and forms an elongated structure comprising seven α -helices, a three-stranded antiparallel β -sheet, a five-stranded antiparallel β -sheet, and a two-stranded antiparallel β -sheet.

[0281] In some cases, where the nucleic acid-guided nucleases do have a PAM requirement, the precise sequence and length requirements for the PAM will differ depending on the nucleic acid-guided nucleases used. In some examples, PAMs are typically 2-5 base pair sequences adjacent the protospacer (that is, the target sequence). Examples of the natural PAM sequences for different nucleic acid-guided nucleases orthologs have been identified and the skilled person will be able to identify further PAM sequences for use with a given nucleic acid-guided nucleases.

[0282] Further, associating a PAM Interacting (PI) domain (e.g., attaching or fusing) to a nucleic acid-guided nuclease may allow programming of PAM specificity, improve target site recognition fidelity, and increase the versatility of the IscB, genome engineering platform. nucleic acid-guided nucleases may be engineered to alter their PAM specificity, for example as described in Kleinstiver BP et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature*. 2015 Jul 23;523(7561):481-5. doi: 10.1038/nature14592. The skilled person will understand that other IscB proteins may be modified analogously.

[0283] The crystal structure information (described in U.S. Provisional Patent Application Nos. 61/915,251 filed December 12, 2013, 61/930,214 filed on January 22, 2014, 61/980,012 filed April 15, 2014; and Nishimasu et al, "Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA," *Cell* 156(5):935–949, DOI: dx.doi.org/10.1016/j.cell.2014.02.001

(2014), each and all of which are incorporated herein by reference) provides structural information to truncate and create modular or multi-part CRISPR enzymes which may be incorporated into inducible composition. In particular, structural information is provided for *S. pyogenes* Cas9 (SpCas9), and this may be extrapolated to other Cas9 orthologs or IscB proteins (as well as homologs and orthologs thereof) or other nucleic acid-guided nucleases. In one embodiment, the conformational variations in the crystal structures of the CRISPR-Cas9 system or of components of the CRISPR-Cas9 provide important and critical information about the flexibility or movement of protein structure regions relative to nucleotide (RNA or DNA) structure regions that may be important for the function of other nucleic acid-guided nucleases and related systems. The structural information provided for Cas9 (e.g. *S. pyogenes* Cas9) as the nucleic acid-guided nuclease in the present application may be used to further engineer and optimize the other nucleic acid-guided nucleases and related system and this may be extrapolated to interrogate structure-function relationships in other nucleic acid-guided nucleases and related systems.

Protein modifications

[0284] The nucleic acid-guided nucleases may comprise one or more modifications. As used herein, the term “modified” with regard to a nucleic acid-guided nuclease generally refers to a nucleic acid-guided nuclease having one or more modifications or mutations (including point mutations, truncations, insertions, deletions, chimeras, fusion proteins, etc.) compared to the wild type counterpart from which it is derived. By derived is meant that the derived enzyme is largely based, in the sense of having a high degree of sequence homology with, a wildtype enzyme, but that it has been mutated (modified) in some way as known in the art or as described herein.

[0285] The modified proteins, e.g., modified nucleic acid-guided nuclease may be catalytically inactive (also referred as dead). As used herein, a catalytically inactive or dead nuclease may have reduced or no nuclease activity compared to a wildtype counterpart nuclease. In some cases, a catalytically inactive or dead nuclease may have nickase activity. In some cases, a catalytically inactive or dead nuclease may not have nickase. Such a catalytically inactive or dead nuclease may not make either double-strand or single-strand break on a target polynucleotide, but may still bind or otherwise form complex with the target polynucleotide.

[0286] In one embodiment, the modifications of the nucleic acid-guided nuclease may or may not cause an altered functionality. By means of example, modifications which do not result

in an altered functionality include for instance codon optimization for expression into a particular host, or providing the nuclease with a particular marker (e.g. for visualization). Modifications which may result in altered functionality may also include mutations, including point mutations, insertions, deletions, truncations (including split nucleases), etc., as well as chimeric nucleases (e.g. comprising domains from different orthologues or homologues) or fusion proteins. Fusion proteins may without limitation include, for instance, fusions with heterologous domains or functional domains (e.g. localization signals, catalytic domains, etc.). In an embodiment, various different modifications may be combined (e.g. a mutated nuclease which is catalytically inactive and which further is fused to a functional domain, such as for instance to induce DNA methylation or another nucleic acid modification, such as including without limitation, a break (e.g. by a different nuclease (domain)), a mutation, a deletion, an insertion, a replacement, a ligation, a digestion, a break or a recombination). As used herein, “altered functionality” includes without limitation an altered specificity (e.g. altered target recognition, increased (e.g. “enhanced” nucleic acid-guided nuclease) or decreased specificity, or altered PAM recognition), altered activity (e.g. increased or decreased catalytic activity, including catalytically inactive nucleases or nickases), and/or altered stability (e.g. fusions with destabilization domains). Examples of all these modifications are known in the art. It will be understood that a “modified” nuclease as referred to herein, and in particular a “modified” nucleic acid-guided nuclease or system or complex preferably still has the capacity to interact with or bind to the polynucleic acid (e.g. in complex with the guide molecule). Such modified nucleic acid-guided nuclease can be combined with the deaminase protein or active domain thereof as described herein.

[0287] In one embodiment, an unmodified nucleic acid-guided nucleases may have cleavage activity. In one embodiment, the nucleic acid-guided nucleases may direct cleavage of one or both nucleic acid (DNA or RNA) strands at the location of or near a target sequence, such as within the target sequence and/or within the complement of the target sequence or at sequences associated with the target sequence. In one embodiment, the nucleic acid-guided nucleases may direct cleavage of one or both DNA or RNA strands within about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 50, 100, 200, 500, or more base pairs or nucleotides from the first or last nucleotide of a target sequence. In one embodiment, the cleavage may be staggered, i.e. generating sticky ends. In one embodiment, the cleavage is a staggered cut with a 5' overhang.

In one embodiment, the cleavage is a staggered cut with a 5' overhang of 1 to 5 nucleotides, preferably of 4 or 5 nucleotides.

[0288] In one embodiment, the cleavage site is distant from the PAM, e.g., the cleavage occurs after the 18th nucleotide on the non-target strand and after the 23rd nucleotide on the targeted strand. In one embodiment, the cleavage site occurs after the 18th nucleotide (counted from the PAM) on the non-target strand and after the 23rd nucleotide (counted from the PAM) on the targeted strand. In one embodiment, a vector encodes a nucleic acid-targeting effector protein that may be mutated with respect to a corresponding wild-type enzyme such that the mutated nucleic acid-targeting effector protein lacks the ability to cleave one or both DNA and RNA strands of a target polynucleotide containing a target sequence. As a further example, two or more catalytic domains of a nucleic acid-guided nuclease (e.g. RuvC I, RuvC II, and RuvC III or the HNH domain) may be mutated to produce a mutated nucleic acid-guided nuclease substantially lacking all DNA cleavage activity. As described herein, corresponding catalytic domains of a nucleic acid-guided nuclease may also be mutated to produce a mutated nucleic acid-guided nuclease lacking all DNA cleavage activity or having substantially reduced DNA cleavage activity. In one embodiment, a nucleic acid-guided nuclease may be considered to substantially lack all polynucleotide cleavage activity when the polynucleotide cleavage activity of the mutated enzyme is no more than 25%, no more than 10%, no more than 5%, no more than 1%, no more than 0.1%, no more than 0.01% of the nucleic acid cleavage activity of the non-mutated form of the enzyme; an example can be when the nucleic acid cleavage activity of the mutated form is nil or negligible as compared with the non-mutated form. A nucleic acid-guided nuclease may be identified with reference to the general class of enzymes that share homology to the biggest nuclease with multiple nuclease domains from the Type I, II, III, IV, V, or VI CRISPR systems.

[0289] In an embodiment, the nuclease domains of the nucleic acid-guided nuclease are catalytically inactive, or modified to be catalytically inactive, or when the protein is a nickase. In an embodiment, both nuclease domains are catalytically inactive.

[0290] In an embodiment, the nucleic acid-guided nuclease may comprise one or more modifications resulting in enhanced activity and/or specificity, such as including mutating residues that stabilize the targeted or non-targeted strand (e.g. eCas9; “Rationally engineered Cas9 nucleases with improved specificity”, Slaymaker et al. (2016), *Science*, 351(6268):84-88, incorporated herewith in its entirety by reference). In an embodiment, the altered or

modified activity of the engineered nucleic acid-guided nuclease comprises increased targeting efficiency or decreased off-target binding. In an embodiment, the altered activity of the engineered nucleic acid-guided nuclease comprises modified cleavage activity. In an embodiment, the altered activity comprises increased cleavage activity as to the target polynucleotide loci. In an embodiment, the altered activity comprises decreased cleavage activity as to the target polynucleotide loci. In an embodiment, the altered activity comprises decreased cleavage activity as to off-target polynucleotide loci. In an embodiment, the altered or modified activity of the modified nuclease comprises altered helicase kinetics. In an embodiment, the modified nuclease comprises a modification that alters association of the protein with the nucleic acid molecule comprising RNA, or a strand of the target polynucleotide loci, or a strand of off-target polynucleotide loci. In an aspect of the invention, the engineered nucleic acid-guided nuclease comprises a modification that alters formation of the nucleic acid-guided nuclease and related complex. In an embodiment, the altered activity comprises increased cleavage activity as to off-target polynucleotide loci. Accordingly, in an embodiment, there is increased specificity for target polynucleotide loci as compared to off-target polynucleotide loci. In other embodiments, there is reduced specificity for target polynucleotide loci as compared to off-target polynucleotide loci. In an embodiment, the mutations result in decreased off-target effects (e.g. cleavage or binding properties, activity, or kinetics), such as in case for nucleic acid-guided nuclease for instance resulting in a lower tolerance for mismatches between target and guide RNA. Other mutations may lead to increased off-target effects (e.g. cleavage or binding properties, activity, or kinetics). Other mutations may lead to increased or decreased on-target effects (e.g. cleavage or binding properties, activity, or kinetics). In an embodiment, the mutations result in altered (e.g. increased or decreased) helicase activity, association or formation of the functional nuclease complex. In an embodiment, the mutations result in an altered PAM recognition, i.e. a different PAM may be (in addition or in the alternative) be recognized, compared to the unmodified nucleic acid-guided nuclease. Examples mutations include positively charged residues and/or (evolutionary) conserved residues, such as conserved positively charged residues, in order to enhance specificity. In an embodiment, such residues may be mutated to uncharged residues, such as alanine.

GUIDE SEQUENCES

[0291] The systems herein may further comprise one or more CRISPR-associated guide molecules. A CRISPR-associated guide molecule may form a complex with a nucleic acid-guided nuclease, and direct the complex to bind with a target sequence. In some examples, the CRISPR-associated guide molecule may comprise a first and second nucleic acid molecules, the first and second nucleic acid molecules capable of forming a duplex, the duplex capable of forming a complex with the nucleic acid-guided nuclease, wherein the second nucleic acid molecule is a recombinant molecule comprising a heterologous CRISPR-associated guide sequence capable of directing site-specific binding of the complex to a target sequence of a target polynucleotide. In some examples, the single CRISPR-associated guide molecule capable of forming a complex with the nucleic acid-guided nuclease and directing site-specific binding of the complex to a target sequence of a target polynucleotide.

[0292] As used herein, a heterologous CRISPR-associated guide molecule is a CRISPR-associated guide molecule that is not derived from the same species as the nucleic acid-guided nuclease. For example, a heterologous CRISPR-associated guide molecule of a nucleic acid-guided nuclease derived from species A is a polynucleotide derived from a species different from species A, or an artificial polynucleotide.

[0293] As used herein, the term “CRISPR-associated guide sequence” or “CRISPR-associated guide molecules” has the meaning as used herein elsewhere and comprises any polynucleotide sequence having sufficient complementarity with a target nucleic acid sequence to hybridize with the target nucleic acid sequence and direct sequence-specific binding of a nucleic acid-targeting complex to the target nucleic acid sequence. In one embodiment, the degree of complementarity of the CRISPR-associated guide sequence to a given target sequence, when optimally aligned using a suitable alignment algorithm, is about or more than 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99%, or more. In certain example embodiments, the CRISPR-associated guide molecule comprises a CRISPR-associated guide sequence that may be designed to have at least one mismatch with the target sequence, such that a RNA duplex formed between the CRISPR-associated guide sequence and the target sequence. Accordingly, the degree of complementarity is less than 99%. For instance, where the CRISPR-associated guide sequence consists of 24 nucleotides, the degree of complementarity is more particularly about 96% or less. In one embodiment, the CRISPR-associated guide sequence is designed to have a stretch of two or more adjacent mismatching

nucleotides, such that the degree of complementarity over the entire CRISPR-associated guide sequence is further reduced. For instance, where the CRISPR-associated guide sequence consists of 24 nucleotides, the degree of complementarity is more particularly about 96% or less, more particularly, about 92% or less, more particularly about 88% or less, more particularly about 84% or less, more particularly about 80% or less, more particularly about 76% or less, more particularly about 72% or less, depending on whether the stretch of two or more mismatching nucleotides encompasses 2, 3, 4, 5, 6 or 7 nucleotides, etc. In one embodiment, aside from the stretch of one or more mismatching nucleotides, the degree of complementarity, when optimally aligned using a suitable alignment algorithm, is about or more than about 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99%, or more. Optimal alignment may be determined with the use of any suitable algorithm for aligning sequences, non-limiting example of which include the Smith-Waterman algorithm, the Needleman-Wunsch algorithm, algorithms based on the Burrows-Wheeler Transform (e.g., the Burrows Wheeler Aligner), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies; available at www.novocraft.com), ELAND (Illumina, San Diego, CA), SOAP (available at soap.genomics.org.cn), and Maq (available at maq.sourceforge.net). The ability of a CRISPR-associated guide sequence (within a nucleic acid-targeting guide RNA) to direct sequence-specific binding of a nucleic acid -targeting complex to a target nucleic acid sequence may be assessed by any suitable assay. For example, the components of a nucleic acid-guided nuclease-guide system sufficient to form a nucleic acid-targeting complex, including the CRISPR-associated guide sequence to be tested, may be provided to a host cell having the corresponding target nucleic acid sequence, such as by transfection with vectors encoding the components of the nucleic acid-targeting complex, followed by an assessment of preferential targeting (e.g., cleavage) within the target nucleic acid sequence, such as by Surveyor assay as described herein. Similarly, cleavage of a target nucleic acid sequence (or a sequence in the vicinity thereof) may be evaluated in a test tube by providing the target nucleic acid sequence, components of a nucleic acid-targeting complex, including the CRISPR-associated guide sequence to be tested and a control guide sequence different from the test CRISPR-associated guide sequence, and comparing binding or rate of cleavage at or in the vicinity of the target sequence between the test CRISPR-associated and control guide sequence reactions. Other assays are possible, and will occur to those skilled in the art. A CRISPR-associated guide

sequence, and hence a nucleic acid-targeting guide RNA may be selected to target any target nucleic acid sequence.

[0294] A CRISPR-associated guide sequence, and hence a nucleic acid-targeting guide, may be selected to target any target nucleic acid sequence. The target sequence may be DNA. The target sequence may be any RNA sequence. In one embodiment, the target sequence may be a sequence within a RNA molecule selected from the group consisting of messenger RNA (mRNA), pre-mRNA, ribosomal RNA (rRNA), transfer RNA (tRNA), micro-RNA (miRNA), small interfering RNA (siRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), double stranded RNA (dsRNA), non-coding RNA (ncRNA), long non-coding RNA (lncRNA), and small cytoplasmatic RNA (scRNA). In some preferred embodiments, the target sequence may be a sequence within a RNA molecule selected from the group consisting of mRNA, pre-mRNA, and rRNA. In some preferred embodiments, the target sequence may be a sequence within a RNA molecule selected from the group consisting of ncRNA, and lncRNA. In some more preferred embodiments, the target sequence may be a sequence within an mRNA molecule or a pre-mRNA molecule.

[0295] In an embodiment, the CRISPR-associated guide sequence or spacer length of the CRISPR-associated guide molecules is from 15 to 50 nt. In an embodiment, the spacer length of the CRISPR-associated guide RNA is at least 15 nucleotides. In an embodiment, the spacer length is from 15 to 17 nt, e.g., 15, 16, or 17 nt, from 17 to 20 nt, e.g., 17, 18, 19, or 20 nt, from 20 to 24 nt, e.g., 20, 21, 22, 23, or 24 nt, from 23 to 25 nt, e.g., 23, 24, or 25 nt, from 24 to 27 nt, e.g., 24, 25, 26, or 27 nt, from 27 to 30 nt, e.g., 27, 28, 29, or 30 nt, from 30 to 35 nt, e.g., 30, 31, 32, 33, 34, or 35 nt, or 35 nt or longer. In certain example embodiment, the CRISPR-associated guide sequence is 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nt.

[0296] In one embodiment, the sequence of the CRISPR-associated guide molecule (direct repeat and/or spacer) is selected to reduce the degree secondary structure within the guide molecule. In one embodiment, about or less than about 75%, 50%, 40%, 30%, 25%, 20%, 15%, 10%, 5%, 1%, or fewer of the nucleotides of the nucleic acid-targeting guide RNA participate in self-complementary base pairing when optimally folded. Optimal folding may be determined

by any suitable polynucleotide folding algorithm. Some programs are based on calculating the minimal Gibbs free energy. An example of one such algorithm is mFold, as described by Zuker and Stiegler (Nucleic Acids Res. 9 (1981), 133-148). Another example of a folding algorithm is the online webserver RNAfold, developed at Institute for Theoretical Chemistry at the University of Vienna, using the centroid structure prediction algorithm (see e.g., A.R. Gruber et al., 2008, Cell 106(1): 23-24; and PA Carr and GM Church, 2009, Nature Biotechnology 27(12): 1151-62). Further algorithms may be found in U.S. application Serial No. TBA (attorney docket 44790.11.2022; Broad Reference BI-2013/004A); incorporated herein by reference.

[0297] In a particular embodiment, the CRISPR-associated guide molecule comprises a guide sequence linked to a direct repeat sequence, wherein the direct repeat sequence comprises one or more stem loops or optimized secondary structures. In one embodiment, the direct repeat has a minimum length of 16 nts and a single stem loop. In further embodiments the direct repeat has a length longer than 16 nts, preferably more than 17 nts, and has more than one stem loops or optimized secondary structures. In one embodiment, the CRISPR-associated guide molecule comprises or consists of the guide sequence linked to all or part of the natural direct repeat sequence. In one embodiment, certain aspects of the CRISPR-associated guide architecture can be modified, for example by addition, subtraction, or substitution of features, whereas certain other aspects of CRISPR-associated guide architecture are maintained. Preferred locations for engineered CRISPR-associated guide molecule modifications, including but not limited to insertions, deletions, and substitutions include CRISPR-associated guide termini and regions of the CRISPR-associated guide molecule that are exposed when complexed with nucleic acid-guided nuclease and/or target, for example the tetraloop and/or loop2.

[0298] In one embodiment, a loop in the CRISPR-associated guide RNA is provided. This may be a stem loop or a tetra loop. The loop is preferably GAAA, but it is not limited to this sequence or indeed to being only 4bp in length. Indeed, preferred loop forming sequences for use in hairpin structures are four nucleotides in length, and most preferably have the sequence GAAA. However, longer or shorter loop sequences may be used, as may alternative sequences. The sequences preferably include a nucleotide triplet (for example, AAA), and an additional nucleotide (for example C or G). Examples of loop forming sequences include CAAA and AAAG.

[0299] In one embodiment, the CRISPR-associated guide molecule forms a stemloop with a separate non-covalently linked sequence, which can be DNA or RNA. In one embodiment, the sequences forming the CRISPR-associated guide are first synthesized using the standard phosphoramidite synthetic protocol (Herdewijn, P., ed., *Methods in Molecular Biology* Col 288, *Oligonucleotide Synthesis: Methods and Applications*, Humana Press, New Jersey (2012)). In one embodiment, these sequences can be functionalized to contain an appropriate functional group for ligation using the standard protocol known in the art (Hermanson, G. T., *Bioconjugate Techniques*, Academic Press (2013)). Examples of functional groups include, but are not limited to, hydroxyl, amine, carboxylic acid, carboxylic acid halide, carboxylic acid active ester, aldehyde, carbonyl, chlorocarbonyl, imidazolylcarbonyl, hydrozide, semicarbazide, thio semicarbazide, thiol, maleimide, haloalkyl, sulfonyl, ally, propargyl, diene, alkyne, and azide. Once this sequence is functionalized, a covalent chemical bond or linkage can be formed between this sequence and the direct repeat sequence. Examples of chemical bonds include, but are not limited to, those based on carbamates, ethers, esters, amides, imines, amidines, aminotrizines, hydrozone, disulfides, thioethers, thioesters, phosphorothioates, phosphorodithioates, sulfonamides, sulfonates, fulfones, sulfoxides, ureas, thioureas, hydrazide, oxime, triazole, photolabile linkages, C-C bond forming groups such as Diels-Alder cyclo-addition pairs or ring-closing metathesis pairs, and Michael reaction pairs.

[0300] In one embodiment, these stem-loop forming sequences can be chemically synthesized. In one embodiment, the chemical synthesis uses automated, solid-phase oligonucleotide synthesis machines with 2'-acetoxyethyl orthoester (2'-ACE) (Scaringe et al., *J. Am. Chem. Soc.* (1998) 120: 11820-11821; Scaringe, *Methods Enzymol.* (2000) 317: 3-18) or 2'-thionocarbamate (2'-TC) chemistry (Dellinger et al., *J. Am. Chem. Soc.* (2011) 133: 11540-11546; Hendel et al., *Nat. Biotechnol.* (2015) 33:985-989).

[0301] The repeat:anti repeat duplex will be apparent from the secondary structure of the sgRNA. It may be typically a first complimentary stretch after (in 5' to 3' direction) the poly U tract and before the tetraloop; and a second complimentary stretch after (in 5' to 3' direction) the tetraloop and before the poly A tract. The first complimentary stretch (the "repeat") is complimentary to the second complimentary stretch (the "anti-repeat"). As such, they Watson-Crick base pair to form a duplex of dsRNA when folded back on one another. As such, the anti-repeat sequence is the complimentary sequence of the repeat and in terms to A-U or C-G

base pairing, but also in terms of the fact that the anti-repeat is in the reverse orientation due to the tetraloop.

[0302] In an embodiment of the invention, modification of CRISPR-associated guide architecture comprises replacing bases in stemloop 2. For example, In one embodiment, “actt” (“acuu” in RNA) and “aagt” (“aagu” in RNA) bases in stemloop2 are replaced with “cgcc” and “gcgg”. In one embodiment, “actt” and “aagt” bases in stemloop2 are replaced with complimentary GC-rich regions of 4 nucleotides. In one embodiment, the complimentary GC-rich regions of 4 nucleotides are “cgcc” and “gcgg” (both in 5’ to 3’ direction). In one embodiment, the complimentary GC-rich regions of 4 nucleotides are “gcgg” and “cgcc” (both in 5’ to 3’ direction). Other combination of C and G in the complimentary GC-rich regions of 4 nucleotides will be apparent including CCCC and GGGG.

[0303] In one aspect, the stemloop 2, e.g., “ACTTgtttAAGT” can be replaced by any “XXXXgtttYYYY”, e.g., where XXXX and YYYY represent any complementary sets of nucleotides that together will base pair to each other to create a stem.

[0304] In one aspect, the stem comprises at least about 4 bp comprising complementary X and Y sequences, although stems of more, e.g., 5, 6, 7, 8, 9, 10, 11 or 12 or fewer, e.g., 3, 2, base pairs are also contemplated. Thus, for example X2-12 and Y2-12 (wherein X and Y represent any complementary set of nucleotides) may be contemplated. In one aspect, the stem made of the X and Y nucleotides, together with the “gttt,” will form a complete hairpin in the overall secondary structure; and, this may be advantageous and the amount of base pairs can be any amount that forms a complete hairpin. In one aspect, any complementary X:Y basepairing sequence (e.g., as to length) is tolerated, so long as the secondary structure of the entire CRISPR-associated sgRNA is preserved. In one aspect, the stem can be a form of X:Y basepairing that does not disrupt the secondary structure of the whole CRISPR-associated sgRNA in that it has a DR:tracr duplex, and 3 stemloops. In one aspect, the “gttt” tetraloop that connects ACTT and AAGT (or any alternative stem made of X:Y basepairs) can be any sequence of the same length (e.g., 4 basepair) or longer that does not interrupt the overall secondary structure of the sgRNA. In one aspect, the stemloop can be something that further lengthens stemloop2, e.g. can be MS2 aptamer. In one aspect, the stemloop3 “GGCACCgagtCGGTGC” can likewise take on a “XXXXXXXXagtYYYYYYY” form, e.g., wherein X7 and Y7 represent any complementary sets of nucleotides that together will base pair to each other to create a stem. In one aspect, the stem comprises about 7bp comprising

complementary X and Y sequences, although stems of more or fewer basepairs are also contemplated. In one aspect, the stem made of the X and Y nucleotides, together with the “agt”, will form a complete hairpin in the overall secondary structure. In one aspect, any complementary X:Y basepairing sequence is tolerated, so long as the secondary structure of the entire sgRNA is preserved. In one aspect, the stem can be a form of X:Y basepairing that doesn't disrupt the secondary structure of the whole sgRNA in that it has a DR:tracr duplex, and 3 stemloops. In one aspect, the “agt” sequence of the stemloop 3 can be extended or be replaced by an aptamer, e.g., a MS2 aptamer or sequence that otherwise generally preserves the architecture of stemloop3. In one aspect for alternative Stemloops 2 and/or 3, each X and Y pair can refer to any basepair. In one aspect, non-Watson Crick basepairing is contemplated, where such pairing otherwise generally preserves the architecture of the stemloop at that position.

[0305] In one aspect, the DR:tracrRNA duplex can be replaced with the form: gYYYYag(N)NNNNxxxxNNNN(AAN)uuRRRRu (using standard IUPAC nomenclature for nucleotides), wherein (N) and (AAN) represent part of the bulge in the duplex, and “xxxx” represents a linker sequence. NNNN on the direct repeat can be anything so long as it basepairs with the corresponding NNNN portion of the tracrRNA. In one aspect, the DR:tracrRNA duplex can be connected by a linker of any length, any base composition, as long as it doesn't alter the overall structure.

[0306] In one embodiment, the natural hairpin or stemloop structure of the CRISPR-associated guide molecule is extended or replaced by an extended stemloop. Extension of the stem can enhance the assembly of the CRISPR-associated guide molecule with the nucleic acid-guided nuclease. In one embodiment the stem of the stemloop is extended by at least 1, 2, 3, 4, 5 or more complementary basepairs (i.e. corresponding to the addition of 2,4, 6, 8, 10 or more nucleotides in the CRISPR-associated guide molecule). In one embodiment these are located at the end of the stem, adjacent to the loop of the stemloop.

[0307] In one embodiment, the susceptibility of the CRISPR-associated guide molecule to RNases or to decreased expression can be reduced by slight modifications of the sequence of the CRISPR-associated guide molecule which do not affect its function. For instance, in one embodiment, premature termination of transcription, such as premature transcription of U6 Pol-III, can be removed by modifying a putative Pol-III terminator (4 consecutive U's) in the CRISPR-associated guide molecules sequence. Where such sequence modification is required

in the stemloop of the CRISPR-associated guide molecule, it is preferably ensured by a basepair flip.

[0308] In an embodiment, the CRISPR-associated guide molecule comprises non-naturally occurring nucleic acids and/or non-naturally occurring nucleotides and/or nucleotide analogs, and/or chemically modifications. Preferably, these non-naturally occurring nucleic acids and non-naturally occurring nucleotides are located outside the CRISPR-associated guide sequence. Non-naturally occurring nucleic acids can include, for example, mixtures of naturally and non-naturally occurring nucleotides. Non-naturally occurring nucleotides and/or nucleotide analogs may be modified at the ribose, phosphate, and/or base moiety. In an embodiment of the invention, a CRISPR-associated guide nucleic acid comprises ribonucleotides and non-ribonucleotides. In one such embodiment, a CRISPR-associated guide comprises one or more ribonucleotides and one or more deoxyribonucleotides. In an embodiment of the invention, the CRISPR-associated guide comprises one or more non-naturally occurring nucleotide or nucleotide analog such as a nucleotide with phosphorothioate linkage, a locked nucleic acid (LNA) nucleotides comprising a methylene bridge between the 2' and 4' carbons of the ribose ring, or bridged nucleic acids (BNA). Other examples of modified nucleotides include 2'-O-methyl analogs, 2'-deoxy analogs, or 2'-fluoro analogs. Further examples of modified bases include, but are not limited to, 2-aminopurine, 5-bromouridine, pseudouridine, inosine, 7-methylguanosine. Examples of guide RNA chemical modifications include, without limitation, incorporation of 2'-O-methyl (M), 2'-O-methyl 3'phosphorothioate (MS), S-constrained ethyl(cEt), or 2'-O-methyl 3'thioPACE (MSP) at one or more terminal nucleotides. Such chemically modified CRISPR-associated guides can comprise increased stability and increased activity as compared to unmodified CRISPR-associated guides, though on-target vs. off-target specificity is not predictable. (See, Hendel, 2015, Nat Biotechnol. 33(9):985-9, doi: 10.1038/nbt.3290, published online 29 June 2015 Ragdarm et al., 0215, PNAS, E7110-E7111; Allerson et al., J. Med. Chem. 2005, 48:901-904; Bramsen et al., Front. Genet., 2012, 3:154; Deng et al., PNAS, 2015, 112:11870-11875; Sharma et al., MedChemComm., 2014, 5:1454-1471; Hendel et al., Nat. Biotechnol. (2015) 33(9): 985-989; Li et al., Nature Biomedical Engineering, 2017, 1, 0066 DOI:10.1038/s41551-017-0066). In one embodiment, the 5' and/or 3' end of a CRISPR-associated guide RNA is modified by a variety of functional moieties including fluorescent dyes, polyethylene glycol, cholesterol, proteins, or detection tags. (See Kelly et al., 2016, J. Biotech. 233:74-83). In an

embodiment, a CRISPR-associated guide comprises ribonucleotides in a region that binds to a target sequence and one or more deoxyribonucleotides and/or nucleotide analogs in a region that binds to the nucleic acid-guided nuclease. In an embodiment, deoxyribonucleotides and/or nucleotide analogs are incorporated in engineered guide structures, such as, without limitation, stem-loop regions, and the seed region. In one embodiment, 3-5 nucleotides at either the 3' or the 5' end of a guide is chemically modified. In one embodiment, only minor modifications are introduced in the seed region, such as 2'-F modifications. In one embodiment, 2'-F modification is introduced at the 3' end of a guide. In an embodiment, three to five nucleotides at the 5' and/or the 3' end of the CRISPR-associated guide are chemically modified with 2'-O-methyl (M), 2'-O-methyl 3' phosphorothioate (MS), S-constrained ethyl(cEt), or 2'-O-methyl 3' thioPACE (MSP). Such modification can enhance genome editing efficiency (see Hendel et al., *Nat. Biotechnol.* (2015) 33(9): 985-989). In an embodiment, all of the phosphodiester bonds of a CRISPR-associated guide are substituted with phosphorothioates (PS) for enhancing levels of gene disruption. In an embodiment, more than five nucleotides at the 5' and/or the 3' end of the CRISPR-associated guide are chemically modified with 2'-O-Me, 2'-F or S-constrained ethyl(cEt). Such chemically modified CRISPR-associated guide can mediate enhanced levels of gene disruption (see Ragdarm et al., 0215, *PNAS*, E7110-E7111). In an embodiment of the invention, a CRISPR-associated guide is modified to comprise a chemical moiety at its 3' and/or 5' end. Such moieties include, but are not limited to amine, azide, alkyne, thio, dibenzocyclooctyne (DBCO), or Rhodamine. In certain embodiment, the chemical moiety is conjugated to the CRISPR-associated guide by a linker, such as an alkyl chain. In an embodiment, the chemical moiety of the modified CRISPR-associated guide can be used to attach the CRISPR-associated guide to another molecule, such as DNA, RNA, protein, or nanoparticles. Such chemically modified CRISPR-associated guide can be used to identify or enrich cells generically edited by a nucleic acid-guided nuclease and related systems (see Lee et al., *eLife*, 2017, 6:e25312, DOI:10.7554).

[0309] In a particular embodiment, the direct repeat may be modified to comprise one or more protein-binding RNA aptamers. In a particular embodiment, one or more aptamers may be included such as part of optimized secondary structure. Such aptamers may be capable of binding a bacteriophage coat protein as detailed further herein.

[0310] In one embodiment, the nucleic acid-guided nuclease may need a tracr sequence. The "tracrRNA" sequence or analogous terms includes any polynucleotide sequence that has

sufficient complementarity with a crRNA sequence to hybridize. In one embodiment, the degree of complementarity between the tracrRNA sequence and crRNA sequence along the length of the shorter of the two when optimally aligned is about or more than about 25%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97.5%, 99%, or higher. In one embodiment, the tracr sequence is about or more than about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50, or more nucleotides in length. In one embodiment, the tracr sequence and CRISPR-associated guide sequence are contained within a single transcript, such that hybridization between the two produces a transcript having a secondary structure, such as a hairpin. In an embodiment of the invention, the transcript or transcribed polynucleotide sequence has at least two or more hairpins. In preferred embodiments, the transcript has two, three, four or five hairpins. In a further embodiment of the invention, the transcript has at most five hairpins. In a hairpin structure the portion of the sequence 5' of the final "N" and upstream of the loop may correspond to the tracr mate sequence, and the portion of the sequence 3' of the loop then corresponds to the tracr sequence. In a hairpin structure the portion of the sequence 5' of the final "N" and upstream of the loop may alternatively correspond to the tracr sequence, and the portion of the sequence 3' of the loop corresponds to the tracr mate sequence.

[0311] In one embodiment, the tracr and tracr mate sequences can be chemically synthesized. In one embodiment, the chemical synthesis uses automated, solid-phase oligonucleotide synthesis machines with 2'-acetoxyethyl orthoester (2'-ACE) (Scaringe et al., J. Am. Chem. Soc. (1998) 120: 11820-11821; Scaringe, Methods Enzymol. (2000) 317: 3-18) or 2'-thionocarbamate (2'-TC) chemistry (Dellinger et al., J. Am. Chem. Soc. (2011) 133: 11540-11546; Hendel et al., Nat. Biotechnol. (2015) 33:985-989).

[0312] In one embodiment, the tracr and tracr mate sequences can be covalently linked using various bioconjugation reactions, loops, bridges, and non-nucleotide links via modifications of sugar, internucleotide phosphodiester bonds, purine and pyrimidine residues. Sletten et al., Angew. Chem. Int. Ed. (2009) 48:6974-6998; Manoharan, M. Curr. Opin. Chem. Biol. (2004) 8: 570-9; Behlke et al., Oligonucleotides (2008) 18: 305-19; Watts, et al., Drug. Discov. Today (2008) 13: 842-55; Shukla, et al., ChemMedChem (2010) 5: 328-49.

[0313] In one embodiment, the tracr and tracr mate sequences can be covalently linked using click chemistry. In one embodiment, the tracr and tracr mate sequences can be covalently linked using a triazole linker. In one embodiment, the tracr and tracr mate sequences can be covalently linked using Huisgen 1,3-dipolar cycloaddition reaction involving an alkyne and

azide to yield a highly stable triazole linker (He et al., ChemBioChem (2015) 17: 1809-1812; WO 2016/186745). In one embodiment, the tracr and tracr mate sequences are covalently linked by ligating a 5'-hexyne tracrRNA and a 3'-azide crRNA. In one embodiment, either or both of the 5'-hexyne tracrRNA and a 3'-azide crRNA can be protected with 2'-acetoxyethyl orthoester (2'-ACE) group, which can be subsequently removed using Dharmacon protocol (Scaringe et al., J. Am. Chem. Soc. (1998) 120: 11820-11821; Scaringe, Methods Enzymol. (2000) 317: 3-18).

[0314] In one embodiment, the tracr and tracr mate sequences can be covalently linked via a linker (e.g., a non-nucleotide loop) that comprises a moiety such as spacers, attachments, bioconjugates, chromophores, reporter groups, dye labeled RNAs, and non-naturally occurring nucleotide analogues. More specifically, suitable spacers for purposes of this invention include, but are not limited to, polyethers (e.g., polyethylene glycols, polyalcohols, polypropylene glycol or mixtures of ethylene and propylene glycols), polyamines group (e.g., spermine, spermidine and polymeric derivatives thereof), polyesters (e.g., poly(ethyl acrylate)), polyphosphodiester, alkylenes, and combinations thereof. Suitable attachments include any moiety that can be added to the linker to add additional properties to the linker, such as but not limited to, fluorescent labels. Suitable bioconjugates include, but are not limited to, peptides, glycosides, lipids, cholesterol, phospholipids, diacyl glycerols and dialkyl glycerols, fatty acids, hydrocarbons, enzyme substrates, steroids, biotin, digoxigenin, carbohydrates, polysaccharides. Suitable chromophores, reporter groups, and dye-labeled RNAs include, but are not limited to, fluorescent dyes such as fluorescein and rhodamine, chemiluminescent, electrochemiluminescent, and bioluminescent marker compounds. The design of example linkers conjugating two RNA components are also described in WO 2004/015075.

[0315] The linker (e.g., a non-nucleotide loop) can be of any length. In one embodiment, the linker has a length equivalent to about 0-16 nucleotides. In one embodiment, the linker has a length equivalent to about 0-8 nucleotides. In one embodiment, the linker has a length equivalent to about 0-4 nucleotides. In one embodiment, the linker has a length equivalent to about 2 nucleotides. Example linker design is also described in International Patent Publication No. WO 2011/008730.

[0316] In an embodiment, the nucleic acid-guided nuclease uses of a tracrRNA, the CRISPR-associated guide sequence, tracr mate, and tracr sequence may reside in a single RNA, i.e. an sgRNA (arranged in a 5' to 3' orientation or alternatively arranged in a 3' to 5'

orientation), or the tracr RNA may be a different RNA than the RNA containing the CRISPR-associated guide and tracr mate sequence. In these embodiments, the tracr hybridizes to the tracr mate sequence and directs the nucleic acid-guided nuclease-guide molecule complex to the target sequence. In some examples, a CRISPR-associated sgRNA comprises (in 5' to 3' direction): a CRISPR-associated guide sequence, a poly U tract, a first complimentary stretch (the "repeat"), a loop (tetraloop), a second complimentary stretch (the "anti-repeat" being complimentary to the repeat), a stem, and further stem loops and stems and a poly A (often poly U in RNA) tail (terminator). In preferred embodiments, certain aspects of CRISPR-associated guide architecture are retained, certain aspect of CRISPR-associated guide architecture can be modified, for example by addition, subtraction, or substitution of features, whereas certain other aspects of CRISPR-associated guide architecture are maintained. Preferred locations for engineered CRISPR-associated sgRNA modifications, including but not limited to insertions, deletions, and substitutions include CRISPR-associated guide termini and regions of the CRISPR-associated sgRNA that are exposed when complexed with nucleic acid-guided nuclease and/or target, for example the tetraloop and/or loop2.

[0317] In one embodiment, the CRISPR-associated guide molecule comprises, in addition the CRISPR-associated guide sequence, a sequence corresponding to a direct repeat in the CRISPR locus. In one embodiment, this sequence comprises at least one hairpin, i.e., a region of self-complementarity. In one embodiment, the CRISPR-associated guide sequence is 3' of the direct repeat comprising at least one hairpin. In further embodiments, the CRISPR-associated guide sequence is 5' of the direct repeat comprising at least one hairpin. In one embodiment, a hairpin is located in the middle of the CRISPR-associated guide sequence, i.e. the CRISPR-associated guide sequence is in part 5' and in part 3' of the direct repeat. The hairpin in the middle of the CRISPR-associated guide sequence may be involved in recognition or processing of the guide molecule. In one embodiment, the hairpin structure comprises at least 5, preferably 7-20 nucleotides.

Escorted guides

[0318] In one embodiment, the compositions or complexes have a CRISPR-associated guide molecule with a functional structure designed to improve CRISPR-associated guide molecule structure, architecture, stability, genetic expression, or any combination thereof. Such a structure can include an aptamer.

[0319] Aptamers are biomolecules that can be designed or selected to bind tightly to other ligands, for example using a technique called systematic evolution of ligands by exponential enrichment (SELEX; Tuerk C, Gold L: "Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase." *Science* 1990, 249:505-510). Nucleic acid aptamers can for example be selected from pools of random-sequence oligonucleotides, with high binding affinities and specificities for a wide range of biomedically relevant targets, suggesting a wide range of therapeutic utilities for aptamers (Keefe, Anthony D., Supriya Pai, and Andrew Ellington. "Aptamers as therapeutics." *Nature Reviews Drug Discovery* 9.7 (2010): 537-550). These characteristics also suggest a wide range of uses for aptamers as drug delivery vehicles (Levy-Nissenbaum, Etgar, et al. "Nanotechnology and aptamers: applications in drug delivery." *Trends in biotechnology* 26.8 (2008): 442-449; and, Hicke BJ, Stephens AW. "Escort aptamers: a delivery service for diagnosis and therapy." *J Clin Invest* 2000, 106:923-928.). Aptamers may also be constructed that function as molecular switches, responding to a cue by changing properties, such as RNA aptamers that bind fluorophores to mimic the activity of green fluorescent protein (Paige, Jeremy S., Karen Y. Wu, and Samie R. Jaffrey. "RNA mimics of green fluorescent protein." *Science* 333.6042 (2011): 642-646). It has also been suggested that aptamers may be used as components of targeted siRNA therapeutic delivery systems, for example targeting cell surface proteins (Zhou, Jiehua, and John J. Rossi. "Aptamer-targeted cell-specific RNA interference." *Silence* 1.1 (2010): 4).

[0320] Accordingly, in one embodiment, the CRISPR-associated guide molecule is modified, e.g., by one or more aptamer(s) designed to improve CRISPR-associated guide molecule delivery, including delivery across the cellular membrane, to intracellular compartments, or into the nucleus. Such a structure can include, either in addition to the one or more aptamer(s) or without such one or more aptamer(s), moiety(ies) so as to render the CRISPR-associated guide molecule deliverable, inducible or responsive to a selected effector. The invention accordingly comprehends a CRISPR-associated guide molecule that responds to normal or pathological physiological conditions, including without limitation pH, hypoxia, O₂ concentration, temperature, protein concentration, enzymatic concentration, lipid structure, light exposure, mechanical disruption (e.g. ultrasound waves), magnetic fields, electric fields, or electromagnetic radiation.

[0321] Light responsiveness of an inducible system may be achieved via the activation and binding of cryptochrome-2 and CIB1. Blue light stimulation induces an activating conformational change in cryptochrome-2, resulting in recruitment of its binding partner CIB1. This binding is fast and reversible, achieving saturation in <15 sec following pulsed stimulation and returning to baseline <15 min after the end of stimulation. These rapid binding kinetics result in a system temporally bound only by the speed of transcription/translation and transcript/protein degradation, rather than uptake and clearance of inducing agents. Cryptochrome-2 activation is also highly sensitive, allowing for the use of low light intensity stimulation and mitigating the risks of phototoxicity. Further, in a context such as the intact mammalian brain, variable light intensity may be used to control the size of a stimulated region, allowing for greater precision than vector delivery alone may offer.

[0322] Energy sources such as electromagnetic radiation, sound energy or thermal energy may induce the CRISPR-associated guide. Advantageously, the electromagnetic radiation is a component of visible light. In a preferred embodiment, the light is a blue light with a wavelength of about 450 to about 495 nm. In an especially preferred embodiment, the wavelength is about 488 nm. In another preferred embodiment, the light stimulation is via pulses. The light power may range from about 0-9 mW/cm². In a preferred embodiment, a stimulation paradigm of as low as 0.25 sec every 15 sec should result in maximal activation.

[0323] The chemical or energy sensitive CRISPR-associated guide may undergo a conformational change upon induction by the binding of a chemical source or by the energy allowing it act as a CRISPR-associated guide and have the nucleic acid-guided nuclease system or complex function. The invention can involve applying the chemical source or energy so as to have the CRISPR-associated guide function and the nucleic acid-guided nuclease system or complex function; and optionally further determining that the expression of the genomic locus is altered.

[0324] There are several different designs of this chemical inducible system: 1. ABI-PYL based system inducible by Absciscic Acid (ABA) (see, e.g., stke.sciencemag.org/cgi/content/abstract/sigtrans;4/164/rs2), 2. FKBP-FRB based system inducible by rapamycin (or related chemicals based on rapamycin) (see, e.g., www.nature.com/nmeth/journal/v2/n6/full/nmeth763.html), 3. GID1-GAI based system inducible by Gibberellin (GA) (see, e.g., www.nature.com/nchembio/journal/v8/n5/full/nchembio.922.html).

[0325] A chemical inducible system can be an estrogen receptor (ER) based system inducible by 4-hydroxytamoxifen (4OHT) (see, e.g., www.pnas.org/content/104/3/1027.abstract). A mutated ligand-binding domain of the estrogen receptor called ERT2 translocates into the nucleus of cells upon binding of 4-hydroxytamoxifen. In further embodiments of the invention any naturally occurring or engineered derivative of any nuclear receptor, thyroid hormone receptor, retinoic acid receptor, estrogen receptor, estrogen-related receptor, glucocorticoid receptor, progesterone receptor, androgen receptor may be used in inducible systems analogous to the ER based inducible system.

[0326] Another inducible system is based on the design using Transient receptor potential (TRP) ion channel-based system inducible by energy, heat or radio-wave (see, e.g., www.sciencemag.org/content/336/6081/604). These TRP family proteins respond to different stimuli, including light and heat. When this protein is activated by light or heat, the ion channel will open and allow the entering of ions such as calcium into the plasma membrane. This influx of ions will bind to intracellular ion interacting partners linked to a polypeptide including the guide and the other components of the nucleic acid-guided nuclease/ CRISPR-associated guide molecule complex or system, and the binding will induce the change of sub-cellular localization of the polypeptide, leading to the entire polypeptide entering the nucleus of cells. Once inside the nucleus, the guide protein and the other components of the nucleic acid-guided nuclease/ CRISPR-associated guide molecule complex will be active and modulating target gene expression in cells.

[0327] While light activation may be an advantageous embodiment, sometimes it may be disadvantageous especially for *in vivo* applications in which the light may not penetrate the skin or other organs. In this instance, other methods of energy activation are contemplated, in particular, electric field energy and/or ultrasound which have a similar effect.

[0328] Electric field energy is preferably administered substantially as described in the art, using one or more electric pulses of from about 1 Volt/cm to about 10 kVolts/cm under *in vivo* conditions. Instead of or in addition to the pulses, the electric field may be delivered in a continuous manner. The electric pulse may be applied for between 1 μ s and 500 milliseconds, preferably between 1 μ s and 100 milliseconds. The electric field may be applied continuously or in a pulsed manner for 5 about minutes.

[0329] As used herein, ‘electric field energy’ is the electrical energy to which a cell is exposed. Preferably the electric field has a strength of from about 1 Volt/cm to about 10 kVolts/cm or more under *in vivo* conditions (see WO97/49450).

[0330] As used herein, the term “electric field” includes one or more pulses at variable capacitance and voltage and including exponential and/or square wave and/or modulated wave and/or modulated square wave forms. References to electric fields and electricity should be taken to include reference the presence of an electric potential difference in the environment of a cell. Such an environment may be set up by way of static electricity, alternating current (AC), direct current (DC), etc., as known in the art. The electric field may be uniform, non-uniform or otherwise, and may vary in strength and/or direction in a time dependent manner.

[0331] Single or multiple applications of electric field, as well as single or multiple applications of ultrasound are also possible, in any order and in any combination. The ultrasound and/or the electric field may be delivered as single or multiple continuous applications, or as pulses (pulsatile delivery).

[0332] Electroporation has been used in both *in vitro* and *in vivo* procedures to introduce foreign material into living cells. With *in vitro* applications, a sample of live cells is first mixed with the agent of interest and placed between electrodes such as parallel plates. Then, the electrodes apply an electrical field to the cell/implant mixture. Examples of systems that perform *in vitro* electroporation include the Electro Cell Manipulator ECM600 product, and the Electro Square Porator T820, both made by the BTX Division of Genetronics, Inc (see U.S. Pat. No 5,869,326).

[0333] The known electroporation techniques (both *in vitro* and *in vivo*) function by applying a brief high voltage pulse to electrodes positioned around the treatment region. The electric field generated between the electrodes causes the cell membranes to temporarily become porous, whereupon molecules of the agent of interest enter the cells. In known electroporation applications, this electric field comprises a single square wave pulse on the order of 1000 V/cm, of about 100 .mu.s duration. Such a pulse may be generated, for example, in known applications of the Electro Square Porator T820.

[0334] Preferably, the electric field has a strength of from about 1 V/cm to about 10 kV/cm under *in vitro* conditions. Thus, the electric field may have a strength of 1 V/cm, 2 V/cm, 3 V/cm, 4 V/cm, 5 V/cm, 6 V/cm, 7 V/cm, 8 V/cm, 9 V/cm, 10 V/cm, 20 V/cm, 50 V/cm, 100 V/cm, 200 V/cm, 300 V/cm, 400 V/cm, 500 V/cm, 600 V/cm, 700 V/cm, 800 V/cm, 900 V/cm,

1 kV/cm, 2 kV/cm, 5 kV/cm, 10 kV/cm, 20 kV/cm, 50 kV/cm or more. More preferably from about 0.5 kV/cm to about 4.0 kV/cm under *in vitro* conditions. Preferably the electric field has a strength of from about 1 V/cm to about 10 kV/cm under *in vivo* conditions. However, the electric field strengths may be lowered where the number of pulses delivered to the target site are increased. Thus, pulsatile delivery of electric fields at lower field strengths is envisaged.

[0335] Preferably, the application of the electric field is in the form of multiple pulses such as double pulses of the same strength and capacitance or sequential pulses of varying strength and/or capacitance. As used herein, the term “pulse” includes one or more electric pulses at variable capacitance and voltage and including exponential and/or square wave and/or modulated wave/square wave forms.

[0336] Preferably, the electric pulse is delivered as a waveform selected from an exponential wave form, a square wave form, a modulated wave form and a modulated square wave form.

[0337] A preferred embodiment employs direct current at low voltage. Thus, Applicants disclose the use of an electric field which is applied to the cell, tissue or tissue mass at a field strength of between 1V/cm and 20V/cm, for a period of 100 milliseconds or more, preferably 15 minutes or more.

[0338] Ultrasound is advantageously administered at a power level of from about 0.05 W/cm² to about 100 W/cm². Diagnostic or therapeutic ultrasound may be used, or combinations thereof.

[0339] As used herein, the term “ultrasound” refers to a form of energy which consists of mechanical vibrations the frequencies of which are so high they are above the range of human hearing. Lower frequency limit of the ultrasonic spectrum may generally be taken as about 20 kHz. Most diagnostic applications of ultrasound employ frequencies in the range 1 and 15 MHz' (From Ultrasonics in Clinical Diagnosis, P. N. T. Wells, ed., 2nd. Edition, Publ. Churchill Livingstone [Edinburgh, London & NY, 1977]).

[0340] Ultrasound has been used in both diagnostic and therapeutic applications. When used as a diagnostic tool ("diagnostic ultrasound"), ultrasound is typically used in an energy density range of up to about 100 mW/cm² (FDA recommendation), although energy densities of up to 750 mW/cm² have been used. In physiotherapy, ultrasound is typically used as an energy source in a range up to about 3 to 4 W/cm² (WHO recommendation). In other therapeutic applications, higher intensities of ultrasound may be employed, for example, HIFU

at 100 W/cm up to 1 kW/cm² (or even higher) for short periods of time. The term "ultrasound" as used in this specification is intended to encompass diagnostic, therapeutic and focused ultrasound.

[0341] Focused ultrasound (FUS) allows thermal energy to be delivered without an invasive probe (see Morocz et al 1998 Journal of Magnetic Resonance Imaging Vol.8, No. 1, pp.136-142. Another form of focused ultrasound is high intensity focused ultrasound (HIFU) which is reviewed by Moussatov et al in Ultrasonics (1998) Vol.36, No.8, pp.893-900 and TranHuuHue et al in Acustica (1997) Vol.83, No.6, pp.1103-1106.

[0342] Preferably, a combination of diagnostic ultrasound and a therapeutic ultrasound is employed. This combination is not intended to be limiting, however, and the skilled reader will appreciate that any variety of combinations of ultrasound may be used. Additionally, the energy density, frequency of ultrasound, and period of exposure may be varied.

[0343] Preferably, the exposure to an ultrasound energy source is at a power density of from about 0.05 to about 100 Wcm⁻². Even more preferably, the exposure to an ultrasound energy source is at a power density of from about 1 to about 15 Wcm⁻².

[0344] Preferably, the exposure to an ultrasound energy source is at a frequency of from about 0.015 to about 10.0 MHz. More preferably the exposure to an ultrasound energy source is at a frequency of from about 0.02 to about 5.0 MHz or about 6.0 MHz. Most preferably, the ultrasound is applied at a frequency of 3 MHz.

[0345] Preferably the exposure is for periods of from about 10 milliseconds to about 60 minutes. Preferably the exposure is for periods of from about 1 second to about 5 minutes. More preferably, the ultrasound is applied for about 2 minutes. Depending on the particular target cell to be disrupted, however, the exposure may be for a longer duration, for example, for 15 minutes.

[0346] Advantageously, the target tissue is exposed to an ultrasound energy source at an acoustic power density of from about 0.05 Wcm⁻² to about 10 Wcm⁻² with a frequency ranging from about 0.015 to about 10 MHz (see WO 98/52609). However, alternatives are also possible, for example, exposure to an ultrasound energy source at an acoustic power density of above 100 Wcm⁻², but for reduced periods of time, for example, 1000 Wcm⁻² for periods in the millisecond range or less.

[0347] Preferably, the application of the ultrasound is in the form of multiple pulses; thus, both continuous wave and pulsed wave (pulsatile delivery of ultrasound) may be employed in

any combination. For example, continuous wave ultrasound may be applied, followed by pulsed wave ultrasound, or vice versa. This may be repeated any number of times, in any order and combination. The pulsed wave ultrasound may be applied against a background of continuous wave ultrasound, and any number of pulses may be used in any number of groups.

[0348] Preferably, the ultrasound may comprise pulsed wave ultrasound. In a highly preferred embodiment, the ultrasound is applied at a power density of 0.7 Wcm^{-2} or 1.25 Wcm^{-2} as a continuous wave. Higher power densities may be employed if pulsed wave ultrasound is used.

[0349] Use of ultrasound is advantageous as, like light, it may be focused accurately on a target. Moreover, ultrasound is advantageous as it may be focused more deeply into tissues unlike light. It is therefore better suited to whole-tissue penetration (such as but not limited to a lobe of the liver) or whole organ (such as but not limited to the entire liver or an entire muscle, such as the heart) therapy. Another important advantage is that ultrasound is a non-invasive stimulus which is used in a wide variety of diagnostic and therapeutic applications. By way of example, ultrasound is well known in medical imaging techniques and, additionally, in orthopedic therapy. Furthermore, instruments suitable for the application of ultrasound to a subject vertebrate are widely available and their use is well known in the art.

[0350] In one embodiment, the CRISPR-associated guide molecule is modified by a secondary structure to increase the specificity of the nucleic acid-guided nuclease and related system and the secondary structure can protect against exonuclease activity and allow for 5' additions to the guide sequence also referred to herein as a protected CRISPR-associated guide molecule.

[0351] In one aspect, the invention provides for hybridizing a "protector RNA" to a sequence of the CRISPR-associated guide molecule, wherein the "protector RNA" is an RNA strand complementary to the 3' end of the guide molecule to thereby generate a partially double-stranded CRISPR-associated guide RNA. In an embodiment of the invention, protecting mismatched bases (i.e., the bases of the guide molecule which do not form part of the guide sequence) with a perfectly complementary protector sequence decreases the likelihood of target DNA binding to the mismatched basepairs at the 3' end. In one embodiment of the invention, additional sequences comprising an extended length may also be present within the CRISPR-associated guide molecule such that the CRISPR-associated guide comprises a protector sequence within the CRISPR-associated guide molecule. This "protector

sequence” ensures that the CRISPR-associated guide molecule comprises a “protected sequence” in addition to an “exposed sequence” (comprising the part of the CRISPR-associated guide sequence hybridizing to the target sequence). In one embodiment, the CRISPR-associated guide molecule is modified by the presence of the protector guide to comprise a secondary structure such as a hairpin. Advantageously there are three or four to thirty or more, e.g., about 10 or more, contiguous base pairs having complementarity to the protected sequence, the CRISPR-associated guide sequence or both. It is advantageous that the protected portion does not impede thermodynamics of the nucleic acid-guided nuclease and related system interacting with its target. By providing such an extension including a partially double stranded CRISPR-associated guide molecule, the CRISPR-associated guide molecule is considered protected and results in improved specific binding of the nucleic acid-guided nuclease/ CRISPR-associated guide molecule complex, while maintaining specific activity.

[0352] In one embodiment, use is made of a truncated CRISPR-associated guide (truncated CRISPR-associated guide), i.e. a CRISPR-associated guide molecule which comprises a CRISPR-associated guide sequence which is truncated in length with respect to the canonical CRISPR-associated guide sequence length. As described by Nowak et al. (Nucleic Acids Res (2016) 44 (20): 9555-9564), such guides may allow catalytically active nucleic acid-guided nuclease to bind its target without cleaving the target DNA. In one embodiment, a truncated CRISPR-associated guide is used which allows the binding of the target but retains only nickase activity of the nucleic acid-guided nuclease .

[0353] In one embodiment, conjugation of triantennary N-acetyl galactosamine (GalNAc) to oligonucleotide components may be used to improve delivery, for example delivery to select cell types, for example hepatocytes (see International Patent Publication No. WO 2014/118272 incorporated herein by reference; Nair, JK et al., 2014, Journal of the American Chemical Society 136 (49), 16958-16961). This is considered to be a sugar-based particle and further details on other particle delivery systems and/or formulations are provided herein. GalNAc can therefore be considered to be a particle in the sense of the other particles described herein, such that general uses and other considerations, for instance delivery of said particles, apply to GalNAc particles as well. A solution-phase conjugation strategy may for example be used to attach triantennary GalNAc clusters (mol. wt. ~2000) activated as PFP (pentafluorophenyl) esters onto 5'-hexylamino modified oligonucleotides (5'-HA ASOs, mol. wt. ~8000 Da; Østergaard et al., Bioconjugate Chem., 2015, 26 (8), pp 1451–1455). Similarly, poly(acrylate)

polymers have been described for *in vivo* nucleic acid delivery (see WO2013158141 incorporated herein by reference). In further alternative embodiments, pre-mixing nucleic acid-guided nuclease nanoparticles (or protein complexes) with naturally occurring serum proteins may be used in order to improve delivery (Akinc A et al, 2010, Molecular Therapy vol. 18 no. 7, 1357–1364).

[0354] Screening techniques are available to identify delivery enhancers, for example by screening chemical libraries (Gilleron J. et al., 2015, Nucl. Acids Res. 43 (16): 7984-8001). Approaches have also been described for assessing the efficiency of delivery vehicles, such as lipid nanoparticles, which may be employed to identify effective delivery vehicles for components (see Sahay G. et al., 2013, Nature Biotechnology 31, 653–658).

Functional domains

[0355] The nucleic acid-guided nuclease (including variants such as a catalytically inactive form) may be associated with one or more functional domains (e.g., via fusion protein or suitable linkers). In an embodiment, the nucleic acid-guided nuclease, or an ortholog or homolog thereof, may be used as a generic nucleic acid binding protein with fusion to or being operably linked to one or more functional domains. In one example, the functional domain is a deaminase. In another example, the functional domain is a transposase. In another example, the functional domain is a reverse transcriptase. In some cases, a functional domain may be associated with (e.g., fuse to) the nucleic acid-guided nuclease. In some cases, a functional domain may be a protein different from the nucleic acid-guided nuclease. In such cases, a functional domain and the nucleic acid-guided nuclease may form a protein complex.

[0356] It is also envisaged that the nucleic acid-guided nuclease-guide molecule complex as a whole may be associated with two or more functional domains. For example, there may be two or more functional domains associated with the nucleic acid-guided nuclease, or there may be two or more functional domains associated with the guide RNA or crRNA (via one or more adaptor proteins), or there may be one or more functional domains associated with the RNA-targeting effector protein and one or more functional domains associated with the guide RNA or crRNA (via one or more adaptor proteins).

[0357] In one embodiment, the nucleic acid-guided nuclease is associated with one or more functional domains. The association can be by direct linkage of the effector protein to the functional domain, or by association with the crRNA. In a non-limiting example, the crRNA comprises an added or inserted sequence that can be associated with a functional domain of

interest, including, for example, an aptamer or a nucleotide that binds to a nucleic acid binding adapter protein. The functional domain may be a functional heterologous domain.

[0358] In one embodiment, the invention also provides for the one or more heterologous functional domains to have one or more of the following activities: methylase activity, demethylase activity, transcription activation activity, transcription repression activity, transcription release factor activity, histone modification activity, nuclease activity, single-strand RNA cleavage activity, double-strand RNA cleavage activity, single-strand DNA cleavage activity, double-strand DNA cleavage activity and nucleic acid binding activity. At least one or more heterologous functional domains may be at or near the amino-terminus of the effector protein and/or wherein at least one or more heterologous functional domains is at or near the carboxy-terminus of the effector protein. The one or more heterologous functional domains may be fused to the effector protein. The one or more heterologous functional domains may be tethered to the effector protein. The one or more heterologous functional domains may be linked to the effector protein by a linker moiety.

[0359] In an embodiment, the nucleic acid-guided nuclease or an ortholog or homolog thereof, may be used as a generic nucleic acid binding protein with fusion to or being operably linked to a functional domain. Exemplary functional domains may include but are not limited to translational initiator, translational activator, translational repressor, nucleases, in particular ribonucleases, a spliceosome, beads, a light inducible/controllable domain or a chemically inducible/controllable domain. In an embodiment, the one or more functional domains are controllable, e.g., inducible.

[0360] In one embodiment, one or more functional domains are associated with a nucleic acid-guided nuclease via an adaptor protein, for example as used with the modified guides of Konnerman et al. (Nature 517, 583–588, 29 January 2015). In one embodiment, the one or more functional domains is attached to the adaptor protein so that upon binding of the nucleic acid-guided nuclease to the guide molecule and target, the functional domain is in a spatial orientation allowing for the functional domain to function in its attributed function.

[0361] In one embodiment, one or more functional domains are associated with a dead guide molecule, e.g., gRNA (dRNA). In one embodiment, a dRNA complex with active Nucleic acid-guided nuclease directs gene regulation by a functional domain at on gene locus while an gRNA directs DNA cleavage by the active Nucleic acid-guided nuclease at another locus, for example as described analogously in CRISPR-Cas systems by Dahlman et al.,

‘Orthogonal gene control with a catalytically active Cas9 nuclease’. In one embodiment, dRNAs are selected to maximize selectivity of regulation for a gene locus of interest compared to off-target regulation. In one embodiment, dRNAs are selected to maximize target gene regulation and minimize target cleavage

[0362] For the purposes of the following discussion, reference to a functional domain could be a functional domain associated with the Nucleic acid-guided nuclease or a functional domain associated with the adaptor protein. In one embodiment, the one or more functional domains is attached to the adaptor protein so that upon binding of the Nucleic acid-guided nuclease to the guide molecule and target, the functional domain is in a spatial orientation allowing for the functional domain to function in its attributed function.

[0363] In the practice of the invention, loops of the guide RNA may be extended, without colliding with the Nucleic acid-guided nuclease by the insertion of distinct RNA loop(s) or distinct sequence(s) that may recruit adaptor proteins that can bind to the distinct RNA loop(s) or distinct sequence(s). The adaptor proteins may include but are not limited to orthogonal RNA-binding protein / aptamer combinations that exist within the diversity of bacteriophage coat proteins. A list of such coat proteins includes, but is not limited to: Q β , F2, GA, fr, JP501, M12, R17, BZ13, JP34, JP500, KU1, M11, MX1, TW18, VK, SP, FI, ID2, NL95, TW19, AP205, ϕ Cb5, ϕ Cb8r, ϕ Cb12r, ϕ Cb23r, 7s and PRR1. These adaptor proteins or orthogonal RNA binding proteins can further recruit effector proteins or fusions which comprise one or more functional domains.

[0364] Examples of functional domains include deaminase domain, transposase domain, reverse transcriptase domain, integrase domain, recombinase domain, resolvase domain, invertase domain, protease domain, DNA methyltransferase domain, DNA hydroxymethylase domain, DNA demethylase domain, histone acetylase domain, histone deacetylases domain, nuclease domain, repressor domain, activator domain, nuclear-localization signal domains, transcription-regulatory protein (or transcription complex recruiting) domain, cellular uptake activity associated domain, nucleic acid binding domain, antibody presentation domain, histone modifying enzymes, recruiter of histone modifying enzymes; inhibitor of histone modifying enzymes, histone methyltransferase, histone demethylase, histone kinase, histone phosphatase, histone ribosylase, histone deribosylase, histone ubiquitinase, histone deubiquitinase, histone biotinase and histone tail protease. In some preferred embodiments, the functional domain is a transcriptional activation domain, such as, without limitation, VP64,

p65, MyoD1, HSF1, RTA, SET7/9 or a histone acetyltransferase. In one embodiment, the functional domain is a transcription repression domain, preferably KRAB. In one embodiment, the transcription repression domain is SID, or concatemers of SID (e.g. SID4X). In one embodiment, the functional domain is an epigenetic modifying domain, such that an epigenetic modifying enzyme is provided. In one embodiment, the functional domain is an activation domain, which may be the P65 activation domain.

[0365] In some examples, the Nucleic acid-guided nuclease is associated with a ligase or functional fragment thereof. The ligase may ligate a single-strand break (a nick) generated by the Nucleic acid-guided nuclease. In certain cases, the ligase may ligate a double-strand break generated by the nucleic acid-guided nuclease. In certain examples, the nucleic acid-guided nuclease is associated with a reverse transcriptase or functional fragment thereof.

[0366] In one embodiment, the one or more functional domains is a transcriptional repressor domain. In one embodiment, the transcriptional repressor domain is a KRAB domain. In one embodiment, the transcriptional repressor domain is a NuE domain, NcoR domain, SID domain or a SID4X domain.

[0367] In one embodiment, the one or more functional domains have one or more activities, e.g., one or more of transposase activity, methylase activity, demethylase activity, translation activation activity, translation repression activity, transcription activation activity, transcription repression activity, transcription release factor activity, chromatin modifying or remodeling activity, histone modification activity, nuclease activity, single-strand RNA cleavage activity, double-strand RNA cleavage activity, single-strand DNA cleavage activity, double-strand DNA cleavage activity, nucleic acid binding activity, and detectable activity.

[0368] Histone modifying domains are also preferred. In one embodiment, Exemplary histone modifying domains are discussed below. Transposase domains, HR (Homologous Recombination) machinery domains, recombinase domains, and/or integrase domains are also preferred as the present functional domains. In one embodiment, DNA integration activity includes HR machinery domains, integrase domains, recombinase domains and/or transposase domains.

[0369] In one embodiment, the DNA cleavage activity is due to a nuclease. In one embodiment, the nuclease comprises a FokI nuclease. See, “Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing”, Shengdar Q. Tsai, Nicolas Wyvekens, Cyd Khayter, Jennifer A. Foden, Vishal Thapar, Deepak Reyon, Mathew J. Goodwin, Martin J.

Aryee, J. Keith Joung *Nature Biotechnology* 32(6): 569-77 (2014), relates to dimeric RNA-guided FokI Nucleases that recognize extended sequences and can edit endogenous genes with high efficiencies in human cells.

[0370] In one embodiment, the one or more functional domains is attached to the nucleic acid-guided nuclease so that upon binding to the sgRNA and target the functional domain is in a spatial orientation allowing for the functional domain to function in its attributed function.

[0371] In one embodiment, the nucleic acid-guided nuclease comprise one or more heterologous functional domains. As used herein, a heterologous functional domain is a polypeptide that is not derived from the same species as the nucleic acid-guided nuclease. For example, a heterologous functional domain of a nucleic acid-guided nuclease derived from species A is a polypeptide derived from a species different from species A, or an artificial polypeptide. The one or more heterologous functional domains may comprise one or more nuclear localization signal (NLS) domains. The one or more heterologous functional domains may comprise at least two or more NLSs. The one or more heterologous functional domains may comprise one or more transcriptional activation domains. A transcriptional activation domain may comprise VP64. The one or more heterologous functional domains may comprise one or more transcriptional repression domains. A transcriptional repression domain may comprise a KRAB domain or a SID domain. The one or more heterologous functional domain may comprise one or more nuclease domains. The one or more nuclease domains may comprise FokI.

[0372] Functional domains may be used to regulate transcription, e.g., transcriptional repression. Transcriptional repression is often mediated by chromatin modifying enzymes such as histone methyltransferases (HMTs) and deacetylases (HDACs). Repressive histone effector domains are known and an exemplary list is provided below. In the exemplary table, preference was given to proteins and functional truncations of small size to facilitate efficient viral packaging (for instance via AAV). In general, however, the domains may include HDACs, histone methyltransferases (HMTs), and histone acetyltransferase (HAT) inhibitors, as well as HDAC and HMT recruiting proteins. The functional domain may be or include, In one embodiment, HDAC Effector Domains, HDAC Recruiter Effector Domains, Histone Methyltransferase (HMT) Effector Domains, Histone Methyltransferase (HMT) Recruiter Effector Domains, or Histone Acetyltransferase Inhibitor Effector Domains.

[0373] In one embodiment, the functional domain may be a Methyltransferase (HMT) Effector Domain. Preferred examples include NUE, vSET, EHMT2/G9A, SUV39H1, dim-5, KYP, SUVR4, SET4, SET1, SETD8, and TgSET8. NUE is exemplified in the present Examples and, although preferred, it is envisaged that others in the class will also be useful.

[0374] In one embodiment, the functional domain may be a Histone Methyltransferase (HMT) Recruiter Effector Domain. Preferred examples include Hp1a, PHF19, and NIPP1.

[0375] In one embodiment, the functional domain may be Histone Acetyltransferase Inhibitor Effector Domain. Preferred examples include SET/TAF-1 β .

[0376] In some cases, the target endogenous (regulatory) control elements (such as enhancers and silencers) in addition to a promoter or promoter-proximal elements. Thus, the invention can also be used to target endogenous control elements (including enhancers and silencers) in addition to targeting of the promoter. These control elements can be located upstream and downstream of the transcriptional start site (TSS), starting from 200bp from the TSS to 100kb away. Targeting of known control elements can be used to activate or repress the gene of interest. In some cases, a single control element can influence the transcription of multiple target genes. Targeting of a single control element could therefore be used to control the transcription of multiple genes simultaneously.

[0377] Targeting of putative control elements on the other hand (e.g. by tiling the region of the putative control element as well as 200bp up to 100kB around the element) can be used as a means to verify such elements (by measuring the transcription of the gene of interest) or to detect novel control elements (e.g. by tiling 100kb upstream and downstream of the TSS of the gene of interest). In addition, targeting of putative control elements can be useful in the context of understanding genetic causes of disease. Many mutations and common SNP variants associated with disease phenotypes are located outside coding regions. Targeting of such regions with either the activation or repression systems described herein can be followed by readout of transcription of either a) a set of putative targets (e.g. a set of genes located in closest proximity to the control element) or b) whole-transcriptome readout by e.g. RNAseq or microarray. This would allow for the identification of likely candidate genes involved in the disease phenotype. Such candidate genes could be useful as novel drug targets.

[0378] In one embodiment is for the one or more functional domains to comprise an acetyltransferase, preferably a histone acetyltransferase. These are useful in the field of epigenomics, for example in methods of interrogating the epigenome. Methods of interrogating

the epigenome may include, for example, targeting epigenomic sequences. Targeting epigenomic sequences may include the guide being directed to an epigenomic target sequence. Epigenomic target sequence may include, In one embodiment, include a promoter, silencer or an enhancer sequence.

[0379] The functional domains may be acetyltransferases domains. Examples of acetyltransferases are known but may include, In one embodiment, histone acetyltransferases. In one embodiment, the histone acetyltransferase may comprise the catalytic core of the human acetyltransferase p300 (Gerbasch & Reddy, Nature Biotech 6th April 2015).

Nuclear localization sequences

[0380] In one embodiment, the nucleic acid-guided nuclease is fused to one or more nuclear localization sequences (NLSs), such as about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more NLSs. In one embodiment, the Nucleic acid-guided nuclease comprises about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more NLSs at or near the amino-terminus, about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more NLSs at or near the carboxy-terminus, or a combination of these (e.g. zero or at least one or more NLS at the amino-terminus and zero or at one or more NLS at the carboxy terminus).

[0381] In one embodiment, the IscB polypeptide nuclease is fused to one or more nuclear localization sequences (NLSs), such as about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more NLSs. In one embodiment, the IscB polypeptide nuclease comprises about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more NLSs at or near the amino-terminus, about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more NLSs at or near the carboxy-terminus, or a combination of these (e.g. zero or at least one or more NLS at the amino-terminus and zero or at one or more NLS at the carboxy terminus).

[0382] When more than one NLS is present, each may be selected independently of the others, such that a single NLS may be present in more than one copy and/or in combination with one or more other NLSs present in one or more copies. In a preferred embodiment of the invention, the Nucleic acid-guided nuclease comprises at most 6 NLSs. In a preferred embodiment of the invention, the IscB polypeptide nuclease comprises at most 6 NLSs.

[0383] In one embodiment, an NLS is considered near the N- or C-terminus when the nearest amino acid of the NLS is within about 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50, or more

amino acids along the polypeptide chain from the N- or C-terminus. Non-limiting examples of NLSs include an NLS sequence derived from: the NLS of the SV40 virus large T-antigen, having the amino acid sequence PKKKRKV (SEQ ID NO: 2002); the NLS from nucleoplasmin (e.g. the nucleoplasmin bipartite NLS with the sequence KRPAATKKAGQAKKKK (SEQ ID NO: 2003); the c-myc NLS having the amino acid sequence PAAKRVKLD (SEQ ID NO: 2004) or RQRRNELKRSP (SEQ ID NO: 2005); the hRNPA1 M9 NLS having the sequence NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID NO: 2006); the sequence RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO: 2007) of the IBB domain from importin-alpha; the sequences VSRKRPRP (SEQ ID NO: 2008) and PPKKARED (SEQ ID NO: 2009) of the myoma T protein; the sequence PQPKKKPL (SEQ ID NO: 2010) of human p53; the sequence SALIKKKKKMAP (SEQ ID NO: 2011) of mouse c-abl IV; the sequences DRLRR (SEQ ID NO: 2012) and PKQKKRK (SEQ ID NO: 2013) of the influenza virus NS1; the sequence RKLKKIKKL (SEQ ID NO: 2014) of the Hepatitis virus delta antigen; the sequence REKKKFLKRR (SEQ ID NO: 2015) of the mouse Mx1 protein; the sequence KRKGDEV DGVDEVAKKSKK (SEQ ID NO: 2016) of the human poly(ADP-ribose) polymerase; and the sequence RKCLQAGMNLEARKTKK (SEQ ID NO: 2017) of the steroid hormone receptors (human) glucocorticoid.

[0384] In general, the one or more NLSs are of sufficient strength to drive accumulation of the nucleic acid-guided nuclease in a detectable amount in the nucleus of a eukaryotic cell. In general, strength of nuclear localization activity may derive from the number of NLSs in the nucleic acid-guided nuclease, the particular NLS(s) used, or a combination of these factors. Detection of accumulation in the nucleus may be performed by any suitable technique. For example, a detectable marker may be fused to the nucleic acid-guided nuclease, such that location within a cell may be visualized, such as in combination with a means for detecting the location of the nucleus (e.g. a stain specific for the nucleus such as DAPI). Cell nuclei may also be isolated from cells, the contents of which may then be analyzed by any suitable process for detecting protein, such as immunohistochemistry, Western blot, or enzyme activity assay. Accumulation in the nucleus may also be determined indirectly, such as by an assay for the effect of complex formation (e.g. assay for DNA cleavage or mutation at the target sequence, or assay for altered gene expression activity affected by complex formation and/or nucleic acid-guided nuclease activity), as compared to a control not exposed to the nucleic acid-guided nuclease or complex, or exposed to a nucleic acid-guided nuclease lacking the one or more

NLSs. In an embodiment of the herein described nucleic acid-guided nuclease protein complexes and systems the codon optimized nucleic acid-guided nuclease proteins comprise an NLS attached to the C-terminal of the protein. In an embodiment, other localization tags may be fused to the nucleic acid-guided nuclease, such as without limitation for localizing the nucleic acid-guided nuclease to particular sites in a cell, such as organelles, such as mitochondria, plastids, chloroplast, vesicles, golgi, (nuclear or cellular) membranes, ribosomes, nucleolus, ER, cytoskeleton, vacuoles, centrosome, nucleosome, granules, centrioles, etc.

[0385] In general, the one or more NLSs are of sufficient strength to drive accumulation of the IscB polypeptide nuclease in a detectable amount in the nucleus of a eukaryotic cell. In general, strength of nuclear localization activity may derive from the number of NLSs in the IscB polypeptide nuclease, the particular NLS(s) used, or a combination of these factors. Detection of accumulation in the nucleus may be performed by any suitable technique. For example, a detectable marker may be fused to the IscB polypeptide nuclease, such that location within a cell may be visualized, such as in combination with a means for detecting the location of the nucleus (e.g. a stain specific for the nucleus such as DAPI). Cell nuclei may also be isolated from cells, the contents of which may then be analyzed by any suitable process for detecting protein, such as immunohistochemistry, Western blot, or enzyme activity assay. Accumulation in the nucleus may also be determined indirectly, such as by an assay for the effect of complex formation (e.g. assay for DNA cleavage or mutation at the target sequence, or assay for altered gene expression activity affected by complex formation and/or IscB polypeptide nuclease activity), as compared to a control not exposed to the IscB polypeptide nuclease or complex, or exposed to a IscB polypeptide nuclease lacking the one or more NLSs. In an embodiment of the herein described IscB polypeptide nuclease protein complexes and systems the codon optimized IscB polypeptide nuclease proteins comprise an NLS attached to the C-terminal of the protein. In an embodiment, other localization tags may be fused to the IscB polypeptide nuclease, such as without limitation for localizing the IscB polypeptide nuclease to particular sites in a cell, such as organelles, such as mitochondria, plastids, chloroplast, vesicles, golgi, (nuclear or cellular) membranes, ribosomes, nucleolus, ER, cytoskeleton, vacuoles, centrosome, nucleosome, granules, centrioles, etc.

[0386] In an embodiment of the invention, at least one nuclear localization signal (NLS) is attached to the nucleic acid sequences encoding the IscB polypeptide nuclease. In preferred

embodiments at least one or more C-terminal or N-terminal NLSs are attached (and hence nucleic acid molecule(s) coding for the IscB polypeptide nuclease can include coding for NLS(s) so that the expressed product has the NLS(s) attached or connected). In a preferred embodiment a C-terminal NLS is attached for optimal expression and nuclear targeting in eukaryotic cells, preferably human cells. The invention also encompasses methods for delivering multiple nucleic acid components, wherein each nucleic acid component is specific for a different target locus of interest thereby modifying multiple target loci of interest. The nucleic acid component of the complex may comprise one or more protein-binding RNA aptamers. The one or more aptamers may be capable of binding a bacteriophage coat protein.

Linkers

[0387] In an embodiment of the invention, at least one nuclear localization signal (NLS) is attached to the nucleic acid sequences encoding the nucleic acid-guided nuclease or the IscB polypeptide nuclease. In preferred embodiments at least one or more C-terminal or N-terminal NLSs are attached (and hence nucleic acid molecule(s) coding for the nucleic acid-guided nuclease or IscB polypeptide nuclease can include coding for NLS(s) so that the expressed product has the NLS(s) attached or connected). In a preferred embodiment a C-terminal NLS is attached for optimal expression and nuclear targeting in eukaryotic cells, preferably human cells. The invention also encompasses methods for delivering multiple nucleic acid components, wherein each nucleic acid component is specific for a different target locus of interest thereby modifying multiple target loci of interest. The nucleic acid component of the complex may comprise one or more protein-binding RNA aptamers. The one or more aptamers may be capable of binding a bacteriophage coat protein.

[0388] In some preferred embodiments, the functional domain is linked to a nucleic acid-guided nuclease (e.g., an active or a dead nucleic acid-guided nuclease) to target and activate epigenomic sequences such as promoters or enhancers. One or more guides directed to such promoters or enhancers may also be provided to direct the binding of the nucleic acid-guided nuclease to such promoters or enhancers.

[0389] In some preferred embodiments, the functional domain is linked to a IscB polypeptide nuclease (e.g., an active or a dead IscB polypeptide nuclease) to target and activate epigenomic sequences such as promoters or enhancers. One or more guides directed to such promoters or enhancers may also be provided to direct the binding of the IscB polypeptide nuclease to such promoters or enhancers.

[0390] The term “associated with” is used here in relation to the association of the functional domain to the IscB polypeptide nuclease protein, nucleic acid-guided nuclease, or the adaptor protein. It is used in respect of how one molecule ‘associates’ with respect to another, for example between an adaptor protein and a functional domain, between the IscB polypeptide nuclease protein and a functional domain, or between the nucleic acid guided nuclease protein and a functional domain. In the case of such protein-protein interactions, this association may be viewed in terms of recognition in the way an antibody recognizes an epitope. Alternatively, one protein may be associated with another protein via a fusion of the two, for instance one subunit being fused to another subunit. Fusion typically occurs by addition of the amino acid sequence of one to that of the other, for instance via splicing together of the nucleotide sequences that encode each protein or subunit. Alternatively, this may essentially be viewed as binding between two molecules or direct linkage, such as a fusion protein. In any event, the fusion protein may include a linker between the two subunits of interest (i.e. between the enzyme and the functional domain or between the adaptor protein and the functional domain). Thus, in one embodiment, the IscB polypeptide nuclease protein, nucleic acid-guided nuclease, or adaptor protein is associated with a functional domain by binding thereto. In other embodiments, the IscB polypeptide nuclease, nucleic acid-guided nuclease, or adaptor protein is associated with a functional domain because the two are fused together, optionally via an intermediate linker.

[0391] The term “linker” as used in reference to a fusion protein refers to a molecule which joins the proteins to form a fusion protein. Generally, such molecules have no specific biological activity other than to join or to preserve some minimum distance or other spatial relationship between the proteins. However, in an embodiment, the linker may be selected to influence some property of the linker and/or the fusion protein such as the folding, net charge, or hydrophobicity of the linker.

[0392] Suitable linkers for use in the methods of the present invention are well known to those of skill in the art and include, but are not limited to, straight or branched-chain carbon linkers, heterocyclic carbon linkers, or peptide linkers. However, as used herein the linker may also be a covalent bond (carbon-carbon bond or carbon-heteroatom bond).

[0393] In one embodiment, the linker is used to separate the IscB polypeptide nuclease and the nucleotide deaminase by a distance sufficient to ensure that each protein retains its required functional property. In one embodiment, the linker is used to separate the nucleic acid-guided

nuclease and the nucleotide deaminase by a distance sufficient to ensure that each protein retains its required functional property.

[0394] Preferred peptide linker sequences adopt a flexible extended conformation and do not exhibit a propensity for developing an ordered secondary structure. In an embodiment, the linker can be a chemical moiety which can be monomeric, dimeric, multimeric or polymeric. Preferably, the linker comprises amino acids. Typical amino acids in flexible linkers include Gly, Asn and Ser. Accordingly, in one embodiment, the linker comprises a combination of one or more of Gly, Asn and Ser amino acids. Other near neutral amino acids, such as Thr and Ala, also may be used in the linker sequence. Exemplary linkers are disclosed in Maratea et al. (1985), Gene 40: 39-46; Murphy et al. (1986) Proc. Nat'l. Acad. Sci. USA 83: 8258-62; U.S. Pat. No. 4,935,233; and U.S. Pat. No. 4,751,180. For example, GlySer linkers GGS, GGGS (SEQ ID NO: 2018) or GSG can be used. GGS, GSG, GGGS (SEQ ID NO: 2018) or GGGGS (SEQ ID NO: 2019) linkers can be used in repeats of 3 (such as (GGS)₃, (SEQ ID NO: 2020) (GGGGS)₃ (SEQ ID NO: 2021) or 5, 6, 7, 9 or even 12 or more, to provide suitable lengths. In some cases, the linker may be (GGGGS)₃₋₁₅. For example, in some cases, the linker may be (GGGGS)₃₋₁₁, e.g., GGGGS (SEQ ID NO: 2022), (GGGGS)₂ (SEQ ID NO: 2023), (GGGGS)₃ (SEQ ID NO: 2021), (GGGGS)₄ (SEQ ID NO: 2024), (GGGGS)₅ (SEQ ID NO: 2025), (GGGGS)₆ (SEQ ID NO: 2026), (GGGGS)₇ (SEQ ID NO: 2027), (GGGGS)₈ (SEQ ID NO: 2028), (GGGGS)₉ (SEQ ID NO: 2029), (GGGGS)₁₀ (SEQ ID NO: 2030), or (GGGGS)₁₁ (SEQ ID NO: 2031).

[0395] In one embodiment, linkers such as (GGGGS)₃ (SEQ ID NO: 2021) are preferably used herein. (GGGGS)₆ (SEQ ID NO: 2026), (GGGGS)₉ (SEQ ID NO: 2029) or (GGGGS)₁₂ (SEQ ID NO: 2032) may preferably be used as alternatives. Other preferred alternatives are (GGGGS)₁ (SEQ ID NO: 2022), (GGGGS)₂ (SEQ ID NO: 2023), (GGGGS)₄ (SEQ ID NO: 2024), (GGGGS)₅ (SEQ ID NO: 2025), (GGGGS)₇ (SEQ ID NO: 2027), (GGGGS)₈ (SEQ ID NO: 2028), (GGGGS)₁₀ (SEQ ID NO: 2030), or (GGGGS)₁₁ (SEQ ID NO: 2031). In yet a further embodiment, LEPGEKPYKCPECGKSFSQSGALTRHQRTHTR (SEQ ID NO: 2033) is used as a linker. In yet an additional embodiment, the linker is an XTEN linker. In one embodiment, the IscB polypeptide nuclease or the nucleic acid-guided nuclease is linked to the deaminase protein or its catalytic domain by means of an LEPGEKPYKCPECGKSFSQSGALTRHQRTHTR (SEQ ID NO: 2033) linker. In further one embodiment, IscB polypeptide nuclease is linked C-terminally to the N-terminus of a

deaminase protein or its catalytic domain by means of an LEPGEKPYKCPECCKSFSQSGALTRHQRTHTR (SEQ ID NO: 2033) linker. In addition, N- and C-terminal NLSs can also function as linker (e.g., PKKKRKVEASSPKKRKVEAS (SEQ ID NO: 2034)).

Table 4. Examples of linkers used in the invention are shown.

GGs	GGTGGTAGT
GGs _{x3} (9) (SEQ ID NO: 2020)	GGTGGTAGTGGAGGGAGCGGCGGTTCA (SEQ ID NO: 2036)
GGs _{x7} (21) (SEQ ID NO: 2035)	ggtggaggaggctctggtggaggcggtagcggaggcgagggtcgGGTGGTAGTGGAGGGAGCGGCGGTTCA (SEQ ID NO: 2037)
XTEN	TCGGGATCTGAGACGCCTGGGACCTCGGAATCGGCTACGCCCCGAAAGT (SEQ ID NO: 2038)
Z-EGFR_Short	Gtgataacaaatttaacaaagaaatgtggggcggtgggaagaaattcgtaacctgccgaacctgaacggctggcagatgaccgcgtttattgcgagcctggtggatgatccgagccagagcggaacctgctggcggaagcgaaaaaactgaacgatgcgcaggcgccgaaaaccggcggtggttctggt (SEQ ID NO: 2039)
GSAT	Ggtggttctgccggtggctccggttctggctccagcggtggcagctctggtgcgtccggcacgggtactgcgggtggcactggcagcggttccggtactggctctggc (SEQ ID NO: 2040)

[0396] Linkers may be used between the hRNA molecules and the functional domain (activator or repressor), or between the IscB polypeptide nuclease and the functional domain. In an embodiment, linkers may be used between the guide molecules and the functional domain (e.g. activator or repressor), or between the Cas IscB polypeptide nuclease and the functional domain. The linkers may be used to engineer appropriate amounts of “mechanical flexibility”.

[0397] In an embodiment, the one or more functional domains are controllable, e.g., inducible.

BASE EDITING

[0398] The present disclosure also provides for base editing systems. In general, such a system may comprise a deaminase (e.g., an adenosine deaminase or cytidine deaminase) associated (e.g., fused) with a IscB polypeptide nuclease, e.g., IscB protein. The IscB polypeptide nuclease may be a dead IscB polypeptide nuclease (such as a IscB polypeptide nickase, e.g., engineered from a IscB polypeptide nuclease). In certain examples, the nucleotide deaminase is a mutated form of an adenosine deaminase. The mutated form of the adenosine deaminase may have both adenosine deaminase and cytidine deaminase activities.

[0399] In some examples, the present disclosure provides an engineered, non-naturally occurring composition comprising: the nuclei acid-guided nuclease that is catalytically inactive, a nucleotide deaminase associated with or otherwise capable of forming a complex with the IscB protein, and a single hRNA molecule or single guide RNA molecule capable of forming a complex with the IscB protein and directing site-specific binding at a target sequence.

[0400] In one aspect, the present disclosure provides an engineered adenosine deaminase. The engineered adenosine deaminase may comprise one or more mutations herein. In one embodiment, the engineered adenosine deaminase has cytidine deaminase activity. In certain examples, the engineered adenosine deaminase has both cytidine deaminase activity and adenosine deaminase. In some cases, the modifications by base editors herein may be used for targeting post-translational signaling or catalysis. In one embodiment, compositions herein comprise nucleotide sequence comprising encoding sequences for one or more components of a base editing system. A base-editing system may comprise a deaminase (e.g., an adenosine deaminase or cytidine deaminase) fused with a IscB polypeptide nuclease or a variant thereof. In some cases, the target polynucleotide is edited at one or more bases to introduce a G→A or C→T mutation.

[0401] In some cases, the adenosine deaminase is double-stranded RNA-specific adenosine deaminase (ADAR). Examples of ADARs include those described Yiannis A Savva et al., The ADAR protein family, *Genome Biol.* 2012; 13(12): 252, which is incorporated by reference in its entirety. In some examples, the ADAR may be hADAR1. In certain examples, the ADAR may be hADAR2. The sequence of hADAR2 may be that described under Accession No. AF525422.1.

[0402] In some cases, the deaminase may be a deaminase domain, e.g., a deaminase domain of ADAR (“ADAR-D”). In one example, the deaminase may be the deaminase domain of hADAR2 (“hADAR2-D”), e.g., as described in Phelps KJ et al., Recognition of duplex RNA by the deaminase domain of the RNA editing enzyme ADAR2. *Nucleic Acids Res.* 2015 Jan;43(2):1123-32, which is incorporated by reference herein in its entirety. In a particular example, the hADAR2-D has a sequence comprising amino acid 299-701 of hADAR2-D, e.g., amino acid 299-701 of the sequence under Accession No. AF525422.1.

[0403] In certain examples, the system comprises a mutated form of an adenosine deaminase fused with a dead IscB polypeptide nuclease (e.g., a IscB polypeptide nickase). The

mutated form of the adenosine deaminase may have both adenosine deaminase and cytidine deaminase activities. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: E488Q based on amino acid sequence positions of hADAR2-D, and mutations in a homologous ADAR protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: E488Q, V351G, based on amino acid sequence positions of hADAR2-D, and mutations in a homologous ADAR protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: E488Q, V351G, S486A, based on amino acid sequence positions of hADAR2-D, and mutations in a homologous ADAR protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: E488Q, V351G, S486A, T375S, based on amino acid sequence positions of hADAR2-D, and mutations in a homologous ADAR protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: E488Q, V351G, S486A, T375S, S370C, based on amino acid sequence positions of hADAR2-D, and mutations in a homologous ADAR protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: E488Q, V351G, S486A, T375S, S370C, P462A, based on amino acid sequence positions of hADAR2-D, and mutations in a homologous ADAR protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: E488Q, V351G, S486A, T375S, S370C, P462A, N597I, based on amino acid sequence positions of hADAR2-D, and mutations in a homologous ADAR protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: E488Q, V351G, S486A, T375S, S370C, P462A, N597I, L332I, based on amino acid sequence positions of hADAR2-D, and mutations in a homologous ADAR protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: E488Q, V351G, S486A, T375S, S370C, P462A, N597I, L332I, I398V, based on amino acid sequence positions of hADAR2-D, and mutations in a homologous ADAR protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: E488Q, V351G, S486A, T375S, S370C, P462A, N597I, L332I, I398V, K350I, based on amino acid sequence positions of hADAR2-D, and mutations in a homologous ADAR protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: E488Q, V351G, S486A, T375S,

S370C, P462A, N597I, L332I, I398V, K350I, M383L, based on amino acid sequence positions of hADAR2-D, and mutations in a homologous ADAR protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: E488Q, V351G, S486A, T375S, S370C, P462A, N597I, L332I, I398V, K350I, M383L, D619G, based on amino acid sequence positions of hADAR2-D, and mutations in a homologous ADAR protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: E488Q, V351G, S486A, T375S, S370C, P462A, N597I, L332I, I398V, K350I, M383L, D619G, S582T, based on amino acid sequence positions of hADAR2-D, and mutations in a homologous ADAR protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: E488Q, V351G, S486A, T375S, S370C, P462A, N597I, L332I, I398V, K350I, M383L, D619G, S582T, V440I based on amino acid sequence positions of hADAR2-D, and mutations in a homologous ADAR protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: E488Q, V351G, S486A, T375S, S370C, P462A, N597I, L332I, I398V, K350I, M383L, D619G, S582T, V440I, S495N based on amino acid sequence positions of hADAR2-D, and mutations in a homologous ADAR protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: E488Q, V351G, S486A, T375S, S370C, P462A, N597I, L332I, I398V, K350I, M383L, D619G, S582T, V440I, S495N, K418E based on amino acid sequence positions of hADAR2-D, and mutations in a homologous ADAR protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: E488Q, V351G, S486A, T375S, S370C, P462A, N597I, L332I, I398V, K350I, M383L, D619G, S582T, V440I, S495N, K418E, S661T based on amino acid sequence positions of hADAR2-D, and mutations in a homologous ADAR protein corresponding to the above. In some examples, provided herein includes a mutated adenosine deaminase e.g., an adenosine deaminase comprising one or more mutations of E488Q, V351G, S486A, T375S, S370C, P462A, N597I, L332I, I398V, K350I, M383L, D619G, S582T, V440I, S495N, K418E, S661T, fused with a dead IscB polypeptide nuclease or IscB polypeptide nickase. In some examples, provided herein includes a mutated adenosine deaminase e.g., an adenosine deaminase comprising E488Q, V351G, S486A, T375S, S370C, P462A, N597I, L332I, I398V, K350I, M383L, D619G, S582T, V440I, S495N, K418E, and S661T, fused with a dead IscB polypeptide nuclease or IscB polypeptide nickase. In some

examples, provided herein includes a mutated adenosine deaminase e.g., an adenosine deaminase comprising E488Q, V351G, S486A, T375S, S370C, P462A, N597I, L332I, I398V, K350I, M383L, D619G, S582T, V440I, S495N, K418E, S661T, and S375N fused with a dead IscB polypeptide nuclease or IscB polypeptide nickase.

[0404] In one embodiment, the adenosine deaminase may be a tRNA-specific adenosine deaminase or a variant thereof. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: W23L, W23R, R26G, H36L, N37S, P48S, P48T, P48A, I49V, R51L, N72D, L84F, S97C, A106V, D108N, H123Y, G125A, A142N, S146C, D147Y, R152H, R152P, E155V, I156F, K157N, K161T, based on amino acid sequence positions of *E. coli* TadA, and mutations in a homologous deaminase protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: D108N based on amino acid sequence positions of *E. coli* TadA, and mutations in a homologous deaminase protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: A106V, D108N, based on amino acid sequence positions of *E. coli* TadA, and mutations in a homologous deaminase protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: A106V, D108N, D147Y, E155V, based on amino acid sequence positions of *E. coli* TadA, and mutations in a homologous deaminase protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: A106V, D108N, based on amino acid sequence positions of *E. coli* TadA, and mutations in a homologous deaminase protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: A106V, D108N, D147Y, E155V, L84F, H123Y, I156F, based on amino acid sequence positions of *E. coli* TadA, and mutations in a homologous deaminase protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: A106V, D108N, D147Y, E155V, L84F, H123Y, I156F, A142N, based on amino acid sequence positions of *E. coli* TadA, and mutations in a homologous deaminase protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: A106V, D108N, D147Y, E155V, L84F, H123Y, I156F, H36L, R51L, S146C, K157N, based on amino acid sequence positions of *E. coli* TadA, and mutations in a homologous deaminase protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: A106V, D108N, D147Y, E155V, L84F, H123Y, I156F, H36L,

R51L, S146C, K157N, P48S, based on amino acid sequence positions of *E. coli* TadA, and mutations in a homologous deaminase protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: A106V, D108N, D147Y, E155V, L84F, H123Y, I156F, H36L, R51L, S146C, K157N, P48S, A142N, based on amino acid sequence positions of *E. coli* TadA, and mutations in a homologous deaminase protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: A106V, D108N, D147Y, E155V, L84F, H123Y, I156F, H36L, R51L, S146C, K157N, P48S, W23R, P48A, based on amino acid sequence positions of *E. coli* TadA, and mutations in a homologous deaminase protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: A106V, D108N, D147Y, E155V, L84F, H123Y, I156F, H36L, R51L, S146C, K157N, P48S, W23R, P48A, A142N, based on amino acid sequence positions of *E. coli* TadA, and mutations in a homologous deaminase protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: A106V, D108N, D147Y, E155V, L84F, H123Y, I156F, H36L, R51L, S146C, K157N, P48S, W23R, P48A, R152P, based on amino acid sequence positions of *E. coli* TadA, and mutations in a homologous deaminase protein corresponding to the above. In one embodiment, the adenosine deaminase may comprise one or more of the mutations: A106V, D108N, D147Y, E155V, L84F, H123Y, I156F, H36L, R51L, S146C, K157N, P48S, W23R, P48A, R152P, A142N, based on amino acid sequence positions of *E. coli* TadA, and mutations in a homologous deaminase protein corresponding to the above.

[0405] In some examples, the base editing systems may comprise an intein-mediated trans-splicing system that enables *in vivo* delivery of a base editor, e.g., a split-intein cytidine base editors (CBE) or adenine base editor (ABE) engineered to trans-splice. Examples of the such base editing systems include those described in Colin K.W. Lim et al., Treatment of a Mouse Model of ALS by *In Vivo* Base Editing, *Mol Ther.* 2020 Jan 14. pii: S1525-0016(20)30011-3. doi: 10.1016/j.ymthe.2020.01.005; and Jonathan M. Levy et al., Cytosine and adenine base editing of the brain, liver, retina, heart and skeletal muscle of mice via adeno-associated viruses, *Nature Biomedical Engineering* volume 4, pages97–110(2020), which are incorporated by reference herein in their entirety.

[0406] Examples of base editing systems include those described in International Patent Publication Nos. WO 2019/071048 (e.g. paragraphs [0933]-[0938]), WO 2019/084063 (e.g.,

paragraphs [0173]-[0186], [0323]-[0475], [0893]-[1094]), WO 2019/126716 (e.g., paragraphs [0290]-[0425], [1077]-[1084]), WO 2019/126709 (e.g., paragraphs [0294]-[0453]), WO 2019/126762 (e.g., paragraphs [0309]-[0438]), WO 2019/126774 (e.g., paragraphs [0511]-[0670]), Cox DBT, et al., RNA editing with CRISPR-Cas13, *Science*. 2017 Nov 24;358(6366):1019-1027; Abudayyeh OO, et al., A cytosine deaminase for programmable single-base RNA editing, *Science* 26 Jul 2019: Vol. 365, Issue 6451, pp. 382-386; Gaudelli NM et al., Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage, *Nature* volume 551, pages 464–471 (23 November 2017); Komor AC, et al., Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*. 2016 May 19;533(7603):420-4; Jordan L. Doman et al., Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors, *Nat Biotechnol* (2020). doi.org/10.1038/s41587-020-0414-6; and Richter MF et al., Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity, *Nat Biotechnol* (2020). doi.org/10.1038/s41587-020-0453-z, which are incorporated by reference herein in their entireties and can be used to adapt to the IscB or CRISPR-associated IscB polypeptides.

PRIME EDITING

[0407] In one embodiment, the present disclosure provides compositions and systems may comprise a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease or a catalytically inactive form, one or more ω RNA or guide molecules, and a reverse transcriptase. The systems may be used to insert a donor polynucleotide to a target polynucleotide. In some examples, the composition or system comprises a catalytically inactive IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, a reverse transcriptase associated with or otherwise capable of forming a complex with the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, and a ω RNA or guide molecule capable of forming a complex with the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and directing site-specific binding of the complex to a target sequence of a target polynucleotide, the ω RNA or guide molecule further comprising a donor sequence for insertion into the target polynucleotide.

[0408] In some cases, the catalytically inactive IscB polypeptide or CRISPR-associated IscB polypeptide nuclease may be a nickase, e.g., a DNA nickase. In some cases, the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease has one or more mutations. In

some examples, the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease comprises mutations corresponding to the mutations in the RuvC or HNH nuclease.

[0409] The IscB polypeptide or CRISPR-associated IscB polypeptide nuclease may be associated with a reverse transcriptase. A reverse transcriptase domain may be a reverse transcriptase or a fragment thereof. In certain aspects, the reverse transcriptase is Human immunodeficiency virus (HIV) RT, Avian myoblastosis virus (AMV) RT, Moloney murine leukemia virus (M-MLV) RT a group II intron RT, a group II intron-like RT, or a chimeric RT. In an embodiment, the RT comprises modified forms of these RTs, such as, engineered variants of Avian myoblastosis virus (AMV) RT, Moloney murine leukemia virus (M-MLV) RT, or Human immunodeficiency virus (HIV) RT (see, e.g., Anzalone, et al., Search-and-replace genome editing without double-strand breaks or donor DNA, *Nature*. 2019 Dec;576(7785):149-157).

[0410] In some examples, the compositions and systems may comprise the IscB or CRISPR-associated protein disclosed herein; a reverse transcriptase (RT) polypeptide connected to or otherwise capable of forming a complex with the IscB polypeptide or CRISPR-associated IscB polypeptide; and a ω RNA or guide molecule capable of forming a complex with the IscB polypeptide or CRISPR-associated IscB polypeptide and comprising: a ω RNA or guide sequence capable of directing site-specific binding of the IscB polypeptide or CRISPR-associated IscB polypeptide complex to a target sequence of a target polynucleotide; a 3' binding site region capable of binding to a cleaved upstream strand of the target polynucleotide; and a RT template sequence encoding an extended sequence, wherein the extended sequence comprises a variant region and a 3' homologous sequence capable of hybridization to the downstream cleaved strand of the target polynucleotide.

[0411] A reverse transcriptase domain may be a reverse transcriptase or a fragment thereof. A wide variety of reverse transcriptases (RT) may be used in alternative embodiments of the present invention, including prokaryotic and eukaryotic RT, provided that the RT functions within the host to generate a donor polynucleotide sequence from the RNA template. If desired, the nucleotide sequence of a native RT may be modified, for example using known codon optimization techniques, so that expression within the desired host is optimized. A reverse transcriptase (RT) is an enzyme used to generate complementary DNA (cDNA) from an RNA template, a process termed reverse transcription. Reverse transcriptases are used by

retroviruses to replicate their genomes, by retrotransposon mobile genetic elements to proliferate within the host genome, by eukaryotic cells to extend the telomeres at the ends of their linear chromosomes, and by some non-retroviruses such as the hepatitis B virus, a member of the Hepadnaviridae, which are dsDNA-RT viruses. Retroviral RT has three sequential biochemical activities: RNA-dependent DNA polymerase activity, ribonuclease H, and DNA-dependent DNA polymerase activity. Collectively, these activities enable the enzyme to convert single-stranded RNA into double-stranded cDNA. In an embodiment, the RT domain of a reverse transcriptase is used in the present invention. The domain may include only the RNA-dependent DNA polymerase activity. In some examples, the RT domain is non-mutagenic, i.e., does not cause mutation in the donor polynucleotide (e.g., during the reverse transcriptase process). In some cases, In some examples, the RT domain may be non-retron RT, e.g., a viral RT or a human endogenous RTs. In some examples, the RT domain may be retron RT or DGRs RT. In some example, the RT may be less mutagenic than a counterpart wildtype RT. In one embodiment, the RT herein is not mutagenic.

[0412] The reverse transcriptase may be fused to the C-terminus of a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease. Alternatively or additionally, the reverse transcriptase may be fused to the N-terminus of a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease. The fusion may be via a linker and/or an adaptor protein. In some examples, the reverse transcriptase may be an M-MLV reverse transcriptase or variant thereof. The M-MLV reverse transcriptase variant may comprise one or more mutations. For the examples, the M-MLV reverse transcriptase may comprise D200N, L603W, and T330P. In another example, the M-MLV reverse transcriptase may comprise D200N, L603W, T330P, T306K, and W313F. In a particular example, the fusion of IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and reverse transcriptase is IscB polypeptide or CRISPR-associated IscB polypeptide nuclease (with a mutation corresponding to H840A of SpCas9) fused with M-MLV reverse transcriptase (D200N+L603W+T330P+T306K+W313F).

[0413] In one embodiment, the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease herein may target DNA using a ω RNA or guide RNA containing a binding sequence that hybridizes to the target sequence on the DNA. The ω RNA or guide RNA may further comprise an editing sequence that contains new genetic information that replaces target DNA nucleotides. The small sizes of the IscB polypeptide or CRISPR-associated IscB polypeptide

nuclease herein may allow easier packaging and delivery of the prime editing system, e.g., with a viral vector, e.g., AAV or lentiviral vector.

[0414] A single-strand break (a nick) may be generated on the target DNA by the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease at the target site to expose a 3'-hydroxyl group, thus priming the reverse transcription of an edit-encoding extension on the ω RNA or guide directly into the target site. These steps may result in a branched intermediate with two redundant single-stranded DNA flaps: a 5' flap that contains the unedited DNA sequence, and a 3' flap that contains the edited sequence copied from the hRNA. The 5' flaps may be removed by a structure-specific endonuclease, e.g., FEN122, which excises 5' flaps generated during lagging-strand DNA synthesis and long-patch base excision repair. The non-edited DNA strand may be nicked to induce bias DNA repair to preferentially replace the non-edited strand. Examples of prime editing systems and methods include those described in Anzalone AV *et al.*, Search-and-replace genome editing without double-strand breaks or donor DNA, Nature. 2019 Oct 21. doi: 10.1038/s41586-019-1711-4, which is incorporated by reference herein in its entirety.

[0415] The IscB polypeptide or CRISPR-associated IscB polypeptide (e.g., the nickase form) may be used to prime-edit a single nucleotide on a target DNA. Alternatively or additionally, the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease may be used to prime-edit at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 200, at least 300, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, or at least 1000 nucleotides on a target DNA.

[0416] In yet another embodiment, PRIME editing is used first to create a longer 3' region (e.g. 20 nucleotides). Examples of prime editing systems and methods include those described in Anzalone AV *et al.*, Search-and-replace genome editing without double-strand breaks or donor DNA, Nature. 2019 Oct 21. doi: 10.1038/s41586-019-1711-4, which is incorporated by reference herein in its entirety. In such cases, the system comprises a IscB polypeptide or CRISPR-associated IscB polypeptide with nickase activity, a reverse transcriptase domain, and a DNA polymerase, and a ω RNA or guide molecule comprising a binding sequence capable of

hybridizing to the target polynucleotide and a editing sequence. The generated region may be further extended on a DNA template as described herein. The latter may allow generation of a target-independent sequence, compatible with a generic donor sequence.

[0417] The IscB polypeptide or CRISPR-associated IscB polypeptide is capable of generating a first cleavage of in the target sequence and a second cleavage outside the target sequence on the target polynucleotide. In some variations, a second IscB polypeptide or CRISPR-associated IscB polypeptide -mediated cleavage in vicinity to the target site may be made, which may enable more efficient invasion of the extended DNA.

[0418] In some examples, the compositions and systems of the IscB polypeptide or CRISPR-associated IscB polypeptide herein comprise: a reverse transcriptase (RT) polypeptide connected to or otherwise capable of forming a complex with the IscB polypeptide or CRISPR-associated IscB polypeptide; a first ω RNA or guide molecule capable of forming a first IscB polypeptide or CRISPR-associated IscB polypeptide-Reverse transcriptase complex with the IscB polypeptide or CRISPR-associated IscB polypeptide and comprising: a ω RNA or guide sequence capable of directing site-specific binding of the first IscB polypeptide or CRISPR-associated IscB polypeptide-Reverse transcriptase complex to a first target sequence of a target polynucleotide; a first binding site region capable of binding to a cleaved or nicked strand of the target polynucleotide; and a RT template sequence encoding a first extended sequence; a second ω RNA or guide molecule capable of forming a second IscB polypeptide or CRISPR-associated IscB polypeptide-Reverse transcriptase complex with the IscB polypeptide or CRISPR-associated IscB polypeptide and comprising: a ω RNA or guide sequence capable of directing site specific binding of the second IscB polypeptide or CRISPR-associated IscB polypeptide-Reverse transcriptase complex to a second target sequence of the target polynucleotide; a second binding site region capable of binding to a cleaved or nicked strand of the target polynucleotide; and a RT template sequence encoding a second extended sequence.

[0419] In some cases, the compositions and systems may further comprise: a donor template; a third ω RNA or guide sequence capable of forming a IscB polypeptide or CRISPR-associated IscB polypeptide-Reverse transcriptase complex- ω RNA or guide with the IscB polypeptide or CRISPR-associated IscB polypeptide and comprising: a ω RNA or guide sequence capable of directing site-specific binding to a target sequence on the donor template;

a third binding region capable of binding to a cleaved or nicked strand of the donor template; and a RT template encoding a third extended region complementary to the first extended region generated on the target polynucleotide; and a fourth ω RNA or guide sequence capable of forming a IscB polypeptide or CRISPR-associated IscB polypeptide-Reverse transcriptase complex with the IscB polypeptide or CRISPR-associated IscB polypeptide and comprising: a ω RNA or guide sequence capable of directing site-specific binding to a second target sequence on the donor template; a fourth binding region capable of binding to a cleaved or nicked strand of the donor template; and a RT template encoding a fourth extended region complementary to the second extended region generated on the target polynucleotide.

[0420] In some cases, the compositions and systems may further comprise a site-specific recombinase, and wherein the first and second extended regions are complementary to each other and introduce a serine integrase recombination site; and a donor molecule comprising a donor sequence for insertion into the target polypeptide and the complementary recombination site to the serine integrase recombination site.

[0421] In some examples, the compositions and systems may further comprise a recombinase. The recombinase is connected to or otherwise capable of forming a complex with the IscB polypeptide or CRISPR-associated IscB polypeptide. In an embodiment, the complex is capable of inserting a recombination site in the DNA loci of interest by extension of RT templates that encode for the recombination site on the 3' extension of the ω RNA or guide sequences by the reverse transcriptase. In an embodiment, a donor template comprising a compatible recombination site is provided that can recombine unidirectionally with the inserted recombination site when a recombinase specific for the recombination site is also provided. In an embodiment, the donor template is a plasmid comprising the complementary recombination site and any sequence for insertion at the DNA loci of interest. In an embodiment, the recombinase is connected to or capable of forming a complex with the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, such that all of the enzymatic proteins are brought into contact at the loci of interest. In an embodiment, the recombinase is codon optimized for eukaryotic cells (described further herein). In an embodiment, the recombinase includes a NLS (described further herein). In an embodiment, the recombinase is provided as a separate protein. The separate recombinase may form a dimer and bind to the donor template recombination site. The recombinase may be targeted to the loci of interest as a result of the

insertion of the compatible recombination site that is also recognized by the recombinase. Thus, the recombinase may recognize the recombination site inserted at the DNA loci of interest and the recombination site on the donor and be targeted to the DNA loci of interest without any additional modifications to the recombinase.

[0422] In an embodiment, a second IscB complex connected to a recombinase is targeted to the DNA loci of interest. In an embodiment, the second TnpB complex comprises a dead IscB protein (dIscB, described further herein), such that the recombinase is targeted to the DNA loci of interest, but the target sequence is not further cleaved. In an embodiment, the dIscB targets a sequence generated only after the insertion of the recombination site. In an embodiment, the recombinase recognizes and binds to the donor template recombination site and the inserted recombination site. In an embodiment, the recombinase forms a dimer with a recombinase provided as a separate protein.

[0423] As used herein, the term “Recombinase” refers to an enzyme that catalyzes recombination between two or more recombination sites (e.g., an acceptor and donor site). Recombinases useful in the present invention catalyze recombination at specific recombination sites which are specific polynucleotide sequences that are recognized by a particular recombinase. “Uni-directional recombinases” or “integrases” refer to recombinase enzymes whose recognition sites are destroyed after the recombination has taken place. The term “integrase” refers to a type of recombinase. In other words, the sequence recognized by the recombinase is changed into one that is not recognized by the recombinase upon recombination. As a result, once a sequence is subjected to recombination by the uni-directional recombinase, the continued presence of the recombinase cannot reverse the previous recombination event.

[0424] “Recombination sites” are specific polynucleotide sequences that are recognized by the recombinase enzymes described herein. Typically, two different sites are involved (in regards to recombination termed “complementary sites”), one present in the target nucleic acid (e.g., a chromosome or episome of a eukaryote) and another on the nucleic acid that is to be integrated at the target recombination site. The terms “attB” and “attP,” which refer to attachment (or recombination) sites originally from a bacterial target (attachment site of bacteria) and a phage donor (attachment site of phage), respectively, are used herein although recombination sites for particular enzymes may have different names. The two attachment sites can share as little sequence identity as a few base pairs. The recombination sites typically

include left and right arms separated by a core or spacer region. Thus, an attB recombination site consists of BOB', where B and B' are the left and right arms, respectively, and O is the core region. Similarly, attP is POP', where P and P' are the arms and O is again the core region. Upon recombination between the attB and attP sites, and concomitant integration of a nucleic acid at the target, the recombination sites that flank the integrated DNA are referred to as "attL" and "aatR." The attL and attR sites, using the terminology above, thus consist of BOP' and POB', respectively. In some representations herein, the "O" is omitted and attB and attP, for example, are designated as BB' and PP', respectively.

GUIDED EXCISION-TRANSPOSITION SYSTEMS

[0425] Embodiments disclosed herein provide an engineered or non-natural guided excision-transposition system. The engineered or non-natural guided excision-transposition system may comprise one or more components of a ω RNA-IscB or guide-CRISPR-associated IscB system and one or more components of a Class II transposon. The components of the ω RNA-IscB or guide-CRISPR-associated IscB system can direct the Class II transposon component(s) to retrotranspose to a target nucleic acid sequence and direct its transposition into a recipient polynucleotide.

[0426] For example, the engineered or non-natural guided excision-transposition systems that can include (a) a first IscB polypeptide or CRISPR-associated IscB polypeptide; (b) a first Class II transposon polypeptide coupled to or otherwise capable of complexing with the first IscB polypeptide or CRISPR-associated IscB polypeptide; (c) a first guide molecule capable of forming a first ω RNA-IscB or guide-CRISPR-associated IscB complex with the first IscB protein or CRISPR-associated IscB polypeptide and directing site-specific binding to a first target sequence of a first target polynucleotide; (d) a second IscB polypeptide or CRISPR-associated IscB polypeptide; (e) a second Class II transposon polypeptide coupled to or otherwise capable of complexing with the second IscB polypeptide or CRISPR-associated IscB polypeptide; (f) a second guide molecule capable of forming a second ω RNA-IscB complex with the first IscB protein and directing site-specific binding to a second target sequence of the first target polynucleotide; and (g) a Class II transposon polynucleotide comprising the first target polynucleotide and is capable of forming a complex with the first and second IscB polypeptide or CRISPR-associated IscB polypeptide, the first and second guide molecules, and the first and second Class II transposon polypeptides.

[0427] In one embodiment, the engineered or non-natural guided excision-transposition system can include (h) a third guide molecule capable of complexing with the first IscB polypeptide or CRISPR-associated IscB polypeptide and directing site-specific binding to a first target sequence of a second target polynucleotide, wherein the third guide molecule is optionally coupled to the first IscB polypeptide or CRISPR-associated IscB polypeptide; (i) optionally, a first ω RNA or guide molecule polynucleotide that encodes the third ω RNA or guide molecule; (j) a fourth ω RNA or guide molecule capable of complexing with the second IscB polypeptide or CRISPR-associated IscB polypeptide and directing site-specific binding to a second target sequence of the second target polynucleotide, wherein the fourth guide molecule is optionally coupled to the second IscB polypeptide or CRISPR-associated IscB polypeptide; and (k) optionally, a second ω RNA or guide molecule polynucleotide that encodes the fourth ω RNA or guide molecule.

[0428] In one embodiment, the first and the second Class II transposon polypeptides are capable of excising the first target polynucleotide from the Class II transposon polynucleotide. In one embodiment, the first and the second Class II transposon polypeptides are capable of transposing the first target polynucleotide in the second target polynucleotide. In one embodiment, the first target polynucleotide does not include one or more Class II transposon long terminal repeats.

[0054] The engineered or non-natural guided excision-transposition systems described herein can be based on a Class II transposon or Class II transposon system. The engineered or non-natural guided excision-transposition system may include a first target polynucleotide, also referred to as a donor polynucleotide or transposon and a second target polynucleotide, which is also referred to herein as a recipient polynucleotide. As used herein, “transposon” (also referred to as transposable element) refers to a polynucleotide sequence that is capable of moving from location in a genome to another. There are several classes of transposons. Transposons include retrotransposons (Class I transposons) and DNA transposons (Class II transposons). In some cases, retrotransposons require the transcription of the polynucleotide that is moved (or transposed) in order to transpose the polynucleotide to a new genome or polynucleotide. DNA transposons are those that do not require reverse transcription of the polynucleotide that is moved (or transposed) in order to transpose the polynucleotide to a new genome or polynucleotide.

[0429] Any suitable transposon system can be used. Suitable transposon and systems thereof can include, but are not limited, to Sleeping Beauty transposon system (Tc1/mariner superfamily) (see e.g. Ivics et al. 1997. *Cell*. 91(4): 501-510), piggyBac (piggyBac superfamily) (see e.g. Li et al. 2013 110(25): E2279-E2287 and Yusa et al. 2011. *PNAS*. 108(4): 1531-1536), Tol2 (superfamily hAT), Frog Prince (Tc1/mariner superfamily) (see e.g. Miskey et al. 2003 *Nucleic Acid Res.* 31(23):6873-6881) and variants thereof.

[0430] In one embodiment, the first and/or second Class II transposon polypeptide is a DD[E/D] transposon or transposon polypeptide. In one embodiment, the first and/or the second Class II transposon polynucleotide is a Tc1/mariner, PiggyBac, Frog Prince, Tn3, Tn5, hAT, CACTA, P, Mutator, PIF/Harbinger, Transib, or a Merlin/IS1016 transposon polynucleotide. In one embodiment, the first and/or second Class II transposon polypeptide is a Tc1/mariner, PiggyBac, Frog Prince, Tn3, Tn5, hAT, CACTA, P, Mutator, PIF/Harbinger, Transib, or a Merlin/IS1016 transposon polypeptide.

[0058] Suitable Class II transposon systems and components that can be utilized can also be and are not limited to those described in e.g. and without limitation, Han et al., 2013. *BMC Genomics*. 14:71, doi: 10.1186/1471-2164-14-71, Lopez and Garcia-Perez. 2010. *Curr. Genomics*. 11(2):115-128; Wessler. 2006. *PNAS*. 103(47): 176000-17601; Gao et al., 2017. *Marine Genomics*. 34:67-77; Bradic et al. 2014. *Mobile DNA*. 5(12) doi:10.1186/1759-8753-5-12; Li et al., 2013. *PNAS*. 110(25)E2279-E2287; Kebriaei et al. 2017. *Trends in Genetics*. 33(11): 852-870; Miskey et al. 2003. *Nucleic Acid res.* 31(23):6873-6881; Nicolas et al. 2015. *Microbiol Spectr.* 3(4) doi: 10.1128/microbiolspec.MDNA3-0060-2014; W.S. Reznikoff. 1993. *Annu Rev. Microbiol.* 47:945-963; Rubin et al. 2001. *Genetics*. 158(3): 949-957; Wicker et al. 2003. *Plant Physiol.* 132(1): 52-63; Majumdar and Rio. 2015. *Microbiol. Spectr.* 3(2) doi: 10.1128/microbiolspec.MDNA3-0004-2014; D. Lisch. 2002. *Trends in Plant Sci.* 7(11): 498-504; Sinzelle et al. 2007. *PNAS*. 105(12): 4715-4720; Han et al. 2014; *Genome Biol. Evol.* 6(7):1748-1757; Grzebelus et al. 2006; *Mol. Genet. Genomics*. 275(5):450-459; Zhang et al. 2004. *Genetics*. 166(2):971-986; Chen and Li. 2008. *Gene*. 408(1-2):51-63; and C. Feschotte. 2004. *Mol. Biol. Evol.* 21(9):1769-1780.

RETROTRANSPOSONS

[0431] The systems and compositions herein may comprise a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, one or more ω RNAs or guide RNAs, and one or more

components of a retrotransposon, e.g., a non-LTR retrotransposon. The one or more components of a retrotransposon include a retrotransposon protein and retrotransposon RNA. The systems and compositions may be used to insert a donor polynucleotide to a target polynucleotide. The systems and compositions may further comprise a donor polynucleotide.

[0432] In some examples, the present disclosure provides an engineered, non-naturally occurring composition comprising: a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, a non-LTR retrotransposon protein associated with or otherwise capable of forming a complex with the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease; a single ω RNA or guide capable of forming a complex with the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and directing site-specific binding to a target sequence of a target polynucleotide. The composition may further comprise a donor construct comprising a donor polynucleotide for insertion to the target polynucleotide and located between two binding elements capable of forming a complex with the non-LTR retrotransposon protein. In some cases, the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease is engineered to have nickase activity.

[0433] In some examples, the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease is fused to the N-terminus of the non-LTR retrotransposon protein. In some examples, the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease is fused to the C-terminus of the non-LTR retrotransposon protein.

[0434] The guides may direct the fusion protein to a target sequence 5' of the targeted insertion site, and wherein the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease generates a double-strand break at the targeted insertion site. The guides may direct the fusion protein to a target sequence 3' of the targeted insertion site, and wherein the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease generates a double-strand break at the targeted insertion site.

[0435] The donor polynucleotide may further comprise a polymerase processing element to facilitate 3' end processing of the donor polynucleotide sequence. The polymerase may be a DNA polymerase, e.g., DNA polymerase I. In some examples, the polymerase may be an RNA polymerase.

[0436] In some examples, the donor polynucleotide further comprises a homology region to the target sequence on the 5' end of the donor construct, the 3' end of the donor construct,

or both. In some examples, the homology region is from 1 to 50, from 5 to 30, from 8 to 25, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 base pairs in length.

[0437] Native or wild-type non-LTR retrotransposons encode the protein machinery necessary for their self-mobilization. The non-LTR retrotransposon element comprises a DNA element integrated into a host genome. This DNA element may encode one or two open reading frames (ORFs). For example, the R2 element of *Bombyx mori* encodes a single ORF containing reverse transcriptase (RT) activity and a restriction enzyme-like (REL) domain. L1 elements encode two ORFs, ORF1 and ORF2. ORF1 contains a leucine zipper domain involved in protein-protein interactions and a C-terminal nucleic acid binding domain. ORF2 has a N-terminal apurinic/apyrimidinic endonuclease (APE), a central RT domain, and a C-terminal cysteine histidine rich domain. An example replicative cycle of a non-LTR retrotransposon may comprise transcription of the full-length retrotransposon element to generate an mRNA active element (retrotransposon RNA). The active element mRNA is translated to generate the encoded retrotransposon proteins or polypeptides. A ribonucleoprotein complex comprising the active element and retrotransposon protein or polypeptide is formed and this RNP facilitates integration of the active element into the genome. The RNA-transposase complex nicks the genome. The 3' end of the nicked DNA serves as a primer to allow the reverse transcription of the transposon RNA into cDNA. Fourth, the transposase proteins integrate the cDNA into the genome.

[0438] Elements of these systems may be engineered to work within the context of the invention. For example a non-LTR retrotransposon polypeptide may be fused to a site-specific nuclease. The binding elements that allow a non-LTR retrotransposon polypeptide to bind to the native retrotransposon DNA element, may be engineered into a donor construct to facilitate entry of a donor polynucleotide sequence into a target polypeptide.

[0439] In the present invention the protein component of the non-LTR retrotransposon may be connected to or otherwise engineered to form a complex with a site-specific nuclease. The retrotransposon RNA may be engineered to encode a donor polynucleotide sequence. Thus, in certain example embodiments, the IscB polypeptide nuclease, via formation of a IscB polypeptide nuclease complex with a guide sequence, directs the retrotransposon complex (e.g. the retrotransposon polypeptide(s) and retrotransposon RNA) to a target sequence in a target

polynucleotide, where the retrotransposon RNP complex facilitates integration of the donor polynucleotide sequence into the target polynucleotide. Accordingly, the one or more non-LTR retrotransposon components may comprise retrotransposon polypeptides, or function domains thereof, that facilitate binding of the retrotransposon RNA, reverse transcription of the retrotransposon RNA into cDNA, and/or integration of the donor polynucleotide into the target polynucleotide, as well as retrotransposon RNA elements modified to encode the donor polynucleotide sequence.

[0440] Examples of non-LTR retrotransposons include CRE, R2, R4, L1, RTE, Tad, R1, LOA, I, Jockey, CR1. In one example, the non-LTR retrotransposon is R2. In another example, the non-LTR retrotransposon is L1. Examples of non-LTR retrotransposons may include those described in Christensen SM et al., RNA from the 5' end of the R2 retrotransposon controls R2 protein binding to and cleavage of its DNA target site, *Proc Natl Acad Sci U S A.* 2006 Nov 21;103(47):17602-7; Eickbush TH et al, Integration, Regulation, and Long-Term Stability of R2 Retrotransposons, *Microbiol Spectr.* 2015 Apr;3(2):MDNA3-0011-2014. doi: 10.1128/microbiolspec.MDNA3-0011-2014; Han JS, Non-long terminal repeat (non-LTR) retrotransposons: mechanisms, recent developments, and unanswered questions, *Mob DNA.* 2010 May 12;1(1):15. doi: 10.1186/1759-8753-1-15; Malik HS et al., The age and evolution of non-LTR retrotransposable elements, *Mol Biol Evol.* 1999 Jun;16(6):793-805, which are incorporated by reference herein in their entireties.

[0441] Examples of the non-LTR retrotransposon polypeptides also include R2 from *Clonorchis sinensis*, or *Zonotrichia albicollis*.

[0442] A non-LTR retrotransposon may comprise multiple retrotransposon polypeptides or polynucleotides encoding same. In one embodiment, the retrotransposon polypeptides may form a complex. For example, a non-LTR retrotransposon is a dimer, e.g., comprising two retrotransposon polypeptides forming a dimer. The dimer subunits may be connected or form a tandem fusion. A IscB polypeptide nuclease may be associated with (e.g., connected to) one or more subunits of such complex. In some examples, the non-LTR retrotransposon is a dimer of two retrotransposon polypeptides; one of the retrotransposon polypeptides comprises nuclease or nickase activity and is connected with a IscB polypeptide nuclease.

[0443] The retrotransposon polypeptides may comprise one or more modifications to, for example, enhance specificity or efficiency of donor polynucleotide recognition, target-primed template recognition (TPTR). The retrotransposon polypeptides may also comprise one or

more truncations or excisions to remove domains or regions of wild-type protein to arrive at a minimal polypeptide that retain donor polynucleotide recognition and TPTR. In some example embodiments, the native endonuclease activity may be mutated to eliminate endonuclease activity.

[0444] In certain example embodiments, the modifications or truncations of the non-LTR retrotransposon peptide may be in a zinc finger region, a Myb region, a basic region, a reverse transcriptase domain, a cysteine-histidine rich motif, or an endonuclease domain.

[0445] A non-LTR retrotransposon may comprise polynucleotide encoding one or more retrotransposon RNA molecules. The polynucleotide may comprise one or more regulatory elements. The regulatory elements may be promoters. The regulatory elements and promoters on the polynucleotides include those described throughout this application. For example, the polynucleotide may comprise a pol2 promoter, a pol3 promoter, or a T7 promoter.

[0446] In some cases, the polynucleotide encodes a retrotransposon RNA with at least a portion of its sequence complementary to a target sequence. For example, the 3' end of the retrotransposon RNA may be complementary to a target sequence. The RNA may be complementary to a portion of a nicked target sequence. In one embodiment, a retrotransposon RNA may comprise one or more donor polynucleotides. In certain cases, a retrotransposon RNA may encode one or more donor polynucleotides.

[0447] A retrotransposon RNA may be capable of binding to a retrotransposon polypeptide. Such retrotransposon RNA may comprise one or more elements for binding to the retrotransposon polypeptide. Examples of binding elements include hairpin structures, pseudoknots (e.g., a nucleic acid secondary structure containing at least two stem-loop structures in which half of one stem is intercalated between the two halves of another stem), stem loops, and bulges (e.g., unpaired stretches of nucleotides located within one strand of a nucleic acid duplex). In certain examples, the retrotransposon RNA comprises one or more hairpin structures. In some examples, the retrotransposon RNA comprises one or more pseudoknots. In certain examples, a retrotransposon RNA comprises a sequence encoding a donor polynucleotide and one or more binding elements for forming a complex with the retrotransposon polypeptide. The binding elements may be located on the 5' end or the 3' end.

[0448] In one embodiment, a retrotransposon RNA comprises a region capable of hybridizing with an overhang of a target polynucleotide at the target site. The overhang may be a stretch of single-stranded DNA. The overhang may function as a primer for reverse

transcription of at least a portion of the retrotransposon RNA to a cDNA. In some cases, a region of the cDNA may be capable of hybridizing a second overhang of the target polynucleotide. The second overhang may function as a primer for the synthesis of a second strand to generate a double-stranded cDNA. The cDNA may comprise a donor polynucleotide sequence. The two overhangs may be from different strands of the target polynucleotide.

Reverse Transcriptase Domain

[0449] The one or more functional domains may be one or more reverse transcriptase domains. In one embodiment, the systems comprise an engineered system for modifying a target polynucleotide comprising: a IscB polypeptide or CRISPR-associated IscB polypeptide or a variant thereof (e.g., dIscB); a reverse transcriptase (RT) domain; a RNA template comprising or encoding a donor polynucleotide to be inserted to a target sequence of the target polynucleotide; and an ω RNA or guide RNA molecule (i.e., a naturally single guide RNA molecule comprising a scaffold for reprogramming).

[0450] The reverse transcriptase may generate single-strand DNA based on the RNA template. The single-strand DNA may be generated by a non-retron, retron, or diversity generating retroelement (DGR). In some examples, the single-strand DNA may be generated from a self-priming RNA template. A self-priming RNA template may be used to generate a DNA without the need of a separate primer.

[0451] A reverse transcriptase domain may be a reverse transcriptase or a fragment thereof. A wide variety of reverse transcriptases (RT) may be used in alternative embodiments of the present invention, including prokaryotic and eukaryotic RT, provided that the RT functions within the host to generate a donor polynucleotide sequence from the RNA template. If desired, the nucleotide sequence of a native RT may be modified, for example using known codon optimization techniques, so that expression within the desired host is optimized. A reverse transcriptase (RT) is an enzyme used to generate complementary DNA (cDNA) from an RNA template, a process termed reverse transcription. Reverse transcriptases are used by retroviruses to replicate their genomes, by retrotransposon mobile genetic elements to proliferate within the host genome, by eukaryotic cells to extend the telomeres at the ends of their linear chromosomes, and by some non-retroviruses such as the hepatitis B virus, a member of the Hepadnaviridae, which are dsDNA-RT viruses. Retroviral RT has three sequential biochemical activities: RNA-dependent DNA polymerase activity, ribonuclease H, and DNA-

dependent DNA polymerase activity. Collectively, these activities enable the enzyme to convert single-stranded RNA into double-stranded cDNA. In an embodiment, the RT domain of a reverse transcriptase is used in the present invention. The domain may include only the RNA-dependent DNA polymerase activity. In some examples, the RT domain is non-mutagenic, i.e., does not cause mutation in the donor polynucleotide (e.g., during the reverse transcriptase process). In some examples, the RT domain may be non-retron RT, e.g., a viral RT or a human endogenous RT. In some examples, the RT domain may be retron RT or DGRs RT. In some examples, the RT may be less mutagenic than a counterpart wildtype RT. In one embodiment, the RT herein is not mutagenic.

Retrons

[0452] In an embodiment, a donor template for homologous recombination is generated by use of a self-priming RNA template for reverse transcription. A non-limiting example of a self-priming reverse transcription system is the retron system. By the term “retron” it is meant a genetic element which encodes components enabling the synthesis of branched RNA-linked single stranded DNA (msDNA) and a reverse transcriptase. Retrons which encode msDNA are known in the art, for example, but not limited to U.S. Pat. No. 6,017,737; U.S. Pat. No. 5,849,563; U.S. Pat. No. 5,780,269; U.S. Pat. No. 5,436,141; U.S. Pat. No. 5,405,775; U.S. Pat. No. 5,320,958; CA 2,075,515; all of which are herein incorporated by reference.

[0453] In an embodiment, the reverse transcriptase domain is a retron RT domain. In an embodiment, the RNA template encodes a retron RNA template that is recognized and reverse transcribed by the retron reverse transcriptase domain. Conserved across many bacterial species, retrons are highly efficient reverse transcription systems of relatively unknown function. The retron system consists of the retron RT protein, as well as the msr and msd transcripts, which function as the primer and template sequences, respectively. All components of the retron system are expressed from a single open reading frame as a single transcript including the msr-msd and encoding the retron RT protein (Lampson, et al., 2005, Retrons, msDNA, and the bacterial genome. Cytogenet Genome Res 110:491–499). The msr element ORF of a retron provides for the RNA portion of the msDNA molecule, while the msd element ORF provides for the DNA portion of the msDNA molecule. The primary transcript from the msr-msd region is thought to serve as both a template and a primer to produce the msDNA. Synthesis of msDNA is primed from an internal rG residue of the RNA transcript using its 2'-

OH group. Modification of msd, or msr may also be made to permit insertion of a RNA template encoding a donor polynucleotide within the msd without altering the functioning of or the production of msDNA. The RNA template encoding a donor polynucleotide sequence may be any length but is preferably less than about 5 kb nucleotides, or also less than about 2 kb, or also less than 500 bases, provided that an msDNA product is produced.

Diversity Generating Retroelements (DGRs)

[0454] In an embodiment, the one or more functional domains may be a diversity generating retroelement(s) (e.g., DGR described in US20100041033A1). In one embodiment, the DGR may insert a donor polynucleotide with its homing mechanism. For example, the DGR may be associated with a catalytically inactive IscB protein (e.g., a dead IscB), and integrate the single-strand DNA using a homing mechanism. In some examples, the DGR may be less mutagenic than a counterpart wild type DGR. In some examples, the DGR is not error-prone. In one embodiment, the DGR herein is not mutagenic. The non-mutagenic DGR may be a mutant of a wild type DGR. As used herein, the term “DGR” encompasses both diversity generating retroelement polynucleotides and proteins encoded by diversity generating retroelement polynucleotides. In some examples, DGR may be proteins encoded by diversity generating retroelement polynucleotides having reverse transcriptase activity. In some examples, DGR may be proteins encoded by diversity generating retroelement polynucleotides having reverse transcriptase activity and integrase activity. In some cases, the template or donor polynucleotide may be encoded by a diversity generating retroelement polynucleotide. In certain cases, the template may be a polynucleotide different from the diversity generating retroelement polynucleotide, e.g., provided as a separate construct or molecule.

[0455] In one embodiment, the DGR herein may also include a Group II intron (and any proteins and polynucleotides encoded), which are mobile ribozymes that self-splice from precursor RNAs to yield excised intron lariat RNAs, which then invade new genomic DNA sites by reverse splicing. Examples of Group II intron include those described in Lambowitz AM et al., Group II Introns: Mobile Ribozymes that Invade DNA, Cold Spring Harb Perspect Biol. 2011 Aug; 3(8): a003616.

[0456] In one embodiment, the diversity-generating retroelements (DGRs) are genetic elements that can produce targeted, massive variations in the genomes that carry these elements. In one embodiment, the DGR systems rely on error-prone reverse transcriptases to produce mutagenized cDNA (containing A-to-N mutations) from a template region (TR), to

replace a segment called a variable region (VR) that is similar to the TR region—this process is called mutagenic retrohoming (see, e.g., Sharifi and Ye, MyDGR: a server for identification and characterization of diversity-generating retroelements. *Nucleic Acids Res.* 2019 Jul 2; 47(W1): W289–W294). DGRs may include a unique family of retroelements that generate sequence diversity of DNA. They exist widely in bacteria, archaea, phage and plasmid, and benefit their hosts by introducing variations and accelerating the evolution of target proteins (see, e.g., Yan et al., Discovery and characterization of the evolution, variation and functions of diversity-generating retroelements using thousands of genomes and metagenomes. *BMC Genomics.* 2019; 20: 595). The first DGR was discovered in a *Bordetella* phage, BPP-1. *Bordetella* causes the respiratory infection in humans and many other mammals, controlled by the BvgAS signal transduction system. The surface of *Bordetella* is highly variable owing to the dynamic gene expression in the infectious cycle. The invasion of BPP-1 to *Bordetella* relies on the phage tail fiber protein Mtd. With the process of mutagenic reverse transcription and cDNA integration, DGR may introduce multiple nucleotide substitutions to Mtd gene and generates different receptor-binding molecules, thus making BPP-1 the ability to invade *Bordetellae* with diverse cell surfaces.

[0457] The systems may be used to generate an ssDNA donor using a retron- or DGR RT, which is then integrated by homologous recombination upon target cleavage or nicking using a IscB nuclease. In one embodiment, the systems may comprise DGRs and/or Group-II intron reverse transcriptases. The homing mechanism of DGRs or Group-II introns may be used in modifying a target polynucleotide. The DGRs or Group-II introns reverse transcriptase may be guided to a target polynucleotide by tethering to a nuclease-dead IscB nuclease, TALE, or ZF protein. In another embodiment, a non-retron/DGR reverse transcriptase (e.g. a viral RT) may be used for generating cDNA off of a self-priming RNA. In one embodiment, a ssDNA may be generated by an RT, but integrate it using a dead IscB polypeptide or CRISPR-associated IscB polypeptide, creating an accessible R-loop instead of nicking/cleaving.

Topoisomerases

[0458] The one or more functional domains may be one or more topoisomerase domains. In one embodiment, an engineered system for modifying a target polynucleotide comprising: a IscB polypeptide or CRISPR-associated IscB polypeptide; a topoisomerase domain; and a nucleic acid template comprising or encoding a donor polynucleotide to be inserted to a target sequence of the target polynucleotide. In some examples, two or more of: the IscB polypeptide

or CRISPR-associated IscB polypeptide; topoisomerase domain; and nucleic acid template may form a complex. In some examples, two or more of: the IscB polypeptide or CRISPR-associated IscB polypeptide; topoisomerase domain, may be comprised in a fusion protein.

[0459] Topoisomerases are a class of enzymes that modify the topological state of DNA via the breakage and rejoining of nucleic acid strands. In some cases, a topoisomerase may be a DNA topoisomerase, which is an enzyme that controls and alters the topologic states of DNA during transcription, and catalyzes the transient breaking and rejoining of a single strand of DNA which allows the strands to pass through one another, thus altering the topology of DNA.

[0460] In one embodiment, the topoisomerase domain is capable of ligating the donor polynucleotide with the target polynucleotide. The ligation may be achieved by sticky end or blunt end ligation. In an example, the donor polynucleotide may comprise an overhang comprising a sequence complementary to a region of the target polynucleotide. Examples of ligating the donor polynucleotide with the target polynucleotide include those of TOPO cloning, e.g., those described in “The Technology Behind TOPO Cloning,” at www.thermofisher.com/us/en/home/life-science/cloning/topo/topo-resources/the-technology-behind-topo-cloning.html.

[0461] In one embodiment, the topoisomerase domain may be associated with the donor polynucleotide. For example, the topoisomerase domain is covalently linked to the donor polynucleotide.

[0462] In one embodiment, a topoisomerase domain may be provided together with, e.g., associated (e.g., fused) with a IscB polypeptide or CRISPR-associated IscB polypeptide (e.g., a IscB polypeptide or CRISPR-associated IscB polypeptide or a variant thereof such as a dead IscB or a IscB nickase). Alternatively or additionally, the topoisomerase domain may be on a molecule different from the IscB polypeptide or CRISPR-associated IscB polypeptide. In some cases, the topoisomerase domain may be associated with a donor polynucleotide. For example, the topoisomerase domain may be pre-loaded covalently with a donor DNA molecule. Such design may allow for efficient ligation of only a specific cargo. The topoisomerase domain may ligate the donor polynucleotide (e.g., a DNA molecule) to a target site on a target polynucleotide (e.g., a free double-stranded DNA end). In one embodiment, the donor polynucleotide may have an overhang that comprises a sequence complementary to a region of the target polynucleotide. For example, the overhang may invade into the target polynucleotide at a cut site generated by the IscB polypeptide or CRISPR-associated IscB polypeptide.

[0463] Examples of topoisomerases include type I, including type IA and type IB topoisomerases, which cleave a single strand of a double-stranded nucleic acid molecule, and type II topoisomerases (e.g., gyrases), which cleave both strands of a double-stranded nucleic acid molecule.

[0464] Type IA and IB topoisomerases cleave one strand of a double-stranded nucleic acid molecule. In some examples, the cleavage of a double-stranded nucleic acid molecule by type IA topoisomerases generates a 5' phosphate and a 3' hydroxyl at the cleavage site, with the type IA topoisomerase covalently binding to the 5' terminus of a cleaved strand. Cleavage of a double-stranded nucleic acid molecule by type IB topoisomerases may generate a 3' phosphate and a 5' hydroxyl at the cleavage site, with the type IB topoisomerase covalently binding to the 3' terminus of a cleaved strand.

[0465] Examples of Type IA topoisomerases include *E. coli* topoisomerase I, *E. coli* topoisomerase III, eukaryotic topoisomerase II, archeal reverse gyrase, yeast topoisomerase III, *Drosophila* topoisomerase III, human topoisomerase III, *Streptococcus pneumoniae* topoisomerase III, and the like, including other type IA topoisomerases. A DNA-protein adduct is formed with the enzyme covalently binding to the 5'-thymidine residue, with cleavage occurring between the two thymidine residues.

[0466] Examples of Type IB topoisomerases include the nuclear type I topoisomerases present in all eukaryotic cells and those encoded by *Vaccinia* and other cellular poxviruses. The eukaryotic type IB topoisomerases are exemplified by those expressed in yeast, *Drosophila* and mammalian cells, including human cells. Viral type IB topoisomerases are exemplified by those produced by the vertebrate poxviruses (*Vaccinia*, Shope fibroma virus, ORF virus, fowlpox virus, and molluscum contagiosum virus), and the insect poxvirus (*Amsacta moorei entomopoxvirus*).

[0467] Examples of Type II topoisomerases include, bacterial gyrase, bacterial DNA topoisomerase IV, eukaryotic DNA topoisomerase II, and T-even phage encoded DNA topoisomerases. Type II topoisomerases may have both cleaving and ligating activities. Substrate double-stranded nucleic acid molecules of type II topoisomerase can be prepared such that the type II topoisomerase can form a covalent linkage to one strand at a cleavage site. For example, calf thymus type II topoisomerase can cleave a substrate ds nucleic acid molecule containing a 5' recessed topoisomerase recognition site positioned three nucleotides from the 5' end, resulting in dissociation of the three nucleic acid molecule 5' to the cleavage site and

covalent binding of the topoisomerase to the 5' terminus of the ds nucleic acid molecule. Furthermore, upon contacting such a type II topoisomerase-charged ds nucleic acid molecule with a second nucleic acid molecule containing a 3' hydroxyl group, the type II topoisomerase can ligate the sequences together, and then is released from the recombinant nucleic acid molecule.

[0468] In some examples, the topoisomerase is DNA topoisomerase I, e.g., a Vaccinia virus topoisomerase I. The topoisomerase may be pre-loaded with a donor polynucleotide. The Vaccinia virus topoisomerase may need a target comprising a 5' -OH group.

Phosphatases

[0469] The systems herein may further comprise a phosphatase domain. A phosphatase is an enzyme capable of removing a phosphate group from a molecule e.g., a nucleic acid such as DNA. Examples of phosphatases include calf intestinal phosphatase, shrimp alkaline phosphatase, Antarctic phosphatase, and APEX alkaline phosphatase.

[0470] In some examples, the 5' -OH group of in the target polynucleotide may be generated by a phosphatase. A topoisomerase compatible with a 5' phosphate target may be used to generate stable loaded intermediates. In some cases, a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease that leaves a 5' OH after cleaving the target polynucleotide may be used. In some cases, the phosphatase domain may be associated with (e.g., fused to) the IscB protein. The phosphatase domain may be capable of generating a -OH group at a 5' end of the target polynucleotide. The phosphatase may be delivered separated from other components in the system, e.g., as a separate protein, on a separate vector from other components.

Polymerases

[0471] The systems herein may further comprise a polymerase domain. A polymerase refers to an enzyme that synthesizes chains of nucleic acids. The polymerase may be a DNA polymerase or an RNA polymerase.

[0472] In one embodiment, the systems comprise an engineered system for modifying a target polynucleotide comprising: a IscB polypeptide or CRISPR-associated IscB polypeptide; a DNA polymerase domain; and a DNA template comprising a donor polynucleotide to be inserted to a target sequence of the target polynucleotide. In some examples, two or more of: the IscB protein; DNA polymerase domain; and DNA template may form a complex. In some examples, two or more of: the IscB protein; DNA polymerase domain; are comprised in a

fusion protein. For example, the IscB polypeptide or CRISPR-associated IscB polypeptide and DNA polymerase domain may be comprised in a fusion protein.

[0473] In one embodiment, the systems may comprise a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease (or variant thereof such as a dIscB polypeptide or CRISPR-associated IscB polypeptide or IscB polypeptide or CRISPR-associated IscB polypeptide nickase) and a DNA polymerase (e.g. phi29, T4, T7 DNA polymerase). The systems may further comprise a single-stranded DNA or double-stranded DNA template. The DNA template may comprise i) a first sequence homologous to a target site of the IscB protein on the target polynucleotide, and/or ii) a second sequence homologous to another region of the target polynucleotide. In one embodiment, the template may be a synthetic single-stranded or PCR-generated DNA molecule, (optionally end-protected by modified nucleotides), or a viral genome (e.g. AAV). In another embodiment, the template is generated using a reverse transcriptase. When the system is delivered into a cell, an endogenous DNA polymerase in the cell may be used. Alternatively or additionally, an exogenous DNA polymerase may be expressed in the cell.

[0474] The DNA template may be end-protected by one or more modified nucleotides, or comprises a portion of a viral genome. In some embodiment, the DNA template comprises LNA or other modifications (e.g., at the 3' end). The presence of LNA and/or the modifications may lead to more efficient annealing with the 3' flap generated by IscB polypeptide or CRISPR-associated IscB polypeptide cleavage.

[0475] Examples of DNA polymerase include Taq, Tne (exo −), Tma (exo −), Pfu (exo −), Pwo (exo −), Thermoanaerobacter thermohydrosulfuricus DNA polymerase, Thermococcus litoralis DNA polymerase I, E. coli DNA polymerase I, Taq DNA polymerase I, Tth DNA polymerase I, Bacillus stearothermophilus (Bst) DNA polymerase I, E. coli DNA polymerase III, bacteriophage T5 DNA polymerase, bacteriophage M2 DNA polymerase, bacteriophage T4 DNA polymerase, bacteriophage T7 DNA polymerase, bacteriophage phi29 DNA polymerase, bacteriophage PRD1 DNA polymerase, bacteriophage phi15 DNA polymerase, bacteriophage phi21DNA polymerase, bacteriophage PZE DNA polymerase, bacteriophage PZA DNA polymerase, bacteriophage Nf DNA polymerase, bacteriophage M2Y DNA polymerase, bacteriophage B103 DNA polymerase, bacteriophage SF5 DNA polymerase, bacteriophage GA-1 DNA polymerase, bacteriophage Cp-5 DNA polymerase, bacteriophage

Cp-7 DNA polymerase, bacteriophage PR4 DNA polymerase, bacteriophage PR5 DNA polymerase, bacteriophage PR722 DNA polymerase and bacteriophage L17 DNA polymerase.

Ligases

[0476] In general, the systems comprise a IscB polypeptide or CRISPR-associated IscB polypeptide and a ligase associated with the IscB protein. The IscB polypeptide or CRISPR-associated IscB polypeptide may be recruited to the target sequence by an ω RNA or guide RNA, and generate a break on the target sequence. The ω RNA or guide RNA may further comprise a template sequence with desired mutations or other sequence elements. The template sequence may be ligated to the target sequence to introduce the mutations or other sequence elements to the nucleic acid molecule. The IscB polypeptide or CRISPR-associated IscB polypeptide may be a nickase that generates a single-strand break on nucleic acid molecule, and the ligase may be a single-strand DNA ligase. In one embodiment, the systems comprise a pair of IscB polypeptide or CRISPR-associated IscB polypeptide-ligases complexes with two distinct ω RNA sequences. Each IscB polypeptide or CRISPR-associated IscB polypeptide-ligase complex can target one strand of a double-stranded polynucleotide, and work together to effectively modify the sequence of the double-stranded polynucleotides.

[0477] In some examples, the IscB polypeptide or CRISPR-associated IscB polypeptide is associated with a ligase or functional fragment thereof. The ligase may ligate a single-strand break (a nick) generated by the IscB polypeptide or CRISPR-associated IscB polypeptide. In certain cases, the ligase may ligate a double-strand break generated by the IscB polypeptide or CRISPR-associated IscB polypeptide. In certain examples, the IscB polypeptide or CRISPR-associated IscB polypeptide is associated with a reverse transcriptase or functional fragment thereof.

[0478] The present invention further provides systems and methods of modifying a nucleic acid sequence using a pair of distinct IscB polypeptide or CRISPR-associated IscB polypeptide-ligase- ω RNA or guide RNA complexes, said systems and methods comprising: (a) an engineered IscB polypeptide or CRISPR-associated IscB polypeptide connected to or complexed with a ligase; (b) two distinct ω RNA or guide RNA sequences complexed with such IscB polypeptide or CRISPR-associated IscB polypeptide-ligase protein complex to form a first and a second distinct IscB-ligase ω RNA complexes; (c) the first IscB-ligase- ω RNA or guide RNA complex binding to one strand of a target double-stranded polynucleotide

sequence, and the second IscB polypeptide or CRISPR-associated IscB polypeptide-ligase- ω RNA or guide RNA complex binding to another strand of the target double-stranded polynucleotide sequence; (d) upon binding of the said complexes to the locus of interest the effector protein induces the modification of the sequences associated with or at the target locus of interest, whereby the two IscB polypeptide or CRISPR-associated IscB polypeptide-ligase- ω RNA or guide RNA complexes work together on different strands of the double-stranded target sequence and modify the sequence.

[0479] One of the advantages of using such a “pair” of IscB polypeptide or CRISPR-associated IscB polypeptide-ligase- ω RNA or guide RNA complexes includes high efficiency in modifying the sequence associated with or at the locus of interest of target double-stranded polynucleotides.

[0480] In one embodiment, the IscB polypeptide or CRISPR-associated IscB polypeptide can be a nickase. In a preferred embodiment, a ligase is linked to the IscB polypeptide or CRISPR-associated IscB polypeptide. The ligase can ligate the donor sequence to the target sequence. The ligase can be a single-strand DNA ligase or a double-strand DNA ligase. The ligase can be fused to the carboxyl-terminus of a IscB polypeptide or CRISPR-associated IscB polypeptide, or to the amino-terminus of a IscB polypeptide or CRISPR-associated IscB polypeptide.

[0481] As used herein the term “ligase” refers to an enzyme, which catalyzes the joining of breaks (e.g., double-stranded breaks or single-stranded breaks (“nicks”) between adjacent bases of nucleic acids. For example, a ligase may be an enzyme capable of forming intra- or inter-molecular covalent bonds between a 5' phosphate group and a 3' hydroxyl group. The term “ligate” refers to the reaction of covalently joining adjacent oligonucleotides through formation of an internucleotide linkage.

[0482] DNA ligases fall into two general categories: ATP-dependent DNA ligases (EC 6.5.1.1), and NAD (+) dependent DNA ligases (EC 6.5.1.2). NAD (+) dependent DNA ligases are found only in bacteria (and some viruses) while ATP-dependent DNA ligases are ubiquitous. The ATP-dependent DNA ligases can be divided into four classes: DNA ligase I, II, III, and IV. DNA ligase I links Okazaki fragments to form a continuous strand of DNA; DNA ligase II is an alternatively spliced form of DNA ligase III, found only in non-dividing cells; DNA ligase III is involved in base excision repair; and DNA ligase IV is involved in the

repair of DNA double-strand breaks by non-homologous end joining (NHEJ). Amongst all ligases, there are two types of prokaryotic and one type of eukaryotic ligases that are particularly well suited for facilitating the blunt-ended, double-stranded DNA ligation: Prokaryotic DNA ligases (T3 and T4) and Eukaryotic DNA ligase (Ligase 1).

[0483] In some cases, the ligase is specific for double-stranded nucleic acids (e.g., dsDNA, dsRNA, RNA/DNA duplex). An example of a ligase specific for double-stranded DNA and DNA/RNA hybrids is T4 DNA ligase. In some cases, the ligase is specific for single-stranded nucleic acids (e.g., ssDNA, ssRNA). An example of such ligase is CircLigase II. In some cases, the ligase is specific for RNA/DNA duplexes. In some cases, the ligase is able to work on single-stranded, double-stranded, and/or RNA/DNA nucleic acids in any combination.

[0484] In some cases, the ligase may be a pan-ligase, which is a single ligase with the ability to ligate both DNA and RNA targets. The ligase may be specific for a target (e.g., DNA-specific or RNA-specific). In some cases, the ligase may be a dual ligase system that include DNA-specific, RNA-specific, and/or pan-ligases, in any combination.

[0485] Examples of ligases that can be used with the disclosure include T4 DNA Ligase, T3 DNA Ligase, T7 DNA Ligase, E. coli DNA Ligase, HiFi Taq DNA Ligase, 9° NTM DNA Ligase, Taq DNA Ligase, SplintR® Ligase (also known as PBCV-1 DNA Ligase or Chlorella virus DNA Ligase), Thermostable 5' AppDNA/RNA Ligase, T4 RNA Ligase, T4 RNA Ligase 2, T4 RNA Ligase 2 Truncated, T4 RNA Ligase 2 Truncated K227Q, T4 RNA Ligase 2, Truncated KQ, RtcB Ligase (joins single stranded RNA with a 3''-phosphate or 2',3'-cyclic phosphate to another RNA), CircLigase II, CircLigase ssDNA Ligase, CircLigase RNA Ligase, or Ampligase® Thermostable DNA Ligase, NAD-dependent ligases including Taq DNA ligase, Thermus filiformis DNA ligase, Escherichia coli DNA ligase, Tth DNA ligase, Thermus scotoductus DNA ligase (I and II), thermostable ligase, Ampligase thermostable DNA ligase, VanC-type ligase, 9° N DNA Ligase, Tsp DNA ligase, and novel ligases discovered by bioprospecting; ATP-dependent ligases including T4 RNA ligase, T4 DNA ligase, T3 DNA ligase, T7 DNA ligase, Pfu DNA ligase, DNA ligase I, DNA ligase III, DNA ligase IV, and novel ligases discovered by bioprospecting, and wild-type, mutant isoforms, and genetically engineered variants thereof. In a particular example, the ligase is a

[0486] In one embodiment, the examples of the ligases include those used in sequencing by synthesis or sequencing by ligation reactions.

HELITRONS

[0487] The systems and compositions herein may comprise a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, one or more ω RNAs or guide RNAs, and one or more components of a helitron. The systems and compositions may be used to insert a donor polynucleotide to a target polynucleotide. The systems and compositions may further comprise a donor polynucleotide.

[0488] The term “helitron”, as used herein, refers to a polynucleotide (or nucleic acid segment), recognized as a transposon that captures and mobilizes gene fragments in eukaryotes. The term “helitron” as used herein refers to transposase that comprises an endonuclease domain and a C-terminal helicase domain. Helitrons are rolling-circle RNA transposons. In one embodiment, the helitron encodes a 1400 to about 2000 amino acid, or about 1800 amino acid multidomain transposase. In embodiments, the helitron comprises a hairpin near the 3' end to function as a transposition terminator. In embodiments, the transposon comprises a RepHel motif comprising a replication initiator (Rep) and a DNA helicase (hel) domain. See, Thomas J. & Pritham E. J. Helitrons, the eukaryotic rolling-circle transposable elements. *Microbiol. Spectr.* 3, 893–926 (2015). In embodiments, the helitron comprises a Rep nuclease domain and C-terminal helicase domain and inserts between an AT dinucleotide in single strand DNA. In an aspect, the C-terminal helicase unwinds the DNA in a 5' to 3' direction. The HUH nuclease domain may comprise one or two active site tyrosine residues, in embodiments, is a 2 Tyrosine (Y2) HUH endonuclease domain. Helitrons can encompass helentron, proto-helentron and helitron2 type proteins, structures of which can be as described in Thomas et al., 2015 at Figures 1 and 3, incorporated specifically by reference. Particular organisms in which the helitron or helentrons have been found can include those in Table 1 of Thomas J. & Pritham E. J. Helitrons, the eukaryotic rolling-circle transposable elements. *Microbiol. Spectr.* 3, 893–926 (2015), incorporated herein by reference. Similarly, helitrons can be identified based at least in part on the Rep motif, and conserved residues in the helitrons, and according to the alignment sequence of Figure 2 of Thomas J. & Pritham E. J. Helitrons, the eukaryotic rolling-circle transposable elements. *Microbiol. Spectr.* 3, 893–926 (2015), specifically incorporated herein by reference.

[0489] The expression “helitron reaction” used herein refers to a reaction wherein a transposase inserts a donor polynucleotide sequence in or adjacent to an insertion site on a

target polynucleotide. The insertion site may contain a sequence or secondary structure recognized by the helitron and/or an insertion motif sequence in the target polynucleotide into which the donor polynucleotide sequence may be inserted.

[0490] As described in Grabundzija 2018, the helitron terminal sequences contains a distinct ~150 base pairs (bp) long sequence with an absolutely conserved dinucleotide at the end of left terminal sequence (LTS), and a tetranucleotide at the end of right terminal sequence (RTS) which is preceded by a palindromic sequence that can form a hairpin structure. Grabundzija et al., Nat. Commun. 2018; 9: 1278; doi:10.1035/s41467-018-03688-w.

[0491] The helitron end sequences may be responsible for identifying the donor polynucleotide for transposition. The helitron end sequences may be the DNA sequences used to perform a transposition reaction, the end sequences may be referred to herein as right terminal sequences and left terminal sequence. The donor polynucleotide can be configured to comprise a first and second helitron recognition sequence that are at least 80%, 85%, 90%, 95% 96%, 97%, 98%, 99% or 100% complementary to a left terminal sequence and/or a right terminal sequence of a polynucleotide encoding the helitron polypeptide.

[0492] In an aspect, the palindromic sequence may be located upstream of the right terminal sequence, for example, about 5, 10, 15, 20, 25, 30, 35 nucleotides upstream of the right terminal sequence end, or about 10 to 15 nucleotides upstream of the right terminal sequence end, about 10 to 12 nucleotides or about 11 nucleotides upstream of the right terminal sequence end. Ivana Grabundzija, Nat Commun. 2016; 7:10716, doi:10.1038/ncomms10716, incorporated herein by reference.

[0493] Exemplary helitrons can be identified using software, for example (EAHelitron) that has been used to identify Helitrons in a wide range of plant genomes. See, Hu, K., Xu, K., Wen, J. *et al.* *Helitron* distribution in Brassicaceae and whole Genome Helitron density as a character for distinguishing plant species. *BMC Bioinformatics* 20, 354 (2019). doi: 10.1186/s12859-019-2945-8, incorporated herein by reference.

[0494] The helitron may be derived from a eukaryote. In an aspect, the helitron is derived from a mammalian genome, in an aspect, vespertilionid bats, e.g. Helibat. In embodiments, the helitron is derived from derived from a Helibat1 transposon. In embodiments, the helitron is Helraiser, the full DNA sequence of the consensus transposon, including left terminal and right terminal sequences as well as hairpin identified is provided in Grabundzija, 2016 at Supplementary Figure 1, specifically incorporated herein by reference. In an aspect, the

helitron is flanked by left and right terminal sequences of the transposon. In an aspect, the left terminal sequence and right terminal sequence terminates with the conserved 5'-TC/CTAG-3' motif. In an embodiment, the helitron may comprise a palindromic sequence that is about 10 to about 35, or about 5-25 bp or about 19-bp-long palindromic sequence with the potential to form a hairpin structure.

[0495] Elements of these systems may be engineered to work within the context of the invention. For example, a helitron polypeptide may be fused to a polypeptide capable of generating an R-loop. Fusion may be by any appropriate linker, in an exemplary embodiment, XTEN16. The binding elements that allow a helitron polypeptide to bind, for example, the use of sequences complementary to the right terminal sequence and the left terminal sequence of the helitron may be engineered into a donor construct to facilitate entry of a donor polynucleotide sequence into a target polynucleotide.

[0496] In certain example embodiments, the Isc polypeptide, via formation of complex with a hRNA sequence, directs the helitron polypeptide to a target sequence in a target polynucleotide, where the helitron facilitates integration of a donor polynucleotide sequence into the target polynucleotide.

[0497] The helitron polypeptides may also comprise one or more truncations or excisions to remove domains or regions of wild-type protein to arrive at a minimal polypeptide, alter functionality according to the system in which the helitron is used, or mutated to enhance or diminish particular activities associated with the helitron, i.e. nuclease activity or helicase activity.

ISCB OR CRISPR-ASSOCIATED ISCB RECOMBINASE

[0498] The systems and compositions herein may comprise a IscB polypeptide or CRISPR-associated IscB polypeptide system, and one or more components of a recombinase. In an aspect, the IscB polypeptide or CRISPR-associated IscB polypeptide is naturally catalytically inactive and utilized with one or more nucleic acid components to provide site-specific targetings, and the one or more components of the recombinase to introduce a modification. In an aspect, the IscB polypeptide or CRISPR-associated IscB polypeptide polypeptide may be catalytically inactivated via mutation of one or more residues of a catalytic domain or via truncation, and utilized with one or more RNA components to provide site-specific targeting, and the one or more components of the recombinase introduce a modification. In an aspect, the IscB polypeptide or CRISPR-associated IscB polypeptide may be catalytically inactivated, and

utilized with one or more RNA components to provide site-specific targeting, with the one or more components of the recombinase introduce a modification. In preferred embodiments, the recombinase mediates unidirectional site-specific recombination. In one embodiment, the recombinase is a serine recombinase (SR), encoded, for example, by IS607 family, Tn4451, and bacteriophage phiC31. See, generally, Smith MC, Thorpe HM: Diversity in the serine recombinases. *Mol Microbiol.* 2002, 44: 299-307. 10.1046/j.1365-2958.2002.02891.x; Li et al., (2018) *J. Mol. Biol.* 430:21, 4401-4418. In an embodiment, the recombinase is a tyrosine recombinase (YR) encoded by IS91, Helitron, IS200/IS605, Crypton or DIRS-retrotransposon families. See, generally, Goodwin TJ, Butler MI, Poulter RT: Cryptons: a group of tyrosine-recombinase-encoding DNA transposons from pathogenic fungi. *Microbiology.* 2003, 149: 3099-3109. Doi:10.1099/mic.0.26529-0; Cappello J, Handelsman K, Lodish HF: Sequence of *Dictyostelium* DIRS-1: an apparent retrotransposon with inverted terminal repeats and an internal circle junction sequence. *Cell.* 1985, 43: 105-115. 10.1016/0092-8674(85)90016-9; In an aspect, the recombinase provides site-specific integration of a template that can be provided with the composition, e.g. a donor oligonucleotide. Without being bound by theory, the recombinase allows for integration independent of payload size and can coordinate strand exchange and re-ligation across multiple cell types, allowing integration of long stretches of polynucleotides. In an exemplary embodiment, the serine recombinase is PhiC31 and the target is DNA. In an aspect, the phiC31 allows for integration of a target site comprising an attP or pseudoattP recognition site. See, e.g. systembio.com/wp-content/uploads/phiC31_productsheet-1.pdf. In an embodiment utilizing phiC231, a donor oligonucleotide would be provided with an attB at sequence that facilitates attachment at the attP site of the target genome. Similar approaches of designing donor oligonucleotides with sequences complementary to attachment sites for a recombinase can be designed for use with the present invention. See, e.g. Li et al., (2018) *J. Mol. Biol.* 430:21, 4401-4418.

[0499] In one embodiment, a naturally inactive IscB is provided with an IS630 transposon. IS630 transposons comprise the DDE motif, which is an RNase H-like fold that draws three catalytically active residues DDE signature and average about 1100 bp in length. See, Gao, B., Wang, Y., Diaby, M. et al. Evolution of pogo, a separate superfamily of IS630-Tc1-mariner transposons, revealing recurrent domestication events in vertebrates. *Mobile DNA* 11, 25 (2020) doi: 10.1186/s13100-020-00220-0. In an aspect, the IS630 transposon provides high

target specificity inserting into a TA dinucleotide. In an aspect, the insertion is at a 5'-NTAN-3'.

SYSTEMS AND COMPLEXES

[0500] In one aspect, the present disclosure provides nucleic acid-targeting systems. Such systems may be used to target, modify, and otherwise manipulate a nucleic acid. In one embodiment, the systems comprise the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and one or more gRNAs or guide RNAs. The IscB polypeptide or CRISPR-associated IscB polypeptide nuclease may have nuclease activity, e.g., capable of cleaving DNA or RNA. The IscB polypeptide or CRISPR-associated IscB polypeptide nuclease may have nickase activity, e.g., capable of generating a single-strand break on a double-strand nucleic acid such as dsDNA or dsRNA. The IscB polypeptide or CRISPR-associated IscB polypeptide nuclease may be in a dead form, e.g., has nickase activity, or does not have nuclease or nickase activity. In one embodiment, the systems further comprising one or more functional domains, e.g., nucleotide deaminase, reverse transcriptase, non-LTR retrotransposon (and protein encoded), polymerase, diversity generating element (and protein encoded). In some examples, the systems further comprise one or more donor polynucleotides. The donor polynucleotides may be inserted to a target polynucleotide by the systems. The donor polynucleotide may be comprised in or coded by a nucleic acid template.

[0501] In some examples, two or more of the components in a system herein may form a complex. For example, the components are separate molecules but interact with each other directly or indirectly. In certain two or more of the components in a system herein may be comprised in a fusion protein.

[0502] As used herein, “target sequence” refers to a sequence to which a guide sequence is designed to have complementarity, where hybridization between a target sequence and a guide RNA promotes the formation of a DNA or RNA-targeting complex. Full complementarity is not necessarily required, provided there is sufficient complementarity to cause hybridization and promote formation of a nucleic acid-targeting complex. A target sequence may comprise RNA polynucleotides. In one embodiment, a target sequence is located in the nucleus or cytoplasm of a cell. In one embodiment, the target sequence may be within an organelle of a eukaryotic cell, for example, mitochondrion or chloroplast. A sequence or template that may be used for recombination into the targeted locus comprising the target sequences is referred

to as an “editing template” or “editing sequence”. In aspects of the invention, an exogenous template may be referred to as an editing template. In an aspect the recombination is homologous recombination.

[0503] In one embodiment, formation of a nucleic acid-targeting complex (comprising a guide RNA hybridized to a target sequence and complexed with one or more nucleic acid-targeting effector proteins) results in cleavage of one or both nucleic acid strands in or near (e.g. within 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or more base pairs from) the target sequence. In one embodiment, one or more vectors driving expression of one or more elements of a nucleic acid-targeting system are introduced into a host cell such that expression of the elements of the nucleic acid-targeting system direct formation of a nucleic acid-targeting complex at one or more target sites. For example, a nucleic acid-targeting effector protein and a ω RNA or guide RNA could each be operably linked to separate regulatory elements on separate vectors. Alternatively, two or more of the elements expressed from the same or different regulatory elements, may be combined in a single vector, with one or more additional vectors providing any components of the nucleic acid-targeting system not included in the first vector. nucleic acid-targeting system elements that are combined in a single vector may be arranged in any suitable orientation, such as one element located 5' with respect to (“upstream” of) or 3' with respect to (“downstream” of) a second element. The coding sequence of one element may be located on the same or opposite strand of the coding sequence of a second element, and oriented in the same or opposite direction. In one embodiment, a single promoter drives expression of a transcript encoding a nucleic acid-targeting effector protein and a guide RNA embedded within one or more intron sequences (e.g. each in a different intron, two or more in at least one intron, or all in a single intron). In one embodiment, the nucleic acid-targeting effector protein and guide RNA are operably linked to and expressed from the same promoter.

[0504] The present disclosure encompasses computational methods and algorithms to predict new IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, identify the components, and new nucleic acid-targeting systems therein. In some examples, a computational method of identifying novel IscB polypeptide or CRISPR-associated IscB polypeptide nuclease loci analysis of the candidates may be conducted by searching metagenomics databases for additional homologs.

[0505] In one aspect the identifying all predicted protein coding genes is carried out by comparing the identified genes with IscB polypeptide or CRISPR-associated IscB polypeptide-specific profiles and annotating them according to NCBI Conserved Domain Database (CDD) which is a protein annotation resource that consists of a collection of well-annotated multiple sequence alignment models for ancient domains and full-length proteins. These are available as position-specific score matrices (PSSMs) for fast identification of conserved domains in protein sequences via RPS-BLAST. CDD content includes NCBI-curated domains, which use 3D-structure information to explicitly define domain boundaries and provide insights into sequence/structure/function relationships, as well as domain models imported from a number of external source databases (Pfam, SMART, COG, PRK, TIGRFAM).

[0506] In a further aspect, the case-by-case analysis is performed using PSI-BLAST (Position-Specific Iterative Basic Local Alignment Search Tool). PSI-BLAST derives a position-specific scoring matrix (PSSM) or profile from the multiple sequence alignment of sequences detected above a given score threshold using protein-protein BLAST. This PSSM is used to further search the database for new matches, and is updated for subsequent iterations with these newly detected sequences. Thus, PSI-BLAST provides a means of detecting distant relationships between proteins.

[0507] In another aspect, the case-by-case analysis is performed using HHpred, a method for sequence database searching and structure prediction that is as easy to use as BLAST or PSI-BLAST and that is at the same time much more sensitive in finding remote homologs. In fact, HHpred's sensitivity is competitive with the most powerful servers for structure prediction currently available. HHpred is the first server that is based on the pairwise comparison of profile hidden Markov models (HMMs). Whereas most conventional sequence search methods search sequence databases such as UniProt or the NR, HHpred searches alignment databases, like Pfam or SMART. This greatly simplifies the list of hits to a number of sequence families instead of a clutter of single sequences. All major publicly available profile and alignment databases are available through HHpred. HHpred accepts a single query sequence or a multiple alignment as input. Within only a few minutes it returns the search results in an easy-to-read format similar to that of PSI-BLAST. Search options include local or global alignment and scoring secondary structure similarity. HHpred can produce pairwise query-template sequence alignments, merged query-template multiple alignments (e.g. for transitive searches), as well as 3D structural models calculated by the MODELLER software from HHpred alignments.

Multiplexing

[0508] In one embodiment, IscB polypeptide or CRISPR-associated IscB polypeptide nucleases may be used in a multiplex (tandem) targeting approach. For example, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease herein can employ more than one RNA guide without losing activity. This may enable the use of the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, systems or complexes as defined herein for targeting multiple DNA targets, genes or gene loci, with a single enzyme, system or complex as defined herein. The ω RNA or guide RNAs may be tandemly arranged, optionally separated by a nucleotide sequence such as a conserved nucleotide sequence as defined herein. The position of the different ω RNA or guide RNAs in the tandem does not influence the activity.

[0509] In one aspect, the IscB polypeptide or CRISPR-associated IscB polypeptide nucleases may be used for tandem or multiplex targeting. It is to be understood that any of the IscB polypeptide or CRISPR-associated IscB polypeptide nucleases, complexes, or compositions herein elsewhere may be used in such an approach. Any of the methods, products, compositions and uses as described herein elsewhere are equally applicable with the multiplex or tandem targeting approach further detailed below. By means of further guidance, the following particular aspects and embodiments are provided.

[0510] In one aspect, the invention provides for the use of a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, complex or system as defined herein for targeting multiple gene loci. In one embodiment, this can be established by using multiple (tandem or multiplex) ω RNA or guide RNA (gRNA) sequences.

[0511] In one aspect, the invention provides methods for using one or more elements of a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, complex or system as defined herein for tandem or multiplex targeting, wherein said system herein comprises multiple ω RNA or guide RNA sequences. Said ω RNA or gRNA sequences are separated by a nucleotide sequence, such as a conserved nucleotide sequence as defined herein elsewhere.

[0512] The IscB polypeptide or CRISPR-associated IscB polypeptide nucleases, compositions, systems or complexes as defined herein provides an effective means for modifying multiple target polynucleotides. The IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, system or complex as defined herein has a wide variety of utility including modifying (e.g., deleting, inserting, translocating, inactivating, activating) one or

more target polynucleotides in a multiplicity of cell types. As such the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, system or complex as defined herein of the invention has a broad spectrum of applications in, e.g., gene therapy, drug screening, disease diagnosis, and prognosis, including targeting multiple gene loci within a single system.

[0513] In one aspect, the present disclosure provides a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, system or complex as defined herein, having a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease having at least one destabilization domain associated therewith, and multiple guide RNAs that target multiple nucleic acid molecules such as DNA molecules, whereby each of said multiple guide RNAs specifically targets its corresponding nucleic acid molecule, e.g., DNA molecule. Each nucleic acid molecule target, e.g., DNA molecule can encode a gene product or encompass a gene locus. Using multiple ω RNA or guide RNAs hence enables the targeting of multiple gene loci or multiple genes. In one embodiment the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease may cleave the DNA molecule encoding the gene product. In one embodiment expression of the gene product is altered. The IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and the guide RNAs do not naturally occur together. The present disclosure comprehends the ω RNA or guide RNAs comprising tandemly arranged guide sequences. The present disclosure further comprehends coding sequences for the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease being codon optimized for expression in a eukaryotic cell. In an embodiment the eukaryotic cell is a mammalian cell, a plant cell or a yeast cell and in a more preferred embodiment the mammalian cell is a human cell. Expression of the gene product may be decreased. The IscB polypeptide or CRISPR-associated IscB polypeptide nuclease may form part of a system or complex, which further comprises tandemly arranged ω RNA or guide RNAs (gRNAs) comprising a series of 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 25, 25, 30, or more than 30 guide sequences, each capable of specifically hybridizing to a target sequence in a genomic locus of interest in a cell. In one embodiment, the functional system or complex binds to the multiple target sequences. In one embodiment, the functional system or complex may edit the multiple target sequences, e.g., the target sequences may comprise a genomic locus, and In one embodiment, there may be an alteration of gene expression. In one embodiment, the functional system or complex may comprise further functional domains. In one embodiment, the invention provides a method for altering

or modifying expression of multiple gene products. The method may comprise introducing into a cell containing said target nucleic acids, e.g., DNA molecules, or containing and expressing target nucleic acid, e.g., DNA molecules; for instance, the target nucleic acids may encode gene products or provide for expression of gene products (e.g., regulatory sequences).

[0514] In one embodiment, the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease used for multiplex targeting is associated with one or more functional domains. In some more specific embodiments, the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease used for multiplex targeting is a dead IscB polypeptide nuclease. The inventors have found that the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease as described herein may enable improved and/or direct access to one or more nucleotides involved in the DNA:RNA duplex.

Donor polynucleotides

[0515] In one embodiment, the compositions and systems herein may comprise one or more nucleic acid templates. In some cases, the nucleic acid template may comprise one or more polynucleotides. In certain cases, the nucleic acid template may comprise coding sequences for one or more polynucleotides. The nucleic acid template may be an RNA template. The nucleic acid template may be a DNA template.

[0516] The donor polynucleotide may be used for editing the target polynucleotide. In some cases, the donor polynucleotide comprises one or more mutations to be introduced into the target polynucleotide. Examples of such mutations include substitutions, deletions, insertions, or a combination thereof. The mutations may cause a shift in an open reading frame on the target polynucleotide. In some cases, the donor polynucleotide alters a stop codon in the target polynucleotide. For example, the donor polynucleotide may correct a premature stop codon. The correction may be achieved by deleting the stop codon or introduces one or more mutations to the stop codon. In other example embodiments, the donor polynucleotide addresses loss of function mutations, deletions, or translocations that may occur, for example, in certain disease contexts by inserting or restoring a functional copy of a gene, or functional fragment thereof, or a functional regulatory sequence or functional fragment of a regulatory sequence. A functional fragment refers to less than the entire copy of a gene by providing sufficient nucleotide sequence to restore the functionality of a wild type gene or non-coding regulatory sequence (e.g. sequences encoding long non-coding RNA). In certain example embodiments, the systems disclosed herein may be used to replace a single allele of a defective

gene or defective fragment thereof. In another example embodiment, the systems disclosed herein may be used to replace both alleles of a defective gene or defective gene fragment. A “defective gene” or “defective gene fragment” is a gene or portion of a gene that when expressed fails to generate a functioning protein or non-coding RNA with functionality of a the corresponding wild-type gene. In certain example embodiments, these defective genes may be associated with one or more disease phenotypes. In certain example embodiments, the defective gene or gene fragment is not replaced but the systems described herein are used to insert donor polynucleotides that encode gene or gene fragments that compensate for or override defective gene expression such that cell phenotypes associated with defective gene expression are eliminated or changed to a different or desired cellular phenotype.

[0517] In an embodiment of the invention, the donor polynucleotide may include, but not be limited to, genes or gene fragments, encoding proteins or RNA transcripts to be expressed, regulatory elements, repair templates, and the like. According to the invention, the donor polynucleotides may comprise left end and right end sequence elements that function with transposition components that mediate insertion.

[0518] In certain cases, the donor polynucleotide manipulates a splicing site on the target polynucleotide. In some examples, the donor polynucleotide disrupts a splicing site. The disruption may be achieved by inserting the polynucleotide to a splicing site and/or introducing one or more mutations to the splicing site. In certain examples, the donor polynucleotide may restore a splicing site. For example, the polynucleotide may comprise a splicing site sequence.

[0519] The donor polynucleotide to be inserted may has a size from 10 basepair or nucleotides to 50 kb in length, e.g., from 50 to 40k, from 100 and 30 k, from 100 to 10000, from 100 to 300, from 200 to 400, from 300 to 500, from 400 to 600, from 500 to 700, from 600 to 800, from 700 to 900, from 800 to 1000, from 900 to from 1100, from 1000 to 1200, from 1100 to 1300, from 1200 to 1400, from 1300 to 1500, from 1400 to 1600, from 1500 to 1700, from 600 to 1800, from 1700 to 1900, from 1800 to 2000 base pairs (bp) or nucleotides in length.

Inducible Systems

[0520] In one embodiment, a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease may form a component of an inducible system. The inducible nature of the system would allow for spatiotemporal control of gene editing or gene expression using a form of energy. The form of energy may include but is not limited to electromagnetic radiation, sound

energy, chemical energy and thermal energy. Examples of inducible system include tetracycline inducible promoters (Tet-On or Tet-Off), small molecule two-hybrid transcription activations systems (FKBP, ABA, etc.), or light inducible systems (Phytochrome, LOV domains, or cryptochrome). In one embodiment, the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease may be a part of a Light Inducible Transcriptional Effector (LITE) to direct changes in transcriptional activity in a sequence-specific manner. The components of a light may include a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, a light-responsive cytochrome heterodimer (e.g. from *Arabidopsis thaliana*), and a transcriptional activation/repression domain. Further examples of inducible DNA binding proteins and methods for their use are provided in US Provisional Application Nos. 61/736,465 and US 61/721,283, and International Patent Publication No. WO 2014/018423 A2 which is hereby incorporated by reference in its entirety.

Self-Inactivating Systems

[0521] Once all copies of a gene in the genome of a cell have been edited, continued expression of the system in that cell is no longer necessary. Indeed, sustained expression would be undesirable in case of off-target effects at unintended genomic sites, etc. Thus time-limited expression would be useful. Inducible expression offers one approach, but in addition Applicants have engineered a self-Inactivating system that relies on the use of a non-coding guide target sequence within the vector itself. Thus, after expression begins, the system will lead to its own destruction, but before destruction is complete it will have time to edit the genomic copies of the target gene (which, with a normal point mutation in a diploid cell, requires at most two edits). Simply, the self-inactivating system includes additional RNA (e.g., guide RNA) that targets the coding sequence for the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease itself or that targets one or more non-coding guide target sequences complementary to unique sequences present in one or more of the following: (a) within the promoter driving expression of the non-coding RNA elements, (b) within the promoter driving expression of the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease gene, (c) within 100bp of the ATG translational start codon in the IscB polypeptide nuclease coding sequence, (d) within the inverted terminal repeat (iTR) of a viral delivery vector, e.g., in the AAV genome.

[0522] In some aspects, a single ω RNA or gRNA is provided that is capable of hybridization to a sequence downstream of a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease start codon, whereby after a period of time there is a loss of the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease expression. In some aspects, one or more ω RNA or gRNA(s) are provided that are capable of hybridization to one or more coding or non-coding regions of the polynucleotide encoding the system, whereby after a period of time there is a inactivation of one or more, or in some cases all, of the system. In some aspects of the system, and not to be limited by theory, the cell may comprise a plurality of complexes, wherein a first subset of complexes comprise a first ω RNA or guide RNA capable of targeting a genomic locus or loci to be edited, and a second subset of complexes comprise at least one ω RNA or second guide RNA capable of targeting the polynucleotide encoding the system, wherein the first subset of complexes mediate editing of the targeted genomic locus or loci and the second subset of complexes eventually inactivate the system, thereby inactivating further expression in the cell.

[0523] The various coding sequences (IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and guide RNAs) can be included on a single vector or on multiple vectors. For instance, it is possible to encode the enzyme on one vector and the various RNA sequences on another vector, or to encode the enzyme and one ω RNA or guide RNA on one vector, and the remaining ω RNA or guide RNA on another vector, or any other permutation. In general, a system using a total of one or two different vectors is preferred.

[0524] Where multiple vectors are used, it is possible to deliver them in unequal numbers, and ideally with an excess of a vector which encodes the first ω RNA or guide RNA relative to the second ω RNA or guide RNA, thereby assisting in delaying final inactivation of the system until genome editing has had a chance to occur.

[0525] The first ω RNA or guide RNA can target any target sequence of interest within a genome, as described elsewhere herein. The second ω RNA or guide RNA targets a sequence within the vector which encodes the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, and thereby inactivates the enzyme's expression from that vector. Thus the target sequence in the vector must be capable of inactivating expression. Suitable target sequences can be, for instance, near to or within the translational start codon for the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease coding sequence, in a non-coding sequence in

the promoter driving expression of the non-coding RNA elements, within the promoter driving expression of the IscB polypeptide nuclease gene, within 100bp of the ATG translational start codon in the IscB polypeptide nuclease coding sequence, and/or within the inverted terminal repeat (iTR) of a viral delivery vector, e.g., in the AAV genome. A double stranded break near this region can induce a frame shift in the IscB polypeptide nuclease coding sequence, causing a loss of protein expression. An alternative target sequence for the “self-inactivating” ω RNA or guide RNA would aim to edit/inactivate regulatory regions/sequences needed for the expression of the system or for the stability of the vector. For instance, if the promoter for the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease coding sequence is disrupted then transcription can be inhibited or prevented. Similarly, if a vector includes sequences for replication, maintenance or stability then it is possible to target these. For instance, in a AAV vector a useful target sequence is within the iTR. Other useful sequences to target can be promoter sequences, polyadenylation sites, etc.

[0526] Furthermore, if the ω RNA or guide RNAs are expressed in array format, the “self-inactivating” ω RNA or guide RNAs that target both promoters simultaneously will result in the excision of the intervening nucleotides from within the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease expression construct, effectively leading to its complete inactivation. Similarly, excision of the intervening nucleotides will result where the ω RNA or guide RNAs target both ITRs, or targets two or more other components simultaneously. Self-inactivation as explained herein is applicable, in general, with systems in order to provide regulation of the systems. For example, self-inactivation as explained herein may be applied to the repair of mutations, for example expansion disorders, as explained herein. As a result of this self-inactivation, repair may be only transiently active.

[0527] Addition of non-targeting nucleotides to the 5' end (e.g. 1-10 nucleotides, preferably 1-5 nucleotides) of the “self-inactivating” ω RNA or guide RNA can be used to delay its processing and/or modify its efficiency as a means of ensuring editing at the targeted genomic locus prior to shut down.

[0528] In one aspect of the self-inactivating AAV system, plasmids that co-express one or more ω RNA or guide RNA targeting genomic sequences of interest (e.g. 1-2, 1-5, 1-10, 1-15, 1-20, 1-30) may be established with “self-inactivating” ω RNA or guide RNAs that target an IscB polypeptide or CRISPR-associated IscB polypeptide nuclease sequence at or near the

engineered ATG start site (e.g. within 5 nucleotides, within 15 nucleotides, within 30 nucleotides, within 50 nucleotides, within 100 nucleotides). A regulatory sequence in the U6 promoter region can also be targeted with an ω RNA or guide RNA. The U6-driven guide RNAs may be designed in an array format such that multiple ω RNA or guide RNA sequences can be simultaneously released. When first delivered into target tissue/cells (left cell) ω RNA or guide RNAs begin to accumulate while IscB polypeptide or CRISPR-associated IscB polypeptide nuclease levels rise in the nucleus. IscB polypeptide or CRISPR-associated IscB polypeptide nuclease complexes with all of the guide RNAs to mediate genome editing and self-inactivation of the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease plasmids.

[0529] One aspect of a self-inactivating system is expression of singly or in tandem array format from 1 up to 4 or more different ω RNA or guide sequences; e.g. up to about 20 or about 30 ω RNA or guide sequences. Each individual self-inactivating ω RNA or guide sequence may target a different target. Such may be processed from, e.g. one chimeric pol3 transcript. Pol3 promoters such as U6 or H1 promoters may be used. Pol2 promoters such as those mentioned throughout herein. Inverted terminal repeat (iTR) sequences may flank the Pol3 promoter - ω RNA or guide RNA(s)-Pol2 promoter- IscB polypeptide or CRISPR-associated IscB polypeptide nuclease.

[0530] One aspect of a tandem array transcript is that one or more ω RNA or guide(s) edit the one or more target(s) while one or more self-inactivating ω RNA or guides inactivate the system. Thus, for example, the described system for repairing expansion disorders may be directly combined with the self-inactivating system described herein. Such a system may, for example, have two ω RNA or guides directed to the target region for repair as well as at least a third ω RNA or guide directed to self-inactivation of the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease or systems.

[0531] The ω RNA or guide RNA may be a control guide. For example it may be engineered to target a nucleic acid sequence encoding the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease itself, as described in U.S. Patent Publication No. US2015232881A1, the disclosure of which is hereby incorporated by reference. In one embodiment, a system or composition may be provided with just the ω RNA or guide RNA engineered to target the nucleic acid sequence encoding the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease. In addition, the system or composition may be provided

with the ω RNA or guide RNA engineered to target the nucleic acid sequence encoding the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, as well as nucleic acid sequence encoding the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and, optionally a second ω RNA or guide RNA and, further optionally, a repair template. The second ω RNA or guide RNA may be the primary target of the system or composition (such a therapeutic, diagnostic, knock out etc. as defined herein). In this way, the system or composition is self-inactivating. This is exemplified in relation to Cas in US2015232881A1 (also published as WO2015070083 (A1) referenced elsewhere herein, and may be extrapolated to other IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, e.g. IscB polypeptides.

Polynucleotides

[0532] The systems herein may comprise one or more polynucleotides. The polynucleotide(s) may comprise coding sequences of components of the systems herein, e.g., IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, ω RNA(s), functional domain(s), donor polynucleotide(s), and/or other components in the systems. The present disclosure further provides vectors or vector systems comprising one or more polynucleotides herein. The vectors or vector systems include those described in the delivery sections herein.

[0533] The terms “polynucleotide”, “nucleotide”, “nucleotide sequence”, “nucleic acid” and “oligonucleotide” are used interchangeably. They refer to a polymeric form of nucleotides of any length, either deoxyribonucleotides or ribonucleotides, or analogs thereof. Polynucleotides may have any three-dimensional structure, and may perform any function, known or unknown. The following are non-limiting examples of polynucleotides: coding or non-coding regions of a gene or gene fragment, loci (locus) defined from linkage analysis, exons, introns, messenger RNA (mRNA), transfer RNA, ribosomal RNA, short interfering RNA (siRNA), short-hairpin RNA (shRNA), micro-RNA (miRNA), ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, nucleic acid probes, and primers. The term also encompasses nucleic-acid-like structures with synthetic backbones, see, e.g., Eckstein, 1991; Baserga et al., 1992; Milligan, 1993; WO 97/03211; WO 96/39154; Mata, 1997; Strauss-Soukup, 1997; and Samstag, 1996. A polynucleotide may comprise one or more modified nucleotides, such as methylated nucleotides and nucleotide analogs. If present, modifications

to the nucleotide structure may be imparted before or after assembly of the polymer. The sequence of nucleotides may be interrupted by non-nucleotide components. A polynucleotide may be further modified after polymerization, such as by conjugation with a labeling component. As used herein the term “wild type” is a term of the art understood by skilled persons and means the typical form of an organism, strain, gene or characteristic as it occurs in nature as distinguished from mutant or variant forms. A “wild type” can be a base line. As used herein the term “variant” should be taken to mean the exhibition of qualities that have a pattern that deviates from what occurs in nature. The terms “non-naturally occurring” or “engineered” are used interchangeably and indicate the involvement of the hand of man. The terms, when referring to nucleic acid molecules or polypeptides mean that the nucleic acid molecule or the polypeptide is at least substantially free from at least one other component with which they are naturally associated in nature and as found in nature. “Complementarity” refers to the ability of a nucleic acid to form hydrogen bond(s) with another nucleic acid sequence by either traditional Watson-Crick base pairing or other non-traditional types. A percent complementarity indicates the percentage of residues in a nucleic acid molecule which can form hydrogen bonds (e.g., Watson-Crick base pairing) with a second nucleic acid sequence (e.g., 5, 6, 7, 8, 9, 10 out of 10 being 50%, 60%, 70%, 80%, 90%, and 100% complementary). “Perfectly complementary” means that all the contiguous residues of a nucleic acid sequence will hydrogen bond with the same number of contiguous residues in a second nucleic acid sequence. “Substantially complementary” as used herein refers to a degree of complementarity that is at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100% over a region of 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, or more nucleotides, or refers to two nucleic acids that hybridize under stringent conditions. As used herein, “stringent conditions” for hybridization refer to conditions under which a nucleic acid having complementarity to a target sequence predominantly hybridizes with the target sequence, and substantially does not hybridize to non-target sequences. Stringent conditions are generally sequence-dependent, and vary depending on a number of factors. In general, the longer the sequence, the higher the temperature at which the sequence specifically hybridizes to its target sequence. Non-limiting examples of stringent conditions are described in detail in Tijssen (1993), *Laboratory Techniques In Biochemistry And Molecular Biology-Hybridization With Nucleic Acid Probes Part I*, Second Chapter “Overview of principles of hybridization and the strategy of nucleic acid probe assay”, Elsevier, N.Y. Where reference is

made to a polynucleotide sequence, then complementary or partially complementary sequences are also envisaged. These are preferably capable of hybridizing to the reference sequence under highly stringent conditions. "Hybridization" refers to a reaction in which one or more polynucleotides react to form a complex that is stabilized via hydrogen bonding between the bases of the nucleotide residues. The hydrogen bonding may occur by Watson Crick base pairing, Hoogsteen binding, or in any other sequence specific manner. The complex may comprise two strands forming a duplex structure, three or more strands forming a multi stranded complex, a single self-hybridizing strand, or any combination of these. A hybridization reaction may constitute a step in a more extensive process, such as the initiation of PCR, or the cleavage of a polynucleotide by an enzyme. A sequence capable of hybridizing with a given sequence is referred to as the "complement" of the given sequence. As used herein, the term "genomic locus" or "locus" (plural loci) is the specific location of a gene or DNA sequence on a chromosome. A "gene" refers to stretches of DNA or RNA that encode a polypeptide or an RNA chain that has functional role to play in an organism and hence is the molecular unit of heredity in living organisms. For the purpose of this invention, it may be considered that genes include regions which regulate the production of the gene product, whether or not such regulatory sequences are adjacent to coding and/or transcribed sequences. Accordingly, a gene includes, but is not necessarily limited to, promoter sequences, terminators, translational regulatory sequences such as ribosome binding sites and internal ribosome entry sites, enhancers, silencers, insulators, boundary elements, replication origins, matrix attachment sites and locus control regions. As used herein, "expression of a genomic locus" or "gene expression" is the process by which information from a gene is used in the synthesis of a functional gene product. The products of gene expression are often proteins, but in non-protein coding genes such as rRNA genes or tRNA genes, the product is functional RNA. The process of gene expression is used by all known life - eukaryotes (including multicellular organisms), prokaryotes (bacteria and archaea) and viruses to generate functional products to survive. As used herein "expression" of a gene or nucleic acid encompasses not only cellular gene expression, but also the transcription and translation of nucleic acid(s) in cloning systems and in any other context. As used herein, "expression" also refers to the process by which a polynucleotide is transcribed from a DNA template (such as into and mRNA or other RNA transcript) and/or the process by which a transcribed mRNA is subsequently translated into peptides, polypeptides, or proteins. Transcripts and encoded

polypeptides may be collectively referred to as “gene product.” If the polynucleotide is derived from genomic DNA, expression may include splicing of the mRNA in a eukaryotic cell. The terms “polypeptide”, “peptide” and “protein” are used interchangeably herein to refer to polymers of amino acids of any length. The polymer may be linear or branched, it may comprise modified amino acids, and it may be interrupted by non amino acids. The terms also encompass an amino acid polymer that has been modified; for example, disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation, such as conjugation with a labeling component. As used herein the term “amino acid” includes natural and/or unnatural or synthetic amino acids, including glycine and both the D or L optical isomers, and amino acid analogs and peptidomimetics. As used herein, the term “domain” or “protein domain” refers to a part of a protein sequence that may exist and function independently of the rest of the protein chain. As described in aspects of the invention, sequence identity is related to sequence homology. Homology comparisons may be conducted by eye, or more usually, with the aid of readily available sequence comparison programs. These commercially available computer programs may calculate percent (%) homology between two or more sequences and may also calculate the sequence identity shared by two or more amino acid or nucleic acid sequences.

[0534] In an embodiment, the polynucleotide sequence is recombinant DNA. In further embodiments, the polynucleotide sequence further comprises additional sequences as described elsewhere herein. In an embodiment, the nucleic acid sequence is synthesized *in vitro*.

[0535] The present disclosure provides polynucleotide molecules that encode one or more components of the system or IscB polypeptide nuclease as referred to in any embodiment herein. In an embodiment, the polynucleotide molecules may comprise further regulatory sequences. By means of guidance and not limitation, the polynucleotide sequence can be part of an expression plasmid, a minicircle, a lentiviral vector, a retroviral vector, an adenoviral or adeno-associated viral vector, a piggyback vector, or a tol2 vector. In an embodiment, the polynucleotide sequence may be a bicistronic expression construct. In further embodiments, the isolated polynucleotide sequence may be incorporated in a cellular genome. In yet further embodiments, the isolated polynucleotide sequence may be part of a cellular genome. In further embodiments, the isolated polynucleotide sequence may be comprised in an artificial chromosome. In an embodiment, the 5' and/or 3' end of the isolated polynucleotide sequence

may be modified to improve the stability of the sequence of actively avoid degradation. In an embodiment, the isolated polynucleotide sequence may be comprised in a bacteriophage. In other embodiments, the isolated polynucleotide sequence may be contained in agrobacterium species. In an embodiment, the isolated polynucleotide sequence is lyophilized.

[0536] Aspects of the invention relate to polynucleotide molecules that encode one or more components of one or more systems as described in any of the embodiments herein, wherein at least one or more regions of the polynucleotide molecule may be codon optimized for expression in a eukaryotic cells. In an embodiment, the polynucleotide molecules that encode one or more components of one or more systems as described in any of the embodiments herein are optimized for expression in a mammalian cell or a plant cell.

[0537] An example of a codon optimized sequence is in this instance a sequence optimized for expression in a eukaryote, e.g., humans (i.e. being optimized for expression in humans), or for another eukaryote, animal or mammal as herein discussed. In one embodiment, an enzyme coding sequence encoding a DNA/RNA-targeting IscB polypeptide or CRISPR-associated IscB polypeptide nuclease is codon optimized for expression in particular cells, such as eukaryotic cells. The eukaryotic cells may be those of or derived from a particular organism, such as a plant or a mammal, including but not limited to human, or non-human eukaryote or animal or mammal as herein discussed, e.g., mouse, rat, rabbit, dog, livestock, or non-human mammal or primate. In one embodiment, processes for modifying the germ line genetic identity of human beings and/or processes for modifying the genetic identity of animals which are likely to cause them suffering without any substantial medical benefit to man or animal, and also animals resulting from such processes, may be excluded. In general, codon optimization refers to a process of modifying a nucleic acid sequence for enhanced expression in the host cells of interest by replacing at least one codon (e.g., about or more than about 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, or more codons) of the native sequence with codons that are more frequently or most frequently used in the genes of that host cell while maintaining the native amino acid sequence. Various species exhibit particular bias for certain codons of a particular amino acid. Codon bias (differences in codon usage between organisms) often correlates with the efficiency of translation of messenger RNA (mRNA), which is in turn believed to be dependent on, among other things, the properties of the codons being translated and the availability of particular transfer RNA (tRNA) molecules. The predominance of selected tRNAs in a cell is generally a reflection of the codons used most frequently in peptide synthesis. Accordingly, genes can be

tailored for optimal gene expression in a given organism based on codon optimization. Codon usage tables are readily available, for example, at the “Codon Usage Database” available at www.kazusa.or.jp/codon/ and these tables can be adapted in a number of ways. See Nakamura, Y., et al. “Codon usage tabulated from the international DNA sequence databases: status for the year 2000” *Nucl. Acids Res.* 28:292 (2000). Computer algorithms for codon optimizing a particular sequence for expression in a particular host cell are also available, such as Gene Forge (Aptagen; Jacobus, PA), are also available. In one embodiment, one or more codons (e.g., 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, or more, or all codons) in a sequence encoding a IscB polypeptide nuclease corresponds to the most frequently used codon for a particular amino acid.

DELIVERY

[0538] The present disclosure also provides delivery systems for introducing components of the systems and compositions herein to cells, tissues, organs, or organisms. A delivery system may comprise one or more delivery vehicles and/or cargos. Exemplary delivery systems and methods include those described in paragraphs [00117] to [00278] of Feng Zhang et al., (WO2016106236A1), and pages 1241-1251 and Table 1 of Lino CA et al., *Delivering CRISPR: a review of the challenges and approaches*, *DRUG DELIVERY*, 2018, VOL. 25, NO. 1, 1234–1257, which are incorporated by reference herein in their entireties and can be adapted for use with the IscB proteins disclosed herein.

[0539] In one embodiment, the delivery systems may be used to introduce the components of the systems and compositions to plant cells. For example, the components may be delivered to plant using electroporation, microinjection, aerosol beam injection of plant cell protoplasts, biolistic methods, DNA particle bombardment, and/or *Agrobacterium*-mediated transformation. Examples of methods and delivery systems for plants include those described in Fu et al., *Transgenic Res.* 2000 Feb;9(1):11-9; Klein RM, et al., *Biotechnology*. 1992;24:384-6; Casas AM et al., *Proc Natl Acad Sci U S A.* 1993 Dec 1; 90(23): 11212–11216; and U.S. Pat. No. 5,563,055, Davey MR et al., *Plant Mol Biol.* 1989 Sep;13(3):273-85, which are incorporated by reference herein in their entireties.

[0540] The example delivery compositions, systems, and methods described herein related to composition or IscB polypeptide or CRISPR-associated IscB polypeptide nuclease also apply to functional domains and other components (e.g., other proteins and polynucleotides

related to the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, such as reverse transcriptase, nucleotide deaminase, retrotransposon, donor polynucleotide, etc.).

Cargos

[0541] The delivery systems may comprise one or more cargos. The cargos may comprise one or more components of the systems and compositions herein. A cargo may comprise one or more of the following: i) a plasmid encoding one or more proteins components in the compositions and systems such as the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and/or functional domains; ii) a plasmid encoding one or more hRNAs, iii) mRNA of one or more one or more proteins components in the compositions and systems such as the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and/or functional domains; iv) one or more guide RNAs; v) one or more proteins components in the compositions and systems such as the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and/or functional domains; vi) any combination thereof. The one or more protein components may include the nucleic acid-guided nuclease (e.g., Cas), reverse transcriptase, nucleotide deaminase, retrotransposon protein, other functional domain, or any combination thereof.

[0542] In some examples, a cargo may comprise a plasmid encoding one or more proteins components in the compositions and systems such as the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and/or functional domains and one or more (e.g., a plurality of) guide RNAs. In some cases, the plasmid may also encode a recombination template (e.g., for HDR). In one embodiment, a cargo may comprise mRNA encoding one or more protein components and one or more ω RNA or guide RNAs.

[0543] In some examples, a cargo may comprise one or more protein components and one or more ω RNA or guide RNAs, e.g., in the form of ribonucleoprotein complexes (RNP). The ribonucleoprotein complexes may be delivered by methods and systems herein. In some cases, the ribonucleoprotein may be delivered by way of a polypeptide-based shuttle agent. In one example, the ribonucleoprotein may be delivered using synthetic peptides comprising an endosome leakage domain (ELD) operably linked to a cell penetrating domain (CPD), to a histidine-rich domain and a CPD, e.g., as describe in WO2016161516. RNP may also be used for delivering the compositions and systems to plant cells, e.g., as described in Wu JW, et al., Nat Biotechnol. 2015 Nov;33(11):1162-4.

Physical delivery

[0544] In one embodiment, the cargos may be introduced to cells by physical delivery methods. Examples of physical methods include microinjection, electroporation, and hydrodynamic delivery. Both nucleic acid and proteins may be delivered using such methods. For example, one or more protein components may be prepared *in vitro*, isolated, (refolded, purified if needed), and introduced to cells.

Microinjection

[0545] Microinjection of the cargo directly to cells can achieve high efficiency, e.g., above 90% or about 100%. In one embodiment, microinjection may be performed using a microscope and a needle (e.g., with 0.5–5.0 μm in diameter) to pierce a cell membrane and deliver the cargo directly to a target site within the cell. Microinjection may be used for *in vitro* and *ex vivo* delivery.

[0546] Plasmids comprising coding sequences for one or more protein components and/or ω RNAs, mRNAs, and/or guide RNAs, may be microinjected. In some cases, microinjection may be used i) to deliver DNA directly to a cell nucleus, and/or ii) to deliver mRNA (e.g., *in vitro* transcribed) to a cell nucleus or cytoplasm. In certain examples, microinjection may be used to delivery ω RNA directly to the nucleus and mRNA to the cytoplasm, e.g., facilitating translation and shuttling of one or more protein components to the nucleus.

[0547] Microinjection may be used to generate genetically modified animals. For example, gene editing cargos may be injected into zygotes to allow for efficient germline modification. Such approach can yield normal embryos and full-term mouse pups harboring the desired modification(s). Microinjection can also be used to provide transiently up- or down- regulate a specific gene within the genome of a cell, e.g., using IscB polypeptide or CRISPR-associated IscB polypeptide.

Electroporation

[0548] In one embodiment, the cargos and/or delivery vehicles may be delivered by electroporation. Electroporation may use pulsed high-voltage electrical currents to transiently open nanometer-sized pores within the cellular membrane of cells suspended in buffer, allowing for components with hydrodynamic diameters of tens of nanometers to flow into the cell. In some cases, electroporation may be used on various cell types and efficiently transfer cargo into cells. Electroporation may be used for *in vitro* and *ex vivo* delivery.

[0549] Electroporation may also be used to deliver the cargo to into the nuclei of mammalian cells by applying specific voltage and reagents, e.g., by nucleofection. Such approaches include those described in Wu Y, et al. (2015). Cell Res 25:67–79; Ye L, et al. (2014). Proc Natl Acad Sci USA 111:9591–6; Choi PS, Meyerson M. (2014). Nat Commun 5:3728; Wang J, Quake SR. (2014). Proc Natl Acad Sci 111:13157–62. Electroporation may also be used to deliver the cargo *in vivo*, e.g., with methods described in Zuckermann M, et al. (2015). Nat Commun 6:7391.

Hydrodynamic delivery

[0550] Hydrodynamic delivery may also be used for delivering the cargos, e.g., for *in vivo* delivery. In some examples, hydrodynamic delivery may be performed by rapidly pushing a large volume (8–10% body weight) solution containing the gene editing cargo into the bloodstream of a subject (e.g., an animal or human), e.g., for mice, via the tail vein. As blood is incompressible, the large bolus of liquid may result in an increase in hydrodynamic pressure that temporarily enhances permeability into endothelial and parenchymal cells, allowing for cargo not normally capable of crossing a cellular membrane to pass into cells. This approach may be used for delivering naked DNA plasmids and proteins. The delivered cargos may be enriched in liver, kidney, lung, muscle, and/or heart.

Transfection

[0551] The cargos, e.g., nucleic acids, may be introduced to cells by transfection methods for introducing nucleic acids into cells. Examples of transfection methods include calcium phosphate-mediated transfection, cationic transfection, liposome transfection, dendrimer transfection, heat shock transfection, magnetofection, lipofection, impalefection, optical transfection, proprietary agent-enhanced uptake of nucleic acid.

Delivery vehicles

[0552] The delivery systems may comprise one or more delivery vehicles. The delivery vehicles may deliver the cargo into cells, tissues, organs, or organisms (e.g., animals or plants). The cargos may be packaged, carried, or otherwise associated with the delivery vehicles. The delivery vehicles may be selected based on the types of cargo to be delivered, and/or the delivery is *in vitro* and/or *in vivo*. Examples of delivery vehicles include vectors, viruses, non-viral vehicles, and other delivery reagents described herein.

[0553] The delivery vehicles in accordance with the present invention may have a greatest dimension (e.g. diameter) of less than 100 microns (μm). In one embodiment, the delivery

vehicles have a greatest dimension of less than 10 μm . In one embodiment, the delivery vehicles may have a greatest dimension of less than 2000 nanometers (nm). In one embodiment, the delivery vehicles may have a greatest dimension of less than 1000 nanometers (nm). In one embodiment, the delivery vehicles may have a greatest dimension (e.g., diameter) of less than 900 nm, less than 800 nm, less than 700 nm, less than 600 nm, less than 500 nm, less than 400 nm, less than 300 nm, less than 200 nm, less than 150nm, or less than 100nm, less than 50nm. In one embodiment, the delivery vehicles may have a greatest dimension ranging between 25 nm and 200 nm.

[0554] In one embodiment, the delivery vehicles may be or comprise particles. For example, the delivery vehicle may be or comprise nanoparticles (e.g., particles with a greatest dimension (e.g., diameter) no greater than 1000nm. The particles may be provided in different forms, e.g., as solid particles (e.g., metal such as silver, gold, iron, titanium), non-metal, lipid-based solids, polymers), suspensions of particles, or combinations thereof. Metal, dielectric, and semiconductor particles may be prepared, as well as hybrid structures (e.g., core-shell particles). Nanoparticles may also be used to deliver the compositions and systems to plant cells, e.g., as described in International Patent Publication No. WO 2008042156, US Publication Application No. US 20130185823, and International Patent Publication No WO 2015/089419.

Vectors

[0555] The systems, compositions, and/or delivery systems may comprise one or more vectors. The present disclosure also include vector systems. A vector system may comprise one or more vectors. In one embodiment, a vector refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. Vectors include nucleic acid molecules that are single-stranded, double-stranded, or partially double-stranded; nucleic acid molecules that comprise one or more free ends, no free ends (e.g., circular); nucleic acid molecules that comprise DNA, RNA, or both; and other varieties of polynucleotides known in the art. A vector may be a plasmid, e.g., a circular double stranded DNA loop into which additional DNA segments can be inserted, such as by standard molecular cloning techniques. Certain vectors may be capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Some vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along

with the host genome. In certain examples, vectors may be expression vectors, e.g., capable of directing the expression of genes to which they are operatively-linked. In some cases, the expression vectors may be for expression in eukaryotic cells. Common expression vectors of utility in recombinant DNA techniques are often in the form of plasmids.

[0556] Examples of vectors include pGEX, pMAL, pRIT5, E. coli expression vectors (e.g., pTrc, pET 11d, yeast expression vectors (e.g., pYepSec1, pMFa, pJRY88, pYES2, and picZ, Baculovirus vectors (e.g., for expression in insect cells such as SF9 cells) (e.g., pAc series and the pVL series), mammalian expression vectors (e.g., pCDM8 and pMT2PC.

[0557] A vector may comprise i) one or more protein components encoding sequence(s), and/or ii) a single, or at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 12, at least 14, at least 16, at least 32, at least 48, at least 50 guide RNA(s) encoding sequences. In a single vector there can be a promoter for each RNA coding sequence. Alternatively or additionally, in a single vector, there may be a promoter controlling (e.g., driving transcription and/or expression) multiple RNA encoding sequences.

[0558] Furthermore, that compositions or systems may be delivered via a vector, e.g., a separate vector or the same vector that is encoding the complex. When provided by a separate vector, the RNA that targets IscB polypeptide or CRISPR-associated IscB polypeptide nuclease expression can be administered sequentially or simultaneously. When administered sequentially, the RNA that targets IscB polypeptide or CRISPR-associated IscB polypeptide nuclease expression is to be delivered after the RNA that is intended for e.g. gene editing or gene engineering. This period may be a period of minutes (e.g. 5 minutes, 10 minutes, 20 minutes, 30 minutes, 45 minutes, 60 minutes). This period may be a period of hours (e.g. 2 hours, 4 hours, 6 hours, 8 hours, 12 hours, 24 hours). This period may be a period of days (e.g. 2 days, 3 days, 4 days, 7 days). This period may be a period of weeks (e.g. 2 weeks, 3 weeks, 4 weeks). This period may be a period of months (e.g. 2 months, 4 months, 8 months, 12 months). This period may be a period of years (2 years, 3 years, 4 years). In this fashion, the IscB polypeptide nuclease associates with a first hRNA molecule capable of hybridizing to a first target, such as a genomic locus or loci of interest and undertakes the function(s) desired of the system (e.g., gene engineering); and subsequently the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease may then associate with the second hRNA molecule capable of hybridizing to the sequence comprising at least part of the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease. Where the guide RNA targets the sequences

encoding expression of the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, the enzyme becomes impeded and the system becomes self-inactivating. In the same manner, RNA that targets IscB polypeptide or CRISPR-associated IscB polypeptidenuclease expression applied via, for example liposome, lipofection, particles, microvesicles as explained herein, may be administered sequentially or simultaneously. Similarly, self-inactivation may be used for inactivation of one or more guide RNA used to target one or more targets.

Regulatory elements

[0559] A vector may comprise one or more regulatory elements. The regulatory element(s) may be operably linked to coding sequences of IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, accessory proteins, ω RNA scaffold and/or guide RNA or combination thereof. The term “operably linked” is intended to mean that the nucleotide sequence of interest is linked to the regulatory element(s) in a manner that allows for expression of the nucleotide sequence (e.g. in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). In certain examples, a vector may comprise: a first regulatory element operably linked to a nucleotide sequence encoding a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, and a second regulatory element operably linked to a nucleotide sequence encoding a ω RNA or guide RNA.

[0560] Examples of regulatory elements include promoters, enhancers, internal ribosomal entry sites (IRES), and other expression control elements (e.g., transcription termination signals, such as polyadenylation signals and poly-U sequences). Such regulatory elements are described, for example, in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Regulatory elements include those that direct constitutive expression of a nucleotide sequence in many types of host cell and those that direct expression of the nucleotide sequence only in certain host cells (e.g., tissue-specific regulatory sequences). A tissue-specific promoter may direct expression primarily in a desired tissue of interest, such as muscle, neuron, bone, skin, blood, specific organs (e.g., liver, pancreas), or particular cell types (e.g., lymphocytes). Regulatory elements may also direct expression in a temporal-dependent manner, such as in a cell-cycle dependent or developmental stage-dependent manner, which may or may not also be tissue or cell-type specific.

[0561] Examples of promoters include one or more pol III promoter (e.g., 1, 2, 3, 4, 5, or more pol III promoters), one or more pol II promoters (e.g., 1, 2, 3, 4, 5, or more pol II promoters), one or more pol I promoters (e.g., 1, 2, 3, 4, 5, or more pol I promoters), or combinations thereof. Examples of pol III promoters include, but are not limited to, U6 and H1 promoters. Examples of pol II promoters include, but are not limited to, the retroviral Rous sarcoma virus (RSV) LTR promoter (optionally with the RSV enhancer), the cytomegalovirus (CMV) promoter (optionally with the CMV enhancer), the SV40 promoter, the dihydrofolate reductase promoter, the β -actin promoter, the phosphoglycerol kinase (PGK) promoter, and the EF1 α promoter.

Viral vectors

[0562] The cargos may be delivered by viruses. In one embodiment, viral vectors are used. A viral vector may comprise virally-derived DNA or RNA sequences for packaging into a virus (e.g., retroviruses, replication defective retroviruses, adenoviruses, replication defective adenoviruses, and adeno-associated viruses). Viral vectors also include polynucleotides carried by a virus for transfection into a host cell. Viruses and viral vectors may be used for *in vitro*, *ex vivo*, and/or *in vivo* deliveries.

Adeno associated virus (AAV)

[0563] The systems and compositions herein may be delivered by adeno associated virus (AAV). AAV vectors may be used for such delivery. AAV, of the Dependovirus genus and Parvoviridae family, is a single stranded DNA virus. In one embodiment, AAV may provide a persistent source of the provided DNA, as AAV delivered genomic material can exist indefinitely in cells, e.g., either as exogenous DNA or, with some modification, be directly integrated into the host DNA. In one embodiment, AAV do not cause or relate with any diseases in humans. The virus itself is able to efficiently infect cells while provoking little to no innate or adaptive immune response or associated toxicity.

[0564] Examples of AAV that can be used herein include AAV-1, AAV-2, AAV-3, AAV-4, AAV-5, AAV-6, AAV-8, and AAV-9. The type of AAV may be selected with regard to the cells to be targeted; e.g., one can select AAV serotypes 1, 2, 5 or a hybrid capsid AAV1, AAV2, AAV5 or any combination thereof for targeting brain or neuronal cells; and one can select AAV4 for targeting cardiac tissue. AAV8 is useful for delivery to the liver. AAV-2-based vectors were originally proposed for CFTR delivery to CF airways, other serotypes such as AAV-1, AAV-5, AAV-6, and AAV-9 exhibit improved gene transfer efficiency in a variety of

models of the lung epithelium. Examples of cell types targeted by AAV are described in Grimm, D. et al, J. Virol. 82: 5887-5911 (2008)), and shown in Table 5 as follows:

Table 5. Examples of cell types targeted by AAV.

Cell Line	AAV-1	AAV-2	AAV-3	AAV-4	AAV-5	AAV-6	AAV-8	AAV-9
Huh-7	13	100	2.5	0.0	0.1	10	0.7	0.0
HEK293	25	100	2.5	0.1	0.1	5	0.7	0.1
HeLa	3	100	2.0	0.1	6.7	1	0.2	0.1
HepG2	3	100	16.7	0.3	1.7	5	0.3	ND
Hep1A	20	100	0.2	1.0	0.1	1	0.2	0.0
911	17	100	11	0.2	0.1	17	0.1	ND
CHO	100	100	14	1.4	333	50	10	1.0
COS	33	100	33	3.3	5.0	14	2.0	0.5
MeWo	10	100	20	0.3	6.7	10	1.0	0.2
NIH3T3	10	100	2.9	2.9	0.3	10	0.3	ND
A549	14	100	20	ND	0.5	10	0.5	0.1
HT1180	20	100	10	0.1	0.3	33	0.5	0.1
Monocytes	1111	100	ND	ND	125	1429	ND	ND
Immature DC	2500	100	ND	ND	222	2857	ND	ND
Mature DC	2222	100	ND	ND	333	3333	ND	ND

[0565] The AAV particles may be created in HEK 293 T cells. Once particles with specific tropism have been created, they are used to infect the target cell line much in the same way that native viral particles do. This may allow for persistent presence of the components in the infected cell type, and what makes this version of delivery particularly suited to cases where long-term expression is desirable. Examples of doses and formulations for AAV that can be used include those describe in US Patent Nos. 8,454,972 and 8,404,658.

[0566] Various strategies may be used for delivery the systems and compositions herein with AAVs. In some examples, coding sequences of IscB polypeptide nuclease and ω RNA may be packaged directly onto one DNA plasmid vector and delivered via one AAV particle. In some examples, AAVs may be used to deliver ω RNAs into cells that have been previously engineered to express IscB polypeptide or CRISPR-associated IscB polypeptide nuclease. In some examples, coding sequences of IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and ω RNA may be made into two separate AAV particles, which are used for co-transfection of target cells. In some examples, markers, tags, and other sequences may be

packaged in the same AAV particles as coding sequences of IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and/or ω RNAs.

Lentiviruses

[0567] The systems and compositions herein may be delivered by lentiviruses. Lentiviral vectors may be used for such delivery. Lentiviruses are complex retroviruses that have the ability to infect and express their genes in both mitotic and post-mitotic cells.

[0568] Examples of lentiviruses include human immunodeficiency virus (HIV), which may use its envelope glycoproteins of other viruses to target a broad range of cell types; minimal non-primate lentiviral vectors based on the equine infectious anemia virus (EIAV), which may be used for ocular therapies. In an embodiment, self-inactivating lentiviral vectors with an siRNA targeting a common exon shared by HIV tat/rev, a nucleolar-localizing TAR decoy, and an anti-CCR5-specific hammerhead ribozyme (see, e.g., DiGiusto et al. (2010) *Sci Transl Med* 2:36ra43) may be used/and or adapted to the nucleic acid-targeting system herein.

[0569] Lentiviruses may be pseudo-typed with other viral proteins, such as the G protein of vesicular stomatitis virus. In doing so, the cellular tropism of the lentiviruses can be altered to be as broad or narrow as desired. In some cases, to improve safety, second- and third-generation lentiviral systems may split essential genes across three plasmids, which may reduce the likelihood of accidental reconstitution of viable viral particles within cells.

[0570] In some examples, leveraging the integration ability, lentiviruses may be used to create libraries of cells comprising various genetic modifications, e.g., for screening and/or studying genes and signaling pathways.

Adenoviruses

[0571] The systems and compositions herein may be delivered by adenoviruses. Adenoviral vectors may be used for such delivery. Adenoviruses include nonenveloped viruses with an icosahedral nucleocapsid containing a double stranded DNA genome. Adenoviruses may infect dividing and non-dividing cells. In one embodiment, adenoviruses do not integrate into the genome of host cells, which may be used for limiting off-target effects of systems in gene editing applications.

Viral vehicles for delivery to plants

[0572] The systems and compositions may be delivered to plant cells using viral vehicles. In one embodiment, the compositions and systems may be introduced in the plant cells using

a plant viral vector (e.g., as described in Scholthof et al. 1996, *Annu Rev Phytopathol.* 1996;34:299-323). Such viral vector may be a vector from a DNA virus, e.g., geminivirus (e.g., cabbage leaf curl virus, bean yellow dwarf virus, wheat dwarf virus, tomato leaf curl virus, maize streak virus, tobacco leaf curl virus, or tomato golden mosaic virus) or nanovirus (e.g., Faba bean necrotic yellow virus). The viral vector may be a vector from an RNA virus, e.g., tobnavirus (e.g., tobacco rattle virus, tobacco mosaic virus), potexvirus (e.g., potato virus X), or hordevirus (e.g., barley stripe mosaic virus). The replicating genomes of plant viruses may be non-integrative vectors.

Non-viral vehicles

[0573] The delivery vehicles may comprise non-viral vehicles. In general, methods and vehicles capable of delivering nucleic acids and/or proteins may be used for delivering the systems compositions herein. Examples of non-viral vehicles include lipid nanoparticles, cell-penetrating peptides (CPPs), DNA nanoclews, gold nanoparticles, streptolysin O, multifunctional envelope-type nanodevices (MENDs), lipid-coated mesoporous silica particles, and other inorganic nanoparticles.

Lipid particles

[0574] The delivery vehicles may comprise lipid particles, e.g., lipid nanoparticles (LNPs) and liposomes.

Lipid nanoparticles (LNPs)

[0575] LNPs may encapsulate nucleic acids within cationic lipid particles (e.g., liposomes), and may be delivered to cells with relative ease. In some examples, lipid nanoparticles do not contain any viral components, which helps minimize safety and immunogenicity concerns. Lipid particles may be used for *in vitro*, *ex vivo*, and *in vivo* deliveries. Lipid particles may be used for various scales of cell populations.

[0576] In some examples, LNPs may be used for delivering DNA molecules (e.g., those comprising coding sequences of IscB polypeptide nuclease and/or hRNA) and/or RNA molecules (e.g., mRNA of IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, ωRNA or gRNAs). In certain cases, LNPs may be used for delivering RNP complexes of IscB polypeptide or CRISPR-associated IscB polypeptide /ωRNA.

[0577] Components in LNPs may comprise cationic lipids 1,2- dilineoyl-3-dimethylammonium-propane (DLinDAP), 1,2-dilinoleyloxy-3-N,N- dimethylaminopropane (DLinDMA), 1,2-dilinoleyloxyketo-N,N-dimethyl-3-aminopropane (DLinK-DMA), 1,2-

dilinoleyl-4-(2-dimethylaminoethyl)-[1,3]-dioxolane (DLinKC2-DMA), (3-*o*-[2"-(methoxypolyethyleneglycol 2000) succinoyl]-1,2-dimyristoyl-sn-glycol (PEG-S-DMG), R-3-[(*ro*-methoxy-poly(ethylene glycol)2000) carbamoyl]-1,2-dimyristyloxylpropyl-3-amine (PEG-C-DMG, and any combination thereof. Preparation of LNPs and encapsulation may be adapted from Rosin et al, Molecular Therapy, vol. 19, no. 12, pages 1286-2200, Dec. 2011).

Liposomes

[0578] In one embodiment, a lipid particle may be liposome. Liposomes are spherical vesicle structures composed of a uni- or multilamellar lipid bilayer surrounding internal aqueous compartments and a relatively impermeable outer lipophilic phospholipid bilayer. In one embodiment, liposomes are biocompatible, nontoxic, can deliver both hydrophilic and lipophilic drug molecules, protect their cargo from degradation by plasma enzymes, and transport their load across biological membranes and the blood brain barrier (BBB).

[0579] Liposomes can be made from several different types of lipids, e.g., phospholipids. A liposome may comprise natural phospholipids and lipids such as 1,2-distearoyl-sn-glycero-3-phosphatidyl choline (DSPC), sphingomyelin, egg phosphatidylcholines, monosialoganglioside, or any combination thereof.

[0580] Several other additives may be added to liposomes in order to modify their structure and properties. For instance, liposomes may further comprise cholesterol, sphingomyelin, and/or 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), e.g., to increase stability and/or to prevent the leakage of the liposomal inner cargo.

Stable nucleic-acid-lipid particles (SNALPs)

[0581] In one embodiment, the lipid particles may be stable nucleic acid lipid particles (SNALPs). SNALPs may comprise an ionizable lipid (DLinDMA) (e.g., cationic at low pH), a neutral helper lipid, cholesterol, a diffusible polyethylene glycol (PEG)-lipid, or any combination thereof. In some examples, SNALPs may comprise synthetic cholesterol, dipalmitoylphosphatidylcholine, 3-N-[(*w*-methoxy polyethylene glycol)2000]carbamoyl]-1,2-dimyrestyloxypropylamine, and cationic 1,2-dilinoleyloxy-3-N,Ndimethylaminopropane. In some examples, SNALPs may comprise synthetic cholesterol, 1,2-distearoyl-sn-glycero-3-phosphocholine, PEG-cDMA, and 1,2-dilinoleyloxy-3-(N,N-dimethyl)aminopropane (DLinDMA)

Other lipids

[0582] The lipid particles may also comprise one or more other types of lipids, e.g., cationic lipids, such as amino lipid 2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]- dioxolane (DLin-KC2-DMA), DLin-KC2-DMA4, C12- 200 and colipids distearylphosphatidyl choline, cholesterol, and PEG-DMG.

Lipoplexes/polyplexes

[0583] In one embodiment, the delivery vehicles comprise lipoplexes and/or polyplexes. Lipoplexes may bind to negatively charged cell membrane and induce endocytosis into the cells. Examples of lipoplexes may be complexes comprising lipid(s) and non-lipid components. Examples of lipoplexes and polyplexes include FuGENE-6 reagent, a non-liposomal solution containing lipids and other components, zwitterionic amino lipids (ZALs), Ca²⁺ (e.g., forming DNA/Ca²⁺ microcomplexes), polyethenimine (PEI) (e.g., branched PEI), and poly(L-lysine) (PLL).

Cell penetrating peptides

[0584] In one embodiment, the delivery vehicles comprise cell penetrating peptides (CPPs). CPPs are short peptides that facilitate cellular uptake of various molecular cargo (e.g., from nanosized particles to small chemical molecules and large fragments of DNA).

[0585] CPPs may be of different sizes, amino acid sequences, and charges. In some examples, CPPs can translocate the plasma membrane and facilitate the delivery of various molecular cargoes to the cytoplasm or an organelle. CPPs may be introduced into cells via different mechanisms, e.g., direct penetration in the membrane, endocytosis-mediated entry, and translocation through the formation of a transitory structure.

[0586] CPPs may have an amino acid composition that either contains a high relative abundance of positively charged amino acids such as lysine or arginine or has sequences that contain an alternating pattern of polar/charged amino acids and non-polar, hydrophobic amino acids. These two types of structures are referred to as polycationic or amphipathic, respectively. A third class of CPPs are the hydrophobic peptides, containing only apolar residues, with low net charge or have hydrophobic amino acid groups that are crucial for cellular uptake. Another type of CPPs is the trans-activating transcriptional activator (Tat) from Human Immunodeficiency Virus 1 (HIV-1). Examples of CPPs include to Penetratin, Tat (48-60), Transportan, and (R-Ahx-R4) (Ahx refers to aminohexanoyl), Kaposi fibroblast growth factor (FGF) signal peptide sequence, integrin β 3 signal peptide sequence, polyarginine peptide Args

sequence, Guanine rich-molecular transporters, and sweet arrow peptide. Examples of CPPs and related applications also include those described in US Patent No. 8,372,951.

[0587] CPPs can be used for *in vitro* and *ex vivo* work quite readily, and extensive optimization for each cargo and cell type is usually required. In some examples, CPPs may be covalently attached to the IscB polypeptide nuclease directly, which is then complexed with the hRNA and delivered to cells. In some examples, separate delivery of CPP–IscB and CPP–hRNA to multiple cells may be performed. CPP may also be used to delivery RNPs.

[0588] CPPs may be used to deliver the compositions and systems to plants. In some examples, CPPs may be used to deliver the components to plant protoplasts, which are then regenerated to plant cells and further to plants.

DNA nanoclews

[0589] In one embodiment, the delivery vehicles comprise DNA nanoclews. A DNA nanoclew refers to a sphere-like structure of DNA (e.g., with a shape of a ball of yarn). The nanoclew may be synthesized by rolling circle amplification with palindromic sequences that aide in the self-assembly of the structure. The sphere may then be loaded with a payload. An example of DNA nanoclew is described in Sun W et al, J Am Chem Soc. 2014 Oct 22;136(42):14722-5; and Sun W et al, Angew Chem Int Ed Engl. 2015 Oct 5;54(41):12029-33. DNA nanoclew may have a palindromic sequences to be partially complementary to the guide RNA within the IscB polypeptide nuclease:hRNA ribonucleoprotein complex. A DNA nanoclew may be coated, e.g., coated with PEI to induce endosomal escape.

Gold nanoparticles

[0590] In one embodiment, the delivery vehicles comprise gold nanoparticles (also referred to AuNPs or colloidal gold). Gold nanoparticles may form complex with cargos, e.g., IscB polypeptide nuclease:hRNA RNP. Gold nanoparticles may be coated, e.g., coated in a silicate and an endosomal disruptive polymer, PAsp(DET). Examples of gold nanoparticles include AuraSense Therapeutics' Spherical Nucleic Acid (SNA™) constructs, and those described in Mout R, et al. (2017). ACS Nano 11:2452–8; Lee K, et al. (2017). Nat Biomed Eng 1:889–901.

iTOP

[0591] In one embodiment, the delivery vehicles comprise iTOP. iTOP refers to a combination of small molecules drives the highly efficient intracellular delivery of native proteins, independent of any transduction peptide. iTOP may be used for induced transduction

by osmocytosis and propanebetaine, using NaCl-mediated hyperosmolality together with a transduction compound (propanebetaine) to trigger macropinocytotic uptake into cells of extracellular macromolecules. Examples of iTOP methods and reagents include those described in D'Astolfo DS, Pagliero RJ, Pras A, et al. (2015). *Cell* 161:674–690.

Polymer-based particles

[0592] In one embodiment, the delivery vehicles may comprise polymer-based particles (e.g., nanoparticles). In one embodiment, the polymer-based particles may mimic a viral mechanism of membrane fusion. The polymer-based particles may be a synthetic copy of Influenza virus machinery and form transfection complexes with various types of nucleic acids ((siRNA, miRNA, plasmid DNA or shRNA, mRNA) that cells take up via the endocytosis pathway, a process that involves the formation of an acidic compartment. The low pH in late endosomes acts as a chemical switch that renders the particle surface hydrophobic and facilitates membrane crossing. Once in the cytosol, the particle releases its payload for cellular action. This Active Endosome Escape technology is safe and maximizes transfection efficiency as it is using a natural uptake pathway. In one embodiment, the polymer-based particles may comprise alkylated and carboxyalkylated branched polyethylenimine. In some examples, the polymer-based particles are VIROMER, e.g., VIROMER RNAi, VIROMER RED, VIROMER mRNA. Example methods of delivering the systems and compositions herein include those described in Bawage SS et al., Synthetic mRNA expressed Cas13a mitigates RNA virus infections, www.biorxiv.org/content/10.1101/370460v1.full doi: doi.org/10.1101/370460, Viromer® RED, a powerful tool for transfection of keratinocytes. doi: [10.13140/RG.2.2.16993.61281](https://doi.org/10.13140/RG.2.2.16993.61281), Viromer® Transfection - Factbook 2018: technology, product overview, users' data., doi: [10.13140/RG.2.2.23912.16642](https://doi.org/10.13140/RG.2.2.23912.16642).

Streptolysin O (SLO)

[0593] The delivery vehicles may be streptolysin O (SLO). SLO is a toxin produced by Group A streptococci that works by creating pores in mammalian cell membranes. SLO may act in a reversible manner, which allows for the delivery of proteins (e.g., up to 100 kDa) to the cytosol of cells without compromising overall viability. Examples of SLO include those described in Sierig G, et al. (2003). *Infect Immun* 71:446–55; Walev I, et al. (2001). *Proc Natl Acad Sci U S A* 98:3185–90; Teng KW, et al. (2017). *Elife* 6:e25460.

Multifunctional envelope-type nanodevice (MEND)

[0594] The delivery vehicles may comprise multifunctional envelope-type nanodevice (MENDs). MENDs may comprise condensed plasmid DNA, a PLL core, and a lipid film shell. A MEND may further comprise cell-penetrating peptide (e.g., stearyl octaarginine). The cell penetrating peptide may be in the lipid shell. The lipid envelope may be modified with one or more functional components, e.g., one or more of: polyethylene glycol (e.g., to increase vascular circulation time), ligands for targeting of specific tissues/cells, additional cell-penetrating peptides (e.g., for greater cellular delivery), lipids to enhance endosomal escape, and nuclear delivery tags. In some examples, the MEND may be a tetra-lamellar MEND (T-MEND), which may target the cellular nucleus and mitochondria. In certain examples, a MEND may be a PEG-peptide-DOPE-conjugated MEND (PPD-MEND), which may target bladder cancer cells. Examples of MENDs include those described in Kogure K, et al. (2004). J Control Release 98:317–23; Nakamura T, et al. (2012). Acc Chem Res 45:1113–21.

Lipid-coated mesoporous silica particles

[0595] The delivery vehicles may comprise lipid-coated mesoporous silica particles. Lipid-coated mesoporous silica particles may comprise a mesoporous silica nanoparticle core and a lipid membrane shell. The silica core may have a large internal surface area, leading to high cargo loading capacities. In one embodiment, pore sizes, pore chemistry, and overall particle sizes may be modified for loading different types of cargos. The lipid coating of the particle may also be modified to maximize cargo loading, increase circulation times, and provide precise targeting and cargo release. Examples of lipid-coated mesoporous silica particles include those described in Du X, et al. (2014). Biomaterials 35:5580–90; Durfee PN, et al. (2016). ACS Nano 10:8325–45.

Inorganic nanoparticles

[0596] The delivery vehicles may comprise inorganic nanoparticles. Examples of inorganic nanoparticles include carbon nanotubes (CNTs) (e.g., as described in Bates K and Kostarelos K. (2013). Adv Drug Deliv Rev 65:2023–33.), bare mesoporous silica nanoparticles (MSNPs) (e.g., as described in Luo GF, et al. (2014). Sci Rep 4:6064), and dense silica nanoparticles (SiNPs) (as described in Luo D and Saltzman WM. (2000). Nat Biotechnol 18:893–5).

Exosomes

[0597] The delivery vehicles may comprise exosomes. Exosomes include membrane bound extracellular vesicles, which can be used to contain and delivery various types of biomolecules, such as proteins, carbohydrates, lipids, and nucleic acids, and complexes thereof (e.g., RNPs). Examples of exosomes include those described in Schroeder A, et al., J Intern Med. 2010 Jan;267(1):9-21; El-Andaloussi S, et al., Nat Protoc. 2012 Dec;7(12):2112-26; Uno Y, et al., Hum Gene Ther. 2011 Jun;22(6):711-9; Zou W, et al., Hum Gene Ther. 2011 Apr;22(4):465-75.

[0598] In some examples, the exosome may form a complex (e.g., by binding directly or indirectly) to one or more components of the cargo. In certain examples, a molecule of an exosome may be fused with first adapter protein and a component of the cargo may be fused with a second adapter protein. The first and the second adapter protein may specifically bind each other, thus associating the cargo with the exosome. Examples of such exosomes include those described in Ye Y, et al., Biomater Sci. 2020 Apr 28. doi: 10.1039/d0bm00427h.

GENETICALLY MODIFIED CELLS AND ORGANISMS

[0599] The present disclosure further provides cells comprising one or more components of the compositions and systems herein, e.g., the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and/or ω RNA(s). Also provided include cells modified by the systems and methods herein, and cell cultures, tissues, organs, organism comprising such cells or progeny thereof. In one embodiment, the present disclosure provides a method of modifying an cell or organism. The cell may be a prokaryotic cell or a eukaryotic cell. The cell may be a mammalian cell. The mammalian cell may be a non-human primate, bovine, porcine, rodent or mouse cell. The cell may be a non-mammalian eukaryotic cell such as poultry, fish or shrimp. The cell may be a therapeutic T cell or antibody-producing B-cell. The cell may also be a plant cell. The plant cell may be of a crop plant such as cassava, corn, sorghum, wheat, or rice. The plant cell may also be of an algae, tree or vegetable. The modification introduced to the cell by the present invention may be such that the cell and progeny of the cell are altered for improved production of biologic products such as an antibody, starch, alcohol or other desired cellular output. The modification introduced to the cell by the present invention may be such that the cell and progeny of the cell include an alteration that changes the biologic product produced.

[0600] In one embodiment, one or more polynucleotide molecules, vectors, or vector systems driving expression of one or more elements of the compositions, systems, or delivery systems comprising one or more elements of the nucleic acid-targeting system are introduced into a host cell such that expression of the elements of the nucleic acid-targeting system direct formation of a nucleic acid-targeting complex at one or more target sites. In an embodiment of the invention the host cell may be a eukaryotic cell, a prokaryotic cell, or a plant cell.

[0601] In one embodiment, the host cell is a cell of a cell line. Cell lines are available from a variety of sources known to those with skill in the art (see, e.g., the American Type Culture Collection (ATCC) (Manassus, Va.)). In one embodiment, a cell transfected with one or more vectors described herein is used to establish a new cell line comprising one or more vector-derived sequences. In one embodiment, a cell transiently transfected with the components of a system as described herein (such as by transient transfection of one or more vectors, or transfection with RNA), and modified through the activity of a complex, is used to establish a new cell line comprising cells containing the modification but lacking any other exogenous sequence. In one embodiment, cells transiently or non-transiently transfected with one or more vectors described herein, or cell lines derived from such cells are used in assessing one or more test compounds.

[0602] Further intended are isolated human cells or tissues, plants or non-human animals comprising one or more of the polynucleotide molecules, vectors, vector systems, or cells described in any of the embodiments herein. In an aspect, host cells and cell lines modified by or comprising the compositions, systems or modified enzymes of present invention are provided, including (isolated) stem cells, and progeny thereof.

[0603] In an embodiment, the plants or non-human animals comprise at least one of the system components, polynucleotide molecules, vectors, vector systems, or cells described in any of the embodiments herein at least one tissue type of the plant or non-human animal. In an embodiment, non-human animals comprise at least one of the system components, polynucleotide molecules, vectors, vector systems, or cells described in any of the embodiments herein in at least one tissue type. In an embodiment, the presence of the system components is transient, in that they are degraded over time. In an embodiment, expression of the components of the systems and compositions described in any of the embodiments comprised in polynucleotide molecules, vectors, vector systems, or cells is limited to certain tissue types or regions in the plant or non-human animal. In an embodiment, the expression of

the components of the systems and compositions described in any of the embodiments comprised in polynucleotide molecules, vectors, vector systems, or cells is dependent of a physiological cue. In an embodiment, expression of the components of the systems and compositions described in any of the embodiments comprised in polynucleotide molecules, vectors, vector systems, or cells may be triggered by an exogenous molecule. In an embodiment, expression of the components of the systems and compositions described in any of the embodiments comprised in polynucleotide molecules, vectors, vector systems, or cells is dependent on the expression of a non-cas molecule in the plant or non-human animal.

APPLICATIONS AND USES IN GENERAL

[0604] The systems, the vector systems, the vectors and the compositions described herein may be used in various nucleic acids-targeting applications, altering or modifying synthesis of a gene product, such as a protein, nucleic acids cleavage, nucleic acids editing, nucleic acids splicing; trafficking of target nucleic acids, tracing of target nucleic acids, isolation of target nucleic acids, visualization of target nucleic acids, etc.

[0605] Aspects of the invention thus also encompass methods and uses of the compositions and systems described herein in genome engineering, e.g. for altering or manipulating the expression of one or more genes or the one or more gene products, in prokaryotic or eukaryotic cells, *in vitro*, *in vivo* or *ex vivo*. In some examples, the target polynucleotides are target sequences within genomic DNA, including nuclear genomic DNA, mitochondrial DNA, or chloroplast DNA.

[0606] Typically, in the context of a nucleic acid-targeting system, formation of a nucleic acid-targeting complex (comprising a crRNA or guide RNA hybridized to a target sequence and complexed with one or more nucleic acid-targeting effector proteins) results in cleavage of one or both DNA or RNA strands in or near (e.g., within 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or more base pairs from) the target sequence. As used herein the term “sequence(s) associated with a target locus of interest” refers to sequences near the vicinity of the target sequence (e.g. within 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or more base pairs from the target sequence, wherein the target sequence is comprised within a target locus of interest).

[0607] In one embodiment, the present disclosure provides a method of targeting a polynucleotide, comprising contacting a sample (such as cell, population of cells, tissue, organ, or an organism) that comprises a target polynucleotide with the composition, systems,

polynucleotide(s), or vector(s). The contacting may result in modification of a gene product or modification of the amount or expression of a gene product. In some examples, the target sequence of the polynucleotide is a disease-associated target sequence.

[0608] In one embodiment, the present disclosure provides a method of modifying target polynucleotides comprising delivering the composition, the one or more polynucleotides of 2, or one or more vectors to a cell or population of cells comprising the target polynucleotides, wherein the complex directs the reverse transcriptase to the target sequence and the reverse transcriptase facilitates insertion of the donor sequence from the ω RNA into the target polynucleotide.

[0609] Examples of target polynucleotides include a sequence associated with a signaling biochemical pathway, e.g., a signaling biochemical pathway-associated gene or polynucleotide. Examples of target polynucleotides include a disease associated gene or polynucleotide. A “disease-associated” gene or polynucleotide refers to any gene or polynucleotide which is yielding transcription or translation products at an abnormal level or in an abnormal form in cells derived from a disease-affected tissues compared with tissues or cells of a non-disease control. It may be a gene that becomes expressed at an abnormally high level; it may be a gene that becomes expressed at an abnormally low level, where the altered expression correlates with the occurrence and/or progression of the disease. A disease-associated gene also refers to a gene possessing mutation(s) or genetic variation that is directly responsible or is in linkage disequilibrium with a gene(s) that is responsible for the etiology of a disease. The transcribed or translated products may be known or unknown, and may be at a normal or abnormal level.

[0610] The target polynucleotide of a complex can be any polynucleotide endogenous or exogenous to the eukaryotic cell. For example, the target polynucleotide can be a polynucleotide residing in the nucleus of the eukaryotic cell. The target polynucleotide can be a sequence coding a gene product (e.g., a protein) or a non-coding sequence (e.g., a regulatory polynucleotide or a junk DNA). Without wishing to be bound by theory, it is believed that the target sequence should be associated with a TAM (target adjacent motif); that is, a short sequence recognized by the complex. The precise sequence and length requirements for the TAM differ depending on the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease used, but TAMs are typically 2-5 base pair sequences adjacent the protospacer (that

is, the target sequence) A skilled person will be able to identify further TAM sequences for use with a given IscB polypeptide or CRISPR-associated IscB polypeptide nuclease. Further, engineering of the TAM Interacting domain may allow programming of TAM specificity, improve target site recognition fidelity, and increase the versatility of the IscB polypeptide nuclease, genome engineering platform. IscB polypeptide or CRISPR-associated IscB polypeptide nuclease may be engineered to alter their TAM specificity.

[0611] In an embodiment the IscB TAM is ATNA where N is any nucleotide. In an embodiment, the IscB TAM is ATGA, ATAA, ATAAA, or ATN. In one embodiment, the IscB is *Ignatius tetrasporus* and the TAM is NNG.

[0612] Examples of target polynucleotides include a sequence associated with a signaling biochemical pathway, e.g., a signaling biochemical pathway-associated gene or polynucleotide. Examples of target polynucleotides include a disease associated gene or polynucleotide. A “disease-associated” gene or polynucleotide refers to any gene or polynucleotide which is yielding transcription or translation products at an abnormal level or in an abnormal form in cells derived from a disease-affected tissues compared with tissues or cells of a non-disease control. It may be a gene that becomes expressed at an abnormally high level; it may be a gene that becomes expressed at an abnormally low level, where the altered expression correlates with the occurrence and/or progression of the disease. A disease-associated gene also refers to a gene possessing mutation(s) or genetic variation that is directly responsible or is in linkage disequilibrium with a gene(s) that is responsible for the etiology of a disease. The transcribed or translated products may be known or unknown, and may be at a normal or abnormal level.

[0613] Aspects of the invention relate to a method of targeting a polynucleotide, comprising contacting a sample that comprises the polynucleotide with a composition, system or IscB polypeptide or CRISPR-associated IscB polypeptide nuclease as described in any embodiment herein, a delivery system comprising a composition, system or IscB polypeptide or CRISPR-associated IscB polypeptide nuclease as described in any embodiment herein, a polynucleotide comprising a composition, system or IscB polypeptide or CRISPR-associated IscB polypeptide nuclease as described in any embodiment herein, a vector comprising a composition, system or IscB polypeptide or CRISPR-associated IscB polypeptide nuclease as described in any embodiment herein, or a vector system comprising a composition, system or IscB polypeptide or CRISPR-associated IscB polypeptide nuclease as described in any

embodiment herein. In an embodiment, a target polynucleotide is contacted with at least two different composition, system or IscB polypeptide or CRISPR-associated IscB polypeptide nucleases. In further embodiments, the two different IscB polypeptide nuclease have different target polynucleotide specificities, or degrees of specificity. In an embodiment, the two different IscB polypeptide or CRISPR-associated IscB polypeptide nuclease have a different TAM specificity.

[0614] Also envisaged are methods of targeting a polynucleotide, comprising contacting a sample that comprises the polynucleotide with the composition and systems, vectors, polynucleotides, herein wherein contacting results in modification of a gene product or modification of the amount or expression of a gene product. In an embodiment, the expression of the targeted gene product is increased by the method. In an embodiment, the expression of the targeted gene product is increased by at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, 100%. In an embodiment, the expression of the targeted gene product is increased at least 1.5-fold, at least 2-fold, at least 2.5-fold, at least 3-fold, at least 3.5-fold, at least 3.5-fold, at least 4-fold, at least 4.5-fold, at least 5-fold, at least 10-fold, at least 10-fold, at least 15-fold, at least 20-fold, at least 25-fold, at least 50-fold, at least 100-fold. In an embodiment, the expression of the targeted gene product is reduced by at least 10%, by at least 15%, by at least 20%, by at least 25%, by at least 30%, by at least 35%, by at least 40%, by at least 45%, by at least 50%, by at least 55%, by at least 60%, by at least 65%, by at least 70%, by at least 75%, by at least 80%, by at least 85%, by at least 90%, by at least 95%, by at least 100%. In an embodiment, the expression of the targeted gene product is reduced at least 1.5-fold, at least 2-fold, at least 2.5-fold, at least 3-fold, at least 3.5-fold, at least 3.5-fold, at least 4-fold, at least 4.5-fold, at least 5-fold, at least 10-fold, at least 10-fold, at least 15-fold, at least 20-fold, at least 25-fold, at least 50-fold, at least 100-fold. In alternative embodiments, the expression of the targeted gene product is reduced by the method. In further embodiments, expression of the targeted gene may be completely eliminated, or may be considered eliminated as remnant expression levels of the targeted gene fall below the detection limit of methods known in the art that are used to quantify, detect, or monitor expression levels of genes.

[0615] In one embodiment, one or more polynucleotide molecules, vectors, or vector systems driving expression of one or more elements of a nucleic acid-targeting system or

delivery systems comprising one or more elements of the nucleic acid-targeting system are introduced into a host cell such that expression of the elements of the nucleic acid-targeting system direct formation of a nucleic acid-targeting complex at one or more target sites. In an embodiment of the invention the host cell may be a eukaryotic cell, a prokaryotic cell, or a plant cell.

[0616] In one embodiment, the host cell is a cell of a cell line. Cell lines are available from a variety of sources known to those with skill in the art (see, e.g., the American Type Culture Collection (ATCC) (Manassus, Va.)). In one embodiment, a cell transfected with one or more vectors described herein is used to establish a new cell line comprising one or more vector-derived sequences. In one embodiment, a cell transiently transfected with the components of a composition or system as described herein (such as by transient transfection of one or more vectors, or transfection with RNA), and modified through the activity of a complex, is used to establish a new cell line comprising cells containing the modification but lacking any other exogenous sequence. In one embodiment, cells transiently or non-transiently transfected with one or more vectors described herein, or cell lines derived from such cells are used in assessing one or more test compounds.

[0617] Further intended are isolated human cells or tissues, plants or non-human animals comprising one or more of the polynucleotide molecules, vectors, vector systems, or cells described in any of the embodiments herein. In an aspect, host cells and cell lines modified by or comprising the compositions, systems or modified enzymes of present invention are provided, including (isolated) stem cells, and progeny thereof.

[0618] In an embodiment, the plants or non-human animals comprise at least one of the compositions, polynucleotide molecules, vectors, vector systems, or cells described in any of the embodiments herein at least one tissue type of the plant or non-human animal. In certain embodiment, non-human animals comprise at least one of the compositions, polynucleotide molecules, vectors, vector systems, or cells described in any of the embodiments herein in at least one tissue type. In an embodiment, the presence of the compositions is transient, in that they are degraded over time. In an embodiment, expression of the compositions described in any of the embodiments comprised in polynucleotide molecules, vectors, vector systems, or cells is limited to certain tissue types or regions in the plant or non-human animal. In an embodiment, the expression of the compositions described in any of the embodiments comprised in polynucleotide molecules, vectors, vector systems, or cells is dependent of a

physiological cue. In an embodiment, expression of the compositions described in any of the embodiments comprised in polynucleotide molecules, vectors, vector systems, or cells may be triggered by an exogenous molecule. In an embodiment, expression of the compositions described in any of the embodiments comprised in polynucleotide molecules, vectors, vector systems, or cells is dependent on the expression of a non-Cas molecule in the plant or non-human animal.

[0619] In one aspect, the invention provides methods for using one or more elements of a nucleic acid-targeting system. The nucleic acid-targeting complex of the invention provides an effective means for modifying a target DNA or RNA (single or double stranded, linear or supercoiled). The nucleic acid-targeting complex of the invention has a wide variety of utility including modifying (e.g., deleting, inserting, translocating, inactivating, activating) a target DNA or RNA in a multiplicity of cell types. As such, the nucleic acid-targeting complex of the invention has a broad spectrum of applications in, e.g., gene therapy, drug screening, disease diagnosis, and prognosis. An exemplary nucleic acid-targeting complex comprises a DNA or RNA-targeting effector protein complexed with a crRNA or guide RNA hybridized to a target sequence within the target locus of interest.

[0620] In one embodiment, this invention provides a method of cleaving a target polynucleotide. The method may comprise modifying a target polynucleotide using a nucleic acid-targeting complex that binds to the target polynucleotide and effect cleavage of said target polynucleotide. In an embodiment, the nucleic acid-targeting complex of the invention, when introduced into a cell, may create a break (e.g., a single or a double strand break) in the polynucleotide sequence. For example, the method can be used to cleave a disease polynucleotide in a cell. For example, an exogenous template comprising a sequence to be integrated flanked by an upstream sequence and a downstream sequence may be introduced into a cell. The upstream and downstream sequences share sequence similarity with either side of the site of integration in the polynucleotide. The exogenous template comprises a sequence to be integrated (e.g., a mutated RNA). The sequence for integration may be a sequence endogenous or exogenous to the cell. Examples of a sequence to be integrated include polynucleotide encoding a protein or a non-coding RNA (e.g., a microRNA). Thus, the sequence for integration may be operably linked to an appropriate control sequence or sequences. Alternatively, the sequence to be integrated may provide a regulatory function. The

upstream and downstream sequences in the recombination template are selected to promote recombination between the RNA sequence of interest and the recombination. The upstream sequence is a polynucleotide sequence that shares sequence similarity with the sequence upstream of the targeted site for integration. Similarly, the downstream sequence is a polynucleotide sequence that shares sequence similarity with the polynucleotide sequence downstream of the targeted site of integration. The upstream and downstream sequences in the recombination template can have 75%, 80%, 85%, 90%, 95%, or 100% sequence identity with the targeted sequence. Preferably, the upstream and downstream sequences in the recombination template have about 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with the targeted sequence. In some methods, the upstream and downstream sequences in the recombination template have about 99% or 100% sequence identity with the targeted sequence. An upstream or downstream sequence may comprise from about 20 bp to about 2500 bp, for example, about 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, or 2500 bp. In some methods, the exemplary upstream or downstream sequence have about 200 bp to about 2000 bp, about 600 bp to about 1000 bp, or more particularly about 700 bp to about 1000 bp. In some methods, the recombination template may further comprise a marker. Such a marker may make it easy to screen for targeted integrations. Examples of suitable markers include restriction sites, fluorescent proteins, or selectable markers. The recombination template of the invention can be constructed using recombinant techniques (see, for example, Sambrook et al., 2001 and Ausubel et al., 1996). In a method for modifying a target sequence by integrating an recombination template, a break (e.g., double or single stranded break in double or single stranded DNA or RNA) is introduced into the DNA or RNA sequence by the nucleic acid-targeting complex, the break is repaired via homologous recombination with an recombination template such that the template is integrated into the target. The presence of a double-stranded break facilitates integration of the template. In other embodiments, this invention provides a method of modifying expression of a RNA in a eukaryotic cell. The method comprises increasing or decreasing expression of a target polynucleotide by using a nucleic acid-targeting complex that binds to the DNA or RNA (e.g., mRNA or pre-mRNA). In some methods, a target can be inactivated to affect the modification of the expression in a cell. For example, upon the binding of a nucleic acid-targeting complex to a target sequence in a cell, the target is inactivated such that the sequence is not translated, the coded protein is not produced, or the

sequence does not function as the wild-type sequence does. For example, a protein or microRNA coding sequence may be inactivated such that the protein or microRNA or pre-microRNA transcript is not produced. The target of a nucleic acid-targeting complex can be any polynucleotide endogenous or exogenous to the eukaryotic cell. For example, the target polynucleotide can be a polynucleotide residing in the nucleus of the eukaryotic cell. The target polynucleotide can be a sequence coding a gene product (e.g., a protein) or a non-coding sequence (e.g., ncRNA, lncRNA, tRNA, or rRNA). Examples of target RNA include a sequence associated with a signaling biochemical pathway, e.g., a signaling biochemical pathway-associated polynucleotide. Examples of target polynucleotide include a disease associated polynucleotide. A “disease-associated” polynucleotide refers to any polynucleotide which is yielding translation products at an abnormal level or in an abnormal form in cells derived from a disease-affected tissues compared with tissues or cells of a non-disease control. It may be a gene that becomes expressed at an abnormally high level; it may be a gene that becomes expressed at an abnormally low level, where the altered expression correlates with the occurrence and/or progression of the disease. A disease-associated polynucleotide also refers to a gene possessing mutation(s) or genetic variation that is directly responsible or is in linkage disequilibrium with a gene(s) that is responsible for the etiology of a disease. The translated products may be known or unknown, and may be at a normal or abnormal level. The target RNA of a nucleic acid-targeting complex can be any polynucleotide endogenous or exogenous to the eukaryotic cell. For example, the target RNA can be a RNA residing in the nucleus of the eukaryotic cell. The target polynucleotide can be a sequence coding a gene product (e.g., a protein) or a non-coding sequence (e.g., ncRNA, lncRNA, tRNA, or rRNA).

[0621] In one embodiment, the method may comprise allowing a compositions to bind to the target DNA or RNA to effect cleavage of said target DNA or RNA thereby modifying the target DNA or RNA, wherein the nucleic acid-targeting complex comprises a nucleic acid-targeting effector protein complexed with a guide RNA hybridized to a target sequence within said target DNA or RNA. In one aspect, the invention provides a method of modifying expression of DNA or RNA in a eukaryotic cell. In one embodiment, the method comprises allowing a nucleic acid-targeting complex to bind to the DNA or RNA such that said binding results in increased or decreased expression of said DNA or RNA; wherein the nucleic acid-targeting complex comprises a nucleic acid-targeting effector protein complexed with a gRNA

or guide RNA. Similar considerations and conditions apply as above for methods of modifying a target DNA or RNA. In fact, these sampling, culturing and re-introduction options apply across the aspects of the present invention. In one aspect, the invention provides for methods of modifying a target DNA or RNA in a eukaryotic cell, which may be *in vivo*, *ex vivo* or *in vitro*. In one embodiment, the method comprises sampling a cell or population of cells from a human or non-human animal, and modifying the cell or cells. Culturing may occur at any stage *ex vivo*. The cell or cells may even be re-introduced into the non-human animal or plant. For re-introduced cells it is particularly preferred that the cells are stem cells. The compositions as described in any embodiment herein may be used to detect nucleic acid identifiers. Nucleic acid identifiers are non-coding nucleic acids that may be used to identify a particular article. Example nucleic acid identifiers, such as DNA watermarks, are described in Heider and Barnekow. "DNA watermarks: A proof of concept" BMC Molecular Biology 9:40 (2008). The nucleic acid identifiers may also be a nucleic acid barcode. A nucleic-acid based barcode is a short sequence of nucleotides (for example, DNA, RNA, or combinations thereof) that is used as an identifier for an associated molecule, such as a target molecule and/or target nucleic acid. A nucleic acid barcode can have a length of at least, for example, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 60, 70, 80, 90, or 100 nucleotides, and can be in single- or double-stranded form. One or more nucleic acid barcodes can be attached, or "tagged," to a target molecule and/or target nucleic acid. This attachment can be direct (for example, covalent or non-covalent binding of the barcode to the target molecule) or indirect (for example, via an additional molecule, for example, a specific binding agent, such as an antibody (or other protein) or a barcode receiving adaptor (or other nucleic acid molecule). Target molecule and/or target nucleic acids can be labeled with multiple nucleic acid barcodes in combinatorial fashion, such as a nucleic acid barcode concatemer. Typically, a nucleic acid barcode is used to identify target molecules and/or target nucleic acids as being from a particular compartment (for example a discrete volume), having a particular physical property (for example, affinity, length, sequence, etc.), or having been subject to certain treatment conditions. Target molecule and/or target nucleic acid can be associated with multiple nucleic acid barcodes to provide information about all of these features (and more). Methods of generating nucleic acid- barcodes are disclosed, for example, in International Patent Application Publication No. WO/2014/047561.

[0622] In an embodiment, compositions induce a double strand break for the purpose of inducing HDR-mediated correction. In a further embodiment, two or more guide RNAs complexing with IscB polypeptide nuclease or an ortholog or homolog thereof, may be used to induce multiplexed breaks for purpose of inducing HDR-mediated correction.

[0623] A recombination template nucleic acid, as that term is used herein, refers to a nucleic acid sequence which can be used in conjunction with compositions disclosed herein to alter the structure of a target position. In an embodiment, the target nucleic acid is modified to have some or all of the sequence of the recombination template nucleic acid, typically at or near cleavage site(s). In an embodiment, the recombination template nucleic acid is single stranded. In an alternate embodiment, the recombination template nucleic acid is double stranded. In an embodiment, the recombination template nucleic acid is DNA, e.g., double stranded DNA. In an alternate embodiment, the recombination template nucleic acid is single stranded DNA.

[0624] In one embodiment, a recombination template is provided to serve as a template in homologous recombination, such as within or near a target sequence nicked or cleaved by a nucleic acid-targeting effector protein as a part of a nucleic acid-targeting complex.

[0625] A recombination template may be a component of another vector as described herein, contained in a separate vector, or provided as a separate polynucleotide. A recombination template polynucleotide may be of any suitable length, such as about or more than about 10, 15, 20, 25, 50, 75, 100, 150, 200, 500, 1000, or more nucleotides in length. In one embodiment, the recombination template polynucleotide is complementary to a portion of a polynucleotide comprising the target sequence. When optimally aligned, a recombination template polynucleotide might overlap with one or more nucleotides of a target sequences (e.g. about or more than about 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 or more nucleotides). In one embodiment, when a recombination template sequence and a polynucleotide comprising a target sequence are optimally aligned, the nearest nucleotide of the recombination template polynucleotide is within about 1, 5, 10, 15, 20, 25, 50, 75, 100, 200, 300, 400, 500, 1000, 5000, 10000, or more nucleotides from the target sequence.

[0626] In an embodiment, the recombination template nucleic acid alters the structure of the target position by participating in homologous recombination. In an embodiment, the recombination template nucleic acid alters the sequence of the target position. In an

embodiment, the recombination template nucleic acid results in the incorporation of a modified, or non-naturally occurring base into the target nucleic acid.

[0627] The recombination template sequence may undergo a breakage mediated or catalyzed recombination with the target sequence. In an embodiment, the recombination template nucleic acid may include sequence that corresponds to a site on the target sequence that is cleaved by an IscB polypeptide nuclease mediated cleavage event. In an embodiment, the recombination template nucleic acid may include sequence that corresponds to both, a first site on the target sequence that is cleaved in a first IscB polypeptide nuclease mediated event and a second site on the target sequence that is cleaved in a second IscB polypeptide nuclease mediated event.

[0628] In an embodiment, the recombination template nucleic acid can include sequence which results in an alteration in the coding sequence of a translated sequence, e.g., one which results in the substitution of one amino acid for another in a protein product, e.g., transforming a mutant allele into a wild type allele, transforming a wild type allele into a mutant allele, and/or introducing a stop codon, insertion of an amino acid residue, deletion of an amino acid residue, or a nonsense mutation. In an embodiment, the recombination template nucleic acid can include sequence which results in an alteration in a non-coding sequence, e.g., an alteration in an exon or in a 5' or 3' non-translated or non-transcribed region. Such alterations include an alteration in a control element, e.g., a promoter, enhancer, and an alteration in a cis-acting or trans-acting control element.

[0629] A recombination template nucleic acid having homology with a target position in a target gene may be used to alter the structure of a target sequence. The recombination template sequence may be used to alter an unwanted structure, e.g., an unwanted or mutant nucleotide. The recombination template nucleic acid may include sequence which, when integrated, results in: decreasing the activity of a positive control element; increasing the activity of a positive control element; decreasing the activity of a negative control element; increasing the activity of a negative control element; decreasing the expression of a gene; increasing the expression of a gene; increasing resistance to a disorder or disease; increasing resistance to viral entry; correcting a mutation or altering an unwanted amino acid residue conferring, increasing, abolishing or decreasing a biological property of a gene product, e.g., increasing the enzymatic activity of an enzyme, or increasing the ability of a gene product to interact with another molecule.

[0630] The recombination template nucleic acid may include sequence which results in: a change in sequence of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 or more nucleotides of the target sequence. In an embodiment, the recombination template nucleic acid may be 20+/- 10, 30+/- 10, 40+/- 10, 50+/- 10, 60+/- 10, 70+/- 10, 80+/- 10, 90+/- 10, 100+/- 10, 110+/- 10, 120+/- 10, 130+/- 10, 140+/- 10, 150+/- 10, 160+/- 10, 170+/- 10, 180+/- 10, 190+/- 10, 200+/- 10, 210+/- 10, of 220+/- 10 nucleotides in length. In an embodiment, the recombination template nucleic acid may be 30+/-20, 40+/-20, 50+/-20, 60+/-20, 70+/- 20, 80+/-20, 90+/-20, 100+/- 20, 110+/-20, 120+/-20, 130+/-20, 140+/-20, 150+/-20, 160+/-20, 170+/-20, 180+/-20, 190+/- 20, 200+/-20, 210+/-20, of 220+/-20 nucleotides in length. In an embodiment, the recombination template nucleic acid is 10 to 1,000, 20 to 900, 30 to 800, 40 to 700, 50 to 600, 50 to 500, 50 to 400, 50 to 300, 50 to 200, or 50 to 100 nucleotides in length.

[0631] A recombination template nucleic acid comprises the following components: [5' homology arm]-[replacement sequence]-[3' homology arm]. The homology arms provide for recombination into the chromosome, thus replacing the undesired element, e.g., a mutation or signature, with the replacement sequence. In an embodiment, the homology arms flank the most distal cleavage sites. In an embodiment, the 3' end of the 5' homology arm is the position next to the 5' end of the replacement sequence. In an embodiment, the 5' homology arm can extend at least 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, or 2000 nucleotides 5' from the 5' end of the replacement sequence. In an embodiment, the 5' end of the 3' homology arm is the position next to the 3' end of the replacement sequence. In an embodiment, the 3' homology arm can extend at least 10, 20, 30, 40, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, or 2000 nucleotides 3' from the 3' end of the replacement sequence.

[0632] In an embodiment, one or both homology arms may be shortened to avoid including certain sequence repeat elements. For example, a 5' homology arm may be shortened to avoid a sequence repeat element. In other embodiments, a 3' homology arm may be shortened to avoid a sequence repeat element. In one embodiment, both the 5' and the 3' homology arms may be shortened to avoid including certain sequence repeat elements.

[0633] In an embodiment, a recombination template nucleic acid for correcting a mutation may be designed for use as a single-stranded oligonucleotide. When using a single-stranded oligonucleotide, 5' and 3' homology arms may range up to about 200 base pairs (bp) in length, e.g., at least 25, 50, 75, 100, 125, 150, 175, or 200 bp in length.

[0634] Unlike IscB polypeptide or CRISPR-associated IscB polypeptide nuclease - mediated gene knockout, which permanently eliminates expression by mutating the gene at the DNA level, IscB polypeptide nuclease knockdown allows for temporary reduction of gene expression through the use of artificial transcription factors. Mutating key residues in both DNA cleavage domains of the IscB polypeptide nuclease, results in the generation of a catalytically inactive IscB polypeptide nuclease. A catalytically inactive IscB polypeptide or CRISPR-associated IscB polypeptide nuclease complexes with a guide RNA and localizes to the DNA sequence specified by that guide RNA's targeting domain, however, it does not cleave the target DNA. Fusion of the inactive IscB polypeptide or CRISPR-associated IscB polypeptide nuclease protein to an effector domain, e.g., a transcription repression domain, enables recruitment of the effector to any DNA site specified by the guide RNA. In an embodiment, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease may be fused to a transcriptional repression domain and recruited to the promoter region of a gene. Especially for gene repression, it is contemplated herein that blocking the binding site of an endogenous transcription factor would aid in downregulating gene expression. In another embodiment, an inactive IscB polypeptide or CRISPR-associated IscB polypeptide nuclease can be fused to a chromatin modifying protein. Altering chromatin status can result in decreased expression of the target gene.

[0635] In an embodiment, a guide RNA molecule can be targeted to a known transcription response elements (e.g., promoters, enhancers, etc.), a known upstream activating sequences, and/or sequences of unknown or known function that are suspected of being able to control expression of the target DNA.

[0636] In some methods, a target polynucleotide can be inactivated to affect the modification of the expression in a cell. For example, upon the binding of a composition to a target sequence in a cell, the target polynucleotide is inactivated such that the sequence is not transcribed, the coded protein is not produced, or the sequence does not function as the wild-type sequence does. For example, a protein or microRNA coding sequence may be inactivated such that the protein is not produced.

Non-Homologous End-Joining

[0637] In an embodiment, nuclease-induced non-homologous end-joining (NHEJ) can be used to target gene-specific knockouts. Nuclease-induced NHEJ can also be used to remove (e.g., delete) sequence in a gene of interest. Generally, NHEJ repairs a double-strand break in

the DNA by joining together the two ends; however, generally, the original sequence is restored only if two compatible ends, exactly as they were formed by the double-strand break, are perfectly ligated. The DNA ends of the double-strand break are frequently the subject of enzymatic processing, resulting in the addition or removal of nucleotides, at one or both strands, prior to rejoining of the ends. This results in the presence of insertion and/or deletion (indel) mutations in the DNA sequence at the site of the NHEJ repair. Two-thirds of these mutations typically alter the reading frame and, therefore, produce a non-functional protein. Additionally, mutations that maintain the reading frame, but which insert or delete a significant amount of sequence, can destroy functionality of the protein. This is locus dependent as mutations in critical functional domains are likely less tolerable than mutations in non-critical regions of the protein. The indel mutations generated by NHEJ are unpredictable in nature; however, at a given break site certain indel sequences are favored and are over-represented in the population, likely due to small regions of microhomology. The lengths of deletions can vary widely; most commonly in the 1-50 bp range, but they can easily be greater than 50 bp, e.g., they can easily reach greater than about 100-200 bp. Insertions tend to be shorter and often include short duplications of the sequence immediately surrounding the break site. However, it is possible to obtain large insertions, and in these cases, the inserted sequence has often been traced to other regions of the genome or to plasmid DNA present in the cells.

[0638] Because NHEJ is a mutagenic process, it may also be used to delete small sequence motifs as long as the generation of a specific final sequence is not required. If a double-strand break is targeted near to a short target sequence, the deletion mutations caused by the NHEJ repair often span, and therefore remove, the unwanted nucleotides. For the deletion of larger DNA segments, introducing two double-strand breaks, one on each side of the sequence, can result in NHEJ between the ends with removal of the entire intervening sequence. Both of these approaches can be used to delete specific DNA sequences; however, the error-prone nature of NHEJ may still produce indel mutations at the site of repair.

[0639] Both double strand cleaving IscB polypeptide nuclease, or an ortholog or homolog thereof, and single strand, or nickase, IscB polypeptide nuclease, or an ortholog or homolog thereof, molecules can be used in the methods and compositions described herein to generate NHEJ-mediated indels. NHEJ-mediated indels targeted to the gene, e.g., a coding region, e.g., an early coding region of a gene of interest can be used to knockout (i.e., eliminate expression of) a gene of interest. For example, early coding region of a gene of interest includes sequence

immediately following a transcription start site, within a first exon of the coding sequence, or within 500 bp of the transcription start site (e.g., less than 500, 450, 400, 350, 300, 250, 200, 150, 100 or 50 bp).

[0640] In an embodiment, in which a guide RNA and IscB polypeptide nuclease, or an ortholog or homolog thereof, generate a double strand break for the purpose of inducing NHEJ-mediated indels, a guide RNA may be configured to position one double-strand break in close proximity to a nucleotide of the target position. In an embodiment, the cleavage site may be between 0-500 bp away from the target position (e.g., less than 500, 400, 300, 200, 100, 50, 40, 30, 25, 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 bp from the target position).

[0641] In an embodiment, in which two guide RNAs complexing with IscB polypeptide nuclease, or an ortholog or homolog thereof, e.g., IscB polypeptide nickases induce two single strand breaks for the purpose of inducing NHEJ-mediated indels, two guide RNAs may be configured to position two single-strand breaks to provide for NHEJ repair a nucleotide of the target position.

[0642] In some examples, the systems herein may introduce one or more indels via NHEJ pathway and insert sequence from a combination template via HDR.

EXEMPLARY APPLICATIONS

[0643] The invention provides a non-naturally occurring or engineered composition, or one or more polynucleotides encoding components of said composition, or vector or delivery systems comprising one or more polynucleotides encoding components of said composition for use in a modifying a target cell *in vivo*, *ex vivo* or *in vitro* and, may be conducted in a manner alters the cell such that once modified the progeny or cell line of the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease modified cell retains the altered phenotype. The modified cells and progeny may be part of a multi-cellular organism such as a plant or animal with *ex vivo* or *in vivo* application of composition to desired cell types. The methods herein include a therapeutic method of treatment. The therapeutic method of treatment may comprise gene or genome editing, or gene therapy.

[0644] In one embodiment, one or more vectors described herein are used to produce a non-human transgenic animal or transgenic plant. In one embodiment, the transgenic animal is a mammal, such as a mouse, rat, or rabbit. Methods for producing transgenic animals and plants are known in the art, and generally begin with a method of cell transfection, such as described herein.

Use of orthogonal catalytically inactive IscB polypeptide or CRISPR-associated IscB polypeptide nucleases

[0645] In one embodiment, the IscB polypeptide nickase is used in combination with an orthogonal catalytically inactive IscB polypeptide or CRISPR-associated IscB polypeptide nuclease to increase efficiency of said nickase (e.g., as described in Chen et al. 2017, Nature Communications 8:14958; doi:10.1038/ncomms14958). More particularly, the orthogonal catalytically inactive IscB polypeptide nuclease is characterized by a different TAM recognition site than the IscB nickase used in the AD-functionalized composition and the corresponding guide sequence is selected to bind to a target sequence proximal to that of the nickase of the functionalized IscB polypeptide or CRISPR-associated IscB polypeptide nuclease. The orthogonal catalytically inactive IscB polypeptide or CRISPR-associated IscB polypeptide nuclease as used in the context of the present invention does not form part of the functionalized composition but merely functions to increase the efficiency of said nickase and is used in combination with a standard hRNA as described in the art for said IscB polypeptide or CRISPR-associated IscB polypeptide nuclease. In one embodiment, said orthogonal catalytically inactive IscB polypeptide or CRISPR-associated IscB polypeptide nuclease is a dead IscB polypeptide nuclease, i.e. comprising one or more mutations which abolishes the nuclease activity of said IscB polypeptide nuclease. In one embodiment, the catalytically inactive orthogonal IscB polypeptide or CRISPR-associated IscB polypeptide nuclease is provided with two or more ω RNAs or guide RNAs which are capable of hybridizing to target sequences which are proximal to the target sequence of the nickase. In one embodiment, at least two ω RNAs are used to target said catalytically inactive IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, of which at least one ω RNA or guide RNAs is capable of hybridizing to a target sequence 5' of the target sequence of the nickase and at least one ω RNA is capable of hybridizing to a target sequence 3' of the target sequence of the nickase of the functionalized composition, whereby said one or more target sequences may be on the same or the opposite DNA strand as the target sequence of the IscB polypeptide or CRISPR-associated IscB polypeptide nickase. In one embodiment, the guide sequences for the one or more hRNAs of the orthogonal catalytically inactive IscB polypeptide or CRISPR-associated IscB polypeptide nuclease are selected such that the target sequences are proximal to that of the hRNA for the targeting of the functionalized composition, e.g. for the targeting of the

nickase. In one embodiment, the one or more target sequences of the orthogonal catalytically inactive IscB polypeptide or CRISPR-associated IscB polypeptide nuclease are each separated from the target sequence of the nickase by more than 5 but less than 450 basepairs. Optimal distances between the target sequences of the guides for use with the orthogonal catalytically inactive IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and the target sequence of the functionalized composition can be determined by the skilled person. In one embodiment, the catalytically inactive orthogonal IscB polypeptide or CRISPR-associated IscB polypeptide nuclease has been modified to alter its TAM specificity as described elsewhere herein. In one embodiment, the IscB polypeptide or CRISPR-associated IscB polypeptide nickase is a nickase which, by itself has limited activity in human cells, but which, in combination with an inactive orthogonal IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and one or more corresponding proximal guides ensures the required nickase activity.

Detection methods such as FISH

[0646] In one aspect, the invention provides an engineered, non-naturally occurring composition comprising a catalytically inactivate IscB polypeptide or CRISPR-associated IscB polypeptide nuclease described herein, and use this system in detection methods such as fluorescence in situ hybridization (FISH). A dead IscB polypeptide or CRISPR-associated IscB polypeptide nuclease which lacks the ability to produce DNA double-strand breaks may be fused with a marker, such as fluorescent protein, such as the enhanced green fluorescent protein (eGFP) and co-expressed with small guide RNAs to target pericentric, centric and teleomeric repeats *in vivo*. The dead IscB polypeptide or CRISPR-associated IscB polypeptide nuclease system can be used to visualize both repetitive sequences and individual genes in the human genome. Such new applications of labelled dead IscB polypeptide or CRISPR-associated IscB polypeptide nuclease may be important in imaging cells and studying the functional nuclear architecture, especially in cases with a small nucleus volume or complex 3-D structures. (Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, Park J, Blackburn EH, Weissman JS, Qi LS, Huang B. 2013. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* 155(7):1479-91. doi: 10.1016/j.cell.2013.12.001.)

Patient-specific screening methods

[0647] A nucleic acid-targeting system that targets DNA, e.g., trinucleotide repeats can be used to screen patients or patent samples for the presence of such repeats. The repeats can be

the target of the RNA of the nucleic acid-targeting system, and if there is binding thereto by the nucleic acid-targeting system, that binding can be detected, to thereby indicate that such a repeat is present. Thus, a nucleic acid-targeting system can be used to screen patients or patient samples for the presence of the repeat. The patient can then be administered suitable compound(s) to address the condition; or, can be administered a nucleic acid-targeting system to bind to and cause insertion, deletion or mutation and alleviate the condition.

Models of Genetic and Epigenetic Conditions

[0648] A method of the invention may be used to create a plant, an animal or cell that may be used to model and/or study genetic or epigenetic conditions of interest, such as a through a model of mutations of interest or a disease model. As used herein, “disease” refers to a disease, disorder, or indication in a subject. For example, a method of the invention may be used to create an animal or cell that comprises a modification in one or more nucleic acid sequences associated with a disease, or a plant, animal or cell in which the expression of one or more nucleic acid sequences associated with a disease are altered. Such a nucleic acid sequence may encode a disease associated protein sequence or may be a disease associated control sequence. Accordingly, it is understood that in embodiments of the invention, a plant, subject, patient, organism or cell can be a non-human subject, patient, organism or cell. Thus, the invention provides a plant, animal or cell, produced by the present methods, or a progeny thereof. The progeny may be a clone of the produced plant or animal, or may result from sexual reproduction by crossing with other individuals of the same species to introgress further desirable traits into their offspring. The cell may be *in vivo* or *ex vivo* in the cases of multicellular organisms, particularly animals or plants. In the instance where the cell is in cultured, a cell line may be established if appropriate culturing conditions are met and preferably if the cell is suitably adapted for this purpose (for instance a stem cell). Bacterial cell lines produced by the invention are also envisaged. Hence, cell lines are also envisaged.

[0649] In some methods, the disease model can be used to study the effects of mutations on the animal or cell and development and/or progression of the disease using measures commonly used in the study of the disease. Alternatively, such a disease model is useful for studying the effect of a pharmaceutically active compound on the disease.

[0650] In some methods, the disease model can be used to assess the efficacy of a potential gene therapy strategy. That is, a disease-associated gene or polynucleotide can be modified such that the disease development and/or progression is inhibited or reduced. In particular, the

method comprises modifying a disease-associated gene or polynucleotide such that an altered protein is produced and, as a result, the animal or cell has an altered response. Accordingly, in some methods, a genetically modified animal may be compared with an animal predisposed to development of the disease such that the effect of the gene therapy event may be assessed.

[0651] In another embodiment, this invention provides a method of developing a biologically active agent that modulates a cell signaling event associated with a disease gene. The method comprises contacting a test compound with a cell comprising one or more vectors that drive expression of one or more of a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, and a conserved nucleotide sequence linked to a guide sequence; and detecting a change in a readout that is indicative of a reduction or an augmentation of a cell signaling event associated with, e.g., a mutation in a disease gene contained in the cell.

[0652] A cell model or animal model can be constructed in combination with the method of the invention for screening a cellular function change. Such a model may be used to study the effects of a genome sequence modified by the complex of the invention on a cellular function of interest. For example, a cellular function model may be used to study the effect of a modified genome sequence on intracellular signaling or extracellular signaling. Alternatively, a cellular function model may be used to study the effects of a modified genome sequence on sensory perception. In some such models, one or more genome sequences associated with a signaling biochemical pathway in the model are modified.

[0653] Several disease models have been specifically investigated. These include de novo autism risk genes CHD8, KATNAL2, and SCN2A; and the syndromic autism (Angelman Syndrome) gene UBE3A. These genes and resulting autism models are of course preferred, but serve to show the broad applicability of the invention across genes and corresponding models. An altered expression of one or more genome sequences associated with a signaling biochemical pathway can be determined by assaying for a difference in the mRNA levels of the corresponding genes between the test model cell and a control cell, when they are contacted with a candidate agent. Alternatively, the differential expression of the sequences associated with a signaling biochemical pathway is determined by detecting a difference in the level of the encoded polypeptide or gene product.

[0654] To assay for an agent-induced alteration in the level of mRNA transcripts or corresponding polynucleotides, nucleic acid contained in a sample is first extracted according to standard methods in the art. For instance, mRNA can be isolated using various lytic enzymes

or chemical solutions according to the procedures set forth in Sambrook et al. (1989), or extracted by nucleic-acid-binding resins following the accompanying instructions provided by the manufacturers. The mRNA contained in the extracted nucleic acid sample is then detected by amplification procedures or conventional hybridization assays (e.g. Northern blot analysis) according to methods widely known in the art or based on the methods exemplified herein.

[0655] For purpose of this invention, amplification means any method employing a primer and a polymerase capable of replicating a target sequence with reasonable fidelity. Amplification may be carried out by natural or recombinant DNA polymerases such as TaqGold™, T7 DNA polymerase, Klenow fragment of *E.coli* DNA polymerase, and reverse transcriptase. A preferred amplification method is PCR. In particular, the isolated RNA can be subjected to a reverse transcription assay that is coupled with a quantitative polymerase chain reaction (RT-PCR) in order to quantify the expression level of a sequence associated with a signaling biochemical pathway.

[0656] Detection of the gene expression level can be conducted in real time in an amplification assay. In one aspect, the amplified products can be directly visualized with fluorescent DNA-binding agents including but not limited to DNA intercalators and DNA groove binders. Because the amount of the intercalators incorporated into the double-stranded DNA molecules is typically proportional to the amount of the amplified DNA products, one can conveniently determine the amount of the amplified products by quantifying the fluorescence of the intercalated dye using conventional optical systems in the art. DNA-binding dye suitable for this application include SYBR green, SYBR blue, DAPI, propidium iodine, Hoeste, SYBR gold, ethidium bromide, acridines, proflavine, acridine orange, acriflavine, fluorcoumanin, ellipticine, daunomycin, chloroquine, distamycin D, chromomycin, homidium, mithramycin, ruthenium polypyridyls, anthramycin, and the like.

[0657] In another aspect, other fluorescent labels such as sequence specific probes can be employed in the amplification reaction to facilitate the detection and quantification of the amplified products. Probe-based quantitative amplification relies on the sequence-specific detection of a desired amplified product. It utilizes fluorescent, target-specific probes (e.g., TaqMan® probes) resulting in increased specificity and sensitivity. Methods for performing probe-based quantitative amplification are well established in the art and are taught in U.S. Patent No. 5,210,015.

[0658] In yet another aspect, conventional hybridization assays using hybridization probes that share sequence homology with sequences associated with a signaling biochemical pathway can be performed. Typically, probes are allowed to form stable complexes with the sequences associated with a signaling biochemical pathway contained within the biological sample derived from the test subject in a hybridization reaction. It will be appreciated by one of skill in the art that where antisense is used as the probe nucleic acid, the target polynucleotides provided in the sample are chosen to be complementary to sequences of the antisense nucleic acids. Conversely, where the nucleotide probe is a sense nucleic acid, the target polynucleotide is selected to be complementary to sequences of the sense nucleic acid.

[0659] Hybridization can be performed under conditions of various stringency. Suitable hybridization conditions for the practice of the present invention are such that the recognition interaction between the probe and sequences associated with a signaling biochemical pathway is both sufficiently specific and sufficiently stable. Conditions that increase the stringency of a hybridization reaction are widely known and published in the art. See, for example, (Sambrook, et al., (1989); *Nonradioactive In Situ Hybridization Application Manual*, Boehringer Mannheim, second edition). The hybridization assay can be formed using probes immobilized on any solid support, including but are not limited to nitrocellulose, glass, silicon, and a variety of gene arrays. A preferred hybridization assay is conducted on high-density gene chips as described in U.S. Patent No. 5,445,934.

[0660] For a convenient detection of the probe-target complexes formed during the hybridization assay, the nucleotide probes are conjugated to a detectable label. Detectable labels suitable for use in the present invention include any composition detectable by photochemical, biochemical, spectroscopic, immunochemical, electrical, optical or chemical means. A wide variety of appropriate detectable labels are known in the art, which include fluorescent or chemiluminescent labels, radioactive isotope labels, enzymatic or other ligands. In preferred embodiments, one will likely desire to employ a fluorescent label or an enzyme tag, such as digoxigenin, β -galactosidase, urease, alkaline phosphatase or peroxidase, avidin/biotin complex.

[0661] The detection methods used to detect or quantify the hybridization intensity will typically depend upon the label selected above. For example, radiolabels may be detected using photographic film or a phosphorimager. Fluorescent markers may be detected and quantified using a photodetector to detect emitted light. Enzymatic labels are typically detected by

providing the enzyme with a substrate and measuring the reaction product produced by the action of the enzyme on the substrate; and finally colorimetric labels are detected by simply visualizing the colored label.

[0662] An agent-induced change in expression of sequences associated with a signaling biochemical pathway can also be determined by examining the corresponding gene products. Determining the protein level typically involves a) contacting the protein contained in a biological sample with an agent that specifically bind to a protein associated with a signaling biochemical pathway; and (b) identifying any agent:protein complex so formed. In one aspect of this embodiment, the agent that specifically binds a protein associated with a signaling biochemical pathway is an antibody, preferably a monoclonal antibody.

[0663] The reaction is performed by contacting the agent with a sample of the proteins associated with a signaling biochemical pathway derived from the test samples under conditions that will allow a complex to form between the agent and the proteins associated with a signaling biochemical pathway. The formation of the complex can be detected directly or indirectly according to standard procedures in the art. In the direct detection method, the agents are supplied with a detectable label and unreacted agents may be removed from the complex; the amount of remaining label thereby indicating the amount of complex formed. For such method, it is preferable to select labels that remain attached to the agents even during stringent washing conditions. It is preferable that the label does not interfere with the binding reaction. In the alternative, an indirect detection procedure may use an agent that contains a label introduced either chemically or enzymatically. A desirable label generally does not interfere with binding or the stability of the resulting agent:polypeptide complex. However, the label is typically designed to be accessible to an antibody for an effective binding and, hence generating a detectable signal.

[0664] A wide variety of labels suitable for detecting protein levels are known in the art. Non-limiting examples include radioisotopes, enzymes, colloidal metals, fluorescent compounds, bioluminescent compounds, and chemiluminescent compounds.

[0665] The amount of agent:polypeptide complexes formed during the binding reaction can be quantified by standard quantitative assays. As illustrated above, the formation of agent:polypeptide complex can be measured directly by the amount of label remained at the site of binding. In an alternative, the protein associated with a signaling biochemical pathway is tested for its ability to compete with a labeled analog for binding sites on the specific agent.

In this competitive assay, the amount of label captured is inversely proportional to the amount of protein sequences associated with a signaling biochemical pathway present in a test sample.

[0666] A number of techniques for protein analysis based on the general principles outlined above are available in the art. They include but are not limited to radioimmunoassay, ELISA (enzyme linked immunoradiometric assays), “sandwich” immunoassays, immunoradiometric assays, in situ immunoassays (using e.g., colloidal gold, enzyme or radioisotope labels), western blot analysis, immunoprecipitation assays, immunofluorescent assays, and SDS-PAGE.

[0667] Antibodies that specifically recognize or bind to proteins associated with a signaling biochemical pathway are preferable for conducting the aforementioned protein analyses. Where desired, antibodies that recognize a specific type of post-translational modifications (e.g., signaling biochemical pathway inducible modifications) can be used. Post-translational modifications include but are not limited to glycosylation, lipidation, acetylation, and phosphorylation. These antibodies may be purchased from commercial vendors. For example, anti-phosphotyrosine antibodies that specifically recognize tyrosine-phosphorylated proteins are available from a number of vendors including Invitrogen and Perkin Elmer. Anti-phosphotyrosine antibodies are particularly useful in detecting proteins that are differentially phosphorylated on their tyrosine residues in response to an ER stress. Such proteins include but are not limited to eukaryotic translation initiation factor 2 alpha (eIF-2 α). Alternatively, these antibodies can be generated using conventional polyclonal or monoclonal antibody technologies by immunizing a host animal or an antibody-producing cell with a target protein that exhibits the desired post-translational modification.

Genome Wide Knock-out Screening

[0668] The IscB polypeptide nuclease and systems described herein can be used to perform efficient and cost effective functional genomic screens. Such screens can utilize IscB polypeptide nuclease-based genome wide libraries. Such screens and libraries can provide for determining the function of genes, cellular pathways genes are involved in, and how any alteration in gene expression can result in a particular biological process. An advantage of the present invention is that the composition avoids off-target binding and its resulting side effects. This is achieved using systems arranged to have a high degree of sequence specificity for the target DNA. In preferred embodiments of the invention, the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease complexes are IscB polypeptide nuclease complexes.

[0669] In embodiments of the invention, a genome wide library may comprise a plurality of IscB polypeptide nuclease guide RNAs, as described herein, comprising guide sequences that are capable of targeting a plurality of target sequences in a plurality of genomic loci in a population of eukaryotic cells. The population of cells may be a population of embryonic stem (ES) cells. The target sequence in the genomic locus may be a non-coding sequence. The non-coding sequence may be an intron, regulatory sequence, splice site, 3' UTR, 5' UTR, or polyadenylation signal. Gene function of one or more gene products may be altered by said targeting. The targeting may result in a knockout of gene function. The targeting of a gene product may comprise more than one guide RNA. A gene product may be targeted by 2, 3, 4, 5, 6, 7, 8, 9, or 10 guide RNAs, preferably 3 to 4 per gene. Off-target modifications may be minimized by exploiting the staggered double strand breaks generated by IscB polypeptide nuclease complexes or by utilizing methods analogous to those used in composition (See, e.g., DNA targeting specificity of RNA-guided Cas nucleases. Hsu, P., Scott, D., Weinstein, J., Ran, FA., Konermann, S., Agarwala, V., Li, Y., Fine, E., Wu, X., Shalem, O., Cradick, TJ., Marraffini, LA., Bao, G., & Zhang, F. Nat Biotechnol doi:10.1038/nbt.2647 (2013)), incorporated herein by reference. The targeting may be of about 100 or more sequences. The targeting may be of about 1000 or more sequences. The targeting may be of about 20,000 or more sequences. The targeting may be of the entire genome. The targeting may be of a panel of target sequences focused on a relevant or desirable pathway. The pathway may be an immune pathway. The pathway may be a cell division pathway.

[0670] One aspect of the invention comprehends a genome wide library that may comprise a plurality of ω RNAs or guide RNAs that may comprise guide sequences that are capable of targeting a plurality of target sequences in a plurality of genomic loci, wherein said targeting results in a knockout of gene function. This library may potentially comprise ω guide RNAs that target each and every gene in the genome of an organism.

[0671] In one embodiment of the invention, the organism or subject is a eukaryote (including mammal including human) or a non-human eukaryote or a non-human animal or a non-human mammal. In one embodiment, the organism or subject is a non-human animal, and may be an arthropod, for example, an insect, or may be a nematode. In some methods of the invention the organism or subject is a plant. In some methods of the invention, the organism or subject is a mammal or a non-human mammal. A non-human mammal may be for example a

rodent (preferably a mouse or a rat), an ungulate, or a primate. In some methods of the invention the organism or subject is algae, including microalgae, or is a fungus.

[0672] The knockout of gene function may comprise introducing into each cell in the population of cells a vector system of one or more vectors comprising an engineered, non-naturally occurring composition herein. The guide sequence may target a unique gene in each cell, wherein the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease is operably linked to a regulatory element, wherein when transcribed, the ω RNAs or guide RNA comprising the guide sequence directs sequence-specific binding of the IscB polypeptide nuclease to a target sequence in the genomic loci of the unique gene, inducing cleavage of the genomic loci by the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, and confirming different knockout mutations in a plurality of unique genes in each cell of the population of cells thereby generating a gene knockout cell library. The invention comprehends that the population of cells is a population of eukaryotic cells, and in a preferred embodiment, the population of cells is a population of embryonic stem (ES) cells.

[0673] The one or more vectors may be plasmid vectors. The vector may be a single vector comprising a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, a ω RNA, and optionally, a selection marker into target cells. Not being bound by a theory, the ability to simultaneously deliver a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and hRNA through a single vector enables application to any cell type of interest, without the need to first generate cell lines that express the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease. The regulatory element may be an inducible promoter. The inducible promoter may be a doxycycline inducible promoter. In some methods of the invention the expression of the guide sequence is under the control of the T7 promoter and is driven by the expression of T7 polymerase. The confirming of different knockout mutations may be by whole exome sequencing. The knockout mutation may be achieved in 100 or more unique genes. The knockout mutation may be achieved in 1000 or more unique genes. The knockout mutation may be achieved in 20,000 or more unique genes. The knockout mutation may be achieved in the entire genome. The knockout of gene function may be achieved in a plurality of unique genes which function in a particular physiological pathway or condition. The pathway or condition may be an immune pathway or condition. The pathway or condition may be a cell division pathway or condition.

[0674] Useful in the practice of the instant invention utilizing IscB polypeptide or CRISPR-associated IscB polypeptide nuclease complexes are methods used in compositions and reference is made to: Genome-Scale CRISPR-Cas Knockout Screening in Human Cells. Shalem, O., Sanjana, NE., Hartenian, E., Shi, X., Scott, DA., Mikkelsen, T., Heckl, D., Ebert, BL., Root, DE., Doench, JG., Zhang, F. Science Dec 12. (2013). [Epub ahead of print]; Published in final edited form as: Science. 2014 Jan 3; 343(6166): 84–87. Shalem et al. involves a new way to interrogate gene function on a genome-wide scale. Their studies showed that delivery of a genome-scale CRISPR-Cas knockout (GeCKO) library targeted 18,080 genes with 64,751 unique guide sequences enabled both negative and positive selection screening in human cells. First, the authors showed use of the GeCKO library to identify genes essential for cell viability in cancer and pluripotent stem cells. Next, in a melanoma model, the authors screened for genes whose loss is involved in resistance to vemurafenib, a therapeutic that inhibits mutant protein kinase BRAF. Their studies showed that the highest-ranking candidates included previously validated genes NF1 and MED12 as well as novel hits NF2, CUL3, TADA2B, and TADA1. The authors observed a high level of consistency between independent guide RNAs targeting the same gene and a high rate of hit confirmation, and thus demonstrated the promise of genome-scale screening with IscB polypeptide nuclease.

[0675] Reference is also made to US patent publication number US20140357530; and PCT Patent Publication WO2014093701, hereby incorporated herein by reference. Reference is also made to NIH Press Release of Oct. 22, 2015 entitled, “Researchers identify potential alternative to CRISPR-Cas genome editing tools: New Cas enzymes shed light on evolution of CRISPR-Cas systems, which is incorporated by reference.

Functional Alteration and Screening

[0676] In another aspect, the present invention provides for a method of functional evaluation and screening of genes. The use of the compositions to precisely deliver functional domains, to activate or repress genes or to alter epigenetic state by precisely altering the methylation site on a specific locus of interest, can be with one or more gRNAs or guide RNAs applied to a single cell or population of cells or with a library applied to genome in a pool of cells *ex vivo* or *in vivo* comprising the administration or expression of a library comprising a plurality of hRNAs (comprising guide molecules) and wherein the screening further comprises use of a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, wherein the

complex comprising the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease is modified to comprise a heterologous functional domain. In an aspect the invention provides a method for screening a genome comprising the administration to a host or expression in a host *in vivo* of a library. In an aspect the invention provides a method as herein discussed further comprising an activator administered to the host or expressed in the host. In an aspect the invention provides a method as herein discussed wherein the activator is attached to a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease. In an aspect the invention provides a method as herein discussed wherein the activator is attached to the N terminus or the C terminus of the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease. In an aspect the invention provides a method as herein discussed wherein the activator is attached to a ω RNA loop. In an aspect the invention provides a method as herein discussed further comprising a repressor administered to the host or expressed in the host. In an aspect the invention provides a method as herein discussed, wherein the screening comprises affecting and detecting gene activation, gene inhibition, or cleavage in the locus.

[0677] It is also preferred to target endogenous (regulatory) control elements (such as enhancers and silencers) e.g. in addition to a promoter or promoter-proximal elements. Thus, the invention can also be used to target endogenous control elements (including enhancers and silencers) in addition to targeting of the promoter. These control elements can be located upstream and downstream of the transcriptional start site (TSS), starting from 200bp from the TSS to 100kb away. Targeting of known control elements can be used to activate or repress the gene of interest. In some cases, a single control element can influence the transcription of multiple target genes. Targeting of a single control element could therefore be used to control the transcription of multiple genes simultaneously.

[0678] Targeting of putative control elements on the other hand (e.g. by tiling the region of the putative control element as well as 200bp up to 100kB around the element) can be used as a means to verify such elements (by measuring the transcription of the gene of interest) or to detect novel control elements (e.g. by tiling 100kb upstream and downstream of the TSS of the gene of interest). In addition, targeting of putative control elements can be useful in the context of understanding genetic causes of disease. Many mutations and common SNP variants associated with disease phenotypes are located outside coding regions. Targeting of such regions with either the activation or repression systems described herein can be followed by

readout of transcription of either a) a set of putative targets (e.g. a set of genes located in closest proximity to the control element) or b) whole-transcriptome readout by e.g. RNAseq or microarray. This would allow for the identification of likely candidate genes involved in the disease phenotype. Such candidate genes could be useful as novel drug targets.

[0679] Histone acetyltransferase (HAT) inhibitors are mentioned herein. However, an alternative In one embodiment is for the one or more functional domains to comprise an acetyltransferase, preferably a histone acetyltransferase. These are useful in the field of epigenomics, for example in methods of interrogating the epigenome. Methods of interrogating the epigenome may include, for example, targeting epigenomic sequences. Targeting epigenomic sequences may include the guide being directed to an epigenomic target sequence. Epigenomic target sequence may include, In one embodiment, include a promoter, silencer or an enhancer sequence.

Saturating Mutagenesis

[0680] The compositions herein can be used to perform saturating or deep scanning mutagenesis of genomic loci in conjunction with a cellular phenotype—for instance, for determining critical minimal features and discrete vulnerabilities of functional elements required for gene expression, drug resistance, and reversal of disease. By saturating or deep scanning mutagenesis is meant that every or essentially every DNA base is cut within the genomic loci. A library of Cas1 effector protein guide RNAs may be introduced into a population of cells. The library may be introduced, such that each cell receives a single hRNA. In the case where the library is introduced by transduction of a viral vector, as described herein, a low multiplicity of infection (MOI) is used. The library may include ω RNAs targeting every sequence upstream of a TAM sequence in a genomic locus. The library may include at least 100 non-overlapping genomic sequences upstream of a TAM sequence for every 1000 base pairs within the genomic locus. The library may include hRNAs targeting sequences upstream of at least one different TAM sequence. The composition may include more than one IscB polypeptide nuclease. Any IscB polypeptide or CRISPR-associated IscB polypeptide nuclease protein as described herein, including orthologues or engineered IscB polypeptide or CRISPR-associated IscB polypeptide nuclease that recognize different TAM sequences may be used. The frequency of off target sites for a ω RNA may be less than 500. Off target scores may be generated to select ω RNAs with the lowest off target sites. Any phenotype determined to be

associated with cutting at a hRNA target site may be confirmed by using sgRNAs targeting the same site in a single experiment. Validation of a target site may also be performed by using a modified IscB polypeptide or CRISPR-associated IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, as described herein, and two hRNAs targeting the genomic site of interest. Not being bound by a theory, a target site is a true hit if the change in phenotype is observed in validation experiments.

[0681] The genomic loci may include at least one continuous genomic region. The at least one continuous genomic region may comprise up to the entire genome. The at least one continuous genomic region may comprise a functional element of the genome. The functional element may be within a non-coding region, coding gene, intronic region, promoter, or enhancer. The at least one continuous genomic region may comprise at least 1 kb, preferably at least 50 kb of genomic DNA. The at least one continuous genomic region may comprise a transcription factor binding site. The at least one continuous genomic region may comprise a region of DNase I hypersensitivity. The at least one continuous genomic region may comprise a transcription enhancer or repressor element. The at least one continuous genomic region may comprise a site enriched for an epigenetic signature. The at least one continuous genomic DNA region may comprise an epigenetic insulator. The at least one continuous genomic region may comprise two or more continuous genomic regions that physically interact. Genomic regions that interact may be determined by '4C technology'. 4C technology allows the screening of the entire genome in an unbiased manner for DNA segments that physically interact with a DNA fragment of choice, as is described in Zhao et al. ((2006) Nat Genet 38, 1341-7) and in U.S. patent 8,642,295, both incorporated herein by reference in its entirety. The epigenetic signature may be histone acetylation, histone methylation, histone ubiquitination, histone phosphorylation, DNA methylation, or a lack thereof.

[0682] The compositions for saturating or deep scanning mutagenesis can be used in a population of cells. The compositions can be used in eukaryotic cells, including but not limited to mammalian and plant cells. The population of cells may be prokaryotic cells. The population of eukaryotic cells may be a population of embryonic stem (ES) cells, neuronal cells, epithelial cells, immune cells, endocrine cells, muscle cells, erythrocytes, lymphocytes, plant cells, or yeast cells.

[0683] In one aspect, the present invention provides for a method of screening for functional elements associated with a change in a phenotype. The library may be introduced into a population of cells that are adapted to contain a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease. The cells may be sorted into at least two groups based on the phenotype. The phenotype may be expression of a gene, cell growth, or cell viability. The relative representation of the ω RNAs or guide RNAs present in each group are determined, whereby genomic sites associated with the change in phenotype are determined by the representation of ω RNAs or guide RNAs present in each group. The change in phenotype may be a change in expression of a gene of interest. The gene of interest may be upregulated, downregulated, or knocked out. The cells may be sorted into a high expression group and a low expression group. The population of cells may include a reporter construct that is used to determine the phenotype. The reporter construct may include a detectable marker. Cells may be sorted by use of the detectable marker.

[0684] In another aspect, the present invention provides for a method of screening for genomic sites associated with resistance to a chemical compound. The chemical compound may be a drug or pesticide. The library may be introduced into a population of cells that are adapted to contain a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, wherein each cell of the population contains no more than one ω RNA or guide RNA; the population of cells are treated with the chemical compound; and the representation of ω RNA or guide RNAs are determined after treatment with the chemical compound at a later time point as compared to an early time point, whereby genomic sites associated with resistance to the chemical compound are determined by enrichment of ω RNAs. Representation of ω RNAs may be determined by deep sequencing methods.

[0685] Useful in the practice of the instant invention utilizing compositions are methods used in compositions and reference is made to the article entitled BCL11A enhancer dissection by Cas-mediated in situ saturating mutagenesis. Canver, M.C., Smith, E.C., Sher, F., Pinello, L., Sanjana, N.E., Shalem, O., Chen, D.D., Schupp, P.G., Vinjamur, D.S., Garcia, S.P., Luc, S., Kurita, R., Nakamura, Y., Fujiwara, Y., Maeda, T., Yuan, G., Zhang, F., Orkin, S.H., & Bauer, D.E. DOI:10.1038/nature15521, published online September 16, 2015, the article is herein incorporated by reference and discussed briefly below:

[0686] Canver et al. involves novel pooled guide RNA libraries to perform in situ saturating mutagenesis of the human and mouse BCL11A erythroid enhancers previously identified as an enhancer associated with fetal hemoglobin (HbF) level and whose mouse ortholog is necessary for erythroid BCL11A expression. This approach revealed critical minimal features and discrete vulnerabilities of these enhancers. Through editing of primary human progenitors and mouse transgenesis, the authors validated the BCL11A erythroid enhancer as a target for HbF reinduction. The authors generated a detailed enhancer map that informs therapeutic genome editing.

Modification of a Cell or Organism

[0687] The present disclosure further provides cells comprising one or more components of the systems herein, e.g., the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and/or ω RNA(s). Also provided include cells modified by the systems and methods herein, and cell cultures, tissues, organs, organism comprising such cells or progeny thereof. The invention In one embodiment comprehends a method of modifying a cell or organism. The cell may be a prokaryotic cell or a eukaryotic cell. The cell may be a mammalian cell. The mammalian cell may be a non-human primate, bovine, porcine, rodent or mouse cell. The cell may be a non-mammalian eukaryotic cell such as poultry, fish or shrimp. The cell may also be a plant cell. The plant cell may be of a crop plant such as cassava, corn, sorghum, wheat, or rice. The plant cell may also be of an algae, tree or vegetable. The modification introduced to the cell by the present invention may be such that the cell and progeny of the cell are altered for improved production of biologic products such as an antibody, starch, alcohol or other desired cellular output. The modification introduced to the cell by the present invention may be such that the cell and progeny of the cell include an alteration that changes the biologic product produced.

THERAPEUTIC USES AND METHODS OF TREATMENT

[0688] Also provided herein are methods of diagnosing, prognosing, treating, and/or preventing a disease, state, or condition in or of a subject. Generally, the methods of diagnosing, prognosing, treating, and/or preventing a disease, state, or condition in or of a subject can include modifying a polynucleotide in a subject or cell thereof using a composition, system, or component thereof described herein and/or include detecting a diseased or healthy polynucleotide in a subject or cell thereof using a composition, system, or component thereof

described herein. In one embodiment, the method of treatment or prevention can include using a composition, system, or component thereof to modify a polynucleotide of an infectious organism (e.g. bacterial or virus) within a subject or cell thereof. In one embodiment, the method of treatment or prevention can include using a composition, system, or component thereof to modify a polynucleotide of an infectious organism or symbiotic organism within a subject. The composition, system, and components thereof can be used to develop models of diseases, states, or conditions. The composition, system, and components thereof can be used to detect a disease state or correction thereof, such as by a method of treatment or prevention described herein. The composition, system, and components thereof can be used to screen and select cells that can be used, for example, as treatments or preventions described herein. The composition, system, and components thereof can be used to develop biologically active agents that can be used to modify one or more biologic functions or activities in a subject or a cell thereof.

[0689] In general, the method can include delivering a composition, system, and/or component thereof to a subject or cell thereof, or to an infectious or symbiotic organism by a suitable delivery technique and/or composition. Once administered the components can operate as described elsewhere herein to elicit a nucleic acid modification event. In some aspects, the nucleic acid modification event can occur at the genomic, epigenomic, and/or transcriptomic level. DNA and/or RNA cleavage, gene activation, and/or gene deactivation can occur. Additional features, uses, and advantages are described in greater detail below. On the basis of this concept, several variations are appropriate to elicit a genomic locus event, including DNA cleavage, gene activation, or gene deactivation. Using the provided compositions, the person skilled in the art can advantageously and specifically target single or multiple loci with the same or different functional domains to elicit one or more genomic locus events. In addition to treating and/or preventing a disease in a subject, the compositions may be applied in a wide variety of methods for screening in libraries in cells and functional modeling *in vivo* (e.g. gene activation of lincRNA and identification of function; gain-of-function modeling; loss-of-function modeling; the use the compositions of the invention to establish cell lines and transgenic animals for optimization and screening purposes).

[0690] The composition, system, and components thereof described elsewhere herein can be used to treat and/or prevent a disease, such as a genetic and/or epigenetic disease, in a subject. The composition, system, and components thereof described elsewhere herein can be

used to treat and/or prevent genetic infectious diseases in a subject, such as bacterial infections, viral infections, fungal infections, parasite infections, and combinations thereof. The composition, system, and components thereof described elsewhere herein can be used to modify the composition or profile of a microbiome in a subject, which can in turn modify the health status of the subject. The composition, system, described herein can be used to modify cells *ex vivo*, which can then be administered to the subject whereby the modified cells can treat or prevent a disease or symptom thereof. This is also referred to in some contexts as adoptive therapy. The composition, system, described herein can be used to treat mitochondrial diseases, where the mitochondrial disease etiology involves a mutation in the mitochondrial DNA.

[0691] Also provided is a method of treating a subject, e.g., a subject in need thereof, comprising inducing gene editing by transforming the subject with the polynucleotide encoding one or more components of the composition, system, or complex or any of polynucleotides or vectors described herein and administering them to the subject. A suitable repair template may also be provided, for example delivered by a vector comprising said repair template. The repair template may be a recombination template herein. Also provided is a method of treating a subject, e.g., a subject in need thereof, comprising inducing transcriptional activation or repression of multiple target gene loci by transforming the subject with the polynucleotides or vectors described herein, wherein said polynucleotide or vector encodes or comprises one or more components of composition, system, complex or component thereof comprising multiple IscB polypeptide nucleases. Where any treatment is occurring *ex vivo*, for example in a cell culture, then it will be appreciated that the term ‘subject’ may be replaced by the phrase “cell or cell culture.”

[0692] Also provided is a method of treating a subject, e.g., a subject in need thereof, comprising inducing gene editing by transforming the subject with the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease(s), advantageously encoding and expressing *in vivo* the remaining portions of the composition, system, (e.g., RNA, guides). A suitable repair template may also be provided, for example delivered by a vector comprising said repair template. Also provided is a method of treating a subject, e.g., a subject in need thereof, comprising inducing transcriptional activation or repression by transforming the subject with the IscB polypeptide nuclease(s) advantageously encoding and expressing *in vivo* the remaining portions of the composition, system, (e.g., RNA, guides); advantageously In one

embodiment the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease is a catalytically inactive IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and includes one or more associated functional domains. Where any treatment is occurring *ex vivo*, for example in a cell culture, then it will be appreciated that the term ‘subject’ may be replaced by the phrase “cell or cell culture.”

[0693] One or more components of the composition and system described herein can be included in a composition, such as a pharmaceutical composition, and administered to a host individually or collectively. Alternatively, these components may be provided in a single composition for administration to a host. Administration to a host may be performed via viral vectors known to the skilled person or described herein for delivery to a host (e.g. lentiviral vector, adenoviral vector, AAV vector). As explained herein, use of different selection markers (e.g. for lentiviral hRNA selection) and concentration of ω RNA (e.g. dependent on whether multiple ω RNAs are used) may be advantageous for eliciting an improved effect.

[0694] Thus, also described herein are methods of inducing one or more polynucleotide modifications in a eukaryotic or prokaryotic cell or component thereof (e.g., a mitochondria) of a subject, infectious organism, and/or organism of the microbiome of the subject. The modification can include the introduction, deletion, or substitution of one or more nucleotides at a target sequence of a polynucleotide of one or more cell(s). The modification can occur *in vitro*, *ex vivo*, *in situ*, or *in vivo*.

[0695] In one embodiment, the method of treating or inhibiting a condition or a disease caused by one or more mutations in a genomic locus in a eukaryotic organism or a non-human organism can include manipulation of a target sequence within a coding, non-coding or regulatory element of said genomic locus in a target sequence in a subject or a non-human subject in need thereof comprising modifying the subject or a non-human subject by manipulation of the target sequence and wherein the condition or disease is susceptible to treatment or inhibition by manipulation of the target sequence including providing treatment comprising delivering a composition comprising the particle delivery system or the delivery system or the virus particle of any one of the above embodiment or the cell of any one of the above embodiment.

[0696] Also provided herein is the use of the particle delivery system or the delivery system or the virus particle of any one of the above embodiments or the cell of any one of the above

embodiment in *ex vivo* or *in vivo* gene or genome editing; or for use in *in vitro*, *ex vivo* or *in vivo* gene therapy. Also provided herein are particle delivery systems, non-viral delivery systems, and/or the virus particle of any one of the above embodiments or the cell of any one of the above embodiments used in the manufacture of a medicament for *in vitro*, *ex vivo* or *in vivo* gene or genome editing or for use in *in vitro*, *ex vivo* or *in vivo* gene therapy or for use in a method of modifying an organism or a non-human organism by manipulation of a target sequence in a genomic locus associated with a disease or in a method of treating or inhibiting a condition or disease caused by one or more mutations in a genomic locus in a eukaryotic organism or a non- human organism.

[0697] In one embodiment, polynucleotide modification can include the introduction, deletion, or substitution of 1-75 nucleotides at each target sequence of said polynucleotide of said cell(s). The modification can include the introduction, deletion, or substitution of at least 1, 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, or 75 nucleotides at each target sequence. The modification can include the introduction, deletion, or substitution of at least 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, or 75 nucleotides at each target sequence of said cell(s). The modification can include the introduction, deletion, or substitution of at least 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, or 75 nucleotides at each target sequence of said cell(s). The modification can include the introduction, deletion, or substitution of at least 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, or 75 nucleotides at each target sequence of said cell(s). The modification can include the introduction, deletion, or substitution of at least 40, 45, 50, 75, 100, 200, 300, 400 or 500 nucleotides at each target sequence of said cell(s). The modification can include the introduction, deletion, or substitution of at least 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 3000, 3100, 3200, 3300, 3400, 3500, 3600, 3700, 3800, 3900, 4000, 4100, 4200, 4300, 4400, 4500, 4600, 4700, 4800, 4900, 5000, 5100, 5200, 5300, 5400, 5500, 5600, 5700, 5800, 5900, 6000, 6100, 6200, 6300, 6400, 6500, 6600, 6700, 6800, 6900, 7000, 7100, 7200, 7300, 7400, 7500, 7600, 7700, 7800, 7900, 8000, 8100, 8200, 8300, 8400, 8500, 8600, 8700, 8800, 8900, 9000, 9100, 9200, 9300, 9400, 9500, 9600, 9700, 9800, or 9900 to 10000 nucleotides at each target sequence of said cell(s).

[0698] In one embodiment, the modifications can include the introduction, deletion, or substitution of nucleotides at each target sequence of said cell(s) via nucleic acid components (e.g., guide(s) RNA(s) or ω RNA(s)), such as those mediated by a composition, system, or a component thereof described elsewhere herein. In one embodiment, the modifications can include the introduction, deletion, or substitution of nucleotides at a target or random sequence of said cell(s) via a composition, system, or technique.

[0699] In one embodiment, the composition, system, or component thereof can promote Non-Homologous End-Joining (NHEJ). In one embodiment, modification of a polynucleotide by a composition, system, or a component thereof, such as a diseased polynucleotide, can include NHEJ. In one embodiment, promotion of this repair pathway by the composition, system, or a component thereof can be used to target gene or polynucleotide specific knock-outs and/or knock-ins. In one embodiment, promotion of this repair pathway by the composition, system, or a component thereof can be used to generate NHEJ-mediated indels. Nuclease-induced NHEJ can also be used to remove (e.g., delete) sequence in a gene of interest. Generally, NHEJ repairs a double-strand break in the DNA by joining together the two ends; however, generally, the original sequence is restored only if two compatible ends, exactly as they were formed by the double-strand break, are perfectly ligated. The DNA ends of the double-strand break are frequently the subject of enzymatic processing, resulting in the addition or removal of nucleotides, at one or both strands, prior to rejoining of the ends. This results in the presence of insertion and/or deletion (indel) mutations in the DNA sequence at the site of the NHEJ repair. The indel can range in size from 1-50 or more base pairs. In one embodiment the indel can be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210,

211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, or 500 base pairs or more. If a double-strand break is targeted near to a short target sequence, the deletion mutations caused by the NHEJ repair often span, and therefore remove, the unwanted nucleotides. For the deletion of larger DNA segments, introducing two double-strand breaks, one on each side of the sequence, can result in NHEJ between the ends with removal of the entire intervening sequence. Both of these approaches can be used to delete specific DNA sequences.

[0700] In one embodiment, composition, system, mediated NHEJ can be used in the method to delete small sequence motifs. In one embodiment, composition, system, mediated NHEJ can be used in the method to generate NHEJ-mediate indels that can be targeted to the gene, e.g., a coding region, e.g., an early coding region of a gene of interest can be used to knockout (i.e., eliminate expression of) a gene of interest. For example, early coding region of a gene of interest includes sequence immediately following a transcription start site, within a first exon of the coding sequence, or within 500 bp of the transcription start site (e.g., less than 500, 450, 400, 350, 300, 250, 200, 150, 100 or 50 bp). In an embodiment, in which a ω RNA or guide RNA and IscB polypeptide nuclease generate a double strand break for the purpose of inducing NHEJ-mediated indels, a ω RNA or guide RNA may be configured to position one

double-strand break in close proximity to a nucleotide of the target position. In an embodiment, the cleavage site may be between 0-500 bp away from the target position (e.g., less than 500, 400, 300, 200, 100, 50, 40, 30, 25, 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 bp from the target position). In an embodiment, in which two ω RNA or guide RNAs complexing with one or more nickases induce two single strand breaks for the purpose of inducing NHEJ-mediated indels, two guide RNAs may be configured to position two single-strand breaks to provide for NHEJ repair a nucleotide of the target position.

[0701] For minimization of toxicity and off-target effect, it may be important to control the concentration of IscB polypeptide or CRISPR-associated IscB polypeptide nuclease mRNA and guide RNA delivered. Optimal concentrations of IscB polypeptide or CRISPR-associated IscB polypeptide nuclease mRNA and guide RNA can be determined by testing different concentrations in a cellular or non-human eukaryote animal model and using deep sequencing to analyze the extent of modification at potential off-target genomic loci. Alternatively, to minimize the level of toxicity and off-target effect, nickase mRNA (for example *S. pyogenes* Cas9 with the D10A mutation) can be delivered with a pair of guide RNAs targeting a site of interest. Guide sequences and strategies to minimize toxicity and off-target effects can be as in International Patent Publication No. WO 2014/093622 (PCT/US2013/074667); or, via mutation. Others are as described elsewhere herein.

[0702] Typically, in the context of an endogenous IscB polypeptide nuclease, formation of a IscB polypeptide nuclease or complex (comprising a guide sequence hybridized to a target sequence and complexed with one or more IscB polypeptide or CRISPR-associated IscB polypeptide nucleases) results in cleavage, nicking, and/or another modification of one or both strands in or near (e.g. within 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or more base pairs from) the target sequence.

[0703] In one embodiment, a method of modifying a target polynucleotide in a cell to treat or prevent a disease can include allowing a composition, system, or component thereof to bind to the target polynucleotide, e.g., to effect cleavage, nicking, or other modification as the composition, system, is capable of said target polynucleotide, thereby modifying the target polynucleotide, wherein the composition, system, or component thereof, complex with a guide sequence, and hybridize said guide sequence to a target sequence within the target polynucleotide, wherein said guide sequence is optionally linked to a ω RNA scaffold sequence.

In some of these embodiments, the composition, system, or component thereof can be or include a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease complexed with a guide sequence. In one embodiment, modification can include cleaving or nicking one or two strands at the location of the target sequence by one or more components of the composition, system, or component thereof.

[0704] The cleavage, nicking, or other modification capable of being performed by the composition, system, can modify transcription of a target polynucleotide. In one embodiment, modification of transcription can include decreasing transcription of a target polynucleotide. In one embodiment, modification can include increasing transcription of a target polynucleotide. In one embodiment, the method includes repairing said cleaved target polynucleotide by homologous recombination with a recombination template polynucleotide, wherein said repair results in a modification such as, but not limited to, an insertion, deletion, or substitution of one or more nucleotides of said target polynucleotide. In one embodiment, said modification results in one or more amino acid changes in a protein expressed from a gene comprising the target sequence. In one embodiment, the modification imparted by the composition, system, or component thereof provides a transcript and/or protein that can correct a disease or a symptom thereof, including but not limited to, any of those described in greater detail elsewhere herein.

[0705] In one embodiment, the method of treating or preventing a disease can include delivering one or more vectors or vector systems to a cell, such as a eukaryotic or prokaryotic cell, wherein one or more vectors or vector systems include the composition, system, or component thereof. In one embodiment, the vector(s) or vector system(s) can be a viral vector or vector system, such as an AAV or lentiviral vector system, which are described in greater detail elsewhere herein. In one embodiment, the method of treating or preventing a disease can include delivering one or more viral particles, such as an AAV or lentiviral particle, containing the composition, system, or component thereof. In one embodiment, the viral particle has a tissue specific tropism. In one embodiment, the viral particle has a liver, muscle, eye, heart, pancreas, kidney, neuron, epithelial cell, endothelial cell, astrocyte, glial cell, immune cell, or red blood cell specific tropism.

[0706] It will be understood that the composition and system, according to the invention as described herein, such as the composition and system, for use in the methods according to the invention as described herein, may be suitably used for any type of application known for

composition, system, preferably in eukaryotes. In certain aspects, the application is therapeutic, preferably therapeutic in a eukaryote organism, such as including but not limited to animals (including human), plants, algae, fungi (including yeasts), etc. Alternatively, or in addition, in certain aspects, the application may involve accomplishing or inducing one or more particular traits or characteristics, such as genotypic and/or phenotypic traits or characteristics, as also described elsewhere herein.

Treating Diseases of the Circulatory System

[0707] In one embodiment, the composition, system, and/or component thereof described herein can be used to treat and/or prevent a circulatory system disease. Exemplary disease is provided, for example, in Table 6. In one embodiment the plasma exosomes of Wahlgren et al. (Nucleic Acids Research, 2012, Vol. 40, No. 17 e130) can be used to deliver the composition, system, and/or component thereof described herein to the blood. In one embodiment, the circulatory system disease can be treated by using a lentivirus to deliver the composition, system, described herein to modify hematopoietic stem cells (HSCs) *in vivo* or *ex vivo* (see e.g. Drakopoulou, “Review Article, The Ongoing Challenge of Hematopoietic Stem Cell-Based Gene Therapy for β -Thalassemia,” Stem Cells International, Volume 2011, Article ID 987980, 10 pages, doi:10.4061/2011/987980, which can be adapted for use with the composition, system, herein in view of the description herein). In one embodiment, the circulatory system disorder can be treated by correcting HSCs as to the disease using a composition, system, herein or a component thereof, wherein the composition, system, optionally includes a suitable HDR repair template (see e.g. Cavazzana, “Outcomes of Gene Therapy for β -Thalassemia Major via Transplantation of Autologous Hematopoietic Stem Cells Transduced *Ex Vivo* with a Lentiviral β A-T87Q-Globin Vector.”; Cavazzana-Calvo, “Transfusion independence and HMGA2 activation after gene therapy of human β -thalassaemia”, Nature 467, 318–322 (16 September 2010) doi:10.1038/nature09328; Nienhuis, “Development of Gene Therapy for Thalassemia, Cold Spring Harbor Perspectives in Medicine, doi: 10.1101/cshperspect.a011833 (2012), LentiGlobin BB305, a lentiviral vector containing an engineered β -globin gene (β A-T87Q); and Xie et al., “Seamless gene correction of β -thalassaemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyback” Genome Research gr.173427.114 (2014) www.genome.org/cgi/doi/10.1101/gr.173427.114 (Cold Spring Harbor Laboratory Press; [1599] Watts, “Hematopoietic Stem Cell Expansion and Gene Therapy” Cytotherapy 13(10):1164–1171. doi:10.3109/14653249.2011.620748 (2011), which can be adapted for use

with the composition, system, herein in view of the description herein). In one embodiment, iPSCs can be modified using a composition, system, described herein to correct a disease polynucleotide associated with a circulatory disease. In this regard, the teachings of Xu et al. (Sci Rep. 2015 Jul 9;5:12065. doi: 10.1038/srep12065) and Song et al. (Stem Cells Dev. 2015 May 1;24(9):1053-65. doi: 10.1089/scd.2014.0347. Epub 2015 Feb 5) with respect to modifying iPSCs can be adapted for use in view of the description herein with the composition, system, described herein.

[0708] The term “Hematopoietic Stem Cell” or “HSC” refers broadly those cells considered to be an HSC, e.g., blood cells that give rise to all the other blood cells and are derived from mesoderm; located in the red bone marrow, which is contained in the core of most bones. HSCs of the invention include cells having a phenotype of hematopoietic stem cells, identified by small size, lack of lineage (lin) markers, and markers that belong to the cluster of differentiation series, like: CD34, CD38, CD90, CD133, CD105, CD45, and also c-kit, - the receptor for stem cell factor. Hematopoietic stem cells are negative for the markers that are used for detection of lineage commitment, and are, thus, called Lin⁻; and, during their purification by FACS, a number of up to 14 different mature blood-lineage markers, e.g., CD13 & CD33 for myeloid, CD71 for erythroid, CD19 for B cells, CD61 for megakaryocytic, etc. for humans; and, B220 (murine CD45) for B cells, Mac-1 (CD11b/CD18) for monocytes, Gr-1 for Granulocytes, Ter119 for erythroid cells, Il7Ra, CD3, CD4, CD5, CD8 for T cells, etc. Mouse HSC markers: CD34^{lo/-}, SCA-1⁺, Thy1.1^{+/lo}, CD38⁺, C-kit⁺, lin⁻, and Human HSC markers: CD34⁺, CD59⁺, Thy1/CD90⁺, CD38^{lo/-}, C-kit/CD117⁺, and lin⁻. HSCs are identified by markers. Hence in embodiments discussed herein, the HSCs can be CD34⁺ cells. HSCs can also be hematopoietic stem cells that are CD34⁻/CD38⁻. Stem cells that may lack c-kit on the cell surface that are considered in the art as HSCs are within the ambit of the invention, as well as CD133⁺ cells likewise considered HSCs in the art.

[0709] In one embodiment, the treatment or prevention for treating a circulatory system or blood disease can include modifying a human cord blood cell with any modification described herein. In one embodiment, the treatment or prevention for treating a circulatory system or blood disease can include modifying a granulocyte colony-stimulating factor-mobilized peripheral blood cell (mPB) with any modification described herein. In one embodiment, the human cord blood cell or mPB can be CD34⁺. In one embodiment, the cord blood cell(s) or mPB cell(s) modified can be autologous. In one embodiment, the cord blood cell(s) or mPB

cell(s) can be allogenic. In addition to the modification of the disease gene(s), allogenic cells can be further modified using the composition, system, described herein to reduce the immunogenicity of the cells when delivered to the recipient. Such techniques are described elsewhere herein and e.g. Cartier, “MINI-SYMPOSIUM: X-Linked Adrenoleukodystrophy, Hematopoietic Stem Cell Transplantation and Hematopoietic Stem Cell Gene Therapy in X-Linked Adrenoleukodystrophy,” *Brain Pathology* 20 (2010) 857–862, which can be adapted for use with the composition, system, herein. The modified cord blood cell(s) or mPB cell(s) can be optionally expanded *in vitro*. The modified cord blood cell(s) or mPB cell(s) can be derived to a subject in need thereof using any suitable delivery technique.

[0710] The compositions may be engineered to target genetic locus or loci in HSCs. In one embodiment, the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease(s) can be codon-optimized for a eukaryotic cell and especially a mammalian cell, e.g., a human cell, for instance, HSC, or iPSC and hRNA targeting a locus or loci in HSC, such as circulatory disease, can be prepared. These may be delivered via particles. The particles may be formed by the IscB polypeptide nuclease and the ω RNA being admixed. The ω RNA and IscB polypeptide nuclease mixture can be, for example, admixed with a mixture comprising or consisting essentially of or consisting of surfactant, phospholipid, biodegradable polymer, lipoprotein and alcohol, whereby particles containing the ω RNA and IscB polypeptide or CRISPR-associated IscB polypeptide nuclease may be formed. The invention comprehends so making particles and particles from such a method as well as uses thereof. Particles suitable delivery of the composition in the context of blood or circulatory system or HSC delivery to the blood or circulatory system are described in greater detail elsewhere herein.

[0711] In one embodiment, after *ex vivo* modification the HSCs or iPCS can be expanded prior to administration to the subject. Expansion of HSCs can be via any suitable method such as that described by, Lee, “Improved *ex vivo* expansion of adult hematopoietic stem cells by overcoming CUL4-mediated degradation of HOXB4.” *Blood*. 2013 May 16;121(20):4082-9. doi: 10.1182/blood-2012-09-455204. Epub 2013 Mar 21.

[0712] In one embodiment, the HSCs or iPSCs modified can be autologous. In one embodiment, the HSCs or iPSCs can be allogenic. In addition to the modification of the disease gene(s), allogenic cells can be further modified using the composition, system, described herein to reduce the immunogenicity of the cells when delivered to the recipient. Such techniques are

described elsewhere herein and e.g. Cartier, "MINI-SYMPOSIUM: X-Linked Adrenoleukodystrophypa, Hematopoietic Stem Cell Transplantation and Hematopoietic Stem Cell Gene Therapy in X-Linked Adrenoleukodystrophy," Brain Pathology 20 (2010) 857–862, which can be adapted for use with the composition, system, herein.

Treating Neurological Diseases

[0713] In one embodiment, the compositions, systems, described herein can be used to treat diseases of the brain and CNS. Delivery options for the brain include encapsulation of IscB polypeptide nuclease and guide RNA in the form of either DNA or RNA into liposomes and conjugating to molecular Trojan horses for trans-blood brain barrier (BBB) delivery. Molecular Trojan horses have been shown to be effective for delivery of B-gal expression vectors into the brain of non-human primates. The same approach can be used to delivery vectors containing IscB polypeptide nuclease and guide RNA. For instance, Xia CF and Boado RJ, Pardridge WM ("Antibody-mediated targeting of siRNA via the human insulin receptor using avidin-biotin technology." Mol Pharm. 2009 May-Jun;6(3):747-51. doi: 10.1021/mp800194) describes how delivery of short interfering RNA (siRNA) to cells in culture, and *in vivo*, is possible with combined use of a receptor-specific monoclonal antibody (mAb) and avidin-biotin technology. The authors also report that because the bond between the targeting mAb and the siRNA is stable with avidin-biotin technology, and RNAi effects at distant sites such as brain are observed *in vivo* following an intravenous administration of the targeted siRNA, the teachings of which can be adapted for use with the compositions, systems, herein. In other embodiments, an artificial virus can be generated for CNS and/or brain delivery. See e.g. Zhang et al. (Mol Ther. 2003 Jan;7(1):11-8.)), the teachings of which can be adapted for use with the compositions, systems, herein.

Treating Hearing Diseases

[0714] In one embodiment the composition and system described herein can be used to treat a hearing disease or hearing loss in one or both ears. Deafness is often caused by lost or damaged hair cells that cannot relay signals to auditory neurons. In such cases, cochlear implants may be used to respond to sound and transmit electrical signals to the nerve cells. But these neurons often degenerate and retract from the cochlea as fewer growth factors are released by impaired hair cells.

[0715] In one embodiment, the composition, system, or modified cells can be delivered to one or both ears for treating or preventing hearing disease or loss by any suitable method or

technique. Suitable methods and techniques include, but are not limited to those set forth in US Patent Publication No. 20120328580 describes injection of a pharmaceutical composition into the ear (e.g., auricular administration), such as into the luminae of the cochlea (e.g., the Scala media, Sc vestibulae, and Sc tympani), e.g., using a syringe, e.g., a single-dose syringe. For example, one or more of the compounds described herein can be administered by intratympanic injection (e.g., into the middle ear), and/or injections into the outer, middle, and/or inner ear; administration in situ, via a catheter or pump (see e.g. McKenna et al., (U.S. Patent Publication No. 2006/0030837) and Jacobsen et al., (U.S. Pat. No. 7,206,639); administration in combination with a mechanical device such as a cochlear implant or a hearing aid, which is worn in the outer ear (see e.g. U.S. Patent Publication No. 2007/0093878, which provides an exemplary cochlear implant suitable for delivery of the compositions, systems, described herein to the ear). Such methods are routinely used in the art, for example, for the administration of steroids and antibiotics into human ears. Injection can be, for example, through the round window of the ear or through the cochlear capsule. Other inner ear administration methods are known in the art (see, e.g., Salt and Plontke, Drug Discovery Today, 10:1299-1306, 2005). In one embodiment, a catheter or pump can be positioned, e.g., in the ear (e.g., the outer, middle, and/or inner ear) of a patient during a surgical procedure. In one embodiment, a catheter or pump can be positioned, e.g., in the ear (e.g., the outer, middle, and/or inner ear) of a patient without the need for a surgical procedure.

[0716] In general, the cell therapy methods described in US Patent Publication No. 20120328580 can be used to promote complete or partial differentiation of a cell to or towards a mature cell type of the inner ear (e.g., a hair cell) *in vitro*. Cells resulting from such methods can then be transplanted or implanted into a patient in need of such treatment. The cell culture methods required to practice these methods, including methods for identifying and selecting suitable cell types, methods for promoting complete or partial differentiation of selected cells, methods for identifying complete or partially differentiated cell types, and methods for implanting complete or partially differentiated cells are described below.

[0717] Cells suitable for use in the present invention include, but are not limited to, cells that are capable of differentiating completely or partially into a mature cell of the inner ear, e.g., a hair cell (e.g., an inner and/or outer hair cell), when contacted, e.g., *in vitro*, with one or more of the compounds described herein. Exemplary cells that are capable of differentiating into a hair cell include, but are not limited to stem cells (e.g., inner ear stem cells, adult stem

cells, bone marrow derived stem cells, embryonic stem cells, mesenchymal stem cells, skin stem cells, iPS cells, and fat derived stem cells), progenitor cells (e.g., inner ear progenitor cells), support cells (e.g., Deiters' cells, pillar cells, inner phalangeal cells, tectal cells and Hensen's cells), and/or germ cells. The use of stem cells for the replacement of inner ear sensory cells is described in Li et al., (U.S. Patent Publication No. 2005/0287127) and Li et al., (U.S. Patent Application No. 11/953,797). The use of bone marrow derived stem cells for the replacement of inner ear sensory cells is described in Edge et al., PCT/US2007/084654. iPS cells are described, e.g., at Takahashi et al., *Cell*, Volume 131, Issue 5, Pages 861-872 (2007); Takahashi and Yamanaka, *Cell* 126, 663-76 (2006); Okita et al., *Nature* 448, 260-262 (2007); Yu, J. et al., *Science* 318(5858):1917-1920 (2007); Nakagawa et al., *Nat. Biotechnol.* 26:101-106 (2008); and Zaehres and Scholer, *Cell* 131(5):834-835 (2007). Such suitable cells can be identified by analyzing (e.g., qualitatively or quantitatively) the presence of one or more tissue specific genes. For example, gene expression can be detected by detecting the protein product of one or more tissue-specific genes. Protein detection techniques involve staining proteins (e.g., using cell extracts or whole cells) using antibodies against the appropriate antigen. In this case, the appropriate antigen is the protein product of the tissue-specific gene expression. Although, in principle, a first antibody (i.e., the antibody that binds the antigen) can be labeled, it is more common (and improves the visualization) to use a second antibody directed against the first (e.g., an anti-IgG). This second antibody is conjugated either with fluorochromes, or appropriate enzymes for colorimetric reactions, or gold beads (for electron microscopy), or with the biotin-avidin system, so that the location of the primary antibody, and thus the antigen, can be recognized.

[0718] The composition and system may be delivered to the ear by direct application of pharmaceutical composition to the outer ear, with compositions modified from US Patent Publication No. 20110142917. In one embodiment the pharmaceutical composition is applied to the ear canal. Delivery to the ear may also be referred to as aural or otic delivery.

[0719] In one embodiment, the compositions, systems, or components thereof and/or vectors or vector systems can be delivered to ear via a transfection to the inner ear through the intact round window by a novel proteidic delivery technology which may be applied to the nucleic acid-targeting system of the present invention (see, e.g., Qi et al., *Gene Therapy* (2013), 1-9). About 40 µl of 10mM RNA may be contemplated as the dosage for administration to the ear.

[0720] According to Rejali et al. (Hear Res. 2007 Jun;228(1-2):180-7), cochlear implant function can be improved by good preservation of the spiral ganglion neurons, which are the target of electrical stimulation by the implant and brain derived neurotrophic factor (BDNF) has previously been shown to enhance spiral ganglion survival in experimentally deafened ears. Rejali et al. tested a modified design of the cochlear implant electrode that includes a coating of fibroblast cells transduced by a viral vector with a BDNF gene insert. To accomplish this type of *ex vivo* gene transfer, Rejali et al. transduced guinea pig fibroblasts with an adenovirus with a BDNF gene cassette insert, and determined that these cells secreted BDNF and then attached BDNF-secreting cells to the cochlear implant electrode via an agarose gel, and implanted the electrode in the scala tympani. Rejali et al. determined that the BDNF expressing electrodes were able to preserve significantly more spiral ganglion neurons in the basal turns of the cochlea after 48 days of implantation when compared to control electrodes and demonstrated the feasibility of combining cochlear implant therapy with *ex vivo* gene transfer for enhancing spiral ganglion neuron survival. Such a system may be applied to the nucleic acid-targeting system of the present invention for delivery to the ear.

[0721] In one embodiment, the system set forth in Mukherjea et al. (Antioxidants & Redox Signaling, Volume 13, Number 5, 2010) can be adapted for transtympanic administration of the composition, system, or component thereof to the ear. In one embodiment, a dosage of about 2 mg to about 4 mg of IscB polypeptide or CRISPR-associated IscB polypeptide nuclease for administration to a human.

[0722] In one embodiment, the system set forth in [Jung et al. (Molecular Therapy, vol. 21 no. 4, 834–841 apr. 2013) can be adapted for vestibular epithelial delivery of the composition, system, or component thereof to the ear. In one embodiment, a dosage of about 1 to about 30 mg of IscB polypeptide nuclease for administration to a human.

Treating Diseases in Non-Dividing Cells

[0723] In one embodiment, the gene or transcript to be corrected is in a non-dividing cell. Exemplary non-dividing cells are muscle cells or neurons. Non-dividing (especially non-dividing, fully differentiated) cell types present issues for gene targeting or genome engineering, for example because homologous recombination (HR) is generally suppressed in the G1 cell-cycle phase. However, while studying the mechanisms by which cells control normal DNA repair systems, Durocher discovered a previously unknown switch that keeps HR “off” in non-dividing cells and devised a strategy to toggle this switch back on. Orthwein et al.

(Daniel Durocher's lab at the Mount Sinai Hospital in Ottawa, Canada) recently reported (Nature 16142, published online 9 Dec 2015) have shown that the suppression of HR can be lifted and gene targeting successfully concluded in both kidney (293T) and osteosarcoma (U2OS) cells. Tumor suppressors, BRCA1, PALB2 and BRAC2 are known to promote DNA DSB repair by HR. They found that formation of a complex of BRCA1 with PALB2 - BRAC2 is governed by a ubiquitin site on PALB2, such that action on the site by an E3 ubiquitin ligase. This E3 ubiquitin ligase is composed of KEAP1 (a PALB2 -interacting protein) in complex with cullin-3 (CUL3)–RBX1. PALB2 ubiquitylation suppresses its interaction with BRCA1 and is counteracted by the deubiquitylase USP11, which is itself under cell cycle control. Restoration of the BRCA1–PALB2 interaction combined with the activation of DNA-end resection is sufficient to induce homologous recombination in G1, as measured by a number of methods including a IscB polypeptide nuclease-based gene-targeting assay directed at USP11 or KEAP1 (expressed from a pX459 vector). However, when the BRCA1–PALB2 interaction was restored in resection-competent G1 cells using either KEAP1 depletion or expression of the PALB2-KR mutant, a robust increase in gene-targeting events was detected. These teachings can be adapted for and/or applied to the compositions, systems, described herein.

[0724] Thus, reactivation of HR in cells, especially non-dividing, fully differentiated cell types is preferred, In one embodiment. In one embodiment, promotion of the BRCA1–PALB2 interaction is preferred In one embodiment. In one embodiment, the target cell is a non-dividing cell. In one embodiment, the target cell is a neuron or muscle cell. In one embodiment, the target cell is targeted *in vivo*. In one embodiment, the cell is in G1 and HR is suppressed. In one embodiment, use of KEAP1 depletion, for example inhibition of expression of KEAP1 activity, is preferred. KEAP1 depletion may be achieved through siRNA, for example as shown in Orthwein et al. Alternatively, expression of the PALB2-KR mutant (lacking all eight Lys residues in the BRCA1-interaction domain is preferred, either in combination with KEAP1 depletion or alone. PALB2-KR interacts with BRCA1 irrespective of cell cycle position. Thus, promotion or restoration of the BRCA1-PALB2 interaction, especially in G1 cells, is preferred In one embodiment, especially where the target cells are non-dividing, or where removal and return (*ex vivo* gene targeting) is problematic, for example neuron or muscle cells. KEAP1 siRNA is available from ThermoFischer. In one embodiment, a BRCA1–PALB2 complex may be delivered to the G1 cell. In one embodiment, PALB2 deubiquitylation may be promoted for

example by increased expression of the deubiquitylase USP11, so it is envisaged that a construct may be provided to promote or up-regulate expression or activity of the deubiquitylase USP11.

Treating Diseases of the Eye

[0725] In one embodiment, the disease to be treated is a disease that affects the eyes. Thus, In one embodiment, the composition, system, or component thereof described herein is delivered to one or both eyes.

[0726] The composition, system, can be used to correct ocular defects that arise from several genetic mutations further described in Genetic Diseases of the Eye, Second Edition, edited by Elias I. Traboulsi, Oxford University Press, 2012.

[0727] In one embodiment, the condition to be treated or targeted is an eye disorder. In one embodiment, the eye disorder may include glaucoma. In one embodiment, the eye disorder includes a retinal degenerative disease. In one embodiment, the retinal degenerative disease is selected from Stargardt disease, Bardet-Biedl Syndrome, Best disease, Blue Cone Monochromacy, Choroideremia, Cone-rod dystrophy, Congenital Stationary Night Blindness, Enhanced S-Cone Syndrome, Juvenile X-Linked Retinoschisis, Leber Congenital Amaurosis, Malattia Leventinese, Norrie Disease or X-linked Familial Exudative Vitreoretinopathy, Pattern Dystrophy, Sorsby Dystrophy, Usher Syndrome, Retinitis Pigmentosa, Achromatopsia or Macular dystrophies or degeneration, Retinitis Pigmentosa, Achromatopsia, and age related macular degeneration. In one embodiment, the retinal degenerative disease is Leber Congenital Amaurosis (LCA) or Retinitis Pigmentosa. Other exemplary eye diseases are described in greater detail elsewhere herein.

[0728] In one embodiment, the composition, system, is delivered to the eye, optionally via intravitreal injection or subretinal injection. Intraocular injections may be performed with the aid of an operating microscope. For subretinal and intravitreal injections, eyes may be prolapsed by gentle digital pressure and fundi visualized using a contact lens system consisting of a drop of a coupling medium solution on the cornea covered with a glass microscope slide coverslip. For subretinal injections, the tip of a 10-mm 34-gauge needle, mounted on a 5- μ l Hamilton syringe may be advanced under direct visualization through the superior equatorial sclera tangentially towards the posterior pole until the aperture of the needle was visible in the subretinal space. Then, 2 μ l of vector suspension may be injected to produce a superior bullous retinal detachment, thus confirming subretinal vector administration. This approach creates a

self-sealing sclerotomy allowing the vector suspension to be retained in the subretinal space until it is absorbed by the RPE, usually within 48 h of the procedure. This procedure may be repeated in the inferior hemisphere to produce an inferior retinal detachment. This technique results in the exposure of approximately 70% of neurosensory retina and RPE to the vector suspension. For intravitreal injections, the needle tip may be advanced through the sclera 1 mm posterior to the corneoscleral limbus and 2 μ l of vector suspension injected into the vitreous cavity. For intracameral injections, the needle tip may be advanced through a corneoscleral limbal paracentesis, directed towards the central cornea, and 2 μ l of vector suspension may be injected. For intracameral injections, the needle tip may be advanced through a corneoscleral limbal paracentesis, directed towards the central cornea, and 2 μ l of vector suspension may be injected. These vectors may be injected at titers of either $1.0\text{--}1.4 \times 10^{10}$ or $1.0\text{--}1.4 \times 10^9$ transducing units (TU)/ml.

[0729] In one embodiment, for administration to the eye, lentiviral vectors. In one embodiment, the lentiviral vector is an equine infectious anemia virus (EIAV) vector. Exemplary EIAV vectors for eye delivery are described in Balagaan, J Gene Med 2006; 8: 275 – 285, Published online 21 November 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jgm.845; Binley et al., HUMAN GENE THERAPY 23:980–991 (September 2012), which can be adapted for use with the composition, system, described herein. In one embodiment, the dosage can be 1.1×10^5 transducing units per eye (TU/eye) in a total volume of 100 μ l.

[0730] Other viral vectors can also be used for delivery to the eye, such as AAV vectors, such as those described in Campochiaro et al., Human Gene Therapy 17:167-176 (February 2006), Millington-Ward et al. (Molecular Therapy, vol. 19 no. 4, 642–649 apr. 2011; Dalkara et al. (Sci Transl Med 5, 189ra76 (2013)), which can be adapted for use with the composition, system, described herein. In one embodiment, the dose can range from about 10^6 to $10^{9.5}$ particle units. In the context of the Millington-Ward AAV vectors, a dose of about 2×10^{11} to about 6×10^{13} virus particles can be administered. In the context of Dalkara vectors, a dose of about 1×10^{15} to about 1×10^{16} vg/ml administered to a human.

[0731] In one embodiment, the sd-rxRNA® system of RXi Pharmaceuticals may be used/and or adapted for delivering composition, system, to the eye. In this system, a single intravitreal administration of 3 μ g of sd-rxRNA results in sequence-specific reduction of PPIB mRNA levels for 14 days. The sd-rxRNA® system may be applied to the nucleic acid-targeting

system of the present invention, contemplating a dose of about 3 to 20 mg of composition administered to a human.

[0732] In other embodiments, the methods of US Patent Publication No. 20130183282, which is directed to methods of cleaving a target sequence from the human rhodopsin gene, may also be modified to the nucleic acid-targeting system of the present invention.

[0733] In other embodiments, the methods of US Patent Publication No. 20130202678 for treating retinopathies and sight-threatening ophthalmologic disorders relating to delivering of the Puf-A gene (which is expressed in retinal ganglion and pigmented cells of eye tissues and displays a unique anti-apoptotic activity) to the sub-retinal or intravitreal space in the eye may be used or adapted. In particular, desirable targets are *zgc:193933*, *prdm1a*, *spata2*, *tex10*, *rbb4*, *ddx3*, *zp2.2*, *Blimp-1* and *HtrA2*, all of which may be targeted by the composition, system, of the present invention.

[0734] Wu (Cell Stem Cell,13:659–62, 2013) designed a guide RNA that led Cas9 to a single base pair mutation that causes cataracts in mice, where it induced DNA cleavage. Then using either the other wild-type allele or oligos given to the zygotes repair mechanisms corrected the sequence of the broken allele and corrected the cataract-causing genetic defect in mutant mouse. This approach can be adapted to and/or applied to the compositions, systems, described herein.

[0735] US Patent Publication No. 20120159653, describes use of zinc finger nucleases to genetically modify cells, animals and proteins associated with macular degeneration (MD), the teachings of which can be applied to and/or adapted for the compositions, systems, described herein.

[0736] One aspect of US Patent Publication No. 20120159653 relates to editing of any chromosomal sequences that encode proteins associated with MD which may be applied to the nucleic acid-targeting system of the present invention.

Treating Muscle Diseases and Cardiovascular Diseases

[0737] In one embodiment, the composition, system can be used to treat and/or prevent a muscle disease and associated circulatory or cardiovascular disease or disorder. The present invention also contemplates delivering the composition, system, described herein, e.g. IscB polypeptide or CRISPR-associated IscB polypeptide systems, to the heart. For the heart, a myocardium tropic adeno-associated virus (AAVM) is preferred, in particular AAVM41 which showed preferential gene transfer in the heart (see, e.g., Lin-Yanga et al., PNAS, March 10,

2009, vol. 106, no. 10). Administration may be systemic or local. A dosage of about $1-10 \times 10^{14}$ vector genomes are contemplated for systemic administration. See also, e.g., Eulalio et al. (2012) Nature 492: 376 and Somasuntharam et al. (2013) Biomaterials 34: 7790, the teachings of which can be adapted for and/or applied to the compositions, systems, described herein.

[0738] For example, US Patent Publication No. 20110023139, the teachings of which can be adapted for and/or applied to the compositions, systems, described herein describes use of zinc finger nucleases to genetically modify cells, animals and proteins associated with cardiovascular disease. Cardiovascular diseases generally include high blood pressure, heart attacks, heart failure, and stroke and TIA. Any chromosomal sequence involved in cardiovascular disease or the protein encoded by any chromosomal sequence involved in cardiovascular disease may be utilized in the methods described in this disclosure. The cardiovascular-related proteins are typically selected based on an experimental association of the cardiovascular-related protein to the development of cardiovascular disease. For example, the production rate or circulating concentration of a cardiovascular-related protein may be elevated or depressed in a population having a cardiovascular disorder relative to a population lacking the cardiovascular disorder. Differences in protein levels may be assessed using proteomic techniques including but not limited to Western blot, immunohistochemical staining, enzyme linked immunosorbent assay (ELISA), and mass spectrometry. Alternatively, the cardiovascular-related proteins may be identified by obtaining gene expression profiles of the genes encoding the proteins using genomic techniques including but not limited to DNA microarray analysis, serial analysis of gene expression (SAGE), and quantitative real-time polymerase chain reaction (Q-PCR).

[0739] The compositions, systems, herein can be used for treating diseases of the muscular system. The present invention also contemplates delivering the composition, system, described herein, effector protein systems, to muscle(s).

[0740] In one embodiment, the muscle disease to be treated is a muscle dystrophy such as DMD. In one embodiment, the composition, system, such as a system capable of RNA modification, described herein can be used to achieve exon skipping to achieve correction of the diseased gene. As used herein, the term “exon skipping” refers to the modification of pre-mRNA splicing by the targeting of splice donor and/or acceptor sites within a pre-mRNA with one or more complementary antisense oligonucleotide(s) (AONs). By blocking access of a spliceosome to one or more splice donor or acceptor site, an AON may prevent a splicing

reaction thereby causing the deletion of one or more exons from a fully-processed mRNA. Exon skipping may be achieved in the nucleus during the maturation process of pre-mRNAs. In some examples, exon skipping may include the masking of key sequences involved in the splicing of targeted exons by using a composition, system, described herein capable of RNA modification. In one embodiment, exon skipping can be achieved in dystrophin mRNA. In one embodiment, the composition, system, can induce exon skipping at exon 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 45, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, or any combination thereof of the dystrophin mRNA. In one embodiment, the composition, system, can induce exon skipping at exon 43, 44, 50, 51, 52, 55, or any combination thereof of the dystrophin mRNA. Mutations in these exons, can also be corrected using non-exon skipping polynucleotide modification methods.

[0741] In one embodiment, for treatment of a muscle disease, the method of Bortolanza et al. (*Molecular Therapy* vol. 19 no. 11, 2055–2064 Nov. 2011) may be applied to an AAV expressing IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and injected into humans at a dosage of about 2×10^{15} or 2×10^{16} vg of vector. The teachings of Bortolanza et al., can be adapted for and/or applied to the compositions, systems, described herein.

[0742] In one embodiment, the method of Dumonceaux et al. (*Molecular Therapy* vol. 18 no. 5, 881–887 May 2010) may be applied to an AAV expressing IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and injected into humans, for example, at a dosage of about 10^{14} to about 10^{15} vg of vector. The teachings of Dumonceaux described herein can be adapted for and/or applied to the compositions, systems, described herein.

[0743] In one embodiment, the method of Kinouchi et al. (*Gene Therapy* (2008) 15, 1126–1130) may be applied to compositions described herein and injected into a human, for example, at a dosage of about 500 to 1000 ml of a 40 μ M solution into the muscle.

[0744] In one embodiment, the method of Hagstrom et al. (*Molecular Therapy* Vol. 10, No. 2, August 2004) can be adapted for and/or applied to the compositions, systems, herein and injected at a dose of about 15 to about 50 mg into the great saphenous vein of a human.

[0745] In one embodiment, the method comprise treating a sickle cell related disease, e.g., sickle cell trait, sickle cell disease such as sickle cell anemia, β -thalassaemia. For example, the method and system may be used to modify the genome of the sickle cell, e.g., by correcting

one or more mutations of the β -globin gene. In the case of β -thalassaemia, sickle cell anemia can be corrected by modifying HSCs with the systems. The system allows the specific editing of the cell's genome by cutting its DNA and then letting it repair itself. The IscB polypeptide nuclease is inserted and directed by a RNA guide to the mutated point and then it cuts the DNA at that point. Simultaneously, a healthy version of the sequence is inserted. This sequence is used by the cell's own repair system to fix the induced cut. In this way, the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease allows the correction of the mutation in the previously obtained stem cells. The methods and systems may be used to correct HSCs as to sickle cell anemia using a systems that targets and corrects the mutation (e.g., with a suitable HDR template that delivers a coding sequence for β -globin, advantageously non-sickling β -globin); specifically, the guide RNA can target mutation that give rise to sickle cell anemia, and the HDR can provide coding for proper expression of β -globin. An ω RNA or guide RNA that targets the mutation-and- IscB polypeptide nuclease containing particle is contacted with HSCs carrying the mutation. The particle also can contain a suitable HDR template to correct the mutation for proper expression of β -globin; or the HSC can be contacted with a second particle or a vector that contains or delivers the HDR template. The so contacted cells can be administered; and optionally treated / expanded; cf. Cartier. The HDR template can provide for the HSC to express an engineered β -globin gene (e.g., β A-T87Q), or β -globin.

Treating Diseases of the Liver and Kidney

[0746] In one embodiment, the composition, system, or component thereof described herein can be used to treat a disease of the kidney or liver. Thus, In one embodiment, delivery of the composition or component thereof described herein is to the liver or kidney.

[0747] Delivery strategies to induce cellular uptake of the therapeutic nucleic acid include physical force or vector systems such as viral-, lipid- or complex- based delivery, or nanocarriers. From the initial applications with less possible clinical relevance, when nucleic acids were addressed to renal cells with hydrodynamic high-pressure injection systemically, a wide range of gene therapeutic viral and non-viral carriers have been applied already to target posttranscriptional events in different animal kidney disease models *in vivo* (Csaba Révész and Péter Hamar (2011). *Delivery Methods to Target RNAs in the Kidney*, Gene Therapy Applications, Prof. Chunsheng Kang (Ed.), ISBN: 978-953-307-541-9, InTech, Available from: www.intechopen.com/books/gene-therapy-applications/delivery-methods-to-target-

rnas-in-the-kidney). Delivery methods to the kidney may include those in Yuan et al. (*Am J Physiol Renal Physiol* 295: F605–F617, 2008). The method of Yuang et al. may be applied to the composition of the present invention contemplating a 1-2 g subcutaneous injection of IscB polypeptide nuclease conjugated with cholesterol to a human for delivery to the kidneys. In one embodiment, the method of Molitoris et al. (*J Am Soc Nephrol* 20: 1754–1764, 2009) can be adapted to the composition and a cumulative dose of 12- 20 mg/kg to a human can be used for delivery to the proximal tubule cells of the kidneys. In one embodiment, the methods of Thompson et al. (*Nucleic Acid Therapeutics*, Volume 22, Number 4, 2012) can be adapted to the compositions and a dose of up to 25 mg/kg can be delivered via i.v. administration. In one embodiment, the method of Shimizu et al. (*J Am Soc Nephrol* 21: 622–633, 2010) can be adapted to the compositions and a dose of about of 10-20 μ mol compositions complexed with nanocarriers in about 1-2 liters of a physiologic fluid for i.p. administration can be used.

[0748] Other various delivery vehicles can be used to deliver the composition, system to the kidney such as viral, hydrodynamic, lipid, polymer nanoparticles, aptamers and various combinations thereof (see e.g. Larson et al., *Surgery*, (Aug 2007), Vol. 142, No. 2, pp. (262-269); Hamar et al., *Proc Natl Acad Sci*, (Oct 2004), Vol. 101, No. 41, pp. (14883-14888); Zheng et al., *Am J Pathol*, (Oct 2008), Vol. 173, No. 4, pp. (973–980); Feng et al., *Transplantation*, (May 2009), Vol. 87, No. 9, pp. (1283–1289); Q. Zhang et al., *PloS ONE*, (Jul 2010), Vol. 5, No. 7, e11709, pp. (1-13); Kushibikia et al., *J Controlled Release*, (Jul 2005), Vol. 105, No. 3, pp. (318-331); Wang et al., *Gene Therapy*, (Jul 2006), Vol. 13, No. 14, pp. (1097-1103); Kobayashi et al., *Journal of Pharmacology and Experimental Therapeutics*, (Feb 2004), Vol. 308, No. 2, pp. (688-693); Wolfrum et al., *Nature Biotechnology*, (Sep 2007), Vol. 25, No. 10, pp. (1149-1157); Molitoris et al., *J Am Soc Nephrol*, (Aug 2009), Vol. 20, No. 8 pp. (1754-1764); Mikhaylova et al., *Cancer Gene Therapy*, (Mar 2011), Vol. 16, No. 3, pp. (217-226); Y. Zhang et al., *J Am Soc Nephrol*, (Apr 2006), Vol. 17, No. 4, pp. (1090–1101); Singhal et al., *Cancer Res*, (May 2009), Vol. 69, No. 10, pp. (4244-4251); Malek et al., *Toxicology and Applied Pharmacology*, (Apr 2009), Vol. 236, No. 1, pp. (97-108); Shimizu et al., *J Am Soc Nephrology*, (Apr 2010), Vol. 21, No. 4, pp. (622-633); Jiang et al., *Molecular Pharmaceutics*, (May-Jun 2009), Vol. 6, No. 3, pp. (727-737); Cao et al., *J Controlled Release*, (Jun 2010), Vol. 144, No. 2, pp. (203-212); Ninichuk et al., *Am J Pathol*, (Mar 2008), Vol. 172, No. 3, pp. (628-637); Purschke et al., *Proc Natl Acad Sci*, (Mar 2006), Vol. 103, No. 13, pp. (5173-5178).

[0749] In one embodiment, delivery is to liver cells. In one embodiment, the liver cell is a hepatocyte. Delivery of the composition and system herein may be via viral vectors, especially AAV (and in particular AAV2/6) vectors. These can be administered by intravenous injection. A preferred target for the liver, whether *in vitro* or *in vivo*, is the albumin gene. This is a so-called ‘safe harbor’ as albumin is expressed at very high levels and so some reduction in the production of albumin following successful gene editing is tolerated. It is also preferred as the high levels of expression seen from the albumin promoter/enhancer allows for useful levels of correct or transgene production (from the inserted recombination template) to be achieved even if only a small fraction of hepatocytes are edited. See sites identified by Wechsler et al. (reported at the 57th Annual Meeting and Exposition of the American Society of Hematology - abstract available online at ash.confex.com/ash/2015/webprogram/Paper86495.html and presented on 6th December 2015) which can be adapted for use with the compositions, systems, herein.

[0750] Exemplary liver and kidney diseases that can be treated and/or prevented are described elsewhere herein.

Treating Epithelial and Lung Diseases

[0751] In one embodiment, the disease treated or prevented by the composition and system described herein can be a lung or epithelial disease. The compositions and systems described herein can be used for treating epithelial and/or lung diseases. The present invention also contemplates delivering the composition, system, described herein, to one or both lungs.

[0752] In one embodiment, as viral vector can be used to deliver the composition, system, or component thereof to the lungs. In one embodiment, the AAV is an AAV-1, AAV-2, AAV-5, AAV-6, and/or AAV-9 for delivery to the lungs. (see, e.g., Li et al., *Molecular Therapy*, vol. 17 no. 12, 2067-2077 Dec 2009). In one embodiment, the MOI can vary from 1×10^3 to 4×10^5 vector genomes/cell. In one embodiment, the delivery vector can be an RSV vector as in Zamora et al. (*Am J Respir Crit Care Med* Vol 183, pp 531–538, 2011. The method of Zamora et al. may be applied to the nucleic acid-targeting system of the present invention and an aerosolized composition, for example with a dosage of 0.6 mg/kg, may be contemplated for the present invention.

[0753] Subjects treated for a lung disease may for example receive pharmaceutically effective amount of aerosolized AAV vector system per lung endobronchially delivered while spontaneously breathing. As such, aerosolized delivery is preferred for AAV delivery in

general. An adenovirus or an AAV particle may be used for delivery. Suitable gene constructs, each operably linked to one or more regulatory sequences, may be cloned into the delivery vector. In this instance, the following constructs are provided as examples: Cbh or EF1a promoter for Cas, U6 or H1 promoter for guide RNA); A preferred arrangement is to use a CFTRdelta508 targeting guide, a repair template for deltaF508 mutation and a codon optimized composition, with optionally one or more nuclear localization signal or sequence(s) (NLS(s)), e.g., two (2) NLSs.

Treating Diseases of the Skin

[0754] The compositions and systems described herein can be used for the treatment of skin diseases. The present invention also contemplates delivering the composition and system, described herein, to the skin.

[0755] In one embodiment, delivery to the skin (intra dermal delivery) of the composition, system, or component thereof can be via one or more microneedles or microneedle containing device. For example, In one embodiment the device and methods of Hickerson et al. (Molecular Therapy—Nucleic Acids (2013) 2, e129) can be used and/or adapted to deliver the composition, system, described herein, for example, at a dosage of up to 300 µl of 0.1 mg/ml compositions to the skin.

[0756] In one embodiment, the methods and techniques of Leachman et al. (Molecular Therapy, vol. 18 no. 2, 442–446 Feb. 2010) can be used and/or adapted for delivery of a compositions described herein to the skin.

[0757] In one embodiment, the methods and techniques of Zheng et al. (PNAS, July 24, 2012, vol. 109, no. 30, 11975–11980) can be used and/or adapted for nanoparticle delivery of a compositions described herein to the skin. In one embodiment, as dosage of about 25 nM applied in a single application can achieve gene knockdown in the skin.

Treating Cancer

[0758] The compositions, systems, described herein can be used for the treatment of cancer. The present invention also contemplates delivering the composition, system, described herein, to a cancer cell. Also, as is described elsewhere herein the compositions, systems, can be used to modify an immune cell, such as a CAR or CAR T cell, which can then in turn be used to treat and/or prevent cancer. This is also described in International Patent Publication No. WO 2015/161276, the disclosure of which is hereby incorporated by reference and described herein below.

[0759] Target genes suitable for the treatment or prophylaxis of cancer can include those set forth in Tables 7 and 8. In one embodiment, target genes for cancer treatment and prevention can also include those described in International Patent Publication No. WO 2015/048577 the disclosure of which is hereby incorporated by reference and can be adapted for and/or applied to the composition, system, described herein.

Adoptive Cell Therapy

[0760] The compositions, systems, and components thereof described herein can be used to modify cells for an adoptive cell therapy. In an aspect of the invention, methods and compositions which involve editing a target nucleic acid sequence, or modulating expression of a target nucleic acid sequence, and applications thereof in connection with cancer immunotherapy are comprehended by adapting the composition, system, of the present invention. In some examples, the compositions, systems, and methods may be used to modify a stem cell (e.g., induced pluripotent cell) to derive modified natural killer cells, gamma delta T cells, and alpha beta T cells, which can be used for the adoptive cell therapy. In certain examples, the compositions, systems, and methods may be used to modify modified natural killer cells, gamma delta T cells, and alpha beta T cells.

[0761] As used herein, “ACT”, “adoptive cell therapy” and “adoptive cell transfer” may be used interchangeably. In an embodiment, Adoptive cell therapy (ACT) can refer to the transfer of cells to a patient with the goal of transferring the functionality and characteristics into the new host by engraftment of the cells (see, e.g., Mettananda et al., Editing an α -globin enhancer in primary human hematopoietic stem cells as a treatment for β -thalassemia, Nat Commun. 2017 Sep 4;8(1):424). As used herein, the term "engraft" or "engraftment" refers to the process of cell incorporation into a tissue of interest *in vivo* through contact with existing cells of the tissue. Adoptive cell therapy (ACT) can refer to the transfer of cells, most commonly immune-derived cells, back into the same patient or into a new recipient host with the goal of transferring the immunologic functionality and characteristics into the new host. If possible, use of autologous cells helps the recipient by minimizing GVHD issues. The adoptive transfer of autologous tumor infiltrating lymphocytes (TIL) (Zacharakis et al., (2018) Nat Med. 2018 Jun;24(6):724-730; Besser et al., (2010) Clin. Cancer Res 16 (9) 2646–55; Dudley et al., (2002) Science 298 (5594): 850–4; and Dudley et al., (2005) Journal of Clinical Oncology 23 (10): 2346–57.) or genetically re-directed peripheral blood mononuclear cells (Johnson et al., (2009) Blood 114 (3): 535–46; and Morgan et al., (2006) Science 314(5796) 126-9) has been

used to successfully treat patients with advanced solid tumors, including melanoma, metastatic breast cancer and colorectal carcinoma, as well as patients with CD19-expressing hematologic malignancies (Kalos et al., (2011) *Science Translational Medicine* 3 (95): 95ra73). In an embodiment, allogenic cells immune cells are transferred (see, e.g., Ren et al., (2017) *Clin Cancer Res* 23 (9) 2255-2266). As described further herein, allogenic cells can be edited to reduce alloreactivity and prevent graft-versus-host disease. Thus, use of allogenic cells allows for cells to be obtained from healthy donors and prepared for use in patients as opposed to preparing autologous cells from a patient after diagnosis.

[0762] Aspects of the invention involve the adoptive transfer of immune system cells, such as T cells, specific for selected antigens, such as tumor associated antigens or tumor specific neoantigens (see, e.g., Maus et al., 2014, Adoptive Immunotherapy for Cancer or Viruses, *Annual Review of Immunology*, Vol. 32: 189-225; Rosenberg and Restifo, 2015, Adoptive cell transfer as personalized immunotherapy for human cancer, *Science* Vol. 348 no. 6230 pp. 62-68; Restifo et al., 2015, Adoptive immunotherapy for cancer: harnessing the T cell response. *Nat. Rev. Immunol.* 12(4): 269-281; and Jenson and Riddell, 2014, Design and implementation of adoptive therapy with chimeric antigen receptor-modified T cells. *Immunol Rev.* 257(1): 127-144; and Rajasagi et al., 2014, Systematic identification of personal tumor-specific neoantigens in chronic lymphocytic leukemia. *Blood.* 2014 Jul 17;124(3):453-62).

[0763] In an embodiment, an antigen (such as a tumor antigen) to be targeted in adoptive cell therapy (such as particularly CAR or TCR T-cell therapy) of a disease (such as particularly of tumor or cancer) may be selected from a group consisting of: MR1 (see, e.g., Crowther, et al., 2020, Genome-wide CRISPR-Cas9 screening reveals ubiquitous T cell cancer targeting via the monomorphic MHC class I-related protein MR1, *Nature Immunology* vol. 21, pages178-185), B cell maturation antigen (BCMA) (see, e.g., Friedman et al., Effective Targeting of Multiple BCMA-Expressing Hematological Malignancies by Anti-BCMA CAR T Cells, *Hum Gene Ther.* 2018 Mar 8; Berdeja JG, et al. Durable clinical responses in heavily pretreated patients with relapsed/refractory multiple myeloma: updated results from a multicenter study of bb2121 anti-Bcma CAR T cell therapy. *Blood.* 2017;130:740; and Mouhieddine and Ghobrial, Immunotherapy in Multiple Myeloma: The Era of CAR T Cell Therapy, *Hematologist*, May-June 2018, Volume 15, issue 3); PSA (prostate-specific antigen); prostate-specific membrane antigen (PSMA); PSCA (Prostate stem cell antigen); Tyrosine-protein kinase transmembrane receptor ROR1; fibroblast activation protein (FAP); Tumor-

associated glycoprotein 72 (TAG72); Carcinoembryonic antigen (CEA); Epithelial cell adhesion molecule (EPCAM); Mesothelin; Human Epidermal growth factor Receptor 2 (ERBB2 (Her2/neu)); Prostase; Prostatic acid phosphatase (PAP); elongation factor 2 mutant (ELF2M); Insulin-like growth factor 1 receptor (IGF-1R); gp100; BCR-ABL (breakpoint cluster region-Abelson); tyrosinase; New York esophageal squamous cell carcinoma 1 (NY-ESO-1); κ -light chain, LAGE (L antigen); MAGE (melanoma antigen); Melanoma-associated antigen 1 (MAGE-A1); MAGE A3; MAGE A6; legumain; Human papillomavirus (HPV) E6; HPV E7; prostatein; survivin; PCTA1 (Galectin 8); Melan-A/MART-1; Ras mutant; TRP-1 (tyrosinase related protein 1, or gp75); Tyrosinase-related Protein 2 (TRP2); TRP-2/INT2 (TRP-2/intron 2); RAGE (renal antigen); receptor for advanced glycation end products 1 (RAGE1); Renal ubiquitous 1, 2 (RU1, RU2); intestinal carboxyl esterase (iCE); Heat shock protein 70-2 (HSP70-2) mutant; thyroid stimulating hormone receptor (TSHR); CD123; CD171; CD19; CD20; CD22; CD26; CD30; CD33; CD44v7/8 (cluster of differentiation 44, exons 7/8); CD53; CD92; CD100; CD148; CD150; CD200; CD261; CD262; CD362; CS-1 (CD2 subset 1, CRACC, SLAMF7, CD319, and 19A24); C-type lectin-like molecule-1 (CLL-1); ganglioside GD3 (aNeu5Ac(2-8)aNeu5Ac(2-3)bDGalp(1-4)bDGlc(1-1)Cer); Tn antigen (Tn Ag); Fms-Like Tyrosine Kinase 3 (FLT3); CD38; CD138; CD44v6; B7H3 (CD276); KIT (CD117); Interleukin-13 receptor subunit alpha-2 (IL-13Ra2); Interleukin 11 receptor alpha (IL-11Ra); prostate stem cell antigen (PSCA); Protease Serine 21 (PRSS21); vascular endothelial growth factor receptor 2 (VEGFR2); Lewis(Y) antigen; CD24; Platelet-derived growth factor receptor beta (PDGFR-beta); stage-specific embryonic antigen-4 (SSEA-4); Mucin 1, cell surface associated (MUC1); mucin 16 (MUC16); epidermal growth factor receptor (EGFR); epidermal growth factor receptor variant III (EGFRvIII); neural cell adhesion molecule (NCAM); carbonic anhydrase IX (CAIX); Proteasome (Prosome, Macropain) Subunit, Beta Type, 9 (LMP2); ephrin type-A receptor 2 (EphA2); Ephrin B2; Fucosyl GM1; sialyl Lewis adhesion molecule (sLe); ganglioside GM3 (aNeu5Ac(2-3)bDGalp(1-4)bDGlc(1-1)Cer); TGS5; high molecular weight-melanoma-associated antigen (HMWMAA); o-acetyl-GD2 ganglioside (OAcGD2); Folate receptor alpha; Folate receptor beta; tumor endothelial marker 1 (TEM1/CD248); tumor endothelial marker 7-related (TEM7R); claudin 6 (CLDN6); G protein-coupled receptor class C group 5, member D (GPRC5D); chromosome X open reading frame 61 (CXORF61); CD97; CD179a; anaplastic lymphoma kinase (ALK); Polysialic acid; placenta-specific 1 (PLAC1); hexasaccharide

portion of globoH glycosphingolipid (GloboH); mammary gland differentiation antigen (NY-BR-1); uroplakin 2 (UPK2); Hepatitis A virus cellular receptor 1 (HAVCR1); adrenoceptor beta 3 (ADRB3); pannexin 3 (PANX3); G protein-coupled receptor 20 (GPR20); lymphocyte antigen 6 complex, locus K 9 (LY6K); Olfactory receptor 51E2 (OR51E2); TCR Gamma Alternate Reading Frame Protein (TARP); Wilms tumor protein (WT1); ETS translocation-variant gene 6, located on chromosome 12p (ETV6-AML); sperm protein 17 (SPA17); X Antigen Family, Member 1A (XAGE1); angiopoietin-binding cell surface receptor 2 (Tie 2); CT (cancer/testis (antigen)); melanoma cancer testis antigen-1 (MAD-CT-1); melanoma cancer testis antigen-2 (MAD-CT-2); Fos-related antigen 1; p53; p53 mutant; human Telomerase reverse transcriptase (hTERT); sarcoma translocation breakpoints; melanoma inhibitor of apoptosis (ML-IAP); ERG (transmembrane protease, serine 2 (TMPRSS2) ETS fusion gene); N-Acetyl glucosaminyl-transferase V (NA17); paired box protein Pax-3 (PAX3); Androgen receptor; Cyclin B1; Cyclin D1; v-myc avian myelocytomatosis viral oncogene neuroblastoma derived homolog (MYCN); Ras Homolog Family Member C (RhoC); Cytochrome P450 1B1 (CYP1B1); CCCTC-Binding Factor (Zinc Finger Protein)-Like (BORIS); Squamous Cell Carcinoma Antigen Recognized By T Cells-1 or 3 (SART1, SART3); Paired box protein Pax-5 (PAX5); proacrosin binding protein sp32 (OY-TES1); lymphocyte-specific protein tyrosine kinase (LCK); A kinase anchor protein 4 (AKAP-4); synovial sarcoma, X breakpoint-1, -2, -3 or -4 (SSX1, SSX2, SSX3, SSX4); CD79a; CD79b; CD72; Leukocyte-associated immunoglobulin-like receptor 1 (LAIR1); Fc fragment of IgA receptor (FCAR); Leukocyte immunoglobulin-like receptor subfamily A member 2 (LILRA2); CD300 molecule-like family member f (CD300LF); C-type lectin domain family 12 member A (CLEC12A); bone marrow stromal cell antigen 2 (BST2); EGF-like module-containing mucin-like hormone receptor-like 2 (EMR2); lymphocyte antigen 75 (LY75); Glypican-3 (GPC3); Fc receptor-like 5 (FCRL5); mouse double minute 2 homolog (MDM2); livin; alphafetoprotein (AFP); transmembrane activator and CAML Interactor (TACI); B-cell activating factor receptor (BAFF-R); V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (KRAS); immunoglobulin lambda-like polypeptide 1 (IGLL1); 707-AP (707 alanine proline); ART-4 (adenocarcinoma antigen recognized by T4 cells); BAGE (B antigen; b-catenin/m, b-catenin/mutated); CAMEL (CTL-recognized antigen on melanoma); CAP1 (carcinoembryonic antigen peptide 1); CASP-8 (caspase-8); CDC27m (cell-division cycle 27 mutated); CDK4/m (cycline-dependent kinase 4 mutated); Cyp-B (cyclophilin B); DAM (differentiation antigen melanoma); EGP-2 (epithelial

glycoprotein 2); EGP-40 (epithelial glycoprotein 40); ErbB2, 3, 4 (erythroblastic leukemia viral oncogene homolog-2, -3, 4); FBP (folate binding protein); fAChR (Fetal acetylcholine receptor); G250 (glycoprotein 250); GAGE (G antigen); GnT-V (N-acetylglucosaminyltransferase V); HAGE (helicose antigen); ULA-A (human leukocyte antigen-A); HST2 (human signet ring tumor 2); KIAA0205; KDR (kinase insert domain receptor); LDLR/FUT (low density lipid receptor/GDP L-fucose: b-D-galactosidase 2-a-L fucosyltransferase); L1CAM (L1 cell adhesion molecule); MC1R (melanocortin 1 receptor); Myosin/m (myosin mutated); MUM-1, -2, -3 (melanoma ubiquitous mutated 1, 2, 3); NA88-A (NA cDNA clone of patient M88); KG2D (Natural killer group 2, member D) ligands; oncofetal antigen (h5T4); p190 minor bcr-abl (protein of 190KD bcr-abl); Pml/RARa (promyelocytic leukemia/retinoic acid receptor a); PRAME (preferentially expressed antigen of melanoma); SAGE (sarcoma antigen); TEL/AML1 (translocation Ets-family leukemia/acute myeloid leukemia 1); TPI/m (triosephosphate isomerase mutated); CD70; and any combination thereof.

[0764] In an embodiment, an antigen to be targeted in adoptive cell therapy (such as particularly CAR or TCR T-cell therapy) of a disease (such as particularly of tumor or cancer) is a tumor-specific antigen (TSA).

[0765] In an embodiment, an antigen to be targeted in adoptive cell therapy (such as particularly CAR or TCR T-cell therapy) of a disease (such as particularly of tumor or cancer) is a neoantigen.

[0766] In an embodiment, an antigen to be targeted in adoptive cell therapy (such as particularly CAR or TCR T-cell therapy) of a disease (such as particularly of tumor or cancer) is a tumor-associated antigen (TAA).

[0767] In an embodiment, an antigen to be targeted in adoptive cell therapy (such as particularly CAR or TCR T-cell therapy) of a disease (such as particularly of tumor or cancer) is a universal tumor antigen. In certain preferred embodiments, the universal tumor antigen is selected from the group consisting of: a human telomerase reverse transcriptase (hTERT), survivin, mouse double minute 2 homolog (MDM2), cytochrome P450 1B 1 (CYP1B), HER2/neu, Wilms' tumor gene 1 (WT1), livin, alphafetoprotein (AFP), carcinoembryonic antigen (CEA), mucin 16 (MUC16), MUC1, prostate-specific membrane antigen (PSMA), p53, cyclin (D1), and any combinations thereof.

[0768] In an embodiment, an antigen (such as a tumor antigen) to be targeted in adoptive cell therapy (such as particularly CAR or TCR T-cell therapy) of a disease (such as particularly of tumor or cancer) may be selected from a group consisting of: CD19, BCMA, CD70, CLL-1, MAGE A3, MAGE A6, HPV E6, HPV E7, WT1, CD22, CD171, ROR1, MUC16, and SSX2. In certain preferred embodiments, the antigen may be CD19. For example, CD19 may be targeted in hematologic malignancies, such as in lymphomas, more particularly in B-cell lymphomas, such as without limitation in diffuse large B-cell lymphoma, primary mediastinal b-cell lymphoma, transformed follicular lymphoma, marginal zone lymphoma, mantle cell lymphoma, acute lymphoblastic leukemia including adult and pediatric ALL, non-Hodgkin lymphoma, indolent non-Hodgkin lymphoma, or chronic lymphocytic leukemia. For example, BCMA may be targeted in multiple myeloma or plasma cell leukemia (see, e.g., 2018 American Association for Cancer Research (AACR) Annual meeting Poster: Allogeneic Chimeric Antigen Receptor T Cells Targeting B Cell Maturation Antigen). For example, CLL1 may be targeted in acute myeloid leukemia. For example, MAGE A3, MAGE A6, SSX2, and/or KRAS may be targeted in solid tumors. For example, HPV E6 and/or HPV E7 may be targeted in cervical cancer or head and neck cancer. For example, WT1 may be targeted in acute myeloid leukemia (AML), myelodysplastic syndromes (MDS), chronic myeloid leukemia (CML), non-small cell lung cancer, breast, pancreatic, ovarian or colorectal cancers, or mesothelioma. For example, CD22 may be targeted in B cell malignancies, including non-Hodgkin lymphoma, diffuse large B-cell lymphoma, or acute lymphoblastic leukemia. For example, CD171 may be targeted in neuroblastoma, glioblastoma, or lung, pancreatic, or ovarian cancers. For example, ROR1 may be targeted in ROR1+ malignancies, including non-small cell lung cancer, triple negative breast cancer, pancreatic cancer, prostate cancer, ALL, chronic lymphocytic leukemia, or mantle cell lymphoma. For example, MUC16 may be targeted in MUC16ecto+ epithelial ovarian, fallopian tube or primary peritoneal cancer. For example, CD70 may be targeted in both hematologic malignancies as well as in solid cancers such as renal cell carcinoma (RCC), gliomas (e.g., GBM), and head and neck cancers (HNSCC). CD70 is expressed in both hematologic malignancies as well as in solid cancers, while its expression in normal tissues is restricted to a subset of lymphoid cell types (see, e.g., 2018 American Association for Cancer Research (AACR) Annual meeting Poster: Allogeneic CRISPR Engineered Anti-CD70 CAR-T Cells Demonstrate Potent Preclinical Activity Against Both Solid and Hematological Cancer Cells).

[0769] Various strategies may for example be employed to genetically modify T cells by altering the specificity of the T cell receptor (TCR) for example by introducing new TCR α and β chains with selected peptide specificity (see U.S. Patent No. 8,697,854; PCT Patent Publications: WO2003020763, WO2004033685, WO2004044004, WO2005114215, WO2006000830, WO2008038002, WO2008039818, WO2004074322, WO2005113595, WO2006125962, WO2013166321, WO2013039889, WO2014018863, WO2014083173; U.S. Patent No. 8,088,379).

[0770] As an alternative to, or addition to, TCR modifications, chimeric antigen receptors (CARs) may be used in order to generate immunoresponsive cells, such as T cells, specific for selected targets, such as malignant cells, with a wide variety of receptor chimera constructs having been described (see U.S. Patent Nos. 5,843,728; 5,851,828; 5,912,170; 6,004,811; 6,284,240; 6,392,013; 6,410,014; 6,753,162; 8,211,422; and, PCT Publication WO 9215322).

[0771] In general, CARs are comprised of an extracellular domain, a transmembrane domain, and an intracellular domain, wherein the extracellular domain comprises an antigen-binding domain that is specific for a predetermined target. While the antigen-binding domain of a CAR is often an antibody or antibody fragment (e.g., a single chain variable fragment, scFv), the binding domain is not particularly limited so long as it results in specific recognition of a target. For example, In one embodiment, the antigen-binding domain may comprise a receptor, such that the CAR is capable of binding to the ligand of the receptor. Alternatively, the antigen-binding domain may comprise a ligand, such that the CAR is capable of binding the endogenous receptor of that ligand.

[0772] The antigen-binding domain of a CAR is generally separated from the transmembrane domain by a hinge or spacer. The spacer is also not particularly limited, and it is designed to provide the CAR with flexibility. For example, a spacer domain may comprise a portion of a human Fc domain, including a portion of the CH3 domain, or the hinge region of any immunoglobulin, such as IgA, IgD, IgE, IgG, or IgM, or variants thereof. Furthermore, the hinge region may be modified so as to prevent off-target binding by FcRs or other potential interfering objects. For example, the hinge may comprise an IgG4 Fc domain with or without a S228P, L235E, and/or N297Q mutation (according to Kabat numbering) in order to decrease binding to FcRs. Additional spacers/hinges include, but are not limited to, CD4, CD8, and CD28 hinge regions.

[0773] The transmembrane domain of a CAR may be derived either from a natural or from a synthetic source. Where the source is natural, the domain may be derived from any membrane bound or transmembrane protein. Transmembrane regions of particular use in this disclosure may be derived from CD8, CD28, CD3, CD45, CD4, CD5, CDS, CD9, CD 16, CD22, CD33, CD37, CD64, CD80, CD86, CD 134, CD137, CD 154, TCR. Alternatively, the transmembrane domain may be synthetic, in which case it will comprise predominantly hydrophobic residues such as leucine and valine. Preferably a triplet of phenylalanine, tryptophan and valine will be found at each end of a synthetic transmembrane domain. Optionally, a short oligo- or polypeptide linker, preferably between 2 and 10 amino acids in length may form the linkage between the transmembrane domain and the cytoplasmic signaling domain of the CAR. A glycine-serine doublet provides a particularly suitable linker.

[0774] Alternative CAR constructs may be characterized as belonging to successive generations. First-generation CARs typically consist of a single-chain variable fragment of an antibody specific for an antigen, for example comprising a VL linked to a VH of a specific antibody, linked by a flexible linker, for example by a CD8 α hinge domain and a CD8 α transmembrane domain, to the transmembrane and intracellular signaling domains of either CD3 ζ or FcR γ (scFv-CD3 ζ or scFv-FcR γ ; see U.S. Patent No. 7,741,465; U.S. Patent No. 5,912,172; U.S. Patent No. 5,906,936). Second-generation CARs incorporate the intracellular domains of one or more costimulatory molecules, such as CD28, OX40 (CD134), or 4-1BB (CD137) within the endodomain (for example scFv-CD28/OX40/4-1BB-CD3 ζ ; see U.S. Patent Nos. 8,911,993; 8,916,381; 8,975,071; 9,101,584; 9,102,760; 9,102,761). Third-generation CARs include a combination of costimulatory endodomains, such a CD3 ζ -chain, CD97, GDI 1a-CD18, CD2, ICOS, CD27, CD154, CDS, OX40, 4-1BB, CD2, CD7, LIGHT, LFA-1, NKG2C, B7-H3, CD30, CD40, PD-1, or CD28 signaling domains (for example scFv-CD28-4-1BB-CD3 ζ or scFv-CD28-OX40-CD3 ζ ; see U.S. Patent No. 8,906,682; U.S. Patent No. 8,399,645; U.S. Pat. No. 5,686,281; PCT Publication No. WO 2014/134165; PCT Publication No. WO 2012/079000). In an embodiment, the primary signaling domain comprises a functional signaling domain of a protein selected from the group consisting of CD3 zeta, CD3 gamma, CD3 delta, CD3 epsilon, common FcR gamma (FCERIG), FcR beta (Fc Epsilon R1b), CD79a, CD79b, Fc gamma RIIa, DAP10, and DAP12. In certain preferred embodiments, the primary signaling domain comprises a functional signaling domain of CD3 ζ or FcR γ . In an embodiment, the one or more costimulatory signaling domains comprise a

functional signaling domain of a protein selected, each independently, from the group consisting of: CD27, CD28, 4-1BB (CD137), OX40, CD30, CD40, PD-1, ICOS, lymphocyte function-associated antigen-1 (LFA-1), CD2, CD7, LIGHT, NKG2C, B7-H3, a ligand that specifically binds with CD83, CDS, ICAM-1, GITR, BAFFR, HVEM (LIGHTR), SLAMF7, NKp80 (KLRF1), CD160, CD19, CD4, CD8 alpha, CD8 beta, IL2R beta, IL2R gamma, IL7R alpha, ITGA4, VLA1, CD49a, ITGA4, IA4, CD49D, ITGA6, VLA-6, CD49f, ITGAD, CD11d, ITGAE, CD103, ITGAL, CD11a, LFA-1, ITGAM, CD11b, ITGAX, CD11c, ITGB1, CD29, ITGB2, CD18, ITGB7, TNFR2, TRANCE/RANKL, DNAM1 (CD226), SLAMF4 (CD244, 2B4), CD84, CD96 (Tactile), CEACAM1, CRTAM, Ly9 (CD229), CD160 (BY55), PSGL1, CD100 (SEMA4D), CD69, SLAMF6 (NTB-A, Lyl08), SLAM (SLAMF1, CD150, IPO-3), BLAME (SLAMF8), SELPLG (CD162), LTBR, LAT, GADS, SLP-76, PAG/Cbp, NKp44, NKp30, NKp46, and NKG2D. In an embodiment, the one or more costimulatory signaling domains comprise a functional signaling domain of a protein selected, each independently, from the group consisting of: 4-1BB, CD27, and CD28. In an embodiment, a chimeric antigen receptor may have the design as described in U.S. Patent No. 7,446,190, comprising an intracellular domain of CD3 ζ chain (such as amino acid residues 52-163 of the human CD3 zeta chain, as shown in SEQ ID NO: 14 of US 7,446,190), a signaling region from CD28 and an antigen-binding element (or portion or domain; such as scFv). The CD28 portion, when between the zeta chain portion and the antigen-binding element, may suitably include the transmembrane and signaling domains of CD28 (such as amino acid residues 114-220 of SEQ ID NO: 10, full sequence shown in SEQ ID NO: 6 of US 7,446,190; these can include the following portion of CD28 as set forth in Genbank identifier NM_006139. Alternatively, when the zeta sequence lies between the CD28 sequence and the antigen-binding element, intracellular domain of CD28 can be used alone (such as amino sequence set forth in SEQ ID NO: 9 of US 7,446,190). Hence, an embodiment employ a CAR comprising (a) a zeta chain portion comprising the intracellular domain of human CD3 ζ chain, (b) a costimulatory signaling region, and (c) an antigen-binding element (or portion or domain), wherein the costimulatory signaling region comprises the amino acid sequence encoded by SEQ ID NO: 6 of US 7,446,190.

[0775] Alternatively, costimulation may be orchestrated by expressing CARs in antigen-specific T cells, chosen so as to be activated and expanded following engagement of their native $\alpha\beta$ TCR, for example by antigen on professional antigen-presenting cells, with attendant

costimulation. In addition, additional engineered receptors may be provided on the immunoresponsive cells, for example to improve targeting of a T-cell attack and/or minimize side effects.

[0776] By means of an example and without limitation, Kochenderfer et al., (2009) *J Immunother.* 32 (7): 689-702 described anti-CD19 chimeric antigen receptors (CAR). FMC63-28Z CAR contained a single chain variable region moiety (scFv) recognizing CD19 derived from the FMC63 mouse hybridoma (described in Nicholson et al., (1997) *Molecular Immunology* 34: 1157-1165), a portion of the human CD28 molecule, and the intracellular component of the human TCR- ζ molecule. FMC63-CD828BBZ CAR contained the FMC63 scFv, the hinge and transmembrane regions of the CD8 molecule, the cytoplasmic portions of CD28 and 4-1BB, and the cytoplasmic component of the TCR- ζ molecule. The exact sequence of the CD28 molecule included in the FMC63-28Z CAR corresponded to Genbank identifier NM_006139; the sequence included all amino acids starting with the amino acid sequence IEVMYPPPY (SEQ ID NO: 2058) and continuing all the way to the carboxy-terminus of the protein. To encode the anti-CD19 scFv component of the vector, the authors designed a DNA sequence which was based on a portion of a previously published CAR (Cooper et al., (2003) *Blood* 101: 1637-1644). This sequence encoded the following components in frame from the 5' end to the 3' end: an XhoI site, the human granulocyte-macrophage colony-stimulating factor (GM-CSF) receptor α -chain signal sequence, the FMC63 light chain variable region (as in Nicholson et al., supra), a linker peptide (as in Cooper et al., supra), the FMC63 heavy chain variable region (as in Nicholson et al., supra), and a NotI site. A plasmid encoding this sequence was digested with XhoI and NotI. To form the MSGV-FMC63-28Z retroviral vector, the XhoI and NotI-digested fragment encoding the FMC63 scFv was ligated into a second XhoI and NotI-digested fragment that encoded the MSGV retroviral backbone (as in Hughes et al., (2005) *Human Gene Therapy* 16: 457-472) as well as part of the extracellular portion of human CD28, the entire transmembrane and cytoplasmic portion of human CD28, and the cytoplasmic portion of the human TCR- ζ molecule (as in Maher et al., 2002) *Nature Biotechnology* 20: 70-75). The FMC63-28Z CAR is included in the KTE-C19 (axicabtagene ciloleucel) anti-CD19 CAR-T therapy product in development by Kite Pharma, Inc. for the treatment of inter alia patients with relapsed/refractory aggressive B-cell non-Hodgkin lymphoma (NHL). Accordingly, in an embodiment, cells intended for adoptive cell therapies, more particularly immunoresponsive cells such as T cells, may express the FMC63-28Z CAR as described by

Kochenderfer et al. (supra). Hence, in an embodiment, cells intended for adoptive cell therapies, more particularly immunoresponsive cells such as T cells, may comprise a CAR comprising an extracellular antigen-binding element (or portion or domain; such as scFv) that specifically binds to an antigen, an intracellular signaling domain comprising an intracellular domain of a CD3 ζ chain, and a costimulatory signaling region comprising a signaling domain of CD28. Preferably, the CD28 amino acid sequence is as set forth in Genbank identifier NM_006139 (sequence version 1, 2 or 3) starting with the amino acid sequence IEVMYPPPY and continuing all the way to the carboxy-terminus of the protein. Preferably, the antigen is CD19, more preferably the antigen-binding element is an anti-CD19 scFv, even more preferably the anti-CD19 scFv as described by Kochenderfer et al. (supra).

[0777] Additional anti-CD19 CARs are further described in International Patent Publication No. WO 2015/187528. More particularly Example 1 and Table 1 of WO2015187528, incorporated by reference herein, demonstrate the generation of anti-CD19 CARs based on a fully human anti-CD19 monoclonal antibody (47G4, as described in US20100104509) and murine anti-CD19 monoclonal antibody (as described in Nicholson et al. and explained above). Various combinations of a signal sequence (human CD8-alpha or GM-CSF receptor), extracellular and transmembrane regions (human CD8-alpha) and intracellular T-cell signaling domains (CD28-CD3 ζ ; 4-1BB-CD3 ζ ; CD27-CD3 ζ ; CD28-CD27-CD3 ζ ; 4-1BB-CD27-CD3 ζ ; CD27-4-1BB-CD3 ζ ; CD28-CD27-Fc RI gamma chain; or CD28-Fc RI gamma chain) were disclosed. Hence, in an embodiment, cells intended for adoptive cell therapies, more particularly immunoresponsive cells such as T cells, may comprise a CAR comprising an extracellular antigen-binding element that specifically binds to an antigen, an extracellular and transmembrane region as set forth in Table 1 of WO2015187528 and an intracellular T-cell signaling domain as set forth in Table 1 of No. WO 2015/187528. Preferably, the antigen is CD19, more preferably the antigen-binding element is an anti-CD19 scFv, even more preferably the mouse or human anti-CD19 scFv as described in Example 1 of WO 2015/187528. In an embodiment, the CAR comprises, consists essentially of or consists of an amino acid sequence of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, or SEQ ID NO: 13 as set forth in Table 1 of WO2015187528.

[0778] By means of an example and without limitation, chimeric antigen receptor that recognizes the CD70 antigen is described in WO2012058460A2 (see also, Park et al., CD70 as a target for chimeric antigen receptor T cells in head and neck squamous cell carcinoma, *Oral Oncol.* 2018 Mar;78:145-150; and Jin et al., CD70, a novel target of CAR T-cell therapy for gliomas, *Neuro Oncol.* 2018 Jan 10;20(1):55-65). CD70 is expressed by diffuse large B-cell and follicular lymphoma and also by the malignant cells of Hodgkins lymphoma, Waldenstrom's macroglobulinemia and multiple myeloma, and by HTLV-1- and EBV-associated malignancies. (Agathangelou et al. *Am.J.Pathol.* 1995;147: 1152-1160; Hunter et al., *Blood* 2004; 104:4881. 26; Lens et al., *J Immunol.* 2005;174:6212-6219; Baba et al., *J Virol.* 2008;82:3843-3852.) In addition, CD70 is expressed by non-hematological malignancies such as renal cell carcinoma and glioblastoma. (Junker et al., *J Urol.* 2005;173:2150-2153; Chahlavi et al., *Cancer Res* 2005;65:5428-5438) Physiologically, CD70 expression is transient and restricted to a subset of highly activated T, B, and dendritic cells.

[0779] By means of an example and without limitation, chimeric antigen receptor that recognizes BCMA has been described (see, e.g., US20160046724A1; WO2016014789A2; WO2017211900A1; WO2015158671A1; US20180085444A1; WO2018028647A1; US20170283504A1; and WO2013154760A1).

[0780] In an embodiment, the immune cell may, in addition to a CAR or exogenous TCR as described herein, further comprise a chimeric inhibitory receptor (inhibitory CAR) that specifically binds to a second target antigen and is capable of inducing an inhibitory or immunosuppressive or repressive signal to the cell upon recognition of the second target antigen. In an embodiment, the chimeric inhibitory receptor comprises an extracellular antigen-binding element (or portion or domain) configured to specifically bind to a target antigen, a transmembrane domain, and an intracellular immunosuppressive or repressive signaling domain. In an embodiment, the second target antigen is an antigen that is not expressed on the surface of a cancer cell or infected cell or the expression of which is downregulated on a cancer cell or an infected cell. In an embodiment, the second target antigen is an MHC-class I molecule. In an embodiment, the intracellular signaling domain comprises a functional signaling portion of an immune checkpoint molecule, such as for example PD-1 or CTLA4. Advantageously, the inclusion of such inhibitory CAR reduces the chance of the engineered immune cells attacking non-target (e.g., non-cancer) tissues.

[0781] Alternatively, T-cells expressing CARs may be further modified to reduce or eliminate expression of endogenous TCRs in order to reduce off-target effects. Reduction or elimination of endogenous TCRs can reduce off-target effects and increase the effectiveness of the T cells (U.S. 9,181,527). T cells stably lacking expression of a functional TCR may be produced using a variety of approaches. T cells internalize, sort, and degrade the entire T cell receptor as a complex, with a half-life of about 10 hours in resting T cells and 3 hours in stimulated T cells (von Essen, M. et al. 2004. J. Immunol. 173:384-393). Proper functioning of the TCR complex requires the proper stoichiometric ratio of the proteins that compose the TCR complex. TCR function also requires two functioning TCR zeta proteins with ITAM motifs. The activation of the TCR upon engagement of its MHC-peptide ligand requires the engagement of several TCRs on the same T cell, which all must signal properly. Thus, if a TCR complex is destabilized with proteins that do not associate properly or cannot signal optimally, the T cell will not become activated sufficiently to begin a cellular response.

[0782] Accordingly, In one embodiment, TCR expression may be eliminated using RNA interference (e.g., shRNA, siRNA, miRNA, etc.), IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, or other methods that target the nucleic acids encoding specific TCRs (e.g., TCR- α and TCR- β) and/or CD3 chains in primary T cells. By blocking expression of one or more of these proteins, the T cell will no longer produce one or more of the key components of the TCR complex, thereby destabilizing the TCR complex and preventing cell surface expression of a functional TCR.

[0783] In some instances, CAR may also comprise a switch mechanism for controlling expression and/or activation of the CAR. For example, a CAR may comprise an extracellular, transmembrane, and intracellular domain, in which the extracellular domain comprises a target-specific binding element that comprises a label, binding domain, or tag that is specific for a molecule other than the target antigen that is expressed on or by a target cell. In such embodiments, the specificity of the CAR is provided by a second construct that comprises a target antigen binding domain (e.g., an scFv or a bispecific antibody that is specific for both the target antigen and the label or tag on the CAR) and a domain that is recognized by or binds to the label, binding domain, or tag on the CAR. See, e.g., WO 2013/044225, WO 2016/000304, WO 2015/057834, WO 2015/057852, WO 2016/070061, US 9,233,125, US 2016/0129109. In this way, a T-cell that expresses the CAR can be administered to a subject,

but the CAR cannot bind its target antigen until the second composition comprising an antigen-specific binding domain is administered.

[0784] Alternative switch mechanisms include CARs that require multimerization in order to activate their signaling function (see, e.g., US Patent Publication Nos. US 2015/0368342, US 2016/0175359, US 2015/0368360) and/or an exogenous signal, such as a small molecule drug (US 2016/0166613, Yung et al., Science, 2015), in order to elicit a T-cell response. Some CARs may also comprise a “suicide switch” to induce cell death of the CAR T-cells following treatment (Buddee et al., PLoS One, 2013) or to downregulate expression of the CAR following binding to the target antigen (International Patent Publication No. WO 2016/011210).

[0785] Alternative techniques may be used to transform target immunoresponsive cells, such as protoplast fusion, lipofection, transfection or electroporation. A wide variety of vectors may be used, such as retroviral vectors, lentiviral vectors, adenoviral vectors, adeno-associated viral vectors, plasmids or transposons, such as a Sleeping Beauty transposon (see U.S. Patent Nos. 6,489,458; 7,148,203; 7,160,682; 7,985,739; 8,227,432), may be used to introduce CARs, for example using 2nd generation antigen-specific CARs signaling through CD3 ζ and either CD28 or CD137. Viral vectors may for example include vectors based on HIV, SV40, EBV, HSV or BPV.

[0786] Cells that are targeted for transformation may for example include T cells, Natural Killer (NK) cells, cytotoxic T lymphocytes (CTL), regulatory T cells, human embryonic stem cells, tumor-infiltrating lymphocytes (TIL) or a pluripotent stem cell from which lymphoid cells may be differentiated. T cells expressing a desired CAR may for example be selected through co-culture with γ -irradiated activating and propagating cells (AaPC), which co-express the cancer antigen and co-stimulatory molecules. The engineered CAR T-cells may be expanded, for example by co-culture on AaPC in presence of soluble factors, such as IL-2 and IL-21. This expansion may for example be carried out so as to provide memory CAR⁺ T cells (which may for example be assayed by non-enzymatic digital array and/or multi-panel flow cytometry). In this way, CAR T cells may be provided that have specific cytotoxic activity against antigen-bearing tumors (optionally in conjunction with production of desired chemokines such as interferon- γ). CAR T cells of this kind may for example be used in animal models, for example to treat tumor xenografts.

[0787] In an embodiment, ACT includes co-transferring CD4⁺ Th1 cells and CD8⁺ CTLs to induce a synergistic antitumor response (see, e.g., Li et al., Adoptive cell therapy with CD4⁺

T helper 1 cells and CD8⁺ cytotoxic T cells enhances complete rejection of an established tumor, leading to generation of endogenous memory responses to non-targeted tumor epitopes. Clin Transl Immunology. 2017 Oct; 6(10): e160).

[0788] In an embodiment, Th17 cells are transferred to a subject in need thereof. Th17 cells have been reported to directly eradicate melanoma tumors in mice to a greater extent than Th1 cells (Muranski P, et al., Tumor-specific Th17-polarized cells eradicate large established melanoma. Blood. 2008 Jul 15; 112(2):362-73; and Martin-Orozco N, et al., T helper 17 cells promote cytotoxic T cell activation in tumor immunity. Immunity. 2009 Nov 20; 31(5):787-98). Those studies involved an adoptive T cell transfer (ACT) therapy approach, which takes advantage of CD4⁺ T cells that express a TCR recognizing tyrosinase tumor antigen. Exploitation of the TCR leads to rapid expansion of Th17 populations to large numbers *ex vivo* for reinfusion into the autologous tumor-bearing hosts.

[0789] In an embodiment, ACT may include autologous iPSC-based vaccines, such as irradiated iPSCs in autologous anti-tumor vaccines (see e.g., Kooreman, Nigel G. et al., Autologous iPSC-Based Vaccines Elicit Anti-tumor Responses *In Vivo*, Cell Stem Cell 22, 1–13, 2018, doi.org/10.1016/j.stem.2018.01.016).

[0790] Unlike T-cell receptors (TCRs) that are MHC restricted, CARs can potentially bind any cell surface-expressed antigen and can thus be more universally used to treat patients (see Irving et al., Engineering Chimeric Antigen Receptor T-Cells for Racing in Solid Tumors: Don't Forget the Fuel, Front. Immunol., 03 April 2017, doi.org/10.3389/fimmu.2017.00267). In an embodiment, in the absence of endogenous T-cell infiltrate (e.g., due to aberrant antigen processing and presentation), which precludes the use of TIL therapy and immune checkpoint blockade, the transfer of CAR T-cells may be used to treat patients (see, e.g., Hinrichs CS, Rosenberg SA. Exploiting the curative potential of adoptive T-cell therapy for cancer. Immunol Rev (2014) 257(1):56–71. doi:10.1111/imr.12132).

[0791] Approaches such as the foregoing may be adapted to provide methods of treating and/or increasing survival of a subject having a disease, such as a neoplasia, for example by administering an effective amount of an immunoresponsive cell comprising an antigen recognizing receptor that binds a selected antigen, wherein the binding activates the immunoresponsive cell, thereby treating or preventing the disease (such as a neoplasia, a pathogen infection, an autoimmune disorder, or an allogeneic transplant reaction).

[0792] In an embodiment, the treatment can be administered after lymphodepleting pretreatment in the form of chemotherapy (typically a combination of cyclophosphamide and fludarabine) or radiation therapy. Initial studies in ACT had short lived responses and the transferred cells did not persist *in vivo* for very long (Houot et al., T-cell-based immunotherapy: adoptive cell transfer and checkpoint inhibition. *Cancer Immunol Res* (2015) 3(10):1115–22; and Kamta et al., Advancing Cancer Therapy with Present and Emerging Immuno-Oncology Approaches. *Front. Oncol.* (2017) 7:64). Immune suppressor cells like Tregs and MDSCs may attenuate the activity of transferred cells by outcompeting them for the necessary cytokines. Not being bound by a theory lymphodepleting pretreatment may eliminate the suppressor cells allowing the TILs to persist.

[0793] In one embodiment, the treatment can be administered into patients undergoing an immunosuppressive treatment (e.g., glucocorticoid treatment). The cells or population of cells, may be made resistant to at least one immunosuppressive agent due to the inactivation of a gene encoding a receptor for such immunosuppressive agent. In an embodiment, the immunosuppressive treatment provides for the selection and expansion of the immunoresponsive T cells within the patient.

[0794] In an embodiment, the treatment can be administered before primary treatment (e.g., surgery or radiation therapy) to shrink a tumor before the primary treatment. In another embodiment, the treatment can be administered after primary treatment to remove any remaining cancer cells.

[0795] In an embodiment, immunometabolic barriers can be targeted therapeutically prior to and/or during ACT to enhance responses to ACT or CAR T-cell therapy and to support endogenous immunity (see, e.g., Irving et al., Engineering Chimeric Antigen Receptor T-Cells for Racing in Solid Tumors: Don't Forget the Fuel, *Front. Immunol.*, 03 April 2017, doi.org/10.3389/fimmu.2017.00267).

[0796] The administration of cells or population of cells, such as immune system cells or cell populations, such as more particularly immunoresponsive cells or cell populations, as disclosed herein may be carried out in any convenient manner, including by aerosol inhalation, injection, ingestion, transfusion, implantation or transplantation. The cells or population of cells may be administered to a patient subcutaneously, intradermally, intratumorally, intranodally, intramedullary, intramuscularly, intrathecally, by intravenous or intralymphatic injection, or intraperitoneally. In one embodiment, the disclosed CARs may be delivered or

administered into a cavity formed by the resection of tumor tissue (i.e. intracavity delivery) or directly into a tumor prior to resection (i.e. intratumoral delivery). In one embodiment, the cell compositions of the present invention are preferably administered by intravenous injection.

[0797] The administration of the cells or population of cells can consist of the administration of 104- 109 cells per kg body weight, preferably 105 to 106 cells/kg body weight including all integer values of cell numbers within those ranges. Dosing in CAR T cell therapies may for example involve administration of from 106 to 109 cells/kg, with or without a course of lymphodepletion, for example with cyclophosphamide. The cells or population of cells can be administered in one or more doses. In another embodiment, the effective amount of cells are administered as a single dose. In another embodiment, the effective amount of cells are administered as more than one dose over a period time. Timing of administration is within the judgment of managing physician and depends on the clinical condition of the patient. The cells or population of cells may be obtained from any source, such as a blood bank or a donor. While individual needs vary, determination of optimal ranges of effective amounts of a given cell type for a particular disease or conditions are within the skill of one in the art. An effective amount means an amount which provides a therapeutic or prophylactic benefit. The dosage administered will be dependent upon the age, health and weight of the recipient, kind of concurrent treatment, if any, frequency of treatment and the nature of the effect desired.

[0798] In another embodiment, the effective amount of cells or composition comprising those cells are administered parenterally. The administration can be an intravenous administration. The administration can be directly done by injection within a tumor.

[0799] To guard against possible adverse reactions, engineered immunoresponsive cells may be equipped with a transgenic safety switch, in the form of a transgene that renders the cells vulnerable to exposure to a specific signal. For example, the herpes simplex viral thymidine kinase (TK) gene may be used in this way, for example by introduction into allogeneic T lymphocytes used as donor lymphocyte infusions following stem cell transplantation (Greco, et al., Improving the safety of cell therapy with the TK-suicide gene. Front. Pharmacol. 2015; 6: 95). In such cells, administration of a nucleoside prodrug such as ganciclovir or acyclovir causes cell death. Alternative safety switch constructs include inducible caspase 9, for example triggered by administration of a small-molecule dimerizer that brings together two nonfunctional icasp9 molecules to form the active enzyme. A wide variety of alternative approaches to implementing cellular proliferation controls have been

described (see U.S. Patent Publication No. 20130071414; International Patent Publication WO 2011/146862; International Patent Publication WO 2014/011987; International Patent Publication WO 2013/040371; Zhou et al. BLOOD, 2014, 123/25:3895 – 3905; Di Stasi et al., The New England Journal of Medicine 2011; 365:1673-1683; Sadelain M, The New England Journal of Medicine 2011; 365:1735-173; Ramos et al., Stem Cells 28(6):1107-15 (2010)).

[0800] In a further refinement of adoptive therapies, genome editing may be used to tailor immunoresponsive cells to alternative implementations, for example providing edited CAR T cells (see Poirot et al., 2015, Multiplex genome edited T-cell manufacturing platform for "off-the-shelf" adoptive T-cell immunotherapies, Cancer Res 75 (18): 3853; Ren et al., 2017, Multiplex genome editing to generate universal CAR T cells resistant to PD1 inhibition, Clin Cancer Res. 2017 May 1;23(9):2255-2266. doi: 10.1158/1078-0432.CCR-16-1300. Epub 2016 Nov 4; Qasim et al., 2017, Molecular remission of infant B-ALL after infusion of universal TALEN gene-edited CAR T cells, Sci Transl Med. 2017 Jan 25;9(374); Legut, et al., 2018, CRISPR-mediated TCR replacement generates superior anticancer transgenic T cells. Blood, 131(3), 311-322; and Georgiadis et al., Long Terminal Repeat CRISPR-CAR-Coupled "Universal" T Cells Mediate Potent Anti-leukemic Effects, Molecular Therapy, In Press, Corrected Proof, Available online 6 March 2018). Cells may be edited using any CRISPR system and method of use thereof as described herein. The composition and systems may be delivered to an immune cell by any method described herein. In preferred embodiments, cells are edited *ex vivo* and transferred to a subject in need thereof. Immunoresponsive cells, CAR T cells or any cells used for adoptive cell transfer may be edited. Editing may be performed for example to insert or knock-in an exogenous gene, such as an exogenous gene encoding a CAR or a TCR, at a preselected locus in a cell (e.g. TRAC locus); to eliminate potential alloreactive T-cell receptors (TCR) or to prevent inappropriate pairing between endogenous and exogenous TCR chains, such as to knock-out or knock-down expression of an endogenous TCR in a cell; to disrupt the target of a chemotherapeutic agent in a cell; to block an immune checkpoint, such as to knock-out or knock-down expression of an immune checkpoint protein or receptor in a cell; to knock-out or knock-down expression of other gene or genes in a cell, the reduced expression or lack of expression of which can enhance the efficacy of adoptive therapies using the cell; to knock-out or knock-down expression of an endogenous gene in a cell, said endogenous gene encoding an antigen targeted by an exogenous CAR or TCR; to knock-out or knock-down expression of one or more MHC constituent proteins in a cell; to activate a T cell;

to modulate cells such that the cells are resistant to exhaustion or dysfunction; and/or increase the differentiation and/or proliferation of functionally exhausted or dysfunctional CD8⁺ T-cells (see International Patent Publication Nos. WO 2013/176915, WO 2014/059173, WO 2014/172606, WO 2014/184744, and WO 2014/191128).

[0801] In an embodiment, editing may result in inactivation of a gene. By inactivating a gene, it is intended that the gene of interest is not expressed in a functional protein form. In a particular embodiment, the system specifically catalyzes cleavage in one targeted gene thereby inactivating said targeted gene. The nucleic acid strand breaks caused are commonly repaired through the distinct mechanisms of homologous recombination or non-homologous end joining (NHEJ). However, NHEJ is an imperfect repair process that often results in changes to the DNA sequence at the site of the cleavage. Repair via non-homologous end joining (NHEJ) often results in small insertions or deletions (Indel) and can be used for the creation of specific gene knockouts. Cells in which a cleavage induced mutagenesis event has occurred can be identified and/or selected by well-known methods in the art. In an embodiment, homology directed repair (HDR) is used to concurrently inactivate a gene (e.g., TRAC) and insert an endogenous TCR or CAR into the inactivated locus.

[0802] Hence, in an embodiment, editing of cells, particularly cells intended for adoptive cell therapies, more particularly immunoresponsive cells such as T cells, may be performed to insert or knock-in an exogenous gene, such as an exogenous gene encoding a CAR or a TCR, at a preselected locus in a cell. Conventionally, nucleic acid molecules encoding CARs or TCRs are transfected or transduced to cells using randomly integrating vectors, which, depending on the site of integration, may lead to clonal expansion, oncogenic transformation, variegated transgene expression and/or transcriptional silencing of the transgene. Directing of transgene(s) to a specific locus in a cell can minimize or avoid such risks and advantageously provide for uniform expression of the transgene(s) by the cells. Without limitation, suitable ‘safe harbor’ loci for directed transgene integration include CCR5 or AAVS1. Homology-directed repair (HDR) strategies are known and described elsewhere in this specification allowing to insert transgenes into desired loci (e.g., TRAC locus).

[0803] Further suitable loci for insertion of transgenes, in particular CAR or exogenous TCR transgenes, include without limitation loci comprising genes coding for constituents of endogenous T-cell receptor, such as T-cell receptor alpha locus (TRA) or T-cell receptor beta locus (TRB), for example T-cell receptor alpha constant (TRAC) locus, T-cell receptor beta

constant 1 (TRBC1) locus or T-cell receptor beta constant 2 (TRBC1) locus. Advantageously, insertion of a transgene into such locus can simultaneously achieve expression of the transgene, potentially controlled by the endogenous promoter, and knock-out expression of the endogenous TCR. This approach has been exemplified in Eyquem et al., (2017) Nature 543: 113-117, wherein the authors used CRISPR/Cas9 gene editing to knock-in a DNA molecule encoding a CD19-specific CAR into the TRAC locus downstream of the endogenous promoter; the CAR-T cells obtained by CRISPR were significantly superior in terms of reduced tonic CAR signaling and exhaustion.

[0804] T cell receptors (TCR) are cell surface receptors that participate in the activation of T cells in response to the presentation of antigen. The TCR is generally made from two chains, α and β , which assemble to form a heterodimer and associates with the CD3-transducing subunits to form the T cell receptor complex present on the cell surface. Each α and β chain of the TCR consists of an immunoglobulin-like N-terminal variable (V) and constant (C) region, a hydrophobic transmembrane domain, and a short cytoplasmic region. As for immunoglobulin molecules, the variable region of the α and β chains are generated by V(D)J recombination, creating a large diversity of antigen specificities within the population of T cells. However, in contrast to immunoglobulins that recognize intact antigen, T cells are activated by processed peptide fragments in association with an MHC molecule, introducing an extra dimension to antigen recognition by T cells, known as MHC restriction. Recognition of MHC disparities between the donor and recipient through the T cell receptor leads to T cell proliferation and the potential development of graft versus host disease (GVHD). The inactivation of TCR α or TCR β can result in the elimination of the TCR from the surface of T cells preventing recognition of alloantigen and thus GVHD. However, TCR disruption generally results in the elimination of the CD3 signaling component and alters the means of further T cell expansion.

[0805] Hence, in an embodiment, editing of cells, particularly cells intended for adoptive cell therapies, more particularly immunoresponsive cells such as T cells, may be performed to knock-out or knock-down expression of an endogenous TCR in a cell. For example, NHEJ-based or HDR-based gene editing approaches can be employed to disrupt the endogenous TCR alpha and/or beta chain genes. For example, gene editing system or systems, such as IScB system or systems, can be designed to target a sequence found within the TCR beta chain conserved between the beta 1 and beta 2 constant region genes (TRBC1 and TRBC2) and/or to target the constant region of the TCR alpha chain (TRAC) gene.

[0806] Allogeneic cells are rapidly rejected by the host immune system. It has been demonstrated that, allogeneic leukocytes present in non-irradiated blood products will persist for no more than 5 to 6 days (Boni, Muranski et al. 2008 Blood 1;112(12):4746-54). Thus, to prevent rejection of allogeneic cells, the host's immune system usually has to be suppressed to some extent. However, in the case of adoptive cell transfer the use of immunosuppressive drugs also have a detrimental effect on the introduced therapeutic T cells. Therefore, to effectively use an adoptive immunotherapy approach in these conditions, the introduced cells would need to be resistant to the immunosuppressive treatment. Thus, in a particular embodiment, the present invention further comprises a step of modifying T cells to make them resistant to an immunosuppressive agent, preferably by inactivating at least one gene encoding a target for an immunosuppressive agent. An immunosuppressive agent is an agent that suppresses immune function by one of several mechanisms of action. An immunosuppressive agent can be, but is not limited to a calcineurin inhibitor, a target of rapamycin, an interleukin-2 receptor α -chain blocker, an inhibitor of inosine monophosphate dehydrogenase, an inhibitor of dihydrofolic acid reductase, a corticosteroid or an immunosuppressive antimetabolite. The present invention allows conferring immunosuppressive resistance to T cells for immunotherapy by inactivating the target of the immunosuppressive agent in T cells. As non-limiting examples, targets for an immunosuppressive agent can be a receptor for an immunosuppressive agent such as: CD52, glucocorticoid receptor (GR), a FKBP family gene member and a cyclophilin family gene member.

[0807] In an embodiment, editing of cells, particularly cells intended for adoptive cell therapies, more particularly immunoresponsive cells such as T cells, may be performed to block an immune checkpoint, such as to knock-out or knock-down expression of an immune checkpoint protein or receptor in a cell. Immune checkpoints are inhibitory pathways that slow down or stop immune reactions and prevent excessive tissue damage from uncontrolled activity of immune cells. In an embodiment, the immune checkpoint targeted is the programmed death-1 (PD-1 or CD279) gene (PDCD1). In other embodiments, the immune checkpoint targeted is cytotoxic T-lymphocyte-associated antigen (CTLA-4). In additional embodiments, the immune checkpoint targeted is another member of the CD28 and CTLA4 Ig superfamily such as BTLA, LAG3, ICOS, PDL1 or KIR. In further additional embodiments, the immune checkpoint targeted is a member of the TNFR superfamily such as CD40, OX40, CD137, GITR, CD27 or TIM-3.

[0808] Additional immune checkpoints include Src homology 2 domain-containing protein tyrosine phosphatase 1 (SHP-1) (Watson HA, et al., SHP-1: the next checkpoint target for cancer immunotherapy? *Biochem Soc Trans.* 2016 Apr 15;44(2):356-62). SHP-1 is a widely expressed inhibitory protein tyrosine phosphatase (PTP). In T-cells, it is a negative regulator of antigen-dependent activation and proliferation. It is a cytosolic protein, and therefore not amenable to antibody-mediated therapies, but its role in activation and proliferation makes it an attractive target for genetic manipulation in adoptive transfer strategies, such as chimeric antigen receptor (CAR) T cells. Immune checkpoints may also include T cell immunoreceptor with Ig and ITIM domains (TIGIT/Vstm3/WUCAM/VSIG9) and VISTA (Le Mercier I, et al., (2015) Beyond CTLA-4 and PD-1, the generation Z of negative checkpoint regulators. *Front. Immunol.* 6:418).

[0809] International Patent Publication No. WO 2014/172606 relates to the use of MT1 and/or MT2 inhibitors to increase proliferation and/or activity of exhausted CD8⁺ T-cells and to decrease CD8⁺ T-cell exhaustion (e.g., decrease functionally exhausted or unresponsive CD8⁺ immune cells). In an embodiment, metallothioneins are targeted by gene editing in adoptively transferred T cells.

[0810] In an embodiment, targets of gene editing may be at least one targeted locus involved in the expression of an immune checkpoint protein. Such targets may include, but are not limited to CTLA4, PPP2CA, PPP2CB, PTPN6, PTPN22, PDCD1, ICOS (CD278), PDL1, KIR, LAG3, HAVCR2, BTLA, CD160, TIGIT, CD96, CRTAM, LAIR1, SIGLEC7, SIGLEC9, CD244 (2B4), TNFRSF10B, TNFRSF10A, CASP8, CASP10, CASP3, CASP6, CASP7, FADD, FAS, TGFBR1, TGFBR2, SMAD2, SMAD3, SMAD4, SMAD10, SKI, SKIL, TGIF1, IL10RA, IL10RB, HMOX2, IL6R, IL6ST, EIF2AK4, CSK, PAG1, SIT1, FOXP3, PRDM1, BATF, VISTA, GUCY1A2, GUCY1A3, GUCY1B2, GUCY1B3, MT1, MT2, CD40, OX40, CD137, GITR, CD27, SHP-1, TIM-3, CEACAM-1, CEACAM-3, or CEACAM-5. In preferred embodiments, the gene locus involved in the expression of PD-1 or CTLA-4 genes is targeted. In other preferred embodiments, combinations of genes are targeted, such as but not limited to PD-1 and TIGIT.

[0811] By means of an example and without limitation, International Patent Publication No. WO 2016/196388 concerns an engineered T cell comprising (a) a genetically engineered antigen receptor that specifically binds to an antigen, which receptor may be a CAR; and (b) a disrupted gene encoding a PD-L1, an agent for disruption of a gene encoding a PD-L1, and/or

disruption of a gene encoding PD-L1, wherein the disruption of the gene may be mediated by a gene editing nuclease, a zinc finger nuclease (ZFN), CRISPR/Cas9 and/or TALEN. WO2015142675 relates to immune effector cells comprising a CAR in combination with an agent (such as the composition or system herein) that increases the efficacy of the immune effector cells in the treatment of cancer, wherein the agent may inhibit an immune inhibitory molecule, such as PD1, PD-L1, CTLA-4, TIM-3, LAG-3, VISTA, BTLA, TIGIT, LAIR1, CD160, 2B4, TGFR beta, CEACAM-1, CEACAM-3, or CEACAM-5. Ren et al., (2017) Clin Cancer Res 23 (9) 2255-2266 performed lentiviral delivery of CAR and electro-transfer of Cas9 mRNA and gRNAs targeting endogenous TCR, β -2 microglobulin (B2M) and PD1 simultaneously, to generate gene-disrupted allogeneic CAR T cells deficient of TCR, HLA class I molecule and PD1.

[0812] In an embodiment, cells may be engineered to express a CAR, wherein expression and/or function of methylcytosine dioxygenase genes (TET1, TET2 and/or TET3) in the cells has been reduced or eliminated, (such as the composition or system herein) (for example, as described in WO201704916).

[0813] In an embodiment, editing of cells, particularly cells intended for adoptive cell therapies, more particularly immunoresponsive cells such as T cells, may be performed to knock-out or knock-down expression of an endogenous gene in a cell, said endogenous gene encoding an antigen targeted by an exogenous CAR or TCR, thereby reducing the likelihood of targeting of the engineered cells. In an embodiment, the targeted antigen may be one or more antigen selected from the group consisting of CD38, CD138, CS-1, CD33, CD26, CD30, CD53, CD92, CD100, CD148, CD150, CD200, CD261, CD262, CD362, human telomerase reverse transcriptase (hTERT), survivin, mouse double minute 2 homolog (MDM2), cytochrome P450 1B1 (CYP1B), HER2/neu, Wilms' tumor gene 1 (WT1), livin, alphafetoprotein (AFP), carcinoembryonic antigen (CEA), mucin 16 (MUC16), MUC1, prostate-specific membrane antigen (PSMA), p53, cyclin (D1), B cell maturation antigen (BCMA), transmembrane activator and CAML Interactor (TACI), and B-cell activating factor receptor (BAFF-R) (for example, as described in International Patent Publication Nos. WO 2016/011210 and WO 2017/011804).

[0814] In an embodiment, editing of cells, particularly cells intended for adoptive cell therapies, more particularly immunoresponsive cells such as T cells, may be performed to knock-out or knock-down expression of one or more MHC constituent proteins, such as one or

more HLA proteins and/or beta-2 microglobulin (B2M), in a cell, whereby rejection of non-autologous (e.g., allogeneic) cells by the recipient's immune system can be reduced or avoided. In preferred embodiments, one or more HLA class I proteins, such as HLA-A, B and/or C, and/or B2M may be knocked-out or knocked-down. Preferably, B2M may be knocked-out or knocked-down. By means of an example, Ren et al., (2017) Clin Cancer Res 23 (9) 2255-2266 performed lentiviral delivery of CAR and electro-transfer of Cas mRNA and gRNAs targeting endogenous TCR, β -2 microglobulin (B2M) and PD1 simultaneously, to generate gene-disrupted allogeneic CAR T cells deficient of TCR, HLA class I molecule and PD1.

[0815] In other embodiments, at least two genes are edited. Pairs of genes may include, but are not limited to PD1 and TCR α , PD1 and TCR β , CTLA-4 and TCR α , CTLA-4 and TCR β , LAG3 and TCR α , LAG3 and TCR β , Tim3 and TCR α , Tim3 and TCR β , BTLA and TCR α , BTLA and TCR β , BY55 and TCR α , BY55 and TCR β , TIGIT and TCR α , TIGIT and TCR β , B7H5 and TCR α , B7H5 and TCR β , LAIR1 and TCR α , LAIR1 and TCR β , SIGLEC10 and TCR α , SIGLEC10 and TCR β , 2B4 and TCR α , 2B4 and TCR β , B2M and TCR α , B2M and TCR β .

[0816] In an embodiment, a cell may be multiplied edited (multiplex genome editing) as taught herein to (1) knock-out or knock-down expression of an endogenous TCR (for example, TRBC1, TRBC2 and/or TRAC), (2) knock-out or knock-down expression of an immune checkpoint protein or receptor (for example PD1, PD-L1 and/or CTLA4); and (3) knock-out or knock-down expression of one or more MHC constituent proteins (for example, HLA-A, B and/or C, and/or B2M, preferably B2M).

[0817] Whether prior to or after genetic modification of the T cells, the T cells can be activated and expanded generally using methods as described, for example, in U.S. Patent Nos. 6,352,694; 6,534,055; 6,905,680; 5,858,358; 6,887,466; 6,905,681; 7,144,575; 7,232,566; 7,175,843; 5,883,223; 6,905,874; 6,797,514; 6,867,041; and 7,572,631. T cells can be expanded *in vitro* or *in vivo*.

[0818] Immune cells may be obtained using any method known in the art. In one embodiment, allogeneic T cells may be obtained from healthy subjects. In one embodiment T cells that have infiltrated a tumor are isolated. T cells may be removed during surgery. T cells may be isolated after removal of tumor tissue by biopsy. T cells may be isolated by any means known in the art. In one embodiment, T cells are obtained by apheresis. In one embodiment, the method may comprise obtaining a bulk population of T cells from a tumor sample by any

suitable method known in the art. For example, a bulk population of T cells can be obtained from a tumor sample by dissociating the tumor sample into a cell suspension from which specific cell populations can be selected. Suitable methods of obtaining a bulk population of T cells may include, but are not limited to, any one or more of mechanically dissociating (e.g., mincing) the tumor, enzymatically dissociating (e.g., digesting) the tumor, and aspiration (e.g., as with a needle).

[0819] The bulk population of T cells obtained from a tumor sample may comprise any suitable type of T cell. Preferably, the bulk population of T cells obtained from a tumor sample comprises tumor infiltrating lymphocytes (TILs).

[0820] The tumor sample may be obtained from any mammal. Unless stated otherwise, as used herein, the term "mammal" refers to any mammal including, but not limited to, mammals of the order Logomorpha, such as rabbits; the order Carnivora, including Felines (cats) and Canines (dogs); the order Artiodactyla, including Bovines (cows) and Swines (pigs); or of the order Perssodactyla, including Equines (horses). The mammals may be non-human primates, e.g., of the order Primates, Ceboids, or Simoids (monkeys) or of the order Anthropoids (humans and apes). In one embodiment, the mammal may be a mammal of the order Rodentia, such as mice and hamsters. Preferably, the mammal is a non-human primate or a human. An especially preferred mammal is the human.

[0821] T cells can be obtained from a number of sources, including peripheral blood mononuclear cells (PBMC), bone marrow, lymph node tissue, spleen tissue, and tumors. In an embodiment of the present invention, T cells can be obtained from a unit of blood collected from a subject using any number of techniques known to the skilled artisan, such as Ficoll separation. In one preferred embodiment, cells from the circulating blood of an individual are obtained by apheresis or leukapheresis. The apheresis product typically contains lymphocytes, including T cells, monocytes, granulocytes, B cells, other nucleated white blood cells, red blood cells, and platelets. In one embodiment, the cells collected by apheresis may be washed to remove the plasma fraction and to place the cells in an appropriate buffer or media for subsequent processing steps. In one embodiment of the invention, the cells are washed with phosphate buffered saline (PBS). In an alternative embodiment, the wash solution lacks calcium and may lack magnesium or may lack many if not all divalent cations. Initial activation steps in the absence of calcium lead to magnified activation. As those of ordinary skill in the art would readily appreciate a washing step may be accomplished by methods known to those

in the art, such as by using a semi-automated “flow-through” centrifuge (for example, the Cobe 2991 cell processor) according to the manufacturer's instructions. After washing, the cells may be resuspended in a variety of biocompatible buffers, such as, for example, Ca-free, Mg-free PBS. Alternatively, the undesirable components of the apheresis sample may be removed and the cells directly resuspended in culture media.

[0822] In another embodiment, T cells are isolated from peripheral blood lymphocytes by lysing the red blood cells and depleting the monocytes, for example, by centrifugation through a PERCOLL™ gradient. A specific subpopulation of T cells, such as CD28+, CD4+, CDC, CD45RA+, and CD45RO+ T cells, can be further isolated by positive or negative selection techniques. For example, in one preferred embodiment, T cells are isolated by incubation with anti-CD3/anti-CD28 (i.e., 3×28)-conjugated beads, such as DYNABEADS® M-450 CD3/CD28 T, or XCYTE DYNABEADS™ for a time period sufficient for positive selection of the desired T cells. In one embodiment, the time period is about 30 minutes. In a further embodiment, the time period ranges from 30 minutes to 36 hours or longer and all integer values there between. In a further embodiment, the time period is at least 1, 2, 3, 4, 5, or 6 hours. In yet another preferred embodiment, the time period is 10 to 24 hours. In one preferred embodiment, the incubation time period is 24 hours. For isolation of T cells from patients with leukemia, use of longer incubation times, such as 24 hours, can increase cell yield. Longer incubation times may be used to isolate T cells in any situation where there are few T cells as compared to other cell types, such in isolating tumor infiltrating lymphocytes (TIL) from tumor tissue or from immunocompromised individuals. Further, use of longer incubation times can increase the efficiency of capture of CD8+ T cells.

[0823] Enrichment of a T cell population by negative selection can be accomplished with a combination of antibodies directed to surface markers unique to the negatively selected cells. A preferred method is cell sorting and/or selection via negative magnetic immunoadherence or flow cytometry that uses a cocktail of monoclonal antibodies directed to cell surface markers present on the cells negatively selected. For example, to enrich for CD4+ cells by negative selection, a monoclonal antibody cocktail typically includes antibodies to CD14, CD20, CD11b, CD16, HLA-DR, and CD8.

[0824] Further, monocyte populations (e.g., CD14+ cells) may be depleted from blood preparations by a variety of methodologies, including anti-CD14 coated beads or columns, or utilization of the phagocytotic activity of these cells to facilitate removal. Accordingly, in one

embodiment, the invention uses paramagnetic particles of a size sufficient to be engulfed by phagocytotic monocytes. In an embodiment, the paramagnetic particles are commercially available beads, for example, those produced by Life Technologies under the trade name Dynabeads™. In one embodiment, other non-specific cells are removed by coating the paramagnetic particles with “irrelevant” proteins (e.g., serum proteins or antibodies). Irrelevant proteins and antibodies include those proteins and antibodies or fragments thereof that do not specifically target the T cells to be isolated. In an embodiment, the irrelevant beads include beads coated with sheep anti-mouse antibodies, goat anti-mouse antibodies, and human serum albumin.

[0825] In brief, such depletion of monocytes is performed by preincubating T cells isolated from whole blood, apheresed peripheral blood, or tumors with one or more varieties of irrelevant or non-antibody coupled paramagnetic particles at any amount that allows for removal of monocytes (approximately a 20:1 bead:cell ratio) for about 30 minutes to 2 hours at 22 to 37 degrees C., followed by magnetic removal of cells which have attached to or engulfed the paramagnetic particles. Such separation can be performed using standard methods available in the art. For example, any magnetic separation methodology may be used including a variety of which are commercially available, (e.g., DYNAL® Magnetic Particle Concentrator (DYNAL MPC®)). Assurance of requisite depletion can be monitored by a variety of methodologies known to those of ordinary skill in the art, including flow cytometric analysis of CD14 positive cells, before and after depletion.

[0826] For isolation of a desired population of cells by positive or negative selection, the concentration of cells and surface (e.g., particles such as beads) can be varied. In an embodiment, it may be desirable to significantly decrease the volume in which beads and cells are mixed together (i.e., increase the concentration of cells), to ensure maximum contact of cells and beads. For example, in one embodiment, a concentration of 2 billion cells/ml is used. In one embodiment, a concentration of 1 billion cells/ml is used. In a further embodiment, greater than 100 million cells/ml is used. In a further embodiment, a concentration of cells of 10, 15, 20, 25, 30, 35, 40, 45, or 50 million cells/ml is used. In yet another embodiment, a concentration of cells from 75, 80, 85, 90, 95, or 100 million cells/ml is used. In further embodiments, concentrations of 125 or 150 million cells/ml can be used. Using high concentrations can result in increased cell yield, cell activation, and cell expansion. Further, use of high cell concentrations allows more efficient capture of cells that may weakly express

target antigens of interest, such as CD28-negative T cells, or from samples where there are many tumor cells present (i.e., leukemic blood, tumor tissue, etc). Such populations of cells may have therapeutic value and would be desirable to obtain. For example, using high concentration of cells allows more efficient selection of CD8⁺ T cells that normally have weaker CD28 expression.

[0827] In a related embodiment, it may be desirable to use lower concentrations of cells. By significantly diluting the mixture of T cells and surface (e.g., particles such as beads), interactions between the particles and cells is minimized. This selects for cells that express high amounts of desired antigens to be bound to the particles. For example, CD4⁺ T cells express higher levels of CD28 and are more efficiently captured than CD8⁺ T cells in dilute concentrations. In one embodiment, the concentration of cells used is $5 \times 10^6/\text{ml}$. In other embodiments, the concentration used can be from about $1 \times 10^5/\text{ml}$ to $1 \times 10^6/\text{ml}$, and any integer value in between.

[0828] T cells can also be frozen. Wishing not to be bound by theory, the freeze and subsequent thaw step provides a more uniform product by removing granulocytes and to some extent monocytes in the cell population. After a washing step to remove plasma and platelets, the cells may be suspended in a freezing solution. While many freezing solutions and parameters are known in the art and will be useful in this context, one method involves using PBS containing 20% DMSO and 8% human serum albumin, or other suitable cell freezing media, the cells then are frozen to -80°C at a rate of 1° per minute and stored in the vapor phase of a liquid nitrogen storage tank. Other methods of controlled freezing may be used as well as uncontrolled freezing immediately at -20°C . or in liquid nitrogen.

[0829] T cells for use in the present invention may also be antigen-specific T cells. For example, tumor-specific T cells can be used. In an embodiment, antigen-specific T cells can be isolated from a patient of interest, such as a patient afflicted with a cancer or an infectious disease. In one embodiment, neoepitopes are determined for a subject and T cells specific to these antigens are isolated. Antigen-specific cells for use in expansion may also be generated *in vitro* using any number of methods known in the art, for example, as described in U.S. Patent Publication No. US 20040224402 entitled, Generation and Isolation of Antigen-Specific T Cells, or in U.S. Pat. No. 6,040,177. Antigen-specific cells for use in the present invention may also be generated using any number of methods known in the art, for example, as described in

Current Protocols in Immunology, or Current Protocols in Cell Biology, both published by John Wiley & Sons, Inc., Boston, Mass.

[0830] In a related embodiment, it may be desirable to sort or otherwise positively select (e.g. via magnetic selection) the antigen specific cells prior to or following one or two rounds of expansion. Sorting or positively selecting antigen-specific cells can be carried out using peptide-MHC tetramers (Altman, et al., Science. 1996 Oct. 4; 274(5284):94-6). In another embodiment, the adaptable tetramer technology approach is used (Andersen et al., 2012 Nat Protoc. 7:891-902). Tetramers are limited by the need to utilize predicted binding peptides based on prior hypotheses, and the restriction to specific HLAs. Peptide-MHC tetramers can be generated using techniques known in the art and can be made with any MHC molecule of interest and any antigen of interest as described herein. Specific epitopes to be used in this context can be identified using numerous assays known in the art. For example, the ability of a polypeptide to bind to MHC class I may be evaluated indirectly by monitoring the ability to promote incorporation of ¹²⁵I labeled β 2-microglobulin (β 2m) into MHC class I/ β 2m/peptide heterotrimeric complexes (see Parker et al., J. Immunol. 152:163, 1994).

[0831] In one embodiment cells are directly labeled with an epitope-specific reagent for isolation by flow cytometry followed by characterization of phenotype and TCRs. In one embodiment, T cells are isolated by contacting with T cell specific antibodies. Sorting of antigen-specific T cells, or generally any cells of the present invention, can be carried out using any of a variety of commercially available cell sorters, including, but not limited to, MoFlo sorter (DakoCytomation, Fort Collins, Colo.), FACSAria™, FACSArray™, FACSVantage™, BD™ LSR II, and FACSCalibur™ (BD Biosciences, San Jose, Calif.).

[0832] In a preferred embodiment, the method comprises selecting cells that also express CD3. The method may comprise specifically selecting the cells in any suitable manner. Preferably, the selecting is carried out using flow cytometry. The flow cytometry may be carried out using any suitable method known in the art. The flow cytometry may employ any suitable antibodies and stains. Preferably, the antibody is chosen such that it specifically recognizes and binds to the particular biomarker being selected. For example, the specific selection of CD3, CD8, TIM-3, LAG-3, 4-1BB, or PD-1 may be carried out using anti-CD3, anti-CD8, anti-TIM-3, anti-LAG-3, anti-4-1BB, or anti-PD-1 antibodies, respectively. The antibody or antibodies may be conjugated to a bead (e.g., a magnetic bead) or to a fluorochrome. Preferably, the flow cytometry is fluorescence-activated cell sorting (FACS).

TCRs expressed on T cells can be selected based on reactivity to autologous tumors. Additionally, T cells that are reactive to tumors can be selected for based on markers using the methods described in patent publication Nos. WO2014133567 and WO2014133568, herein incorporated by reference in their entirety. Additionally, activated T cells can be selected for based on surface expression of CD107a.

[0833] In one embodiment of the invention, the method further comprises expanding the numbers of T cells in the enriched cell population. Such methods are described in U.S. Patent No. 8,637,307 and is herein incorporated by reference in its entirety. The numbers of T cells may be increased at least about 3-fold (or 4-, 5-, 6-, 7-, 8-, or 9-fold), more preferably at least about 10-fold (or 20-, 30-, 40-, 50-, 60-, 70-, 80-, or 90-fold), more preferably at least about 100-fold, more preferably at least about 1,000 fold, or most preferably at least about 100,000-fold. The numbers of T cells may be expanded using any suitable method known in the art. Exemplary methods of expanding the numbers of cells are described in patent publication No. WO 2003/057171, U.S. Patent No. 8,034,334, and U.S. Patent Publication No. 2012/0244133, each of which is incorporated herein by reference.

[0834] In one embodiment, *ex vivo* T cell expansion can be performed by isolation of T cells and subsequent stimulation or activation followed by further expansion. In one embodiment of the invention, the T cells may be stimulated or activated by a single agent. In another embodiment, T cells are stimulated or activated with two agents, one that induces a primary signal and a second that is a co-stimulatory signal. Ligands useful for stimulating a single signal or stimulating a primary signal and an accessory molecule that stimulates a second signal may be used in soluble form. Ligands may be attached to the surface of a cell, to an Engineered Multivalent Signaling Platform (EMSP), or immobilized on a surface. In a preferred embodiment both primary and secondary agents are co-immobilized on a surface, for example a bead or a cell. In one embodiment, the molecule providing the primary activation signal may be a CD3 ligand, and the co-stimulatory molecule may be a CD28 ligand or 4-1BB ligand.

[0835] In an embodiment, T cells comprising a CAR or an exogenous TCR, may be manufactured as described in International Patent Publication No. WO 2015/120096, by a method comprising enriching a population of lymphocytes obtained from a donor subject; stimulating the population of lymphocytes with one or more T-cell stimulating agents to produce a population of activated T cells, wherein the stimulation is performed in a closed

system using serum-free culture medium; transducing the population of activated T cells with a viral vector comprising a nucleic acid molecule which encodes the CAR or TCR, using a single cycle transduction to produce a population of transduced T cells, wherein the transduction is performed in a closed system using serum-free culture medium; and expanding the population of transduced T cells for a predetermined time to produce a population of engineered T cells, wherein the expansion is performed in a closed system using serum-free culture medium. In an embodiment, T cells comprising a CAR or an exogenous TCR, may be manufactured as described in WO 2015/120096, by a method comprising: obtaining a population of lymphocytes; stimulating the population of lymphocytes with one or more stimulating agents to produce a population of activated T cells, wherein the stimulation is performed in a closed system using serum-free culture medium; transducing the population of activated T cells with a viral vector comprising a nucleic acid molecule which encodes the CAR or TCR, using at least one cycle transduction to produce a population of transduced T cells, wherein the transduction is performed in a closed system using serum-free culture medium; and expanding the population of transduced T cells to produce a population of engineered T cells, wherein the expansion is performed in a closed system using serum-free culture medium. The predetermined time for expanding the population of transduced T cells may be 3 days. The time from enriching the population of lymphocytes to producing the engineered T cells may be 6 days. The closed system may be a closed bag system. Further provided is population of T cells comprising a CAR or an exogenous TCR obtainable or obtained by said method, and a pharmaceutical composition comprising such cells.

[0836] In an embodiment, T cell maturation or differentiation *in vitro* may be delayed or inhibited by the method as described in International Patent Publication No. WO 2017/070395, comprising contacting one or more T cells from a subject in need of a T cell therapy with an AKT inhibitor (such as, e.g., one or a combination of two or more AKT inhibitors disclosed in claim 8 of WO2017070395) and at least one of exogenous Interleukin-7 (IL-7) and exogenous Interleukin-15 (IL-15), wherein the resulting T cells exhibit delayed maturation or differentiation, and/or wherein the resulting T cells exhibit improved T cell function (such as, e.g., increased T cell proliferation; increased cytokine production; and/or increased cytolytic activity) relative to a T cell function of a T cell cultured in the absence of an AKT inhibitor.

[0837] In an embodiment, a patient in need of a T cell therapy may be conditioned by a method as described in International Patent Publication No. WO 2016/191756 comprising

administering to the patient a dose of cyclophosphamide between 200 mg/m²/day and 2000 mg/m²/day and a dose of fludarabine between 20 mg/m²/day and 900 mg/m²/day.

Diseases

Genetic Diseases and Diseases with a Genetic and/or Epigenetic Aspect

[0838] The compositions, systems, or components thereof can be used to treat and/or prevent a genetic disease or a disease with a genetic and/or epigenetic aspect. The genes and conditions exemplified herein are not exhaustive. In one embodiment, a method of treating and/or preventing a genetic disease can include administering a composition, system, and/or one or more components thereof to a subject, where the composition, system, and/or one or more components thereof is capable of modifying one or more copies of one or more genes associated with the genetic disease or a disease with a genetic and/or epigenetic aspect in one or more cells of the subject. In one embodiment, modifying one or more copies of one or more genes associated with a genetic disease or a disease with a genetic and/or epigenetic aspect in the subject can eliminate a genetic disease or a symptom thereof in the subject. In one embodiment, modifying one or more copies of one or more genes associated with a genetic disease or a disease with a genetic and/or epigenetic aspect in the subject can decrease the severity of a genetic disease or a symptom thereof in the subject. In one embodiment, the compositions, systems, or components thereof can modify one or more genes or polynucleotides associated with one or more diseases, including genetic diseases and/or those having a genetic aspect and/or epigenetic aspect, including but not limited to, any one or more set forth in Table 6. It will be appreciated that those diseases and associated genes listed herein are non-exhaustive and non-limiting. Further some genes play roles in the development of multiple diseases.

Table 6. Exemplary Genetic and Other Diseases and Associated Genes			
Disease Name	Primary Tissues or System Affected	Additional Tissues/Systems Affected	Genes
Achondroplasia	Bone and Muscle		fibroblast growth factor receptor 3 (FGFR3)
Achromatopsia	eye		CNGA3, CNGB3, GNAT2, PDE6C, PDE6H, ACHM2, ACHM3,
Acute Renal Injury	kidney		NFkappaB, AATF, p85alpha, FAS, Apoptosis cascade elements (e.g. FASR, Caspase 2, 3, 4, 6, 7, 8, 9, 10, AKT, TNF alpha, IGF1, IGF1R, RIPK1), p53

Age Related Macular Degeneration	eye		Abcr; CCL2; CC2; CP (ceruloplasmin); Timp3; cathepsinD; VLDLR, CCR2
AIDS	Immune System		KIR3DL1, NKAT3, NKB1, AMB11, KIR3DS1, IFNG, CXCL12, SDF1
Albinism (including oculocutaneous albinism (types 1-7) and ocular albinism)	Skin, hair, eyes,		TYR, OCA2, TYRP1, and SLC45A2, SLC24A5 and C10orf11
Alkaptonuria	Metabolism of amino acids	Tissues/organs where homogentisic acid accumulates, particularly cartilage (joints), heart valves, kidneys	HGD
alpha-1 antitrypsin deficiency (AATD or A1AD)	Lung	Liver, skin, vascular system, kidneys, GI	SERPINA1, those set forth in WO2017165862, PiZ allele
ALS	CNS		SOD1; ALS2; ALS3; ALS5; ALS7;STEX; FUS; TARDBP; VEGF (VEGF-a; VEGF-b; VEGF-c); DPP6; NEFH, PTGS1, SLC1A2, TNFRSF10B, PRPH, HSP90AA1, CRIA2, IFNG, AMPA2 S100B, FGF2, AOX1, CS, TXN, RAPHJ1, MAP3K5, NBEAL1, GPX1, ICA1L, RAC1, MAPT, ITPR2, ALS2CR4, GLS, ALS2CR8, CNTFR, ALS2CR11, FOLH1, FAM117B, P4HB, CNTF, SQSTM1, STRADB, NAIP, NLR, YWHAQ, SLC33A1, TRAK2, SCA1, NIF3L1, NIF3, PARD3B, COX8A, CDK15, HECW1, HECT, C2, WW 15, NOS1, MET, SOD2, HSPB1, NEFL, CTSB, ANG, HSPA8, RNase A, VAPB, VAMP, SNCA, alpha HGF, CAT, ACTB, NEFM, TH, BCL2, FAS, CASP3, CLU, SMN1, G6PD, BAX, HSF1, RNF19A, JUN, ALS2CR12, HSPA5, MAPK14, APEX1, TXNRD1, NOS2, TIMP1, CASP9, XIAP, GLG1, EPO, VEGFA, ELN, GDNF, NFE2L2, SLC6A3, HSPA4, APOE, PSMB8, DCTN2, TIMP3, KIFAP3, SLC1A1, SMN2, CCNC, STUB1, ALS2, PRDX6, SYP, CABIN1, CASP1, GART, CDK5, ATXN3, RTN4, C1QB, VEGFC, HTT, PARK7, XDH, GFAP, MAP2, CYCS, FCGR3B, CCS, UBL5, MMP9m SLC18A3, TRPM7, HSPB2, AKT1, DEERL1, CCL2, NGRN, GSR, TPPP3, APAF1, BTBD10, GLUD1, CXCR4, S:C1A3,

			FLT1, PON1, AR, LIF, ERBB3, :GA:S1, CD44, TP53, TLR3, GRIA1, GAPDH, AMPA, GRIK1, DES, CHAT, FLT4, CHMP2B, BAG1, CHRNA4, GSS, BAK1, KDR, GSTP1, OGG1, IL6
Alzheimer's Disease	Brain		E1; CHIP; UCH; UBB; Tau; LRP; PICALM; CLU; PS1; SORL1; CR1; VLDLR; UBA1; UBA3; CHIP28; AQP1; UCHL1; UCHL3; APP, AAA, CVAP, AD1, APOE, AD2, DCP1, ACE1, MPO, PACIP1, PAXIP1L, PTIP, A2M, BLMH, BMH, PSEN1, AD3, ALAS2, ABCA1, BIN1, BDNF, BTNL8, C1ORF49, CDH4, CHRNA2, CKLFSF2, CLEC4E, CR1L, CSF3R, CST3, CYP2C, DAPK1, ESR1, FCAR, FCGR3B, FFA2, FGA, GAB2, GALP, GAPDHS, GMPB, HP, HTR7, IDE, IF127, IFI6, IFIT2, IL1RN, IL-1RA, IL8RA, IL8RB, JAG1, KCNJ15, LRP6, MAPT, MARK4, MPHOSPH1, MTHFR, NBN, NCSTN, NIACR2, NMNAT3, NTM, ORM1, P2RY13, PBEF1, PCK1, PICALM, PLAUI, PLXNC1, PRNP, PSEN1, PSEN2, PTPRA, RALGPS2, RGSL2, SELENBP1, SLC25A37, SORL1, Mitoferrin-1, TF, TFAM, TNF, TNFRSF10C, UBE1C
Amyloidosis			APOA1, APP, AAA, CVAP, AD1, GSN, FGA, LYZ, TTR, PALB
Amyloid neuropathy			TTR, PALB
Anemia	Blood		CDAN1, CDA1, RPS19, DBA, PKLR, PK1, NT5C3, UMPH1, PSN1, RHAG, RH50A, NRAMP2, SPTB, ALAS2, ANH1, ASB, ABCB7, ABC7, ASAT
Angelman Syndrome	Nervous system, brain		UBE3A
Attention Deficit Hyperactivity Disorder (ADHD)	Brain		PTCHD1
Autoimmune lymphoproliferative syndrome	Immune system		TNFRSF6, APT1, FAS, CD95, ALPS1A
Autism, Autism spectrum disorders (ASDs), including Asperger's and a general diagnostic category called Pervasive Developmental Disorders (PDDs)	Brain		PTCHD1; Mecp2; BZRAP1; MDGA2; Sema5A; Neurexin 1; GLO1, RTT, PPMX, MRX16, RX79, NLGN3, NLGN4, KIAA1260, AUTSX2, FMR1, FMR2; FXR1; FXR2; MGLUR5, ATP10C, CDH10, GRM6, MGLUR6, CDH9, CNTN4, NLGN2, CNTNAP2, SEMA5A, DHCR7, NLGN4X, NLGN4Y, DPP6, NLGN5, EN2, NRCAM, MDGA2,

			NRXN1, FMR2, AFF2, FOXP2, OR4M2, OXTR, FXR1, FXR2, PAH, GABRA1, PTEN, GABRA5, PTPRZ1, GABRB3, GABRG1, HIRIP3, SEZ6L2, HOXA1, SHANK3, IL6, SHBZRAP1, LAMB1, SLC6A4, SERT, MAPK3, TAS2R1, MAZ, TSC1, MDGA2, TSC2, MECP2, UBE3A, WNT2, see also 20110023145
autosomal dominant polycystic kidney disease (ADPKD) - (includes diseases such as von Hippel-Lindau disease and tubereous sclerosis complex disease)	kidney	liver	PKD1, PKD2
Autosomal Recessive Polycystic Kidney Disease (ARPKD)	kidney	liver	PKDH1
Ataxia-Telangiectasia (a.k.a Louis Bar syndrome)	Nervous system, immune system	various	ATM
B-Cell Non-Hodgkin Lymphoma			BCL7A, BCL7
Bardet-Biedl syndrome	Eye, musculoskeletal system, kidney, reproductive organs	Liver, ear, gastrointestinal system, brain	ARL6, BBS1, BBS2, BBS4, BBS5, BBS7, BBS9, BBS10, BBS12, CEP290, INPP5E, LZTFL1, MKKS, MKS1, SDCCAG8, TRIM32, TTC8
Bare Lymphocyte Syndrome	blood		TAPBP, TPSN, TAP2, ABCB3, PSF2, RING11, MHC2TA, C2TA, RFX5, RFXAP, RFX5
Bartter's Syndrome (types I, II, III, IVA and B, and V)	kidney		SLC12A1 (type I), KCNJ1 (type II), CLCNKB (type III), BSND (type IV A), or both the CLCNKA CLCNKB genes (type IV B), CASR (type V).
Becker muscular dystrophy	Muscle		DMD, BMD, MYF6
Best Disease (Vitelliform Macular Dystrophy type 2)	eye		VMD2
Bleeding Disorders	blood		TBXA2R, P2RX1, P2X1
Blue Cone Monochromacy	eye		OPN1LW, OPN1MW, and LCR
Breast Cancer	Breast tissue		BRCA1, BRCA2, COX-2
Bruton's Disease (aka X-linked Agammaglobulinemia)	Immune system, specifically B cells		BTK
Cancers (e.g., lymphoma, chronic lymphocytic leukemia (CLL), B cell acute lymphocytic leukemia (B-ALL), acute lymphoblastic leukemia, acute myeloid leukemia, non-Hodgkin's lymphoma (NHL), diffuse large cell lymphoma (DLCL), multiple myeloma, renal cell carcinoma (RCC), neuroblastoma, colorectal cancer, breast cancer, ovarian cancer, melanoma, sarcoma,	Various		FAS, BID, CTLA4, PDCD1, CBLB, PTPN6, TRAC, TRBC, those described in WO2015048577

prostate cancer, lung cancer, esophageal cancer, hepatocellular carcinoma, pancreatic cancer, astrocytoma, mesothelioma, head and neck cancer, and medulloblastoma			
Cardiovascular Diseases	heart	Vascular system	<p>IL1B, XDH, TP53, PTGS, MB, IL4, ANGPT1, ABCG8, CTSK, PTGIR, KCNJ11, INS, CRP, PDGFRB, CCNA2, PDGFB, KCNJ5, KCNN3, CAPN10, ADRA2B, ABCG5, PRDX2, CPAN5, PARP14, MEX3C, ACE, RNF, IL6, TNF, STN, SERPINE1, ALB, ADIPOQ, APOB, APOE, LEP, MTHFR, APOA1, EDN1, NPPB, NOS3, PPARG, PLAT, PTGS2, CETP, AGTR1, HMGR, IGF1, SELE, REN, PPARA, PON1, KNG1, CCL2, LPL, VWF, F2, ICAM1, TGFB, NPPA, IL10, EPO, SOD1, VCAM1, IFNG, LPA, MPO, ESR1, MAPK, HP, F3, CST3, COG2, MMP9, SERPINC1, F8, HMOX1, APOC3, IL8, PROL1, CBS, NOS2, TLR4, SELP, ABCA1, AGT, LDLR, GPT, VEGFA, NR3C2, IL18, NOS1, NR3C1, FGB, HGF, IL1A, AKT1, LIPC, HSPD1, MAPK14, SPP1, ITGB3, CAT, UTS2, THBD, F10, CP, TNFRSF11B, EGFR, MMP2, PLG, NPY, RHOD, MAPK8, MYC, FN1, CMA1, PLA2, GNB3, ADRB2, SOD2, F5, VDR, ALOX5, HLA-DRB1, PARP1, CD40LG, PON2, AGER, IRS1, PTGS1, ECE1, F7, IRMN, EPHX2, IGFBP1, MAPK10, FAS, ABCB1, JUN, IGFBP3, CD14, PDE5A, AGTR2, CD40, LCAT, CCR5, MMP1, TIMP1, ADM, DYT10, STAT3, MMP3, ELN, USF1, CFH, HSPA4, MMP12, MME, F2R, SELL, CTSB, ANXA5, ADRB1, CYBA, FGA, GGT1, LIPG, HIF1A, CXCR4, PROC, SCARB1, CD79A, PLTP, ADD1, FGG, SAA1, KCNH2, DPP4, NPR1, VTN, KIAA0101, FOS, TLR2, PPIG, IL1R1, AR, CYP1A1, SERPINA1, MTR, RBP4, APOA4, CDKN2A, FGF2, EDNRB, ITGA2, VLA-2, CABIN1, SHBG, HMGB1, HSP90B2P, CYP3A4, GJA1, CAV1, ESR2, LTA, GDF15, BDNF, CYP2D6, NGF, SP1, TGIF1, SRC, EGF, PIK3CG, HLA-A, KCNQ1,</p>

		<p> CNR1, FBN1, CHKA, BEST1, CTNNB1, IL2, CD36, PRKAB1, TPO, ALDH7A1, CX3CR1, TH, F9, CH1, TF, HFE, IL17A, PTEN, GSTM1, DMD, GATA4, F13A1, TTR, FABP4, PON3, APOC1, INSR, TNFRSF1B, HTR2A, CSF3, CYP2C9, TXN, CYP11B2, PTH, CSF2, KDR, PLA2G2A, THBS1, GCG, RHOA, ALDH2, TCF7L2, NFE2L2, NOTCH1, UGT1A1, IFNA1, PPARD, SIRT1, GNHR1, PAPPA, ARR3, NPPC, AHSP, PTK2, IL13, MTOR, ITGB2, GSTT1, IL6ST, CPB2, CYP1A2, HNF4A, SLC64A, PLA2G6, TNFSF11, SLC8A1, F2RL1, AKR1A1, ALDH9A1, BGLAP, MTTP, MTRR, SULT1A3, RAGE, C4B, P2RY12, RNLS, CREB1, POMC, RAC1, LMNA, CD59, SCM5A, CYP1B1, MIF, MMP13, TIMP2, CYP19A1, CUP21A2, PTPN22, MYH14, MBL2, SELPLG, AOC3, CTSL1, PCNA, IGF2, ITGB1, CAST, CXCL12, IGHE, KCNE1, TFRC, COL1A1, COL1A2, IL2RB, PLA2G10, ANGPT2, PROCR, NOX4, HAMP, PTPN11, SLCA1, IL2RA, CCL5, IRF1, CF:AR, CA:CA, EIF4E, GSTP1, JAK2, CYP3A5, HSPG2, CCL3, MYD88, VIP, SOAT1, ADRBK1, NR4A2, MMP8, NPR2, GCH1, EPRS, PPARGC1A, F12, PECAM1, CCL4, CERPINA34, CASR, FABP2, TTF2, PROS1, CTF1, SGCB, YME1L1, CAMP, ZC3H12A, AKR1B1, MMP7, AHR, CSF1, HDAC9, CTGF, KCNMA1, UGT1A, PRKCA, COMT, S100B, EGR1, PRL, IL15, DRD4, CAMK2G, SLC22A2, CCL11, PGF, THPO, GP6, TACR1, NTS, HNF1A, SST, KCDN1, LOC646627, TBXAS1, CUP2J2, TBXA2R, ADH1C, ALOX12, AHSB, BHMT, GJA4, SLC25A4, ACLY, ALOX5AP, NUMA1, CYP27B1, CYSLTR2, SOD3, LTC4S, UCN, GHRL, APOC2, CLEC4A, KBTBD10, TNC, TYMS, SHC1, LRP1, SOCS3, ADH1B, KLK3, HSD11B1, VKORC1, SERPINB2, TNS1, RNF19A, EPOR, ITGAM, PITX2, MAPK7, FCGR3A, LEEPR, ENG, GPX1, GOT2, HRH1, NR112, CRH, HTR1A, VDAC1, HPSE, SFTPD, </p>
--	--	--

			<p>TAP2, RMF123, PTK2Bm NTRK2, IL6R, ACHE, GLP1R, GHR, GSR, NQO1, NR5A1, GJB2, SLC9A1, MAOA, PCSK9, FCGR2A, SERPINF1, EDN3, UCP2, TFAP2A, C4BPA, SERPINF2, TYMP, ALPP, CXCR2, SLC3A3, ABCG2, ADA, JAK3, HSPA1A, FASN, FGF1, F11, ATP7A, CR1, GFPA, ROCK1, MECP2, MYLK, BCHE, LIPE, ADORA1, WRN, CXCR3, CD81, SMAD7, LAMC2, MAP3K5, CHGA, IAPP, RHO, ENPP1, PTHLH, NRG1, VEGFC, ENPEP, CEBPB, NAGLU, F2RL3, CX3CL1, BDKRB1, ADAMTS13, ELANE, ENPP2, CISH, GAST, MYOC, ATP1A2, NF1, GJB1, MEF2A, VCL, BMPR2, TUBB, CDC42, KRT18, HSF1, MYB, PRKAA2, ROCK2, TFP1, PRKG1, BMP2, CTNND1, CTH, CTSS, VAV2, NPY2R, IGFBP2, CD28, GSTA1, PPIA, APOH, S100A8, IL11, ALOX15, FBLN1, NR1H3, SCD, GIP, CHGB, PRKCB, SRD5A1, HSD11B2, CALCRL, GALNT2, ANGPTL4, KCNN4, PIK3C2A, HBEGF, CYP7A1, HLA-DRB5, BNIP3, GCKR, S100A12, PADI4, HSPA14, CXCR1, H19, KRTAP19-3, IDDM2, RAC2, YRY1, CLOCK, NGFR, DBH, CHRNA4, CACNA1C, PRKAG2, CHAT, PTGDS, NR1H2, TEK, VEGFB, MEF2C, MAPKAPK2, TNFRSF11A, HSPA9, CYSLTR1, MAT1A, OPRL1, IMPA1, CLCN2, DLD, PSMA6, PSMB8, CHI3L1, ALDH1B1, PARP2, STAR, LBP, ABCC6, RGS2, EFNB2, GJB6, APOA2, AMPD1, DYSF, FDFT1, EMD2, CCR6, GJB3, IL1RL1, ENTPD1, BBS4, CELSR2, F11R, RAPGEF3, HYAL1, ZNF259, ATOX1, ATF6, KHK, SAT1, GGH, TIMP4, SLC4A4, PDE2A, PDE3B, FADS1, FADS2, TMSB4X, TXNIP, LIMS1, RHOB, LY96, FOXO1, PNPLA2, TRH, GJC1, SLC17A5, FTO, GJD2, PRSC1, CASP12, GPBAR1, PXX, IL33, TRIB1, PBX4, NUPR1, 15-SEP, CILP2, TERC, GGT2, MTCO1, UOX, AVP</p>
Cataract	eye		<p>CRYAA, CRYA1, CRYBB2, CRYB2, PITX3, BFSP2, CP49, CP47, CRYAA, CRYA1, PAX6,</p>

			AN2, MGDA, CRYBA1, CRYB1, CRYGC, CRYG3, CCL, LIM2, MP19, CRYGD, CRYG4, BFSP2, CP49, CP47, HSF4, CTM, HSF4, CTM, MIP, AQP0, CRYAB, CRYA2, CTPP2, CRYBB1, CRYGD, CRYG4, CRYBB2, CRYB2, CRYGC, CRYG3, CCL, CRYAA, CRYA1, GJA8, CX50, CAE1, GJA3, CX46, CZP3, CAE3, CCM1, CAM, KRIT1
CDKL-5 Deficiencies or Mediated Diseases	Brain, CNS		CDKL5
Charcot-Marie-Tooth (CMT) disease (Types 1, 2, 3, 4,)	Nervous system	Muscles (dystrophy)	PMP22 (CMT1A and E), MPZ (CMT1B), LITAF (CMT1C), EGR2 (CMT1D), NEFL (CMT1F), GJB1 (CMT1X), MFN2 (CMT2A), KIF1B (CMT2A2B), RAB7A (CMT2B), TRPV4 (CMT2C), GARS (CMT2D), NEFL (CMT2E), GAPD1 (CMT2K), HSPB8 (CMT2L), DYNC1H1, CMT2O), LRSAM1 (CMT2P), IGHMBP2 (CMT2S), MORC2 (CMT2Z), GDAP1 (CMT4A), MTMR2 or SBF2/MTMR13 (CMT4B), SH3TC2 (CMT4C), NDRG1 (CMT4D), PRX (CMT4F), FIG4 (CMT4J), NT-3
Chédiak-Higashi Syndrome	Immune system	Skin, hair, eyes, neurons	LYST
Choroideremia			CHM, REP1,
Chorioretinal atrophy	eye		PRDM13, RGR, TEAD1
Chronic Granulomatous Disease	Immune system		CYBA, CYBB, NCF1, NCF2, NCF4
Chronic Mucocutaneous Candidiasis	Immune system		AIRE, CARD9, CLEC7A IL12B, IL12B1, IL1F, IL17RA, IL17RC, RORC, STAT1, STAT3, TRAF31P2
Cirrhosis	liver		KRT18, KRT8, CIRH1A, NAIC, TEX292, KIAA1988
Colon cancer (Familial adenomatous polyposis (FAP) and hereditary nonpolyposis colon cancer (HNPCC))	Gastrointestinal		FAP: APC HNPCC: MSH2, MLH1, PMS2, SH6, PMS1
Combined Immunodeficiency	Immune System		IL2RG, SCIDX1, SCIDX, IMD4); HIV-1 (CCL5, SCYA5, D17S136E, TCP228
Cone(-rod) dystrophy	eye		AIPL1, CRX, GUA1A, GUCY2D, PITPM3, PROM1, PRPH2, RIMS1, SEMA4A, ABCA4, ADAM9, ATF6, C21ORF2, C8ORF37, CACNA2D4, CDHR1, CERKL, CNGA3, CNGB3, CNNM4, CNAT2, IFT81, KCNV2, PDE6C, PDE6H, POC1B, RAX2, RDH5, RPGRIP1, TTLL5, RetCG1, GUCY2E
Congenital Stationary Night Blindness	eye		CABP4, CACNA1F, CACNA2D4, GNAT1, CPR179, GRK1, GRM6,

			LRIT3, NYX, PDE6B, RDH5, RHO, RLBP1, RPE65, SAG, SLC24A1, TRPM1,
Congenital Fructose Intolerance	Metabolism		ALDOB
Cori's Disease (Glycogen Storage Disease Type III)	Various- wherever glycogen accumulates, particularly liver, heart, skeletal muscle		AGL
Corneal clouding and dystrophy	eye		APOA1, TGFBI, CSD2, CDGG1, CSD, BIGH3, CDG2, TACSTD2, TROP2, M1S1, VSX1, RINX, PPCD, PPD, KTCN, COL8A2, FECD, PPCD2, PIP5K3, CFD
Cornea plana congenital			KERA, CNA2
Cri du chat Syndrome, also known as 5p syndrome and cat cry syndrome			Deletions involving only band 5p15.2 to the entire short arm of chromosome 5, e.g. CTNND2, TERT,
Cystic Fibrosis (CF)	Lungs and respiratory system	Pancreas, liver, digestive system, reproductive system, exocrine, glands,	CTFR, ABCC7, CF, MRP7, SCNN1A, those described in WO2015157070
Diabetic nephropathy	kidney		Gremlin, 12/15- lipoxygenase, TIM44,
Dent Disease (Types 1 and 2)	Kidney		Type 1: CLCN5, Type 2: ORCL
Dentatorubro-Pallidolusian Atrophy (DRPLA) (aka Haw River and Naito-Oyanagi Disease)	CNS, brain, muscle		Atrophin-1 and Atnl
Down Syndrome	various		Chromosome 21 trisomy
Drug Addiction	Brain		Prkce; Drd2; Drd4; ABAT; GRIA2; Grm5; Grin1; Htr1b; Grin2a; Drd3; Pdyn; Grial
Duane syndrome (Types 1, 2, and 3, including subgroups A, B and C). Other names for this condition include: Duane's Retraction Syndrome (or DR syndrome), Eye Retraction Syndrome, Retraction Syndrome, Congenital retraction syndrome and Stilling-Turk-Duane Syndrome	eye		CHN1, indels on chromosomes 4 and 8
Duchenne muscular dystrophy (DMD)	muscle	Cardiovascular, respiratory	DMD, BMD, dystrophin gene, intron flanking exon 51 of DMD gene, exon 51 mutations in DMD gene, see also WO2013163628 and US Pat. Pub. 20130145487
Edward's Syndrome (Trisomy 18)			Complete or partial trisomy of chromosome 18
Ehlers-Danlos Syndrome (Types I-VI)	Various depending on type: including		COL5A1, COL5A2, COL1A1, COL3A1, TNXB, PLOD1, COL1A2, FKBP14 and ADAMTS2

	musculoskeletal, eye, vasculature, immune, and skin		
Emery-Dreifuss muscular dystrophy	muscle		LMNA, LMN1, EMD2, FPLD, CMD1A, HGPS, LGMD1B, LMNA, LMN1, EMD2, FPLD, CMD1A
Enhanced S-Cone Syndrome	eye		NR2E3, NRL
Fabry's Disease	Various – including skin, eyes, and gastrointestinal system, kidney, heart, brain, nervous system		GLA
Facioscapulohumeral muscular dystrophy	muscles		FSHMD1A, FSHD1A, FRG1,
Factor H and Factor H-like 1	blood		HF1, CFH, HUS
Factor V Leiden thrombophilia and Factor V deficiency	blood		Factor V (F5)
Factor V and Factor VII deficiency	blood		MCFD2
Factor VII deficiency	blood		F7
Factor X deficiency	blood		F10
Factor XI deficiency	blood		F11
Factor XII deficiency	blood		F12, HAF
Factor XIII A deficiency	blood		F13A1, F13A
Factor XIII B deficiency	blood		F13B
Familial Hypercholesterolemia	Cardiovascular system		APOB, LDLR, PCSK9
Familial Mediterranean Fever (FMF) also called recurrent polyserositis or familial paroxysmal polyserositis	Various- organs/tissues with serous or synovial membranes, skin, joints	Heart, kidney, brain/CNS, reproductive organs	MEFV
Fanconi Anemia	Various – blood (anemia), immune system, cognitive, kidneys, eyes, musculoskeletal		FANCA, FACA, FA1, FA, FAA, FAAP95, FAAP90, FLJ34064, FANCC, FANCG, RAD51, BRCA1, BRCA2, BRIP1, BACH1, FANCI, FANCB, FANCD1, FANCD2, FANCD, FAD, FANCE, FACE, FANCF, FANCI, ERCC4, FANCL, FANCM, PALB2, RAD51C, SLX4, UBE2T, FANCB, XRCC9, PHF9, KIAA1596
Fanconi Syndrome Types I (Childhood onset) and II (Adult Onset)	kidneys		FRTS1, GATM
Fragile X syndrome and related disorders	brain		FMR1, FMR2; FXR1; FXR2; mGLUR5
Fragile XE Mental Retardation (aka Martin Bell syndrome)	Brain, nervous system		FMR1
Friedreich Ataxia (FRDA)	Brain, nervous system	heart	FXN/X25

Fuchs endothelial corneal dystrophy	Eye		TCF4; COL8A2
Galactosemia	Carbohydrate metabolism disorder	Various-where galactose accumulates – liver, brain, eyes	GALT, GALK1, and GALE
Gastrointestinal Epithelial Cancer, GI cancer			CISH
Gaucher Disease (Types 1, 2, and 3, as well as other unusual forms that may not fit into these types)	Fat metabolism disorder	Various-liver, spleen, blood, CNS, skeletal system	GBA
GrisCELLI syndrome			
Glaucoma	eye		MYOC, TIGR, GLC1A, JOAG, GPOA, OPTN, GLC1E, FIP2, HYPL, NRP, CYP1B1, GLC3A, OPA1, NTG, NPG, CYP1B1, GLC3A, those described in WO2015153780
Glomerulo sclerosis	kidney		CC chemokine ligand 2
Glycogen Storage Diseases Types I-VI -See also Cori's Disease, Pompe's Disease, McArdle's disease, Hers Disease, and Von Gierke's disease	Metabolism Diseases		SLC2A2, GLUT2, G6PC, G6PT, G6PT1, GAA, LAMP2, LAMPB, AGL, GDE, GBE1, GYS2, PYGL, PFKM, see also Cori's Disease, Pompe's Disease, McArdle's disease, Hers Disease, and Von Gierke's disease
RBC Glycolytic enzyme deficiency	blood		any mutations in a gene for an enzyme in the glycolysis pathway including mutations in genes for hexokinases I and II, glucokinase, phosphoglucose isomerase, phosphofructokinase, aldolase B, triosephosphate isomerase, glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerokinase, phosphoglycerate mutase, enolase I, pyruvate kinase
Hartnup's disease	Malabsorption disease	Various- brain, gastrointestinal, skin,	SLC6A19
Hearing Loss	ear		NOX3, Hes5, BDNF,
Hemochromatosis (HH)	Iron absorption regulation disease	Various-whenever iron accumulates, liver, heart, pancreas, joints, pituitary gland	HFE and H63D
Hemophagocytic lymphohistiocytosis disorders	blood		PRF1, HPLH2, UNC13D, MUNC13-4, HPLH3, HLH3, FHL3
Hemorrhagic disorders	blood		PI, ATT, F5
Hers disease (Glycogen storage disease Type VI)	liver	muscle	PYGL
Hereditary angioedema (HAE)			kalikrein B1
Hereditary Hemorrhagic Telangiectasia (Osler-Weber-Rendu Syndrome)	Skin and mucous membranes		ACVRL1, ENG and SMAD4

Hereditary Spherocytosis	blood		NK1, EPB42, SLC4A1, SPTA1, and SPTB
Hereditary Persistence of Fetal Hemoglobin	blood		HBG1, HBG2, BCL11A, promoter region of HBG 1 and/or 2 (in the CCAAT box)
Hemophilia (hemophilia A (Classic) a B (aka Christmas disease) and C)	blood		A: FVIII, F8C, HEMA B: FVIX, HEMB C: F9, F11
Hepatic adenoma	liver		TCF1, HNF1A, MODY3
Hepatic failure, early onset, and neurologic disorder	liver		SCOD1, SCO1
Hepatic lipase deficiency	liver		LIPC
Hepatoblastoma, cancer and carcinomas	liver		CTNNB1, PDGFRL, PDGRL, PRLTS, AXIN1, AXIN, CTNNB1, TP53, P53, LFS1, IGF2R, MPRI, MET, CASP8, MCH5
Hermansky-Pudlak syndrome	Skin, eyes, blood, lung, kidneys, intestine		HPS1, HPS3, HPS4, HPS5, HPS6, HPS7, DTNBP1, BLOC1, BLOC1S2, BLOC3
HIV susceptibility or infection	Immune system		IL10, CSIF, CMKBR2, CCR2, CMKBR5, CCCKR5 (CCR5), those in WO2015148670A1
Holoprosencephaly (HPE) (Alobar, Semilobar, and Lobar)	brain		ACVRL1, ENG, SMAD4
Homocystinuria	Metabolic disease	Various-connective tissue, muscles, CNS, cardiovascular system	CBS, MTHFR, MTR, MTRR, and MMADHC
HPV			HPV16 and HPV18 E6/E7
HSV1, HSV2, and related keratitis	eye		HSV1 genes (immediate early and late HSV-1 genes (UL1, 1.5, 5, 6, 8, 9, 12, 15, 16, 18, 19, 22, 23, 26, 26.5, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 42, 48, 49.5, 50, 52, 54, S6, RL2, RS1, those described in WO2015153789, WO2015153791
Hunter's Syndrome (aka Mucopolysaccharidosis type II)	Lysosomal storage disease	Various- liver, spleen, eye, joint, heart, brain, skeletal	IDS
Huntington's disease (HD) and HD-like disorders	Brain, nervous system		HD, HTT, IT15, PRNP, PRIP, JPH3, JP3, HDL2, TBP, SCA17, PRKCE; IGF1; EP300; RCOR1; PRKCZ; HDAC4; and TGM2, and those described in WO2013130824, WO2015089354
Hurler's Syndrome (aka mucopolysaccharidosis type I H, MPS IH)	Lysosomal storage disease	Various- liver, spleen, eye, joint, heart, brain, skeletal	IDUA, α -L-iduronidase
Hurler-Scheie syndrome (aka mucopolysaccharidosis type I H-S, MPS I H-S)	Lysosomal storage disease	Various- liver, spleen, eye, joint, heart, brain, skeletal	IDUA, α -L-iduronidase

hyaluronidase deficiency (aka MPS IX)	Soft and connective tissues		HYAL1
Hyper IgM syndrome	Immune system		CD40L
Hyper- tension caused renal damage	kidney		Mineral corticoid receptor
Immunodeficiencies	Immune System		CD3E, CD3G, AICDA, AID, HIGM2, TNFRSF5, CD40, UNG, DGU, HIGM4, TNFSF5, CD40LG, HIGM1, IGM, FOXP3, IPEX, AIID, XPID, PIDX, TNFRSF14B, TACI
Inborn errors of metabolism: including urea cycle disorders, organic acidemias, fatty acid oxidation defects, amino acidopathies, carbohydrate disorders, mitochondrial disorders	Metabolism diseases, liver	Various organs and cells	See also: Carbohydrate metabolism disorders (e.g. galactosemia), Amino acid Metabolism disorders (e.g. phenylketonuria), Fatty acid metabolism (e.g. MCAD deficiency), Urea Cycle disorders (e.g. Citrullinemia), Organic acidemias (e.g. Maple Syrup Urine disease), Mitochondrial disorders (e.g. MELAS), peroxisomal disorders (e.g. Zellweger syndrome)
Inflammation	Various		IL-10; IL-1 (IL-1a; IL-1b); IL-13; IL-17 (IL-17a (CTLA8); IL-17b; IL-17c; IL-17d; IL-17f); IL-23; Cx3cr1; ptpn22; TNFa; NOD2/CARD15 for IBD; IL-6; IL-12 (IL-12a; IL-12b); CTLA4; Cx3cl1
Inflammatory Bowel Diseases (e.g. Ulcerative Colitis and Chron's Disease)	Gastrointestinal	Joints, skin	NOD2, IRGM, LRRK2, ATG5, ATG16L1, IRGM, GATM, ECM1, CDH1, LAMB1, HNF4A, GNA12, IL10, CARD9/15, CCR6, IL2RA, MST1, TNFSF15, REL, STAT3, IL23R, IL12B, FUT2
Interstitial renal fibrosis	kidney		TGF- β type II receptor
Job's Syndrome (aka Hyper IgE Syndrome)	Immune System		STAT3, DOCK8
Juvenile Retinoschisis	eye		RS1, XLRS1
Kabuki Syndrome 1			MLL4, KMT2D
Kennedy Disease (aka Spinobulbar Muscular Atrophy)	Muscles, brain, nervous system		SBMA/SMAX1/AR
Klinefelter syndrome	Various-particularly those involved in development of male characteristics		Extra X chromosome in males
Lafora Disease	Brain, CNS		EMP2A and EMP2B
Leber Congenital Amaurosis	eye		CRB1, RP12, CORD2, CRD, CRX, IMPDH1, OTX2, AIPL1, CABP4, CCT2, CEP290, CLUAP1, CRB1, CRX, DTHD1, GDF6, GUCY2D, IFT140, IQCB1, KCNJ13, LCA5, LRAT, NMNAT1, PRPH2, RD3, RDH12, RPE65, RP20, RPGRIP1,

			SPATA7, TULP1, LCA1, LCA4, GUC2D, CORD6, LCA3,
Lesch-Nyhan Syndrome	Metabolism disease	Various – joints, cognitive, brain, nervous system	HPRT1
Leukocyte deficiencies and disorders	blood		ITGB2, CD18, LCAMB, LAD, EIF2B1, EIF2BA, EIF2B2, EIF2B3, EIF2B5, LVWM, CACH, CLE, EIF2B4
Leukemia	Blood		TAL1, TCL5, SCL, TAL2, FLT3, NBS1, NBS, ZNFN1A1, IK1, LYF1, HOXD4, HOX4B, BCR, CML, PHL, ALL, ARNT, KRAS2, RASK2, GMPS, AF10, ARHGEF12, LARG, KIAA0382, CALM, CLTH, CEBPA, CEBP, CHIC2, BTL, FLT3, KIT, PBT, LPP, NPM1, NUP214, D9S46E, CAN, CAIN, RUNX1, CBFA2, AML1, WHSC1L1, NSD3, FLT3, AF1Q, NPM1, NUMA1, ZNF145, PLZF, PML, MYL, STAT5B, AF10, CALM, CLTH, ARL11, ARLTS1, P2RX7, P2X7, BCR, CML, PHL, ALL, GRAF, NF1, VRNF, WSS, NFNS, PTPN11, PTP2C, SHP2, NS1, BCL2, CCND1, PRAD1, BCL1, TCRA, GATA1, GF1, ERYF1, NFE1, ABL1, NQO1, DIA4, NMOR1, NUP214, D9S46E, CAN, CAIN
Limb-girdle muscular dystrophy diseases	muscle		LGMD
Lowe syndrome	brain, eyes, kidneys		OCRL
Lupus glomerulo- nephritis	kidney		MAPK1
Machado-Joseph's Disease (also known as Spinocerebellar ataxia Type 3)	Brain, CNS, muscle		ATX3
Macular degeneration	eye		ABC4, CBC1, CHM1, APOE, C1QTNF5, C2, C3, CCL2, CCR2, CD36, CFB, CFH, CFHR1, CFHR3, CNGB3, CP, CRP, CST3, CTSD, CX3CR1, ELOVL4, ERCC6, FBLN5, FBLN6, FSCN2, HMCN1, HTRA1, IL6, IL8, PLEKHA1, PROM1, PRPH2, RPGR, SERPING1, TCOF1, TIMP3, TLR3
Macular Dystrophy	eye		BEST1, C1QTNF5, CTNNA1, EFEMP1, ELOVL4, FSCN2, GUCA1B, HMCN1, IMPG1, OTX2, PRDM13, PROM1, PRPH2, RPIL1, TIMP3, ABCA4, CFH, DRAM2, IMG1, MFSD8, ADMD, STGD2, STGD3, RDS, RP7, PRPH, AVMD, AOFMD, VMD2
Malattia Leventinese	eye		EFEMP1, FBLN3

Maple Syrup Urine Disease	Metabolism disease		BCKDHA, BCKDHB, and DBT
Marfan syndrome	Connective tissue	Musculoskeletal	FBN1
Maroteaux-Lamy Syndrome (aka MPS VI)	Musculoskeletal system, nervous system	Liver, spleen	ARSB
McArdle's Disease (Glycogen Storage Disease Type V)	Glycogen storage disease	muscle	PYGM
Medullary cystic kidney disease	kidney		UMOD, HNFJ, FJHN, MCKD2, ADMCKD2
Metachromatic leukodystrophy	Lysosomal storage disease	Nervous system	ARSA
Methylmalonic acidemia (MMA)	Metabolism disease		MMAA, MMAB, MUT, MMACHC, MMADHC, LMBRD1
Morquio Syndrome (aka MPS IV A and B)	Connective tissue, skin, bone, eyes	heart	GALNS
Mucopolysaccharidosis diseases (Types I H/S, I H, II, III A B and C, I S, IVA and B, IX, VII, and VI)	Lysosomal storage disease – affects various organs/tissues		See also Hurler/Scheie syndrome, Hurler disease, Sanfillipo syndrome, Scheie syndrome, Morquio syndrome, hyaluronidase deficiency, Sly syndrome, and Maroteaux-Lamy syndrome
Muscular Atrophy	muscle		VAPB, VAPC, ALS8, SMN1, SMA1, SMA2, SMA3, SMA4, BSCL2, SPG17, GARS, SMAD1, CMT2D, HEXB, IGHMBP2, SMUBP2, CATF1, SMARD1
Muscular dystrophy	muscle		FKRP, MDC1C, LGMD2I, LAMA2, LAMM, LARGE, KIAA0609, MDC1D, FCMD, TTID, MYOT, CAPN3, CANP3, DYSF, LGMD2B, SGCG, LGMD2C, DMDA1, SCG3, SGCA, ADL, DAG2, LGMD2D, DMDA2, SGCB, LGMD2E, SGCD, SGD, LGMD2F, CMD1L, TCAP, LGMD2G, CMD1N, TRIM32, HT2A, LGMD2H, FKRP, MDC1C, LGMD2I, TTN, CMD1G, TMD, LGMD2J, POMT1, CAV3, LGMD1C, SEPN1, SELN, RSMD1, PLEC1, PLTN, EBS1
Myotonic dystrophy (Type 1 and Type 2)	Muscles	Eyes, heart, endocrine	CNBP (Type 2) and DMPK (Type 1)
Neoplasia			PTEN; ATM; ATR; EGFR; ERBB2; ERBB3; ERBB4; Notch1; Notch2; Notch3; Notch4; AKT; AKT2; AKT3; HIF; HIF1a; HIF3a; Met; HRG; Bcl2; PPAR alpha; PPAR gamma; WT1 (Wilms Tumor); FGF Receptor Family members (5 members: 1, 2, 3, 4, 5); CDKN2a; APC; RB

			(retinoblastoma); MEN1; VHL; BRCA1; BRCA2; AR (Androgen Receptor); TSG101; IGF; IGF Receptor; Igf1 (4 variants); Igf2 (3 variants); Igf 1 Receptor; Igf 2 Receptor; Bax; Bcl2; caspases family (9 members: 1, 2, 3, 4, 6, 7, 8, 9, 12); Kras; Apc
Neurofibromatosis (NF) (NF1, formerly Recklinghausen's NF, and NF2)	brain, spinal cord, nerves, and skin		NF1, NF2
Niemann-Pick Lipidosis (Types A, B, and C)	Lysosomal Storage Disease	Various- where sphingomyelin accumulates, particularly spleen, liver, blood, CNS	Types A and B: SMPD1; Type C: NPC1 or NPC2
Noonan Syndrome	Various – musculoskeletal, heart, eyes, reproductive organs, blood		PTPN11, SOS1, RAF1 and KRAS
Norrie Disease or X-linked Familial Exudative Vitreoretinopathy	eye		NDP
North Carolina Macular Dystrophy	eye		MCDR1
Osteogenesis imperfecta (OI) (Types I, II, III, IV, V, VI, VII)	bones, musculoskeletal		COL1A1, COL1A2, CRTAP, P3H
Osteopetrosis	bones		LRP5, BMND1, LRP7, LR3, OPPG, VBCH2, CLCN7, CLC7, OPTA2, OSTM1, GL, TCIRG1, TIRC7, OC116, OPTB1
Patau's Syndrome (Trisomy 13)	Brain, heart, skeletal system		Additional copy of chromosome 13
Parkinson's disease (PD)	Brain, nervous system		SNCA (PARK1), UCHL1 (PARK 5), and LRRK2 (PARK8), (PARK3), PARK2, PARK4, PARK7 (PARK7), PINK1 (PARK6); x-Synuclein, DJ-1, Parkin, NR4A2, NURR1, NOT, TINUR, SNCAIP, TBP, SCA17, NCAP, PRKN, PDJ, DBH, NDUFV2
Pattern Dystrophy of the RPE	eye		RDS/peripherin
Phenylketonuria (PKU)	Metabolism disorder	Various due to build-up of phenylalanine, phenyl ketones in tissues and CNS	PAH, PKU1, QDPR, DHPR, PTS
Polycystic kidney and hepatic disease	Kidney, liver		FCYT, PKHD1, ARPKD, PKD1, PKD2, PKD4, PKDTS, PRKCSH, G19P1, PCLD, SEC63
Pompe's Disease	Glycogen storage disease	Various – heart, liver, spleen	GAA
Porphyria (actually refers to a group of different diseases all	Various- wherever heme		ALAD, ALAS2, CPOX, FECH, HMBS, PPOX, UROD, or UROS

having a specific heme production process abnormality)	precursors accumulate		
posterior polymorphous corneal dystrophy	eyes		TCF4; COL8A2
Primary Hyperoxaluria (e.g. type 1)	Various – eyes, heart, kidneys, skeletal system		LDHA (lactate dehydrogenase A) and hydroxyacid oxidase 1 (HAO1)
Primary Open Angle Glaucoma (POAG)	eyes		MYOC
Primary sclerosing cholangitis	Liver, gallbladder		TCF4; COL8A2
Progeria (also called Hutchinson-Gilford progeria syndrome)	All		LMNA
Prader-Willi Syndrome	Musculoskeletal system, brain, reproductive and endocrine system		Deletion of region of short arm of chromosome 15, including UBE3A
Prostate Cancer	prostate		HOXB13, MSMB, GPRC6A, TP53
Pyruvate Dehydrogenase Deficiency	Brain, nervous system		PDHA1
Kidney/Renal carcinoma	kidney		RLIP76, VEGF
Rett Syndrome	Brain		MECP2, RTT, PPMX, MRX16, MRX79, CDKL5, STK9, MECP2, RTT, PPMX, MRX16, MRX79, x-Synuclein, DJ-1
Retinitis pigmentosa (RP)	eye		ADIPOR1, ABCA4, AGL5, ARHGEF18, ARL2BP, ARL3, ARL6, BEST1, BBS1, BBS2, C2ORF71, C8ORF37, CA4, CERKL, CLRN1, CNGA1, CMGB1, CRB1, CRX, CYP4V2, DHDDS, DHX38, EMC1, EYS, FAM161A, FSCN2, GPR125, GUCA1B, HK1, HPRPF3, HGSNAT, IDH3B, IMPDH1, IMPG2, IFT140, IFT172, KLHL7, KIAA1549, KIZ, LRAT, MAK, MERTK, MVK, NEK2, NUROD1, NR2E3, NRL, OFD1, PDE6A, PDE6B, PDE6G, POMGNT1, PRCD, PROM1, PRPF3, PRPF4, PRPF6, PRPF8, PRPF31, PRPH2, RPB3, RDH12, REEP6, RP39, RGR, RHO, RLBP1, ROM1, RP1, RP1L1, RPY, RP2, RP9, RPE65, RPGR, SAMD11, SAG, SEMA4A, SLC7A14, SNRNP200, SPP2, SPATA7, TRNT1, TOPORS, TTC8, TULP1, USH2A, ZFN408, ZNF513, see also 20120204282
Scheie syndrome (also known as mucopolysaccharidosis type I S(MPS I-S))	Various- liver, spleen, eye, joint, heart, brain, skeletal		IDUA, α -L-iduronidase

Schizophrenia	Brain		Neuregulin1 (Nrg1); Erb4 (receptor for Neuregulin); Complexin1 (Cplx1); Tph1 Tryptophan hydroxylase; Tph2 Tryptophan hydroxylase 2; Neurexin 1; GSK3; GSK3a; GSK3b; 5-HTT (Slc6a4); COMT; DRD (Drd1a); SLC6A3; DAOA; DTNBP1; Dao (Dao1); TCF4; COL8A2
Secretase Related Disorders	Various		APH-1 (alpha and beta); PSEN1; NCSTN; PEN-2; Nos1, Parp1, Nat1, Nat2, CTSB, APP, APH1B, PSEN2, PSENEN, BACE1, ITM2B, CTSD, NOTCH1, TNF, INS, DYT10, ADAM17, APOE, ACE, STN, TP53, IL6, NGFR, IL1B, ACHE, CTNNB1, IGF1, IFNG, NRG1, CASP3, MAPK1, CDH1, APBB1, HMGR, CREB1, PTGS2, HES1, CAT, TGFB1, ENO2, ERBB4, TRAPPC10, MAOB, NGF, MMP12, JAG1, CD40LG, PPARG, FGF2, LRP1, NOTCH4, MAPK8, PREP, NOTCH3, PRNP, CTSG, EGF, REN, CD44, SELP, GHR, ADCYAP1, INSR, GFAP, MMP3, MAPK10, SP1, MYC, CTSE, PPARA, JUN, TIMP1, IL5, IL1A, MMP9, HTR4, HSPG2, KRAS, CYCS, SMG1, IL1R1, PROK1, MAPK3, NTRK1, IL13, MME, TKT, CXCR2, CHRM1, ATXN1, PAWR, NOTCH2, M6PR, CYP46A1, CSNK1D, MAPK14, PRG2, PRKCA, L1 CAM, CD40, NR1H2, JAG2, CTNND1, CMA1, SORT1, DLK1, THEM4, JUP, CD46, CCL11, CAV3, RNASE3, HSPA8, CASP9, CYP3A4, CCR3, TFAP2A, SCP2, CDK4, JOF1A, TCF7L2, B3GALT, MDM2, REL, CASP7, IDE, FANP4, CASK, ADCYAP1R1, ATF4, PDGFA, C21ORF33, SCG5, RMF123, NKFB1, ERBB2, CAV1, MMP7, TGFA, RXRA, STX1A, PSMC4, P2RY2, TNFRSF21, DLG1, NUMBL, SPN, PLSCR1, UBQLN2, UBQLN1, PCSK7, SPON1, SILV, QPCT, HESS, GCC1
Selective IgA Deficiency	Immune system		Type 1: MSH5; Type 2: TNFRSF13B
Severe Combined Immunodeficiency (SCID) and SCID-X1, and ADA-SCID	Immune system		JAK3, JAK1, DCLRE1C, ARTEMIS, SCID1, RAG1, RAG2, ADA, PTPRC, CD45, LCA, IL7R, CD3D, T3D, IL2RG, SCID1X1, SCID1X, IMD4, those identified in US Pat. App. Pub. 20110225664,

			20110091441, 20100229252, 20090271881 and 20090222937;
Sickle cell disease	blood		HBB, BCL11A, BCL11Ae, cis-regulatory elements of the B-globin locus, HBG 1/2 promoter, HBG distal CCAAT box region between -92 and -130 of the HBG Transcription Start Site, those described in WO2015148863, WO 2013/126794, US Pat. Pub. 20110182867
Sly Syndrome (aka MPS VII)			GUSB
Spinocerebellar Ataxias (SCA types 1, 2, 3, 6, 7, 8, 12 and 17)			ATXN1, ATXN2, ATX3
Sorsby Fundus Dystrophy	eye		TIMP3
Stargardt disease	eye		ABCR, ELOVL4, ABCA4, PROM1
Tay-Sachs Disease	Lysosomal Storage disease	Various – CNS, brain, eye	HEX-A
Thalassemia (Alpha, Beta, Delta)	blood		HBA1, HBA2 (Alpha), HBB (Beta), HBB and HBD (delta), LCRB, BCL11A, BCL11Ae, cis-regulatory elements of the B-globin locus, HBG 1/2 promoter, those described in WO2015148860, US Pat. Pub. 20110182867, 2015/148860
Thymic Aplasia (DiGeorge Syndrome; 22q11.2 deletion syndrome)	Immune system, thymus		deletion of 30 to 40 genes in the middle of chromosome 22 at a location known as 22q11.2, including TBX1, DGCR8
Transthyretin amyloidosis (ATTR)	liver		TTR (transthyretin)
trimethylaminuria	Metabolism disease		FMO3
Trinucleotide Repeat Disorders (generally)	Various		HTT; SBMA/SMAX1/AR; FXN/X25 ATX3; ATXN1; ATXN2; DMPK; Atrophin-1 and Atn1 (DRPLA Dx); CBP (Creb-BP - global instability); VLDLR; Atxn7; Atxn10; FEN1, TNRC6A, PABPN1, JPH3, MED15, ATXN1, ATXN3, TBP, CACNA1A, ATXN80S, PPP2R2B, ATXN7, TNRC6B, TNRC6C, CELF3, MAB21L1, MSH2, TMEM185A, SIX5, CNPY3, RAXE, GNB2, RPL14, ATXN8, ISR, TTR, EP400, GIGYF2, OGG1, STC1, CNDP1, C10ORF2, MAML3, DKC1, PAXIP1, CASK, MAPT, SP1, POLG, AFF2, THBS1, TP53, ESR1, CGGBP1, ABT1, KLK3, PRNP, JUN, KCNN3, BAX, FRAXA, KBTBD10, MBNL1, RAD51, NCOA3, ERDA1, TSC1, COMP, GGLC, RRAD, MSH3, DRD2, CD44, CTCF, CCND1, CLSPN, MEF2A, PTPRU, GAPDH, TRIM22,

			WT1, AHR, GPX1, TPMT, NDP, ARX, TYR, EGR1, UNG, NUMBL, FABP2, EN2, CRYGC, SRP14, CRYGB, PDCD1, HOXA1, ATXN2L, PMS2, GLA, CBL, FTH1, IL12RB2, OTX2, HOXA5, POLG2, DLX2, AHRR, MANF, RMEM158, see also 20110016540
Turner's Syndrome (XO)	Various – reproductive organs, and sex characteristics, vasculature		Monosomy X
Tuberous Sclerosis	CNS, heart, kidneys		TSC1, TSC2
Usher syndrome (Types I, II, and III)	Ears, eyes		ABHD12, CDH23, CIB2, CLRN1, DFNB31, GPR98, HARS, MYO7A, PCDH15, USH1C, USH1G, USH2A, USH11A, those described in WO2015134812A1
Velocardiofacial syndrome (aka 22q11.2 deletion syndrome, DiGeorge syndrome, conotruncal anomaly face syndrome (CTAF), autosomal dominant Opitz G/BB syndrome or Cayler cardiofacial syndrome)	Various – skeletal, heart, kidney, immune system, brain		Many genes are deleted, COM, TBX1, and other are associated with symptoms
Von Gierke's Disease (Glycogen Storage Disease type I)	Glycogen Storage disease	Various – liver, kidney	G6PC and SLC37A4
Von Hippel-Lindau Syndrome	Various – cell growth regulation disorder	CNS, Kidney, Eye, visceral organs	VHL
Von Willebrand Disease (Types I, II and III)	blood		VWF
Wilson Disease	Various - Copper Storage Disease	Liver, brains, eyes, other tissues where copper builds up	ATP7B
Wiskott-Aldrich Syndrome	Immune System		WAS
Xeroderma Pigmentosum	Skin	Nervous system	POLH
XXX Syndrome	Endocrine, brain		X chromosome trisomy

[0839] In one embodiment, the compositions, systems, or components thereof can be used to treat or prevent a disease in a subject by modifying one or more genes associated with one or more cellular functions, such as any one or more of those in Table 7. In one embodiment, the disease is a genetic disease or disorder. In some of embodiments, the composition, system, or component thereof can modify one or more genes or polynucleotides associated with one or more genetic diseases such as any set forth in Table 7.

Table 7. Exemplary Genes controlling Cellular Functions

CELLULAR FUNCTION	GENES
PI3K/AKT Signaling	PRKCE; ITGAM; ITGA5; IRAK1; PRKAA2; EIF2AK2; PTEN; EIF4E; PRKCZ; GRK6; MAPK1; TSC1; PLK1; AKT2; IKBKB; PIK3CA; CDK8; CDKN1B; NFKB2; BCL2; PIK3CB; PPP2R1A; MAPK8; BCL2L1; MAPK3; TSC2; ITGA1; KRAS; EIF4EBP1; RELA; PRKCD; NOS3; PRKAA1; MAPK9; CDK2; PPP2CA; PIM1; ITGB7; YWHAZ; ILK; TP53; RAF1; IKBKG; RELB; DYRK1A; CDKN1A; ITGB1; MAP2K2; JAK1; AKT1; JAK2; PIK3R1; CHUK; PDPK1; PPP2R5C; CTNNB1; MAP2K1; NFKB1; PAK3; ITGB3; CCND1; GSK3A; FRAP1; SFN; ITGA2; TTK; CSNK1A1; BRAF; GSK3B; AKT3; FOXO1; SGK; HSP90AA1; RPS6KB1
ERK/MAPK Signaling	PRKCE; ITGAM; ITGA5; HSPB1; IRAK1; PRKAA2; EIF2AK2; RAC1; RAP1A; TLN1; EIF4E; ELK1; GRK6; MAPK1; RAC2; PLK1; AKT2; PIK3CA; CDK8; CREB1; PRKCI; PTK2; FOS; RPS6KA4; PIK3CB; PPP2R1A; PIK3C3; MAPK8; MAPK3; ITGA1; ETS1; KRAS; MYCN; EIF4EBP1; PPARG; PRKCD; PRKAA1; MAPK9; SRC; CDK2; PPP2CA; PIM1; PIK3C2A; ITGB7; YWHAZ; PPP1CC; KSR1; PXN; RAF1; FYN; DYRK1A; ITGB1; MAP2K2; PAK4; PIK3R1; STAT3; PPP2R5C; MAP2K1; PAK3; ITGB3; ESR1; ITGA2; MYC; TTK; CSNK1A1; CRKL; BRAF; ATF4; PRKCA; SRF; STAT1; SGK
Glucocorticoid Receptor Signaling	RAC1; TAF4B; EP300; SMAD2; TRAF6; PCAF; ELK1; MAPK1; SMAD3; AKT2; IKBKB; NCOR2; UBE2I; PIK3CA; CREB1; FOS; HSPA5; NFKB2; BCL2; MAP3K14; STAT5B; PIK3CB; PIK3C3; MAPK8; BCL2L1; MAPK3; TSC22D3; MAPK10; NRIP1; KRAS; MAPK13; RELA; STAT5A; MAPK9; NOS2A; PBX1; NR3C1; PIK3C2A; CDKN1C; TRAF2; SERPINE1; NCOA3; MAPK14; TNF; RAF1; IKBKG; MAP3K7; CREBBP; CDKN1A; MAP2K2; JAK1; IL8; NCOA2; AKT1; JAK2; PIK3R1; CHUK; STAT3; MAP2K1; NFKB1; TGFBR1; ESR1; SMAD4; CEBPB; JUN; AR; AKT3; CCL2; MMP1; STAT1; IL6; HSP90AA1

Axonal Guidance Signaling	PRKCE; ITGAM; ROCK1; ITGA5; CXCR4; ADAM12; IGF1; RAC1; RAP1A; EIF4E; PRKCZ; NRP1; NTRK2; ARHGEF7; SMO; ROCK2; MAPK1; PGF; RAC2; PTPN11; GNAS; AKT2; PIK3CA; ERBB2; PRKCI; PTK2; CFL1; GNAQ; PIK3CB; CXCL12; PIK3C3; WNT11; PRKD1; GNB2L1; ABL1; MAPK3; ITGA1; KRAS; RHOA; PRKCD; PIK3C2A; ITGB7; GLI2; PXN; VASP; RAF1; FYN; ITGB1; MAP2K2; PAK4; ADAM17; AKT1; PIK3R1; GLI1; WNT5A; ADAM10; MAP2K1; PAK3; ITGB3; CDC42; VEGFA; ITGA2; EPHA8; CRKL; RND1; GSK3B; AKT3; PRKCA
Ephrin Receptor Signaling Actin Cytoskeleton Signaling	PRKCE; ITGAM; ROCK1; ITGA5; CXCR4; IRAK1; PRKAA2; EIF2AK2; RAC1; RAP1A; GRK6; ROCK2; MAPK1; PGF; RAC2; PTPN11; GNAS; PLK1; AKT2; DOK1; CDK8; CREB1; PTK2; CFL1; GNAQ; MAP3K14; CXCL12; MAPK8; GNB2L1; ABL1; MAPK3; ITGA1; KRAS; RHOA; PRKCD; PRKAA1; MAPK9; SRC; CDK2; PIM1; ITGB7; PXN; RAF1; FYN; DYRK1A; ITGB1; MAP2K2; PAK4; AKT1; JAK2; STAT3; ADAM10; MAP2K1; PAK3; ITGB3; CDC42; VEGFA; ITGA2; EPHA8; TTK; CSNK1A1; CRKL; BRAF; PTPN13; ATF4; AKT3; SGK
	ACTN4; PRKCE; ITGAM; ROCK1; ITGA5; IRAK1; PRKAA2; EIF2AK2; RAC1; INS; ARHGEF7; GRK6; ROCK2; MAPK1; RAC2; PLK1; AKT2; PIK3CA; CDK8; PTK2; CFL1; PIK3CB; MYH9; DIAPH1; PIK3C3; MAPK8; F2R; MAPK3; SLC9A1; ITGA1; KRAS; RHOA; PRKCD; PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A; ITGB7; PPP1CC; PXN; VIL2; RAF1; GSN; DYRK1A; ITGB1; MAP2K2; PAK4; PIP5K1A; PIK3R1; MAP2K1; PAK3; ITGB3; CDC42; APC; ITGA2; TTK; CSNK1A1; CRKL; BRAF; VAV3; SGK
Huntington's Disease Signaling	PRKCE; IGF1; EP300; RCOR1; PRKCZ; HDAC4; TGM2; MAPK1; CAPNS1; AKT2; EGFR; NCOR2; SP1; CAPN2; PIK3CA; HDAC5; CREB1; PRKCI; HSPA5; REST; GNAQ; PIK3CB; PIK3C3; MAPK8; IGF1R; PRKD1; GNB2L1; BCL2L1; CAPN1; MAPK3; CASP8; HDAC2; HDAC7A; PRKCD; HDAC11; MAPK9; HDAC9; PIK3C2A; HDAC3; TP53; CASP9; CREBBP; AKT1; PIK3R1; PDPK1; CASP1; APAF1; FRAP1; CASP2; JUN; BAX; ATF4; AKT3; PRKCA; CLTC; SGK; HDAC6; CASP3

Apoptosis Signaling	PRKCE; ROCK1; BID; IRAK1; PRKAA2; EIF2AK2; BAK1; BIRC4; GRK6; MAPK1; CAPNS1; PLK1; AKT2; IKBKB; CAPN2; CDK8; FAS; NFKB2; BCL2; MAP3K14; MAPK8; BCL2L1; CAPN1; MAPK3; CASP8; KRAS; RELA; PRKCD; PRKAA1; MAPK9; CDK2; PIM1; TP53; TNF; RAF1; IKBKG; RELB; CASP9; DYRK1A; MAP2K2; CHUK; APAF1; MAP2K1; NFKB1; PAK3; LMNA; CASP2; BIRC2; TTK; CSNK1A1; BRAF; BAX; PRKCA; SGK; CASP3; BIRC3; PARP1
B Cell Receptor Signaling	RAC1; PTEN; LYN; ELK1; MAPK1; RAC2; PTPN11; AKT2; IKBKB; PIK3CA; CREB1; SYK; NFKB2; CAMK2A; MAP3K14; PIK3CB; PIK3C3; MAPK8; BCL2L1; ABL1; MAPK3; ETS1; KRAS; MAPK13; RELA; PTPN6; MAPK9; EGR1; PIK3C2A; BTK; MAPK14; RAF1; IKBKG; RELB; MAP3K7; MAP2K2; AKT1; PIK3R1; CHUK; MAP2K1; NFKB1; CDC42; GSK3A; FRAP1; BCL6; BCL10; JUN; GSK3B; ATF4; AKT3; VAV3; RPS6KB1
Leukocyte Extravasation Signaling	ACTN4; CD44; PRKCE; ITGAM; ROCK1; CXCR4; CYBA; RAC1; RAP1A; PRKCZ; ROCK2; RAC2; PTPN11; MMP14; PIK3CA; PRKCI; PTK2; PIK3CB; CXCL12; PIK3C3; MAPK8; PRKD1; ABL1; MAPK10; CYBB; MAPK13; RHOA; PRKCD; MAPK9; SRC; PIK3C2A; BTK; MAPK14; NOX1; PXN; VIL2; VASP; ITGB1; MAP2K2; CTNND1; PIK3R1; CTNNB1; CLDN1; CDC42; F11R; ITK; CRKL; VAV3; CTTN; PRKCA; MMP1; MMP9
Integrin Signaling	ACTN4; ITGAM; ROCK1; ITGA5; RAC1; PTEN; RAP1A; TLN1; ARHGEF7; MAPK1; RAC2; CAPNS1; AKT2; CAPN2; PIK3CA; PTK2; PIK3CB; PIK3C3; MAPK8; CAV1; CAPN1; ABL1; MAPK3; ITGA1; KRAS; RHOA; SRC; PIK3C2A; ITGB7; PPP1CC; ILK; PXN; VASP; RAF1; FYN; ITGB1; MAP2K2; PAK4; AKT1; PIK3R1; TNK2; MAP2K1; PAK3; ITGB3; CDC42; RND3; ITGA2; CRKL; BRAF; GSK3B; AKT3
Acute Phase Response Signaling	IRAK1; SOD2; MYD88; TRAF6; ELK1; MAPK1; PTPN11; AKT2; IKBKB; PIK3CA; FOS; NFKB2; MAP3K14; PIK3CB; MAPK8; RIPK1; MAPK3; IL6ST; KRAS; MAPK13; IL6R; RELA; SOCS1; MAPK9; FTL; NR3C1; TRAF2; SERPINE1; MAPK14; TNF; RAF1; PDK1; IKBKG; RELB; MAP3K7; MAP2K2; AKT1; JAK2; PIK3R1; CHUK; STAT3; MAP2K1; NFKB1; FRAP1; CEBPB; JUN; AKT3; IL1R1; IL6

PTEN Signaling	ITGAM; ITGA5; RAC1; PTEN; PRKCZ; BCL2L11; MAPK1; RAC2; AKT2; EGFR; IKBKB; CBL; PIK3CA; CDKN1B; PTK2; NFKB2; BCL2; PIK3CB; BCL2L1; MAPK3; ITGA1; KRAS; ITGB7; ILK; PDGFRB; INSR; RAF1; IKBKG; CASP9; CDKN1A; ITGB1; MAP2K2; AKT1; PIK3R1; CHUK; PDGFRA; PDPK1; MAP2K1; NFKB1; ITGB3; CDC42; CCND1; GSK3A; ITGA2; GSK3B; AKT3; FOXO1; CASP3; RPS6KB1
p53 Signaling Aryl Hydrocarbon Receptor Signaling	PTEN; EP300; BBC3; PCAF; FASN; BRCA1; GADD45A; BIRC5; AKT2; PIK3CA; CHEK1; TP53INP1; BCL2; PIK3CB; PIK3C3; MAPK8; THBS1; ATR; BCL2L1; E2F1; PMAIP1; CHEK2; TNFRSF10B; TP73; RB1; HDAC9; CDK2; PIK3C2A; MAPK14; TP53; LRDD; CDKN1A; HIPK2; AKT1; PIK3R1; RRM2B; APAF1; CTNNB1; SIRT1; CCND1; PRKDC; ATM; SFN; CDKN2A; JUN; SNAI2; GSK3B; BAX; AKT3
	HSPB1; EP300; FASN; TGM2; RXRA; MAPK1; NQO1; NCOR2; SP1; ARNT; CDKN1B; FOS; CHEK1; SMARCA4; NFKB2; MAPK8; ALDH1A1; ATR; E2F1; MAPK3; NRIP1; CHEK2; RELA; TP73; GSTP1; RB1; SRC; CDK2; AHR; NFE2L2; NCOA3; TP53; TNF; CDKN1A; NCOA2; APAF1; NFKB1; CCND1; ATM; ESR1; CDKN2A; MYC; JUN; ESR2; BAX; IL6; CYP1B1; HSP90AA1
Xenobiotic Metabolism Signaling	PRKCE; EP300; PRKCZ; RXRA; MAPK1; NQO1; NCOR2; PIK3CA; ARNT; PRKCI; NFKB2; CAMK2A; PIK3CB; PPP2R1A; PIK3C3; MAPK8; PRKD1; ALDH1A1; MAPK3; NRIP1; KRAS; MAPK13; PRKCD; GSTP1; MAPK9; NOS2A; ABCB1; AHR; PPP2CA; FTL; NFE2L2; PIK3C2A; PPARGC1A; MAPK14; TNF; RAF1; CREBBP; MAP2K2; PIK3R1; PPP2R5C; MAP2K1; NFKB1; KEAP1; PRKCA; EIF2AK3; IL6; CYP1B1; HSP90AA1
SAPK/JNK Signaling	PRKCE; IRAK1; PRKAA2; EIF2AK2; RAC1; ELK1; GRK6; MAPK1; GADD45A; RAC2; PLK1; AKT2; PIK3CA; FADD; CDK8; PIK3CB; PIK3C3; MAPK8; RIPK1; GNB2L1; IRS1; MAPK3; MAPK10; DAXX; KRAS; PRKCD; PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A; TRAF2; TP53; LCK; MAP3K7; DYRK1A; MAP2K2; PIK3R1; MAP2K1; PAK3; CDC42; JUN; TTK; CSNK1A1; CRKL; BRAF; SGK

PPAr/RXR Signaling	PRKAA2; EP300; INS; SMAD2; TRAF6; PPARA; FASN; RXRA; MAPK1; SMAD3; GNAS; IKBKB; NCOR2; ABCA1; GNAQ; NFKB2; MAP3K14; STAT5B; MAPK8; IRS1; MAPK3; KRAS; RELA; PRKAA1; PPARGC1A; NCOA3; MAPK14; INSR; RAF1; IKBKG; RELB; MAP3K7; CREBBP; MAP2K2; JAK2; CHUK; MAP2K1; NFKB1; TGFB1; SMAD4; JUN; IL1R1; PRKCA; IL6; HSP90AA1; ADIPOQ
NF-KB Signaling	IRAK1; EIF2AK2; EP300; INS; MYD88; PRKCZ; TRAF6; TBK1; AKT2; EGFR; IKBKB; PIK3CA; BTRC; NFKB2; MAP3K14; PIK3CB; PIK3C3; MAPK8; RIPK1; HDAC2; KRAS; RELA; PIK3C2A; TRAF2; TLR4; PDGFRB; TNF; INSR; LCK; IKBKG; RELB; MAP3K7; CREBBP; AKT1; PIK3R1; CHUK; PDGFRA; NFKB1; TLR2; BCL10; GSK3B; AKT3; TNFAIP3; IL1R1
Neuregulin Signaling Wnt & Beta catenin Signaling	ERBB4; PRKCE; ITGAM; ITGA5; PTEN; PRKCZ; ELK1; MAPK1; PTPN11; AKT2; EGFR; ERBB2; PRKCI; CDKN1B; STAT5B; PRKD1; MAPK3; ITGA1; KRAS; PRKCD; STAT5A; SRC; ITGB7; RAF1; ITGB1; MAP2K2; ADAM17; AKT1; PIK3R1; PDPK1; MAP2K1; ITGB3; EREG; FRAP1; PSEN1; ITGA2; MYC; NRG1; CRKL; AKT3; PRKCA; HSP90AA1; RPS6KB1
	CD44; EP300; LRP6; DVL3; CSNK1E; GJA1; SMO; AKT2; PIN1; CDH1; BTRC; GNAQ; MARK2; PPP2R1A; WNT11; SRC; DKK1; PPP2CA; SOX6; SFRP2; ILK; LEF1; SOX9; TP53; MAP3K7; CREBBP; TCF7L2; AKT1; PPP2R5C; WNT5A; LRP5; CTNNB1; TGFB1; CCND1; GSK3A; DVL1; APC; CDKN2A; MYC; CSNK1A1; GSK3B; AKT3; SOX2
Insulin Receptor Signaling	PTEN; INS; EIF4E; PTPN11; PRKCZ; MAPK1; TSC1; PTPN11; AKT2; CBL; PIK3CA; PRKCI; PIK3CB; PIK3C3; MAPK8; IRS1; MAPK3; TSC2; KRAS; EIF4EBP1; SLC2A4; PIK3C2A; PPP1CC; INSR; RAF1; FYN; MAP2K2; JAK1; AKT1; JAK2; PIK3R1; PDPK1; MAP2K1; GSK3A; FRAP1; CRKL; GSK3B; AKT3; FOXO1; SGK; RPS6KB1
IL-6 Signaling	HSPB1; TRAF6; MAPKAPK2; ELK1; MAPK1; PTPN11; IKBKB; FOS; NFKB2; MAP3K14; MAPK8; MAPK3; MAPK10; IL6ST; KRAS; MAPK13; IL6R; RELA; SOCS1; MAPK9; ABCB1; TRAF2; MAPK14; TNF; RAF1; IKBKG; RELB; MAP3K7; MAP2K2; IL8; JAK2; CHUK; STAT3; MAP2K1; NFKB1; CEBPB; JUN; IL1R1; SRF; IL6

Hepatic Cholestasis	PRKCE; IRAK1; INS; MYD88; PRKCZ; TRAF6; PPARA; RXRA; IKBKB; PRKCI; NFKB2; MAP3K14; MAPK8; PRKD1; MAPK10; RELA; PRKCD; MAPK9; ABCB1; TRAF2; TLR4; TNF; INSR; IKBKG; RELB; MAP3K7; IL8; CHUK; NR1H2; TJP2; NFKB1; ESR1; SREBF1; FGFR4; JUN; IL1R1; PRKCA; IL6
IGF-1 Signaling	IGF1; PRKCZ; ELK1; MAPK1; PTPN11; NEDD4; AKT2; PIK3CA; PRKCI; PTK2; FOS; PIK3CB; PIK3C3; MAPK8; IGF1R; IRS1; MAPK3; IGFBP7; KRAS; PIK3C2A; YWHAZ; PXN; RAF1; CASP9; MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1; IGFBP2; SFN; JUN; CYR61; AKT3; FOXO1; SRF; CTGF; RPS6KB1
NRF2-mediated Oxidative Stress Response	PRKCE; EP300; SOD2; PRKCZ; MAPK1; SQSTM1; NQO1; PIK3CA; PRKCI; FOS; PIK3CB; PIK3C3; MAPK8; PRKD1; MAPK3; KRAS; PRKCD; GSTP1; MAPK9; FTL; NFE2L2; PIK3C2A; MAPK14; RAF1; MAP3K7; CREBBP; MAP2K2; AKT1; PIK3R1; MAP2K1; PPIB; JUN; KEAP1; GSK3B; ATF4; PRKCA; EIF2AK3; HSP90AA1
Hepatic Fibrosis/Hepatic Stellate Cell Activation	EDN1; IGF1; KDR; FLT1; SMAD2; FGFR1; MET; PGF; SMAD3; EGFR; FAS; CSF1; NFKB2; BCL2; MYH9; IGF1R; IL6R; RELA; TLR4; PDGFRB; TNF; RELB; IL8; PDGFRA; NFKB1; TGFB1; SMAD4; VEGFA; BAX; IL1R1; CCL2; HGF; MMP1; STAT1; IL6; CTGF; MMP9
PPAR Signaling	EP300; INS; TRAF6; PPARA; RXRA; MAPK1; IKBKB; NCOR2; FOS; NFKB2; MAP3K14; STAT5B; MAPK3; NRIP1; KRAS; PPARG; RELA; STAT5A; TRAF2; PPARGC1A; PDGFRB; TNF; INSR; RAF1; IKBKG; RELB; MAP3K7; CREBBP; MAP2K2; CHUK; PDGFRA; MAP2K1; NFKB1; JUN; IL1R1; HSP90AA1
Fc Epsilon RI Signaling	PRKCE; RAC1; PRKCZ; LYN; MAPK1; RAC2; PTPN11; AKT2; PIK3CA; SYK; PRKCI; PIK3CB; PIK3C3; MAPK8; PRKD1; MAPK3; MAPK10; KRAS; MAPK13; PRKCD; MAPK9; PIK3C2A; BTK; MAPK14; TNF; RAF1; FYN; MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1; AKT3; VAV3; PRKCA
G-Protein Coupled Receptor Signaling	PRKCE; RAP1A; RGS16; MAPK1; GNAS; AKT2; IKBKB; PIK3CA; CREB1; GNAQ; NFKB2; CAMK2A; PIK3CB; PIK3C3; MAPK3; KRAS; RELA; SRC; PIK3C2A; RAF1; IKBKG; RELB; FYN; MAP2K2; AKT1; PIK3R1; CHUK; PDPK1; STAT3; MAP2K1; NFKB1; BRAF; ATF4; AKT3; PRKCA

Inositol Phosphate Metabolism	PRKCE; IRAK1; PRKAA2; EIF2AK2; PTEN; GRK6; MAPK1; PLK1; AKT2; PIK3CA; CDK8; PIK3CB; PIK3C3; MAPK8; MAPK3; PRKCD; PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A; DYRK1A; MAP2K2; PIP5K1A; PIK3R1; MAP2K1; PAK3; ATM; TTK; CSNK1A1; BRAF; SGK
PDGF Signaling	EIF2AK2; ELK1; ABL2; MAPK1; PIK3CA; FOS; PIK3CB; PIK3C3; MAPK8; CAV1; ABL1; MAPK3; KRAS; SRC; PIK3C2A; PDGFRB; RAF1; MAP2K2; JAK1; JAK2; PIK3R1; PDGFRA; STAT3; SPHK1; MAP2K1; MYC; JUN; CRKL; PRKCA; SRF; STAT1; SPHK2
VEGF Signaling	ACTN4; ROCK1; KDR; FLT1; ROCK2; MAPK1; PGF; AKT2; PIK3CA; ARNT; PTK2; BCL2; PIK3CB; PIK3C3; BCL2L1; MAPK3; KRAS; HIF1A; NOS3; PIK3C2A; PXN; RAF1; MAP2K2; ELAVL1; AKT1; PIK3R1; MAP2K1; SFN; VEGFA; AKT3; FOXO1; PRKCA
Natural Killer Cell Signaling	PRKCE; RAC1; PRKCZ; MAPK1; RAC2; PTPN11; KIR2DL3; AKT2; PIK3CA; SYK; PRKCI; PIK3CB; PIK3C3; PRKD1; MAPK3; KRAS; PRKCD; PTPN6; PIK3C2A; LCK; RAF1; FYN; MAP2K2; PAK4; AKT1; PIK3R1; MAP2K1; PAK3; AKT3; VAV3; PRKCA
Cell Cycle: G1/S Checkpoint Regulation	HDAC4; SMAD3; SUV39H1; HDAC5; CDKN1B; BTRC; ATR; ABL1; E2F1; HDAC2; HDAC7A; RB1; HDAC11; HDAC9; CDK2; E2F2; HDAC3; TP53; CDKN1A; CCND1; E2F4; ATM; RBL2; SMAD4; CDKN2A; MYC; NRG1; GSK3B; RBL1; HDAC6
T Cell Receptor Signaling	RAC1; ELK1; MAPK1; IKBKB; CBL; PIK3CA; FOS; NFKB2; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS; RELA; PIK3C2A; BTK; LCK; RAF1; IKBKG; RELB; FYN; MAP2K2; PIK3R1; CHUK; MAP2K1; NFKB1; ITK; BCL10; JUN; VAV3
Death Receptor Signaling	CRADD; HSPB1; BID; BIRC4; TBK1; IKBKB; FADD; FAS; NFKB2; BCL2; MAP3K14; MAPK8; RIPK1; CASP8; DAXX; TNFRSF10B; RELA; TRAF2; TNF; IKBKG; RELB; CASP9; CHUK; APAF1; NFKB1; CASP2; BIRC2; CASP3; BIRC3
FGF Signaling	RAC1; FGFR1; MET; MAPKAPK2; MAPK1; PTPN11; AKT2; PIK3CA; CREB1; PIK3CB; PIK3C3; MAPK8; MAPK3; MAPK13; PTPN6; PIK3C2A; MAPK14; RAF1; AKT1; PIK3R1; STAT3; MAP2K1; FGFR4; CRKL; ATF4; AKT3; PRKCA; HGF

GM-CSF Signaling	LYN; ELK1; MAPK1; PTPN11; AKT2; PIK3CA; CAMK2A; STAT5B; PIK3CB; PIK3C3; GNB2L1; BCL2L1; MAPK3; ETS1; KRAS; RUNX1; PIM1; PIK3C2A; RAF1; MAP2K2; AKT1; JAK2; PIK3R1; STAT3; MAP2K1; CCND1; AKT3; STAT1
Amyotrophic Lateral Sclerosis Signaling	BID; IGF1; RAC1; BIRC4; PGF; CAPNS1; CAPN2; PIK3CA; BCL2; PIK3CB; PIK3C3; BCL2L1; CAPN1; PIK3C2A; TP53; CASP9; PIK3R1; RAB5A; CASP1; APAF1; VEGFA; BIRC2; BAX; AKT3; CASP3; BIRC3
JAK/Stat Signaling	PTPN1; MAPK1; PTPN11; AKT2; PIK3CA; STAT5B; PIK3CB; PIK3C3; MAPK3; KRAS; SOCS1; STAT5A; PTPN6; PIK3C2A; RAF1; CDKN1A; MAP2K2; JAK1; AKT1; JAK2; PIK3R1; STAT3; MAP2K1; FRAP1; AKT3; STAT1
Nicotinate and Nicotinamide Metabolism	PRKCE; IRAK1; PRKAA2; EIF2AK2; GRK6; MAPK1; PLK1; AKT2; CDK8; MAPK8; MAPK3; PRKCD; PRKAA1; PBEF1; MAPK9; CDK2; PIM1; DYRK1A; MAP2K2; MAP2K1; PAK3; NT5E; TTK; CSNK1A1; BRAF; SGK
Chemokine Signaling	CXCR4; ROCK2; MAPK1; PTK2; FOS; CFL1; GNAQ; CAMK2A; CXCL12; MAPK8; MAPK3; KRAS; MAPK13; RHOA; CCR3; SRC; PPP1CC; MAPK14; NOX1; RAF1; MAP2K2; MAP2K1; JUN; CCL2; PRKCA
IL-2 Signaling	ELK1; MAPK1; PTPN11; AKT2; PIK3CA; SYK; FOS; STAT5B; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS; SOCS1; STAT5A; PIK3C2A; LCK; RAF1; MAP2K2; JAK1; AKT1; PIK3R1; MAP2K1; JUN; AKT3
Synaptic Long Term Depression	PRKCE; IGF1; PRKCZ; PRDX6; LYN; MAPK1; GNAS; PRKCI; GNAQ; PPP2R1A; IGF1R; PRKD1; MAPK3; KRAS; GRN; PRKCD; NOS3; NOS2A; PPP2CA; YWHAZ; RAF1; MAP2K2; PPP2R5C; MAP2K1; PRKCA
Estrogen Receptor Signaling	TAF4B; EP300; CARM1; PCAF; MAPK1; NCOR2; SMARCA4; MAPK3; NRIP1; KRAS; SRC; NR3C1; HDAC3; PPARGC1A; RBM9; NCOA3; RAF1; CREBBP; MAP2K2; NCOA2; MAP2K1; PRKDC; ESR1; ESR2
Protein Ubiquitination Pathway	TRAF6; SMURF1; BIRC4; BRCA1; UCHL1; NEDD4; CBL; UBE2I; BTRC; HSPA5; USP7; USP10; FBXW7; USP9X; STUB1; USP22; B2M; BIRC2; PARK2; USP8; USP1; VHL; HSP90AA1; BIRC3
IL-10 Signaling	TRAF6; CCR1; ELK1; IKBKB; SP1; FOS; NFKB2; MAP3K14; MAPK8; MAPK13; RELA; MAPK14; TNF; IKBKG; RELB; MAP3K7; JAK1; CHUK; STAT3; NFKB1; JUN; IL1R1; IL6

VDR/RXR Activation	PRKCE; EP300; PRKCZ; RXRA; GADD45A; HES1; NCOR2; SP1; PRKCI; CDKN1B; PRKD1; PRKCD; RUNX2; KLF4; YY1; NCOA3; CDKN1A; NCOA2; SPP1; LRP5; CEBPB; FOXO1; PRKCA
TGF-beta Signaling	EP300; SMAD2; SMURF1; MAPK1; SMAD3; SMAD1; FOS; MAPK8; MAPK3; KRAS; MAPK9; RUNX2; SERPINE1; RAF1; MAP3K7; CREBBP; MAP2K2; MAP2K1; TGFB1; SMAD4; JUN; SMAD5
Toll-like Receptor Signaling	IRAK1; EIF2AK2; MYD88; TRAF6; PPARA; ELK1; IKKB; FOS; NFKB2; MAP3K14; MAPK8; MAPK13; RELA; TLR4; MAPK14; IKBK; RELB; MAP3K7; CHUK; NFKB1; TLR2; JUN
p38 MAPK Signaling	HSPB1; IRAK1; TRAF6; MAPKAPK2; ELK1; FADD; FAS; CREB1; DDIT3; RPS6KA4; DAXX; MAPK13; TRAF2; MAPK14; TNF; MAP3K7; TGFB1; MYC; ATF4; IL1R1; SRF; STAT1
Neurotrophin/TRK Signaling	NTRK2; MAPK1; PTPN11; PIK3CA; CREB1; FOS; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS; PIK3C2A; RAF1; MAP2K2; AKT1; PIK3R1; PDK1; MAP2K1; CDC42; JUN; ATF4
FXR/RXR Activation	INS; PPARA; FASN; RXRA; AKT2; SDC1; MAPK8; APOB; MAPK10; PPARG; MTP; MAPK9; PPARGC1A; TNF; CREBBP; AKT1; SREBF1; FGFR4; AKT3; FOXO1
Synaptic Long Term Potentiation	PRKCE; RAP1A; EP300; PRKCZ; MAPK1; CREB1; PRKCI; GNAQ; CAMK2A; PRKD1; MAPK3; KRAS; PRKCD; PPP1CC; RAF1; CREBBP; MAP2K2; MAP2K1; ATF4; PRKCA
Calcium Signaling	RAP1A; EP300; HDAC4; MAPK1; HDAC5; CREB1; CAMK2A; MYH9; MAPK3; HDAC2; HDAC7A; HDAC11; HDAC9; HDAC3; CREBBP; CALR; CAMKK2; ATF4; HDAC6
EGF Signaling	ELK1; MAPK1; EGFR; PIK3CA; FOS; PIK3CB; PIK3C3; MAPK8; MAPK3; PIK3C2A; RAF1; JAK1; PIK3R1; STAT3; MAP2K1; JUN; PRKCA; SRF; STAT1
Hypoxia Signaling in the Cardiovascular System	EDN1; PTEN; EP300; NQO1; UBE2I; CREB1; ARNT; HIF1A; SLC2A4; NOS3; TP53; LDHA; AKT1; ATM; VEGFA; JUN; ATF4; VHL; HSP90AA1
LPS/IL-1 Mediated Inhibition of RXR Function	IRAK1; MYD88; TRAF6; PPARA; RXRA; ABCA1; MAPK8; ALDH1A1; GSTP1; MAPK9; ABCB1; TRAF2; TLR4; TNF; MAP3K7; NR1H2; SREBF1; JUN; IL1R1
LXR/RXR Activation	FASN; RXRA; NCOR2; ABCA1; NFKB2; IRF3; RELA; NOS2A; TLR4; TNF; RELB; LDLR; NR1H2; NFKB1; SREBF1; IL1R1; CCL2; IL6; MMP9

Amyloid Processing	PRKCE; CSNK1E; MAPK1; CAPNS1; AKT2; CAPN2; CAPN1; MAPK3; MAPK13; MAPT; MAPK14; AKT1; PSEN1; CSNK1A1; GSK3B; AKT3; APP
IL-4 Signaling	AKT2; PIK3CA; PIK3CB; PIK3C3; IRS1; KRAS; SOCS1; PTPN6; NR3C1; PIK3C2A; JAK1; AKT1; JAK2; PIK3R1; FRAP1; AKT3; RPS6KB1
Cell Cycle: G2/M DNA Damage Checkpoint Regulation	EP300; PCAF; BRCA1; GADD45A; PLK1; BTRC; CHEK1; ATR; CHEK2; YWHAZ; TP53; CDKN1A; PRKDC; ATM; SFN; CDKN2A
Nitric Oxide Signaling in the Cardiovascular System	KDR; FLT1; PGF; AKT2; PIK3CA; PIK3CB; PIK3C3; CAV1; PRKCD; NOS3; PIK3C2A; AKT1; PIK3R1; VEGFA; AKT3; HSP90AA1
Purine Metabolism	NME2; SMARCA4; MYH9; RRM2; ADAR; EIF2AK4; PKM2; ENTPD1; RAD51; RRM2B; TJP2; RAD51C; NT5E; POLD1; NME1
cAMP-mediated Signaling	RAP1A; MAPK1; GNAS; CREB1; CAMK2A; MAPK3; SRC; RAF1; MAP2K2; STAT3; MAP2K1; BRAF; ATF4
Mitochondrial Dysfunction Notch Signaling	SOD2; MAPK8; CASP8; MAPK10; MAPK9; CASP9; PARK7; PSEN1; PARK2; APP; CASP3 HES1; JAG1; NUMB; NOTCH4; ADAM17; NOTCH2; PSEN1; NOTCH3; NOTCH1; DLL4
Endoplasmic Reticulum Stress Pathway Pyrimidine Metabolism	HSPA5; MAPK8; XBP1; TRAF2; ATF6; CASP9; ATF4; EIF2AK3; CASP3 NME2; AICDA; RRM2; EIF2AK4; ENTPD1; RRM2B; NT5E; POLD1; NME1
Parkinson's Signaling	UCHL1; MAPK8; MAPK13; MAPK14; CASP9; PARK7; PARK2; CASP3
Cardiac & Beta Adrenergic Signaling	GNAS; GNAQ; PPP2R1A; GNB2L1; PPP2CA; PPP1CC; PPP2R5C
Glycolysis/Gluconeogenesis	HK2; GCK; GPI; ALDH1A1; PKM2; LDHA; HK1
Interferon Signaling	IRF1; SOCS1; JAK1; JAK2; IFITM1; STAT1; IFIT3
Sonic Hedgehog Signaling	ARRB2; SMO; GLI2; DYRK1A; GLI1; GSK3B; DYRK1B
Glycerophospholipid Metabolism	PLD1; GRN; GPAM; YWHAZ; SPHK1; SPHK2
Phospholipid Degradation	PRDX6; PLD1; GRN; YWHAZ; SPHK1; SPHK2
Tryptophan Metabolism	SIAH2; PRMT5; NEDD4; ALDH1A1; CYP1B1; SIAH1
Lysine Degradation	SUV39H1; EHMT2; NSD1; SETD7; PPP2R5C
Nucleotide Excision Repair Pathway	ERCC5; ERCC4; XPA; XPC; ERCC1
Starch and Sucrose Metabolism	UCHL1; HK2; GCK; GPI; HK1
Aminosugars Metabolism	NQO1; HK2; GCK; HK1
Arachidonic Acid Metabolism	PRDX6; GRN; YWHAZ; CYP1B1
Circadian Rhythm Signaling	CSNK1E; CREB1; ATF4; NR1D1

Coagulation System	BDKRB1; F2R; SERPINE1; F3
Dopamine Receptor Signaling	PPP2R1A; PPP2CA; PPP1CC; PPP2R5C
Glutathione Metabolism	IDH2; GSTP1; ANPEP; IDH1
Glycerolipid Metabolism	ALDH1A1; GPAM; SPHK1; SPHK2
Linoleic Acid Metabolism	PRDX6; GRN; YWHAZ; CYP1B1
Methionine Metabolism	DNMT1; DNMT3B; AHCY; DNMT3A
Pyruvate Metabolism	GLO1; ALDH1A1; PKM2; LDHA
Arginine and Proline Metabolism	ALDH1A1; NOS3; NOS2A
Eicosanoid Signaling	PRDX6; GRN; YWHAZ
Fructose and Mannose Metabolism	HK2; GCK; HK1
Galactose Metabolism	HK2; GCK; HK1
Stilbene, Coumarine and Lignin Biosynthesis	PRDX6; PRDX1; TYR
Antigen Presentation Pathway	CALR; B2M
Biosynthesis of Steroids	NQO1; DHCR7
Butanoate Metabolism	ALDH1A1; NLGN1
Citrate Cycle	IDH2; IDH1
Fatty Acid Metabolism	ALDH1A1; CYP1B1
Glycerophospholipid Metabolism	PRDX6; CHKA
Histidine Metabolism	PRMT5; ALDH1A1
Inositol Metabolism	ERO1L; APEX1
Metabolism of Xenobiotics by Cytochrome p450	GSTP1; CYP1B1
Methane Metabolism	PRDX6; PRDX1
Phenylalanine Metabolism	PRDX6; PRDX1
Propanoate Metabolism	ALDH1A1; LDHA
Selenoamino Acid Metabolism	PRMT5; AHCY
Sphingolipid Metabolism	SPHK1; SPHK2
Aminophosphonate Metabolism	PRMT5
Androgen and Estrogen Metabolism	PRMT5
Ascorbate and Aldarate Metabolism	ALDH1A1
Bile Acid Biosynthesis	ALDH1A1
Cysteine Metabolism	LDHA
Fatty Acid Biosynthesis	FASN
Glutamate Receptor Signaling	GNB2L1
NRF2-mediated Oxidative Stress Response	PRDX1

Pentose Phosphate Pathway	GPI
Pentose and Glucuronate Interconversions	UCHL1
Retinol Metabolism	ALDH1A1
Riboflavin Metabolism	TYR
Tyrosine Metabolism	PRMT5, TYR
Ubiquinone Biosynthesis	PRMT5
Valine, Leucine and Isoleucine Degradation	ALDH1A1
Glycine, Serine and Threonine Metabolism	CHKA
Lysine Degradation	ALDH1A1
Pain/Taste	TRPM5; TRPA1
Pain	TRPM7; TRPC5; TRPC6; TRPC1; Cnr1; cnr2; Grk2; Trpa1; Pomc; Cgrp; Crf; Pka; Era; Nr2b; TRPM5; Prkaca; Prkacb; Prkar1a; Prkar2a
Mitochondrial Function	AIF; CytC; SMAC (Diablo); Aifm-1; Aifm-2
Developmental Neurology	BMP-4; Chordin (Chrd); Noggin (Nog); WNT (Wnt2; Wnt2b; Wnt3a; Wnt4; Wnt5a; Wnt6; Wnt7b; Wnt8b; Wnt9a; Wnt9b; Wnt10a; Wnt10b; Wnt16); beta-catenin; Dkk-1; Frizzled related proteins; Otx-2; Gbx2; FGF-8; Reelin; Dab1; unc-86 (Pou4f1 or Brn3a); Numb; Reln

[0840] In an aspect, the invention provides a method of individualized or personalized treatment of a genetic disease in a subject in need of such treatment comprising: (a) introducing one or more mutations *ex vivo* in a tissue, organ or a cell line, or *in vivo* in a transgenic non-human mammal, comprising delivering to cell(s) of the tissue, organ, cell or mammal a composition comprising the particle delivery system or the delivery system or the virus particle of any one of the above embodiment or the cell of any one of the above embodiment, wherein the specific mutations or precise sequence substitutions are or have been correlated to the genetic disease; (b) testing treatment(s) for the genetic disease on the cells to which the vector has been delivered that have the specific mutations or precise sequence substitutions correlated to the genetic disease; and (c) treating the subject based on results from the testing of treatment(s) of step (b).

Infectious Diseases

[0841] In one embodiment, the composition, system,(s) or component(s) thereof can be used to diagnose, prognose, treat, and/or prevent an infectious disease caused by a microorganism, such as bacteria, virus, fungi, parasites, or combinations thereof.

[0842] In one embodiment, the system(s) or component(s) thereof can be capable of targeting specific microorganism within a mixed population. Exemplary methods of such techniques are described in e.g. Gomaa AA, Klumpe HE, Luo ML, Selle K, Barrangou R, Beisel CL. 2014. Programmable removal of bacterial strains by use of genome-targeting composition, systems, mBio 5:e00928-13; Citorik RJ, Mimee M, Lu TK. 2014. Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. Nat Biotechnol 32:1141–1145, the teachings of which can be adapted for use with the compositions, systems, and components thereof described herein.

[0843] In one embodiment, the composition, system,(s) and/or components thereof can be capable of targeting pathogenic and/or drug-resistant microorganisms, such as bacteria, virus, parasites, and fungi. In one embodiment, the composition, system,(s) and/or components thereof can be capable of targeting and modifying one or more polynucleotides in a pathogenic microorganism such that the microorganism is less virulent, killed, inhibited, or is otherwise rendered incapable of causing disease and/or infecting and/or replicating in a host cell.

[0844] In one embodiment, the pathogenic bacteria that can be targeted and/or modified by the composition, system,(s) and/or component(s) thereof described herein include, but are not limited to, those of the genus Actinomyces (e.g. *A. israelii*), Bacillus (e.g. *B. anthracis*, *B. cereus*), Bacteroides (e.g. *B. fragilis*), Bartonella (*B. henselae*, *B. quintana*), Bordetella (*B. pertussis*), Borrelia (e.g. *B. burgdorferi*, *B. garinii*, *B. afzelii*, and *B. recurrentis*), Brucella (e.g. *B. abortus*, *B. canis*, *B. melitensis*, and *B. suis*), Campylobacter (e.g. *C. jejuni*), Chlamydia (e.g. *C. pneumoniae* and *C. trachomatis*), Chlamydophila (e.g. *C. psittaci*), Clostridium (e.g. *C. botulinum*, *C. difficile*, *C. perfringens*, *C. tetani*), Corynebacterium (e.g. *C. diphtheriae*), Enterococcus (e.g. *E. Faecalis*, *E. faecium*), Ehrlichia (*E. canis* and *E. chaffensis*) Escherichia (e.g. *E. coli*), Francisella (e.g. *F. tularensis*), Haemophilus (e.g. *H. influenzae*), Helicobacter (*H. pylori*), Klebsiella (E.g. *K. pneumoniae*), Legionella (e.g. *L. pneumophila*), Leptospira (e.g. *L. interrogans*, *L. santarosai*, *L. weilii*, *L. noguchii*), Listeria (e.g. *L. monocytogenes*), Mycobacterium (e.g. *M. leprae*, *M. tuberculosis*, *M. ulcerans*), Mycoplasma (*M. pneumoniae*), Neisseria (*N. gonorrhoeae* and *N. meningitidis*), Nocardia (e.g. *N. asteroides*), Pseudomonas (*P. aeruginosa*), Rickettsia (*R. rickettsia*), Salmonella (*S. typhi* and *S. typhimurium*), Shigella (*S. sonnei* and *S. dysenteriae*), Staphylococcus (*S. aureus*, *S. epidermidis*, and *S. saprophyticus*), Streptococcus (*S. agalactiae*, *S. pneumoniae*, *S. pyogenes*), Treponema (*T.*

pallidum), Ureaplasma (e.g. *U. urealyticum*), Vibrio (e.g. *V. cholerae*), Yersinia (e.g. *Y. pestis*, *Y. enterocolitica*, and *Y. pseudotuberculosis*).

[0845] In one embodiment, the pathogenic virus that can be targeted and/or modified by the composition, system,(s) and/or component(s) thereof described herein include, but are not limited to, a double-stranded DNA virus, a partly double-stranded DNA virus, a single-stranded DNA virus, a positive single-stranded RNA virus, a negative single-stranded RNA virus, or a double stranded RNA virus. In one embodiment, the pathogenic virus can be from the family Adenoviridae (e.g. Adenovirus), Herpesviridae (e.g. Herpes simplex, type 1, Herpes simplex, type 2, Varicella-zoster virus, Epstein–Barr virus, Human cytomegalovirus, Human herpesvirus, type 8), Papillomaviridae (e.g. Human papillomavirus), Polyomaviridae (e.g. BK virus, JC virus), Poxviridae (e.g. smallpox), Hepadnaviridae (e.g. Hepatitis B), Parvoviridae (e.g. Parvovirus B19), Astroviridae (e.g. Human astrovirus), Caliciviridae (e.g. Norwalk virus), Picornaviridae (e.g. coxsackievirus, hepatitis A virus, poliovirus, rhinovirus), Coronaviridae (e.g. Severe acute respiratory syndrome-related coronavirus, strains: Severe acute respiratory syndrome virus, Severe acute respiratory syndrome coronavirus 2 (COVID-19)), Flaviviridae (e.g. Hepatitis C virus, yellow fever virus, dengue virus, West Nile virus, TBE virus), Togaviridae (e.g. Rubella virus), Hepeviridae (e.g. Hepatitis E virus), Retroviridae (Human immunodeficiency virus (HIV)), Orthomyxoviridae (e.g. Influenza virus), Arenaviridae (e.g. Lassa virus), Bunyaviridae (e.g. Crimean-Congo hemorrhagic fever virus, Hantaan virus), Filoviridae (e.g. Ebola virus and Marburg virus), Paramyxoviridae (e.g. Measles virus, Mumps virus, Parainfluenza virus, Respiratory syncytial virus), Rhabdoviridae (Rabies virus), Hepatitis D virus, Reoviridae (e.g. Rotavirus, Orbivirus, Coltivirus, Borna disease virus).

[0846] In one embodiment, the pathogenic fungi that can be targeted and/or modified by the composition, system,(s) and/or component(s) thereof described herein include, but are not limited to, those of the genus Candida (e.g. *C. albicans*), Aspergillus (e.g. *A. fumigatus*, *A. flavus*, *A. clavatus*), Cryptococcus (e.g. *C. neoformans*, *C. gattii*), Histoplasma (*H. capsulatum*), Pneumocystis (e.g. *P. jirovecii*), Stachybotrys (e.g. *S. chartarum*).

[0847] In one embodiment, the pathogenic parasites that can be targeted and/or modified by the composition, system,(s) and/or component(s) thereof described herein include, but are not limited to, protozoa, helminths, and ectoparasites. In one embodiment, the pathogenic protozoa that can be targeted and/or modified by the composition, system,(s) and/or

component(s) thereof described herein include, but are not limited to, those from the groups Sarcodina (e.g. ameba such as *Entamoeba*), Mastigophora (e.g. flagellates such as *Giardia* and *Leishmania*), Cilophora (e.g. ciliates such as *Balantidium*), and sporozoa (e.g. plasmodium and cryptosporidium). In one embodiment, the pathogenic helminths that can be targeted and/or modified by the composition, system, (s) and/or component(s) thereof described herein include, but are not limited to, flatworms (platyhelminths), thorny-headed worms (acanthocephalins), and roundworms (nematodes). In one embodiment, the pathogenic ectoparasites that can be targeted and/or modified by the composition, system, (s) and/or component(s) thereof described herein include, but are not limited to, ticks, fleas, lice, and mites.

[0848] In one embodiment, the pathogenic parasite that can be targeted and/or modified by the composition, system, (s) and/or component(s) thereof described herein include, but are not limited to, *Acanthamoeba* spp., *Balamuthia mandrillaris*, *Babesiosis* spp. (e.g. *Babesia B. divergens*, *B. bigemina*, *B. equi*, *B. microfti*, *B. duncani*), *Balantidiasis* spp. (e.g. *Balantidium coli*), *Blastocystis* spp., *Cryptosporidium* spp., *Cyclosporiasis* spp. (e.g. *Cyclospora cayetanensis*), *Dientamoebiasis* spp. (e.g. *Dientamoeba fragilis*), *Amoebiasis* spp. (e.g. *Entamoeba histolytica*), *Giardiasis* spp. (e.g. *Giardia lamblia*), *Isosporiasis* spp. (e.g. *Isospora belli*), *Leishmania* spp., *Naegleria* spp. (e.g. *Naegleria fowleri*), *Plasmodium* spp. (e.g. *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale curtisi*, *Plasmodium ovale wallikeri*, *Plasmodium malariae*, *Plasmodium knowlesi*), *Rhinosporidiosis* spp. (e.g. *Rhinosporidium seeberi*), *Sarcocystosis* spp. (e.g. *Sarcocystis bovi hominis*, *Sarcocystis sui hominis*), *Toxoplasma* spp. (e.g. *Toxoplasma gondii*), *Trichomonas* spp. (e.g. *Trichomonas vaginalis*), *Trypanosoma* spp. (e.g. *Trypanosoma brucei*), *Trypanosoma* spp. (e.g. *Trypanosoma cruzi*), Tapeworm (e.g. *Cestoda*, *Taenia multiceps*, *Taenia saginata*, *Taenia solium*), *Diphyllobothrium latum* spp., *Echinococcus* spp. (e.g. *Echinococcus granulosus*, *Echinococcus multilocularis*, *E. vogeli*, *E. oligarthrus*), *Hymenolepis* spp. (e.g. *Hymenolepis nana*, *Hymenolepis diminuta*), *Bertiella* spp. (e.g. *Bertiella mucronata*, *Bertiella studeri*), *Spirometra* (e.g. *Spirometra erinaceieuropaei*), *Clonorchis* spp. (e.g. *Clonorchis sinensis*; *Clonorchis viverrini*), *Dicrocoelium* spp. (e.g. *Dicrocoelium dendriticum*), *Fasciola* spp. (e.g. *Fasciola hepatica*, *Fasciola gigantica*), *Fasciolopsis* spp. (e.g. *Fasciolopsis buski*), *Metagonimus* spp. (e.g. *Metagonimus yokogawai*), *Metorchis* spp. (e.g. *Metorchis conjunctus*), *Opisthorchis* spp. (e.g. *Opisthorchis viverrini*, *Opisthorchis felineus*), *Clonorchis* spp. (e.g. *Clonorchis sinensis*), *Paragonimus* spp. (e.g. *Paragonimus westermani*; *Paragonimus*

africanus; *Paragonimus caliensis*; *Paragonimus kellicotti*; *Paragonimus skrjabini*; *Paragonimus uterobilateralis*), *Schistosoma* sp., *Schistosoma* spp. (e.g. *Schistosoma mansoni*, *Schistosoma haematobium*, *Schistosoma japonicum*, *Schistosoma mekongi*, and *Schistosoma intercalatum*), *Echinostoma* spp. (e.g. *E. echinatum*), *Trichobilharzia* spp. (e.g. *Trichobilharzia regent*), *Ancylostoma* spp. (e.g. *Ancylostoma duodenale*), *Necator* spp. (e.g. *Necator americanus*), *Angiostrongylus* spp., *Anisakis* spp., *Ascaris* spp. (e.g. *Ascaris lumbricoides*), *Baylisascaris* spp. (e.g. *Baylisascaris procyonis*), *Brugia* spp. (e.g. *Brugia malayi*, *Brugia timori*), *Dioctophyme* spp. (e.g. *Dioctophyme renale*), *Dracunculus* spp. (e.g. *Dracunculus medinensis*), *Enterobius* spp. (e.g. *Enterobius vermicularis*, *Enterobius gregorii*), *Gnathostoma* spp. (e.g. *Gnathostoma spinigerum*, *Gnathostoma hispidum*), *Halicephalobus* spp. (e.g. *Halicephalobus gingivalis*), *Loa loa* spp. (e.g. *Loa loa filaria*), *Mansonella* spp. (e.g. *Mansonella streptocerca*), *Onchocerca* spp. (e.g. *Onchocerca volvulus*), *Strongyloides* spp. (e.g. *Strongyloides stercoralis*), *Thelazia* spp. (e.g. *Thelazia californiensis*, *Thelazia callipaeda*), *Toxocara* spp. (e.g. *Toxocara canis*, *Toxocara cati*, *Toxascaris leonine*), *Trichinella* spp. (e.g. *Trichinella spiralis*, *Trichinella britovi*, *Trichinella nelsoni*, *Trichinella nativa*), *Trichuris* spp. (e.g. *Trichuris trichiura*, *Trichuris vulpis*), *Wuchereria* spp. (e.g. *Wuchereria bancrofti*), *Dermatobia* spp. (e.g. *Dermatobia hominis*), *Tunga* spp. (e.g. *Tunga penetrans*), *Cochliomyia* spp. (e.g. *Cochliomyia hominivorax*), *Linguatula* spp. (e.g. *Linguatula serrata*), *Archiacanthocephala* sp., *Moniliformis* sp. (e.g. *Moniliformis moniliformis*), *Pediculus* spp. (e.g. *Pediculus humanus capitis*, *Pediculus humanus humanus*), *Pthirus* spp. (e.g. *Pthirus pubis*), *Arachnida* spp. (e.g. *Trombiculidae*, *Ixodidae*, *Argasidae*), *Siphonaptera* spp. (e.g. *Siphonaptera: Pulicinae*), *Cimicidae* spp. (e.g. *Cimex lectularius* and *Cimex hemipterus*), *Diptera* spp., *Demodex* spp. (e.g. *Demodex folliculorum/brevis/canis*), *Sarcoptes* spp. (e.g. *Sarcoptes scabiei*), *Dermanyssus* spp. (e.g. *Dermanyssus gallinae*), *Ornithonyssus* spp. (e.g. *Ornithonyssus sylviae*, *Ornithonyssus bursa*, *Ornithonyssus bacoti*), *Laelaps* spp. (e.g. *Laelaps echidnina*), *Liponyssoides* spp. (e.g. *Liponyssoides sanguineus*).

[0849] In one embodiment the gene targets can be any of those as set forth in Table 1 of Strich and Chertow. 2019. J. Clin. Microbio. 57:4 e01307-18, which is incorporated herein as if expressed in its entirety herein.

[0850] In one embodiment, the method can include delivering a composition, system, and/or component thereof to a pathogenic organism described herein, allowing the

composition, system, and/or component thereof to specifically bind and modify one or more targets in the pathogenic organism, whereby the modification kills, inhibits, reduces the pathogenicity of the pathogenic organism, or otherwise renders the pathogenic organism non-pathogenic. In one embodiment, delivery of the composition, system, occurs *in vivo* (i.e. in the subject being treated). In one embodiment occurs by an intermediary, such as microorganism or phage that is non-pathogenic to the subject but is capable of transferring polynucleotides and/or infecting the pathogenic microorganism. In one embodiment, the intermediary microorganism can be an engineered bacteria, virus, or phage that contains the composition, system,(s) and/or component(s) thereof and/or vectors and/or vector systems. The method can include administering an intermediary microorganism containing the composition, system,(s) and/or component(s) thereof and/or vectors and/or vector systems to the subject to be treated. The intermediary microorganism can then produce the compositions and/or component thereof or transfer a composition, system, polynucleotide to the pathogenic organism. In embodiments, where the compositions and/or component thereof, vector, or vector system is transferred to the pathogenic microorganism, the composition, system, or component thereof is then produced in the pathogenic microorganism and modifies the pathogenic microorganism such that it is less virulent, killed, inhibited, or is otherwise rendered incapable of causing disease and/or infecting and/or replicating in a host or cell thereof.

[0851] In one embodiment, where the pathogenic microorganism inserts its genetic material into the host cell's genome (e.g. a virus), the composition, system, can be designed such that it modifies the host cell's genome such that the viral DNA or cDNA cannot be replicated by the host cell's machinery into a functional virus. In one embodiment, where the pathogenic microorganism inserts its genetic material into the host cell's genome (e.g. a virus), the composition, system, can be designed such that it modifies the host cell's genome such that the viral DNA or cDNA is deleted from the host cell's genome.

[0852] It will be appreciated that inhibiting or killing the pathogenic microorganism, the disease and/or condition that its infection causes in the subject can be treated or prevented. Thus, also provided herein are methods of treating and/or preventing one or more diseases or symptoms thereof caused by any one or more pathogenic microorganisms, such as any of those described herein.

Mitochondrial Diseases

[0853] Some of the most challenging mitochondrial disorders arise from mutations in mitochondrial DNA (mtDNA), a high copy number genome that is maternally inherited. In one embodiment, mtDNA mutations can be modified using a composition, system, described herein. In one embodiment, the mitochondrial disease that can be diagnosed, prognosed, treated, and/or prevented can be MELAS (mitochondrial myopathy encephalopathy, and lactic acidosis and stroke-like episodes), CPEO/PEO (chronic progressive external ophthalmoplegia syndrome/progressive external ophthalmoplegia), KSS (Kearns-Sayre syndrome), MIDD (maternally inherited diabetes and deafness), MERRF (myoclonic epilepsy associated with ragged red fibers), NIDDM (noninsulin-dependent diabetes mellitus), LHON (Leber hereditary optic neuropathy), LS (Leigh Syndrome) an aminoglycoside induced hearing disorder, NARP (neuropathy, ataxia, and pigmentary retinopathy), Extrapyrarnidal disorder with akinesia-rigidity, psychosis and SNHL, Nonsyndromic hearing loss a cardiomyopathy, an encephalomyopathy, Pearson's syndrome, or a combination thereof.

[0854] In one embodiment, the mtDNA of a subject can be modified *in vivo* or *ex vivo*. In one embodiment, where the mtDNA is modified *ex vivo*, after modification the cells containing the modified mitochondria can be administered back to the subject. In one embodiment, the composition, system, or component thereof can be capable of correcting an mtDNA mutation, or a combination thereof.

[0855] In one embodiment, at least one of the one or more mtDNA mutations is selected from the group consisting of: A3243G, C3256T, T3271C, G1019A, A1304T, A15533G, C1494T, C4467A, T1658C, G12315A, A3421G, A8344G, T8356C, G8363A, A13042T, T3200C, G3242A, A3252G, T3264C, G3316A, T3394C, T14577C, A4833G, G3460A, G9804A, G11778A, G14459A, A14484G, G15257A, T8993C, T8993G, G10197A, G13513A, T1095C, C1494T, A1555G, G1541A, C1634T, A3260G, A4269G, T7587C, A8296G, A8348G, G8363A, T9957C, T9997C, G12192A, C12297T, A14484G, G15059A, duplication of CCCCCTCCCC-tandem repeats at positions 305-314 and/or 956-965, deletion at positions from 8,469-13,447, 4,308-14,874, and/or 4,398-14,822, 961ins/delC, the mitochondrial common deletion (e.g. mtDNA 4,977 bp deletion), and combinations thereof.

[0856] In one embodiment, the mitochondrial mutation can be any mutation as set forth in or as identified by use of one or more bioinformatic tools available at Mitomap available at mitomap.org. Such tools include, but are not limited to, "Variant Search, aka Market Finder",

Find Sequences for Any Haplogroup, aka “Sequence Finder”, “Variant Info”, “POLG Pathogenicity Prediction Server”, “MITOMASTER”, “Allele Search”, “Sequence and Variant Downloads”, “Data Downloads”. MitoMap contains reports of mutations in mtDNA that can be associated with disease and maintains a database of reported mitochondrial DNA Base Substitution Diseases: rRNA/tRNA mutations.

[0857] In one embodiment, the method includes delivering a composition, system, and/or a component thereof to a cell, and more specifically one or more mitochondria in a cell, allowing the composition, system, and/or component thereof to modify one or more target polynucleotides in the cell, and more specifically one or more mitochondria in the cell. The target polynucleotides can correspond to a mutation in the mtDNA, such as any one or more of those described herein. In one embodiment, the modification can alter a function of the mitochondria such that the mitochondria functions normally or at least is/are less dysfunctional as compared to an unmodified mitochondria. Modification can occur *in vivo* or *ex vivo*. Where modification is performed *ex vivo*, cells containing modified mitochondria can be administered to a subject in need thereof in an autologous or allogenic manner.

Microbiome Modification

[0858] Microbiomes play important roles in health and disease. For example, the gut microbiome can play a role in health by controlling digestion, preventing growth of pathogenic microorganisms and have been suggested to influence mood and emotion. Imbalanced microbiomes can promote disease and are suggested to contribute to weight gain, unregulated blood sugar, high cholesterol, cancer, and other disorders. A healthy microbiome has a series of joint characteristics that can be distinguished from non-healthy individuals, thus detection and identification of the disease-associated microbiome can be used to diagnose and detect disease in an individual. The compositions, systems, and components thereof can be used to screen the microbiome cell population and be used to identify a disease associated microbiome. Cell screening methods utilizing compositions, systems, and components thereof are described elsewhere herein and can be applied to screening a microbiome, such as a gut, skin, vagina, and/or oral microbiome, of a subject.

[0859] In one embodiment, the microbe population of a microbiome in a subject can be modified using a composition, system, and/or component thereof described herein. In one embodiment, the composition, system, and/or component thereof can be used to identify and select one or more cell types in the microbiome and remove them from the microbiome

population. Exemplary methods of selecting cells using a composition, system, and/or component thereof are described elsewhere herein. In this way the make-up or microorganism profile of the microbiome can be altered. In one embodiment, the alteration causes a change from a diseased microbiome composition to a healthy microbiome composition. In this way the ratio of one type or species of microorganism to another can be modified, such as going from a diseased ratio to a healthy ratio. In one embodiment, the cells selected are pathogenic microorganisms.

[0860] In one embodiment, the compositions and systems described herein can be used to modify a polynucleotide in a microorganism of a microbiome in a subject. In one embodiment, the microorganism is a pathogenic microorganism. In one embodiment, the microorganism is a commensal and non-pathogenic microorganism. Methods of modifying polynucleotides in a cell in the subject are described elsewhere herein and can be applied to these embodiments.

Models of Diseases and Conditions

[0861] In an aspect, the invention provides a method of modeling a disease associated with a genomic locus in a eukaryotic organism or a non-human organism comprising manipulation of a target sequence within a coding, non-coding or regulatory element of said genomic locus comprising delivering a non- naturally occurring or engineered composition comprising a viral vector system comprising one or more viral vectors operably encoding a composition for expression thereof, wherein the composition comprises particle delivery system or the delivery system or the virus particle of any one of the above embodiments or the cell of any one of the above embodiment.

[0862] In one aspect, the invention provides a method of generating a model eukaryotic cell that can include one or more a mutated disease genes and/or infectious microorganisms. In one embodiment, a disease gene is any gene associated an increase in the risk of having or developing a disease. In one embodiment, the method includes (a) introducing one or more vectors into a eukaryotic cell, wherein the one or more vectors comprise a composition, system, and/or component thereof and/or a vector or vector system that is capable of driving expression of a composition, system, and/or component thereof including, but not limited to: a guide sequence, one or more IscB polypeptide nucleases, and combinations thereof and (b) allowing a composition, system, or complex to bind to one or more target polynucleotides, e.g., to effect cleavage, nicking, or other modification of the target polynucleotide within said disease gene, wherein the composition, system, or complex is composed of one or more IscB polypeptide or

CRISPR-associated IscB polypeptide nuclease complexed with (1) one or more ω RNAs or guide sequences that is/are hybridized to the target sequence(s) within the target polynucleotide(s), and optionally (2) the ω RNA scaffold sequence(s), thereby generating a model eukaryotic cell comprising one or more mutated disease gene(s). Thus, In one embodiment the composition and system, contains nucleic acid molecules for and drives expression of one or more of: a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, a guide sequence and/or a Homologous Recombination template and/or a stabilizing ligand if the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease has a destabilization domain. In one embodiment, said cleavage comprises cleaving one or two strands at the location of the target sequence by the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease. In one embodiment, nicking comprises nicking one or two strands at the location of the target sequence by the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease. In one embodiment, said cleavage or nicking results in modified transcription of a target polynucleotide. In one embodiment, modification results in decreased transcription of the target polynucleotide. In one embodiment, the method further comprises repairing said cleaved or nicked target polynucleotide by homologous recombination with a recombination template polynucleotide, wherein said repair results in a mutation comprising an insertion, deletion, or substitution of one or more nucleotides of said target polynucleotide. In one embodiment, said mutation results in one or more amino acid changes in a protein expression from a gene comprising the target sequence.

[0863] The disease modeled can be any disease with a genetic or epigenetic component. In one embodiment, the disease modeled can be any as discussed elsewhere herein.

In situ Disease Detection

[0864] The compositions, systems, and/or components thereof can be used for diagnostic methods of detection such as in CASFISH (see e.g. Deng et al. 2015. PNAS USA 112(38): 11870-11875), CRISPR-Live FISH (see e.g. Wang et al. 2020. Science; 365(6459):1301-1305), sm-FISH (Lee and Jefcoate. 2017. Front. Endocrinol. doi.org/10.3389/fendo.2017.00289), sequential FISH CRISPRainbow (Ma et al. Nat Biotechnol, 34 (2016), pp. 528-530), CRISPR-Sirius (Nat Methods, 15 (2018), pp. 928-931), Casilio (Cheng et al. Cell Res, 26 (2016), pp. 254-257), Halo-Tag based genomic loci visualization techniques (e.g. Deng et al. 2015. PNAS USA 112(38): 11870-11875; Knight et

al., Science, 350 (2015), pp. 823-826), RNA-aptamer based methods (e.g. Ma et al., J Cell Biol, 214 (2016), pp. 529-537), molecular beacon-based methods (e.g. Zhao et al. Biomaterials, 100 (2016), pp. 172-183; Wu et al. Nucleic Acids Res (2018)), Quantum Dot-based systems (e.g. Ma et al. Anal Chem, 89 (2017), pp. 12896-12901), multiplexed methods (e.g. Ma et al., Proc Natl Acad Sci U S A, 112 (2015), pp. 3002-3007; Fu et al. Nat Commun, 7 (2016), p. 11707; Ma et al. Nat Biotechnol, 34 (2016), pp. 528-530; Shao et al. Nucleic Acids Res, 44 (2016), Article e86); Wang et al. Sci Rep, 6 (2016), p. 26857), ζ , and other in situ CRISPR-hybridization based methods (e.g. Chen et al. Cell, 155 (2013), pp. 1479-1491; Gu et al. Science, 359 (2018), pp. 1050-1055; Tanebaum et al. Cell, 159 (2014), pp. 635-646; Ye et al. Protein Cell, 8 (2017), pp. 853-855; Chen et al. Nat Commun, 9 (2018), p. 5065; Shao et al. ACS Synth Biol (2017); Fu et al. Nat Commun, 7 (2016), p. 11707; Shao et al. Nucleic Acids Res, 44 (2016), Article e86; Wang et al., Sci Rep, 6 (2016), p. 26857), all of which are incorporated by reference herein as if expressed in their entirety and whose teachings can be adapted to the compositions, systems, and components thereof described herein in view of the description herein.

[0865] In one embodiment, the composition, system, or component thereof can be used in a detection method, such as an in-situ detection method described herein. In one embodiment, the composition, system, or component thereof can include a catalytically inactivate IscB polypeptide nuclease described herein and use this system in detection methods such as fluorescence in situ hybridization (FISH) or any other described herein. In one embodiment, the inactivated IscB polypeptide nuclease, which lacks the ability to produce DNA double-strand breaks may be fused with a marker, such as fluorescent protein, such as the enhanced green fluorescent protein (eGFP) and co-expressed with small guide RNAs to target pericentric, centric and telomeric repeats *in vivo*. The dead IscB polypeptide or CRISPR-associated IscB polypeptide nuclease or system thereof can be used to visualize both repetitive sequences and individual genes in the human genome. Such new applications of labelled dead IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and compositions, systems, thereof can be important in imaging cells and studying the functional nuclear architecture, especially in cases with a small nucleus volume or complex 3-D structures.

Cell Selection

[0866] In one embodiment, the compositions, systems, and/or components thereof described herein can be used in a method to screen and/or select cells. In one embodiment,

composition, system,-based screening/selection method can be used to identify diseased cells in a cell population. In one embodiment, selection of the cells results in a modification in the cells such that the selected cells die. In this way, diseased cells can be identified, and removed from the healthy cell population. In one embodiment, the diseased cells can be a cancer cell, pre-cancerous cell, a virus or other pathogenic organism infected cells, or otherwise abnormal cell. In one embodiment, the modification can impart another detectable change in the cells to be selected (e.g. a functional change and/or genomic barcode) that facilitates selection of the desired cells. In one embodiment a negative selection scheme can be used to obtain a desired cell population. In these embodiments, the cells to be selected against are modified, thus can be removed from the cell population based on their death or identification or sorting based the detectable change imparted on the cells. Thus, in these embodiments, the remaining cells after selection are the desired cell population.

[0867] In one embodiment, a method of selecting one or more cell(s) containing a polynucleotide modification can include: introducing one or more composition, system,(s) and/or components thereof, and/or vectors or vector systems into the cell(s), wherein the composition, system,(s) and/or components thereof, and/or vectors or vector systems contains and/or is capable of expressing one or more of: a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, an ω RNA sequence, and a recombination template; wherein, for example that which is being expressed is within and expressed *in vivo* by the composition, system, vector or vector system and/or the recombination template comprises the one or more mutations that abolish IscB polypeptide or CRISPR-associated IscB polypeptide nuclease cleavage; allowing homologous recombination of the recombination template with the target polynucleotide in the cell(s) to be selected; allowing a composition, system, or complex to bind to a target polynucleotide to effect cleavage of the target polynucleotide within said gene, wherein the AAV- complex comprises the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease complexed with (1) the ω RNA or guide sequence that is hybridized to the target sequence within the target polynucleotide, and (2) the ω RNA scaffold, wherein binding of the complex to the target polynucleotide induces cell death or imparts some other detectable change to the cell, thereby allowing one or more cell(s) in which one or more mutations have been introduced to be selected. In one embodiment, the cell to be selected may be a eukaryotic cell. In one embodiment, the cell to be selected may be a prokaryotic cell.

Selection of specific cells via the methods herein can be performed without requiring a selection marker or a two-step process that may include a counter-selection system.

Therapeutic Agent Development

[0868] The compositions, systems, and components thereof described herein can be used to develop IscB polypeptide nuclease-based biologically active agents, such as small molecule therapeutics. Thus, described herein are methods for developing a biologically active agent that modulates a cell function and/or signaling event associated with a disease and/or disease gene. In one embodiment, the method comprises (a) contacting a test compound with a diseased cell and/or a cell containing a disease gene cell; and (b) detecting a change in a readout that is indicative of a reduction or an augmentation of a cell signaling event or other cell functionality associated with said disease or disease gene, thereby developing said biologically active agent that modulates said cell signaling event or other functionality associated with said disease gene. In one embodiment, the diseased cell is a model cell described elsewhere herein. In one embodiment, the diseased cell is a diseased cell isolated from a subject in need of treatment. In one embodiment, the test compound is a small molecule agent. In one embodiment, test compound is a small molecule agent. In one embodiment, the test compound is a biologic molecule agent.

[0869] In one embodiment, the method involves developing a therapeutic based on the composition, system, described herein. In one embodiment, the therapeutic comprises a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and/or a ω RNA with a reprogrammable spacer capable of hybridizing to a target sequence of interest. In one embodiment, the therapeutic is a vector or vector system that can contain a) a first regulatory element operably linked to a nucleotide sequence encoding the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease; and b) a second regulatory element operably linked to one or more nucleotide sequences encoding one or more nucleic acid molecules comprising a ω RNA comprising a reprogrammable spacer sequence, a conserved RNA sequence; wherein components (a) and (b) are located on same or different vectors. In one embodiment, the biologically active agent is a composition comprising a delivery system operably configured to deliver composition, system, or components thereof, and/or or one or more polynucleotide sequences, vectors, or vector systems containing or encoding said components into a cell and capable of forming a complex with the components of the composition and system herein, and

wherein said complex is operable in the cell. In one embodiment, the complex can include the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease as described herein, ω RNA scaffold comprising the guide sequence (reprogrammable spacer sequence), and a conserved nucleotide sequence. In any such compositions, the delivery system can be a yeast system, a lipofection system, a microinjection system, a biolistic system, virosomes, liposomes, immunoliposomes, polycations, lipid:nucleic acid conjugates or artificial virions, or any other system as described herein. In one embodiment, the delivery is via a particle, a nanoparticle, a lipid or a cell penetrating peptide (CPP).

[0870] Also described herein are methods for developing or designing a composition, system,, optionally a composition, system, based therapy or therapeutic, comprising (a) selecting for a (therapeutic) locus of interest ω RNA target sites, wherein said target sites have minimal sequence variation across a population, and from said selected target sites subselecting target sites, wherein a ω RNA directed against said target sites recognizes a minimal number of off-target sites across said population, or (b) selecting for a (therapeutic) locus of interest ω RNA target sites, wherein said target sites have minimal sequence variation across a population, or selecting for a (therapeutic) locus of interest ω RNA target sites, wherein a ω RNA directed against said target sites recognizes a minimal number of off-target sites across said population, and optionally estimating the number of (sub)selected target sites needed to treat or otherwise modulate or manipulate a population, and optionally validating one or more of the (sub)selected target sites for an individual subject, optionally designing one or more ω RNA recognizing one or more of said (sub)selected target sites.

[0871] In one embodiment, the method for developing or designing a ω RNA for use in a composition, system,, optionally a composition, system, based therapy or therapeutic, can include (a) selecting for a (therapeutic) locus of interest ω RNA target sites, wherein said target sites have minimal sequence variation across a population, and from said selected target sites subselecting target sites, wherein a gRNA directed against said target sites recognizes a minimal number of off-target sites across said population, or (b) selecting for a (therapeutic) locus of interest gRNA target sites, wherein said target sites have minimal sequence variation across a population, or selecting for a (therapeutic) locus of interest gRNA target sites, wherein a gRNA directed against said target sites recognizes a minimal number of off-target sites across said population, and optionally estimating the number of (sub)selected target sites needed to

treat or otherwise modulate or manipulate a population, optionally validating one or more of the (sub)selected target sites for an individual subject, optionally designing one or more gRNA recognizing one or more of said (sub)selected target sites.

[0872] In one embodiment, the method for developing or designing a composition, system, optionally a composition, system, based therapy or therapeutic in a population, can include (a) selecting for a (therapeutic) locus of interest reprogrammable spacer target sites, wherein said target sites have minimal sequence variation across a population, and from said selected target sites subselecting target sites, wherein a ω RNA directed against said target sites recognizes a minimal number of off-target sites across said population, or (b) selecting for a (therapeutic) locus of interest ω RNA reprogrammable spacer target sites, wherein said target sites have minimal sequence variation across a population, or selecting for a (therapeutic) locus of interest ω RNA reprogrammable spacer target sites, wherein a ω RNA directed against said target sites recognizes a minimal number of off-target sites across said population, and optionally estimating the number of (sub)selected target sites needed to treat or otherwise modulate or manipulate a population, optionally validating one or more of the (sub)selected target sites for an individual subject, optionally designing one or more ω RNA recognizing one or more of said (sub)selected target sites.

[0873] In one embodiment the method for developing or designing a gRNA for use in a composition, system,, optionally a composition, system, based therapy or therapeutic in a population, can include (a) selecting for a (therapeutic) locus of interest gRNA target sites, wherein said target sites have minimal sequence variation across a population, and from said selected target sites subselecting target sites, wherein a gRNA directed against said target sites recognizes a minimal number of off-target sites across said population, or (b) selecting for a (therapeutic) locus of interest gRNA target sites, wherein said target sites have minimal sequence variation across a population, or selecting for a (therapeutic) locus of interest gRNA target sites, wherein a gRNA directed against said target sites recognizes a minimal number of off-target sites across said population, and optionally estimating the number of (sub)selected target sites needed to treat or otherwise modulate or manipulate a population, optionally validating one or more of the (sub)selected target sites for an individual subject, optionally designing one or more ω RNA reprogrammable spacer recognizing one or more of said (sub)selected target sites.

[0874] In one embodiment, the method for developing or designing a composition, system, such as a composition, system, based therapy or therapeutic, optionally in a population; or for developing or designing a ω RNA reprogrammable spacer for use in a composition, system, optionally a composition, system, based therapy or therapeutic, optionally in a population, can include selecting a set of target sequences for one or more loci in a target population, wherein the target sequences do not contain variants occurring above a threshold allele frequency in the target population (i.e. platinum target sequences); removing from said selected (platinum) target sequences any target sequences having high frequency off-target candidates (relative to other (platinum) targets in the set) to define a final target sequence set; preparing one or more, such as a set of compositions, systems, based on the final target sequence set, optionally wherein a number of composition prepared is based (at least in part) on the size of a target population.

[0875] In an embodiment, off-target candidates/off-targets, TAM restrictiveness, target cleavage efficiency, or effector protein specificity is identified or determined using a sequencing-based double-strand break (DSB) detection assay, such as described herein elsewhere. In an embodiment, off-target candidates/off-targets are identified or determined using a sequencing-based double-strand break (DSB) detection assay, such as described herein elsewhere. In an embodiment, off-targets, or off target candidates have at least 1, preferably 1-3, mismatches or (distal) TAM mismatches, such as 1 or more, such as 1, 2, 3, or more (distal) TAM mismatches. In an embodiment, sequencing-based DSB detection assay comprises labeling a site of a DSB with an adapter comprising a primer binding site, labeling a site of a DSB with a barcode or unique molecular identifier, or combination thereof, as described herein elsewhere.

[0876] It will be understood that the reprogrammable spacer sequence of the ω RNA is 100% complementary to the target site, i.e. does not comprise any mismatch with the target site. It will be further understood that “recognition” of an (off-)target site by a reprogrammable spacer presupposes composition, system, functionality, i.e. an (off-)target site is only recognized by a reprogrammable spacer RNA if binding of the reprogrammable spacer RNA to the (off-)target site leads to composition, system, activity (such as induction of single or double strand DNA cleavage, transcriptional modulation, etc.).

[0877] In an embodiment, the target sites having minimal sequence variation across a population are characterized by absence of sequence variation in at least 99%, preferably at least 99.9%, more preferably at least 99.99% of the population. In an embodiment, optimizing target location comprises selecting target sequences or loci having an absence of sequence variation in at least 99%, %, preferably at least 99.9%, more preferably at least 99.99% of a population. These targets are referred to herein elsewhere also as “platinum targets”. In an embodiment, said population comprises at least 1000 individuals, such as at least 5000 individuals, such as at least 10000 individuals, such as at least 50000 individuals.

[0878] In an embodiment, the off-target sites are characterized by at least one mismatch between the off-target site and the ω RNA. In an embodiment, the off-target sites are characterized by at most five, preferably at most four, more preferably at most three mismatches between the off-target site and the ω RNA. In an embodiment, the off-target sites are characterized by at least one mismatch between the off-target site and the ω RNA and by at most five, preferably at most four, more preferably at most three mismatches between the off-target site and the ω RNA.

[0879] In an embodiment, said minimal number of off-target sites across said population is determined for high-frequency haplotypes in said population. In an embodiment, said minimal number of off-target sites across said population is determined for high-frequency haplotypes of the off-target site locus in said population. In an embodiment, said minimal number of off-target sites across said population is determined for high-frequency haplotypes of the target site locus in said population. In an embodiment, the high-frequency haplotypes are characterized by occurrence in at least 0.1% of the population.

[0880] In an embodiment, the number of (sub)selected target sites needed to treat a population is estimated based on based low frequency sequence variation, such as low frequency sequence variation captured in large scale sequencing datasets. In an embodiment, the number of (sub)selected target sites needed to treat a population of a given size is estimated.

[0881] In an embodiment, the method further comprises obtaining genome sequencing data of a subject to be treated; and treating the subject with a composition, system, selected from the set of compositions, systems, wherein the composition, system, selected is based (at least in part) on the genome sequencing data of the individual. In an embodiment, the ((sub)selected) target is validated by genome sequencing, preferably whole genome sequencing.

[0882] In an embodiment, target sequences or loci as described herein are (further) selected based on optimization of one or more parameters, such as TAM type (natural or modified), TAM nucleotide content, TAM length, target sequence length, TAM restrictiveness, target cleavage efficiency, and target sequence position within a gene, a locus or other genomic region. Methods of optimization are discussed in greater detail elsewhere herein.

[0883] In an embodiment, target sequences or loci as described herein are (further) selected based on optimization of one or more of target loci location, target length, target specificity, and TAM characteristics. As used herein, TAM characteristics may comprise for instance TAM sequence, TAM length, and/or TAM GC contents. In an embodiment, optimizing TAM characteristics comprises optimizing nucleotide content of a TAM. In an embodiment, optimizing nucleotide content of TAM is selecting a TAM with a motif that maximizes abundance in the one or more target loci, minimizes mutation frequency, or both. Minimizing mutation frequency can for instance be achieved by selecting TAM sequences devoid of or having low or minimal CpG.

[0884] In an embodiment, the effector protein for each composition and system, in the set of compositions, systems, is selected based on optimization of one or more parameters selected from the group consisting of; effector protein size, ability of effector protein to access regions of high chromatin accessibility, degree of uniform enzyme activity across genomic targets, epigenetic tolerance, mismatch/budge tolerance, effector protein specificity, effector protein stability or half-life, effector protein immunogenicity or toxicity. Methods of optimization are discussed in greater detail elsewhere herein.

Optimization of the Systems

[0885] The methods of the present invention can involve optimization of selected parameters or variables associated with the composition, system, and/or its functionality, as described herein further elsewhere. Optimization of the composition, system, in the methods as described herein may depend on the target(s), such as the therapeutic target or therapeutic targets, the mode or type of composition, system, modulation, such as composition, system, based therapeutic target(s) modulation, modification, or manipulation, as well as the delivery of the composition, system, components. One or more targets may be selected, depending on the genotypic and/or phenotypic outcome. For instance, one or more therapeutic targets may be selected, depending on (genetic) disease etiology or the desired therapeutic outcome. The (therapeutic) target(s) may be a single gene, locus, or other genomic site, or may be multiple

genes, loci or other genomic sites. As is known in the art, a single gene, locus, or other genomic site may be targeted more than once, such as by use of multiple ω RNAs, or ω RNA scaffold and multiple reprogrammable spacers.

[0886] The activity of the composition and/or system, such as IscB polypeptide nuclease-based therapy or therapeutics may involve target disruption, such as target mutation, such as leading to gene knockout. The activity of the composition and/or system, such as IscB polypeptide nuclease-based therapy or therapeutics may involve replacement of particular target sites, such as leading to target correction. IscB polypeptide or CRISPR-associated IscB polypeptide nuclease-based therapy or therapeutics may involve removal of particular target sites, such as leading to target deletion. The activity of the composition and/or system, such as IscB polypeptide or CRISPR-associated IscB polypeptide nuclease-based therapy or therapeutics may involve modulation of target site functionality, such as target site activity or accessibility, leading for instance to (transcriptional and/or epigenetic) gene or genomic region activation or gene or genomic region silencing. The skilled person will understand that modulation of target site functionality may involve IscB polypeptide or CRISPR-associated IscB polypeptide nuclease mutation (such as for instance generation of a catalytically inactive IscB polypeptide or CRISPR-associated IscB polypeptide nuclease) and/or functionalization (such as for instance fusion of the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease with a heterologous functional domain, such as a transcriptional activator or repressor), as described herein elsewhere.

[0887] Accordingly, in an aspect, the invention relates to a method as described herein, comprising selection of one or more (therapeutic) target, selecting one or more functionality of the composition and/or system, and optimization of selected parameters or variables associated with the composition and/or its functionality. In a related aspect, the invention relates to a method as described herein, comprising (a) selecting one or more (therapeutic) target loci, (b) selecting one or more composition functionalities, (c) optionally selecting one or more modes of delivery, and preparing, developing, or designing a composition herein selected based on steps (a)-(c).

[0888] In an embodiment, the functionality of the composition and/or system comprises genomic mutation. In an embodiment, the functionality of the composition and/or system comprises single genomic mutation. In an embodiment, the functionality of the composition

and/or system functionality comprises multiple genomic mutation. In an embodiment, the functionality of the composition and/or system comprises gene knockout. In an embodiment, the functionality of the composition and/or system comprises single gene knockout. In an embodiment, the functionality of the composition and/or system comprises multiple gene knockout. In an embodiment, the functionality of the composition and/or system comprises gene correction. In an embodiment, the functionality of the composition and/or system comprises single gene correction. In an embodiment, the functionality of the composition and/or system comprises multiple gene correction. In an embodiment, the functionality of the composition and/or system comprises genomic region correction. In an embodiment, the functionality of the composition and/or system comprises single genomic region correction. In an embodiment, the functionality of the composition and/or system comprises multiple genomic region correction. In an embodiment, the functionality of the composition and/or system comprises gene deletion. In an embodiment, the functionality of the composition and/or system comprises single gene deletion. In an embodiment, the functionality of the composition and/or system comprises multiple gene deletion. In an embodiment, the functionality of the composition and/or system comprises genomic region deletion. In an embodiment, the functionality of the composition and/or system comprises single genomic region deletion. In an embodiment, the functionality of the composition and/or system comprises multiple genomic region deletion. In an embodiment, the functionality of the composition and/or system comprises modulation of gene or genomic region functionality. In an embodiment, the functionality of the composition and/or system comprises modulation of single gene or genomic region functionality. In an embodiment, the functionality of the composition and/or system comprises modulation of multiple gene or genomic region functionality. In an embodiment, the functionality of the composition and/or system comprises gene or genomic region functionality, such as gene or genomic region activity. In an embodiment, the functionality of the composition and/or system comprises single gene or genomic region functionality, such as gene or genomic region activity. In an embodiment, the functionality of the composition and/or system comprises multiple gene or genomic region functionality, such as gene or genomic region activity. In an embodiment, the functionality of the composition and/or system comprises modulation gene activity or accessibility optionally leading to transcriptional and/or epigenetic gene or genomic region activation or gene or genomic region silencing. In an embodiment, the functionality of the composition and/or system comprises

modulation single gene activity or accessibility optionally leading to transcriptional and/or epigenetic gene or genomic region activation or gene or genomic region silencing. In an embodiment, the functionality of the composition and/or system comprises modulation multiple gene activity or accessibility optionally leading to transcriptional and/or epigenetic gene or genomic region activation or gene or genomic region silencing.

[0889] Optimization of selected parameters or variables in the methods as described herein may result in optimized or improved the system, such as IscB polypeptide or CRISPR-associated IscB polypeptide nuclease-based therapy or therapeutic, specificity, efficacy, and/or safety. In an embodiment, one or more of the following parameters or variables are taken into account, are selected, or are optimized in the methods of the invention as described herein: IscB polypeptide or CRISPR-associated IscB polypeptide nuclease allosteric interactions, IscB polypeptide nuclease functional domains and functional domain interactions, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease specificity, hRNA specificity, composition specificity, TAM restrictiveness, TAM type (natural or modified), TAM nucleotide content, TAM length, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease activity, hRNA activity, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease/guide complex activity, target cleavage efficiency, target site selection, target sequence length, ability of effector protein to access regions of high chromatin accessibility, degree of uniform enzyme activity across genomic targets, epigenetic tolerance, mismatch/budge tolerance, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease stability, IscB polypeptide or CRISPR-associated IscB polypeptide or CRISPR-associated IscB polypeptide nuclease mRNA stability, gRNA stability, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease complex stability, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease protein or mRNA immunogenicity or toxicity, gRNA immunogenicity or toxicity, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease immunogenicity or toxicity, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease or mRNA dose or titer, gRNA dose or titer, dose or titer, IscB polypeptide nuclease protein size, IscB polypeptide nuclease expression level, gRNA expression level, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease expression level, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease spatiotemporal expression, gRNA spatiotemporal expression, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease/hRNA spatiotemporal expression.

[0890] By means of example, and without limitation, parameter or variable optimization may be achieved as follows. IscB polypeptide nuclease specificity may be optimized by selecting the most specific IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, e.g. IscB polypeptide or CRISPR-associated IscB polypeptide. This may be achieved for instance by selecting the most specific IscB polypeptide nuclease orthologue or by specific IscB polypeptide or CRISPR-associated IscB polypeptide or CRISPR-associated IscB polypeptide nuclease mutations which increase specificity. ω RNA specificity may be optimized by selecting the most specific ω RNA. This can be achieved for instance by selecting ω RNA having low homology, i.e. at least one or preferably more, such as at least 2, or preferably at least 3, mismatches to off-target sites. The specificity may be optimized by increasing IscB polypeptide or CRISPR-associated IscB polypeptide nuclease specificity and/or ω RNA specificity as above. TAM restrictiveness may be optimized by selecting a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease having to most restrictive TAM recognition. This can be achieved for instance by selecting a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease ortholog having more restrictive TAM recognition or by specific IscB polypeptide or CRISPR-associated IscB polypeptide or CRISPR-associated IscB polypeptide mutations which increase or alter TAM restrictiveness. TAM type may be optimized for instance by selecting the appropriate IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, such as the appropriate IscB polypeptide or CRISPR-associated IscB polypeptide nuclease recognizing a desired TAM type. The IscB polypeptide or CRISPR-associated IscB polypeptide nuclease or TAM type may be naturally occurring or may for instance be optimized based on IscB polypeptide or CRISPR-associated IscB polypeptide nuclease mutants having an altered TAM recognition, or TAM recognition repertoire. TAM nucleotide content may for instance be optimized by selecting the appropriate IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, such as the appropriate IscB polypeptide or CRISPR-associated IscB polypeptide nuclease recognizing a desired TAM nucleotide content. The IscB polypeptide or CRISPR-associated IscB polypeptide nuclease or TAM type may be naturally occurring or may for instance be optimized based on IscB polypeptide or CRISPR-associated IscB polypeptide nuclease mutants having an altered TAM recognition, or TAM recognition repertoire. TAM length may for instance be optimized by selecting the appropriate IscB polypeptide nuclease, such as the appropriate IscB polypeptide or CRISPR-associated

IscB polypeptide nuclease recognizing a desired TAM nucleotide length. The IscB polypeptide or CRISPR-associated IscB polypeptide nuclease or TAM type may be naturally occurring or may for instance be optimized based on IscB polypeptide nuclease mutants having an altered TAM recognition, or TAM recognition repertoire.

[0891] Target length or target sequence length may be optimized, for instance, by selecting the appropriate IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, such as the appropriate IscB polypeptide nuclease recognizing a desired target or target sequence nucleotide length. Alternatively, or in addition, the target (sequence) length may be optimized by providing a target having a length deviating from the target (sequence) length typically associated with the IscB polypeptide nuclease, such as the naturally occurring IscB polypeptide or CRISPR-associated IscB polypeptide nuclease. The IscB polypeptide or CRISPR-associated IscB polypeptide nuclease or target (sequence) length may be naturally occurring or may for instance be optimized based on IscB polypeptide or CRISPR-associated IscB polypeptide nuclease mutants having an altered target (sequence) length recognition, or target (sequence) length recognition repertoire. For instance, increasing or decreasing target (sequence) length may influence target recognition and/or off-target recognition. IscB polypeptide or CRISPR-associated IscB polypeptide nuclease activity may be optimized by selecting the most active IscB polypeptide or CRISPR-associated IscB polypeptide nuclease. This may be achieved for instance by selecting the most active IscB polypeptide or CRISPR-associated IscB polypeptide nuclease ortholog or by specific IscB polypeptide or CRISPR-associated IscB polypeptide nuclease mutations which increase activity. The ability of the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease protein to access regions of high chromatin accessibility, may be optimized by selecting the appropriate IscB polypeptide or CRISPR-associated IscB polypeptide nuclease or mutant thereof, and can consider the size of the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, charge, or other dimensional variables etc. The degree of uniform IscB polypeptide or CRISPR-associated IscB polypeptide nuclease activity may be optimized by selecting the appropriate IscB polypeptide nuclease or mutant thereof, and can consider IscB polypeptide or CRISPR-associated IscB polypeptide or CRISPR-associated IscB polypeptide nuclease specificity and/or activity, TAM specificity, target length, mismatch tolerance, epigenetic tolerance, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and/or mRNA stability and/or half-life, IscB polypeptide or CRISPR-

associated IscB polypeptide nuclease and/or ω RNA immunogenicity and/or toxicity, etc. ω RNA activity may be optimized by selecting the most active ω RNA. In one embodiment, this can be achieved by increasing ω RNA stability through RNA modification. compositions activity may be optimized by increasing IscB polypeptide or CRISPR-associated IscB polypeptide or CRISPR-associated IscB polypeptide nuclease activity and/or ω RNA activity as above.

[0892] The target site selection may be optimized by selecting the optimal position of the target site within a gene, locus or other genomic region. The target site selection may be optimized by optimizing target location comprises selecting a target sequence with a gene, locus, or other genomic region having low variability. This may be achieved for instance by selecting a target site in an early and/or conserved exon or domain (i.e. having low variability, such as polymorphisms, within a population).

[0893] In an embodiment, optimizing target (sequence) length comprises selecting a target sequence within one or more target loci between 5 and 25 nucleotides. In an embodiment, a target sequence is 20 nucleotides.

[0894] In an embodiment, optimizing target specificity comprises selecting targets loci that minimize off-target candidates.

[0895] In one embodiment, the target site may be selected by minimization of off-target effects (e.g. off-targets qualified as having 1-5, 1-4, or preferably 1-3 mismatches compared to target and/or having one or more TAM mismatches, such as distal TAM mismatches), preferably also considering variability within a population. IscB polypeptide nuclease stability may be optimized by selecting IscB polypeptide nuclease having appropriate half-life, such as preferably a short half-life while still capable of maintaining sufficient activity. In one embodiment, this can be achieved by selecting an appropriate IscB polypeptide nuclease orthologue having a specific half-life or by specific IscB polypeptide nuclease mutations or modifications which affect half-life or stability, such as inclusion (e.g. fusion) of stabilizing or destabilizing domains or sequences. IscB polypeptide nuclease mRNA stability may be optimized by increasing or decreasing IscB polypeptide nuclease mRNA stability. In one embodiment, this can be achieved by increasing or IscB polypeptide or CRISPR-associated IscB polypeptide nuclease mRNA stability through mRNA modification. hRNA stability may be optimized by increasing or decreasing ω RNA stability. In one embodiment, this can be

achieved by increasing or decreasing ω RNA stability through RNA modification. The stability may be optimized by increasing or decreasing IscB polypeptide or CRISPR-associated IscB polypeptide nuclease stability and/or gRNA stability as above. IscB polypeptide nuclease protein or mRNA immunogenicity or toxicity may be optimized by decreasing IscB polypeptide or CRISPR-associated IscB polypeptide nuclease or mRNA immunogenicity or toxicity. In one embodiment, this can be achieved by mRNA or protein modifications. Similarly, in case of DNA based expression systems, DNA immunogenicity or toxicity may be decreased. ω RNA immunogenicity or toxicity may be optimized by decreasing hRNA immunogenicity or toxicity. In one embodiment, this can be achieved by ω RNA modifications. Similarly, in case of DNA based expression systems, DNA immunogenicity or toxicity may be decreased. The immunogenicity or toxicity may be optimized by decreasing IscB polypeptide or CRISPR-associated IscB polypeptide nuclease immunogenicity or toxicity and/or ω RNA immunogenicity or toxicity as above, or by selecting the least immunogenic or toxic IscB polypeptide or CRISPR-associated IscB polypeptide nuclease /hRNA combination. Similarly, in case of DNA based expression systems, DNA immunogenicity or toxicity may be decreased. IscB polypeptide or CRISPR-associated IscB polypeptide nuclease protein or mRNA dose or titer may be optimized by selecting dosage or titer to minimize toxicity and/or maximize specificity and/or efficacy. ω RNA dose or titer may be optimized by selecting dosage or titer to minimize toxicity and/or maximize specificity and/or efficacy. The composition dose or titer may be optimized by selecting dosage or titer to minimize toxicity and/or maximize specificity and/or efficacy. The IscB polypeptide or CRISPR-associated IscB polypeptide nuclease size may be optimized by selecting minimal protein size to increase efficiency of delivery, in particular for virus mediated delivery. IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, ω RNA, or complex thereof expression level may be optimized by limiting (or extending) the duration of expression and/or limiting (or increasing) expression level. This may be achieved for instance by using self-inactivating compositions, systems,, such as including a self-targeting (e.g. IscB polypeptide or CRISPR-associated IscB polypeptide nuclease targeting) gRNA, by using viral vectors having limited expression duration, by using appropriate promoters for low (or high) expression levels, by combining different delivery methods for individual IscB polypeptide or CRISPR-associated IscB polypeptide system components, such as virus mediated delivery of IscB polypeptide or

CRISPR-associated IscB polypeptide nuclease encoding nucleic acid combined with non-virus mediated delivery of ω RNA, or virus mediated delivery of ω RNA combined with non-virus mediated delivery of IscB polypeptide or CRISPR-associated IscB polypeptide or CRISPR-associated IscB polypeptide nuclease or mRNA. IscB polypeptide nuclease, ω RNA, or IscB polypeptide or CRISPR-associated IscB polypeptide complex spatiotemporal expression may be optimized by appropriate choice of conditional and/or inducible expression systems, including controllable IscB polypeptide or CRISPR-associated IscB polypeptide nuclease activity optionally a destabilized IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and/or a split IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, and/or cell- or tissue-specific expression systems.

[0896] In an aspect, the invention relates to a method as described herein, comprising selection of one or more (therapeutic) target, selecting the functionality of the composition and/or system, selecting composition mode of delivery, selecting composition delivery vehicle or expression system, and optimization of selected parameters or variables associated with the composition and/or its functionality, optionally wherein the parameters or variables are one or more selected from IscB polypeptide or CRISPR-associated IscB polypeptide nuclease specificity, ω RNA specificity, IscB polypeptide or CRISPR-associated IscB polypeptide complex specificity, TAM restrictiveness, TAM type (natural or modified), TAM nucleotide content, TAM length, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease activity, gRNA activity, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease/hRNA complex activity, target cleavage efficiency, target site selection, target sequence length, ability of effector protein to access regions of high chromatin accessibility, degree of uniform enzyme activity across genomic targets, epigenetic tolerance, mismatch/budge tolerance, IscB polypeptide or CRISPR-associated IscB polypeptide or CRISPR-associated IscB polypeptide nuclease stability, IscB polypeptide nuclease mRNA stability, ω RNA stability, IscB polypeptide or CRISPR-associated IscB polypeptide complex stability, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease protein or mRNA immunogenicity or toxicity, ω RNA immunogenicity or toxicity, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease/ ω RNA complex immunogenicity or toxicity, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease protein or mRNA dose or titer, ω RNA dose or titer, IscB polypeptide or CRISPR-associated IscB polypeptide complex dose

or titer, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease protein size, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease expression level, ω RNA expression level, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease/ ω RNA molecule complex expression level, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease spatiotemporal expression, ω RNA spatiotemporal expression, IscB polypeptide or CRISPR-associated IscB polypeptide nuclease/ ω RNA complex spatiotemporal expression.

[0897] It will be understood that the parameters or variables to be optimized as well as the nature of optimization may depend on the (therapeutic) target, the functionality of the composition and/or system, the system mode of delivery, and/or the composition delivery vehicle or expression system.

[0898] In an aspect, the invention relates to a method as described herein, comprising optimization of ω RNA specificity at the population level. Preferably, said optimization of ω RNA specificity comprises minimizing ω RNA target site sequence variation across a population and/or minimizing ω RNA off-target incidence across a population.

[0899] In one embodiment, optimization can result in selection of a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease that is naturally occurring or is modified. In one embodiment, optimization can result in selection of a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease that has nuclease, nickase, deaminase, transposase, and/or has one or more effector functionalities deactivated or eliminated. In one embodiment, optimizing a TAM specificity can include selecting a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease with a modified TAM specificity. In one embodiment, optimizing can include selecting a IscB polypeptide or CRISPR-associated IscB polypeptide nuclease having a minimal size. In an embodiment, optimizing effector protein stability comprises selecting an effector protein having a short half-life while maintaining sufficient activity, such as by selecting an appropriate IscB polypeptide nuclease orthologue having a specific half-life or stability. In an embodiment, optimizing immunogenicity or toxicity comprises minimizing effector protein immunogenicity or toxicity by protein modifications. In an embodiment, optimizing functional specific comprises selecting a protein effector with reduced tolerance of mismatches and/or bulges between the guide RNA and one or more target loci.

[0900] In an embodiment, optimizing efficacy comprises optimizing overall efficiency, epigenetic tolerance, or both. In an embodiment, maximizing overall efficiency comprises

selecting an effector protein with uniform enzyme activity across target loci with varying chromatin complexity, selecting an effector protein with enzyme activity limited to areas of open chromatin accessibility. In an embodiment, chromatin accessibility is measured using one or more of ATAC-seq, or a DNA-proximity ligation assay. In an embodiment, optimizing epigenetic tolerance comprises optimizing methylation tolerance, epigenetic mark competition, or both. In an embodiment, optimizing methylation tolerance comprises selecting an effector protein that modify methylated DNA. In an embodiment, optimizing epigenetic tolerance comprises selecting an effector protein unable to modify silenced regions of a chromosome, selecting an effector protein able to modify silenced regions of a chromosome, or selecting target loci not enriched for epigenetic markers

[0901] In an embodiment, selecting an optimized guide RNA comprises optimizing gRNA stability, gRNA immunogenicity, or both, or other gRNA associated parameters or variables as described herein elsewhere.

[0902] In an embodiment, optimizing gRNA stability and/or gRNA immunogenicity comprises RNA modification, or other gRNA associated parameters or variables as described herein elsewhere. In an embodiment, the modification comprises removing 1-3 nucleotides from the 3' end of a target complementarity region of the gRNA. In an embodiment, modification comprises an extended gRNA and/or trans RNA/DNA element that create stable structures in the gRNA that compete with gRNA base pairing at a target of off-target loci, or extended complimentary nucleotides between the gRNA and target sequence, or both.

[0903] In an embodiment, the mode of delivery comprises delivering gRNA and/or IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, delivering gRNA and/or IscB polypeptide or CRISPR-associated IscB polypeptide nuclease mRNA, or delivery gRNA and/or IscB polypeptide nuclease as a DNA based expression system. In an embodiment, the mode of delivery further comprises selecting a delivery vehicle and/or expression systems from the group consisting of liposomes, lipid particles, nanoparticles, biolistics, or viral-based expression/delivery systems. In an embodiment, expression is spatiotemporal expression is optimized by choice of conditional and/or inducible expression systems, including controllable IscB polypeptide nuclease activity optionally a destabilized IscB polypeptide or CRISPR-associated IscB polypeptide or CRISPR-associated IscB polypeptide and/or a split IscB polypeptide nuclease, and/or cell- or tissue-specific expression system.

[0904] The methods as described herein may further involve selection of the mode of delivery. In an embodiment, ω RNA and/or IscB polypeptide or CRISPR-associated IscB polypeptide nuclease are or are to be delivered. In an embodiment, ω RNA and/or IscB polypeptide or CRISPR-associated IscB polypeptide nuclease mRNA are or are to be delivered. In an embodiment, ω RNA and/or IscB polypeptide or CRISPR-associated IscB polypeptide nuclease provided in a DNA-based expression system are or are to be delivered. In an embodiment, delivery of the individual system components comprises a combination of the above modes of delivery. In an embodiment, delivery comprises delivering ω RNA and/or IscB polypeptide or CRISPR-associated IscB polypeptide nuclease protein, delivering ω RNA and/or IscB polypeptide or CRISPR-associated IscB polypeptide nuclease mRNA, or delivering ω RNA and/or IscB polypeptide or CRISPR-associated IscB polypeptide nuclease as a DNA based expression system.

[0905] The methods as described herein may further involve selection of the composition delivery vehicle and/or expression system. Delivery vehicles and expression systems are described herein elsewhere. By means of example, delivery vehicles of nucleic acids and/or proteins include nanoparticles, liposomes, etc. Delivery vehicles for DNA, such as DNA-based expression systems include for instance biolistics, viral based vector systems (e.g. adenoviral, AAV, lentiviral), etc. the skilled person will understand that selection of the mode of delivery, as well as delivery vehicle or expression system may depend on for instance the cell or tissues to be targeted. In an embodiment, the delivery vehicle and/or expression system for delivering the compositions, systems, or components thereof comprises liposomes, lipid particles, nanoparticles, biolistics, or viral-based expression/delivery systems.

Considerations for Therapeutic Applications

[0906] A consideration in genome editing therapy is the choice of sequence-specific nuclease, such as a variant of a IscB polypeptide nuclease. Each nuclease variant may possess its own unique set of strengths and weaknesses, many of which must be balanced in the context of treatment to maximize therapeutic benefit. For a specific editing therapy to be efficacious, a sufficiently high level of modification must be achieved in target cell populations to reverse disease symptoms. This therapeutic modification ‘threshold’ is determined by the fitness of edited cells following treatment and the amount of gene product necessary to reverse symptoms. With regard to fitness, editing creates three potential outcomes for treated cells

relative to their unedited counterparts: increased, neutral, or decreased fitness. In the case of increased fitness, corrected cells may be able and expand relative to their diseased counterparts to mediate therapy. In this case, where edited cells possess a selective advantage, even low numbers of edited cells can be amplified through expansion, providing a therapeutic benefit to the patient. Where the edited cells possess no change in fitness, an increase the therapeutic modification threshold can be warranted. As such, significantly greater levels of editing may be needed to treat diseases, where editing creates a neutral fitness advantage, relative to diseases where editing creates increased fitness for target cells. If editing imposes a fitness disadvantage, as would be the case for restoring function to a tumor suppressor gene in cancer cells, modified cells would be outcompeted by their diseased counterparts, causing the benefit of treatment to be low relative to editing rates. This may be overcome with supplemental therapies to increase the potency and/or fitness of the edited cells relative to the diseased counterparts.

[0907] In addition to cell fitness, the amount of gene product necessary to treat disease can also influence the minimal level of therapeutic genome editing that can treat or prevent a disease or a symptom thereof. In cases where a small change in the gene product levels can result in significant changes in clinical outcome, the minimal level of therapeutic genome editing is less relative to cases where a larger change in the gene product levels are needed to gain a clinically relevant response. In one embodiment, the minimal level of therapeutic genome editing can range from 0.1 to 1 %, 1-5%, 5-10%, 10-15%, 15-20%, 20-25%, 25-30%, 30-35%, 35-40%, 40-45%, 45-50%, or 50-55%. Thus, where a small change in gene product levels can influence clinical outcomes and diseases where there is a fitness advantage for edited cells, are ideal targets for genome editing therapy, as the therapeutic modification threshold is low enough to permit a high chance of success.

[0908] The activity of NHEJ and HDR DSB repair can vary by cell type and cell state. NHEJ is not highly regulated by the cell cycle and is efficient across cell types, allowing for high levels of gene disruption in accessible target cell populations. In contrast, HDR acts primarily during S/G2 phase, and is therefore restricted to cells that are actively dividing, limiting treatments that require precise genome modifications to mitotic cells [Ciccia, A. & Elledge, S.J. *Molecular cell* 40, 179-204 (2010); Chapman, J.R., et al. *Molecular cell* 47, 497-510 (2012)].

[0909] The efficiency of correction via HDR may be controlled by the epigenetic state or sequence of the targeted locus, or the specific repair template configuration (single vs. double stranded, long vs. short homology arms) used [Hacein-Bey-Abina, S., et al. The New England journal of medicine 346, 1185-1193 (2002); Gaspar, H.B., et al. Lancet 364, 2181-2187 (2004); Beumer, K.J., et al. G3 (2013)]. The relative activity of NHEJ and HDR machineries in target cells may also affect gene correction efficiency, as these pathways may compete to resolve DSBs [Beumer, K.J., et al. Proceedings of the National Academy of Sciences of the United States of America 105, 19821-19826 (2008)]. HDR also imposes a delivery challenge not seen with NHEJ strategies, as it uses the concurrent delivery of nucleases and repair templates. Thus, these differences can be kept in mind when designing, optimizing, and/or selecting a IscB polypeptide nuclease based therapeutic as described in greater detail elsewhere herein.

[0910] IscB polypeptide or CRISPR-associated IscB polypeptide nuclease-based polynucleotide modification application can include combinations of proteins, small RNA molecules, and/or repair templates, and can make, In one embodiment, delivery of these multiple parts substantially more challenging than, for example, traditional small molecule therapeutics. Two main strategies for delivery of compositions, systems, and components thereof have been developed: *ex vivo* and *in vivo*. In one embodiment of *ex vivo* treatments, diseased cells are removed from a subject, edited and then transplanted back into the patient. In other embodiments, cells from a healthy allogeneic donor are collected, modified using a composition or component thereof, to impart various functionalities and/or reduce immunogenicity, and administered to an allogeneic recipient in need of treatment. *Ex vivo* editing has the advantage of allowing the target cell population to be well defined and the specific dosage of therapeutic molecules delivered to cells to be specified. The latter consideration may be particularly important when off-target modifications are a concern, as titrating the amount of nuclease may decrease such mutations (Hsu et al., 2013). Another advantage of *ex vivo* approaches is the typically high editing rates that can be achieved, due to the development of efficient delivery systems for proteins and nucleic acids into cells in culture for research and gene therapy applications.

[0911] *In vivo* polynucleotide modification via compositions, systems, and/or components thereof involves direct delivery of the compositions, systems, and/or components thereof to cell types in their native tissues. *In vivo* polynucleotide modification via compositions, systems, and/or components thereof allows diseases in which the affected cell population is not

amenable to *ex vivo* manipulation to be treated. Furthermore, delivering compositions, systems, and/or components thereof to cells in situ allows for the treatment of multiple tissue and cell types.

[0912] In one embodiment, such as those where viral vector systems are used to generate viral particles to deliver the composition and/or component thereof to a cell, the total cargo size of the composition and/or component thereof should be considered as vector systems can have limits on the size of a polynucleotide that can be expressed therefrom and/or packaged into cargo inside of a viral particle. In one embodiment, the tropism of a vector system, such as a viral vector system, should be considered as it can impact the cell type to which the composition or component thereof can be efficiently and/or effectively delivered.

[0913] When delivering a system or component thereof via a viral-based system, it can be important to consider the amount of viral particles that will be needed to achieve a therapeutic effect so as to account for the potential immune response that can be elicited by the viral particles when delivered to a subject or cell(s). When delivering a system or component thereof via a viral based system, it can be important to consider mechanisms of controlling the distribution and/or dosage of the system *in vivo*. Generally, to reduce the potential for off-target effects, it is optimal but not necessarily required, that the amount of the system be as close to the minimum or least effective dose. In practice this can be challenging to do.

[0914] In one embodiment, it can be important to consider the immunogenicity of the system or component thereof. In embodiments, where the immunogenicity of the system or component thereof is of concern, the immunogenicity system or component thereof can be reduced. By way of example only, the immunogenicity of the system or component thereof can be reduced using the approach set out in Tangri et al. Accordingly, directed evolution or rational design may be used to reduce the immunogenicity of the IscB polypeptide nuclease in the host species (human or other species).

Xenotransplantation

[0915] The present invention also contemplates use of the composition described herein, e.g. IscB polypeptide nuclease protein systems, to provide RNA-guided DNA nucleases adapted to be used to provide modified tissues for transplantation. For example, RNA-guided DNA nucleases may be used to knockout, knockdown or disrupt selected genes in an animal, such as a transgenic pig (such as the human heme oxygenase-1 transgenic pig line), for example by disrupting expression of genes that encode epitopes recognized by the human immune

system, i.e. xenoantigen genes. Candidate porcine genes for disruption may for example include $\alpha(1,3)$ -galactosyltransferase and cytidine monophosphate-N-acetylneuraminic acid hydroxylase genes (see PCT Patent Publication WO 2014/066505). In addition, genes encoding endogenous retroviruses may be disrupted, for example the genes encoding all porcine endogenous retroviruses (see Yang et al., 2015, Genome-wide inactivation of porcine endogenous retroviruses (PERVs), *Science* 27 November 2015: Vol. 350 no. 6264 pp. 1101-1104). In addition, RNA-guided DNA nucleases may be used to target a site for integration of additional genes in xenotransplant donor animals, such as a human CD55 gene to improve protection against hyperacute rejection.

[0916] Embodiments of the invention also relate to methods and compositions related to knocking out genes, amplifying genes and repairing particular mutations associated with DNA repeat instability and neurological disorders (Robert D. Wells, Tetsuo Ashizawa, *Genetic Instabilities and Neurological Diseases*, Second Edition, Academic Press, Oct 13, 2011 – Medical). Specific aspects of tandem repeat sequences have been found to be responsible for more than twenty human diseases (New insights into repeat instability: role of RNA•DNA hybrids. McIvor EI, Polak U, Napierala M. *RNA Biol.* 2010 Sep-Oct;7(5):551-8). The present effector protein systems may be harnessed to correct these defects of genomic instability.

[0917] Several further aspects of the invention relate to correcting defects associated with a wide range of genetic diseases which are further described on the website of the National Institutes of Health under the topic subsection Genetic Disorders (website at health.nih.gov/topic/GeneticDisorders). The genetic brain diseases may include but are not limited to Adrenoleukodystrophy, Agenesis of the Corpus Callosum, Aicardi Syndrome, Alpers' Disease, Alzheimer's Disease, Barth Syndrome, Batten Disease, CADASIL, Cerebellar Degeneration, Fabry's Disease, Gerstmann-Straussler-Scheinker Disease, Huntington's Disease and other Triplet Repeat Disorders, Leigh's Disease, Lesch-Nyhan Syndrome, Menkes Disease, Mitochondrial Myopathies and NINDS Colpocephaly. These diseases are further described on the website of the National Institutes of Health under the subsection Genetic Brain Disorders.

APPLICATIONS IN PLANTS AND FUNGI

[0918] The compositions, systems, and methods described herein can be used to perform gene or genome interrogation or editing or manipulation in plants and fungi. For example, the

applications include investigation and/or selection and/or interrogations and/or comparison and/or manipulations and/or transformation of plant genes or genomes; e.g., to create, identify, develop, optimize, or confer trait(s) or characteristic(s) to plant(s) or to transform a plant or fungus genome. There can accordingly be improved production of plants, new plants with new combinations of traits or characteristics or new plants with enhanced traits. The compositions, systems, and methods can be used with regard to plants in Site-Directed Integration (SDI) or Gene Editing (GE) or any Near Reverse Breeding (NRB) or Reverse Breeding (RB) techniques.

[0919] The compositions, systems, and methods herein may be used to confer desired traits (e.g., enhanced nutritional quality, increased resistance to diseases and resistance to biotic and abiotic stress, and increased production of commercially valuable plant products or heterologous compounds) on essentially any plants and fungi, and their cells and tissues. The compositions, systems, and methods may be used to modify endogenous genes or to modify their expression without the permanent introduction into the genome of any foreign gene.

[0920] In one embodiment, compositions, systems, and methods may be used in genome editing in plants or where RNAi or similar genome editing techniques have been used previously; see, e.g., Nekrasov, “Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR-Cas system,” *Plant Methods* 2013, 9:39 (doi:10.1186/1746-4811-9-39); Brooks, “Efficient gene editing in tomato in the first generation using the CRISPR-Cas9 system,” *Plant Physiology* September 2014 pp 114.247577; Shan, “Targeted genome modification of crop plants using a CRISPR-Cas system,” *Nature Biotechnology* 31, 686-688 (2013); Feng, “Efficient genome editing in plants using a CRISPR/Cas system,” *Cell Research* (2013) 23:1229–1232. doi:10.1038/cr.2013.114; published online 20 August 2013; Xie, “RNA-guided genome editing in plants using a CRISPR-Cas system,” *Mol Plant*. 2013 Nov;6(6):1975-83. doi: 10.1093/mp/sst119. Epub 2013 Aug 17; Xu, “Gene targeting using the *Agrobacterium tumefaciens*-mediated CRISPR-Cas system in rice,” *Rice* 2014, 7:5 (2014), Zhou et al., “Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial *Populus* reveals 4-coumarate: CoA ligase specificity and Redundancy,” *New Phytologist* (2015) (Forum) 1-4 (available online only at www.newphytologist.com); Caliendo et al, “Targeted DNA degradation using a CRISPR device stably carried in the host genome, *NATURE COMMUNICATIONS* 6:6989, DOI: 10.1038/ncomms7989, www.nature.com/naturecommunications DOI: 10.1038/ncomms7989; US Patent No. 6,603,061 - *Agrobacterium*-Mediated Plant Transformation Method; US Patent

No. 7,868,149 - Plant Genome Sequences and Uses Thereof and US 2009/0100536 - Transgenic Plants with Enhanced Agronomic Traits, Morrell et al “Crop genomics: advances and applications,” Nat Rev Genet. 2011 Dec 29;13(2):85-96, all the contents and disclosure of each of which are herein incorporated by reference in their entirety. Aspects of utilizing the compositions, systems, and methods may be analogous to the use of the composition in plants, and mention is made of the University of Arizona website “CRISPR-PLANT” (genome.arizona.edu/crispr/) (supported by Penn State and AGI).

[0921] The compositions, systems, and methods may also be used on protoplasts. A “protoplast” refers to a plant cell that has had its protective cell wall completely or partially removed using, for example, mechanical or enzymatic means resulting in an intact biochemical competent unit of living plant that can reform their cell wall, proliferate and regenerate grow into a whole plant under proper growing conditions.

[0922] The compositions, systems, and methods may be used for screening genes (e.g., endogenous, mutations) of interest. In some examples, genes of interest include those encoding enzymes involved in the production of a component of added nutritional value or generally genes affecting agronomic traits of interest, across species, phyla, and plant kingdom. By selectively targeting e.g. genes encoding enzymes of metabolic pathways, the genes responsible for certain nutritional aspects of a plant can be identified. Similarly, by selectively targeting genes which may affect a desirable agronomic trait, the relevant genes can be identified. Accordingly, the present invention encompasses screening methods for genes encoding enzymes involved in the production of compounds with a particular nutritional value and/or agronomic traits.

[0923] It is also understood that reference herein to animal cells may also apply, *mutatis mutandis*, to plant or fungal cells unless otherwise apparent; and, the enzymes herein having reduced off-target effects and systems employing such enzymes can be used in plant applications, including those mentioned herein.

[0924] In some cases, nucleic acids introduced to plants and fungi may be codon optimized for expression in the plants and fungi. Methods of codon optimization include those described in Kwon KC, et al., Codon Optimization to Enhance Expression Yields Insights into Chloroplast Translation, Plant Physiol. 2016 Sep;172(1):62-77.

[0925] The components (e.g., IscB polypeptide nuclease) in the compositions and systems may further comprise one or more functional domains described herein. In some examples, the

functional domains may be an exonuclease. Such exonuclease may increase the efficiency of the IscB polypeptide nuclease' function, e.g., mutagenesis efficiency. An example of the functional domain is Trex2, as described in Weiss T et al., www.biorxiv.org/content/10.1101/2020.04.11.037572v1, doi: doi.org/10.1101/2020.04.11.037572.

Examples of plants

[0926] The compositions, systems, and methods herein can be used to confer desired traits on essentially any plant. A wide variety of plants and plant cell systems may be engineered for the desired physiological and agronomic characteristics. In general, the term “plant” relates to any various photosynthetic, eukaryotic, unicellular or multicellular organism of the kingdom Plantae characteristically growing by cell division, containing chloroplasts, and having cell walls comprised of cellulose. The term plant encompasses monocotyledonous and dicotyledonous plants.

[0927] The compositions, systems, and methods may be used over a broad range of plants, such as for example with dicotyledonous plants belonging to the orders *Magnoliales*, *Illiciales*, *Laurales*, *Piperales*, *Aristolochiales*, *Nymphaeales*, *Ranunculales*, *Papaverales*, *Sarraceniales*, *Trochodendrales*, *Hamamelidales*, *Eucomiales*, *Leitneriales*, *Myricales*, *Fagales*, *Casuarinales*, *Caryophyllales*, *Batales*, *Polygonales*, *Plumbaginales*, *Dilleniales*, *Theales*, *Malvales*, *Urticales*, *Lecythidales*, *Violales*, *Salicales*, *Capparales*, *Ericales*, *Diapensales*, *Ebenales*, *Primulales*, *Rosales*, *Fabales*, *Podostemales*, *Haloragales*, *Myrtales*, *Cornales*, *Proteales*, *Santalales*, *Rafflesiales*, *Celastrales*, *Euphorbiales*, *Rhamnales*, *Sapindales*, *Juglandales*, *Geraniales*, *Polygalales*, *Umbellales*, *Gentianales*, *Polemoniales*, *Lamiales*, *Plantaginales*, *Scrophulariales*, *Campanulales*, *Rubiales*, *Dipsacales*, and *Asterales*; monocotyledonous plants such as those belonging to the orders *Alismatales*, *Hydrocharitales*, *Najadales*, *Triuridales*, *Commelinales*, *Eriocaulales*, *Restionales*, *Poales*, *Juncals*, *Cyperales*, *Typhales*, *Bromeliales*, *Zingiberales*, *Arecales*, *Cyclanthales*, *Pandanales*, *Arales*, *Lilliales*, and *Orchidales*, or with plants belonging to *Gymnospermae*, e.g., those belonging to the orders *Pinales*, *Ginkgoales*, *Cycadales*, *Araucariales*, *Cupressales* and *Gnetales*.

[0928] The compositions, systems, and methods herein can be used over a broad range of plant species, included in the non-limitative list of dicot, monocot or gymnosperm genera hereunder: *Atropa*, *Alseodaphne*, *Anacardium*, *Arachis*, *Beilschmiedia*, *Brassica*, *Carthamus*,

Cocculus, Croton, Cucumis, Citrus, Citrullus, Capsicum, Catharanthus, Cocos, Coffea, Cucurbita, Daucus, Duguetia, Eschscholzia, Ficus, Fragaria, Glaucium, Glycine, Gossypium, Helianthus, Hevea, Hyoscyamus, Lactuca, Landolphia, Linum, Litsea, Lycopersicon, Lupinus, Manihot, Majorana, Malus, Medicago, Nicotiana, Olea, Parthenium, Papaver, Persea, Phaseolus, Pistacia, Pisum, Pyrus, Prunus, Raphanus, Ricinus, Senecio, Sinomenium, Stephania, Sinapis, Solanum, Theobroma, Trifolium, Trigonella, Vicia, Vinca, Vilis, and Vigna; and the genera *Allium, Andropogon, Aragrostis, Asparagus, Avena, Cynodon, Elaeis, Festuca, Festulolium, Heterocallis, Hordeum, Lemna, Lolium, Musa, Oryza, Panicum, Pannisetum, Phleum, Poa, Secale, Sorghum, Triticum, Zea, Abies, Cunninghamia, Ephedra, Picea, Pinus, and Pseudotsuga*.

[0929] In one embodiment, target plants and plant cells for engineering include those monocotyledonous and dicotyledonous plants, such as crops including grain crops (e.g., wheat, maize, rice, millet, barley), fruit crops (e.g., tomato, apple, pear, strawberry, orange), forage crops (e.g., alfalfa), root vegetable crops (e.g., carrot, potato, sugar beets, yam), leafy vegetable crops (e.g., lettuce, spinach); flowering plants (e.g., petunia, rose, chrysanthemum), conifers and pine trees (e.g., pine fir, spruce); plants used in phytoremediation (e.g., heavy metal accumulating plants); oil crops (e.g., sunflower, rape seed) and plants used for experimental purposes (e.g., *Arabidopsis*). Specifically, the plants are intended to comprise without limitation angiosperm and gymnosperm plants such as acacia, alfalfa, amaranth, apple, apricot, artichoke, ash tree, asparagus, avocado, banana, barley, beans, beet, birch, beech, blackberry, blueberry, broccoli, Brussel's sprouts, cabbage, canola, cantaloupe, carrot, cassava, cauliflower, cedar, a cereal, celery, chestnut, cherry, Chinese cabbage, citrus, clementine, clover, coffee, corn, cotton, cowpea, cucumber, cypress, eggplant, elm, endive, eucalyptus, fennel, figs, fir, geranium, grape, grapefruit, groundnuts, ground cherry, gum hemlock, hickory, kale, kiwifruit, kohlrabi, larch, lettuce, leek, lemon, lime, locust, pine, maidenhair, maize, mango, maple, melon, millet, mushroom, mustard, nuts, oak, oats, oil palm, okra, onion, orange, an ornamental plant or flower or tree, papaya, palm, parsley, parsnip, pea, peach, peanut, pear, peat, pepper, persimmon, pigeon pea, pine, pineapple, plantain, plum, pomegranate, potato, pumpkin, radicchio, radish, rapeseed, raspberry, rice, rye, sorghum, safflower, sallow, soybean, spinach, spruce, squash, strawberry, sugar beet, sugarcane, sunflower, sweet potato, sweet corn, tangerine, tea, tobacco, tomato, trees, triticale, turf grasses, turnips, vine, walnut, watercress, watermelon, wheat, yams, yew, and zucchini.

[0930] The term plant also encompasses Algae, which are mainly photoautotrophs unified primarily by their lack of roots, leaves and other organs that characterize higher plants. The compositions, systems, and methods can be used over a broad range of "algae" or "algae cells." Examples of algae include eukaryotic phyla, including the *Rhodophyta* (red algae), *Chlorophyta* (green algae), *Phaeophyta* (brown algae), *Bacillariophyta* (diatoms), *Eustigmatophyta* and dinoflagellates as well as the prokaryotic phylum *Cyanobacteria* (blue-green algae). Examples of algae species include those of *Amphora*, *Anabaena*, *Anikstrodesmis*, *Botryococcus*, *Chaetoceros*, *Chlamydomonas*, *Chlorella*, *Chlorococcum*, *Cyclotella*, *Cylindrotheca*, *Dunaliella*, *Emiliana*, *Euglena*, *Hematococcus*, *Isochrysis*, *Monochrysis*, *Monoraphidium*, *Nannochloris*, *Nannochloropsis*, *Navicula*, *Nephrochloris*, *Nephroselmis*, *Nitzschia*, *Nodularia*, *Nostoc*, *Oochromonas*, *Oocystis*, *Oscillatoria*, *Pavlova*, *Phaeodactylum*, *Playtmonas*, *Pleurochrysis*, *Porphyra*, *Pseudoanabaena*, *Pyramimonas*, *Stichococcus*, *Synechococcus*, *Synechocystis*, *Tetraselmis*, *Thalassiosira*, and *Trichodesmium*.

Plant promoters

[0931] In order to ensure appropriate expression in a plant cell, the components of the components and systems herein may be placed under control of a plant promoter. A plant promoter is a promoter operable in plant cells. A plant promoter is capable of initiating transcription in plant cells, whether or not its origin is a plant cell. The use of different types of promoters is envisaged.

[0932] In some examples, the plant promoter is a constitutive plant promoter, which is a promoter that is able to express the open reading frame (ORF) that it controls in all or nearly all of the plant tissues during all or nearly all developmental stages of the plant (referred to as "constitutive expression"). One example of a constitutive promoter is the cauliflower mosaic virus 35S promoter. In some examples, the plant promoter is a regulated promoter, which directs gene expression not constitutively, but in a temporally- and/or spatially-regulated manner, and includes tissue-specific, tissue-preferred and inducible promoters. Different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental conditions. In some examples, the plant promoter is a tissue-preferred promoters, which can be utilized to target enhanced expression in certain cell types within a particular plant tissue, for instance vascular cells in leaves or roots or in specific cells of the seed.

[0933] Exemplary plant promoters include those obtained from plants, plant viruses, and bacteria such as *Agrobacterium* or *Rhizobium* which comprise genes expressed in plant cells. Additional examples of promoters include those described in Kawamata et al., (1997) *Plant Cell Physiol* 38:792-803; Yamamoto et al., (1997) *Plant J* 12:255-65; Hire et al, (1992) *Plant Mol Biol* 20:207-18, Kuster et al, (1995) *Plant Mol Biol* 29:759-72, and Capana et al., (1994) *Plant Mol Biol* 25:681 -91.

[0934] In some examples, a plant promoter may be an inducible promoter, which is inducible and allows for spatiotemporal control of gene editing or gene expression may use a form of energy. The form of energy may include sound energy, electromagnetic radiation, chemical energy and/or thermal energy. Examples of inducible systems include tetracycline inducible promoters (Tet-On or Tet-Off), small molecule two-hybrid transcription activations systems (FKBP, ABA, etc.), or light inducible systems (Phytochrome, LOV domains, or cryptochrome), such as a Light Inducible Transcriptional Effector (LITE) that direct changes in transcriptional activity in a sequence-specific manner. In a particular example, of the components of a light inducible system include a IscB polypeptide nuclease, a light-responsive cytochrome heterodimer (e.g. from *Arabidopsis thaliana*), and a transcriptional activation/repression domain.

[0935] In some examples, the promoter may be a chemical-regulated promoter (where the application of an exogenous chemical induces gene expression) or a chemical-repressible promoter (where application of the chemical represses gene expression). Examples of chemical-inducible promoters include maize In2-2 promoter (activated by benzene sulfonamide herbicide safeners), the maize GST promoter (activated by hydrophobic electrophilic compounds used as pre-emergent herbicides), the tobacco PR-1 a promoter (activated by salicylic acid), promoters regulated by antibiotics (such as tetracycline-inducible and tetracycline-repressible promoters).

Stable integration in the genome of plants

[0936] In one embodiment, polynucleotides encoding the components of the compositions and systems may be introduced for stable integration into the genome of a plant cell. In some cases, vectors or expression systems may be used for such integration. The design of the vector or the expression system can be adjusted depending on for when, where and under what conditions the guide RNA and/or the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease gene are expressed. In some cases, the polynucleotides may be integrated into an

organelle of a plant, such as a plastid, mitochondrion or a chloroplast. The elements of the expression system may be on one or more expression constructs which are either circular such as a plasmid or transformation vector, or non-circular such as linear double stranded DNA.

[0937] In one embodiment, the method of integration generally comprises the steps of selecting a suitable host cell or host tissue, introducing the construct(s) into the host cell or host tissue, and regenerating plant cells or plants therefrom. In some examples, the expression system for stable integration into the genome of a plant cell may contain one or more of the following elements: a promoter element that can be used to express the RNA and/or IscB polypeptide nuclease in a plant cell; a 5' untranslated region to enhance expression; an intron element to further enhance expression in certain cells, such as monocot cells; a multiple-cloning site to provide convenient restriction sites for inserting the guide RNA and/or the IscB polypeptide nuclease gene sequences and other desired elements; and a 3' untranslated region to provide for efficient termination of the expressed transcript.

Transient expression in plants

[0938] In one embodiment, the components of the compositions and systems may be transiently expressed in the plant cell. In some examples, the compositions and systems may modify a target nucleic acid only when both the guide RNA and the IscB polypeptide or CRISPR-associated IscB polypeptide or CRISPR-associated IscB polypeptide nuclease are present in a cell, such that genomic modification can further be controlled. As the expression of the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease is transient, plants regenerated from such plant cells typically contain no foreign DNA. In certain examples, the IscB polypeptide or CRISPR-associated IscB polypeptide is stably expressed and the guide sequence is transiently expressed.

[0939] DNA and/or RNA (e.g., mRNA) may be introduced to plant cells for transient expression. In such cases, the introduced nucleic acid may be provided in sufficient quantity to modify the cell but do not persist after a contemplated period of time has passed or after one or more cell divisions.

[0940] The transient expression may be achieved using suitable vectors. Exemplary vectors that may be used for transient expression include a pEAQ vector (may be tailored for Agrobacterium-mediated transient expression) and Cabbage Leaf Curl virus (CaLCuV), and vectors described in Sainsbury F. et al., *Plant Biotechnol J.* 2009 Sep;7(7):682-93; and Yin K et al., *Scientific Reports* volume 5, Article number: 14926 (2015).

[0941] Combinations of the different methods described above are also envisaged.

Translocation to and/or expression in specific plant organelles

[0942] The compositions and systems herein may comprise elements for translocation to and/or expression in a specific plant organelle.

Chloroplast targeting

[0943] In one embodiment, it is envisaged that the compositions and systems are used to specifically modify chloroplast genes or to ensure expression in the chloroplast. The compositions and systems (e.g., IscB polypeptide or CRISPR-associated IscB polypeptide nuclease, ωRNAs, or their encoding polynucleotides) may be transformed, compartmentalized, and/or targeted to the chloroplast. In an example, the introduction of genetic modifications in the plastid genome can reduce biosafety issues such as gene flow through pollen.

[0944] Examples of methods of chloroplast transformation include Particle bombardment, PEG treatment, and microinjection, and the translocation of transformation cassettes from the nuclear genome to the plastid. In some examples, targeting of chloroplasts may be achieved by incorporating in chloroplast localization sequence, and/or the expression construct a sequence encoding a chloroplast transit peptide (CTP) or plastid transit peptide, operably linked to the 5' region of the sequence encoding the components of the compositions and systems. Additional examples of transforming, targeting and localization of chloroplasts include those described in WO2010061186, Protein Transport into Chloroplasts, 2010, Annual Review of Plant Biology, Vol. 61: 157-180, and US 20040142476, which are incorporated by reference herein in their entireties.

Exemplary applications in plants

[0945] The compositions, systems, and methods may be used to generate genetic variation(s) in a plant (e.g., crop) of interest. One or more, e.g., a library of, ωRNAs targeting one or more locations in a genome may be provided and introduced into plant cells together with the IscB polypeptide nuclease. For example, a collection of genome-scale point mutations and gene knock-outs can be generated. In some examples, the compositions, systems, and methods may be used to generate a plant part or plant from the cells so obtained and screening the cells for a trait of interest. The target genes may include both coding and non-coding regions. In some cases, the trait is stress tolerance and the method is a method for the generation of stress-tolerant crop varieties.

[0946] In one embodiment, the compositions, systems, and methods are used to modify endogenous genes or to modify their expression. The expression of the components may induce targeted modification of the genome, either by direct activity of the IscB polypeptide or CRISPR-associated IscB polypeptide nuclease and optionally introduction of recombination template DNA, or by modification of genes targeted. The different strategies described herein above allow IscB polypeptide or CRISPR-associated IscB polypeptide nuclease-mediated targeted genome editing without requiring the introduction of the components into the plant genome.

[0947] In some cases, the modification may be performed without the permanent introduction into the genome of the plant of any foreign gene, including those encoding components of the composition herein, so as to avoid the presence of foreign DNA in the genome of the plant. This can be of interest as the regulatory requirements for non-transgenic plants are less rigorous. Components which are transiently introduced into the plant cell are typically removed upon crossing.

[0948] For example, the modification may be performed by transient expression of the components of the compositions and systems. The transient expression may be performed by delivering the components of the compositions and systems with viral vectors, delivery into protoplasts, with the aid of particulate molecules such as nanoparticles or CPPs.

Generation of plants with desired traits

[0949] The compositions, systems, and methods herein may be used to introduce desired traits to plants. The approaches include introduction of one or more foreign genes to confer a trait of interest, editing or modulating endogenous genes to confer a trait of interest.

Agronomic traits

[0950] In one embodiment, crop plants can be improved by influencing specific plant traits. Examples of the traits include improved agronomic traits such as herbicide resistance, disease resistance, abiotic stress tolerance, high yield, and superior quality, pesticide-resistance, disease resistance, insect and nematode resistance, resistance against parasitic weeds, drought tolerance, nutritional value, stress tolerance, self-pollination avoidance, forage digestibility biomass, and grain yield.

[0951] In one embodiment, genes that confer resistance to pests or diseases may be introduced to plants. In cases there are endogenous genes that confer such resistance in a plants,

their expression and function may be enhanced (e.g., by introducing extra copies, modifications that enhance expression and/or activity).

[0952] Examples of genes that confer resistance include plant disease resistance genes (e.g., Cf- 9, Pto, RSP2, SIDMR6-1), genes conferring resistance to a pest (e.g., those described in WO96/30517), *Bacillus thuringiensis* proteins, lectins, Vitamin-binding proteins (e.g., avidin), enzyme inhibitors (e.g., protease or proteinase inhibitors or amylase inhibitors), insect-specific hormones or pheromones (e.g., ecdysteroid or a juvenile hormone, variant thereof, a mimetic based thereon, or an antagonist or agonist thereof) or genes involved in the production and regulation of such hormone and pheromones, insect-specific peptides or neuropeptide, Insect-specific venom (e.g., produced by a snake, a wasp, etc., or analog thereof), Enzymes responsible for a hyperaccumulation of a monoterpene, a sesquiterpene, a steroid, hydroxamic acid, a phenylpropanoid derivative or another nonprotein molecule with insecticidal activity, Enzymes involved in the modification of biologically active molecule (e.g., a glycolytic enzyme, a proteolytic enzyme, a lipolytic enzyme, a nuclease, a cyclase, a transaminase, an esterase, a hydrolase, a phosphatase, a kinase, a phosphorylase, a polymerase, an elastase, a chitinase and a glucanase, whether natural or synthetic), molecules that stimulates signal transduction, Viral-invasive proteins or a complex toxin derived therefrom, Developmental-arrestive proteins produced in nature by a pathogen or a parasite, a developmental-arrestive protein produced in nature by a plant, or any combination thereof.

[0953] The compositions, systems, and methods may be used to identify, screen, introduce or remove mutations or sequences lead to genetic variability that give rise to susceptibility to certain pathogens, e.g., host specific pathogens. Such approach may generate plants that are non-host resistance, e.g., the host and pathogen are incompatible or there can be partial resistance against all races of a pathogen, typically controlled by many genes and/or also complete resistance to some races of a pathogen but not to other races.

[0954] In one embodiment, the compositions, systems, and methods may be used to modify genes involved in plant diseases. Such genes may be removed, inactivated, or otherwise regulated or modified. Examples of plant diseases include those described in [0045]-[0080] of US20140213619A1, which is incorporated by reference herein in its entirety.

[0955] In one embodiment, genes that confer resistance to herbicides may be introduced to plants. Examples of genes that confer resistance to herbicides include genes conferring resistance to herbicides that inhibit the growing point or meristem, such as an imidazolinone

or a sulfonylurea, genes conferring glyphosate tolerance (e.g., resistance conferred by, e.g., mutant 5-enolpyruvylshikimate-3-phosphate synthase genes, *aroA* genes and glyphosate acetyl transferase (GAT) genes, respectively), or resistance to other phosphono compounds such as by glufosinate (phosphinothricin acetyl transferase (PAT) genes from *Streptomyces* species, including *Streptomyces hygroscopicus* and *Streptomyces viridichromogenes*), and to pyridinoxy or phenoxy propionic acids and cyclohexones by ACCase inhibitor-encoding genes), genes conferring resistance to herbicides that inhibit photosynthesis (such as a triazine (*psbA* and *gs+* genes) or a benzonitrile (nitrilase gene), and glutathione S-transferase), genes encoding enzymes detoxifying the herbicide or a mutant glutamine synthase enzyme that is resistant to inhibition, genes encoding a detoxifying enzyme is an enzyme encoding a phosphinothricin acetyltransferase (such as the *bar* or *pat* protein from *Streptomyces* species), genes encoding hydroxyphenylpyruvate dioxygenases (HPPD) inhibitors, e.g., naturally occurring HPPD resistant enzymes, and genes encoding a mutated or chimeric HPPD enzyme.

[0956] In one embodiment, genes involved in Abiotic stress tolerance may be introduced to plants. Examples of genes include those capable of reducing the expression and/or the activity of poly(ADP-ribose) polymerase (PARP) gene, transgenes capable of reducing the expression and/or the activity of the PARG encoding genes, genes coding for a plant-functional enzyme of the nicotinamide adenine dinucleotide salvage synthesis pathway including nicotinamidase, nicotinate phosphoribosyltransferase, nicotinic acid mononucleotide adenyl transferase, nicotinamide adenine dinucleotide synthetase or nicotine amide phosphorybosyltransferase, enzymes involved in carbohydrate biosynthesis, enzymes involved in the production of polyfructose (e.g., the inulin and levan-type), the production of alpha-1,6 branched alpha-1,4-glucans, the production of alternan, the production of hyaluronan.

[0957] In one embodiment, genes that improve drought resistance may be introduced to plants. Examples of genes Ubiquitin Protein Ligase protein (UPL) protein (UPL3), DR02, DR03, ABC transporter, and DREB1A.

Nutritionally improved plants

[0958] In one embodiment, the compositions, systems, and methods may be used to produce nutritionally improved plants. In some examples, such plants may provide functional foods, e.g., a modified food or food ingredient that may provide a health benefit beyond the traditional nutrients it contains. In certain examples, such plants may provide nutraceuticals foods, e.g., substances that may be considered a food or part of a food and provides health

benefits, including the prevention and treatment of disease. The nutraceutical foods may be useful in the prevention and/or treatment of diseases in animals and humans, e.g., cancers, diabetes, cardiovascular disease, and hypertension.

[0959] An improved plant may naturally produce one or more desired compounds and the modification may enhance the level or activity or quality of the compounds. In some cases, the improved plant may not naturally produce the compound(s), while the modification enables the plant to produce such compound(s). In some cases, the compositions, systems, and methods used to modify the endogenous synthesis of these compounds indirectly, e.g. by modifying one or more transcription factors that controls the metabolism of this compound.

[0960] Examples of nutritionally improved plants include plants comprising modified protein quality, content and/or amino acid composition, essential amino acid contents, oils and fatty acids, carbohydrates, vitamins and carotenoids, functional secondary metabolites, and minerals. In some examples, the improved plants may comprise or produce compounds with health benefits. Examples of nutritionally improved plants include those described in Newell-McGloughlin, *Plant Physiology*, July 2008, Vol. 147, pp. 939–953.

[0961] Examples of compounds that can be produced include carotenoids (e.g., α -Carotene or β -Carotene), lutein, lycopene, Zeaxanthin, Dietary fiber (e.g., insoluble fibers, β -Glucan, soluble fibers, fatty acids (e.g., ω -3 fatty acids, Conjugated linoleic acid, GLA,), Flavonoids (e.g., Hydroxycinnamates, flavonols, catechins and tannins), Glucosinolates, indoles, isothiocyanates (e.g., Sulforaphane), Phenolics (e.g., stilbenes, caffeic acid and ferulic acid, epicatechin), Plant stanols/sterols, Fructans, inulins, fructo-oligosaccharides, Saponins, Soybean proteins, Phytoestrogens (e.g., isoflavones, lignans), Sulfides and thiols such as diallyl sulphide, Allyl methyl trisulfide, dithiolthiones, Tannins, such as proanthocyanidins, or any combination thereof.

[0962] The compositions, systems, and methods may also be used to modify protein/starch functionality, shelf life, taste/aesthetics, fiber quality, and allergen, antinutrient, and toxin reduction traits.

[0963] Examples of genes and nucleic acids that can be modified to introduce the traits include stearyl-ACP desaturase, DNA associated with the single allele which may be responsible for maize mutants characterized by low levels of phytic acid, Tf RAP2.2 and its interacting partner SINAT2, Tf Dof1, and DOF Tf AtDof1.1 (OBP2).

Modification of polyploid plants

[0964] The compositions, systems, and methods may be used to modify polyploid plants. Polyploid plants carry duplicate copies of their genomes (e.g. as many as six, such as in wheat). In some cases, the compositions, systems, and methods may be/can be multiplexed to affect all copies of a gene, or to target dozens of genes at once. For instance, the compositions, systems, and methods may be used to simultaneously ensure a loss of function mutation in different genes responsible for suppressing defenses against a disease. The modification may be simultaneous suppression the expression of the TaMLO-A1, TaMLO-B1 and TaMLO-D1 nucleic acid sequence in a wheat plant cell and regenerating a wheat plant therefrom, in order to ensure that the wheat plant is resistant to powdery mildew (e.g., as described in WO2015109752).

Regulation of fruit-ripening

[0965] The compositions, systems, and methods may be used to regulate ripening of fruits. Ripening is a normal phase in the maturation process of fruits and vegetables. Only a few days after it starts it may render a fruit or vegetable inedible, which can bring significant losses to both farmers and consumers.

[0966] In one embodiment, the compositions, systems, and methods are used to reduce ethylene production. In some examples, the compositions, systems, and methods may be used to suppress the expression and/or activity of ACC synthase, insert a ACC deaminase gene or a functional fragment thereof, insert a SAM hydrolase gene or functional fragment thereof, suppress ACC oxidase gene expression

[0967] Alternatively or additionally, the compositions, systems, and methods may be used to modify ethylene receptors (e.g., suppressing ETR1) and/or Polygalacturonase (PG). Suppression of a gene may be achieved by introducing a mutation, an antisense sequence, and/or a truncated copy of the gene to the genome.

Increasing storage life of plants

[0968] In one embodiment, the compositions, systems, and methods are used to modify genes involved in the production of compounds which affect storage life of the plant or plant part. The modification may be in a gene that prevents the accumulation of reducing sugars in potato tubers. Upon high-temperature processing, these reducing sugars react with free amino acids, resulting in brown, bitter-tasting products and elevated levels of acrylamide, which is a potential carcinogen. In one embodiment, the methods provided herein are used to reduce or

inhibit expression of the vacuolar invertase gene (VInv), which encodes a protein that breaks down sucrose to glucose and fructose.

Reducing allergens in plants

[0969] In one embodiment, the compositions, systems, and methods are used to generate plants with a reduced level of allergens, making them safer for consumers. To this end, the compositions, systems, and methods may be used to identify and modify (e.g., suppress) one or more genes responsible for the production of plant allergens. Examples of such genes include Lol p5, as well as those in peanuts, soybeans, lentils, peas, lupin, green beans, mung beans, such as those described in Nicolaou et al., Current Opinion in Allergy and Clinical Immunology 2011;11(3):222), which is incorporated by reference herein in its entirety.

Generation of male sterile plants

[0970] The compositions, systems, and methods may be used to generate male sterile plants. Hybrid plants typically have advantageous agronomic traits compared to inbred plants. However, for self-pollinating plants, the generation of hybrids can be challenging. In different plant types (e.g., maize and rice), genes have been identified which are important for plant fertility, more particularly male fertility. Plants that are as such genetically altered can be used in hybrid breeding programs.

[0971] The compositions, systems, and methods may be used to modify genes involved male fertility, e.g., inactivating (such as by introducing mutations to) genes required for male fertility. Examples of the genes involved in male fertility include cytochrome P450-like gene (MS26) or the meganuclease gene (MS45), and those described in Wan X et al., Mol Plant. 2019 Mar 4;12(3):321-342; and Kim YJ, et al., Trends Plant Sci. 2018 Jan;23(1):53-65.

Increasing the fertility stage in plants

[0972] In one embodiment, the compositions, systems, and methods may be used to prolong the fertility stage of a plant such as of a rice. For instance, a rice fertility stage gene such as Ehd3 can be targeted in order to generate a mutation in the gene and plantlets can be selected for a prolonged regeneration plant fertility stage.

Production of early yield of products

[0973] In one embodiment, the compositions, systems, and methods may be used to produce early yield of the product. For example, flowering process may be modulated, e.g., by mutating flowering repressor gene such as SP5G. Examples of such approaches include those described in Soyk S, et al., Nat Genet. 2017 Jan;49(1):162-168.

Oil and biofuel production

[0974] The compositions, systems, and methods may be used to generate plants for oil and biofuel production. Biofuels include fuels made from plant and plant-derived resources. Biofuels may be extracted from organic matter whose energy has been obtained through a process of carbon fixation or are made through the use or conversion of biomass. This biomass can be used directly for biofuels or can be converted to convenient energy containing substances by thermal conversion, chemical conversion, and biochemical conversion. This biomass conversion can result in fuel in solid, liquid, or gas form. Biofuels include bioethanol and biodiesel. Bioethanol can be produced by the sugar fermentation process of cellulose (starch), which may be derived from maize and sugar cane. Biodiesel can be produced from oil crops such as rapeseed, palm, and soybean. Biofuels can be used for transportation.

Generation of plants for production of vegetable oils and biofuels

[0975] The compositions, systems, and methods may be used to generate algae (e.g., diatom) and other plants (e.g., grapes) that express or overexpress high levels of oil or biofuels.

[0976] In some cases, the compositions, systems, and methods may be used to modify genes involved in the modification of the quantity of lipids and/or the quality of the lipids. Examples of such genes include those involved in the pathways of fatty acid synthesis, e.g., acetyl-CoA carboxylase, fatty acid synthase, 3-ketoacyl_acyl- carrier protein synthase III, glycerol-3-phosphate deshydrogenase (G3PDH), Enoyl-acyl carrier protein reductase (Enoyl-ACP-reductase), glycerol-3-phosphate acyltransferase, lysophosphatidic acyl transferase or diacylglycerol acyltransferase, phospholipid:diacylglycerol acyltransferase, phosphatidate phosphatase, fatty acid thioesterase such as palmitoyl protein thioesterase, or malic enzyme activities.

[0977] In further embodiments, it is envisaged to generate diatoms that have increased lipid accumulation. This can be achieved by targeting genes that decrease lipid catabolization. Examples of genes include those involved in the activation of triacylglycerol and free fatty acids, β -oxidation of fatty acids, such as genes of acyl-CoA synthetase, 3-ketoacyl-CoA thiolase, acyl-CoA oxidase activity and phosphoglucomutase.

[0978] In some examples, algae may be modified for production of oil and biofuels, including fatty acids (e.g., fatty esters such as acid methyl esters (FAME) and fatty acid ethyl esters (FAEE)). Examples of methods of modifying microalgae include those described in

Stovicek et al. Metab. Eng. Comm., 2015; 2:1; US Patent No. 8,945,839; and International Patent Publication No. WO 2015/086795.

[0979] In some examples, one or more genes may be introduced (e.g., overexpressed) to the plants (e.g., algae) to produce oils and biofuels (e.g., fatty acids) from a carbon source (e.g., alcohol). Examples of the genes include genes encoding acyl-CoA synthases, ester synthases, thioesterases (e.g., *tesA*, *'tesA*, *tesB*, *fatB*, *fatB2*, *fatB3*, *fatA1*, or *fatA*), acyl-CoA synthases (e.g., *fadD*, *JadK*, *BH3103*, *pfl-4354*, *EAV15023*, *fadD1*, *fadD2*, *RPC_4074*, *fadDD35*, *fadDD22*, *faa39*), ester synthases (e.g., synthase/acyl-CoA:diacylglycerol acyltransferase from *Simmondsia chinensis*, *Acinetobacter sp. ADP*, *Alcanivorax borkumensis*, *Pseudomonas aeruginosa*, *Fundibacter jadensis*, *Arabidopsis thaliana*, or *Alkaligenes eutrophus*, or variants thereof).

[0980] Additionally or alternatively, one or more genes in the plants (e.g., algae) may be inactivated (e.g., expression of the genes is decreased). For examples, one or more mutations may be introduced to the genes. Examples of such genes include genes encoding acyl-CoA dehydrogenases (e.g., *fade*), outer membrane protein receptors, and transcriptional regulator (e.g., repressor) of fatty acid biosynthesis (e.g., *fabR*), pyruvate formate lyases (e.g., *pflB*), lactate dehydrogenases (e.g., *IdhA*).

Organic acid production

[0981] In one embodiment, plants may be modified to produce organic acids such as lactic acid. The plants may produce organic acids using sugars, pentose or hexose sugars. To this end, one or more genes may be introduced (e.g., and overexpressed) in the plants. An example of such genes include LDH gene.

[0982] In some examples, one or more genes may be inactivated (e.g., expression of the genes is decreased). For examples, one or more mutations may be introduced to the genes. The genes may include those encoding proteins involved an endogenous metabolic pathway which produces a metabolite other than the organic acid of interest and/or wherein the endogenous metabolic pathway consumes the organic acid.

[0983] Examples of genes that can be modified or introduced include those encoding pyruvate decarboxylases (*pdc*), fumarate reductases, alcohol dehydrogenases (*adh*), acetaldehyde dehydrogenases, phosphoenolpyruvate carboxylases (*ppc*), D-lactate dehydrogenases (*d-ldh*), L-lactate dehydrogenases (*l-ldh*), lactate 2-monooxygenases, lactate

dehydrogenase, cytochrome-dependent lactate dehydrogenases (e.g., cytochrome B2-dependent L-lactate dehydrogenases).

Enhancing plant properties for biofuel production

[0984] In one embodiment, the compositions, systems, and methods are used to alter the properties of the cell wall of plants to facilitate access by key hydrolyzing agents for a more efficient release of sugars for fermentation. By reducing the proportion of lignin in a plant the proportion of cellulose can be increased. In one embodiment, lignin biosynthesis may be downregulated in the plant so as to increase fermentable carbohydrates.

[0985] In some examples, one or more lignin biosynthesis genes may be down regulated. Examples of such genes include 4-coumarate 3-hydroxylases (C3H), phenylalanine ammonia-lyases (PAL), cinnamate 4-hydroxylases (C4H), hydroxycinnamoyl transferases (HCT), caffeic acid O-methyltransferases (COMT), caffeoyl CoA 3-O-methyltransferases (CCoAOMT), ferulate 5- hydroxylases (F5H), cinnamyl alcohol dehydrogenases (CAD), cinnamoyl CoA-reductases (CCR), 4- coumarate-CoA ligases (4CL), monolignol-lignin-specific glycosyltransferases, and aldehyde dehydrogenases (ALDH), and those described in WO 2008064289.

[0986] In some examples, plant mass that produces lower level of acetic acid during fermentation may be reduced. To this end, genes involved in polysaccharide acetylation (e.g., Cas1L and those described in WO 2010096488) may be inactivated.

Other microorganisms for oils and biofuel production

[0987] In one embodiment, microorganisms other than plants may be used for production of oils and biofuels using the compositions, systems, and methods herein. Examples of the microorganisms include those of the genus of *Escherichia*, *Bacillus*, *Lactobacillus*, *Rhodococcus*, *Synechococcus*, *Synechoystis*, *Pseudomonas*, *Aspergillus*, *Trichoderma*, *Neurospora*, *Fusarium*, *Humicola*, *Rhizomucor*, *Kluyveromyces*, *Pichia*, *Mucor*, *Myceliophthora*, *Penicillium*, *Phanerochaete*, *Pleurotus*, *Trametes*, *Chrysosporium*, *Saccharomyces*, *Stenotrophomonas*, *Schizosaccharomyces*, *Yarrowia*, or *Streptomyces*.

Plant cultures and regeneration

[0988] In one embodiment, the modified plants or plant cells may be cultured to regenerate a whole plant which possesses the transformed or modified genotype and thus the desired phenotype. Examples of regeneration techniques include those relying on manipulation of

certain phytohormones in a tissue culture growth medium, relying on a biocide and/or herbicide marker which has been introduced together with the desired nucleotide sequences, obtaining from cultured protoplasts, plant callus, explants, organs, pollens, embryos or parts thereof.

Detecting modifications in the plant genome- selectable markers

[0989] When the compositions, systems, and methods are used to modify a plant, suitable methods may be used to confirm and detect the modification made in the plant. In some examples, when a variety of modifications are made, one or more desired modifications or traits resulting from the modifications may be selected and detected. The detection and confirmation may be performed by biochemical and molecular biology techniques such as Southern analysis, PCR, Northern blot, S1 RNase protection, primer-extension or reverse transcriptase-PCR, enzymatic assays, ribozyme activity, gel electrophoresis, Western blot, immunoprecipitation, enzyme-linked immunoassays, *in situ* hybridization, enzyme staining, and immunostaining.

[0990] In some cases, one or more markers, such as selectable and detectable markers, may be introduced to the plants. Such markers may be used for selecting, monitoring, isolating cells and plants with desired modifications and traits. A selectable marker can confer positive or negative selection and is conditional or non-conditional on the presence of external substrates. Examples of such markers include genes and proteins that confer resistance to antibiotics, such as hygromycin (hpt) and kanamycin (nptII), and genes that confer resistance to herbicides, such as phosphinothricin (bar) and chlorosulfuron (als), enzyme capable of producing or processing a colored substances (e.g., the β -glucuronidase, luciferase, B or C1 genes).

Applications in fungi

[0991] The compositions, systems, and methods described herein can be used to perform efficient and cost-effective gene or genome interrogation or editing or manipulation in fungi or fungal cells, such as yeast. The approaches and applications in plants may be applied to fungi as well.

[0992] A fungal cell may be any type of eukaryotic cell within the kingdom of fungi, such as phyla of *Ascomycota*, *Basidiomycota*, *Blastocladiomycota*, *Chytridiomycota*, *Glomeromycota*, *Microsporidia*, and *Neocallimastigomycota*. Examples of fungi or fungal cells include yeasts, molds, and filamentous fungi.

[0993] In one embodiment, the fungal cell is a yeast cell. A yeast cell refers to any fungal cell within the phyla *Ascomycota* and *Basidiomycota*. Examples of yeasts include budding

yeast, fission yeast, and mold, *S. cerevisiae*, *Kluyveromyces marxianus*, *Issatchenkia orientalis*, *Candida* spp. (e.g., *Candida albicans*), *Yarrowia* spp. (e.g., *Yarrowia lipolytica*), *Pichia* spp. (e.g., *Pichia pastoris*), *Kluyveromyces* spp. (e.g., *Kluyveromyces lactis* and *Kluyveromyces marxianus*), *Neurospora* spp. (e.g., *Neurospora crassa*), *Fusarium* spp. (e.g., *Fusarium oxysporum*), and *Issatchenkia* spp. (e.g., *Issatchenkia orientalis*, *Pichia kudriavzevii* and *Candida acidothermophilum*).

[0994] In one embodiment, the fungal cell is a filamentous fungal cell, which grow in filaments, e.g., hyphae or mycelia. Examples of filamentous fungal cells include *Aspergillus* spp. (e.g., *Aspergillus niger*), *Trichoderma* spp. (e.g., *Trichoderma reesei*), *Rhizopus* spp. (e.g., *Rhizopus oryzae*), and *Mortierella* spp. (e.g., *Mortierella isabellina*).

[0995] In one embodiment, the fungal cell is of an industrial strain. Industrial strains include any strain of fungal cell used in or isolated from an industrial process, e.g., production of a product on a commercial or industrial scale. Industrial strain may refer to a fungal species that is typically used in an industrial process, or it may refer to an isolate of a fungal species that may be also used for non-industrial purposes (e.g., laboratory research). Examples of industrial processes include fermentation (e.g., in production of food or beverage products), distillation, biofuel production, production of a compound, and production of a polypeptide. Examples of industrial strains include, without limitation, JAY270 and ATCC4124.

[0996] In one embodiment, the fungal cell is a polyploid cell whose genome is present in more than one copy. Polyploid cells include cells naturally found in a polyploid state, and cells that has been induced to exist in a polyploid state (e.g., through specific regulation, alteration, inactivation, activation, or modification of meiosis, cytokinesis, or DNA replication). A polyploid cell may be a cell whose entire genome is polyploid, or a cell that is polyploid in a particular genomic locus of interest. In some examples, the abundance of guide RNA may more often be a rate-limiting component in genome engineering of polyploid cells than in haploid cells, and thus the methods using the composition described herein may take advantage of using certain fungal cell types.

[0997] In one embodiment, the fungal cell is a diploid cell, whose genome is present in two copies. Diploid cells include cells naturally found in a diploid state, and cells that have been induced to exist in a diploid state (e.g., through specific regulation, alteration, inactivation, activation, or modification of meiosis, cytokinesis, or DNA replication). A diploid cell may

refer to a cell whose entire genome is diploid, or it may refer to a cell that is diploid in a particular genomic locus of interest.

[0998] In one embodiment, the fungal cell is a haploid cell, whose genome is present in one copy. Haploid cells include cells naturally found in a haploid state, or cells that have been induced to exist in a haploid state (e.g., through specific regulation, alteration, inactivation, activation, or modification of meiosis, cytokinesis, or DNA replication). A haploid cell may refer to a cell whose entire genome is haploid, or it may refer to a cell that is haploid in a particular genomic locus of interest.

[0999] The compositions and systems, and nucleic acid encoding thereof may be introduced to fungi cells using the delivery systems and methods herein. Examples of delivery systems include lithium acetate treatment, bombardment, electroporation, and those described in Kawai et al., 2010, *Bioeng Bugs*. 2010 Nov-Dec; 1(6): 395–403.

[01000] In some examples, a yeast expression vector (e.g., those with one or more regulatory elements) may be used. Examples of such vectors include a centromeric (CEN) sequence, an autonomous replication sequence (ARS), a promoter, such as an RNA Polymerase III promoter, operably linked to a sequence or gene of interest, a terminator such as an RNA polymerase III terminator, an origin of replication, and a marker gene (e.g., auxotrophic, antibiotic, or other selectable markers). Examples of expression vectors for use in yeast may include plasmids, yeast artificial chromosomes, 2 μ plasmids, yeast integrative plasmids, yeast replicative plasmids, shuttle vectors, and episomal plasmids.

Biofuel and materials production by fungi

[01001] In one embodiment, the compositions, systems, and methods may be used for generating modified fungi for biofuel and material productions. For instance, the modified fungi for production of biofuel or biopolymers from fermentable sugars and optionally to be able to degrade plant-derived lignocellulose derived from agricultural waste as a source of fermentable sugars. Foreign genes required for biofuel production and synthesis may be introduced in to fungi. In some examples, the genes may encode enzymes involved in the conversion of pyruvate to ethanol or another product of interest, degrade cellulose (e.g., cellulase), endogenous metabolic pathways which compete with the biofuel production pathway.

[01002] In some examples, the compositions, systems, and methods may be used for generating and/or selecting yeast strains with improved xylose or cellobiose utilization, isoprenoid biosynthesis, and/or lactic acid production. One or more genes involved in the metabolism and synthesis of these compounds may be modified and/or introduced to yeast cells. Examples of the methods and genes include lactate dehydrogenase, PDC1 and PDC5, and those described in Ha, S.J., et al. (2011) Proc. Natl. Acad. Sci. USA 108(2):504-9 and Galazka, J.M., et al. (2010) Science 330(6000):84-6; Jakočiūnas T et al., Metab Eng. 2015 Mar;28:213-222; Stovicek V, et al., FEMS Yeast Res. 2017 Aug 1;17(5).

Improved plants and yeast cells

[01003] The present disclosure further provides improved plants and fungi. The improved and fungi may comprise one or more genes introduced, and/or one or more genes modified by the compositions, systems, and methods herein. The improved plants and fungi may have increased food or feed production (e.g., higher protein, carbohydrate, nutrient or vitamin levels), oil and biofuel production (e.g., methanol, ethanol), tolerance to pests, herbicides, drought, low or high temperatures, excessive water, etc.

[01004] The plants or fungi may have one or more parts that are improved, e.g., leaves, stems, roots, tubers, seeds, endosperm, ovule, and pollen. The parts may be viable, nonviable, regeneratable, and/or non-regeneratable.

[01005] The improved plants and fungi may include gametes, seeds, embryos, either zygotic or somatic, progeny and/or hybrids of improved plants and fungi. The progeny may be a clone of the produced plant or fungi, or may result from sexual reproduction by crossing with other individuals of the same species to introgress further desirable traits into their offspring. The cell may be *in vivo* or *ex vivo* in the cases of multicellular organisms, particularly plants.

Further applications in plants

[01006] Further applications of the compositions, systems, and methods on plants and fungi include visualization of genetic element dynamics (e.g., as described in Chen B, et al., Cell. 2013 Dec 19;155(7):1479-91), targeted gene disruption positive-selection *in vitro* and *in vivo* (as described in Malina A et al., Genes Dev. 2013 Dec 1;27(23):2602-14), epigenetic modification such as using fusion of IscB polypeptide nuclease and histone-modifying enzymes (e.g., as described in Rusk N, Nat Methods. 2014 Jan;11(1):28), identifying transcription regulators (e.g., as described in Waldrip ZJ, Epigenetics. 2014 Sep;9(9):1207-11), anti-virus treatment for both RNA and DNA viruses (e.g., as described in Price AA, et al., Proc

Natl Acad Sci U S A. 2015 May 12;112(19):6164-9; Ramanan V et al., Sci Rep. 2015 Jun 2;5:10833), alteration of genome complexity such as chromosome numbers (e.g., as described in Karimi-Ashtiyani R et al., Proc Natl Acad Sci U S A. 2015 Sep 8;112(36):11211-6; Anton T, et al., Nucleus. 2014 Mar-Apr;5(2):163-72), self-cleavage of the composition for controlled inactivation/activation (e.g., as described Sugano SS et al., Plant Cell Physiol. 2014 Mar;55(3):475-81), multiplexed gene editing (as described in Kabadi AM et al., Nucleic Acids Res. 2014 Oct 29;42(19):e147), development of kits for multiplex genome editing (as described in Xing HL et al., BMC Plant Biol. 2014 Nov 29;14:327), starch production (as described in Hebelstrup KH et al., Front Plant Sci. 2015 Apr 23;6:247), targeting multiple genes in a family or pathway (e.g., as described in Ma X et al., Mol Plant. 2015 Aug;8(8):1274-84), regulation of non-coding genes and sequences (e.g., as described in Lowder LG, et al., Plant Physiol. 2015 Oct;169(2):971-85), editing genes in trees (e.g., as described in Belhaj K et al., Plant Methods. 2013 Oct 11;9(1):39; Harrison MM, et al., Genes Dev. 2014 Sep 1;28(17):1859-72; Zhou X et al., New Phytol. 2015 Oct;208(2):298-301), introduction of mutations for resistance to host-specific pathogens and pests.

[01007] Additional examples of modifications of plants and fungi that may be performed using the compositions, systems, and methods include analogous modifications described in International Patent Publication Nos. WO2016/099887, WO2016/025131, WO2016/073433, WO2017/066175, WO2017/100158, WO 2017/105991, WO2017/106414, WO2016/100272, WO2016/100571, WO 2016/100568, WO 2016/100562, and WO 2017/019867.

APPLICATIONS IN NON-HUMAN ANIMALS

[01008] The compositions, systems, and methods may be used to study and modify non-human animals, e.g., introducing desirable traits and disease resilience, treating diseases, facilitating breeding, etc. In one embodiment, the compositions, systems, and methods may be used to improve breeding and introducing desired traits, e.g., increasing the frequency of trait-associated alleles, introgression of alleles from other breeds/species without linkage drag, and creation of de novo favorable alleles. Genes and other genetic elements that can be targeted may be screened and identified. Examples of application and approaches include those described in Tait-Burkard C, et al., Livestock 2.0 - genome editing for fitter, healthier, and more productive farmed animals. Genome Biol. 2018 Nov 26;19(1):204; Lillico S, Agricultural applications of genome editing in farmed animals. Transgenic Res. 2019 Aug;28(Suppl 2):57-60; Houston RD, et al., Harnessing genomics to fast-track genetic improvement in aquaculture.

Nat Rev Genet. 2020 Apr 16. doi: 10.1038/s41576-020-0227-y, which are incorporated herein by reference in their entireties. Applications described in other sections such as therapeutic, diagnostic, etc. can also be used on the animals herein.

[01009] The compositions, systems, and methods may be used on animals such as fish, amphibians, reptiles, mammals, and birds. The animals may be farm and agriculture animals, or pets. Examples of farm and agriculture animals include horses, goats, sheep, swine, cattle, llamas, alpacas, and birds, e.g., chickens, turkeys, ducks, and geese. The animals may be a non-human primate, e.g., baboons, capuchin monkeys, chimpanzees, lemurs, macaques, marmosets, tamarins, spider monkeys, squirrel monkeys, and vervet monkeys. Examples of pets include dogs, cats horses, wolfs, rabbits, ferrets, gerbils, hamsters, chinchillas, fancy rats, guinea pigs, canaries, parakeets, and parrots.

[01010] In one embodiment, one or more genes may be introduced (e.g., overexpressed) in the animals to obtain or enhance one or more desired traits. Growth hormones, insulin-like growth factors (IGF-1) may be introduced to increase the growth of the animals, e.g., pigs or salmon (such as described in Pursel VG et al., J Reprod Fertil Suppl. 1990;40:235-45; Waltz E, Nature. 2017;548:148). Fat-1 gene (e.g., from *C elegans*) may be introduced for production of larger ratio of n-3 to n-6 fatty acids may be induced, e.g. in pigs (such as described in Li M, et al., Genetics. 2018;8:1747-54). Phytase (e.g., from *E coli*) xylanase (e.g., from *Aspergillus niger*), beta-glucanase (e.g., from *Bacillus licheniformis*) may be introduced to reduce the environmental impact through phosphorous and nitrogen release reduction, e.g. in pigs (such as described in Golovan SP, et al., Nat Biotechnol. 2001;19:741-5; Zhang X et al., elife. 2018). shRNA decoy may be introduced to induce avian influenza resilience e.g. in chicken (such as described in Lyall et al., Science. 2011;331:223-6). Lysozyme or lysostaphin may be introduced to induce mastitis resilience e.g., in goat and cow (such as described in Maga EA et al., Foodborne Pathog Dis. 2006;3:384-92; Wall RJ, et al., Nat Biotechnol. 2005;23:445-51). Histone deacetylase such as HDAC6 may be introduced to induce PRRSV resilience, e.g., in pig (such as described in Lu T., et al., PLoS One. 2017;12:e0169317). CD163 may be modified (e.g., inactivated or removed) to introduce PRRSV resilience in pigs (such as described in Prather RS et al., Sci Rep. 2017 Oct 17;7(1):13371). Similar approaches may be used to inhibit or remove viruses and bacteria (e.g., Swine Influenza Virus (SIV) strains which include influenza C and the subtypes of influenza A known as H1N1, H1N2, H2N1, H3N1, H3N2, and

H2N3, as well as pneumonia, meningitis and oedema) that may be transmitted from animals to humans.

[01011] In one embodiment, one or more genes may be modified or edited for disease resistance and production traits. Myostatin (e.g., GDF8) may be modified to increase muscle growth, e.g., in cow, sheep, goat, catfish, and pig (such as described in Crispo M et al., *PLoS One*. 2015;10:e0136690; Wang X, et al., *Anim Genet*. 2018;49:43–51; Khalil K, et al., *Sci Rep*. 2017;7:7301; Kang J-D, et al., *RSC Adv*. 2017;7:12541–9). *Pc POLLED* may be modified to induce hornlessness, e.g., in cow (such as described in Carlson DF et al., *Nat Biotechnol*. 2016;34:479–81). *KISS1R* may be modified to induce boar taint (hormone release during sexual maturity leading to undesired meat taste), e.g., in pigs. *Dead end protein (dnd)* may be modified to induce sterility, e.g., in salmon (such as described in Wargelius A, et al., *Sci Rep*. 2016;6:21284). *Nano2* and *DDX* may be modified to induce sterility (e.g., in surrogate hosts), e.g., in pigs and chicken (such as described Park K-E, et al., *Sci Rep*. 2017;7:40176; Taylor L et al., *Development*. 2017;144:928–34). *CD163* may be modified to induce PRRSV resistance, e.g., in pigs (such as described in Whitworth KM, et al., *Nat Biotechnol*. 2015;34:20–2). *RELA* may be modified to induce ASFV resilience, e.g., in pigs (such as described in Lillico SG, et al., *Sci Rep*. 2016;6:21645). *CD18* may be modified to induce *Mannheimia (Pasteurella) haemolytica* resilience, e.g., in cows (such as described in Shanthalingam S, et al., *Proc Natl Acad Sci U S A*. 2016;113:13186–90). *NRAMP1* may be modified to induce tuberculosis resilience, e.g., in cows (such as described in Gao Y et al., *Genome Biol*. 2017;18:13). Endogenous retrovirus genes may be modified or removed for xenotransplantation such as described in Yang L, et al. *Science*. 2015;350:1101–4; Niu D et al., *Science*. 2017;357:1303–7). Negative regulators of muscle mass (e.g., Myostatin) may be modified (e.g., inactivated) to increase muscle mass, e.g., in dogs (as described in Zou Q et al., *J Mol Cell Biol*. 2015 Dec;7(6):580–3).

[01012] Animals such as pigs with severe combined immunodeficiency (SCID) may be generated (e.g., by modifying *RAG2*) to provide useful models for regenerative medicine, xenotransplantation (discussed also elsewhere herein), and tumor development. Examples of methods and approaches include those described Lee K, et al., *Proc Natl Acad Sci U S A*. 2014 May 20;111(20):7260–5; and Schomberg et al. *FASEB Journal*, April 2016; 30(1):Suppl 571.1.

[01013] SNPs in the animals may be modified. Examples of methods and approaches include those described Tan W. et al., Proc Natl Acad Sci U S A. 2013 Oct 8;110(41):16526-31; Mali P, et al., Science. 2013 Feb 15;339(6121):823-6.

[01014] Stem cells (e.g., induced pluripotent stem cells) may be modified and differentiated into desired progeny cells, e.g., as described in Heo YT et al., Stem Cells Dev. 2015 Feb 1;24(3):393-402.

[01015] Profile analysis (such as Igenity) may be performed on animals to screen and identify genetic variations related to economic traits. The genetic variations may be modified to introduce or improve the traits, such as carcass composition, carcass quality, maternal and reproductive traits and average daily gain.

KITS

[01016] In one aspect, the invention provides kits containing any one or more of the elements disclosed in the above methods and compositions. In one aspect, the invention provides a kit comprising one or more of the components described herein. In one embodiment, the kit comprises the compositions herein and instructions for using the kit. In one embodiment, the kit comprises a vector system and instructions for using the kit. In one embodiment, the kit comprises a delivery system and instructions for using the kit. In one embodiment, the kit comprises a vector system and instructions for using the kit. Elements may be provided individually or in combinations, and may be provided in any suitable container, such as a vial, a bottle, or a tube. The kits may include the ω RNA and optionally an unbound protector strand as described herein. The kits may include the ω RNA with a protector strand bound to at least partially to a reprogrammable spacer portion of the ω RNA sequence (i.e., phRNA). Thus the kits may include the phRNA in the form of a partially double stranded nucleotide sequence as described here. In one embodiment, the kit includes instructions in one or more languages, for example in more than one language. The instructions may be specific to the applications and methods described herein.

[01017] In one embodiment, a kit comprises one or more reagents for use in a process utilizing one or more of the elements described herein. Reagents may be provided in any suitable container. For example, a kit may provide one or more reaction or storage buffers. Reagents may be provided in a form that is usable in a particular assay, or in a form that requires addition of one or more other components before use (e.g., in concentrate or lyophilized form).

A buffer can be any buffer, including but not limited to a sodium carbonate buffer, a sodium bicarbonate buffer, a borate buffer, a Tris buffer, a MOPS buffer, a HEPES buffer, and combinations thereof. In one embodiment, the buffer is alkaline. In one embodiment, the buffer has a pH from about 7 to about 10. In one embodiment, the kit comprises one or more oligonucleotides corresponding to a ω RNA scaffold, reprogrammable sequence for insertion into a vector so as to operably link the ω RNA sequence and a regulatory element. In one embodiment, the kit comprises a homologous recombination template polynucleotide. In one embodiment, the kit comprises one or more of the vectors and/or one or more of the polynucleotides described herein. The kit may advantageously allow to provide all elements of the systems of the invention.

[01018] The invention is further described in the following examples, which do not limit the scope of the invention described in the claims.

[01019] Further embodiments are illustrated in the following Examples which are given for illustrative purposes only and are not intended to limit the scope of the invention.

EXAMPLES

Example 1 –

[01020] FIG. 1 shows the IscB cleavage of an endogenous target and a non-endogenous target sequence, with sequences comprising TAMs 1-3 showing cleavage when using a complementary spacer using the *K. racemifer* IscB protein, ω RNA scaffold and spacers for endogenous target Kr (FIG. 1, left) and non-endogenous target Fn (FIG. 1, right). Cleavage with TAM sequences 1-3 agrees with TAM weblogo identified for *K. racemifer* IscB in FIG. 2.

[01021] FIG. 2 shows determination of TAM for an IscB polypeptide Polypeptide. Sequences utilized in the experiment

[01022] >IscB protein sequence (from *K. racemifer*)
 MNVVYVLSPERTPLMPCQPAIARLLLKQGKAKVRHRTPFITQLLAQPEHVYTQPLTH
 GVDTGSSIIGSAVANEHGHVVYLSEVEIRNDIANTMKERARARRNRRQRKTRYRPAR
 WL
 NRKKSIIKTGRFSPTMRSKIDTHLREIRFIRSLIPITSTILETGSDPYALRNPEVLQKKW
 LYQRGINYGFANTKAYVLTRDGYLCQQCKGKSKDRRLEVHHIIFRSRNGSDEEANL
 LTLCKTCHDGLHAGTITLKLTKGKKGTQHQATQMNSIRIQLLKRVEAEETWGFVTKE

HRLLVGLPKEHIFDAAVIATRGVKPTFYTTSVLSKHCVSDGDYKQTKGKHGQQRVN
 TGKIMGFRKFDKVVYLGKEYFIKGRMSTGYAILMDIDGNKIEFKPLPKFDKMKRVSA
 R SSWMMKQRTTPNPSFSITSSLASAGKNV* (SEQ ID NO: 2059)

[01023] > ω RNA scaffold sequence

GTGAACTACCACTGAGCTGAAGACGCAGTGGCTTCTTCGGAAGTCACTGAAGAC
 GCAGACCAGGAGCTCCTTCGGAAGCTTGAGTTCACCAGACTCGTTTCCAGAAATG
 GGAACAGCGTTCGATTGGTCATGACACCTGCGGTTGACGCATCAGACCGCTG
 CTCTGTCGCTGAGGGTTAAGTAGGCTTGAGGAAAGGGCCGGTGCTCTCAGCGCA
 A

AAAGCCTTTTGAACACTGTCGAGATGAAGCCGGATTCCCTTCGTGGTCACAGCG
 AAGGGATACGCACCACCCGGCGCTTGCCGGAGCATTTTCCGAAAGGAGTTTT
 (SEQ ID NO: 2060)

[01024] >Kr spacer sequence (endogenous spacer)

GAAGAAGAGGCCGCACCCGTTTGAGGCCGCACCAAAT (SEQ ID NO: 2061)

[01025] >Fn spacer sequence

CAAGCTTTTTTAACAGTGGCCTTATTAAATGACTTCTC (SEQ ID NO: 2062)

[01026] FIG. 3 provides a sequence logo of the N terminal domain from an alignment of the 1004 representative IscB loci with conserved motifs boxed and annotated. While exact residues and domain size may vary, there are identified conserved motifs in the cluster of loci of IscBs at 60% identity and 70% coverage.

[01027] FIG. 4 includes a sequence alignment of this cluster of IscB loci, that span a majority of existing IscB sequences. Provided in the table are approximate conserved portions of the ω RNA scaffold without the reprogrammable spacer, and IscB protein. The conserved region on the 5' region may vary, but generally identifies the approximate location where the spacer and conserved hRNA scaffold region would meet. The hRNA scaffold region and the IscB protein, along with identifying IscB loci information are detailed in Table 1.

[01028] FIG. 5 shows the consensus sequence from IscB loci of Table 1.

Example 2 – IscB genome editing methods

[01029] Mammalian cell culture experiments were performed in the HEK293FT line (American Type Culture Collection (ATCC)), which was grown in Dulbecco's Modified Eagle

Medium with high glucose, sodium pyruvate, and GlutaMAX (Thermo Fisher Scientific), additionally supplemented with 1× penicillin–streptomycin (Thermo Fisher Scientific) and 10% fetal bovine serum (VWR Seradigm). Transfections were performed with Lipofectamine 2000 (Thermo Fisher Scientific) in 96-well plates. Cells were plated at approximately 20,000 cells/well 16 hours prior to transfection to ensure 90% confluency at the time of transfection.

[01030] For each well on the plate, 300 ng guideRNA expression plasmid, pHS0812_Isc_large_27 (FIG. 8A) and 150 ng IscB expression plasmid, pHS0810_IscB_large_27 (FIG. 8B) were combined with OptiMEM I Reduced Serum Medium (Thermo Fisher) to a total of 10 µl. Separately, 9.2 µl of OptiMEM was combined with 0.8 µl of Lipofectamine 2000. Plasmid and Lipofectamine solutions were then combined and incubated for 5 minutes, after which they were pipetted onto cells.

[01031] After 72 hours, genomic DNA was harvested from the cells by adding 50 uL of QuickExtract DNA Extraction Solution (Lucigen) and incubating at 65C for 15 min, 68C for 15 min and 95C for 10 min. The extracted genomic DNA was then subjected to two rounds of PCR to amplify the target site and add Illumina adaptors and sample barcodes using NEBNext High-Fidelity 2X PCR Master Mix (New England Biolabs). The library was then subjected to next generation sequencing on an Illumina MiSeq. Indel rate was evaluated using the CRISPResso2 pipeline (Clement et. al., *Nat. Biotech.* 2019) using a quantification window center of -8 and size 5.

[01032] FIG. 6A-6C shows (6A) Weblogo indicating the TAMs identified in this study; (6B) Indel frequency compared to negative control condition at VEGFA site 2; and (6C) Representative indels at VEGFA site 2.

[01033] FIG. 7A-7B. IscB Protein sequence utilized in this example and the studies shown in Figure 6A-6C, and (7A) OGEU01000025.1
 MMAVVYVISKSGKPLMPTTRCGHVRILLKEGKARVVERKPFTIQLTYESAEETQPLV
 LGIDPGRTNIGMSVVTESGESVFNAQIETRNKDVPKLMKDRKQYRMAHRRLKRRCK
 RRRRAKAAGTAFEERGEKQRLLPGCFKPITCKSIRNKEARFNNRKRVPVGLTPTANHL
 LVTHLNVVKKVQKILPVAKVVLLELNRFSEFMAMNNPKVQRWQYQRGPLYGKGSVEE
 AVSMQQDGHCLFCKHGDHYHHVPPRRKNGSETLENRVGLCEEHHRLVHTDKEWE
 ANLASKKSGMNKKYHALSVLNQIIPYLADQLADMFPGNFCVTSGQDITYLFREEHGIP
 KDHYLDAYCIACSALTDAAKVSSPKGRPYMVHQFRRHDRQACHKANLNRSYMG
 GKL VATNRHKAMDQKTD SLEEYRAAHSAADVSKLTVKHPSAQYKDMSRIMPGSIL

VSGEGKLFTLSRSEGRNKGQVNYFVSTEGIKYWARKCQYLRNNGGLQIYV* (SEQ ID NO: 2063) and (7B) its >ωRNA scaffold sequence

GGCTCTTCCAAC TTTATGGTTGCGACCGTAGGTTGAAAGAGCACAGGCTGAGAC
ATTCGTAAGGCCGAAAGACCGGACGCACCCTGGGATT TCCCCAGTCCCCGGAAC
TGCATAGCGGATGCCAGTTGATGGAGCAATCTATCAGATAAGCCAGGGGGAACA
ATCACCTCTCTGTATCAGAGAGAGTTTTACAAAAGGAGGAACGG. (SEQ ID NO: 2064)

Example 3 – The widespread IS200/605 transposon family encodes diverse re-targetable RNA-guided endonucleases

[01034] The prokaryotic RNA-guided defense system CRISPR-Cas9 (type II CRISPR-Cas), which has been adopted for genome editing in eukaryotic cells (Zhang, F. (2019), *Quarterly Reviews of Biophysics* **52**; Hille, F, *et al.* (2018), *Cell* **172**: 1239-1259), is thought to have evolved from IscB proteins (3). Despite its wide distribution across prokaryotes and shared domain composition and architecture with Cas9, the function of IscB remains unknown (Fig. 43). Moreover, given that IscB has not been reported to be associated with non-coding RNA (ncRNA) or CRISPR arrays, the evolutionary origins of the RNA-guided activity in Cas9 systems are unclear. IscB is encoded by a distinct subset of IS200/605 superfamily transposons that also include transposons encoding *tnpB*, a putative endonuclease distantly related to *iscB* and thought to be the ancestor of Cas12, the type V CRISPR effector (Kapitonav, V. *et al.* (2015), *J. Bacteriol.* **198**, 797-807; Siguir, P. *et al.* (2014), *FEMS Microbiol. Rev.* **38**, 865-891; S. Shmakov, S. *et al.* (2017), *Nat. Rev. Microbiol.* **15**, 169–182). Using phylogenetic analysis, RNA-seq, and biochemical experiments, Applicants sought to elucidate the functions of these proteins and the origin of RNA-guided activity in class 2 CRISPR systems.

IscB is associated with an evolutionarily conserved non-coding RNA

[01035] IscB is ~400 amino acids (aa) long and contains a RuvC endonuclease domain split by the insertion of a bridge helix (BH) and an HNH endonuclease domain, an architecture that is shared with Cas9 (Fig. 1A) (Kapitonav, V. *et al.* (2015), *J. Bacteriol.* **198**, 797-807). Applicants performed a comprehensive search for proteins containing an HNH or a split RuvC endonuclease domain and found that Cas9 and IscB were the only proteins that contained

both domains. This search also showed that IscB contains a previously unidentified N-terminus that lacks clear homology to known domains and is absent in Cas9, which is denoted PLMP after its conserved sequence motifs (Fig. 9A, Fig. 10). Clustering and phylogenetic analysis of the combined RuvC, BH, and HNH domains strongly suggests that all extant Cas9s descended from a single ancestral IscB (Fig. 9B). Applicants searched for CRISPR arrays adjacent to *iscB* genes from each cluster and found 6 distinct groups of IscB, containing 16 clusters (of 603 total), that were CRISPR-associated, contrary to previous observations (Kapitonov, V. et al. (2015), *J. Bacteriol.* 198, 797-807). CRISPR-associated IscBs were scattered around the IscB phylogenetic tree, suggesting they evolved independently, with one association event leading to the Cas9 lineage (Fig. 9B). In total Applicants identified 31 unique CRISPR-associated *iscB* loci (of 2811 total).

[01036] Given their association with CRISPR arrays, Applicants suspected that the rarely occurring CRISPR-associated IscBs may be RNA-guided nucleases. Applicants first examined a cluster of CRISPR-associated IscBs similar to non-CRISPR associated IscBs (at ~50% aa identity). Applicants heterologously expressed a representative locus from this clade in *E. coli* and performed small RNA-seq, which showed expression of not only the CRISPR array, but also a 329-bp intergenic region between the CRISPR array and the IscB open reading frame (ORF) (Fig. 9C). Applicants purified the IscB protein and sequenced the co-purified RNA, demonstrating that this protein interacts with a single ncRNA component, encompassing both the CRISPR array and this intergenic region (Fig. 9C).

[01037] Given its interaction with a ncRNA that includes the CRISPR direct repeat (DR) and spacer, as well as its similar domain architecture to Cas9, Applicants tested this IscB for RNA-guided endonuclease activity. Using a previously established protospacer adjacent motif (PAM)-discovery assay (Table 12) (Zetsche, B. et al. (2015), *Cell.* **163**, 759–771), Applicants observed depletion of specific PAM sequences (Fig. 9D, Fig. 46A-C), indicating that CRISPR-associated IscBs are reprogrammable RNA-guided nucleases. Applicants confirmed this enzymatic activity with an *in vitro* cleavage assay using recombinant ribonucleoprotein (RNP) complexes (Fig. 9E).

[01038] Our finding that IscB functionally associated with CRISPR at least once, and likely on additional occasions, suggested that IscB systems more generally share a core ancestral ncRNA gene that is prone to evolving into a CRISPR array and in some cases a separate trans-acting *tracrRNA* (Deltcheva, E, et al. (2011), *Nature.* **471**, 602–607). To test this

hypothesis, Applicants aligned 563 non-redundant *iscB* loci and searched for conserved nucleotide (nt) sequences either upstream or downstream of the *iscB* ORF. This analysis revealed a highly conserved intergenic region ~300 bp in length upstream of the ORF with a drop in conservation at the 5' end, which corresponds to an IS200/605 transposon end. Secondary structure predictions for individual sequences revealed the presence of multiple G:U pairs (Fig. 11), suggesting that the conserved region encodes an ncRNA containing functionally important hairpins, which Applicants named ω RNA. Small RNA-seq on a sample of *Ktedonobacter racemifer* strain SOSP1-21, a soil bacterium that harbors 49 *IscB* loci in its genome (Kapitonov, V. et al. (2015), *J. Bacteriol.* 198, 797-807), demonstrated expression of the predicted ω RNA in many of these loci (Fig. 9F, Figs. 12, 13A). Moreover, Applicants observed that the transcripts consistently extended beyond the conservation boundary at the 5' end.

[01039] An RFAM search for potential homologs of the ω RNA showed that the conserved region of the ω RNA partially matched the previously reported HEARO RNA, a ncRNA that was found upstream of HNH domain-containing proteins, which at the time were thought to be homing endonucleases (Kalvari E, et al. (2021), *Nucleic Acids Res.* **49**, D192–D200; Weinberg, Z. et al. (2009), *Nature.* **462**, 656–659). However, the RFAM search did not provide any clues about the nature of the 5'-terminal non-conserved portion of these transcripts. Comparison of the consensus CRISPR-associated *IscB* ncRNA and the covariance folded ω RNA secondary structures revealed high degrees of structural and sequence similarity, particularly in shared multi-stem regions and pseudoknots (Fig. 9G, Fig. 47A-B, Supplementary Text). Most importantly, Applicants inferred that the 5'-most non-conserved sequence in the ω RNA might function as a guide sequence, because the sequence immediately downstream was predicted to form hairpins that structurally resembled the hairpins formed by the DR/anti-repeat duplex in the CRISPR-associated *IscB* ncRNA (Fig. 9G).

Table 8. Expression plasmids used in this study.

Expression plasmids		
Plasmid name	Description	Associated Figures
pHS0570	pET45b(+) T7/LacO His14-MBP-bdSUMO-CRISPR associated <i>IscB</i> -TwinStrep	Fig. 9C, E
pHS0578	CRISPR associated <i>IscB</i> non-coding locus in pColA-Duet	Fig. 9C

pHS0663	pET45b(+) KraIscB-1 TwinStrep	Fig. 13A
pHS0667	KraIscB-1 non-coding locus in pCOLADuet-1	Fig. 13A
pHS0781	pcDNA3.1(+) CMV-SV40 NLS-AwaIscB-nucleoplasmin NLS-3xHA	
pHS0806	pET45b(+) His14-bdSUMO- AwaIscB	Fig. 15E-G; Fig. 17; Fig. 18; Fig. 19
pHS0810	pcDNA3.1(+) CMV-SV40 NLS-OgeuIscB-nucleoplasmin NLS-3xHA	Fig. 39F-G; Fig. 40
pHS0812	hU6-BpiI-OgeuIscB omegaRNA scaffold	Fig. 39F-G; Fig. 40
pHS0766	pET45b(+) His14-MBP-TEV- AloTnpB-2 3' extension	Fig. 61A
Locus plasmids		
pHS0537	CRISPR-associated IscB locus (Delaware Bay) in pBR322 with Fn spacer	Fig. 9C-D
pHS0535	CRISPR-associated IscB (Chesapeake Bay) in pBR322 with Fn spacer	Fig. 46A-46C
PAM/TAM libraries		
pUC19 Fn 8N library	Fn	Fig. 15B-D; Fig. 39D,E,I,H; Fig. 48; Fig. 61B
pACYC Fn 8N library	Fn	Fig. 1D; Fig. 46A-46C
pACYC KraIscB-1 endogenous target 8N library	KraIscB-1 endogenous target	Fig. 15B-C

IscB is a reprogrammable RNA-guided DNA endonuclease

[01040] To test whether IscB was capable of cleaving DNA complementary to the putative ω RNA guide, Applicants performed an *in vitro* plasmid cleavage assay with KraIscB-1 using an *in vitro* transcription/translation (IVTT) expression system (Fig. 15A, 15B). Applicants found that KraIscB-1 cleaved the target in an ω RNA-dependent manner, with an ATAAA 3' target-adjacent motif (TAM) (Fig. 2C). Retargeting of KraIscB-1 using a different guide (Fn guide) (Zetsche, B. et al. (2015), *Cell*. **163**, 759–771) also mediated cleavage of the cognate target (Fig. 15C, Fig. 13B), implying that IscB is a reprogrammable RNA-guided nuclease.

[01041] Next, Applicants biochemically characterized IscB *in vitro*. Applicants identified activity in 57/86 (66%) selected phylogenetically diverse systems (Table 9) as determined by the identification of a TAM (Fig. 48). Of these 57 functional IscBs, 5 could be reconstituted

with the respective ω RNA *in vitro* to achieve efficient target cleavage, and from those, Applicants selected AwaIscB (from *Allochromatium warmingii*) for detailed biochemical characterization (Fig. 15D-15G).

[01042] Applicants confirmed the ability of recombinant AwaIscB to cleave multiple dsDNA targets in a programmable manner (Fig. 15E) and showed that the activity of AwaIscB is magnesium- dependent with a temperature optimum from 35-40°C (Fig. 17A, B). Appreciable activity was observed *in vitro* with guide lengths between 15 and 45 nt (Fig. 17D). Mutation of the catalytic RuvC-II residue (E157A) abolished the nucleolytic activity on the non-target DNA strand, whereas the HNH domain catalytic mutant H212A abolished the nucleolytic activity on the target strand (Fig. 15F). Combination of the E157A and H212A mutations (dAwaIscB) abolished all dsDNA nucleolytic activity (Fig. 15F) (Jinek, M, et al. (2012), *Science* **337**, 816–821; Gasiunas, G. et al. (2020), *Nat. Commun.* **11**, 55). Sequencing of the cleavage products showed that AwaIscB cleaves the target strand 3 nt upstream of the TAM, similar to Cas9s (GasiunasG, et al. (2020), *Nat. Commun.* **11**, 5512).

[01043] Cleavage of the non-target strand occurred 8 or 12 nt upstream of the TAM, generating 5- or 9-nt long 5' overhangs (Fig. 15G, Fig. 18). Exonuclease III mapping of a target substrate engaged by the dAwaIscB- ω RNA RNP showed that the RNP hindered exonuclease III treatment 19 nt upstream of the TAM on the target strand and 6 nt downstream of the targeted sequence on the non-target strand (Fig. 19) (Jinek, M. et al. (2014), *Science*. **343**, 1247997). Applicants also found that truncation of more than 4 aa of the PLMP domain of AwaIscB abolished cleavage activity (Fig. 49). Amino acid sequences for the proteins of Table 9 and the DNA sequences of their omegaRNAs correspond to SEQ ID NOs. 2059 to 2530.

Table 9. Contig accessions and sequence information for experimentally tested IscBs.

Contig Accession	Cluster ID	Key name	Source organism/taxonomy (if applicable)	Consensus TAM (if functional)
ADVG01000004		KraIscB-1	Ktedonobacter racemifer DSM 44963	ATAAA
ADVG01000001.1	44123	KraIscB-2	Ktedonobacter racemifer DSM 44963	N/A
ADVG01000001.1	52169	KraIscB-3	Ktedonobacter racemifer DSM 44963	ATG
ADVG01000001.1	30772	KraIscB-4	Ktedonobacter racemifer DSM 44963	ATMAA
ADVG01000004.1	31896	KraIscB-5	Ktedonobacter racemifer DSM 44963	GTGAA
ADVG01000004.1	30772	KraIscB-6	Ktedonobacter racemifer DSM 44963	N/A
AZFH01000048.1	41499	LeqIscB1	Lactobacillus equi DSM 15833 = JCM 10991	N/A

AZRL01000003.1	20578	PolIscB1	Petrogla olearia DSM 13574	NTGAD
AZRL01000004.1	49815	PolIscB2	Petrogla olearia DSM 13574	VTGA
AZRM01000012.1	41273	PmiIscB1	Petrogla miotherma DSM 10691	TTRA
BAMK01000010.1	26262	LcoIscB1	Lactobacillus composti DSM 18527	TTRA
BBBW01000007.1	40542	LecIscB1	Lactobacillus equicursoris DSM 19284 = JCM 14600	TTGAA
CP001393.1	23802	CbeIscB1	Caldicellulosiruptor bescii DSM 6725	GBGDDD
CP001393.1	36486	CbeIscB2	Caldicellulosiruptor bescii DSM 6725	N/A
CP001814.1	12901	SroIscB1	Streptosporangium roseum DSM 43021	DTG
CP001896.1	29247	AviIscB1	Allochromatium vinosum DSM 180	RTG
CP015439.1	49704	AamIscB1	Anoxybacillus amylolyticus strain DSM 15939 plasmid pDSM15939_1	ATGAH
CP026105.1	44770	PhoIscB1	Paraburkholderia hospita strain DSM 17164	ATG
CP026105.1	44770	PhoIscB2	Paraburkholderia hospita strain DSM 17164	ATGA
FMXB01000023.1	65074	MmiIscB1	Methanobrevibacter millerae strain DSM 16643	ATGA
FMXR01000006.1	46359	EoxIscB1	Eubacterium oxidoreducens strain DSM 3217	ATGA
FNDL01000011.1	18054	FchIscB1	Fervidobacterium changbaicum strain DSM 17883	N/A
FNDL01000023.1	32003	FchIscB2	Fervidobacterium changbaicum strain DSM 17883	GTGA
FNOW01000019.1	44770	AwaIscB1	Allochromatium warmingii strain DSM 173	ATGA
FOXRO1000041.1	18054	CfalIscB1	Caldicoprobacter faecalis strain DSM 20678	ATGAH
FPAI01000021.1	27545	HmiIscB1	Halolactibacillus miurensis strain DSM 17074	N/A
FPBV01000023.1	47441	AmaIscB1	Alicyclobacillus macrosporangiidus strain DSM 17980	RTGG
FQZM01000035.1	36667	DthIscB1	Desulfotomaculum thermosubterraneum DSM 16057	GTGA
FRDJ01000006.1	18054	FgoIscB1	Fervidobacterium gondwanense DSM 13020	ATGAH
FWXY01000026.1	22166	DvalIscB1	Desulfobacterium vacuolatum DSM 3385	NTAAA
JAAGRP010000061.1	53804	DhyIscB1	Desulfobacter hydrogenophilus strain DSM 3380	NNGA
JAFA01000051.1	44770	PnoIscB1	Paraburkholderia nodosa DSM 21604	ATGA
JFHC01000001.1	44770	CgIIscB1	Caballeronia glathei strain DSM 50014	ATGA
JMKS01000001.1	71441	EanIscB1	Exiguobacterium antarcticum DSM 14480	ATGA
JQSG02000004.1	44770	AprIscB1	Acidihalobacter prosperus strain DSM 5130	ATGA
KE386494.1	36667	CpoIscB1	Caldanaerobius polysaccharolyticus DSM 13641	GTGA
KE387209.1	17533	AauIscB1	Azohydromonas australica DSM 1124	ATG
KI535313.1	58923	EceIscB1	Enterococcus cecorum DSM 20682 = ATCC 43198	ATG
LMBW01000001.1	30772	BhuIscB1	Bacillus humi strain DSM 16318	ATAAA
LT607411.1	34693	MviIscB1	Micromonospora viridifaciens strain DSM 43909	ATGA
LT607751.1	34693	MsiIscB1	Micromonospora siamensis strain DSM 45097	NTG
LWAE01000002.1	54479	CmaIscB1	Clostridium magnum DSM 2767	N/A
LWAE01000013.1	54479	CmaIscB2	Clostridium magnum DSM 2767	N/A
NWBQ01000011.1	44770	MbiIscB1	Macromonas bipunctata strain DSM 12705	VTGA
NWBQ01000013.1	44770	MbiIscB2	Macromonas bipunctata strain DSM 12705	VTG

SAUN01000001.1	51289	NpoIscB1	Nonomuraea polychroma strain DSM 43925	N/A
SMBX01000003.1	44770	PsoIscB1	Paracandidimonas soli strain DSM 100048	ATGA
SMKY01000017.1	12901	AdaIscB1	Actinomadura darangshiensis strain DSM 45941	GTGAA
SMKY01000166.1	12901	AdaIscB2	Actinomadura darangshiensis strain DSM 45941	NTGAA
UESZ01000001.1	865	BheIscB1	Branchiibius hedensis strain DSM 22951	GGGGHG
VFEU01000027.1	44683	PrhIscB1	Pseudomonas rhodesiae strain DSM 14020	ATGA
QUIR01000011.1	21516	QuirIscB1	Ruminococcus sp. AM40-10AC	ATGA
JH724136.1	14084	BdoIscB1	Bacteroides dorei CL02T12C06	N/A
QSPK01000009.1	32498	EstIscB1	Eubacterium sp. TM05-53	N/A
FMDX01000002.1	29934	FmdxIscB1	uncultured Clostridium sp. isolate 2789STDY5834874	N/A
CP027167.1	44683	PaeIscB1	Pseudomonas aeruginosa strain AR_0356 plasmid unnamed3	ATGA
FMGK01000008.1	58923	FmgkIscB1	uncultured Clostridium sp. isolate 2789STDY5834935	ATMAA
QULK01000019.1	58923	RsoIscB1	Ruminococcus sp. OM05-10BH	N/A
KB976725.1	60626	BceIscB1	Bacillus cereus VD136	TTRA
RCAL01000070.1	40565	EcoIscB1	Escherichia coli strain B8110	N/A
QUIP01000016.1	6379	QuipIscB1	Ruminococcus sp. AM42-10AC	N/A
CP006668.1	44770	RpiIscB1	Ralstonia pickettii DTP0602	N/A
UKSF01000002.1	39989	KpnIscB1	Klebsiella pneumoniae strain EuSCAPE_IT215	NBRA
QUIS01000016.1	21516	QuisIscB1	Ruminococcus sp. AM36-17	NTGA
OEZH01000093.1	63259	CdiIscB1	Clostridioides difficile strain 173070	NTAAR
QTWT01000019.1	1420	QtwIscB1	Ruminococcus sp. AF18	N/A
CABVRB010000040.1	61670	AbuIscB1	Arcobacter butzleri isolate Ab_1711	ATGADRD
LFOD01000003.1	24531	McoIscB1	Mycolicibacterium conceptionense strain MLE	N/A
QUIO01000086.1	26022	QuioIscB1	Ruminococcus sp. AM42-11	N/A
DEGM01000152.1	69876	DegmIscB1	Lachnospiraceae bacterium UBA2826	N/A
OWZO01000071.1	6594	OwzoIscB1	human gut metagenome genome assembly, contig: NODE_724_length_33230_cov_4.750505, whole genome shotgun sequence SOURCE: WGS DATE: 2019-01-27 GENOME_ID: 10741 CONTIG_ID: 723 GENOME_ACC: ORRM01.1	NTAAA
UPEL01000159.1	54268	UpelIscB1	human gut metagenome genome assembly, contig: NODE_159_length_75563_cov_9.847632	N/A
OIZI01004561.1	1922	OiziIscB1	human gut metagenome genome assembly, contig: NODE_4561_length_4588_cov_3.294066	NNAAA
OWFL01000025.1	2221	OwflIscB1	human gut metagenome genome assembly, contig: NODE_25_length_213871_cov_13.923256, whole genome shotgun sequence SOURCE: WGS DATE: 2019-01-27 GENOME_ID: 11957 CONTIG_ID: 24 GENOME_ACC: OWFL01.1	NNAAA
UPFJ01002787.1	14824	UpfjIscB1	human gut metagenome genome assembly, contig: NODE_2787_length_13531_cov_3.087266	NTAAA

ORVI01000029.1	11690	OrviIscB1	human gut metagenome genome assembly, contig: NODE_29_length_154150_cov_4.224790, whole genome shotgun sequence SOURCE: WGS DATE: 2019-01-27 GENOME_ID: 10841 CONTIG_ID: 28 GENOME_ACC: ORVI01.1	N/A
UWST01001880.1	26698	UwstIscB1	mouse gut metagenome genome assembly, contig: NODE_1880_length_30065_cov_4.990836, whole genome shotgun sequence SOURCE: WGS DATE: 2019-06-06 GENOME_ID: 279225 CONTIG_ID: 1879 GENOME_ACC: UWST01.1	ATAA
OMBG01000014.1	52299	OmbgIscB1	human gut metagenome genome assembly, contig: NODE_14_length_122693_cov_23.179227	N/A
OGEU01000025.1	11690	OgeuIscB1	metagenome genome assembly, contig: NODE_25_length_150080_cov_8.882980	NWRRNA
AP019309.1	22690	EbaIscB1	Erysipelotrichaceae bacterium SG0102 DNA, complete genome SOURCE: NCBI_Prokaryotes DATE: 2019-01-27 GENOME_ID: 208690 CONTIG_ID: 0 GENOME_ACC: GCA_003925875.1_ASM392587v1_genomic	N/A
OIXA01.1	46682	HgmIscB1	human gut metagenome genome assembly, contig: NODE_137_length_98666_cov_9.003610, whole genome shotgun sequence SOURCE: WGS DATE: 2019-01-27 GENOME_ID: 6111 CONTIG_ID: 136 GENOME_ACC: OIXA01.1	N/A
OPXW01000461.1	49070	OpxwIscB1	human gut metagenome genome assembly, contig: NODE_461_length_28699_cov_2.357946, whole genome shotgun sequence SOURCE: WGS DATE: 2019-01-27 GENOME_ID: 9575 CONTIG_ID: 460 GENOME_ACC: OPXW01.1	N/A
OMGI01003682.1	49070	OmgIscB1	human gut metagenome genome assembly, contig: NODE_3682_length_6100_cov_3.653598, whole genome shotgun sequence SOURCE: WGS DATE: 2019-01-27 GENOME_ID: 8314 CONTIG_ID: 3681 GENOME_ACC: OMGI01.1	N/A
QUKU01000013.1	6594	QukuIscB1	Ruminococcus sp. OM08-7 OM08-7.Scaf13, whole genome shotgun sequence SOURCE: NCBI_Prokaryotes DATE: 2019-01-27 GENOME_ID: 172916 CONTIG_ID: 4 GENOME_ACC: GCA_003481515.1_ASM348151v1_genomic	N/A
QUIS01000025.1	21516	QuisIscB2	Ruminococcus sp. AM36-17 AM36-17.Scaf25, whole genome shotgun sequence SOURCE: NCBI_Prokaryotes DATE: 2019-01-27 GENOME_ID: 172850 CONTIG_ID: 17 GENOME_ACC: GCA_003480205.1_ASM348020v1_genomic	N/A
JACBGJ010000050	4393	LasIscB1	Lactobacillus salivarius strain KZ-NCB Lactobacillus_salivarius_contig_50, whole genome shotgun sequence SOURCE: NCBI_Prokaryotes DATE: 2020-10-27 GENOME_ID: 358408 CONTIG_ID: 45 GENOME_ACC: GCA_013391745.1_ASM1339174v1_genomic	NTGA
ORRM01000724.1	10049	OrrmIscB1	human gut metagenome genome assembly, contig: NODE_724_length_33230_cov_4.750505, whole genome shotgun sequence SOURCE: WGS DATE: 2019-01-27 GENOME_ID: 10741 CONTIG_ID: 723 GENOME_ACC: ORRM01.1	NNGNNR

Eukaryotic				
KY407659		IteIscB1	Ignatius tetrasporus culture UTEX:2012 chloroplast, complete genome	NG
CRISPR-associated IscB				
Ga0348337_018242			Delaware Bay aquatic sample metagenome	NAC

IscB employ multiple guide-encoding mechanisms

[01044] A distinct advantage of RNA-guided systems is that they allow an effector to target many substrates by simply reprogramming the RNA guide. One-way *IscB* evolved to use multiple guides is association with CRISPR arrays (Fig. 16A). However, given that *iscB* loci typically encode a single ω RNA, it is unclear how or even whether these systems achieve such modularity in general. By searching for ω RNAs not directly adjacent to *iscB* ORFs, Applicants uncovered three additional potential mechanisms for guide encoding and switching: ω RNA arrays, transposon expansion, and standalone, *trans*-acting ω RNAs (Fig. 16A). ω RNA arrays consist of multiple ω RNAs, each encompassing a distinct guide, separated by up to 200 bp, and are found in 15/3356 unique *IscB*/*IscB* loci (0.4%). Transposon expansion involves the insertion of nearly identical IS200/605 superfamily transposons in multiple locations, resulting in multiple loci per genome, each capable of expressing a nearly identical ω RNA scaffold with a unique guide (Fig. 20). By contrast, standalone ω RNAs, which show no detectable genomic associations with *iscB*, were more common and were found in multiple copies in some genomes. *Cis* ω RNAs from 95/3356 (2.8%) unique *IscB*/*IscB* loci were nearly identical ($\geq 95\%$ sequence identity) to distally encoded standalone ω RNAs (Fig. 50), implying that these standalone ω RNAs could encode guides used by *trans*-encoded *IscBs*.

[01045] Applicants tested this possibility by examining 10 standalone ω RNAs in the *K. racemifer* genome, (Fig. 16B), 9 of which were found to be expressed (Fig. 16C, Fig. 21). Of the 6 standalone ω RNAs tested, Applicants found that 5 could mediate RNA-guided DNA cleavage with a distally encoded *IscB* from the same genome (Fig. 16D), demonstrating that a single *IscB* can use multiple *trans*- encoded ω RNAs. Guides from many ω RNAs, both *IscB*-adjacent and *trans*-encoded, mostly target prokaryotic genomic sequences (61.5% genomic, 0.7% plasmid, 2.0% phage, 35.8% unmatched, $N=36323$), suggesting a non-defense function

for *IscB* systems (Fig. 50). In particular, Applicants found that more than a third of the ω RNAs (34.1%) targeted the same locus without the IS200/605 transposon insertion (Fig. 51).

Evolution and diversity of *IscB* systems

[01046] Applicants next investigated the evolutionary relationships between *IscB*, Cas9, and other homologous proteins to gain a broader insight into the evolution of RNA-guided mechanisms. In our search for proteins containing split RuvC domains, Applicants detected another group of shorter, ~350 aa *IscB* homologs that are also encoded in IS200/605 superfamily transposons. These proteins contain a PLMP domain and split RuvC but lack the HNH domain. Applicants renamed these proteins *IsrB* (Insertion sequence RuvC-like OrfB) to emphasize their distinct domain architecture, replacing the previous designation, *IscB1* (Kapitonov, V. et al. (2015), *J. Bacteriol.* **198**, 797–807). In addition to *IscB* and *IsrB*, Applicants identified a family of even smaller (~180 aa) proteins that only contained the PLMP domain and HNH domain but no RuvC domain, which Applicants named *IshB* (Insertion sequence HNH-like OrfB).

[01047] To investigate the relationships between these proteins, Applicants built a maximum likelihood (ML) tree from a multiple alignment of the split RuvC nuclease and BH domains using IQ-TREE 2 (Figs. 22, 31A, 52, Table 11) (Mihn, B. et al (2020), *Mol. Biol. Evol.* **37**, 1530–1534). The topology of the resulting tree was supported by several additional ML and Bayesian phylogenetic and robustness analyses ((Figs. 22, 32–33, 35, 52–56), see Supplementary Text for details). In the resulting tree, *IsrB*, *IscB*, and Cas9 formed distinct, strongly supported clades, suggesting that each of these nucleases originated from a unique evolutionary event (Figs. 31A, 32, 33A, 33C, 35, 54C–54D, and Supplementary Text). Applicants then analyzed the associations between each protein cluster and IS200/605 *tnpA* genes (Kapitonov, V. et al. (2015), *J. Bacteriol.* **198**, 797–807), ω RNAs, CRISPR-Cas adaptation genes (*cas1*, *cas2*, *cas4*, and *csn2*), CRISPR arrays upstream and downstream of the respective ORF, and CRISPR anti-repeats (Fig. 31A). As discussed above, *IscB* and *isrB* were rarely associated with CRISPR arrays and were not found to be associated with CRISPR-Cas adaptation genes. The *isrBs* are associated with structurally distinct ω RNAs. The *iscBs* are flanked by transposon ends similar to those mobilized by TnpA (Kapitonov, V. et al. (2015), *J. Bacteriol.* **198**, 797–807), but are only found near *tnpA* in 56/2811 of unique *IscB* loci (2.0%) (Figs. 31A, 57D).

[01048] Additionally, Applicants identified two distinct groups of Cas9s. The first is a new subtype, II-D, a group of relatively small *cas9*s (~700aa) that are not associated with any other known *cas* genes (Makarova, K., et al. (2020), *Nat. Rev. Microbiol.* **18**, 67–83). The second is a distinct clade branching from within the II-C subtype, which includes exceptionally large *cas9*s (>1700aa) that are associated with *tnpA* (Fig. 31A, Fig. 57). The *tnpA*- associated II-C loci often encompass unusually long DRs (more than 42bp in length) and in some cases encode HIRAN domain proteins between the *cas9* and other *cas* genes (Fig. 31A, Fig. 44). Predicted transposon ends surround various combinations of the *tnpA*, *cas* acquisition genes, and CRISPR arrays in these loci.

[01049] These phylogenetic and association analyses confirm that IS200/605 transposon-encoded IscBs and IsrBs share a common evolutionary history with Cas9 (Supplementary Text). Given the deep position of the IsrB clade in the tree (Fig. 31A) and the lack of the HNH domain, IsrBs likely represent the ancestral state, probably having evolved from the compact RuvC endonuclease (Majorek, K. et al. (2014), *Nucleic Acids Res.* **42**, 4160–4179). Almost all *isrBs* are associated with an ω RNA, suggesting that these systems became RNA- guided at an early stage of evolution, concomitantly with the insertions in the RuvC-like domain that are likely to be involved in complex formation with ω RNA. IsrB subsequently gained the HNH domain, possibly through insertion of another mobile element or recombination with a gene encoding an IshB-like protein, founding the IscB family (gray squares, Fig. 31A-31B, Supplementary Text).

[01050] CRISPR arrays emerged within IscB systems on multiple, independent occasions (black circles, Fig. 31A-31B). These short arrays consist of repeats that could have evolved by duplication of segments of the ancestral ω RNA. The resulting systems encompass a hybrid CRISPR- ω RNA that consists of a CRISPR array preceding a partial ω RNA. These CRISPR-associated IscB proteins likely also gained REC-like insertions between the RuvC-I and RuvC-II subdomains on a number of occasions, often contemporaneously with or shortly after the CRISPR association (white squares, Fig. 31A-31B, Fig. 58). In particular, one CRISPR-associated IscB cluster (cluster 2089) apparently founded the Cas9 family (Fig. 35) upon the loss of the hallmark PLMP domain (gray square, Fig. 31A-31B, Fig. 58). Moreover, the tracrRNAs of Subtype II-D, a deep branch in the Cas9 subtree (ML branch support: $\geq 97/100$, Bayesian posterior probability: 100%, Fig. 54B-D, Fig. 35), shows significant similarity to IscB ω RNAs (E-value 4.1e-8), suggesting that the Cas9 tracrRNA originally evolved from

ω RNA (Fig. 36). The continued evolution of Cas9 apparently involved the gain of additional REC-like insertions between the bridge helix and the RuvC-II domains resulting in increased protein size (Fig. 58). Finally, upon the association with the CRISPR adaptation machinery (*cas1*, *cas2*, and possibly *cas4*) (light blue circles, Fig. 31A-B), a burst of Cas9 diversification and widespread dispersion among bacteria via horizontal gene transfer followed, resulting in the evolution of multiple type II CRISPR subtypes.

[01051] Applicants also explored the evolutionary history of ω RNAs. By iteratively building a set of ω RNA profiles that spanned all major groups of ω RNAs associated with *iscBs* and *isrBs*, Applicants found that diverse ω RNAs are associated with almost all *iscBs* and *isrBs*. Moreover, different IsrB and IscB clades are associated with distinct ω RNA structures (Fig. 4A, 4C, Fig. 37, Fig. 38A, Fig. 52A). The transition from *isrB* to *iscB* was likely accompanied by loss of a second pseudoknot, the adaptor pseudoknot, between the transposon end region and the multi-stem loop in *isrB*-associated ω RNAs (light gray square, Fig. 31A-31C). The inverse relationship between the complexity of the ω RNA structure and the associated protein size is also reflected by the simplified ω RNA structures associated with clades of large IscBs and the even smaller tracrRNAs associated with large Cas9s (Fig. 31C, Fig. 38). Accession numbers and natural targets search data are shown in Table 10.

Table 10A-H. Accession numbers and natural target search data.

Table 10A. RuvC BH

N sites	208				
N proteins	1853				
Model	-LnL	df	AIC	AICc	BIC
LG	503754.123	3703	1014914.25	28446738.2	1027273.15
LG+I	501031.081	3704	1009470.16	28456110.2	1021832.4
LG+G4	484701.308	3704	976810.616	28423450.6	989172.857
LG+I+G4	484248.917	3705	975907.834	28437367.8	988273.413
LG+F+I+G4	486649.626	3724	980747.252	28724547.3	993176.243
WAG+I+G4	485319.069	3705	978048.137	28439508.1	990413.716
WAG+F+I+G4	486966.711	3724	981381.423	28725181.4	993810.414
JTT+I+G4	492218.062	3705	991846.123	28453306.1	1004211.7
JTT+F+I+G4	493619.025	3724	994686.05	28738486.1	1007115.04
JTTDCMut+I+G4	492121.288	3705	991652.576	28453112.6	1004018.15
JTTDCMut+F+I+G4	493571.492	3724	994590.984	28738391	1007019.98

DCMut+I+G4	495427.634	3705	998265.269	28459725.3	1010630.85
DCMut+F+I+G4	493953.2	3724	995354.4	28739154.4	1007783.39
VT+I+G4	484911.291	3705	977232.583	28438692.6	989598.161
VT+F+I+G4	486476.96	3724	980401.921	28724201.9	992830.913
PMB+I+G4	485970.164	3705	979350.329	28440810.3	991715.907
PMB+F+I+G4	487425.47	3724	982298.94	28726098.9	994727.932
Blosum62+I+G4	485128.966	3705	977667.933	28439127.9	990033.511
Blosum62+F+I+G4	487231.102	3724	981910.204	28725710.2	994339.195
Dayhoff+I+G4	495468.881	3705	998347.761	28459807.8	1010713.34
Dayhoff+F+I+G4	494003.659	3724	995455.317	28739255.3	1007884.31
mtREV+I+G4	517365.999	3705	1042142	28503602	1054507.58
mtREV+F+I+G4	497460.523	3724	1002369.05	28746169	1014798.04
mtART+I+G4	525196.939	3705	1057803.88	28519263.9	1070169.46
mtART+F+I+G4	502278.647	3724	1012005.29	28755805.3	1024434.29
mtZOA+I+G4	510552.349	3705	1028514.7	28489974.7	1040880.28
mtZOA+F+I+G4	496519.09	3724	1000486.18	28744286.2	1012915.17
mtMet+I+G4	517023.603	3705	1041457.21	28502917.2	1053822.79
mtMet+F+I+G4	498462.685	3724	1004373.37	28748173.4	1016802.36
mtVer+I+G4	525831.436	3705	1059072.87	28520532.9	1071438.45
mtVer+F+I+G4	509246.342	3724	1025940.69	28769740.7	1038369.68
mtInv+I+G4	516516.275	3705	1040442.55	28501902.6	1052808.13
mtInv+F+I+G4	493117.244	3724	993682.487	28737482.5	1006111.48
mtMAM+I+G4	534945.189	3705	1077300.38	28538760.4	1089665.96
mtMAM+F+I+G4	513646.512	3724	1034741.02	28778541	1047170.02
HIVb+I+G4	507589.225	3705	1022588.45	28484048.4	1034954.03
HIVb+F+I+G4	510843.95	3724	1029135.9	28772935.9	1041564.89
HIVw+I+G4	529355.482	3705	1066120.96	28527581	1078486.54
HIVw+F+I+G4	526209.981	3724	1059867.96	28803668	1072296.95
FLU+I+G4	503465.942	3705	1014341.88	28475801.9	1026707.46
FLU+F+I+G4	506544.44	3724	1020536.88	28764336.9	1032965.87
rtREV+I+G4	488028.849	3705	983467.697	28444927.7	995833.276
rtREV+F+I+G4	486920.363	3724	981288.725	28725088.7	993717.717
cpREV+I+G4	488546.347	3705	984502.694	28445962.7	996868.272
cpREV+F+I+G4	492889.794	3724	993227.587	28737027.6	1005656.58

Table 10B. RuvC BH HNH

N sites	315				
N proteins	1710				
Model	-LnL	df	AIC	AICc	BIC

LG	729020.94	3417	1464875.88	24823487.9	1477698.42
LG+I	725636.107	3418	1458108.21	24830392.2	1470934.51
LG+G4	700605.149	3418	1408046.3	24780330.3	1420872.59
LG+I+G4	699919.215	3419	1406676.43	24792636.4	1419506.48
LG+F+I+G4	701771.88	3438	1410419.76	25056983.8	1423321.1
WAG+I+G4	701096.463	3419	1409030.93	24794990.9	1421860.97
WAG+F+I+G4	701889.424	3438	1410654.85	25057218.8	1423556.19
JTT+I+G4	710934.601	3419	1428707.2	24814667.2	1441537.25
JTT+F+I+G4	710823.602	3438	1428523.2	25075087.2	1441424.55
JTTDCMut+I+G4	710751.156	3419	1428340.31	24814300.3	1441170.36
JTTDCMut+F+I+G4	710735.814	3438	1428347.63	25074911.6	1441248.97
DCMut+I+G4	714711.037	3419	1436260.07	24822220.1	1449090.12
DCMut+F+I+G4	711284.014	3438	1429444.03	25076008	1442345.37
VT+I+G4	700737.235	3419	1408312.47	24794272.5	1421142.52
VT+F+I+G4	701335.287	3438	1409546.58	25056110.6	1422447.92
PMB+I+G4	702726.44	3419	1412290.88	24798250.9	1425120.93
PMB+F+I+G4	702772.429	3438	1412420.86	25058984.9	1425322.2
Blosum62+I+G4	701617.707	3419	1410073.42	24796033.4	1422903.46
Blosum62+F+I+G4	702611.837	3438	1412099.67	25058663.7	1425001.02
Dayhoff+I+G4	714770.68	3419	1436379.36	24822339.4	1449209.41
Dayhoff+F+I+G4	711355.685	3438	1429587.37	25076151.4	1442488.72
mtREV+I+G4	749028.617	3419	1504895.23	24890855.2	1517725.28
mtREV+F+I+G4	720238.257	3438	1447352.51	25093916.5	1460253.86
mtART+I+G4	761576.069	3419	1529990.14	24915950.1	1542820.18
mtART+F+I+G4	728240.83	3438	1463357.66	25109921.7	1476259
mtZOA+I+G4	739982.765	3419	1486803.53	24872763.5	1499633.58
mtZOA+F+I+G4	718607.263	3438	1444090.53	25090654.5	1456991.87
mtMet+I+G4	748088.976	3419	1503015.95	24888976	1515846
mtMet+F+I+G4	719959.273	3438	1446794.55	25093358.5	1459695.89
mtVer+I+G4	758904.234	3419	1524646.47	24910606.5	1537476.51
mtVer+F+I+G4	735119.819	3438	1477115.64	25123679.6	1490016.98
mtInv+I+G4	748323.791	3419	1503485.58	24889445.6	1516315.63
mtInv+F+I+G4	713020.322	3438	1432916.64	25079480.6	1445817.99
mtMAM+I+G4	772787.688	3419	1552413.38	24938373.4	1565243.42
mtMAM+F+I+G4	742830.414	3438	1492536.83	25139100.8	1505438.17
HIVb+I+G4	731661.086	3419	1470160.17	24856120.2	1482990.22
HIVb+F+I+G4	733746.26	3438	1474368.52	25120932.5	1487269.86
HIVw+I+G4	764429.124	3419	1535696.25	24921656.2	1548526.29
HIVw+F+I+G4	755850.659	3438	1518577.32	25165141.3	1531478.66

FLU+I+G4	726306.5	3419	1459451	24845411	1472281.05
FLU+F+I+G4	726675.548	3438	1460227.1	25106791.1	1473128.44
rtREV+I+G4	705812.59	3419	1418463.18	24804423.2	1431293.23
rtREV+F+I+G4	703074.234	3438	1413024.47	25059588.5	1425925.81
cpREV+I+G4	706531.046	3419	1419900.09	24805860.1	1432730.14
cpREV+F+I+G4	711884.209	3438	1430644.42	25077208.4	1443545.76

Table 10C. RuvC BH HNH No IIB Domain

N sites	315				
N proteins	1671				
Model	-LnL	df	AIC	AICc	BIC
LG	710639.991	3339	1427957.98	23732478	1440487.82
LG+I	707348.838	3340	1421377.68	23739257.7	1433911.27
LG+G4	682599.624	3340	1371879.25	23689759.2	1384412.84
LG+I+G4	681961.367	3341	1370604.74	23701848.7	1383142.08
LG+F+I+G4	683776.089	3360	1374272.18	23960192.2	1386880.82
WAG+I+G4	683029.053	3341	1372740.11	23703984.1	1385277.45
WAG+F+I+G4	683787.589	3360	1374295.18	23960215.2	1386903.82
JTT+I+G4	692546.409	3341	1391774.82	23723018.8	1404312.16
JTT+F+I+G4	692512.293	3360	1391744.59	23977664.6	1404353.23
JTTDCMut+I+G4	692373.142	3341	1391428.28	23722672.3	1403965.63
JTTDCMut+F+I+G4	692433.37	3360	1391586.74	23977506.7	1404195.38
DCMut+I+G4	696327.276	3341	1399336.55	23730580.6	1411873.9
DCMut+F+I+G4	692986.323	3360	1392692.65	23978612.6	1405301.29
VT+I+G4	682678.293	3341	1372038.59	23703282.6	1384575.93
VT+F+I+G4	683276.456	3360	1373272.91	23959192.9	1385881.56
PMB+I+G4	684667.591	3341	1376017.18	23707261.2	1388554.53
PMB+F+I+G4	684710.622	3360	1376141.24	23962061.2	1388749.89
Blosum62+I+G4	683517.3	3341	1373716.6	23704960.6	1386253.95
Blosum62+F+I+G4	684507.936	3360	1375735.87	23961655.9	1388344.52
Dayhoff+I+G4	696385.628	3341	1399453.26	23730697.3	1411990.6
Dayhoff+F+I+G4	693056.563	3360	1392833.13	23978753.1	1405441.77
mtREV+I+G4	729901.526	3341	1466485.05	23797729.1	1479022.4
mtREV+F+I+G4	701737.505	3360	1410195.01	23996115	1422803.65
mtART+I+G4	742286.815	3341	1491255.63	23822499.6	1503792.98
mtART+F+I+G4	709647.414	3360	1426014.83	24011934.8	1438623.47
mtZOA+I+G4	721038.084	3341	1448758.17	23780002.2	1461295.51
mtZOA+F+I+G4	700219.93	3360	1407159.86	23993079.9	1419768.5
mtMet+I+G4	729039.574	3341	1464761.15	23796005.1	1477298.49

mtMet+F+I+G4	701509.141	3360	1409738.28	23995658.3	1422346.93
mtVer+I+G4	739625.403	3341	1485932.81	23817176.8	1498470.15
mtVer+F+I+G4	716440.076	3360	1439600.15	24025520.2	1452208.8
mtInv+I+G4	729219.014	3341	1465120.03	23796364	1477657.37
mtInv+F+I+G4	694761.665	3360	1396243.33	23982163.3	1408851.97
mtMAM+I+G4	753025.909	3341	1512733.82	23843977.8	1525271.16
mtMAM+F+I+G4	723890.679	3360	1454501.36	24040421.4	1467110
HIVb+I+G4	712957.289	3341	1432596.58	23763840.6	1445133.92
HIVb+F+I+G4	715070.081	3360	1436860.16	24022780.2	1449468.81
HIVw+I+G4	744863.94	3341	1496409.88	23827653.9	1508947.23
HIVw+F+I+G4	736648.004	3360	1480016.01	24065936	1492624.65
FLU+I+G4	707623.242	3341	1421928.49	23753172.5	1434465.83
FLU+F+I+G4	708080.883	3360	1422881.77	24008801.8	1435490.41
rtREV+I+G4	687596.012	3341	1381874.02	23713118	1394411.37
rtREV+F+I+G4	684941.592	3360	1376603.18	23962523.2	1389211.83
cpREV+I+G4	688288.268	3341	1383258.54	23714502.5	1395795.88
cpREV+F+I+G4	693509.306	3360	1393738.61	23979658.6	1406347.26

Table 10D. RuvC BH HNH IscB Plus early Cas

N sites	315				
N proteins	736				
Model	-LnL	df	AIC	AICc	BIC
LG	268338.373	1469	539614.746	4858474.75	545127.276
LG+I	267002.735	1470	536945.47	4861685.47	542461.751
LG+G4	257727.26	1470	518394.521	4843134.52	523910.803
LG+I+G4	257519.901	1471	517981.802	4848605.8	523501.836
LG+F+I+G4	258407.475	1490	519794.951	4962974.95	525386.284
WAG+I+G4	257657.092	1471	518256.185	4848880.19	523776.219
WAG+F+I+G4	258102.487	1490	519184.974	4962364.97	524776.307
JTT+I+G4	261401.16	1471	525744.319	4856368.32	531264.353
JTT+F+I+G4	261844.307	1490	526668.614	4969848.61	532259.948
JTTDCMut+I+G4	261316.21	1471	525574.42	4856198.42	531094.454
JTTDCMut+F+I+G4	261773.035	1490	526526.07	4969706.07	532117.403
DCMut+I+G4	262678.146	1471	528298.293	4858922.29	533818.327
DCMut+F+I+G4	261705.074	1490	526390.147	4969570.15	531981.48
VT+I+G4	257755.118	1471	518452.237	4849076.24	523972.271
VT+F+I+G4	258222.769	1490	519425.537	4962605.54	525016.871
PMB+I+G4	258496.869	1471	519935.738	4850559.74	525455.773
PMB+F+I+G4	258629.856	1490	520239.713	4963419.71	525831.046

Blosum62+I+G4	258168.988	1471	519279.977	4849903.98	524800.011
Blosum62+F+I+G4	258604.889	1490	520189.778	4963369.78	525781.111
Dayhoff+I+G4	262698.376	1471	528338.752	4858962.75	533858.786
Dayhoff+F+I+G4	261728.954	1490	526437.907	4969617.91	532029.24
mtREV+I+G4	276586.333	1471	556114.665	4886738.67	561634.7
mtREV+F+I+G4	265541.934	1490	534063.868	4977243.87	539655.201
mtART+I+G4	282239.447	1471	567420.895	4898044.9	572940.929
mtART+F+I+G4	269280.548	1490	541541.095	4984721.1	547132.429
mtZOA+I+G4	273850.752	1471	550643.504	4881267.5	556163.539
mtZOA+F+I+G4	265474.249	1490	533928.499	4977108.5	539519.832
mtMet+I+G4	277072.239	1471	557086.478	4887710.48	562606.512
mtMet+F+I+G4	265967.258	1490	534914.516	4978094.52	540505.849
mtVer+I+G4	280924.716	1471	564791.432	4895415.43	570311.466
mtVer+F+I+G4	271340.385	1490	545660.77	4988840.77	551252.103
mtInv+I+G4	277013.161	1471	556968.322	4887592.32	562488.357
mtInv+F+I+G4	263129.733	1490	529239.466	4972419.47	534830.799
mtMAM+I+G4	286245.798	1471	575433.597	4906057.6	580953.631
mtMAM+F+I+G4	274328.526	1490	551637.051	4994817.05	557228.385
HIVb+I+G4	269657.388	1471	542256.775	4872880.78	547776.81
HIVb+F+I+G4	270846.891	1490	544673.782	4987853.78	550265.115
HIVw+I+G4	281628.722	1471	566199.444	4896823.44	571719.478
HIVw+F+I+G4	278988.827	1490	560957.654	5004137.65	566548.987
FLU+I+G4	267371.953	1471	537685.905	4868309.91	543205.94
FLU+F+I+G4	267438.659	1490	537857.319	4981037.32	543448.652
rtREV+I+G4	259334.306	1471	521610.613	4852234.61	527130.647
rtREV+F+I+G4	258979.685	1490	520939.371	4964119.37	526530.704
cpREV+I+G4	260311.41	1471	523564.819	4854188.82	529084.853
cpREV+F+I+G4	261757.896	1490	526495.792	4969675.79	532087.125

Table 10E. Early Cas9 Amino Acid

N sites	1048				
N proteins	128				
Model	-LnL	df	AIC	AICc	BIC
LG	158546.793	253	317599.586	317761.456	318853.11
LG+I	158165.737	254	316839.474	317002.828	318097.952
LG+G4	153534.28	254	307576.559	307739.914	308835.037
LG+I+G4	153450.975	255	307411.95	307576.798	308675.383
LG+F+I+G4	153344.18	274	307236.361	307431.316	308593.932
WAG+I+G4	153438.079	255	307386.158	307551.006	308649.59

WAG+F+I+G4	153281.497	274	307110.994	307305.949	308468.565
JTT+I+G4	154955.321	255	310420.642	310585.491	311684.075
JTT+F+I+G4	154753.553	274	310055.105	310250.06	311412.676
JTTDCMut+I+G4	154922.992	255	310355.983	310520.832	311619.416
JTTDCMut+F+I+G4	154725.109	274	309998.219	310193.173	311355.79
DCMut+I+G4	155870.3	255	312250.6	312415.448	313514.032
DCMut+F+I+G4	155113.884	274	310775.767	310970.722	312133.338
VT+I+G4	153592.848	255	307695.696	307860.544	308959.128
VT+F+I+G4	153448.997	274	307445.995	307640.95	308803.566
PMB+I+G4	154434.683	255	309379.365	309544.214	310642.798
PMB+F+I+G4	154055.781	274	308659.563	308854.517	310017.134
Blosum62+I+G4	154250.423	255	309010.846	309175.695	310274.279
Blosum62+F+I+G4	153937.508	274	308423.015	308617.97	309780.586
Dayhoff+I+G4	155873.824	255	312257.649	312422.497	313521.082
Dayhoff+F+I+G4	155119.656	274	310787.312	310982.267	312144.883
mtREV+I+G4	165269.973	255	331049.945	331214.794	332313.378
mtREV+F+I+G4	157064.584	274	314677.168	314872.123	316034.739
mtART+I+G4	169405.3	255	339320.601	339485.449	340584.034
mtART+F+I+G4	159787.176	274	320122.351	320317.306	321479.923
mtZOA+I+G4	163921.253	255	328352.507	328517.355	329615.94
mtZOA+F+I+G4	157454.96	274	315457.92	315652.874	316815.491
mtMet+I+G4	166109.582	255	332729.164	332894.013	333992.597
mtMet+F+I+G4	157397.596	274	315343.191	315538.146	316700.762
mtVer+I+G4	167284.318	255	335078.637	335243.485	336342.07
mtVer+F+I+G4	159774.207	274	320096.413	320291.368	321453.984
mtInv+I+G4	166773.594	255	334057.187	334222.036	335320.62
mtInv+F+I+G4	156166.225	274	312880.451	313075.405	314238.022
mtMAM+I+G4	170224.016	255	340958.033	341122.881	342221.466
mtMAM+F+I+G4	161394.78	274	323337.56	323532.515	324695.131
HIVb+I+G4	158770.796	255	318051.592	318216.441	319315.025
HIVb+F+I+G4	158521.047	274	317590.094	317785.049	318947.665
HIVw+I+G4	165646.002	255	331802.005	331966.853	333065.438
HIVw+F+I+G4	162883.131	274	326314.262	326509.217	327671.833
FLU+I+G4	158550.963	255	317611.926	317776.774	318875.359
FLU+F+I+G4	157457.023	274	315462.047	315657.001	316819.618
rtREV+I+G4	154713.716	255	309937.431	310102.28	311200.864
rtREV+F+I+G4	153762.52	274	308073.04	308267.995	309430.611
cpREV+I+G4	154803.199	255	310116.399	310281.247	311379.832
cpREV+F+I+G4	154747.45	274	310042.901	310237.855	311400.472

Table 10F. Early Cas9 DNA

N sites	3039				
N proteins	128				
GTR+F	292394.76	261	585311.52	585360.769	586882.553
GTR+F+I	291454.796	262	583433.591	583483.235	585010.644
GTR+F+G4	279093.787	262	558711.574	558761.218	560288.626
GTR+F+I+G4	278956.868	263	558439.736	558489.777	560022.808
SYM+I+G4	279352.506	260	559225.011	559273.867	560790.025
TVM+F+I+G4	278968.492	262	558460.983	558510.627	560038.036
TVMe+I+G4	279407.399	259	559332.797	559381.261	560891.792
TIM3+F+I+G4	279049.294	261	558620.588	558669.837	560191.621
TIM3e+I+G4	279401.196	258	559318.391	559366.465	560871.367
TIM2+F+I+G4	279352.841	261	559227.682	559276.931	560798.715
TIM2e+I+G4	280434.163	258	561384.325	561432.399	562937.3
TIM+F+I+G4	279440.534	261	559403.068	559452.317	560974.101
TIMe+I+G4	280499.184	258	561514.369	561562.442	563067.344
TPM3u+F+I+G4	279064.274	260	558648.548	558697.403	560213.562
TPM3+F+I+G4	279064.274	260	558648.548	558697.403	560213.562
TPM2u+F+I+G4	279363.223	260	559246.446	559295.301	560811.459
TPM2+F+I+G4	279362.542	260	559245.084	559293.939	560810.097
K3Pu+F+I+G4	279454.965	260	559429.93	559478.785	560994.943
K3P+I+G4	280559.351	257	561632.702	561680.387	563179.657
TN+F+I+G4	279455.636	260	559431.272	559480.128	560996.286
TNe+I+G4	280496.481	257	561506.961	561554.646	563053.917
HKY+F+I+G4	279465.418	259	559448.837	559497.3	561007.831
K2P+I+G4	280565.775	256	561643.549	561690.848	563184.486
F81+F+I+G4	281411.443	258	563338.885	563386.958	564891.86
JC+I+G4	282299.492	255	565108.984	565155.897	566643.901

Table 10G. Omega RNA.

N sites	522				
N RNAs	711				
Model	-LnL	df	AIC	AICc	BIC
GTR+F	137969.864	1427	278793.729	4354305.73	284869.42
GTR+F+ASC	137969.861	1427	278793.721	4354305.72	284869.413
GTR+F+G4	132814.717	1428	268485.434	4349709.43	274565.383

GTR+F+ASC+G4	132814.709	1428	268485.418	4349709.42	274565.367
SYM+G4	132847.753	1425	268545.507	4332645.51	274612.683
SYM+ASC+G4	132847.753	1425	268545.507	4332645.51	274612.683
TVM+F+G4	132918.625	1427	268691.251	4344203.25	274766.942
TVM+F+ASC+G4	132918.625	1427	268691.251	4344203.25	274766.942
TVMe+G4	132862.651	1424	268573.302	4326973.3	274636.22
TVMe+ASC+G4	132862.651	1424	268573.302	4326973.3	274636.22
TIM3+F+G4	132865.718	1426	268583.436	4338387.44	274654.87
TIM3+F+ASC+G4	132865.718	1426	268583.436	4338387.44	274654.87
TIM3e+G4	132931.537	1423	268709.074	4321413.07	274767.735
TIM3e+ASC+G4	132931.537	1423	268709.074	4321413.07	274767.735
TIM2+F+G4	132882.127	1426	268616.255	4338420.26	274687.689
TIM2+F+ASC+G4	132882.127	1426	268616.255	4338420.26	274687.689
TIM2e+G4	132910.507	1423	268667.013	4321371.01	274725.674
TIM2e+ASC+G4	132910.507	1423	268667.013	4321371.01	274725.674
TIM+F+G4	132882.794	1426	268617.588	4338421.59	274689.022
TIM+F+ASC+G4	132882.794	1426	268617.588	4338421.59	274689.022
TIMe+G4	132943.006	1423	268732.012	4321436.01	274790.673
TIMe+ASC+G4	132943.006	1423	268732.012	4321436.01	274790.673
TPM3u+F+G4	132975.699	1425	268801.399	4332901.4	274868.575
TPM3u+F+ASC+G4	132975.699	1425	268801.399	4332901.4	274868.575
TPM3+F+G4	132975.699	1425	268801.399	4332901.4	274868.575
TPM3+F+ASC+G4	132975.699	1425	268801.399	4332901.4	274868.575
TPM2u+F+G4	132989.627	1425	268829.254	4332929.25	274896.43
TPM2u+F+ASC+G4	132989.627	1425	268829.254	4332929.25	274896.43
TPM2+F+G4	132989.627	1425	268829.254	4332929.25	274896.43
TPM2+F+ASC+G4	132989.627	1425	268829.254	4332929.25	274896.43
K3Pu+F+G4	132994.102	1425	268838.203	4332938.2	274905.379
K3Pu+F+ASC+G4	132994.102	1425	268838.203	4332938.2	274905.379
K3P+G4	132958.903	1422	268761.806	4315773.81	274816.21
K3P+ASC+G4	132958.903	1422	268761.806	4315773.81	274816.21
TN+F+G4	132909.484	1425	268668.968	4332768.97	274736.145
TN+F+ASC+G4	132909.484	1425	268668.968	4332768.97	274736.145
TNe+G4	132969.855	1422	268783.709	4315795.71	274838.113
TNe+ASC+G4	132969.857	1422	268783.714	4315795.71	274838.117
HKY+F+G4	133021.421	1424	268890.843	4327290.84	274953.761
HKY+F+ASC+G4	133021.421	1424	268890.843	4327290.84	274953.761
K2P+G4	132985.807	1421	268813.613	4310137.61	274863.759
K2P+ASC+G4	132985.807	1421	268813.613	4310137.61	274863.759

F81+F+G4	135134.431	1423	273114.861	4325818.86	279173.522
F81+F+ASC+G4	135134.431	1423	273114.861	4325818.86	279173.522
JC+G4	135161.994	1420	273163.988	4308803.99	279209.876
JC+ASC+G4	135161.994	1420	273163.988	4308803.99	279209.876

Table 10H. PLMP Domain

N sites	49				
N proteins	2050				
Model	-LnL	df	AIC	AICc	BIC
LG	122204.458	4071	252550.916	33406774.9	260252.517
LG+G4	118284.816	4072	244713.633	33415225.6	252417.125
LG+F+G4	118726.584	4091	245635.168	33726379.2	253374.605
WAG+G4	118217.215	4072	244578.43	33415090.4	252281.922
WAG+F+G4	118513.55	4091	245209.1	33725953.1	252948.536
JTT+G4	119140.743	4072	246425.487	33416937.5	254128.979
JTT+F+G4	119731.465	4091	247644.929	33728388.9	255384.366
JTTDCMut+G4	119116.725	4072	246377.451	33416889.5	254080.943
JTTDCMut+F+G4	119688.426	4091	247558.851	33728302.9	255298.288
DCMut+G4	120403.011	4072	248950.023	33419462	256653.515
DCMut+F+G4	120243.902	4091	248669.803	33729413.8	256409.24
VT+G4	117974.209	4072	244092.419	33414604.4	251795.911
VT+F+G4	118386.151	4091	244954.302	33725698.3	252693.739
PMB+G4	118297.184	4072	244738.368	33415250.4	252441.86
PMB+F+G4	118922.423	4091	246026.846	33726770.8	253766.283
Blosum62+G4	118283.672	4072	244711.344	33415223.3	252414.836
Blosum62+F+G4	118856.612	4091	245895.224	33726639.2	253634.661
Dayhoff+G4	120429.145	4072	249002.29	33419514.3	256705.782
Dayhoff+F+G4	120262.14	4091	248706.279	33729450.3	256445.716
mtREV+G4	127008.803	4072	262161.606	33432673.6	269865.098
mtREV+F+G4	121622.94	4091	251427.88	33732171.9	259167.317
mtART+G4	130752.472	4072	269648.944	33440160.9	277352.436
mtART+F+G4	124656.265	4091	257494.53	33738238.5	265233.967
mtZOA+G4	126309.547	4072	260763.094	33431275.1	268466.586
mtZOA+F+G4	122039.936	4091	252261.873	33733005.9	260001.31
mtMet+G4	127023.696	4072	262191.392	33432703.4	269894.884
mtMet+F+G4	121736.553	4091	251655.106	33732399.1	259394.542
mtVer+G4	129268.339	4072	266680.678	33437192.7	274384.171
mtVer+F+G4	124100.301	4091	256382.601	33737126.6	264122.038
mtInv+G4	126775.212	4072	261694.425	33432206.4	269397.917
mtInv+F+G4	120370.355	4091	248922.709	33729666.7	256662.146

mtMAM+G4	131950.378	4072	272044.757	33442556.8	279748.249
mtMAM+F+G4	126196.539	4091	260575.079	33741319.1	268314.515
HIVb+G4	122513.558	4072	253171.116	33423683.1	260874.608
HIVb+F+G4	123247.507	4091	254677.014	33735421	262416.451
HIVw+G4	127543.628	4072	263231.256	33433743.3	270934.749
HIVw+F+G4	126951.314	4091	262084.629	33742828.6	269824.066
FLU+G4	122300.586	4072	252745.171	33423257.2	260448.663
FLU+F+G4	123081.967	4091	254345.934	33735089.9	262085.371
rtREV+G4	119057.976	4072	246259.953	33416772	253963.445
rtREV+F+G4	119052.974	4091	246287.948	33727031.9	254027.385
cpREV+G4	119089.943	4072	246323.886	33416835.9	254027.378
cpREV+F+G4	119416.158	4091	247014.317	33727758.3	254753.754

IS200/IS605 elements encode diverse RNA-guided nucleases

[01052] In addition to the distinct succession of evolutionary events that yielded the abundant and diverse type II CRISPR systems, our phylogenetic analysis revealed several other events in the evolution of IscB and related proteins that led to the extant diversity, which Applicants sought to experimentally explore.

[01053] First, Applicants searched for IscB homologs in eukaryotic genomes and identified multiple *iscB* loci in the chloroplast genome of *Ignatius tetrasporus* UTEX B 2012, a terrestrial green alga (Fig. 39A-B, Fig. 59). Although the ORF is disrupted by multiple stop codons in most of these loci, one locus encodes an intact IscB (~50% aa identity to related prokaryotic IscBs) and a transcriptionally active ω RNA (Fig. 39C). This eukaryotic IscB cleaves DNA with a minimal NNG TAM (Fig. 39D), which differs from other characterized IscB TAMs (Fig. 48).

[01054] Second, Applicant investigated the clade of large IscBs, which contain a BH domain that is split in two by REC domain-like insertions (white squares, Figs. 31A, 39A). Applicants hypothesized that these insertions might enhance DNA unwinding, similarly to the REC lobe of Cas9 (Nishimisu, H. et al. (2014), *Cell*. **156**, 935–949) and would therefore facilitate genome editing in the complex landscape of eukaryotic chromatin structure. Applicants screened 6 large IscB proteins, using a pool of 12 guides each, for their ability to generate insertions/deletions (indels) in HEK293FT cells (see Methods, Table 11); one (OgeuIscB) produced appreciable indels (Fig. 39E-F, Fig. 40A). To further examine OgeuIscB activity, Applicants tested a range of guide lengths targeting 3 loci in the human genome and

found that OgeuIscB achieved the maximum indel rate with a 16 nt guide (Fig. 40B). On a panel of 46 sites in the human genome, Applicants found that OgeuIscB induced indels at 28 of these sites with varying efficiency up to 4.4% (Fig. 39G, Fig. 40C, Table 11). Thus, OgeuIscB seems a promising candidate for further development of IscB-based genome editing tools.

[01055] Third, Applicants experimentally characterized the putative nuclease activity of IsrB, the apparent ancestor of IscB (Fig. 39A). *K. racemifer* contains 5 *isrBs* associated with ω RNAs that are natively expressed (Fig. 39H, Fig. 41). Applicants found that the IsrB- ω RNA RNP nicks the non-target strand of a dsDNA substrate in a guide- and TAM-specific manner (Fig. 39I-J, Fig. 42), which is analogous to the activity of IscB upon inactivation of the HNH domain (Fig. 15F).

[01056] Finally, Applicants sought to determine if IS200/605 transposons in general harbor RNA-guided nucleases. In addition to the distinct IscB and IsrB families, most IS200/IS605 transposons

encode RuvC-like endonucleases of another family, TnpB, which is thought to be the ancestor of Cas12s, the type V CRISPR effectors (Fig. 39A) (Shmakov, S. et al. (2017), *Nat. Rev. Microbiol.* **15**, 169–182). Additionally, TnpB is the likely ancestor of larger proteins, Fanzors, encoded in diverse eukaryotic transposons (Fig. 39A) (Bao, W. et al (2013), *Mob. DNA* **4**, 12). The TnpB family, including Fanzor, is an order of magnitude more diverse than the IscB family; an HMMER search identified more than a million *tnpB* loci in publicly available prokaryotic genomes.

[01057] Applicants identified conserved non-coding regions immediately downstream of the CDS of many *tnpBs*, suggesting the presence of associated ncRNAs that could function as RNA guides (Fig. 60). Previous work has identified ncRNAs overlapping the 3'-end of *tnpB* genes in archaea and bacteria (Gomes-Filho, J. et al. (2015), *RNA Biol.* **12**, 490–500; Weinberg, Z., et al. (2017), *Nucleic Acids Res.* **45**, 10811–10823), but the function of these ncRNAs has not been characterized. Small RNA-seq of *K. racemifer* revealed native expression of a ncRNA overlapping the 3' end of the associated *tnpB* ORF (Fig. 39K), which Applicants classified as a distinct group of ω RNAs. The reverse complement of the KraTnpB ω RNA 3' end is nearly identical to the 5' of the ω RNA associated with some KraIscBs, a region that corresponds to the predicted transposon end in each locus.

[01058] Analysis of non-redundant loci containing *tnpB* genes that clustered with KraTnpB showed a drop of sequence conservation at the 3' end of the loci (Fig. 60), corresponding to the IS200/605 transposon end. Comparison to the small RNA-seq trace revealed expression beyond the conservation drop, indicating possible presence of a guide sequence in the transcript (Fig. 39M). *In vitro* plasmid cleavage assays for multiple TnpB proteins from this cluster using a reprogrammed guide demonstrated RNA-guided cleavage with a 5' TAM (Fig. 39N, Fig. 61). Applicants recombinantly purified a TnpB from *Alicyclobacillus macrosporangiidus* (AmaTnpB) and confirmed its reprogrammable RNA-guided dsDNA endonuclease activity (Fig. 39O, Fig. 61). Applicants also observed that AmaTnpB robustly cleaved target-containing ssDNA substrates (Fig. 39P) and non-specifically cleaved a collateral substrate upon recognition of dsDNA or ssDNA substrates (Fig. 39Q). Sequences corresponding to the 12-guide pooled experiments for OgeuIscB, OwzoIscB, OiziIscB, OwfIscB, UpfjIscB, and UwstIscB guides targeting human EMX1, DNMT1, VEGFA and FANCF genes in Table 11 are provided as SEQ ID NOs: 2551-2574 in the Sequence Listing.

Tables 11A-D. Guide sequences and statistical analysis relating to mammalian genome editing experiments

Table 11A. Guides and statistical analysis pertaining to Figure 40.

Guide	Sequence	Rep 1	Rep 2	Rep 3	Rep 4	p value (2-tailed T test)
DNMT1 12 bp	gAACATATGAGTG (SEQ ID NO: 2065)	0.001906505	0	0.009962144	0.014548629	0.723631176
DNMT1 14 bp	gTTAACATATGAGTG (SEQ ID NO: 2066)	0.077547934	0.067019273	0.084919567	0.060092543	5.29219E-05
DNMT1 16 bp	gGTTTAAACATATGAGTG (SEQ ID NO: 2067)	0.458614455	0.437795105	0.497773537	0.456562279	3.72633E-08
DNMT1 18 bp	gAAGTTTAAACATATGAGTG (SEQ ID NO: 2068)	0.098899209	0.065241889	0.047588542	0.057594754	0.001774195
DNMT1 20 bp	gACAAGTTTAAACATATGAGTG (SEQ ID NO: 2069)	0.030587276	0.029461899	0.031764077	0.062483728	0.008582107
DNMT1 22 bp	gAAACAAGTTTAAACATATGAGTG (SEQ ID NO: 2070)	0.043732575	0.03300502	0.046676261	0.007673861	0.02686037
DNMT1 24 bp	gACAAACAAGTTTAAACATATGAGTG (SEQ ID NO: 2071)	0.013979822	0.01884336	0.009703792	0.038013616	0.078758296
DNMT1 26 bp	gACACAAACAAGTTTAAACATATGAGTG (SEQ ID NO: 2072)	0.027471899	0.011556304	0.018715493	0.015009256	0.037924487
Non-targeting	gATGTCTTCTGGGACGAAGACAA (SEQ ID NO: 2073)	0	0.003144011	0	0.015752063	
TTL11 12 bp	gTAGGCATATGTC (SEQ ID NO: 2074)	0	0	0.012709306	0.008514503	0.146337514
TTL11 14 bp	gTCTAGGCATATGTC (SEQ ID NO: 2075)	0.627943485	0.785402618	0.633681552	0.611796195	3.25846E-06

TTL11 16 bp	gATTCTAGGCATATGTC (SEQ ID NO: 2076)	0.746360468	0.772388627	0.60119664	0.606760197	5.34844E-06
TTL11 18 bp	gAAATTCTAGGCATATGTC (SEQ ID NO: 2077)	0.590834854	0.38052957	0.391436718	0.37120723	0.000172577
TTL11 20 bp	gCAAAATTCTAGGCATATGTC (SEQ ID NO: 2078)	0.473175608	0.37771732	0.258312202	0.284159823	0.000384218
TTL11 22 bp	gACCAAAATTCTAGGCATATGTC (SEQ ID NO: 2079)	0.503107428	0.475723677	0.534404067	0.594416729	8.35011E-07
TTL11 24 bp	gGCACCAAAATTCTAGGCATATGTC (SEQ ID NO: 2080)	0.664339872	0.535442787	0.521896007	0.565888659	2.03162E-06
TTL11 26 bp	gAAGCACCAAAATTCTAGGCATATGT C (SEQ ID NO: 2081)	0.482984511	0.397657213	0.441831108	0.521148952	2.35984E-06
Non-targeting	gATGTCTTCCTGGGACGAAGACAA (SEQ ID NO: 2082)	0	0	0	0	
VEGFA 12 bp	gGAGTGAACGAGA (SEQ ID NO: 2083)	0.020806087	0.010516629	0.005556482	0.027097648	0.01682896
VEGFA 14 bp	gAAGAGTGAACGAGA (SEQ ID NO: 2084)	3.750390016	5.007405318	4.862240882	4.29670913	4.40473E-06
VEGFA 16 bp	gAAAAGAGTGAACGAGA (SEQ ID NO: 2085)	3.937613652	5.037626945	4.71507566	4.029523838	3.03492E-06
VEGFA 18 bp	gCAAAAAGAGTGAACGAGA (SEQ ID NO: 2086)	3.013126492	4.913294798	4.708655706	4.563681991	6.17349E-05
VEGFA 20 bp	gATCAAAAAGAGTGAACGAGA (SEQ ID NO: 2087)	2.61686747	3.960702406	4.613233924	3.3221303	0.000147162
VEGFA 22 bp	gTCATCAAAAAGAGTGAACGAGA (SEQ ID NO: 2088)	3.347833504	5.021651323	4.349930621	4.429301533	1.71866E-05
VEGFA 24 bp	gCATCATCAAAAAGAGTGAACGAGA (SEQ ID NO: 2089)	2.717800404	3.867531264	4.314221941	3.037901182	7.8654E-05
VEGFA 26 bp	gAGCATCATCAAAAAGAGTGAACGAG A (SEQ ID NO: 2090)	2.229058873	3.562953421	3.555705185	2.759688398	8.82134E-05
Non-targeting	gATGTCTTCCTGGGACGAAGACAA (SEQ ID NO: 2091)	0	0	0	0	

Table 11B. Guides and statistical analysis pertaining to Figure 39.

Guide	Sequence	Targeting Rep 1	Targeting Rep 2	Targeting Rep 3	Targeting Rep 4	p value (2-tailed T test)
DNMT1_g1	gAAAAGACGAGGATGA A (SEQ ID NO: 2092)	1.072633964	0.65965914		0.827668403	0.000373076
DNMT1_g2	gAGACAGCTTAACAGA A (SEQ ID NO: 2093)	0.314819763	0.304418458	0.260612729	0.242899	3.73442E-06
DNMT1_g3	gGTTTAACATATGAGT G (SEQ ID NO: 2094)	0.278377062	0.273855622	0.162021476	0.111194092	0.003153651
DNMT1_g4	gCCTAAGGCCCTTTTC (SEQ ID NO: 2095)	0.177124959	0.164299743	0.171190645	0.131435982	0.000102901
DNMT1_g5	gAGAAAAAGAACCTGA A (SEQ ID NO: 2096)	0.163750933	0.143394874	0.203416045	0.178246635	1.03199E-05
DNMT1_g6	gCCCCGGTTGGTCTTAC (SEQ ID NO: 2097)	0.061835518	0.038334611	0.027352765	0.027261752	0.165463548
VEGFA_g1	gAAAAGAGTGAACGAG A (SEQ ID NO: 2098)	2.256517612	1.908146202	2.280491183	1.766814372	3.70939E-06
VEGFA_g2	gATAGAGCAAGACAAG A (SEQ ID NO: 2099)	0.710804855	0.60404738	0.564880423	0.431536468	5.81344E-05
VEGFA_g3	gGGAGGCTCAAAGAGG C (SEQ ID NO: 2100)	0.04783926	0.049239138	0.046950456	0.048631681	4.63616E-07

VEGFA_g4	gCCCTTCAGATCAGCTT (SEQ ID NO: 2101)	0	0.007052311	0	0.005182958	0.141402755
VEGFA_g5	gGCTGTTTCAGGTCTCTG (SEQ ID NO: 2102)	0.030619936	0.005038799	0.014452259	0.020447807	0.020315011
VEGFA_g6	gAAGGCCGCACAGCTA G (SEQ ID NO: 2103)	0	0	0.007665038	0	0.070664369
EMX1_g1	gAATGGTGGAAACACA G (SEQ ID NO: 2104)	1.958384333	1.548672566	1.803849158	1.533513564	3.06047E-06
EMX1_g2	gGGACATGGCAGATAA T (SEQ ID NO: 2105)	0.263425281	0.265922336	0.264294907	0.372616029	3.92584E-05
EMX1_g3	gCCTGACTCTGCAAAG C (SEQ ID NO: 2106)	0.116292592	0.175527426	0.078441216	0.123543357	0.000951497
EMX1_g4	gAATAGCAGATTATTC C (SEQ ID NO: 2107)	0.045621533	0.02685562	0.040665904	0.074069854	0.003261575
EMX1_g5	gCTAGTCCCTTCCCTTT (SEQ ID NO: 2108)	0.027511578	0.096091149	0.043548834	0.030868281	0.029757632
EMX1_g6	gAGCTTTTTCCTGCAG (SEQ ID NO: 2109)	0.016334309	0.017236246	0.013600122	0.004832941	0.017822346
RNF2_g1	gTATTATACCTGCACGA (SEQ ID NO: 2110)	0.154957248	0.178900002	0.227125243	0.143247385	0.002072584
RNF2_g2	gAGAATAAGTTGAGAA A (SEQ ID NO: 2111)	0.022168736	0.035699574	0.024923671	0.028511961	0.000204006
RNF2_g3	gCTTCCTTCCAAGGTCA (SEQ ID NO: 2112)	0.011156564	0	0.006029727	0	0.116248624
RNF2_g4	gTTTCTATCTGTAAAT (SEQ ID NO: 2113)	0.012779961	0.01581878	0.003229766	0.003604123	0.410974349
RNF2_g5	gATTTTACCTTTTTCAA (SEQ ID NO: 2114)	0.012835323	0.023784041	0.032782407	0.01409344	0.000267325
RNF2_g6	gAATTTACTTTTGAAA (SEQ ID NO: 2115)	0.011674878	0.028580851	0.014731446	0.017115961	0.000118098
CXCR4_g1	gCAAAGCCCAAAGTGG T (SEQ ID NO: 2116)	0.134122288	0.108800774	0.108800774	0.100255195	6.4896E-06
CXCR4_g2	gAGTGAATCACGTAAAG (SEQ ID NO: 2117)	0.07815553	0.121625552	0.068431478	0.068431478	0.000784235
CXCR4_g3	gTTACAAAATTCTTTGT (SEQ ID NO: 2118)	0	0.006548038	0.006548038	0.017648443	0.080595539
TTLL11_g1	gATTCTAGGCATATGTC (SEQ ID NO: 2119)	0.366022099	0.437177085	0.429769392	0.403901093	2.41477E-07
TTLL11_g2	gCAAGGCAGAGCCACG G (SEQ ID NO: 2120)	0.051463966	0.039233636	0.053996591	0.039527112	0.014096134
TTLL11_g3	gATTCAATTACTACCCA (SEQ ID NO: 2121)	0	0.023314099	0.004305983	0	0.260846138
TTLL11_g4	gTTTCTATGACATATGC (SEQ ID NO: 2122)	0	0	0	0	N/A
TTLL11_g5	gATAATTGTTTCTATTC (SEQ ID NO: 2123)	0.004805613	0	0	0	0.99710752
TTLL11_g6	gAGGAGGCTGTAAATC T (SEQ ID NO: 2124)	0	0.01531941	0.008456302	0.00873744	0.286942892
DYNC1H1_g1	gCATAAAGTAACAAAA C (SEQ ID NO: 2125)	0.166073072	0.156347128	0.164880462	0.151315555	6.7371E-08
DYNC1H1_g2	gATGTTTACAAGATAG T (SEQ ID NO: 2126)	0.103883537	0.106560636	0.09729619	0.140374957	2.65811E-05
DYNC1H1_g3	gAGTCTGGCAAGGCAG A (SEQ ID NO: 2127)	0.084957968	0.040483004	0.043899639	0.03806735	0.0047032

DYNC1H1_g4	gACCTGCCCTAGAAAT A (SEQ ID NO: 2128)	0.084638172	0.035884094	0.07242628	0.08841733	0.003420442
DYNC1H1_g5	gCTAAACTAACCTGC C (SEQ ID NO: 2129)	0.024642681	0.060848401	0.03506434 3	0.035799523	0.018308875
DYNC1H1_g6	gGCAGTGCATTTCACTA (SEQ ID NO: 2130)	0.005725082	0.003175309	0.00339558 6	0.00561435	0.229661813
DYNC1H1_g7	gTTAATGGTTTTCACAT (SEQ ID NO: 2131)	0.003866527	0.008701706	0.01677782	0.020944602	0.103120809
ALDH1A3_g1	gAGTGAAGAAGGAGA T (SEQ ID NO: 2132)	1.137739046	1.119515885	0.91663029 2	1.940891045	0.001319886
ALDH1A3_g2	gGCTCTGCAGGAACAG G (SEQ ID NO: 2133)	0.234226321	0.184977578	0.22123893 8	0.14900909	4.93864E-05
ALDH1A3_g3	gTAAATAAATTTGCTC (SEQ ID NO: 2134)	0.171969046	0.154294531	0.06777363 6	0.022326412	0.026244571
ALDH1A3_g4	gAGAAGGCAGCTTTCT G (SEQ ID NO: 2135)	0	0	0	0	N/A
ALDH1A3_g5	gTCTGGCAGAAGACAC T (SEQ ID NO: 2136)	0.004161119	0.011395579	0.00623571	0.016026497	0.598222613
ALDH1A3_g6	gCCAGATTTCTTTCTC (SEQ ID NO: 2137)	0.021824928	0.018056653	0	0.006874742	0.463508461

Table 11C. Guides and statistical analysis for Figure 40.

Guide	Sequence	Non-targeting Rep 1	Non-targeting Rep 2	Non- targeting Rep 3	Non- targeting Rep 4	p value (2-tailed T test)
DNMT1_g1	gAAAAGACGAGGATGAA (SEQ ID NO: 2138)	0.002401768	0.002041608	0	0.00278404 2	0.000373
DNMT1_g2	gAGACAGCTTAACAGAA (SEQ ID NO: 2139)	0.003923415	0	0.006171982	0.00466330 9	3.73E-06
DNMT1_g3	gGTTTAACATATGAGTG (SEQ ID NO: 2140)	0.00784683	0.002726653	0.012343965	0.01088105 5	0.003154
DNMT1_g4	gCCTAAGGCCCTTTTC (SEQ ID NO: 2141)	0.062927552	0.042503864	0.066338785	0.04684139 9	0.000103
DNMT1_g5	gAGAAAAAGAACCTGAA (SEQ ID NO: 2142)	0.002401768	0.002041608	0	0.00278404 2	1.03E-05
DNMT1_g6	gCCCCGTTGGTCTTAC (SEQ ID NO: 2143)	0.071047236	0.042503864	0.073201418	0.04258309	0.165464
VEGFA_g1	gAAAAGAGTGAACGAGA (SEQ ID NO: 2144)	0.002895529	0	0	0	3.71E-06
VEGFA_g2	gATAGAGCAAGACAAGA (SEQ ID NO: 2145)	0.001232666	0.003634381	0	0	5.81E-05
VEGFA_g3	gGGAGGCTCAAAGAGGC (SEQ ID NO: 2146)	0.010686615	0.004547625	0.004950863	0.01059294	4.64E-07
VEGFA_g4	gCCCTTCAGATCAGCTT (SEQ ID NO: 2147)	0	0	0	0	0.141403
VEGFA_g5	gGCTGTTTCAGGTCTCTG (SEQ ID NO: 2148)	0	0.002814048	0	0	0.020315
VEGFA_g6	gAAGGCCGCACAGCTAG (SEQ ID NO: 2149)	0.024044884	0.006821438	0.004950863	0.01324117 5	0.070664
EMX1_g1	gAATGGTGGAAACACAG (SEQ ID NO: 2150)	0	0	0	0	3.06E-06
EMX1_g2	gGGACATGGCAGATAAT (SEQ ID NO: 2151)	0	0.005596911	0	0.00235571 3	3.93E-05
EMX1_g3	gCCTGACTCTGCAAAGC (SEQ ID NO: 2152)	0	0	0.009628655	0	0.000951

EMX1_g4	gAATAGCAGATTATTCC (SEQ ID NO: 2153)	0	0	0	0	0.003262
EMX1_g5	gCTAGTCCCTTCCCTTT (SEQ ID NO: 2154)	0	0	0.00947538	0.006171506	0.029758
EMX1_g6	gAGCTTTTTCCCTGCAG (SEQ ID NO: 2155)	0.005697194	0.003893828	0.002665849	0	0.017822
RNF2_g1	gTATTATACCTGCACGA (SEQ ID NO: 2156)	0.05907765	0.083386145	0.08737758	0.063398845	0.002073
RNF2_g2	gAGAATAAGTTGAGAAA (SEQ ID NO: 2157)	0.05907765	0.073379807	0.076455383	0.063398845	0.000204
RNF2_g3	gCTTCCTTCCAAGGTCA (SEQ ID NO: 2158)	0.028760426	0	0.023276564	0.015143485	0.116249
RNF2_g4	gTTTCTATCTGTAAAAT (SEQ ID NO: 2159)	0.017975266	0	0.023276564	0.015143485	0.410974
RNF2_g5	gATTTTACCTTTTTTCAA (SEQ ID NO: 2160)	0.062770003	0.080050699	0.083736848	0.066735627	0.000267
RNF2_g6	gAATTTACTTTTTGAAA (SEQ ID NO: 2161)	0.062770003	0.080050699	0.083736848	0.066735627	0.000118
CXCR4_g1	gCAAAGCCCAAAGTGGT (SEQ ID NO: 2162)	0.008136697	0	0	0	6.49E-06
CXCR4_g2	gAGTGAATCACGTAAAG (SEQ ID NO: 2163)	0.008136697	0	0	0.006918261	0.000784
CXCR4_g3	gTTACAAAATTCTTTGT (SEQ ID NO: 2164)	0	0	0	0	0.080596
TTLL11_g1	gATTCTAGGCATATGTC (SEQ ID NO: 2165)	0	0	0	0	2.41E-07
TTLL11_g2	gCAAGGCAGAGCCACGG (SEQ ID NO: 2166)	0.035417518	0.010157647	0.01585504	0.028010551	0.014096
TTLL11_g3	gATTCAATTACTACCCA (SEQ ID NO: 2167)	0	0	0	0	0.260846
TTLL11_g4	gTTTCTATGACATATGC (SEQ ID NO: 2168)	0	0	0	0	N/A
TTLL11_g5	gATAATTGTTTCTATTC (SEQ ID NO: 2169)	0	0	0	0.00478	0.997108
TTLL11_g6	gAGGAGGCTGTAAATCT (SEQ ID NO: 2170)	0.029968669	0.008126117	0.004530011	0.021007913	0.286943
DYNCH1_g1	gCATAAAGTAACAAAAC (SEQ ID NO: 2171)	0.021546236	0.022641199	0.015949281	0.028466374	6.74E-08
DYNCH1_g2	gATGTTTACAAGATAGT (SEQ ID NO: 2172)	0.001451948	0.001450074	0.001473796	0.001739584	2.66E-05
DYNCH1_g3	gAGTCTGGCAAGGCAGA (SEQ ID NO: 2173)	0.001451948	0.004350222	0.001473796	0.005218753	0.004703
DYNCH1_g4	gACCTGCCCTAGAAATA (SEQ ID NO: 2174)	0.010387991	0.018835939	0.014041493	0.011356324	0.00342
DYNCH1_g5	gCTAAAACTAACCTGCC (SEQ ID NO: 2175)	0.010387991	0.018835939	0.014041493	0.011356324	0.018309
DYNCH1_g6	gGCAGTGCATTTCACTA (SEQ ID NO: 2176)	0	0.011301563	0.010531119	0.011356324	0.229662
DYNCH1_g7	gTTAATGGTTTTTCACAT (SEQ ID NO: 2177)	0.019587488	0.022641199	0.015949281	0.026094176	0.103121
ALDH1A3_g1	gAGTGGAAGAAGGAGAT (SEQ ID NO: 2178)	0	0	0	0	0.00132

ALDH1A3_g2	gGCTCTGCAGGAACAGG (SEQ ID NO: 2179)	0	0	0	0	4.94E-05
ALDH1A3_g3	gTAAAATAAATTTGCTC (SEQ ID NO: 2180)	0	0	0	0	0.026245
ALDH1A3_g4	gAGAAGGCAGCTTTCTG (SEQ ID NO: 2181)	0	0	0	0	N/A
ALDH1A3_g5	gTCTGGCAGAAGACACT (SEQ ID NO: 2182)	0.004822996	0.006558524	0.014718509	0.022568269	0.598223
ALDH1A3_g6	gCCAGATTCTTTCTC (SEQ ID NO: 2183)	0	0.004372349	0.009812339	0.014105168	0.463508

Table 11D. 12-Guide Pool Methods

Key	Gene	Guide Sequence (OgeuIscB)	OgeuIscB indel rate (%)	Guide Sequence (OwzoIscB, OiziIscB, OwflIscB, UpfjIscB, UwstIscB)	OwzoIscB indel rate (%)	OiziIscB indel rate (%)	OwflIscB indel rate (%)	UpfjIscB indel rate (%)	UwstIscB indel rate (%)
1	EMX1	CGGTTCCGCA GGACCCAGGG (SEQ ID NO: 2184)	0.0173869 13	CCCCCTTCCC TATGGGAATA (SEQ ID NO: 2185)	0.018779 343	0.0452488 69	0	0	0
2	EMX1	AGTGGTGCCTG GAAAATAAA (SEQ ID NO: 2186)	0.0328299 41	ACTACAGTGG TGCCTGGAAA (SEQ ID NO: 2187)	0.018779 343	0.0452488 69	0	0.01295 4207	0
3	EMX1	CCCCCTTCCC TATGGGAATA (SEQ ID NO: 2188)	0.0492449 11	ATGCATATAC CAGTTTGTGG (SEQ ID NO: 2189)	0.003716 336	0.0035553 59	0.0037 13492	0.00391 324	0.00 3742
4	DNMT1	ACCGGGAAGT GAATGGACGT (SEQ ID NO: 2190)	0.0690698 59	ACGAATTTCT GCAAACAGAA (SEQ ID NO: 2191)	0.004224 489	0.0052655 6	0.0034 45751	0.00194 2936	0.00 568
5	DNMT1	CTCCAAGGAC AAATCTTTAT (SEQ ID NO: 2192)	0.0624917 77	TTTATTTTAGC TGAAGGGAA (SEQ ID NO: 2193)	0.005869 406	0.0069495 78	0.0045 43917	0.00865 4182	0.00 4322
6	DNMT1	CCAAGCAAGA AGTGAAGCCC (SEQ ID NO: 2194)	0.0581063 9	TTTATTTCCCT TCAGCTAAA (SEQ ID NO: 2195)	0.005869 406	0.0069495 78	0.0045 43917	0.00865 4182	0.00 4322
7	VEGFA	TCCCAAAGAT GCCACCTGC (SEQ ID NO: 2196)	0.0266082 74	GAGAGCAAAA GATACATCTC (SEQ ID NO: 2197)	0	0	0	0.77519 3798	0
8	VEGFA	ATCAAAAAGA GTGAACGAGA (SEQ ID NO: 2198)	0.1969903 6	ATTTGTACCG GTTTTGTAT (SEQ ID NO: 2199)	0	0	0	0.77519 3798	0
9	VEGFA	TGGTGGTCTG GATAAAAGAA (SEQ ID NO: 2200)	0.1812544 44	TGTGCCCATT GGTGGTCTGG (SEQ ID NO: 2201)	0.006639 886	0.0039954 81	0.0078 02074	0.00517 844	0.01 231
10	FANCF	CGCAGAGAGT CGCCGTCTCC (SEQ ID NO: 2202)	0.0014245 47	TATTCCTGAC ACTGCCAGGA (SEQ ID NO: 2203)	0	0.0187492 32	0	0.01737 3176	0.00 3696

11	FANC F	TTTAAAGAAA AAGCAGCTTT (SEQ ID NO: 2204)	0.0649471 14	GGGACAGAAA ACCTAGAAAA (SEQ ID NO: 2205)	0	0.0213353 33	0	0.01895 2555	0.01 1088
12	FANC F	TCTGTCCCTCC CTCAGTAGT (SEQ ID NO: 2206)	0.0417517 16	TTTAAGAGCAT CGAACAATA (SEQ ID NO: 2207)	0	0	0	0	0

Tables 12A-G. Accession numbers and position information for loci displayed in Figures.

Table 12A. Pertains to Figure 9C, 9F.

Subpanel	Contig accession	Feature of interest	start	end	strand
9C	Ga0348337_018242	CRISPR-associated IscB	802	2175	1
9F	NZ_ADVG01000001	IscB	305647	306958	-1

Table 12B. Pertains to Figure 16A.

Figure Subpanel	Contig accession	Feature of interest	start	end	strand
16A (P. excrementihominis strain BIOML-A6)	WNCH01000015.1	IscB	57970	59260	1
16A (Prevotellaceae bacterium UBA3839)	DGHV01000076.1	IscB	29502	30768	-1
16A (K. racemifer - transposon expansion, left)	NZ_ADVG01000003	IscB	1805791	1807081	-1
16A (K. racemifer - transposon expansion, right)	NZ_ADVG01000001	IscB	2708268	2709558	1
16A (K. racemifer - cis ωRNA)	NZ_ADVG01000001	IscB	2516461	2517751	-1
16A (K. racemifer - trans ωRNA), B	NZ_ADVG01000001	ωRNA	704325	704688	1

Table 12C. Pertains to Figure 31B.

Figure Subpanel	Contig accession	Feature of interest	start	end	strand
31B (K. racemifer - IsrB)	NZ_ADVG01000004	IsrB	697010	698075	1
31B (K. racemifer - IscB)	NZ_ADVG01000003	IscB	1400692	1401982	1
31B (Bacteroidales RUG11478)	CACXPM01000003.1	IscB	91847	93110	-1
31B (Tissierellia AS22ysBPME_61)	JAAZKS010000250.1	IscB	1343	2939	-1
31B (Armatimonadota)	Ga0315277_10040887	Cas9	1829	3848	1
31B (Nitrospirae RBG_13_39_12)	MHEF01000135.1	Cas9	1737	4026	1
31B (C. jejuni strain NCTC 11168)	SZUC01000003.1	Cas9	186809	189764	-1

Table 12D. Pertains to Figure 39C,H.

Figure Subpanel	Contig accession	Feature of interest	start	end	strand
39C	KY407659	IscB	21119	22433	-1
39H	NZ_ADVG01000001	IscB	978948	980013	-1

Table 12E. Pertains to Figure 51A-C

Figure Subpanel	Contig accession	Feature of interest	start	end	strand
51A,B,C (top locus)	NDEV01000101	class III alcohol dehydrogenase	5301	6411	1
51A,B,C (bottom locus)	RCAL01000070	IscB	14393	15798	1
51A,B,C (bottom locus)	RCAL01000070	class III alcohol dehydrogenase	15794	16904	1

Table 12F. Pertains to Figure 44.

From top to bottom	Contig accession	Feature of interest	start	end	strand
Locus 1	FZOK01000004.1	Cas9	172500	177072	-1
Locus 2	CP003281.1	Cas9	3445310	3449369	-1
Locus 3	FNXE01000009.1	Cas9	43590	47886	1
Locus 4	AQHR01000065.1	Cas9	18175	22762	1
Locus 5	QQBA01000002.1	Cas9	136771	141220	-1
Locus 6	CP007035.1	Cas9	301950	306231	1

Table 12G. Pertains to Figure 61A.

Figure Subpanel	Database	Contig accession	Feature of interest	start	end	strand
61A	NCBI	JACHNC010000001	tnpB	9479671	9480832	-1

Discussion

[01059] Naturally programmable biological systems offer an efficient solution for diverse organisms to achieve scalable complexity via modularity of their components. RNA-guided defense and regulatory systems, which are widespread in prokaryotes and eukaryotes, are a prominent case in point, and have served as the basis of numerous biotechnology applications thanks to the ease with which they can be engineered and reprogrammed (Hüttenhofer, A., *et al.* (2006), *Nat. Rev. Genet.* **7**, 475–482; Schneider, A. *et al.* (2020), *EMBO Rep.* **21**, e51918; Koonin, E. *et al.* (2017), *Biol. Direct.* **12**, e51918).

[01060] Here, through the exploration of Cas9 evolution, Applicants discovered the programmable RNA-guided mechanism of 3 highly abundant but previously uncharacterized

transposon-encoded nucleases: *IscB*, *IsrB*, and *TnpB*, which Applicants collectively refer to as Ω (OMEGA: Obligate Mobile Element Guided Activity) (Fig. 45) because the mobile element localization and movement likely determines the identity of their guides. Although the biological functions of Ω systems remain unknown, several hypotheses are compatible with the available evidence, including roles in facilitating *TnpA*-catalyzed, RNA-guided transposition, or acting as a toxin, with the transposon acting as the antitoxin, securing maintenance of IS200/605 insertions (Supplementary Text).

[01061] The broad distribution of the Ω systems characterized here indicates that RNA-guided mechanisms are more widespread in prokaryotes than previously suspected and suggests that RNA-guided activities are likely ancient and evolved on multiple, independent occasions, of which only the most common ones have likely been identified so far. The *TnpB* family is far more abundant and diverse than the *IscB* family; indeed, Applicants identified more than a million putative *tnpB* loci in bacterial and archaeal genomes, making it one of the most common prokaryotic genes altogether. These *TnpBs* might represent an untapped wealth of diverse RNA-guided mechanisms present not only in prokaryotes, but also in eukaryotes. Combined with our identification of a chloroplast-encoded *IscB*, these findings suggest that the expansion of RNA-guided systems into eukaryotic genomes could be a general phenomenon, and more broadly, that RNA-guided systems are functionally diverse and permeate all domains of life.

References and Notes

1. F. Zhang, Development of CRISPR-Cas systems for genome editing and beyond. *Quarterly Reviews of Biophysics*. **52** (2019).
2. F. Hille, H. Richter, S. P. Wong, M. Bratovič, S. Ressel, E. Charpentier, The biology of CRISPR-Cas: Backward and forward. *Cell*. **172**, 1239–1259 (2018).
3. V. V. Kapitonov, K. S. Makarova, E. V. Koonin, ISC, a novel group of bacterial and Archaeal DNA transposons that encode Cas9 homologs. *J. Bacteriol.* **198**, 797–807 (2015).
4. P. Siguier, E. Groubeyre, M. Chandler, Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol. Rev.* **38**, 865–891 (2014).

5. S. Shmakov, A. Smargon, D. Scott, D. Cox, N. Pyzocha, W. Yan, O. O. Abudayyeh, J. S. Gootenberg, K. S. Makarova, Y. I. Wolf, K. Severinov, F. Zhang, E. V. Koonin, Diversity and evolution of class 2 CRISPR-Cas systems. *Nat. Rev. Microbiol.* **15**, 169–182 (2017).
6. B. Zetsche, J. S. Gootenberg, O. O. Abudayyeh, I. M. Slaymaker, K. S. Makarova, P. Essletzbichler, S. E. Volz, J. Joung, J. van der Oost, A. Regev, E. V. Koonin, F. Zhang, Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell.* **163**, 759–771 (2015).
7. E. Deltcheva, K. Chylinski, C. M. Sharma, K. Gonzales, Y. Chao, Z. A. Pirzada, M. R. Eckert, J. Vogel, E. Charpentier, CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature.* **471**, 602–607 (2011).
8. I. Kalvari, E. P. Nawrocki, N. Ontiveros-Palacios, J. Argasinska, K. Lamkiewicz, M. Marz, S. Griffiths-Jones, C. Toffano-Nioche, D. Gautheret, Z. Weinberg, E. Rivas, S. R. Eddy, R. D. Finn, A. Bateman, A. I. Petrov. Rfam 14: expanded coverage of metagenomic, viral and microRNA families. *Nucleic Acids Res.* **49**, D192–D200 (2021).
9. Z. Weinberg, J. Perreault, M. M. Meyer, R. R. Breaker, Exceptional structured noncoding RNAs revealed by bacterial metagenome analysis. *Nature.* **462**, 656–659 (2009).
10. M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, E. Charpentier, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science.* **337**, 816–821 (2012).
11. G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E2579-86 (2012).
12. G. Gasiunas, J. K. Young, T. Karvelis, D. Kazlauskas, T. Urbaitis, M. Jasnauskaite, M. M. Grusyte, S. Paulraj, P.-H. Wang, Z. Hou, S. K. Dooley, M. Cigan, C. Alarcon, N. D. Chilcoat, G. Bigelyte, J. L. Curcuru, M. Mabuchi, Z. Sun, R. T. Fuchs, E. Schildkraut, P.

- R. Weigele, W. E. Jack, G. B. Robb, Č. Venclovas, V. Siksnys, A catalogue of biochemically diverse CRISPR-Cas9 orthologs. *Nat. Commun.* **11**, 5512 (2020).
13. M. Jinek, F. Jiang, D. W. Taylor, S. H. Sternberg, E. Kaya, E. Ma, C. Anders, M. Hauer, K. Zhou, S. Lin, M. Kaplan, A. T. Iavarone, E. Charpentier, E. Nogales, J. A. Doudna, Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science*. **343**, 1247997 (2014).
14. B. Q. Minh, H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, R. Lanfear, IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
15. K. S. Makarova, Y. I. Wolf, J. Iranzo, S. A. Shmakov, O. S. Alkhnbashi, S. J. J. Brouns, E. Charpentier, D. Cheng, D. H. Haft, P. Horvath, S. Moineau, F. J. M. Mojica, D. Scott, S. A. Shah, V. Siksnys, M. P. Terns, Č. Venclovas, M. F. White, A. F. Yakunin, W. Yan, F. Zhang, R. A. Garrett, R. Backofen, J. van der Oost, R. Barrangou, E. V. Koonin, Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* **18**, 67–83 (2020).
16. K. A. Majorek, S. Dunin-Horkawicz, K. Steczkiewicz, A. Muszewska, M. Nowotny, K. Ginalski, J. M. Bujnicki, The RNase H-like superfamily: new members, comparative structural analysis and evolutionary classification. *Nucleic Acids Res.* **42**, 4160–4179 (2014).
17. H. Nishimasu, F. A. Ran, P. D. Hsu, S. Konermann, S. I. Shehata, N. Dohmae, R. Ishitani, F. Zhang, O. Nureki, Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*. **156**, 935–949 (2014).
18. W. Bao, J. Jurka, Homologues of bacterial TnpB_IS605 are widespread in diverse eukaryotic transposable elements. *Mob. DNA*. **4**, 12 (2013).
19. J. V. Gomes-Filho, L. S. Zaramela, V. C. da S. Italiani, N. S. Baliga, R. Z. N. Vêncio, T.

- Koide, Sense overlapping transcripts in IS1341-type transposase genes are functional non- coding RNAs in archaea. *RNA Biol.* **12**, 490–500 (2015).
20. Z. Weinberg, C. E. Lünse, K. A. Corbino, T. D. Ames, J. W. Nelson, A. Roth, K. R. Perkins, M. E. Sherlock, R. R. Breaker, Detection of 224 candidate structured RNAs by comparative analysis of specific subsets of intergenic regions. *Nucleic Acids Res.* **45**, 10811–10823 (2017).
 21. A. Hüttenhofer, P. Schattner, The principles of guiding by RNA: chimeric RNA-protein enzymes. *Nat. Rev. Genet.* **7**, 475–482 (2006).
 22. A. Schneider, A short history of guide RNAs: The intricate path that led to the discovery of a basic biological concept. *EMBO Rep.* **21**, e51918 (2020).
 23. E. V. Koonin, Evolution of RNA- and DNA-guided antiviral defense systems in prokaryotes and eukaryotes: common ancestry vs convergence. *Biol. Direct.* **12** (2017).
 24. H. Altae-Tran, S. Kannan, F. Zhang, Code and processed data for: The widespread IS200/605 transposon family encodes diverse programmable RNA-guided endonucleases, Version 1.0, Zenodo (2021); <link to be inserted after official acceptance of manuscript>.
 25. C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, T. L. Madden, BLAST+: architecture and applications. *BMC Bioinformatics.* **10**, 421 (2009).
 26. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
 27. M. Steinegger, J. Söding, MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017).
 28. M. Steinegger, M. Meier, M. Mirdita, H. Vöhringer, S. J. Haunsberger, J. Söding, HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinformatics.* **20**, 473 (2019).
 29. S. R. Eddy, Accelerated Profile HMM Searches. *PLoS Comput. Biol.* **7**, e1002195

(2011).

30. S. A. Shmakov, K. S. Makarova, Y. I. Wolf, K. V. Severinov, E. V. Koonin, Systematic prediction of genes functionally linked to CRISPR-Cas systems by gene neighborhood analysis. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E5307–E5316 (2018).
31. A. B. Crawley, J. R. Henriksen, R. Barrangou, CRISPRdisco: An Automated Pipeline for the Discovery and Analysis of CRISPR-Cas Systems. *CRISPR J.* **1**, 171–181 (2018).
32. D. H. Haft, B. J. Loftus, D. L. Richardson, F. Yang, J. A. Eisen, I. T. Paulsen, O. White, TIGRFAMs: a protein family resource for the functional identification of proteins. *Nucleic Acids Res.* **29**, 41–43 (2001).
33. A. Delcher, Improved microbial gene identification with GLIMMER. *Nucleic Acids Research.* **27** (1999), pp. 4636–4641.
34. S. Naser-Khdour, B. Q. Minh, W. Zhang, E. A. Stone, R. Lanfear, The Prevalence and Impact of Model Violations in Phylogenetic Analysis. *Genome Biol. Evol.* **11**, 3341–3352 (2019).
35. S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. von Haeseler, L. S. Jermini, ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods.* **14**, 587–589 (2017).
36. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One.* **5**, e9490 (2010).
37. G. Altekar, S. Dwarkadas, J. P. Huelsenbeck, F. Ronquist, Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian posterior phylogenetic inference. *Bioinformatics.* **20**, 407–415 (2004).
38. A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* **30**, 1312–1313 (2014).
39. R. Lorenz, S. H. Bernhart, C. H. zu Siederdissen, H. Tafer, C. Flamm, P. F. Stadler, I. L. Hofacker, ViennaRNA Package 2.0. *Algorithms Mol. Biol.* **6**, 1–14 (2011).

40. E. Rivas, J. Clements, S. R. Eddy, A statistical test for conserved RNA structure shows lack of evidence for structure in lncRNAs. *Nat. Methods*. **14**, 45–48 (2017).
41. E. P. Nawrocki, S. R. Eddy, Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics*. **29**, 2933–2935 (2013).
42. Z. Weinberg, R. R. Breaker, R2R--software to speed the depiction of aesthetic consensus RNA secondary structures. *BMC Bioinformatics*. **12**, 3 (2011).
43. J. Mistry, S. Chuguransky, L. Williams, M. Qureshi, G. A. Salazar, E. L. L. Sonnhammer, S. C. E. Tosatto, L. Paladin, S. Raj, L. J. Richardson, R. D. Finn, A. Bateman, Pfam: The protein families database in 2021. *Nucleic Acids Res*. **49**, D412–D419 (2021).
44. C. Bland, T. L. Ramsey, F. Sabree, M. Lowe, K. Brown, N. C. Kyrpides, P. Hugenholtz, CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinformatics*. **8**, 209 (2007).
45. F. Asnicar, G. Weingart, T. L. Tickle, C. Huttenhower, N. Segata, Compact graphical representation of phylogenetic data and metadata with GraPhlAn. *PeerJ*. **3**, e1029 (2015).
46. L. Maaten, G. Hinton, Visualizing Data using t-SNE. *Journal of machine learning research*. **9**, 2579–2605 (2008).
47. P. G. Poličar, M. Stražar, B. Zupan, Embedding to reference t-SNE space addresses batch effects in single-cell classification. *bioRxiv* (2019), , doi:10.1101/671404.
48. P. J. Rousseeuw, Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math*. **20**, 53–65 (1987).
49. R. J. G. B. Campello, D. Moulavi, J. Sander, in *Advances in Knowledge Discovery and Data Mining* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2013), *Lecture notes in*

computer science, pp. 160–172.

50. D. H. Parks, M. Imelfort, C. T. Skennerton, P. Hugenholtz, G. W. Tyson, CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).
51. P.-A. Chaumeil, A. J. Mussig, P. Hugenholtz, D. H. Parks, GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* (2019), doi:10.1093/bioinformatics/btz848.
52. S. Frey, D. Görlich, A new set of highly efficient, tag-cleaving proteases for purifying recombinant proteins. *J. Chromatogr. A.* **1337**, 95–105 (2014).
53. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal.* **17**, 10–12 (2011).
54. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods.* **9**, 357–359 (2012).
55. A. M. Sriramachandran, G. Petrosino, M. Méndez-Lago, A. J. Schäfer, L. S. Batista-Nascimento, N. Zilio, H. D. Ulrich, Genome-wide Nucleotide-Resolution Mapping of DNA Replication Patterns, Single-Strand Breaks, and Lesions by GLOE-Seq. *Mol. Cell.* **78**, 975–985.e7 (2020).
56. K. Clement, H. Rees, M. C. Canver, J. M. Gehrke, R. Farouni, J. Y. Hsu, M. A. Cole, D. R. Liu, J. K. Joung, D. E. Bauer, L. Pinello, CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat. Biotechnol.* **37**, 224–226 (2019).
57. A. Bankevich, S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, V. M. Lesin, S. I. Nikolenko, S. Pham, A. D. Prjibelski, A. V. Pyshkin, A. V. Sirotkin, N. Vyahhi, G. Tesler, M. A. Alekseyev, P. A. Pevzner, SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).

58. M. Turmel, C. Otis, C. Lemieux, Divergent copies of the large inverted repeat in the chloroplast genomes of ulvophyceean green algae. *Sci. Rep.* **7**, 994 (2017).
59. D. T. Hoang, O. Chernomor, A. von Haeseler, B. Q. Minh, L. S. Vinh, UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).
60. P. J. Lockhart, A. W. Larkum, M. Steel, P. J. Waddell, D. Penny, Evolution of chlorophyll and bacteriochlorophyll: the problem of invariant sites in sequence analysis. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 1930–1934 (1996).
61. A. Criscuolo, S. Gribaldo, BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* **10**, 210 (2010).
62. J. Strecker, A. Ladha, Z. Gardner, J. L. Schmid-Burgk, K. S. Makarova, E. V. Koonin, F. Zhang, RNA-guided DNA insertion with CRISPR-associated transposases. *Science*. **365**, 48–53 (2019).
63. S. E. Klompe, P. L. H. Vo, T. S. Halpin-Healy, S. H. Sternberg, Transposon-encoded CRISPR–Cas systems direct RNA-guided DNA integration. *Nature*. **571**, 219–225 (2019).
64. M. Krupovic, K. S. Makarova, P. Forterre, D. Prangishvili, E. V. Koonin, Casposons: a new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity. *BMC Biol.* **12**, 1–12 (2014).
65. S. M. Crotty, B. Q. Minh, N. G. Bean, B. R. Holland, J. Tuke, L. S. Jermini, A. V. Haeseler, GHOST: Recovering Historical Signal from Heterotachously Evolved Sequence Alignments. *Syst. Biol.* **69**, 249–264 (2020).

Example 4

[01062] FIG. 62 shows reclustering CRISPR-associated IscB polypeptide at 60% sequence identity reveals novel IscB proteins. The IscB proteins comprise an X domain and a Y domain. CRISPR-associated IscB proteins from 00644 cluster were functional with an NAC PAM (FIG. 63A showing IscB protein of locus JGI accession Gaa0099850_1002913; FIG. 63B showing

IscB protein of JGI accession Ga0348337_018242; FIG. 63C showing IscB protein of locus JGI accession Ga0208542_1002724).

Materials and Methods

Profile curation

[01063] Initial IscB sequences were curated using NCBI's PSI-BLAST on the NR database with 8 iterations from a starting seed sequence up to a maximum of 20000 target sequences (Camacho, C. et al. (2009), *BMC Bioinformatics*. **10**, 421). Strong filtering parameters (expect threshold of 1e-5 and PSI-BLAST threshold of 1e-6) were selected to reduce the accumulation of unrelated proteins also containing HNH domains, such as restriction enzymes or homing endonucleases. All proteins smaller than 260 aa were discarded. The remaining proteins were then aligned using MAFFT FFT-NS-1 and partial proteins as well as proteins with poor alignment coverage to the HNH domain were discarded (Kotah, K., et al. (2013), *Mol. Biol. Evol.* **30**, 772–780). The filtered set was then clustered using MMSeqs2 at 70% sequence identity with a minimum coverage of 70%. The MMSeqs2 representatives for each cluster were then aligned using MAFFT-einsi (Steinegger, M. et al. (2017), *Nat. Biotechnol.* **35**, 1026–1028). The resulting alignment was further split into multiple domains (PLMP, RuvC-I, RuvC-II, HNH, and RuvC-III) in order to create distinct HHAlign profiles for the respective regions (Steinegger, M. et al. (2019), *BMC Bioinformatics*. **20**, 473). For HMMER profiles, the PLMP, HNH, and RuvC-III were used to create separate profiles. RuvC-I, BH, and RuvC-II were combined into a single HMMER profile to reduce false positives (Eddy, S., et al. (2018), *Proc. Natl. Acad. Sci. U. S. A.* **115**, E5307–E5316).

[01064] Due to the extensive diversity in TnpA, a PSI-BLAST search resulted in more than 20000 homologous sequences. Because TnpA consists of a single contiguous catalytic domain, HHblits was instead used to identify more non-redundant homologs using an E-value cutoff of 1e-3, minimum hit probability of 80% and 8 iterations on the UniRef30_2020_06 database (Steinegger, M. et al. (2019), *BMC Bioinformatics*. **20**, 473). The resulting proteins were aligned using MAFFT-einsi, and partial proteins were removed (Kotah, K., et al. (2013), *Mol. Biol. Evol.* **30**, 772–780). The resulting alignment was used to create an HMMER profile.

Identification of *IscB*, *IsrB*, and *IshA*

[01065] All prokaryotic (bacterial + archaeal) genomes were downloaded from NCBI, and NCBI WGS, as well as JGI projects for which express permission was given, and were combined into a single prokaryotic database. ORFs on all contigs were predicted with TTG, TAG, and TAA stop codons, and ATG start codons (with a minimum size of 55 aa), allowing for alternative start codons GTG (if it produced a minimum ORF size of 100 aa), TTG (if it produced a minimum ORF size of 100 aa) or CTG (if it produced a minimum ORF size of 300 aa). In the case of multiple potential start sites with different start codons for the same potential ORF, all potential start sites were listed in order of increasing corresponding ORF size. A best start site was iteratively selected by traversing the list and selecting a new best start site based on the previous one with the first item in the list selected as the initial best. GTG/TTG start sites were accepted as best if they would make the ORF 20 aa longer than the current best start site, while CTG was accepted only if it would make the protein 40 aa longer than the current best start site. Any ATG start site larger than the current best was automatically selected. ORFs sharing the same stop location and strand (+/-) as an existing protein annotation were discarded in favor of the existing annotation. All ORFs were then searched using HMMER and the 4 *IscB* profiles with a minimum bit score of 18. Any ORF/protein with a hit to any of the 4 domains was considered a protein of interest (POI) and retained for further analysis. For redundancy reduction, all proteins were then clustered using MMSeqs2 at 90% sequence identity with 85% coverage. Within each 90% cluster, proteins smaller than the 80th percentile in length were discarded. Of the remaining proteins, proteins containing X (ambiguous) amino acids were discarded unless their removal would result in an empty set of sequences. Of the remaining proteins, the sequence with the largest minimum distance from either the start or the end of the protein-coding sequence to the edge of the respective contig, that is, the protein sequence farthest away from either edge of the containing contig, was selected as the cluster representative. Representative proteins with start or end sites within 200 bp of contig edges were considered partial and were subsequently discarded. These remaining representative sequences constituted non-redundant POIs. The non-redundant POIs were then further clustered into “clusters” (as opposed to non-redundant sequences) using MMSeqs2 at 65% sequence identity and 70% sequence coverage. For each cluster, the representative sequence was taken to be the sequence with the 90th percentile length in the cluster. The use of 90th

percentile reflect the skew towards protein fragments that are often observed in large clusters. Unless specified otherwise, when referring to a cluster's components, such as the protein ORF (e.g. IscB cluster 2089) or associated RNA sequence (e.g. 2089 ncRNA), it is the respective component of the representative locus for that cluster that is being referenced.

IscB-IscB-Cas9 RuvC/BH and RuvC/BH/HNH domain based phylogenetic analysis

[01066] Cas9s identified from the IscB domain-based search exhibited monophyletic branching from a RuvC-based tree, however not all Cas9s were identified from this search. To expand the space of Cas9s included in the analysis, Cas9 profiles from the Koonin Lab (30), TIGRFAM, and profiles made from MAFFT alignments of Cas9 proteins from CRISPRDisco were used to identify additional Cas9 proteins not found in the initial IscB domain-based search using HMMER with a minimum bit score of 25 and protein length of 500 aa (Kotah, K. et al. (2013), *Mol. Biol. Evol.* **30**, 772–780; Crawley, A. et al. (2018), *CRISPR J.* **1**, 171–181; Haft, D. et al. (2001), *Nucleic Acids Res.* **29**, 41–43). Cas9 proteins from the IscB domain search were combined with the Cas9 proteins from the Cas9 search and deduplicated. The Cas9 ORF start sites were then refined using GLIMMER (Delcher, A, et al. (1999), *Nucleic Acids Res.* **29**, 41–43). The superset of Cas9 proteins were then clustered 90% sequence identity with 85% coverage and redundancy reduced in the same manner as for the IscB search. The superset of non-redundant Cas9 proteins were then reclustered at 50% sequence identity and 60% coverage. The choice of 50% sequence identity reflects the difference in the size of conserved regions between IscBs and Cas9s with approximately 100 aa more slowly evolving in the key regions (RuvC, BH, and HNH). The clustering criteria of 65% minimum sequence identity for 400 aa proteins is functionally similar to $65\% - 100/400 + 100/1000 = 50\%$ sequence identity for larger 1000 aa Cas9s with the same conserved domains spanning approximately 100 aa. For each cluster, the representative sequence was taken to be the sequence with the 90th percentile length in the cluster.

[01067] Clusters were additionally filtered for homology to IscB/Cas9 catalytic domains using profile-profile comparison. Specifically, for each cluster, the non-redundant sequences in the cluster were aligned using MAFFT in order to generate hhm inputs for HHAlign. These resulting alignments were then compared to the 6 IscB domain alignments using HHAlign (for IscB/IscB clusters), or the corresponding 8 Cas9 domain hhm profiles of the equivalent regions

(RuvC-I, Bridge Helix, REC1, REC2, RuvC-II, HNH, RuvC-III). IscB/IsrB clusters were selected for further analysis if at least two hits to any of the IscB RuvC-I, RuvC-II, or RuvC-III domains were identified through HHAlign (with a minimum bit score of 17). Cas9 clusters were selected for further analysis if at least two hits to any of the IscB RuvC-I, RuvC-II, or RuvC-III domains were identified through HHAlign (with a minimum bit score of 17), or if any two hits to the Cas9 RuvC-I, RuvC-II, RuvC-III domains were identified through HHAlign (with a minimum bit score of 17). The passing Cas9 clusters were then pooled with the passing IscB clusters for phylogenetic analysis.

[01068] IscB, IsrB and Cas9 obtained from the above filtering criteria were aligned using MAFFT-x2 (two iterations) and BLOSUM62 scoring (default if not otherwise specified). The RuvC-I and BH regions of Cas9 did not align with the RuvC-I and BH regions of IscB and IsrB due to the different domain architectures and sizes of these divergent protein families. RuvC-I and BH were manually grouped and realigned using MAFFT-x2. Sequences with poor coverage at the RuvC-I domain region of the alignment were removed if their HHAlign score to either the IscB-RuvC-I or Cas9-RuvC-I profiles were less than 21. Small proteins with both Cas9-like and IscB-like sequences containing HHAlign hits to RuvC-I and BH typically did not have correct alignments with RuvC-I and BH relative to all other proteins due to their hybrid nature. For such proteins, all amino acids between the N-terminus and RuvC-II were grouped together as RuvC-I and BH for alignment. RuvC-I and BH regions were then realigned using MAFFT-linsi. IsrB aligned columns from the HNH domain were moved to the RuvC-II column group. RuvC-II, RuvC-III, HNH, and PLMP domains were then sequentially aligned using MAFFT-linsi. Excess regions between the BH domain and RuvC-II were then aligned using MAFFT-einsi and BLOSUM30 for identification of REC-like insertions. Sequences with no or poor alignment to any of the RuvC or BH domains were removed. The resulting alignment was then used for phylogenetic analysis.

[01069] All domains not common to all 3 types of proteins, namely PLMP, REC1, REC2, PI domains, and IscB/IsrB C-terminal domains were removed from the alignment, leaving a trimmed alignment containing only the highly conserved portions of the RuvC-I, BH, RuvC-II, and RuvC-III domains, creating a RuvC/BH alignment containing IscB, IsrB, and Cas9. Another alignment was created containing only RuvC-I, BH, RuvC-II, HNH, and RuvC-III domains for only IscB and Cas9, termed the RuvC/BH/HNH alignment. For both of these alignments, clusters with dead representative sequences (sequences with key catalytic sites

mutated) were removed. Specifically, the positions filtered were the RuvC-I conserved D, RuvC-II conserved E, HNH Conserved H (when applicable), and RuvC-III conserved D and H. Positions (columns) with more than 75% gaps were also removed. Symmetry tests implemented in IQ-Tree 2 were used to identify potential phylogenetic violations for the alignments (Fig. 22) (Naser-Khdour, S. et al. (2019), *Genome Biol. Evol.* **11**, 3341–3352). The RuvC/BH/HNH alignment displayed significant violation to the stationarity assumption of the three main assumptions used in typical phylogenetic analysis (reversibility, stationarity, homogeneity). As the alignment contained too many taxa for the use of heterotachy models in IQ-Tree 2, Applicants use a subtractive approach to identify the source of the stationarity violation. Preliminary analyses showed that the major clade of II-B Cas9s consistently split from the rest of Cas9s with a branch length of ≥ 1 , suggesting its accurate placement along the tree might be difficult. Applicants removed the major clade of II-B Cas9s from the RuvC/BH/HNH alignment, which substantially reduced the stationarity violation as determined by marginal symmetry test p-value (Fig. 22). Applicants also created another alignment only consisting of IscB and Cas9s from the early stages of Cas9 evolution, which completely eliminated any stationarity violation. For both of these additional alignments, catalytically inactive variants were removed as described for the RuvC/BH and RuvC/BH/HNH alignments, and positions (columns) with more than 75% gaps were removed. For each alignment, substitution model selection was performed using the ModelFinder tool implemented in IQ-Tree 2 (Kalyaanamoorthy, S. et al. (2017), *Nat. Methods.* **14**, 587–589). Optimal models were selected using the Akaike Information Criterion corrected for small sample sizes (AICc). In most cases, the AICc best model differed from the Bayesian Information Criterion (BIC) or AIC (standard Akaike Information Criterion), and some analyses were run for both sets of models; however, AICc was generally preferred due the small sample size correction. Then, for each alignment, phylogenetic trees were built using multiple methods for cross comparison (IQ-Tree 2, RAxML, MrBayes). While FastTree2 was used for quick visualizations of phylogenetic information, the likelihood scores obtained using this method were substantially worse than those for IQ-Tree 2, RAxML, or MrBayes (Mihn, B., et al. (2020), *Mol. Biol. Evol.* **37**, 1530–1534; Price, M., et al. (2010), *PLoS One.* **5**, e9490; Altekar, G. et al. (2004), *Bioinformatics.* **20**, 407– 415; Stamatakis, A. et al. (2014), *Bioinformatics.* **30**, 1312–1313). As a result, FastTree2 was not used for comprehensive cross comparisons.

[01070] For Figure 31A, a hybrid tree approach was used in order to condense the information pertinent to Cas9 while maintaining phylogenetic accuracy. For this approach, a subsample of Cas9 clusters were selected from the alignment, in addition to the complete set of IscB and IsrBs. The resulting sub-alignment was then used for phylogenetic inference using IQ-Tree 2 with identical parameters. Due to the potentially skewed Cas9 related information present in the sub- alignment, placement of the Cas9 lineage was inferred from the tree built with the original alignment. This was done by detaching the Cas9 branch from the sub alignment tree and replacing the original Cas9 branch on the original tree with the smaller detached branch. The branch selected for grafting was the branching between Cas9_849 and all other Cas9s, as this region shared the same topology between both trees. The ordering of Cas9 subtype evolution was checked for consistency between each tree to ensure compatibility after having substantially downsampled Cas9 proteins.

IscB/IsrB ω RNA discovery, curation, and analysis

[01071] Using results from just the IscB domain search, all representatives from clusters with at least 3 proteins and at least 2 HHAlign hits to any of RuvC-I, RuvC-II, or RuvC-III with at least 17 bit score were collected. All regions upstream (-300 bp to +200 bp of the start codon), and downstream (-200 bp to +300 bp of the stop codon) for all IscB and IsrB proteins were aligned separately using MAFFT-einsi. The upstream alignment demonstrated large conserved regions outside of the typical CDS boundaries of the IscB/IsrB. The downstream alignment did not contain any large conserved region and was discarded for further analysis. Individual sequences in the conserved upstream region were folded using ViennaRNA RNAFold (Lorenz, R. et al. (2011), *Algorithms Mol. Biol.* **6**, 1–14). Sequences in the alignment were split into separate groups on the basis of conservation to key distinct regions in the alignment. The main group was labeled G1a and spanned a large number of IscB ω RNAs. R-scape was used to infer covariance-folded RNA structures that correct for phylogenetic correlations (Rivas, E. et al. (2017), *Nat. Methods.* **14**, 45–48). The R-scape parameters used for all profiles was an E-value threshold of 1e-2 and a gap threshold of 0.75. CMbuild from Infernal was used to optimize the RNA alignments and build covariance models (CMs) (Nawrocki, E. et al. (2013), *Bioinformatics.* **29**, 2933–2935). R2R was used to visualize the resulting RNA structures (Weinberg, Z. et al. (2011), *BMC Bioinformatics.* **12**, 3). Additional alignment groups for IscB/IsrB ω RNAs (G1b-i) were iteratively created based on conserved upstream regions (relative to the ORF) from clades of IscB/IsrB in the RuvC tree that were not

strongly associated with existing ω RNA groups. Models for these groups were built in the same manner, except the consensus secondary structure identified using ViennaRNA was used in place of the R-scape structure when the sample size was too small to permit an accurate covariance-folded structure. Due to this small sample size, in Fig. 1G, for the CRISPR-associated IscB ncRNA, nucleotides with $\geq 97\%$ identity and nucleotides with $\geq 90\%$ identity are the same by definition and were displayed as being $\geq 90\%$ to avoid overstating their level of conservation in the presence of limited data. Pseudoknots were not explicitly encoded in the covariance models for Infernal. The structure for the hybrid CRISPR/ ω RNA associated with CRISPR-associated IscBs was inferred in the same way, using the consensus secondary structure in lieu of the covariance-folded structure due to small sample size.

[01072] All IscB/IsrB ω RNA CM profiles were then searched against the prokaryotic database using cmsearch with a minimum bit score of 20 to identify potential ω RNAs (Nawrocki, E. et al. (2013), *Bioinformatics*. **29**, 2933–2935).

[01073]). Overlapping profile hits were resolved by assigning the group of overlapping hits to the profile hit that produced the highest bit score. ω RNAs were annotated using the natural target search analysis

Natural Target Search

[01074] For the target search, a separate database was curated consisting of plasmids and bacteriophage/archaeal viruses from the NCBI microbial genomes portal. The 30 bp upstream from each ω RNA scaffold as determined by CMsearch along with the generated profiles was used as the corresponding guide sequence for that ω RNA. Guides were then searched against the prokaryotic database or the plasmid/viral database for potential spacer matches using a minimum bit score cutoff of 44. To remove candidate target hits that are related to the original guide/ ω RNA pair, a separate pruning step was performed. For each target hit, the corresponding ω RNA scaffold of the guide that generated the hit was separately blasted against a 300 bp window around the target. For this second search, if the ω RNA scaffold had a match within the window with a bit score of 44 or higher and was found within 15 bp of the identified target, the target was discarded for being too similar to the original locus.

[01075] Target hits were then searched for evidence of transposition. For each guide/target pair, the sequence consisting of up to 2000 bp downstream from the end of the ω RNA scaffold

was used as a 3' prime seed. This seed region was then searched against the target contig using blastn with a minimum bit score of 44. Hits from this search were discarded if the strand was not the same as the guide/target match strand. Moreover, hits that were not both ≤ 50 bp downstream from the guide/target match site and ≤ 15 bp upstream from the guide/target match site were discarded for being implausible insertion sites. If any 3' seed hits remained at this stage, the guide/target pair was considered to be associated with transposon insertion/transposition. In this case, of the remaining 3' seed hits, the most upstream hit was selected, which along with the guide/target site demarcated the 3' insertion site flanking region and 5' insertion site flanking region respectively. Guide/target pairs that did not have an identifiable 3' insertion site flanking region were not considered insertion linked and were discarded unless their guide/target match bit score was ≥ 50 , which approximately corresponded to an E-value of 0.02 for the given size of the prokaryotic database.

[01076] To determine if prophage sequences might be a target for ω RNAs, all guide/target pairs were analyzed to detect if the guide targeted ORFs. Specifically, a guide was considered ORF targeting if any corresponding target was found to overlap with an ORF in the given database. All targeted ORFs were then searched for possible PFAM hits using HMMER and the Pfam34 database (Mistry, S. et al. (2021), *Nucleic Acids Res.* **49**, D412–D419) and a minimum bit score of 18. All ORFs that had any hit to any phage Pfam profiles listed in Table 12 were considered phage targets, and the corresponding guide was labeled as phage ORF targeting.

[01077] ω RNAs were then further annotated as being near IscB/IsrB fragments based on the presence of any nearby ORF (≤ 2000 bp) with an HMMER hit of 18 bits or higher to any of the 4 IscB HMM profiles. ω RNAs were then clustered as follows. ω RNA scaffold sequences were clustered at 95% sequence identity using cd-hit-est using only global +/- strand comparisons.

[01078] Similarly, all spacers were clustered at 95% sequence identity using cd-hit-est using only global +/- strand comparisons. ω RNAs were then grouped according to both their ω RNA scaffold cluster id (ω RNA_id) and their spacer cluster id (spacer_id). Each (ω RNA_id, spacer_id) group was then classified as being associated near IscB if any ω RNA in the group was considered near IscB/IsrB fragments. Then, for each ω RNA_id, the fraction of all (ω RNA_id, spacer_id) pairs considered associated with IscB/IsrBs was computed. If this fraction was ≤ 0.75 , the ω RNA_id was classified as being associated with standalone ω RNAs,

and this classification was propagated back to all IscB/IsrB ω RNAs according to their ω RNA_id.

Anti-Repeat Determination

[01079] For CRISPR associated IscB, IsrBs, and Cas9s, anti-repeats corresponding to likely tracrRNA regions were identified as follows. For each locus, the closest predicted CRISPR array to the locus within 10 kbp of the protein of interest (POI) (namely IscB, IsrB, or Cas9) was selected. If no such array was identified, the analysis was not performed for the given locus. Next, edit distance between each direct repeat in the selected CRISPR array was computed. The resulting edit distance matrix was used to a medoid direct repeat that had the lowest average edit distance to all other direct repeats in the CRISPR array. The medoid direct repeat was then searched against the 10 kbp window around the POI using blastn with a word size of 4. Hits with bit score < 20 were discarded. Hits that mapped to any predicted CRISPR array were used for calibration. Specifically, all such CRISPR array matching hits were collected into a list, and the 20th quantile bit score of such hits was denoted the min_crispr_score. Any hit that was found ≤ 40 bp of any CRISPR array was discarded as a potential anti-repeat due to its high proximity to CRISPR arrays suggesting the hit could simply be another CRISPR direct repeat. Hits occurring inside the protein of interest were also discarded. Remaining hits occurring ≤ 500 bp from either the POI or any predicted CRISPR array were retained. For each hit in this final list, if the region of the query median direct repeat that produced the hit does not reach ≤ 3 bp of either edge of the direct repeat, the hit was discarded as a potential anti-repeat due to the likely inability to produce a stable lower stem region for the putative repeat-anti-repeat stem-loop. If any hits remained, these hits were considered anti-repeats for the POI.

[01080] Anti-repeats were then used to predict potential tracrRNAs for specific Cas9s. This process involved taking the anti-repeat region and extending it until the boundary of the neighboring feature (either protein gene or CRISPR array). TracrRNAs were predicted in this manner for the tracrRNAs shown in Figure 31B.

Cluster annotation

[01081] All 10 kb genomic frames (10 kb region around the protein-coding gene of interest (POI)) were collected, and encompassed all unique/non-redundant *IscB*, *IsrB*, and *Cas9* genes (according to their 90% redundancy reducing clustering id, denoted *mc_id*). CRISPR arrays were identified using CRT (Bland, C. et al. (2007) *BMC Bioinformatics*. **8**, 209). *TnpA* was predicted for all ORFs within 10 kb of a POI using HMMER with a minimum bit-score cutoff of 18.0. All ω RNA annotations from the natural target search were merged with the corresponding genomic frames when applicable. RNA profile hits with a bit-score below 35 were discarded. For each genomic frame, all ω RNAs predictions were merged with the POIs according to their genomic locations. All ω RNAs further than 250 bp away from either end of the POI were considered to be unrelated to the POI. Any ω RNAs closer than 250 bp were considered related to the POI, and in the case of multiple such ω RNAs, only the closest one to the POI was considered for tabulation purposes. For each POI, a locus boundary was then computed as follows. A minimal locus interval [L, R] was set to encompass the POI and associated ω RNA when applicable. Then, iteratively, all other ORFs and CRISPR arrays in the contig were considered for addition into the locus interval, upon which a new minimal interval [L', R'] would be computed to encompass all elements in the locus. Specifically, an ORF was added to the locus if it was ≤ 1000 bp from the existing interval and contained a *Cas1*, *Cas2*, *Cas4*, or *Csn2* HMM profile hit with a minimum bit score of 25 as determined by *hmmsearch*. CRISPR arrays were included if they were ≤ 500 bp from the existing interval and contained ≥ 3 direct repeats and had a median direct repeat length of ≥ 25 bp. The locus determination process was terminated once no new elements were being added, resulting in a final locus interval [L, R], which Applicants termed the locus bounds. ORFs with HMMER hits to *TnpA* with a minimum bit score of 14 was then considered *tnpA* and a part of the locus if they were ≤ 500 bp away from the locus bounds ([L, R]).

[01082] For tabulation of cluster annotations, all POIs and corresponding locus information were grouped by cluster id (*c_id*). POIs with 'X' amino acids in the sequence translation were discarded. Moreover, POIs with ≤ 250 bp distance to either edge of the contig were also discarded. The resulting locus features (features within the locus bounds [L, R]), and *tnpA* association were then used to tabulate aggregate metrics, such as average CRISPR association (fraction of loci in the cluster that have a CRISPR array within the locus bounds), *tnpA* association, and various *cas* gene associations. Anti-repeat association rates were also

calculated in this way using anti-repeats as determined above, where loci that did not contain a CRISPR directly inside the locus bounds ([L, R]) were considered to not have an anti-repeat. For natural target search tabulations, the associated ω RNA in the locus bounds for POI closest to the POI was used. The fraction of associated ω RNAs in the cluster targeting genomic regions, plasmids, and phage were tabulated. Similarly, the fraction of associated ω RNAs associated with standalone ω RNAs for the cluster was also calculated. Lastly, the fraction of associated ω RNAs in the cluster that demonstrated evidence of insertion/transposition was also computed for each cluster.

[01083] For tabulation of protein domains, the hhm profile for the alignment of non-redundant sequences from the cluster was used to identify domain hits. Specifically, the cluster was considered to have the PLMP domain if the cluster alignment had an HHAlign match to the IscB PLMP domain with a bit score ≥ 30 . Likewise, the cluster was considered to have the HNH domain if the cluster alignment had an HHAlign match to the IscB HNH domain with a bit score ≥ 25 . As the HNH has multiple variants of the motif (often being replaced by HNN as in most Cas9s), the identity of the last amino acid in the HNH motif (e.g. HNH vs HNN) was computed for the cluster according to the corresponding position in alignment of representative sequences used for the phylogenetic analysis.

[01084] The protein length for each cluster was taken to be the non-gapped length of the cluster representative in the alignment containing all IscB, IsrB, and Cas9 cluster representative sequences (IscB_IsrB_Cas9_full_gappy.fasta). The calculated protein length therefore reflected the trimming of large (> 50 aa) non-homologous sequences at the protein termini as determined from the IscB_IsrB_Cas9_full_gappy.fasta file. REC-like regions were taken to be the region directly between the Bridge Helix and the RuvC-II domain as found in the alignment of cluster representatives. The REC-like regions are found in IscB_IsrB_Cas9_REC_only.fasta. Lengths of REC-like insertions for each cluster were taken to be the non-gapped lengths from the REC alignment. GraPhlAn was used to visualize the resulting information on different phylogenetic trees (Asnicar, F. et al. (2015), *PeerJ.* **3**, e1029).

[01085] Throughout the text, where fraction of loci containing a specific feature are computed, the analysis considered only the non-inactivated IsrB loci, as determined by RuvC/BH alignment, and the non-activated IscB and Cas9 loci as determined by the RuvC/BH/HNH. In addition, proteins from cluster 34507 were also included despite being

considered an inactivated cluster because the cluster contains active CRISPR-associated *IscB* variants.

IscB/IsrB ω RNA phylogenetic analysis

[01086] All *iscB/ISR* associated ω RNAs from the loci of cluster representatives were collected.

For CRISPR-associated ω RNAs, the entire region spanning the CRISPR array and the remainder of the ω RNA scaffold was taken to be the corresponding RNA for the system. These RNAs were then converted to DNA sequences and aligned using MAFFT-ginsi to identify conserved regions spanning multiple types of RNAs. A phylogenetic tree was then constructed using IQ-Tree 2 with 5000 rapid bootstraps with hill climbing, a GTR substitution model with 4 Gamma rate categories.

[01087] To analyze the relationship between II-D tracrRNAs and ω RNAs, cmsearch was used to identify if the predicted II-D tracrRNAs contained any homology to *IscB/IsrB* ω RNAs using the corresponding *IscB/IsrB* ω RNA profiles. Based on the potential homology determined this way, a small panel of ω RNAs that were highly related to the II-D tracrRNAs were combined with synthetic single guide sequences from II-D systems constructed by concatenating the DR with 'AAAA' and the predicted tracrRNA. These combined sequences were then aligned using MAFFT. The resulting alignment was used for Bayesian phylogenetic analysis with MrBayes with 2 chains at a delta temperature of 0.025 with 8 independent runs for 5M generations. A standard GTR model with gamma rates and 4 categories was used.

Binning of Yellowstone Lake metagenomes

[01088] Tetranucleotide frequencies of assembled metagenomes (contigs > 5 kb) provided an input for the t-stochastic neighbor embedding (t-SNE) analysis, which reduced the data to two dimensions (Maaten, L. et al. (2008), *Journal of Machine Learning*, **9**, 2579-2605). Applicants used openTSNE (Poličar, P. et al. (2019), *bioRxiv* (2019), , doi:10.1101/671404) because of its speed and scalability. The perplexity parameter was set at 40, while the learning rate and the number of iterations were variable. Typically, both were set to larger values for bigger datasets. A grid search coupled with silhouette scores (Rousseeuw, P. et al. (1987), *J. Comput. Appl. Math.* **20**, 53–65) was used to find optimal parameters for density-based

clustering (Campello, R. et al. (2013), *Advances in Knowledge Discovery and Data Mining, Lecture Notes in Computer Science*, pp. 160-172) (HDBSCAN) of data points from the t-SNE plot (MD and WPI, in preparation). DNA sequences were grouped into metagenome-assembled genomes (MAGs), and their completeness was evaluated using CheckM (Parks, D. et al. (2015), *Genome Res.* **25**, 1043–1055). The GTDB Toolkit (Chaumeil, P. et al. (2019), *Bioinformatics* doi:10.1093/bioinformatics/btz848) was used to classify MAGs into taxonomic groups based on the Genome Taxonomy Database.

Identification of guide encoding mechanisms for *IscB/IsrB* ω RNA

[01089] Upon complete classification and curation of all major ω RNA types according to *IscB* and *IsrB*, the genome of *K. racemifer* was searched for all instances of *IscB* and *IsrB* using HMMER with the previously described profiles. The genome was also searched for all instances of *IscB/IsrB* ω RNAs using CMsearch with the G1a-G1i RNA covariance models. Occurrences of multiple nearly identical *IscBs* associated with nearly identical ω RNAs were identified with BLASTn and was classified as transposon expansion. Occurrences of ω RNAs without detectable *IscBs* or *IsrBs* within 500 bp on the same strand were classified as standalone *trans*-acting ω RNAs. In some instances, ω RNAs and corresponding *IscB/IsrBs* were separated by the insertion of unrelated transposons between them. In such cases, the ω RNA was not considered a *trans*-acting ω RNA.

[01090] All covariance models were searched against our prokaryotic genome database. Examples with multiple ω RNAs on the same strand within 300 bp were retained for further analysis and categorized as ω RNA arrays.

Identification of eukaryotic *IscB* orthologs.

[01091] All eukaryotic genomes were downloaded from NCBI. In order to capture all possible *IscBs*, existing gene models were discarded for this analysis. All DNA sequences were translated into 6 frame amino acids translations and split into ORFs by splitting across stop codons (*).

[01092] Each ORF was then searched for *IscB* domains using the HMMER profiles generated by the *IscB* profile curation step. ORFs with hits to both *IscB* HNH and RuvC domains were retained for further analysis. The regions around the ORFs were then searched

for IscB-linked ω RNAs using CMsearch and the IscB-linked ω RNA covariance models created in this study.

Codon analysis of *I. tetrasporus* chloroplast IscB

[01093] All previously annotated ORFs from the NCBI chloroplast genome KY407659 were obtained and classified as non-IscB or IscB using a BLASTp search with the single intact IscB as the query. ORFs corresponding to IscB fragments were discarded. For each remaining IscB or non-IscB ORF, the average codon usage for each of the 60 codons was computed as a vector, p_i , where p_{ic} is the fraction of codons in ORF i that use codon c . The average and standard deviations of p across i (excluding IscB ORFs) were computed to get distribution information for how the codon usage varies across different non-IscB ORFs. The codon usage of the single IscB ORF for each codon was then compared to the average and standard deviation of codon usage across all the other non-IscB ORFs. To quantitatively assess how different the IscB codon usage is relative to other ORFs in the same genome, p was averaged across the non-IscB ORFs to obtain an average distribution of codon usage, which Applicants referred to as the reference distribution. For each ORF, the Kullback-Leibler divergence between the ORF's codon usage, p_i , and the reference codon usage distribution was computed.

Discovery of IshB

[01094] The powerset of all possible domain combinations of the PLMP, RuvC-I, BH, RuvC-II, HNH, and RuvC-III domains was generated. For each domain combination, the number of clusters from the IscB domain search with a hit to every domain in the combination with a minimum bit score of 21 was computed. Domain combinations that exhibited high levels of protein sequence conservation within the combination were retained for further analysis. Domain combinations that were N-terminal or C-terminal truncations of IscB, Cas9, or IsrB were discarded. From the remaining combinations, PLMP + HNH displayed high cluster counts relative to the other domain combinations, such as RuvC-II + PLMP only, suggesting that the combination corresponded to a true protein family. These proteins were subsequently named IshB due to the presence of an HNH domain, while also containing the PLMP domain present in IscB and IsrB.

Taxonomy analysis of *IscB*, *IsrB*, and *Cas9*

[01095] Taxonomy information for available genomes were obtained from the NCBI microbial genomes portal. Genomes which lacked taxonomic information or metagenomes were discarded for this analysis. Genes were classified into *IsrB*, *IscB*, *Cas9*, or other on the basis of their cluster id. The fraction of genomes in each taxonomic group containing *IsrB*, *IscB*, or *Cas9* was computed. For each gene type (*IscB*, *IsrB*, and *Cas9*) the distribution of the number of genes per genome in each taxonomic group was calculated. Box plots of gene counts per taxonomic group was then computed using Python.

TnpB curation and analysis of ω RNA

[01096] Examples of *IscB*-linked ω RNA from *K. racemifer* were searched in the *K. racemifer* genome using BLASTn. Hits in the vicinity of *IscB* or *IsrB* were discarded. Multiple partial hits were found in the vicinity of *TnpB*, with the hit being always downstream of the *TnpB* gene.

[01097] Exploration of these hits showed that multiple *TnpBs* shared transposon ends with *IscB*. An upstream and downstream locus conservation analysis was performed for related *TnpB* loci as done for *IscB*. As *TnpB* is highly diverse, only *TnpBs* that were identifiable via high similarity in an mmseqs2 search were included. The RNA-seq traces were then used to identify the 5' RNA boundary on the *TnpB* ω RNAs in the conservation analysis. These ω RNAs from the conservation analysis were then extracted and used to create an R-scape secondary structure with an E-value threshold of 1e-2 and a gap threshold of 0.75.

Expression and purification of CRISPR-associated *IscB* RNPs

[01098] To purify the CRISPR-associated *IscB* in complex with the ncRNA of its native locus, the human codon optimized CRISPR-associated *IscB* protein was cloned into a modified pET45b(+) backbone with an N-terminal His14-MBP-bdSUMO tag (Frey, S. et al. (2014), *J. Chromatogr. A.* **1337**, 95–105), and C-terminal twin- strep tag. The non-coding part of the CRISPR-associated *IscB* locus was cloned into a separate pCOLADuet-1 vector, and co-expressed with the CRISPR-associated *IscB* protein in Rosetta(DE3)pLysS strain (EMD Millipore) (Tables 5, 9). Cells were grown at 37°C in the terrific broth (TB) medium supplemented with 100 µg/ml ampicillin, 25 µg/ml kanamycin, and 34 µg/ml chloramphenicol

until reaching an optical density (OD₆₀₀) of 0.3, then shifted to 18°C, and grown further until OD₆₀₀ was 0.7. The bacterial culture was induced overnight at 18°C with 0.2 mM isopropyl β-D-1thiogalactopyranoside (IPTG) and harvested by centrifugation. The cell paste was resuspended in lysis buffer (50 mM Tris pH 8, 200 mM NaCl, 5% glycerol, 5 mM MgCl₂, and 1 mM DTT) supplemented with protease inhibitors (PMSF and Roche cOmplete, EDTA-free), and lysed by two passes with a high-pressure homogenizer (LM20 Microfluidizer, Microfluidics). The lysate was cleared by centrifugation, and the soluble fraction was gently mixed with Strep-Tactin Superflow Plus resin (Qiagen) at 4°C. The resin was first washed with lysis buffer, then briefly with buffer A (50 mM Tris pH 8, 1 M NaCl, 5% glycerol, 5 mM MgCl₂, and 1 mM DTT), and buffer B (50 mM Tris pH 8, 500 mM NaCl, 5% glycerol, 5 mM MgCl₂, and 1 mM DTT) on a gravity flow column. Bound RNP was then eluted in buffer B supplemented with 5 mM desthiobiotin (Sigma), and the N-terminal solubility tag was cleaved using bdSENP1 protease.

[01099] Once the boundaries of the co-purified RNA were defined by small RNA sequencing, the predicted ncRNA sequence was cloned downstream of a T7 promoter in a pCOLADuet-1 vector for inducible expression. The CRISPR-associated IscB-ncRNA complexes used for *in vitro* cleavage assays were then prepared following the same procedure explained above.

Small RNA sequencing: Heterologous expression in E. coli

[01100] Stbl3 chemically competent *E. coli* were transformed with plasmids containing the locus of interest. A single colony was used to seed a 5 mL overnight culture. Following overnight growth, cultures were spun down, resuspended in 750 μL TRI reagent (Zymo) and incubated for 5 min at room temperature. 0.5 mm zirconia/silica beads (BioSpec Products) were added and the culture was vortexed for approximately 1 minute to mechanically lyse cells. 200 μL chloroform (Sigma Aldrich) was then added, culture was inverted gently to mix and incubated at room temperature for 3 min, followed by spinning at 12000xg at 4°C for 15 min. The aqueous phase was used as input for RNA extraction using a Direct-zol RNA miniprep plus kit (Zymo). Extracted RNA was treated with 10 units of DNase I (NEB) for 30 min at 37°C to remove residual DNA and purified again with an RNA Clean & Concentrator-25 kit (Zymo). Ribosomal RNA was removed using a RiboMinus Transcriptome Isolation Kit for

bacteria (Thermo Fisher Scientific) as per the manufacturer's protocol using half-volume reactions. The purified sample was then treated with 20 units of T4 polynucleotide kinase (NEB) for 6 h at 37°C and purified again with an RNA Clean & Concentrator-25 (Zymo) kit. The purified RNA was treated with 20 units of 5' RNA polyphosphatase (Lucigen) for 30 min at 37°C and purified again using an RNA Clean & Concentrator-5 kit (Zymo). Purified RNA was used as input to an NEBNext Small RNA Library Prep for Illumina (NEB) as per the manufacturer's protocol with an extension time of 60 s and 16 cycles in the final PCR. Amplified libraries were gel extracted, quantified by qPCR using a KAPA Library Quantification Kit for Illumina (Roche) on a StepOne Plus machine (Applied Biosystems/Thermo Fisher Scientific) and sequenced on an Illumina NextSeq with Read 1 42 cycles, Read 2 42 cycles and Index 1 6 cycles. Adapters were trimmed using CutAdapt (Martin, M. et al. (2011), *EMBnet.journal*. **17**, 10–12) and mapped to loci of interest using Bowtie2 (Langmead, B. et al. (2012), *Nat. Methods*. **9**, 357–359). Filled reads were obtained and filled reads longer than 200 bp were visualized using a custom Python script.

Ribonucleoprotein

[01101] RNPs were purified as described. 100 µL concentrated RNP was used as input. The above protocol was followed with the following modifications: 300 µL TRI reagent (Zymo) and 60 µL chloroform (Sigma Aldrich) were used for RNA extraction.

K. racemifer

[01102] Freeze-dried *K. racemifer* SOSP1-21 DSM 44963 was obtained from DSMZ (www.dsmz.de/collection/catalogue/details/culture/DSM-44963), resuspended in GYM Streptomyces media (4 g glucose, 10 g malt extract, 4 g yeast extract in 1 L water) and grown in a shaking incubator at 28 C. After 76 days, the culture was spun down, and the above protocol was followed with the following modifications: mechanical lysis using zirconia beads was performed with approximately 30 min of vortexing. Ribosomal RNA was removed using a NEBNext rRNA Depletion Kit (Bacteria) (NEB) as per the manufacturer's protocol and the rRNA-depleted sample was purified using Agencourt RNAClean XP beads (Beckman Coulter) prior to T4 PNK treatment. T4 PNK treatment was performed for 1.5 h and purified with an RNA Clean & Concentrator-5 kit (Zymo). Final PCR in the small RNA library prep contained 15 cycles.

I. tetrasporus

[01103] An agar slant of *I. tetrasporus* UTEX B 2012 was obtained from UTEX, inoculated in Modified Bold 3N Medium (UTEX) and grown at 20°C in a shaking incubator with 12-hour light/dark cycles. After 18 days, the culture was spun down and the above protocol was followed with the following modifications: ribosomal RNA was removed using the manufacturer's protocol from the RiboMinus Transcriptome Isolation Kit for bacteria (Thermo Fisher Scientific) with probes from the RiboMinus Transcriptome Isolation Kit for bacteria, RiboMinus Plant Kit for RNA-seq and RiboMinus Eukaryote Kit for RNA-seq (all Thermo Fisher Scientific) combined in an equimolar ratio. Eluted RNA from ribosomal removal was concentrated using an RNA Clean & Concentrator-25 kit (Zymo) prior to T4 PNK treatment. Final PCR of the small RNA library prep contained 15 cycles.

Cloning PAM/TAM libraries

Target sequences with 8N degenerate flanking sequences were synthesized by IDT and amplified by PCR with NEBNext High Fidelity 2X Master Mix (NEB). Backbone plasmid was digested with restriction enzymes (pACYC: EcoRV; pUC19: Eco88I and HindIII, Thermo Fisher Scientific) and treated with FastAP alkaline phosphatase (Thermo Fisher Scientific). The amplified library fragment was inserted into the backbone plasmid by Gibson assembly at 50°C for 1 hour using 2X Gibson Assembly Master Mix (NEB) with an 8:1 molar ratio of insert:vector. The Gibson assembly reaction was then isopropanol precipitated by addition of an equal volume of isopropanol (Sigma Aldrich), final concentration of 50 mM NaCl, and 1 µL of GlycoBlue nucleic acid co-precipitant (Thermo Fisher Scientific). After a 15 min incubation at room temperature, the solution was spun down at max speed at 4°C for 15 min, then the supernatant was pipetted off and the pelleted DNA was resuspended in 12 µL TE and incubated at 50°C for 10 minutes to dissolve. 2 µL were then transformed by electroporation into Endura Electrocompetent *E. coli* (Lucigen) as per the manufacturer's instructions, recovered by shaking at 37°C for 1 h, then plated across 5 22.7cm x 22.7cm BioAssay plates with the appropriate antibiotic resistance. After 12-16 hours of growth at 37 C, cells were scraped from the plates and midi- or maxi-prepped using a NucleoBond Midi- or Maxi-prep kit (Machery Nagel).

E. coli PAM screen

[01104] 100 ng each of a plasmid containing a locus of interest and a target 8N degenerate flanking library plasmid were transformed by electroporation into 30 μ L Endura Electrocompetent *E. coli* (Lucigen) as per the manufacturer's protocol, with 3 biological replicates per locus of interest as well as 3 biological replicates of an empty control. After recovery by shaking at 37°C for 1 hour, cells were plated across 1 22.7cm x 22.7 cm BioAssay plate with the appropriate antibiotic resistance and grown for 12-16 h at 37°C. Cells were scraped from the plate and mixed well, and 2 mL of the scraped cells were used as input to minipreps (Qiagen). 100 ng miniprepplasmids were input to PCR to amplify the PAM-containing region with a 12-cycle PCR using NEBNext High Fidelity 2X PCR Master Mix (NEB) with an annealing temperature of 63°C, followed by a second 18-cycle round of PCR to further add Illumina adaptors and barcodes. Amplified libraries were gel extracted, quantified by a Qubit dsDNA HS assay (Thermo Fisher Scientific) and subject to single-end sequencing on an Illumina NextSeq with Read 1 75 cycles, Index 1 8 cycles and Index 2 8 cycles. PAMs were extracted and Weblogos depicting depleted PAMs were visualized using a custom Python script (Altae-Tran, H. et al. (2021)).

In vitro cleavage assays

[01105] Double-stranded DNA (dsDNA) substrates were produced by PCR amplification of pUC19 plasmids containing the target sites and the TAM sequences. Cy3 and Cy5-conjugated DNA oligonucleotides (IDT) were used as primers to generate the labeled dsDNA substrates. Single-stranded DNA (ssDNA) substrates were ordered as Cy5.5-conjugated oligonucleotides (IDT). All ω RNA used in the biochemical assays was *in vitro* transcribed using the HiScribe T7 Quick High Yield RNA Synthesis kit (NEB) from the DNA templates purchased from Twist Biosciences. Target cleavage assays performed with AwaIscB contained 10 nM of DNA substrate, 1 μ M of protein, and 4 μ M of ω RNA in a final 1x reaction buffer of 20 mM HEPES pH 7.5, 50 mM NaCl, and 5 mM MgCl₂. Assays were allowed to proceed at 37°C for 1 hour, then briefly shifted to 50°C for 5 min, and immediately placed on ice to help relax the RNA structure prior to RNA digestion. Reactions were then treated with RNase A (Qiagen), and Proteinase K (NEB), and purified using a PCR cleanup kit (Qiagen). DNA was resolved by gel electrophoresis on Novex 10% TBE (dsDNA substrates), 6% TBE-Urea (dsDNA substrates),

and 15% TBE-Urea (ssDNA substrates) polyacrylamide gels (Thermo Fisher Scientific). Target cleavage assays performed with CRISPR-associated IscB RNPs were performed similarly, except that 450 nM RNP replaced the protein and the ω RNA, and the reactions were incubated for 1.5 hours at 37°C. Cleaved reactions were then run on a 4% agarose E-gel (Thermo Fisher Scientific). Target cleavage assays with AmaTnpB were performed with 1 μ M of protein, 3 μ M of ω RNA and 10nM substrate, purified as described, and visualized on 2% agarose E-gels (dsDNA substrates) or 10% TBE-Urea polyacrylamide gels (ssDNA substrates).

[01106] For kinetic analysis of AwaIscB or AmaTnpB activity, cleavage reactions were quenched with 11 mM of EDTA at each time point prior to cleanup. For screening metals, MgCl₂ was eliminated from the reaction buffer, while 2 mM of EDTA, and 7 mM of the indicated metal were added. Collateral cleavage assays for AwaIscB were performed using 10 nM of unlabeled ds/ssDNA substrate along with 10 nM of Cy5.5-labeled collateral ssDNA substrate, and the reactions were allowed to proceed for 3 hours at 37°C. Cleavage of the labeled nontargeted ssDNA was then assessed on a 15% TBE-Urea polyacrylamide gel. Collateral cleavage assays for AmaTnpB were performed with 10 nM of unlabeled ds/ssDNA substrate and 10 nM of Cy5.5-labeled collateral ssDNA substrate for 1 hour at 60°C and the resulting cleavage was visualized on a 10% TBE-Urea gel.

[01107] Single-stranded RNA (ssRNA) substrates were *in vitro* transcribed and labeled with pCp- Cy5 (Jena Bioscience) on their 3' end. For the 3' end labeling, 50 pmol of ssRNA was incubated with 100 pmol of pCp-Cy5 and 50 U of T4 RNA ligase 1 (NEB) in 50 mM Tris pH 7.8, 10 mM MgCl₂, 10 mM DTT, 2 mM ATP, and 10% DMSO at 4°C for 40 hours. Labeling reactions were quenched with 20 mM EDTA and purified using an RNA Clean and Concentrator-25 kit (Zymo). ssRNA cleavage assays were performed similarly to the DNA cleavage assays, quenched with 19 mM EDTA at the end of the reactions, treated with Proteinase K, and visualized on a 6% TBE- Urea polyacrylamide gel.

[01108] For all experiments, all conditions were performed at least twice for replicability.

Cell-free transcription/translation TAM screen

[01109] IscB protein sequences were human codon optimized using the GenScript codon optimization tool, and IscB genes, TnpB genes with endogenous codon optimization, and

ωRNA scaffolds were synthesized by Twist Biosciences. Transcription/translation templates were generated by PCR from custom synthesis products. Cell-free transcription/translation reactions were carried out using a PURExpress *In Vitro* Protein Synthesis Kit (NEB) as per the manufacturer's protocol with half-volume reactions, using 75 ng of template for the protein of interest, 125 ng of template for the corresponding ωRNA with a guide targeting the TAM library and 25 ng of TAM library plasmid. Reactions were incubated at 37°C for 4 hours, then quenched by placing at 4°C or on ice and adding 10 ug RNase A (Qiagen) and 8 units Proteinase K (NEB) each followed by a 5 min incubation at 37°C. DNA was extracted by PCR purification and adaptors were ligated using an NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB) using the NEBNext Adaptor for Illumina (NEB) as per the manufacturer's protocol. Following adaptor ligation, cleaved products were amplified specifically using one primer specific to the TAM library backbone and one primer specific to the NEBNext adaptor with a 12-cycle PCR using NEBNext High Fidelity 2X PCR Master Mix (NEB) with an annealing temperature of 63°C, followed by a second 18-cycle round of PCR to further add the Illumina i5 adaptor. Amplified libraries were gel extracted, quantified by qPCR using a KAPA Library Quantification Kit for Illumina (Roche) on a StepOne Plus machine (Applied Biosystems/Thermo Fisher Scientific) and subject to single-end sequencing on an Illumina MiSeq with Read 1 80 cycles, Index 1 8 cycles and Index 2 8 cycles. TAMs were extracted and enrichment score for each TAM was calculated by filtering for all TAMs present more than once and normalizing to the TAM frequency in the input library subject to the same *in vitro* transcription/translation and quenching reactions. A position weight matrix based on the enrichment score was generated and Weblogos were visualized based on this position weight matrix using a custom Python script (Altae-Tran, H. et al. (2021)).

Expression and purification of KraIscB-1 RNP complex

[01110] Purification of the KraIscB-1 in complex with the ncRNA of its native locus was performed similarly to the CRISPR-associated IscB-ncRNA RNP complex with the following modifications: (1) the KraIscB-1 CDS was not codon optimized; (2) co-expression with the ncRNA was performed in BL21(DE3) cells (NEB) in the presence of 100 μg/ml ampicillin, and 25 μg/ml kanamycin; (3) bdSEN1 protease was not used since KraIscB-1 protein was not attached to an N-terminal tag, but was only twin-strep tagged on its C-terminus. Once the

boundaries of the co-purified RNA were defined by small RNA sequencing, the predicted ω RNA sequence was cloned downstream of a T7 promoter in a pCOLADuet-1 vector for inducible expression, and the KraIscB-1- ω RNA complexes were then prepared following the same procedure.

Cell-free transcription/translation cleavage assays

[01111] ω RNA templates were amplified from custom synthesis products as described and *in vitro* transcribed using a HiScribe T7 Quick High Yield RNA Synthesis Kit (NEB) with 150 ng DNA template in 2 μ L, 2 μ L T7 RNA Polymerase Mix (NEB) and 6.67 mM final concentration each NTP in 30 μ L total volume reactions and purified with an RNA Clean & Concentrator-25 kit (Zymo). Protein sequences were amplified from custom synthesis products or locus plasmid templates. To generate targets, short oligos containing target and PAM sequences with appropriate overhangs were synthesized by Genewiz and cloned into corresponding backbone plasmids by Golden Gate or restriction-ligation cloning. Dye-conjugated primers for generating labeled linear targets with dyes as annotated were synthesized by IDT and linear targets were amplified from target plasmids by PCR using Q5 Hot Start High Fidelity 2x Master Mix (NEB) as per the manufacturer's protocol.

[01112] Cell-free transcription/translation reactions were carried out using a PURExpress *In Vitro* Protein Synthesis Kit (NEB) as per the manufacturer's protocol with half-volume reactions using 75 ng of template for the protein of interest and a final concentration of 1 μ M *in vitro*-transcribed ω RNA to cap the possible final RNP concentration. Reactions were incubated at 37°C for 4 hours to allow for RNP formation, then placed on ice to quench *in vitro* transcription/translation. 50-100 ng of target substrate was then added and the reactions were incubated at the specified temperature for 1 additional hour. Reactions were then quenched by adding 10 μ g RNase A (Qiagen) and 8 units Proteinase K (NEB) each followed by a 5 min incubation at 37°C. DNA was extracted by PCR purification and run on either 10% or 6% Novex TBE-Urea gels, or 10% Novex TBE gels (Thermo Fisher Scientific) as per the manufacturer's protocols, as specified in figures. Gels were stained with 1X SYBR Gold (Thermo Fisher Scientific) where specified for 10-15 min and imaged on a ChemiDoc imager (BioRad) with optimal exposure settings. Each condition was performed twice for replicability.

Mammalian cell culture and transfection

[01113] Mammalian cell culture experiments were performed in the HEK293FT line (American Type Culture Collection (ATCC)) grown in Dulbecco's Modified Eagle Medium with high glucose, sodium pyruvate, and GlutaMAX (Thermo Fisher), additionally supplemented with 1× penicillin–streptomycin (Thermo Fisher), 10 mM HEPES (Thermo Fisher), and 10% fetal bovine serum (VWR Seradigm). All cells were maintained at confluency below 80%.

[01114] All transfections were performed with Lipofectamine 2000 (Thermo Fisher). Cells were plated 16-20 hours prior to transfection to ensure 90% confluency at the time of transfection. For 96-well plates, cells were plated at 20,000 cells/well, and for 24-well plates, cells were plated at 100,000 cells/well. For each well on the plate, transfection plasmids were combined with OptiMEM I Reduced Serum Medium (Thermo Fisher) to a total of 25 µL. Separately, 23 µL of OptiMEM was combined with 2 µL of Lipofectamine 2000. Plasmid and Lipofectamine solutions were then combined and pipetted onto cells.

Mammalian lysate cleavage assays

[01115] Human codon-optimized IscB genes were cloned into a CMV expression backbone by Gibson assembly using 2X Gibson Assembly Master Mix (NEB) to generate pCMV-SV40 NLS- IscB protein-nucleoplasmin NLS-3xHA constructs. 500 ng each protein expression plasmid was transfected in a separate well of a 24-well plate as described. After approximately 48 hours, cells were washed with 500 µL of Dulbecco's phosphate buffered saline (Sigma Aldrich). 50 µL ice- cold lysis buffer (20 mM HEPES 7.5, 100 mM KCl, 5 mM MgCl₂, 0.1% Triton-X 100, 5% glycerol, 1 mM DTT, 1X cOmplete Protease Inhibitor Cocktail) was added, then cells were scraped from the plate, transferred to clean tubes and incubated on ice for 15 min. Cells were then sonicated in a cold water bath with amplitude 30 for 4 cycles of 10 s each. Lysate was then cleared by centrifugation at max speed for 20 min and supernatant was collected and either used fresh in assays or snap frozen in liquid nitrogen for later use. Labeled targets and *in vitro*- transcribed ωRNA was generated as described in "Cell-free transcription/translation cleavage assays" above.

[01116] To perform cleavage assays, 10 µL cell lysate was incubated with 1 µg *in vitro*-transcribed ωRNA or sgRNA, or no RNA for negative controls, and 100 ng target substrate in

1X NEBuffer 3.1 (NEB). Reactions were incubated at 37°C for 1 hour, then quenched by adding 10 ug RNase A (Qiagen) and 8 units Proteinase K (NEB) each followed by a 5 min incubation at 37°C. DNA was extracted by PCR purification and run on a 4% Agarose E-gel EX (Thermo Fisher Scientific) as per the manufacturer's instructions and visualized on a ChemiDoc imager (BioRad).

Expression and purification of AwaIscB

[01117] The human codon optimized AwaIscB protein was expressed from a pET45b(+) plasmid backbone, with a His14-bdSUMO tag attached to its N-terminus, and a twin-strep tag attached to its C-terminus. Rosetta(DE3)pLysS cells transformed with the expression construct were grown at 37°C in the terrific broth (TB) medium supplemented with 100 µg/ml ampicillin, and 34 µg/ml chloramphenicol, and shifted to 18°C at an OD600 of 0.3. Protein production was induced after reaching an OD600 of 0.6-0.8 with 0.2 mM IPTG and continued at 18°C for 16-18 hours. Cells were harvested by centrifugation, and resuspended in lysis buffer (50 mM Tris pH 8, 1 M NaCl, 5% glycerol, 5 mM MgCl₂, 40 mM imidazole, and 5 mM β-mercaptoethanol) supplemented with benzonase (Sigma), and protease inhibitors (PMSF and Roche cOmplete, EDTA-free), and then lysed by two passes with a high-pressure homogenizer (LM20 Microfluidizer, Microfluidics). After clearing the lysate by centrifugation, the soluble fraction was bound to Ni-Sepharose 6 Fast Flow resin (GE Healthcare). The Ni beads were first washed with lysis buffer, then briefly with buffer C (50 mM Tris pH 8, 2 M NaCl, 5% glycerol, 5 mM MgCl₂, 40 mM imidazole, and 5 mM β-mercaptoethanol), and buffer D (50 mM Tris pH 8, 500 mM NaCl, 5% glycerol, 5 mM MgCl₂, 40 mM imidazole, and 5 mM β-mercaptoethanol). The AwaIscB protein was then eluted in elution buffer (buffer D containing 300 mM imidazole), and dialyzed overnight against buffer E (20 mM Tris pH 8, 500 mM NaCl, 5% glycerol, 5 mM MgCl₂, and 0.5 mM TCEP). Following dialysis, the protein was purified through a HiTrap Heparin HP column against a gradient of 0.25–2 M NaCl. Afterwards, peak fractions containing AwaIscB were pooled and dialyzed overnight once against buffer E. Protein was concentrated up to 1.5-2 mg/ml, aliquoted, snap-frozen, and stored at –80°C.

Sequencing of cleavage products

[01118] *In vitro* cleavage assays were performed as described. Purified reactions were subjected to a GLOE-seq library preparation protocol (Sriramachandran, A. et al. (2020), *Mol. Cell.* **78**, 975- 985.e7) as described using 2.5 μ M adapter as input to the proximal adapter annealing step. The final amplification to add Illumina adapters and barcodes was performed with NEBNext High Fidelity 2x PCR Master Mix (NEB) with an annealing temperature of 63°C for 15 s and 12 cycles. Libraries were subjected to paired-end sequencing using an Illumina MiSeq with Read 1 150 cycles, Read 2 150 cycles, Index 1 8 cycles and Index 2 8 cycles. Paired-end reads were mapped to the target substrate using BWA and 3' ends were extracted and plotted using a custom Python script.

Enzymatic footprinting assay

[01119] dsDNA substrate (191 bp) was produced by PCR amplification from a plasmid containing the target site and the TAM sequence. 10 pmol of dAwaIscB and 40 pmol of ω RNA were incubated in reaction buffer (20 mM HEPES pH 7.5, 50 mM NaCl, 10 mM MgCl₂, and 5% glycerol) at 37°C for 30 min. Then, 0.1 pmol of DNA substrate was added and the reaction was allowed to proceed for another 30 min at 37°C. Next, 500 U of Exonuclease III (NEB) was added, the assay was incubated for an additional 10 min at 37°C and quenched with 20 mM EDTA. As negative control, another reaction was run in parallel, in which ω RNA was excluded and the volume was replaced with water. After quenching, both reactions were briefly shifted to 50°C for 5 min, then immediately placed on ice, treated with RNase A (Qiagen), and Proteinase K (NEB), and purified using a PCR cleanup kit (Qiagen). Purified reactions were subjected to a GLOE-seq library preparation protocol (Sriramachandran, A. et al. (2020), *Mol. Cell.* **78**, 975- 985.e7) as described using 2.5 μ M adapter as input to both the proximal and distal adapter annealing steps. Libraries were amplified as described in the “Sequencing of cleavage products” section above and subjected to paired-end sequencing using an Illumina MiSeq with Read 1 100 cycles, Read 2 100 cycles, Index 1 8 cycles and Index 2 8 cycles. Paired-end reads were mapped to the target substrate using BWA and 3' ends were extracted and plotted using a custom Python script (Altae-Tran, H. et al. (2021), Zenodo (2021)).

Mammalian genome editing

[01120] ω RNA scaffold backbones were cloned into a pUC19-based human U6 expression backbone by Gibson Assembly. For initial testing, 12-guide libraries were cloned in a pool mixing primers to add each of the 12 guides in in a given pool at equimolar ratios, and ω RNA scaffold backbones were subject to whole plasmid amplification with guide primers annealing to the U6 promoter and a second primer annealing to the start of the ω RNA scaffold using Phusion Flash High-Fidelity 2X Master Mix (Thermo Fisher Scientific). PCR products were gel extracted and eluted in 30 μ L, then blunt-end ligated to circularize by addition of 5 units T4 PNK (NEB), 200 units T4 DNA Ligase (NEB) and final 1X T4 DNA Ligase Buffer (NEB) and incubation for 1.5 h at room temperature prior to transformation in Stbl3 chemically competent *E. coli* (NEB). For individual guide constructs, oligos with appropriate overhangs were synthesized by Genewiz, annealed and phosphorylated using T4 PNK (NEB) and cloned into ω RNA backbones by restriction-ligation cloning. Protein expression constructs were cloned as described in “Mammalian lysate cleavage assays” above.

[01121] Prior to testing individual guides, each tested IscB protein was screened for activity in HEK293FT using a pool of 12 guides cloned as described. This was done to more rapidly assess many guides for a given protein for a broader survey of possible guides, as Applicants had no prior knowledge of the effect of guide properties such as sequence or genomic location may affect the ability of IscB to generate indels in eukaryotic cells. For this 12-guide pooled initial screening of IscB proteins, 800 ng protein expression construct and 1200 ng of the corresponding guide pool with corresponding ω RNA scaffold were transfected in one well of a 24-well plate as described. After 60-72 hours, genomic DNA was harvested by washing the cells once in 1xDPBS (Sigma Aldrich) and dry trypsinizing cells using TrypLE (Thermo Fisher Scientific). Trypsinized cells were collected in 1 mL 1xDPBS and pelleted by centrifugation at 300xg at 4°C for 5 min. The supernatant was removed and cells were resuspended in 50 μ L QuickExtract DNA Extraction Solution (Lucigen) and cycled at 65°C for 15 min, 68°C for 15 min then 95°C for 10 min to lyse cells. 2.5 μ L of lysed cells were used as input into each PCR reaction. Amplification of each region targeted by a guide in a given guide pool was performed individually. Insertion/deletion (indel) frequency was analyzed using CRISPResso2 (Clement, K. et al. (2019), *Nat. Biotechnol.* **37**, 224–226). Given the low frequency of indel events with IscBs, in order to eliminate noise from PCR and sequencing error, only indels with at least 2 reads or more than 1 base inserted or deleted were counted towards reported indel frequencies.

The allele frequency table generated by CRISPResso2 was used for manual verification of indels (i.e. both insertions and deletions occurring, occurrence of insertions and deletions of diverse sizes, not occurring in homonucleotide stretches), to determine if computed indels were due to PCR or sequencing error given that putative detected indels often occurred at frequencies similar to the detection threshold based on the number of cells input used in the experiment and the number of reads obtained from sequencing. Indel quantification results from the 12-guide pools are available in Table 11. Guides with manually verified indels from this pooled screen were assessed individually to validate indel formation as described below.

[01122] For individual guide sequences, 250 ng guide/ωRNA expression plasmid and 125 ng protein expression plasmid were transfected in each of 4 wells as biological replicates in a 96- well plate for each guide condition as described. After 60-72 hours, genomic DNA was harvested by washing the cells once in 1xDPBS (Sigma Aldrich) and adding 50 μL QuickExtract DNA Extraction Solution (Lucigen). Cells were scraped from the plates to suspend in QuickExtract and cycled at 65°C for 15 min, 68°C for 15 min then 95°C for 10 min to lyse cells. 2.5 μL of lysed cells were used as input into each PCR reaction.

[01123] For library amplification, target genomic regions were amplified and with a 12-cycle PCR using NEBNext High Fidelity 2X PCR Master Mix (NEB) with an annealing temperature of 63°C for 15 s, followed by a second 18-cycle round of PCR to add Illumina adapters and barcodes. The libraries were gel extracted and subject to single-end sequencing on an Illumina MiSeq with Read 1 300 cycles, Index 1 8 cycles, Index 2 8 cycles. Insertion/deletion (indel) frequency was analyzed using CRISPResso2 (Clement, K. et al. (2019), *Nat. Biotechnol.* **37**, 224–226) as described. For individual guide/ωRNA experiments, to assess statistical significance, 2-tailed T-tests were performed using non- targeting guide/ωRNA conditions as a negative control (see Table 11).

I. tetrasporus shotgun sequencing

[01124] An agar slant of *I. tetrasporus* UTEX B 2012 was obtained from UTEX inoculated in Modified Bold 3N Medium (UTEX) and grown at 20°C in a shaking incubator with 12 hour light/dark cycles. After 14 days, cultures were spun down and DNA was extracted using a DNeasy Plant Mini kit (Qiagen). Total DNA was tagmented using Tn5 and tagmentation reactions were quenched by addition of Buffer PB (Qiagen) and PCR purified. 10 μL purified

DNA was used as input to PCR using 2X KOD HotStart Master Mix (Millipore Sigma) as per the manufacturer's instructions with an annealing temperature of 60°C and 30 cycles to add Illumina adapters. Amplified libraries were gel extracted, quantified by qPCR using a KAPA Library Quantification Kit for Illumina (Roche) on a StepOne Plus machine (Applied Biosystems/Thermo Fisher Scientific) and subject to paired-end sequencing on an Illumina NextSeq with Read 1 150 cycles, Read 2 150 cycles, Index 1 8 cycles and Index 2 8 cycles. SPAdes-3.15.2 was used with the "isolate" option and kmer lengths of 21, 33, 55, 67 to create a contig level assembly from the reads (Bankevich, A. et al. (2012), *J. Comput. Biol.* **19**, 455–477). BLASTn was then used to compare the final contigs containing ChlorIscB and IscB fragments with the published UTEX B 2012 chloroplast genome (GenBank accession KY407659.1) (Turmel, M. et al. (2017), *Sci. Rep.* **7**, 994).

Expression and purification of TnpB proteins

[01125] To purify the *A. lobatus* TnpB-2 protein in complex with the putative ω RNA of its native locus, an N-terminally His14-MBP tagged TnpB CDS and the corresponding downstream locus up to 80 bp beyond the end of predicted guide adaptor were cloned as a single piece downstream of a T7 promoter in a pET45b(+) vector. The AloTnpB-2 RNP was expressed and purified similarly to the CRISPR-associated IscB RNPs with the following modifications: (1) the lysis buffer, and the buffers A and B were supplemented with 40 mM imidazole, and 5 mM β -mercaptoethanol, while MgCl₂ and DTT in all of the buffers were eliminated; (2) the RNP was purified on Ni-Sepharose 6 Fast Flow beads (GE Healthcare) instead of the Strep-Tactin resin; the elution buffer was supplemented with 300 mM imidazole and 5 mM β -mercaptoethanol, while MgCl₂, DTT and desthiobiotin were eliminated; (4) the His14-MBP solubility tag was kept attached to the RNP to provide stability.

[01126] To purify the *A. macrosporangiidus* TnpB, the native sequence was expressed from a pET45b(+) plasmid backbone, with a His14-MBP tag attached to its N-terminus via a TEV protease cleavage site. The AmaTnpB was expressed and purified similarly to AwaIscB with the following modifications: (1) expression was performed in BL21(DE3) cells (NEB) in the presence of 100 μ g/ml ampicillin; (2) MgCl₂ was omitted from buffers C, D and E and the elution buffer; (3) TEV protease was added to the dialysis for solubility tag cleavage; (4)

following dialysis, the protein was purified through a Resource S column against a gradient of 0.2–2M NaCl; (5) protein was concentrated up to 5 μ M.

Supplementary Text

Extended discussion of hybrid CRISPR/ ω RNA and ω RNA secondary structure analysis

[01127] Comparing the CRISPR-associated IscB ncRNA and ω RNA covariance model-based

[01128] secondary structure predictions (Fig. 9G), Applicants found that all 10 base positions that were more than 97% conserved in the ω RNA were also highly conserved in the CRISPR-associated IscB ncRNA. Both the ω RNA and CRISPR-associated ncRNA contain a short 3-4 bp nexus, a conserved nexus hairpin, and a large region consisting of two interconnected multi-stem loops. The ω RNA additionally contains a hairpin that encompasses a Shine-Dalgarno (SD) sequence located approximately 10 bp upstream of the CDS start codon, implying that the ω RNA might be involved in the regulation of the translation of IscB. A statistically significant pseudoknot structure was predicted to exist in the ω RNA between the loop of one hairpin in the multi-stem loop region and the region directly downstream of the nexus, which Applicants termed the nexus pseudoknot hairpin (Fig. 9G). Examination of individual CRISPR-associated ncRNA sequences suggested the presence of a homologous pseudoknot structure (Fig. 9G inset). Scrambling the sequences found in the pseudoknot of the ncRNA disrupting the structure abolished the cleavage activity of the CRISPR-associated IscB. However, mutating the pseudoknot sequence while maintaining base pairing had no effect on cleavage activity, suggesting that the pseudoknot structure, but not necessarily sequence, was important for ncRNA function (Fig. 47).

[01129] The majority of *iscBs* are associated with predicted ω RNA scaffolds matching the structure of the covariance model shown in Figure 15B, but several large groups are not. To explore the diversity of the possible ω RNA scaffolds, Applicants iteratively built a set of profiles that span all major groups of ω RNAs associated with *iscBs* and *isrBs*, and found that different *iscB/isrB* clades were associated with structurally distinct ω RNAs (Fig. 31C). The structures differed by the number of hairpins in the multi-stem loop and guide adaptor regions, but the nexus pseudoknot was present in all major *iscB*- and *isrB*-associated ω RNAs. The *IsrB* ω RNAs additionally contained a second conserved pseudoknot between the guide adaptor and

a hairpin loop in the multi-stem loop region. Two of the three main IsrB ω RNA structures contained a third hairpin in the multi-stem loop region that forms the pseudoknot with the guide adaptor.

[01130] This guide adaptor pseudoknot was unique to IsrB ω RNAs and appears to have been lost contemporaneously with the insertion of the HNH domain (Fig. 41A-C).

Extended phylogenetic analysis and discussion

[01131] To assess the robustness of our phylogenetic analyses and interpretations, Applicants describe in this section the additional analyses Applicants performed and the reasoning behind our phylogenetic conclusions.

[01132] To obtain maximally robust trees, Applicants used modified parameters for tree inference software. For all IQ-Tree 2 inferred trees, 5000 ultrafast bootstraps were performed with the bnni option to help reduce the impact of severe model violations on bootstrap support values (Hoang, D. et al. (2018), *Mol. Biol. Evol.* **35**, 518–522).

[01133] Under this option, individual bootstrap trees are optimized using nearest neighbor exchange with hill-climbing on the likelihood function based on the bootstrap alignment. To improve the optimization performance of IQ-Tree 2 relative to default parameters, Applicants used nstop=500 (500 iterations of no improvement before terminating), with ninit=500 (500 initial trees), ntop=100 (100 top trees retained), and nbest=20 (20 best trees retained throughout the search and iteratively improved upon). For RAxML, Applicant used 2000 rapid bootstraps for each tree to obtain accurate bootstrap values. For MrBayes, Applicants typically used 8-16 chains with a temperature difference between 0.01 and 0.025. Applicants also used 4-8 (resources permitting) independent runs to obtain a better estimate of convergence. Due to the occurrence of numerical overflow in MrBayes when computing likelihoods with a large number of taxa, Applicants employed a reaper script to restart from the last checkpoint when a numerical overflow was detected in any of the chains during the runs.

[01134] Applicants analyzed multiple different alignments to ensure robustness of the phylogenetic conclusions. Focusing the alignments only on the core, highly conserved regions containing the catalytic sites provided for comparison of IscB, IsrB, and Cas9s, proteins that widely differ in size and domain composition. Furthermore, analysis of slowly evolving regions enables phylogenetic inferences to be made over long evolutionary larger timescales that

encompass the evolution of entire protein families, such as IscB and Cas9. The presence of multiple decaying, inactivated variants of IscB, IscrB, and Cas9 with high sequence divergence from active forms rendered long branch attraction a potential issue (Lockhart, P. et al. (1996), *Proc. Natl. Acad. Sci. U. S. A.* **93**, 1930–1934). Therefore, all inactivated proteins were removed from the RuvC/BH and RuvC/BH/HNH alignments to ensure the phylogeny tracked the evolution of active nucleases with the maximum possible accuracy, avoiding the artifacts caused by mixing active and inactive nucleases that evolve with substantially different rates. The evolution of Cas9 from IscB through the intermediate CRISPR-associated IscB cluster 2089 was supported by all inferred trees, and the evolution of IscB from IscrB is also strongly supported. However, to ensure robustness of these interpretations, Applicants considered numerous phylogenetic parameters, compared Bayesian and maximum likelihood methods, and also assessed the effect of taxa sampling (Figs. 22, 35C-D, 52-53, 54A-B).

[01135] Next, Applicants tested whether the main assumptions of phylogenetic inference held for the alignments used for our phylogenetic analyses. In some cases, the stationarity condition did not hold for the default alignment, prompting additional modifications to reduce the extent of stationarity violation. The stationarity violations detected for RuvC/BH/HNH alignments were mostly resolved by removing Cas9s of subtype II-B (prompted by the unstable placement of the major II-B Cas9 clade in the phylogenetic trees) and were completely eliminated by removing all Cas9s except for the earliest forms that evolved from IscB. This improvement in the tree quality suggests different evolutionary regimes for the core regions of IscB/IscrB (the RuvC, BH, and HNH domains) are substantially different in IscB/IscrBs and the homologous regions of the late, specialized Cas9s (II-B, II-C, *tmpA*-associated II-C, and the majority of II-A). Such a difference in evolutionary pressures is likely to stem from the shift in biological functionality between Ω systems (IscB/IscrB) and CRISPR-Cas systems (Cas9). Applicants performed additional phylogenetic analyses using these two modified alignments in order to assess the consistency of the inferences. The origin of Cas9 from the IscB cluster 2089 was consistently reproduced in these analyses (Figs. 32-33, 35, 54A,C, 56).

[01136] Due to the removal of clusters that included inactivated nucleases, some CRISPR-associated IscBs were removed. For experimental characterization, Applicants considered all CRISPR-associated clusters regardless of whether or not they were included in the main phylogenetic analyses. In particular, IscB cluster 34507 contained mostly inactivated variants. Applicants studied IscBs from this cluster experimentally, reasoning that the high level of

inactivation implies that the active members of this cluster are likely to have a high level of activity that could render them toxic to the host cell resulting in selection driving inactivation. To determine the likely position of this cluster in the full phylogenetic tree, Applicants performed an additional phylogenetic analysis of full IscB proteins (excluding the C-termini that aligned poorly). Applicants found that the experimentally studied CRISPR associated IscBs belonged to a small clade of CRISPR-associated IscBs that includes 4 non-redundant, CRISPR-associated loci. This analysis also yielded additional information on the relationship between IscB and Cas9. Because the alignment used for this analysis consisted of all IscBs and Cas9_1261 and Cas9_665 (two previously identified early Cas9 clusters from II-D), Applicants were able to construct a MAFFT-einsi alignment that spanned the entire length of these proteins (except for the C-terminal and the PLMP domain that is missing in Cas9). The phylogenetic analysis based on this alignment incorporated information from all homologous regions of IscB and Cas9 rather than the core region only. Clusters lacking coverage to any of the catalytic domains were removed; however, clusters including inactive nucleases were retained in the analysis. Positions in the alignment with more than 75% gaps were removed, and the tree was inferred using IQ-Tree 2 with the modified parameters described above. The exclusion of additional Cas9s in this analysis could potentially skew the placement of the Cas9 family. Nevertheless, the resulting tree supported CRISPR-associated IscB cluster 2089 as the ancestor of Cas9 (Fig. 56).

[01137] Phylogenetic symmetry tests showed that the main RuvC/BH/HNH alignment that included all Cas9s and IscBs likely violated the stationarity condition used by most phylogenetic methods (SRH conditions). Similarly, phylogenetic symmetry tests show that the high resolution (individual proteins as opposed to cluster representatives) codon DNA sequence alignment of early *cas9*s also likely violates the stationarity condition. Applicants considered a number of methods to assess the validity of the phylogenetic trees inferred using these alignments. For the early Cas9 tree, Applicants compared the tree inferred from the amino acid sequence (IQ-Tree 2, WAG+F+I+G4 model) (Fig. 32A) to the tree inferred from the underlying DNA codon sequences (GTR+G4) (Fig. 33A). The two trees were closely congruent with respect to the transition from IscB to Cas9 although despite the DNA tree inference only involved the simple GTR substitution model that does not take codon level information into account (Figs. 32A, 33A).

[01138] To assess the generalizability of the phylogenetic conclusions, in addition to bootstrap analyses in all ML trees, Applicants also performed a sensitivity analysis of the RuvC/BH/HNH phylogenetic inferences. Specifically, Applicants considered the effect of random exclusion of taxa. To this end, Applicants explored a range of taxa dropout rates. For each dropout rate, r , Applicants selected $1000 / (1-r)$ alignments with r taxa randomly dropped out. Because Applicants were interested predominantly in the topology surrounding early Cas9 evolution, Applicants ensured that IscB cluster 2089 (putative Cas9 modern ancestor) and early Cas9 clusters (Cas9_1261, Cas9_665, Cas9_1079, Cas9_849) were retained in every alignment. Applicants then obtained inference for each of these taxon subsampled alignments. All trees were rooted on the clade of IscBs considered to be ancestral, that is evolved from IsrB, as inferred from the RuvC/BH tree. Next, various features (such as percentage of trees where Cas9 is monophyletic) were summarized across all trees for each dropout rate (Fig. 35). Error bars (95% confidence intervals) for the estimates were computed using 2000 bootstraps over the tree samples. Due to the computationally intense nature of ML inference for alignments with many taxa, Applicants limited this dropout analysis to FastTree2 (as opposed to IQ-Tree 2) based phylogenetic inference. The results of this analysis supported that Cas9_1261, Cas9_665 were the oldest Cas9 groups, and that all Cas9s likely originated from the ancestors of these two clusters. Moreover, the tree supported that even with taxon dropout rates of up to 50%, most of the resulting trees supported the CRISPR-associated IscB cluster 2089 as the ancestor to all extant Cas9s.

[01139] Applicants next considered whether or not the coarse graining of proteins (specifically the use of cluster representatives), as opposed to including all non-redundant sequences (at 90% redundancy), affected the phylogenetic inference. Applicants curated two “high resolution” alignments consisting of all non-redundant proteins of interest (Cas9/IscB) from clusters related to the early evolution of Cas9. One version of this alignment consisted of full protein alignments (using MAFFT-einsi) with the REC domains and PLMP domains removed. Positions with 75% or more gaps were removed. Another alignment consisted of translation aligned DNA codons for the ORFs, again removing REC domains and PLMP domains. Highly gapped codons triplet positions (sets of 3 positions) were removed using BMGE (Criscuolo, A. et al. (2010), *BMC Evol. Biol.* **10**, 210). The resulting alignments were analyzed using IQ-Tree 2 and MrBayes. As the DNA alignment suffered from severe violation of the stationarity requirement for phylogenetic analysis, Applicants considered further using

heterotachy models to account for changing modes of evolution throughout the underlying evolutionary tree. Adding a heterotachy-aware model did not change the inferred sequence of events in the evolution of early Cas9 from IscB (Fig. 33B). Phylogenetic analyses of the DNA sequences were agnostic to codons. Nevertheless, the two sets of alignments produced closely similar topologies for early events in Cas9 evolution. These topologies further agreed with previously identified topologies inferred using coarse grained clusters, suggesting that clustering was not detrimental to the inference of phylogenies.

[01140] Comparison of IscB/IsrB ω RNA phylogeny to the protein phylogenies can be useful for identifying co-evolution, and potentially could provide more information regarding the ancestry of Cas9. However, the structural motifs differed across the diversity of ω RNAs. Therefore, the main large-scale phylogenetic analysis across IscB/IsrB ω RNAs used a non-structured nucleotide alignment of representative ω RNAs that was produced using MAFFT-ginsi. Because the nexus pseudoknot is conserved in almost all ω RNAs, greater emphasis was given to this region by removing all sequences with low coverage to the nexus pseudoknot. An IQ-Tree 2 ML inference showed that IscB ω RNAs are a deep branch in the tree of ω RNAs, suggesting that IscB ω RNAs evolved on a single occasion from IsrB ω RNAs (Fig. 55A). Due to the high diversity of ω RNAs, the early Cas9_1261 and Cas9_665 tracrRNAs could not be included in the analysis for accurate placement because as the alignment across all RNAs failed to capture sufficient global sequence conservation. CMsearch of two tracrRNAs from early Cas9s (Cas9_1261 and Cas9_665) yielded significant hits to the G1a ω RNA profile (E-values 3.7e-8, 4.1e-8 respectively), suggesting that at least part of the tracrRNA in early Cas9 systems originated from the IscB ω RNA (Fig. 36). The likely evolutionary link between ω RNAs and tracrRNAs further suggests that Cas9 evolved from IscB. To further address the question of whether or not early Cas9 tracrRNAs were derived from ω RNAs, Applicants constructed an accurate MAFFT-ginsi alignment of the early Cas9 tracrRNAs along with 25 closely related IscB ω RNAs, including the hybrid CRISPR/ ω RNA from IscB cluster 2089. To account for the discrepancy between tracrRNAs and ω RNAs, Applicants concatenated the DR of the corresponding CRISPRs to the tracrRNA with a 4 bp poly-A tetraloop joining the two. Applicants performed a Bayesian phylogenetic analysis to address potential issues with the smaller number of taxa in this analysis. Applicants found that the combined DR/tracrRNAs likely evolved from IscB (2089) like ω RNAs, with high posterior probability (branch posterior

89% $\pm$ 1%, Fig. 55C). This result further suggests that all Cas9s evolved from the CRISPR-associated IscB cluster 2089.

[01141] Additional evidence from protein or RNA alignment comparisons were also useful in identifying the most likely ancestor of Cas9. On the sequence level, the Cas9_1261 and Cas9_665 representative sequences are more similar to the IscB cluster 2089 representative sequence in the split RuvC and HNH domains compared to other example IscBs (Fig. 35). Moreover, the N-terminus of Cas9_1261 potentially contains fragments of the PLMP domain that is characteristic of IscB, suggesting that the PLMP domain gradually degraded during the transition to Cas9. Lastly, the 5' end of the DRs from CRISPR-associated IscB cluster 2089 is closely similar to the 5'-terminal sequences of the DRs from Cas9_1261 and Cas9_665, suggesting further homology. All of these observations, in addition to the phylogenetic analyses, strongly suggest that CRISPR-associated IscB cluster 2089 is the ancestor of the earliest Cas9s. Accordingly, the sequence of events in the evolution of Cas9 was, first, the association of CRISPR arrays with IscB, either through duplication of the 5' of the ω RNA or by insertion of the ω RNA into a CRISPR array. Then, REC-like sequences were inserted between the BH and RuvC-II, or possibly inside the BH, splitting it into two domains, the second of which remains in some early Cas9s. Then, the PLMP domain degraded from the lengthened IscB protein becoming Cas9, which subsequently became associated with the CRISPR adaptation module. Given the rarity of the early Cas9s, such as Subtype II-D, compared to IscB and other Cas9 subtypes, it appears likely that these systems containing small, simply organized Cas9s were outcompeted by the subsequently evolved, more efficient, larger Cas9s. From our phylogenetic analyses, the most likely candidate set of adaptation genes that became associated with Cas9 at first included Cas1, Cas2, and Cas4 genes. The association with Cas4 was then likely lost, forming Type II-C Cas9s, which subsequently gained additional associations to become II-A (associated with Csn2), II-B (re-association with Cas4), and *tnpA*-associated II-C (re-associated with Y1 recombinases).

Extended discussion of potential biological functions of IS200/605 endonucleases.

[01142] The biological functions of Ω systems remain unknown, but several hypotheses appear to be compatible with available evidence.

[01143] First, *IscB*, *IsrB*, and *TnpB* could mediate RNA-guided *TnpA*-catalyzed transposition while suppressing transposition independent of ω RNA. Our phylogenomic analysis together with previous observations (Kapitonov, V. et al. (2015), *J. Bacteriol.* **198**, 797–807) indicate that *iscB*, *isrB*, and a subfamily of *cas9* have all associated with the Y1 transposase *tnpA* of IS200/IS605 transposons on multiple, independent occasions (Fig. 31A) (Kapitonov, V. et al. (2015), *J. Bacteriol.* **198**, 797–807). Moreover, the *iscB*, *isrB* and *tnpB* genes themselves are embedded in non- autonomous transposons of the IS200/IS605 superfamily. Integration events of ω RNA- containing transposons have been observed in locations with an ATGA or GTGA 3' flanking sequences (Weinberg, Z. (2009), *Nature* **462**, 656–659), which is a common TAM sequence for *IscB* proteins (Fig. 48). Many *IscB* TAMs contain start and stop codons, such that *IscB* targeting could mediate transposon insertion at the beginning or end of operons, allowing the transposon to exploit existing transcription initiation and termination signals. Together, these observations suggest the possibility that *IscB*, *IsrB* and *TnpB* are functionally analogous to CRISPR-associated transposases (CAST complexes) (Strecker, J. et al. (2019), *Science* **365**, 48–53; Klompe, S. et al. (2019), *Nature* **571**, 219–225), which mediate RNA-guided transposition of Tn7-like transposons. However, in contrast to CAST effectors such as Cas12k, an inactivated *TnpB* derivative (Strecker, J. et al. (2019), *Science* **365**, 48–53), the great majority of *IscB*, *IsrB* and *TnpB* are predicted to be active nucleases. The role of the nuclease activity of these proteins in RNA-guided transposition (if any) remains to be studied experimentally. One possibility is that these nucleases eliminate aberrant DNA structures emerging upon RNA-guided transposon insertion.

[01144] An orthogonal hypothesis is that these nucleases, together with the ω RNA, are a distinct variety of toxins, to which *TnpA* is the antitoxin. Such toxins would kill cells lacking the inserted IS200/605 locus, thus making the host addicted to the IS200/605 transposase and ensuring that the non-autonomous transposons encoding *IscB*/*TnpB* also maintain their transposition capacity.

[01145] In addition to its direct role in transposition, *TnpA* might function to acquire guides for Ω systems, analogously to the Cas1 protein that is thought to have evolved from the transposase of the casposons, a distinct class of self-synthesizing transposons (Krupovic, M. et al. (2014), *BMC Biol.* **12**, 1–12). The high sequence identity between the ω RNA transposon end-overlapping regions in both *iscB* and *tnpB* loci in *K. racemifer* suggests that, at least in this bacterium, some *IscBs* and *TnpBs* share a mechanism for acquiring guide sequences. More

generally, loci encoding complete Ω systems or standalone ω RNAs could be mobilized and moved to new locations by TnpA, thereby associating with new, diverse flanking sequences that become guides.

References for Extended Discussion of Supplemental Materials

1. F. Zhang, Development of CRISPR-Cas systems for genome editing and beyond. *Quarterly Reviews of Biophysics*. **52** (2019).
2. F. Hille, H. Richter, S. P. Wong, M. Bratovič, S. Ressel, E. Charpentier, The biology of CRISPR-Cas: Backward and forward. *Cell*. **172**, 1239–1259 (2018).
3. V. V. Kapitonov, K. S. Makarova, E. V. Koonin, ISC, a novel group of bacterial and Archaeal DNA transposons that encode Cas9 homologs. *J. Bacteriol.* **198**, 797–807 (2015).
4. P. Siguier, E. Gourbeyre, M. Chandler, Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol. Rev.* **38**, 865–891 (2014).
5. S. Shmakov, A. Smargon, D. Scott, D. Cox, N. Pyzocha, W. Yan, O. O. Abudayyeh, J. S. Gootenberg, K. S. Makarova, Y. I. Wolf, K. Severinov, F. Zhang, E. V. Koonin, Diversity and evolution of class 2 CRISPR-Cas systems. *Nat. Rev. Microbiol.* **15**, 169–182 (2017).
6. B. Zetsche, J. S. Gootenberg, O. O. Abudayyeh, I. M. Slaymaker, K. S. Makarova, P. Essletzbichler, S. E. Volz, J. Joung, J. van der Oost, A. Regev, E. V. Koonin, F. Zhang, Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*. **163**, 759–771 (2015).
7. E. Deltcheva, K. Chylinski, C. M. Sharma, K. Gonzales, Y. Chao, Z. A. Pirzada, M. R. Eckert, J. Vogel, E. Charpentier, CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*. **471**, 602–607 (2011).
8. I. Kalvari, E. P. Nawrocki, N. Ontiveros-Palacios, J. Argasinska, K. Lamkiewicz, M. Marz,

- S. Griffiths-Jones, C. Toffano-Nioche, D. Gautheret, Z. Weinberg, E. Rivas, S. R. Eddy, R.
D. Finn, A. Bateman, A. I. Petrov, Rfam 14: expanded coverage of metagenomic, viral and microRNA families. *Nucleic Acids Res.* **49**, D192–D200 (2021).
9. Z. Weinberg, J. Perreault, M. M. Meyer, R. R. Breaker, Exceptional structured noncoding RNAs revealed by bacterial metagenome analysis. *Nature*. **462**, 656–659 (2009).
 10. M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, E. Charpentier, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. **337**, 816–821 (2012).
 11. G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E2579-86 (2012).
 12. G. Gasiunas, J. K. Young, T. Karvelis, D. Kazlauskas, T. Urbaitis, M. Jasnauskaite, M. M. Grusyte, S. Paulraj, P.-H. Wang, Z. Hou, S. K. Dooley, M. Cigan, C. Alarcon, N. D. Chilcoat, G. Bigelyte, J. L. Curcuru, M. Mabuchi, Z. Sun, R. T. Fuchs, E. Schildkraut, P. R. Weigele, W. E. Jack, G. B. Robb, Č. Venclovas, V. Siksnys, A catalogue of biochemically diverse CRISPR-Cas9 orthologs. *Nat. Commun.* **11**, 5512 (2020).
 13. M. Jinek, F. Jiang, D. W. Taylor, S. H. Sternberg, E. Kaya, E. Ma, C. Anders, M. Hauer, K. Zhou, S. Lin, M. Kaplan, A. T. Iavarone, E. Charpentier, E. Nogales, J. A. Doudna, Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science*. **343**, 1247997 (2014).
 14. B. Q. Minh, H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, R. Lanfear, IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
 15. K. S. Makarova, Y. I. Wolf, J. Iranzo, S. A. Shmakov, O. S. Alkhnbashi, S. J. J. Brouns, E. Charpentier, D. Cheng, D. H. Haft, P. Horvath, S. Moineau, F. J. M. Mojica, D. Scott, S. A. Shah, V. Siksnys, M. P. Terns, Č. Venclovas, M. F. White, A. F. Yakunin, W. Yan,

- F. Zhang, R. A. Garrett, R. Backofen, J. van der Oost, R. Barrangou, E. V. Koonin, Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* **18**, 67–83 (2020).
16. K. A. Majorek, S. Dunin-Horkawicz, K. Steczkiewicz, A. Muszewska, M. Nowotny, K. Ginalski, J. M. Bujnicki, The RNase H-like superfamily: new members, comparative structural analysis and evolutionary classification. *Nucleic Acids Res.* **42**, 4160–4179 (2014).
17. H. Nishimasu, F. A. Ran, P. D. Hsu, S. Konermann, S. I. Shehata, N. Dohmae, R. Ishitani, F. Zhang, O. Nureki, Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell.* **156**, 935–949 (2014).
18. W. Bao, J. Jurka, Homologues of bacterial TnpB_IS605 are widespread in diverse eukaryotic transposable elements. *Mob. DNA.* **4**, 12 (2013).
19. J. V. Gomes-Filho, L. S. Zaramela, V. C. da S. Italiani, N. S. Baliga, R. Z. N. Vêncio, T. Koide, Sense overlapping transcripts in IS1341-type transposase genes are functional non-coding RNAs in archaea. *RNA Biol.* **12**, 490–500 (2015).
20. Z. Weinberg, C. E. Lünse, K. A. Corbino, T. D. Ames, J. W. Nelson, A. Roth, K. R. Perkins, M. E. Sherlock, R. R. Breaker, Detection of 224 candidate structured RNAs by comparative analysis of specific subsets of intergenic regions. *Nucleic Acids Res.* **45**, 10811–10823 (2017).
21. A. Hüttenhofer, P. Schattner, The principles of guiding by RNA: chimeric RNA-protein enzymes. *Nat. Rev. Genet.* **7**, 475–482 (2006).
22. A. Schneider, A short history of guide RNAs: The intricate path that led to the discovery of a basic biological concept. *EMBO Rep.* **21**, e51918 (2020).
23. E. V. Koonin, Evolution of RNA- and DNA-guided antiviral defense systems in

- prokaryotes and eukaryotes: common ancestry vs convergence. *Biol. Direct.* **12** (2017).
24. H. Altae-Tran, S. Kannan, F. Zhang, Code and processed data for: The widespread IS200/605 transposon family encodes diverse programmable RNA-guided endonucleases, Version 1.0, Zenodo (2021); <link to be inserted after official acceptance of manuscript>.
 25. C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, T. L. Madden, BLAST+: architecture and applications. *BMC Bioinformatics.* **10**, 421 (2009).
 26. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
 27. M. Steinegger, J. Söding, MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017).
 28. M. Steinegger, M. Meier, M. Mirdita, H. Vöhringer, S. J. Haunsberger, J. Söding, HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinformatics.* **20**, 473 (2019).
 29. S. R. Eddy, Accelerated Profile HMM Searches. *PLoS Comput. Biol.* **7**, e1002195 (2011).
 30. S. A. Shmakov, K. S. Makarova, Y. I. Wolf, K. V. Severinov, E. V. Koonin, Systematic prediction of genes functionally linked to CRISPR-Cas systems by gene neighborhood analysis. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E5307–E5316 (2018).
 31. A. B. Crawley, J. R. Henriksen, R. Barrangou, CRISPRdisco: An Automated Pipeline for the Discovery and Analysis of CRISPR-Cas Systems. *CRISPR J.* **1**, 171–181 (2018).
 32. D. H. Haft, B. J. Loftus, D. L. Richardson, F. Yang, J. A. Eisen, I. T. Paulsen, O. White, TIGRFAMs: a protein family resource for the functional identification of proteins. *Nucleic Acids Res.* **29**, 41–43 (2001).
 33. A. Delcher, Improved microbial gene identification with GLIMMER. *Nucleic Acids Research.* **27** (1999), pp. 4636–4641.

34. S. Naser-Khdour, B. Q. Minh, W. Zhang, E. A. Stone, R. Lanfear, The Prevalence and Impact of Model Violations in Phylogenetic Analysis. *Genome Biol. Evol.* **11**, 3341–3352 (2019).
35. S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. von Haeseler, L. S. Jermiin, ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods.* **14**, 587–589 (2017).
36. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One.* **5**, e9490 (2010).
37. G. Altekar, S. Dwarkadas, J. P. Huelsenbeck, F. Ronquist, Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. *Bioinformatics.* **20**, 407–415 (2004).
38. A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* **30**, 1312–1313 (2014).
39. R. Lorenz, S. H. Bernhart, C. H. zu Siederdissen, H. Tafer, C. Flamm, P. F. Stadler, I. L. Hofacker, ViennaRNA Package 2.0. *Algorithms Mol. Biol.* **6**, 1–14 (2011).
40. E. Rivas, J. Clements, S. R. Eddy, A statistical test for conserved RNA structure shows lack of evidence for structure in lncRNAs. *Nat. Methods.* **14**, 45–48 (2017).
41. E. P. Nawrocki, S. R. Eddy, Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics.* **29**, 2933–2935 (2013).
42. Z. Weinberg, R. R. Breaker, R2R--software to speed the depiction of aesthetic consensus RNA secondary structures. *BMC Bioinformatics.* **12**, 3 (2011).
43. J. Mistry, S. Chuguransky, L. Williams, M. Qureshi, G. A. Salazar, E. L. L. Sonnhammer, S. C. E. Tosatto, L. Paladin, S. Raj, L. J. Richardson, R. D. Finn, A. Bateman, Pfam: The protein families database in 2021. *Nucleic Acids Res.* **49**, D412–D419 (2021).
44. C. Bland, T. L. Ramsey, F. Sabree, M. Lowe, K. Brown, N. C. Kyrpides, P. Hugenholtz, CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly

- interspaced palindromic repeats. *BMC Bioinformatics*. **8**, 209 (2007).
45. F. Asnicar, G. Weingart, T. L. Tickle, C. Huttenhower, N. Segata, Compact graphical representation of phylogenetic data and metadata with GraPhlAn. *PeerJ*. **3**, e1029 (2015).
 46. L. Maaten, G. Hinton, Visualizing Data using t-SNE. *Journal of machine learning research*. **9**, 2579–2605 (2008).
 47. P. G. Poličar, M. Stražar, B. Zupan, Embedding to reference t-SNE space addresses batch effects in single-cell classification. *bioRxiv* (2019), , doi:10.1101/671404.
 48. P. J. Rousseeuw, Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 53–65 (1987).
 49. R. J. G. B. Campello, D. Moulavi, J. Sander, in *Advances in Knowledge Discovery and Data Mining* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2013), *Lecture notes in computer science*, pp. 160–172.
 50. D. H. Parks, M. Imelfort, C. T. Skennerton, P. Hugenholtz, G. W. Tyson, CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).
 51. P.-A. Chaumeil, A. J. Mussig, P. Hugenholtz, D. H. Parks, GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* (2019), doi:10.1093/bioinformatics/btz848.
 52. S. Frey, D. Görlich, A new set of highly efficient, tag-cleaving proteases for purifying recombinant proteins. *J. Chromatogr. A*. **1337**, 95–105 (2014).
 53. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*. **17**, 10–12 (2011).
 54. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods*. **9**, 357–359 (2012).

55. A. M. Sriramachandran, G. Petrosino, M. Méndez-Lago, A. J. Schäfer, L. S. Batista-Nascimento, N. Zilio, H. D. Ulrich, Genome-wide Nucleotide-Resolution Mapping of DNA Replication Patterns, Single-Strand Breaks, and Lesions by GLOE-Seq. *Mol. Cell.* **78**, 975- 985.e7 (2020).
56. K. Clement, H. Rees, M. C. Canver, J. M. Gehrke, R. Farouni, J. Y. Hsu, M. A. Cole, D. R. Liu, J. K. Joung, D. E. Bauer, L. Pinello, CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat. Biotechnol.* **37**, 224–226 (2019).
57. A. Bankevich, S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, V. M. Lesin, S. I. Nikolenko, S. Pham, A. D. Prjibelski, A. V. Pyshkin, A. V. Sirotkin, N. Vyahhi, G. Tesler, M. A. Alekseyev, P. A. Pevzner, SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
58. M. Turmel, C. Otis, C. Lemieux, Divergent copies of the large inverted repeat in the chloroplast genomes of ulvophycean green algae. *Sci. Rep.* **7**, 994 (2017).
59. D. T. Hoang, O. Chernomor, A. von Haeseler, B. Q. Minh, L. S. Vinh, UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).
60. P. J. Lockhart, A. W. Larkum, M. Steel, P. J. Waddell, D. Penny, Evolution of chlorophyll and bacteriochlorophyll: the problem of invariant sites in sequence analysis. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 1930–1934 (1996).
61. A. Criscuolo, S. Gribaldo, BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* **10**, 210 (2010).
62. J. Strecker, A. Ladha, Z. Gardner, J. L. Schmid-Burgk, K. S. Makarova, E. V. Koonin, F. Zhang, RNA-guided DNA insertion with CRISPR-associated transposases. *Science*. **365**, 48–53 (2019).
63. S. E. Klompe, P. L. H. Vo, T. S. Halpin-Healy, S. H. Sternberg, Transposon-encoded CRISPR–Cas systems direct RNA-guided DNA integration. *Nature*. **571**, 219–225

(2019).

64. M. Krupovic, K. S. Makarova, P. Forterre, D. Prangishvili, E. V. Koonin, Casposons: a new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity. *BMC Biol.* **12**, 1–12 (2014).
65. S. M. Crotty, B. Q. Minh, N. G. Bean, B. R. Holland, J. Tuke, L. S. Jermin, A. V. Haeseler, GHOST: Recovering Historical Signal from Heterotachously Evolved Sequence Alignments. *Syst. Biol.* **69**, 249–264 (2020).

[01146] Various modifications and variations of the described methods, pharmaceutical compositions, and kits of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific embodiments, it will be understood that it is capable of further modifications and that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention that are obvious to those skilled in the art are intended to be within the scope of the invention. This application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure come within known customary practice within the art to which the invention pertains and may be applied to the essential features herein before set forth.

CLAIMS

What is claimed is:

1. A non-naturally occurring, engineered composition comprising a) an IscB polypeptide comprising a split Ruv-C nuclease domain comprising RuvC-I, RuvC-II, and RuvC-III subdomains, an HNH domain or both and b) an ω RNA molecule comprising a scaffold and a reprogrammable spacer sequence, the ω RNA molecule capable of forming a complex with the IscB polypeptide and directing the IscB polypeptide to a target polynucleotide.
2. The composition of claim 1, wherein the IscB polypeptide comprises a PLMP domain and optionally a conserved C-terminal Y domain.
3. The composition of claims 1 or 2, wherein the engineered IscB polypeptide comprises a HNH domain but no RuvC-I, RuvC-II, and RuvC-III subdomains.
4. The composition of claim 1, wherein the HNH domain is located between RuvC-II and RuvC-III subdomains.
5. The composition of claims 1 or 2, wherein the engineered IscB polypeptide comprises a RuvC-I, RuvC-II, and RuvC-III subdomains but no HNH domain.
6. The composition of claim 1, wherein the IscB polypeptide comprises about 170 to about 1000 amino acids.
7. The composition of claim 1, wherein the reprogrammable spacer sequence comprises a spacer of 10 nucleotides to 150 nucleotides in length, preferably 12 to 50 nt, more preferably 15 and 45 nt in length.
8. The composition of any of the previous claims, wherein the target sequence comprises a target adjacent motif (TAM) sequence 3' of the target polynucleotide.

9. The composition of any of the previous claims, wherein the target polynucleotide is DNA.
10. The composition of any of the previous claims wherein the ω RNA further comprises an aptamer.
11. The composition of any of the previous claims wherein the ω RNA molecule further comprises an extension to add an RNA template.
12. The composition of any of the previous claims, further comprising a functional domain associated with the IscB protein.
13. The composition of claim 12, wherein the functional domain has transposase activity, methylase activity, demethylase activity, translation activation activity, translation repression activity, transcription activation activity, transcription repression activity, transcription release factor activity, chromatin modifying or remodeling activity, histone modification activity, nuclease activity, single-strand RNA cleavage activity, double-strand RNA cleavage activity, single-strand DNA cleavage activity, double-strand DNA cleavage activity, nucleic acid binding activity, detectable activity, or any combination thereof.
14. The composition of claim 1, further comprising a serine or tyrosine recombinase.
15. The composition of any one of the preceding claims, further comprising a homologous recombination donor template comprising a donor sequence for insertion into a target polynucleotide.
16. A vector system comprising one or more vectors encoding the Isc polypeptide and the ω RNA molecule of claim 1.
17. An engineered cell comprising the composition of claim 1.

18. A method of modifying a target polynucleotide sequence in a cell, comprising introducing to the cell the composition of any of claims 1 to 15.
19. The method of claim 18, wherein the polypeptide and/or nucleic acid components are provided via one or more polynucleotides encoding the polypeptides and/or nucleic acid component(s), and wherein the one or more polynucleotides are operably configured to express the IscB polypeptide and/or the ω RNA molecule.
20. The method of claim 19, wherein the modifying comprises cleaving a DNA polynucleotide.
21. The method of any of claims 18-20, wherein the cleaving results in 5' overhangs.
22. An engineered, non-naturally occurring composition comprising an IscB protein, wherein the IscB protein comprises an N-terminal X domain, a RuvC domain, a Bridge Helix domain, and a C-terminal Y domain.
23. The composition of claim 22, wherein the X domain has an amino acid sequence that share at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with X domains in Table 2.
24. The composition of claim 22, wherein the Y domain has an amino acid sequence that share at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with Y domains in Table 2.

25. The composition of claim 22, wherein the IscB protein shares at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with a IscB protein selected from Tables 2 and 3.
26. The composition of claim 22, wherein the X domain is no more than 50 amino acids in length.
27. The composition of claim 22, wherein the IscB protein further comprises an HNH domain.
28. The composition of claim 27, wherein the RuvC domain comprises a RuvC I subdomain, a Ruv II subdomain and a Ruv III subdomain, and the HNH is located between the Ruv C II and RuvC III subdomains of the RuvC domain.
29. The composition of claim 22, wherein the IscB protein is no more than 500, no more than 600, no more than 700, or no more than 800 amino acids in length.
30. The composition of any one of claims 22 to 29 further comprising a first and second nucleic acid molecules, the first and second nucleic acid molecules capable of forming a duplex, the duplex capable of forming a complex with the IscB protein, wherein the second nucleic acid molecule is a recombinant molecule comprising a heterologous CRISPR-associated guide sequence capable of directing site-specific binding of the complex to a target sequence of a target polynucleotide.

31. The composition of anyone of claims 22 to 29, wherein comprising a CRISPR-associated single guide molecule capable of forming a complex with the IscB protein and directing site-specific binding of the complex to a target sequence of a target polynucleotide.
32. The composition of anyone of claims 22 to 29, wherein the IscB protein targets DNA.
33. The composition of any one of the proceeding claims, wherein the nuclease domains of the IscB protein are catalytically inactive.
34. The composition of claim 33, wherein the nuclease domain has nickase activity or is engineered to have nickase activity.
35. The composition of claim 33 or 34, further comprising a functional domain associated with the IscB protein.
36. The composition of claim 35, wherein the functional domain has transposase activity, methylase activity, demethylase activity, translation activation activity, translation repression activity, transcription activation activity, transcription repression activity, transcription release factor activity, chromatin modifying or remodeling activity, histone modification activity, nuclease activity, single-strand RNA cleavage activity, double-strand RNA cleavage activity, single-strand DNA cleavage activity, double-strand DNA cleavage activity, nucleic acid binding activity, detectable activity, or any combination thereof.

37. The composition of any one of claims 22-36, further comprising a homologous recombination donor template comprising a donor sequence for insertion into a target polynucleotide.
38. The composition of any one of claims 22-36, the target sequence comprises a PAM of NAC, where N is A, C, G, or T.
39. One or more polynucleotides encoding one or more components of the composition of any of claims 22-38.
40. One or more vectors comprising the one or more polynucleotides of claim 39.
41. A cell or progeny thereof genetically engineered to express one or more components of the compositions of any one of claims 22 to 38.
42. A method of targeting a polynucleotide, comprising contacting a sample that comprises a target polynucleotide with the composition of any one of claims 22 to 38, or the one or more polynucleotides or one or more vectors of claim 39 or 40.
43. The method of claim 42, wherein contacting results in modification of a gene product or modification of the amount or expression of a gene product.

44. The method of claim 43, wherein the target sequence of the polynucleotide is a disease-associated target sequence.
45. An engineered, non-naturally occurring composition comprising:
- a. the IscB protein of any one of claims 22 to 29, wherein the IscB protein is catalytically inactive,
 - b. a nucleotide deaminase associated with or otherwise capable of forming a complex with the IscB protein, and
 - c. a CRISPR-associated single guide molecule capable of forming a complex with the IscB protein and directing site-specific binding at a target sequence.
46. The composition of claim 45, wherein the nucleotide deaminase is an adenosine deaminase or a cytidine deaminase.
47. One or more polynucleotides encoding one or more components of the composition of any one of claims 45 or 46.
48. One or more vectors encoding the one or more polynucleotides of claim 47.
49. A cell or progeny thereof genetically engineered to express one or more components of the composition of any one of claims 45 or 46.
50. A method of editing nucleic acids in target polynucleotides comprising delivering the composition of claim 45 or 46, the one or more polynucleotides of claim 47, or one or

more vectors of claim 48 to a cell or population of cells comprising the target polynucleotides.

51. The method of claim 50, wherein the target polynucleotides are target sequences within genomic DNA.
52. The method of claim 50 or 51, wherein the target polynucleotide is edited at one or more bases to introduce a G→A or C→T mutation.
53. An isolated cell or progeny thereof comprising one or more base edits made using the method of any one of claims 50 to 52.
54. An engineered, non-naturally occurring composition comprising:
 - a. the IscB protein of any one of claims 22 to 29, wherein the IscB is catalytically inactive,
 - b. a reverse transcriptase associated with or otherwise capable of forming a complex with the IscB protein, and
 - c. a CRISPR-associated guide molecule capable of forming a complex with the IscB protein and directing site-specific binding of the complex to a target sequence of a target polynucleotide, the guide molecule further comprising a donor sequence for insertion into the target polynucleotide.
55. One or more polynucleotides encoding one or more components of the composition of claim 54.

56. One or more vectors encoding the one or more polynucleotides of claim 55.

57. A method of modifying target polynucleotides comprising

delivering the composition of claim 54, the one or more polynucleotides of claim 55, or one or more vectors of claim 56 to a cell or population of cells comprising the target polynucleotides, wherein the complex directs the reverse transcriptase to the target sequence and the reverse transcriptase facilitates insertion of the donor sequence from the CRISPR-associated guide molecule into the target polynucleotide.

58. The method of claim 57, wherein insertion of the donor sequence:

- a. introduces one or more base edits;
- b. corrects or introduces a premature stop codon;
- c. disrupts a splice site;
- d. inserts or restores a splice site;
- e. inserts a gene or gene fragment at one or both alleles of the target polynucleotide; or;
- f. a combination thereof.

59. An isolated cell or progeny thereof comprising the modifications made using the method of claim 57 or 58.

60. An engineered, non-naturally occurring composition comprising:

- a. the IscB protein of any one of claims 22 to 29,

- b. a non-LTR retrotransposon protein associated with or otherwise capable of forming a complex with the IscB protein;
 - c. a CRISPR-associated single guide molecule capable of forming a complex with the IscB protein and directing site-specific binding to a target sequence of a target polynucleotide; and
 - d. a donor construct comprising a donor polynucleotide for insertion to the target polynucleotide and located between two binding elements capable of forming a complex with the non-LTR retrotransposon protein.
61. The composition of claim 60, wherein the IscB protein is fused to the N-terminus of the non-LTR retrotransposon protein.
62. The composition of claim 60 or 61, wherein the IscB protein is engineered to have nickase activity.
63. The composition of claim 60, wherein the CRISPR-associated guides direct the fusion protein to a target sequence 5' of the targeted insertion site, and wherein the IscB protein generates a double-strand break at the targeted insertion site.
64. The composition of claim 60, wherein the CRISPR-associated guides direct the fusion protein to a target sequence 3' of the targeted insertion site, and wherein the IscB protein generates a double-strand break at the targeted insertion site.

65. The composition of claim 60, wherein the donor polynucleotide further comprises a polymerase processing element to facilitate 3' end processing of the donor polynucleotide sequence.
66. The composition of claim 60, wherein the donor polynucleotide further comprises a homology region to the target sequence on the 5' end of the donor construct, the 3' end of the donor construct, or both.
67. The composition of claim 66, wherein the homology region is from 8 to 25 base pairs.
68. One or more polynucleotides encoding one or more components of the composition of any one of claims 60 to 67.
69. One or more vectors comprising the one or more polynucleotides of claim 68.
70. A method of modifying target polynucleotides comprising
delivering the composition of any one of claims 60 to 67, the one or more polynucleotides of claim 68, or one or more vectors of claim 69 to a cell or population of cells comprising the target polynucleotides, wherein the complex directs the non-LTR retrotransposon protein to the target sequence and the non-LTR retrotransposon protein facilitates insertion of the donor polynucleotide sequence from the donor construct into the target polynucleotide.
71. The method of claim 70, wherein insertion of the donor sequence:

- a. introduces one or more base edits;
- b. corrects or introduces a premature stop codon;
- c. disrupts a splice site;
- d. inserts or restores a splice site;
- e. inserts a gene or gene fragment at one or both alleles of the target polynucleotide; or;
- f. a combination thereof.

72. An isolated cell or progeny thereof comprising the modifications made using the method of claim 70 or 71.

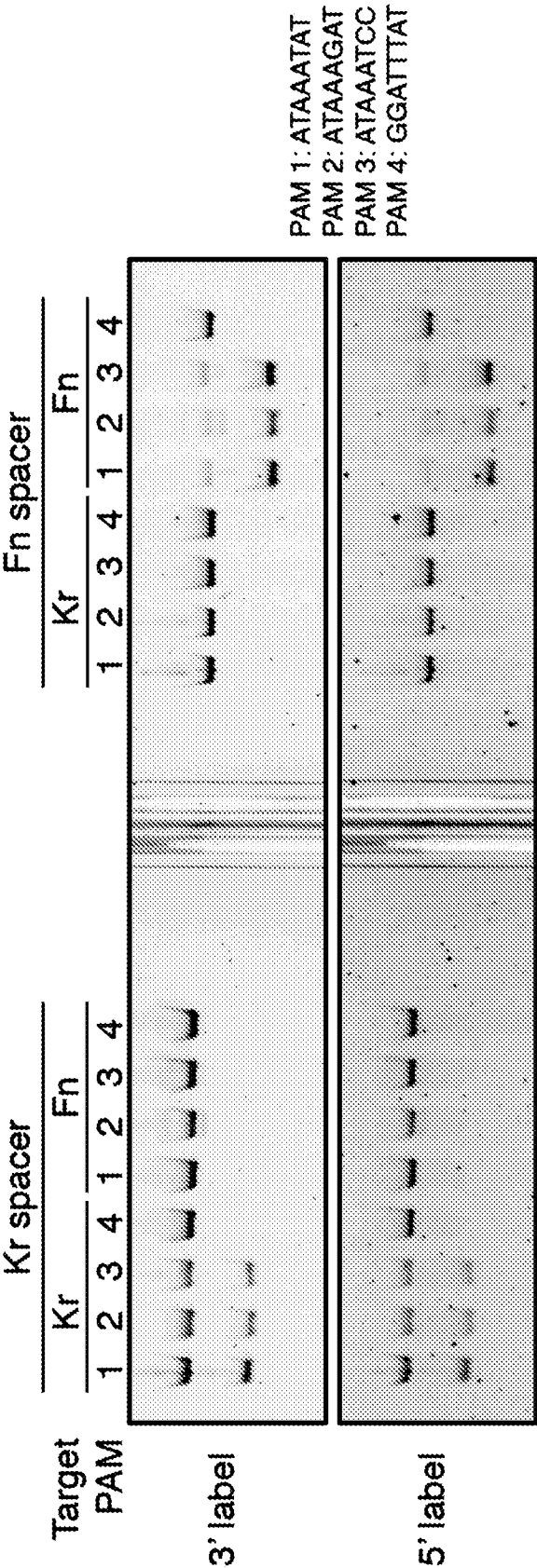


FIG. 1

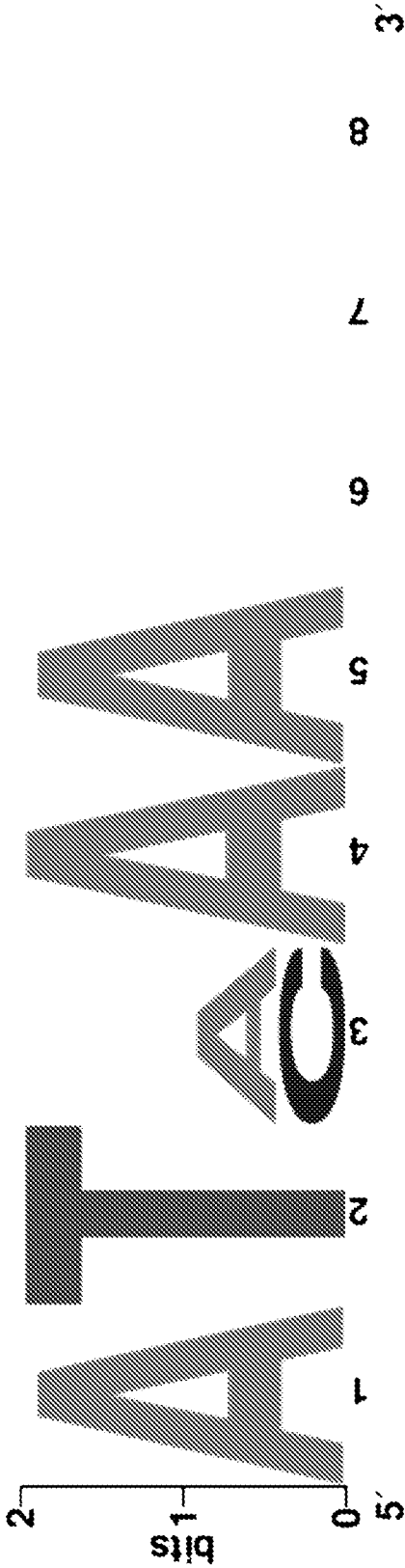


FIG. 2

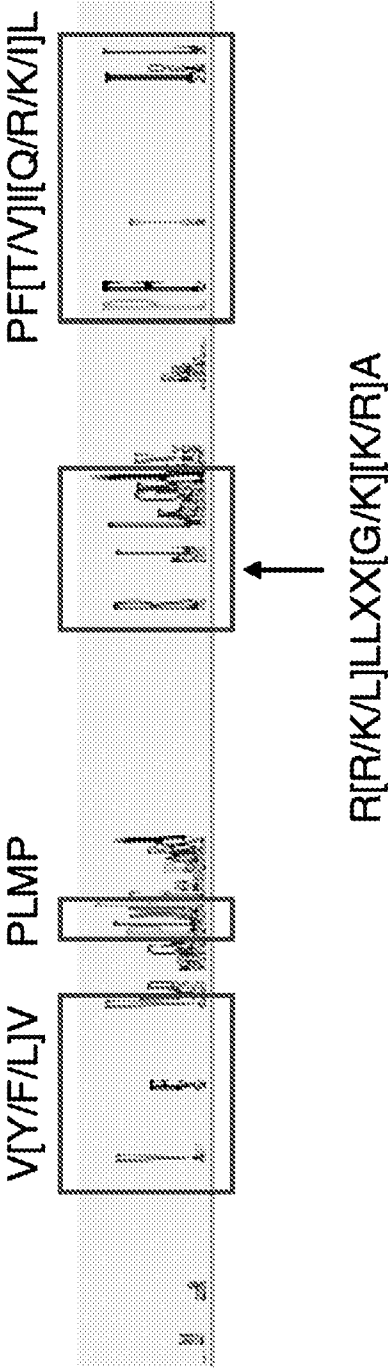


FIG. 3

[illegible]

FIG. 4-1

FIG. 4-2

[illegible]

FIG. 4-3

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80																				

FIG. 4-4

[illegible]

Consensus	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015	1016	1017	1018	1019	1020	1021	1022	1023	1024	1025	1026	1027	1028	1029	1030	1031	1032	1033	1034	1035	1036	1037	1038	1039	1040	1041	1042	1043	1044	1045	1046	1047	1048	1049	1050	1051	1052	1053	1054	1055	1056	1057	1058	1059	1060	1061	1062	1063	1064	1065	1066	1067	1068	1069	1070	1071	1072	1073	1074	1075	1076	1077	1078	1079	1080	1081	1082	1083	1084	1085	1086	1087	1088	1089	1090	1091	1092	1093	1094	1095	1096	1097	1098	1099	1100	1101	1102	1103	1104	1105	1106	1107	1108	1109	1110	1111	1112	1113	1114	1115	1116	1117	1118	1119	1120	1121	1122	1123	1124	1125	1126	1127	1128	1129	1130	1131	1132	1133	1134	1135	1136	1137	1138	1139	1140	1141	1142	1143	1144	1145	1146	1147	1148	1149	1150	1151	1152	1153	1154	1155	1156	1157	1158	1159	1160	1161	1162	1163	1164	1165	1166	1167	1168	1169	1170	1171	1172	1173	1174	1175	1176	1177	1178	1179	1180	1181	1182	1183	1184	1185	1186	1187	1188	1189	1190	1191	1192	1193	1194	1195	1196	1197	1198	1199	1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	1210	1211	1212	1213	1214	1215	1216	1217	1218	1219	1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233	1234	1235	1236	1237	1238	1239	1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254	1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296	1297	1298	1299	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	1479	1480	1481
-----------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

Chromosome	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015	1016	1017	1018	1019	1020	1021	1022	1023	1024	1025	1026	1027	1028	1029	1030	1031	1032	1033	1034	1035	1036	1037	1038	1039	1040	1041	1042	1043	1044	1045	1046	1047	1048	1049	1050	1051	1052	1053	1054	1055	1056	1057	1058	1059	1060	1061	1062	1063	1064	1065	1066	1067	1068	1069	1070	1071	1072	1073	1074	1075	1076	1077	1078	1079	1080	1081	1082	1083	1084	1085	1086	1087	1088	1089	1090	1091	1092	1093	1094	1095	1096	1097	1098	1099	1100	1101	1102	1103	1104	1105	1106	1107	1108	1109	1110	1111	1112	1113	1114	1115	1116	1117	1118	1119	1120	1121	1122	1123	1124	1125	1126	1127	1128	1129	1130	1131	1132	1133	1134	1135	1136	1137	1138	1139	1140	1141	1142	1143	1144	1145	1146	1147	1148	1149	1150	1151	1152	1153	1154	1155	1156	1157	1158	1159	1160	1161	1162	1163	1164	1165	1166	1167	1168	1169	1170	1171	1172	1173	1174	1175	1176	1177	1178	1179	1180	1181	1182	1183	1184	1185	1186	1187	1188	1189	1190	1191	1192	1193	1194	1195	1196	1197	1198	1199	1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	1210	1211	1212	1213	1214	1215	1216	1217	1218	1219	1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233	1234	1235	1236	1237	1238	1239	1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254	1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296	1297	1298	1299	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	1479
------------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

[illegible]

FIG. 4-8

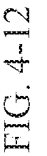
[illegible]

FIG. 4.9

Consensus	88	R
68		
113		
163		
166		
181		
186		
197		
71		
146		
14		
198		
199		
123		
198		
179		
150		
123		
23		
150		
189		
213		
116		
101		
107		
185		
188		
187		
38		
182		
109		
169		
92		
194		
117		
103		
120		
103		
213		
163		
163		
75		
107		
90		
194		
115		
199		
109		
157		
105		
105		
109		
180		
108		
23		
116		
110		
108		
118		
103		
143		
196		
202		
123		
74		

FIG. 4-10

Consensus	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015	1016	1017	1018	1019	1020	1021	1022	1023	1024	1025	1026	1027	1028	1029	1030	1031	1032	1033	1034	1035	1036	1037	1038	1039	1040	1041	1042	1043	1044	1045	1046	1047	1048	1049	1050	1051	1052	1053	1054	1055	1056	1057	1058	1059	1060	1061	1062	1063	1064	1065	1066	1067	1068	1069	1070	1071	1072	1073	1074	1075	1076	1077	1078	1079	1080	1081	1082	1083	1084	1085	1086	1087	1088	1089	1090	1091	1092	1093	1094	1095	1096	1097	1098	1099	1100	1101	1102	1103	1104	1105	1106	1107	1108	1109	1110	1111	1112	1113	1114	1115	1116	1117	1118	1119	1120	1121	1122	1123	1124	1125	1126	1127	1128	1129	1130	1131	1132	1133	1134	1135	1136	1137	1138	1139	1140	1141	1142	1143	1144	1145	1146	1147	1148	1149	1150	1151	1152	1153	1154	1155	1156	1157	1158	1159	1160	1161	1162	1163	1164	1165	1166	1167	1168	1169	1170	1171	1172	1173	1174	1175	1176	1177	1178	1179	1180	1181	1182	1183	1184	1185	1186	1187	1188	1189	1190	1191	1192	1193	1194	1195	1196	1197	1198	1199	1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	1210	1211	1212	1213	1214	1215	1216	1217	1218	1219	1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233	1234	1235	1236	1237	1238	1239	1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254	1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296	1297	1298	1299	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	147
-----------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----



Chromatids	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015	1016	1017	1018	1019	1020	1021	1022	1023	1024	1025	1026	1027	1028	1029	1030	1031	1032	1033	1034	1035	1036	1037	1038	1039	1040	1041	1042	1043	1044	1045	1046	1047	1048	1049	1050	1051	1052	1053	1054	1055	1056	1057	1058	1059	1060	1061	1062	1063	1064	1065	1066	1067	1068	1069	1070	1071	1072	1073	1074	1075	1076	1077	1078	1079	1080	1081	1082	1083	1084	1085	1086	1087	1088	1089	1090	1091	1092	1093	1094	1095	1096	1097	1098	1099	1100	1101	1102	1103	1104	1105	1106	1107	1108	1109	1110	1111	1112	1113	1114	1115	1116	1117	1118	1119	1120	1121	1122	1123	1124	1125	1126	1127	1128	1129	1130	1131	1132	1133	1134	1135	1136	1137	1138	1139	1140	1141	1142	1143	1144	1145	1146	1147	1148	1149	1150	1151	1152	1153	1154	1155	1156	1157	1158	1159	1160	1161	1162	1163	1164	1165	1166	1167	1168	1169	1170	1171	1172	1173	1174	1175	1176	1177	1178	1179	1180	1181	1182	1183	1184	1185	1186	1187	1188	1189	1190	1191	1192	1193	1194	1195	1196	1197	1198	1199	1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	1210	1211	1212	1213	1214	1215	1216	1217	1218	1219	1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233	1234	1235	1236	1237	1238	1239	1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254	1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296	1297	1298	1299	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	1
------------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	---

Consensus	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015	1016	1017	1018	1019	1020	1021	1022	1023	1024	1025	1026	1027	1028	1029	1030	1031	1032	1033	1034	1035	1036	1037	1038	1039	1040	1041	1042	1043	1044	1045	1046	1047	1048	1049	1050	1051	1052	1053	1054	1055	1056	1057	1058	1059	1060	1061	1062	1063	1064	1065	1066	1067	1068	1069	1070	1071	1072	1073	1074	1075	1076	1077	1078	1079	1080	1081	1082	1083	1084	1085	1086	1087	1088	1089	1090	1091	1092	1093	1094	1095	1096	1097	1098	1099	1100	1101	1102	1103	1104	1105	1106	1107	1108	1109	1110	1111	1112	1113	1114	1115	1116	1117	1118	1119	1120	1121	1122	1123	1124	1125	1126	1127	1128	1129	1130	1131	1132	1133	1134	1135	1136	1137	1138	1139	1140	1141	1142	1143	1144	1145	1146	1147	1148	1149	1150	1151	1152	1153	1154	1155	1156	1157	1158	1159	1160	1161	1162	1163	1164	1165	1166	1167	1168	1169	1170	1171	1172	1173	1174	1175	1176	1177	1178	1179	1180	1181	1182	1183	1184	1185	1186	1187	1188	1189	1190	1191	1192	1193	1194	1195	1196	1197	1198	1199	1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	1210	1211	1212	1213	1214	1215	1216	1217	1218	1219	1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233	1234	1235	1236	1237	1238	1239	1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254	1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296	1297	1298	1299	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	147
-----------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----

Accession	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015	1016	1017	1018	1019	1020	1021	1022	1023	1024	1025	1026	1027	1028	1029	1030	1031	1032	1033	1034	1035	1036	1037	1038	1039	1040	1041	1042	1043	1044	1045	1046	1047	1048	1049	1050	1051	1052	1053	1054	1055	1056	1057	1058	1059	1060	1061	1062	1063	1064	1065	1066	1067	1068	1069	1070	1071	1072	1073	1074	1075	1076	1077	1078	1079	1080	1081	1082	1083	1084	1085	1086	1087	1088	1089	1090	1091	1092	1093	1094	1095	1096	1097	1098	1099	1100	1101	1102	1103	1104	1105	1106	1107	1108	1109	1110	1111	1112	1113	1114	1115	1116	1117	1118	1119	1120	1121	1122	1123	1124	1125	1126	1127	1128	1129	1130	1131	1132	1133	1134	1135	1136	1137	1138	1139	1140	1141	1142	1143	1144	1145	1146	1147	1148	1149	1150	1151	1152	1153	1154	1155	1156	1157	1158	1159	1160	1161	1162	1163	1164	1165	1166	1167	1168	1169	1170	1171	1172	1173	1174	1175	1176	1177	1178	1179	1180	1181	1182	1183	1184	1185	1186	1187	1188	1189	1190	1191	1192	1193	1194	1195	1196	1197	1198	1199	1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	1210	1211	1212	1213	1214	1215	1216	1217	1218	1219	1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233	1234	1235	1236	1237	1238	1239	1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254	1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296	1297	1298	1299	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	1479	1480	1481</
-----------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--------

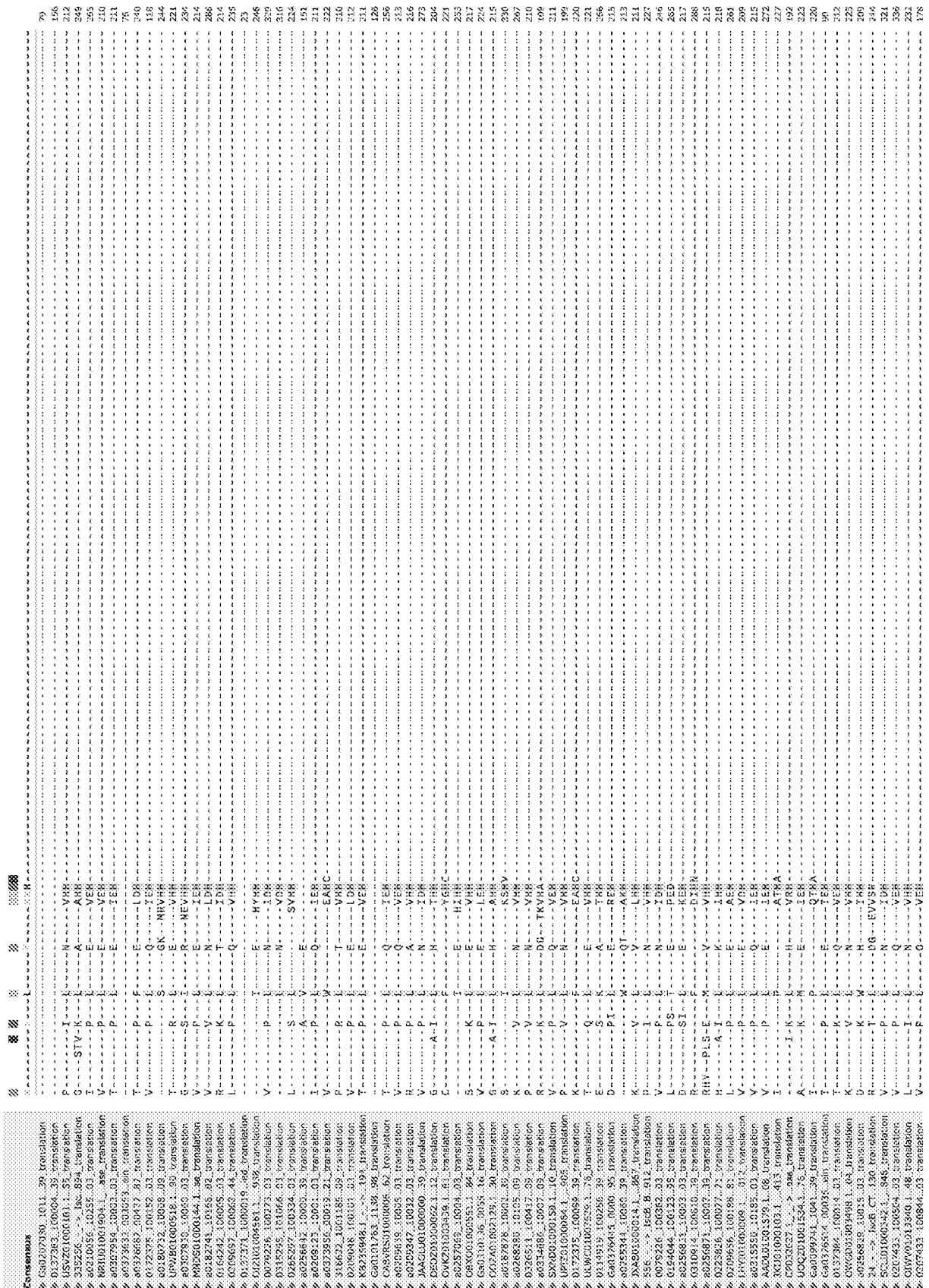


FIG. 4-16

Conserved	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015	1016	1017	1018	1019	1020	1021	1022	1023	1024	1025	1026	1027	1028	1029	1030	1031	1032	1033	1034	1035	1036	1037	1038	1039	1040	1041	1042	1043	1044	1045	1046	1047	1048	1049	1050	1051	1052	1053	1054	1055	1056	1057	1058	1059	1060	1061	1062	1063	1064	1065	1066	1067	1068	1069	1070	1071	1072	1073	1074	1075	1076	1077	1078	1079	1080	1081	1082	1083	1084	1085	1086	1087	1088	1089	1090	1091	1092	1093	1094	1095	1096	1097	1098	1099	1100	1101	1102	1103	1104	1105	1106	1107	1108	1109	1110	1111	1112	1113	1114	1115	1116	1117	1118	1119	1120	1121	1122	1123	1124	1125	1126	1127	1128	1129	1130	1131	1132	1133	1134	1135	1136	1137	1138	1139	1140	1141	1142	1143	1144	1145	1146	1147	1148	1149	1150	1151	1152	1153	1154	1155	1156	1157	1158	1159	1160	1161	1162	1163	1164	1165	1166	1167	1168	1169	1170	1171	1172	1173	1174	1175	1176	1177	1178	1179	1180	1181	1182	1183	1184	1185	1186	1187	1188	1189	1190	1191	1192	1193	1194	1195	1196	1197	1198	1199	1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	1210	1211	1212	1213	1214	1215	1216	1217	1218	1219	1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233	1234	1235	1236	1237	1238	1239	1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254	1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296	1297	1298	1299	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	1479	1480	1481</
-----------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--------

FIG. 4-18[illegible]



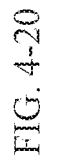


FIG. 4-21

Consensus	83
6-0207040.1011.38, translation	83
6-017281.100604.39, translation	108
6-017281.100604.39, translation	325
6-017281.100604.39, translation	372
6-017281.100604.39, translation	364
6-017281.100604.39, translation	340
6-017281.100604.39, translation	277
6-017281.100604.39, translation	101
6-017281.100604.39, translation	365
6-017281.100604.39, translation	244
6-017281.100604.39, translation	377
6-017281.100604.39, translation	365
6-017281.100604.39, translation	346
6-017281.100604.39, translation	416
6-017281.100604.39, translation	345
6-017281.100604.39, translation	279
6-017281.100604.39, translation	31
6-017281.100604.39, translation	377
6-017281.100604.39, translation	388
6-017281.100604.39, translation	347
6-017281.100604.39, translation	343
6-017281.100604.39, translation	359
6-017281.100604.39, translation	344
6-017281.100604.39, translation	324
6-017281.100604.39, translation	360
6-017281.100604.39, translation	324
6-017281.100604.39, translation	393
6-017281.100604.39, translation	297
6-017281.100604.39, translation	266
6-017281.100604.39, translation	395
6-017281.100604.39, translation	335
6-017281.100604.39, translation	395
6-017281.100604.39, translation	338
6-017281.100604.39, translation	378
6-017281.100604.39, translation	381
6-017281.100604.39, translation	328
6-017281.100604.39, translation	135
6-017281.100604.39, translation	377
6-017281.100604.39, translation	317
6-017281.100604.39, translation	327
6-017281.100604.39, translation	347
6-017281.100604.39, translation	383
6-017281.100604.39, translation	358
6-017281.100604.39, translation	304
6-017281.100604.39, translation	326
6-017281.100604.39, translation	347
6-017281.100604.39, translation	285
6-017281.100604.39, translation	377
6-017281.100604.39, translation	360
6-017281.100604.39, translation	326
6-017281.100604.39, translation	345
6-017281.100604.39, translation	391
6-017281.100604.39, translation	328
6-017281.100604.39, translation	341
6-017281.100604.39, translation	357
6-017281.100604.39, translation	310
6-017281.100604.39, translation	359
6-017281.100604.39, translation	283
6-017281.100604.39, translation	222
6-017281.100604.39, translation	328
6-017281.100604.39, translation	314
6-017281.100604.39, translation	367
6-017281.100604.39, translation	316
6-017281.100604.39, translation	316

FIG. 4-22

FIG. 4-23

[illegible]

FIG. 4-24

FIG. 4-25

Consensus	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015	1016	1017	1018	1019	1020	1021	1022	1023	1024	1025	1026	1027	1028	1029	1030	1031	1032	1033	1034	1035	1036	1037	1038	1039	1040	1041	1042	1043	1044	1045	1046	1047	1048	1049	1050	1051	1052	1053	1054	1055	1056	1057	1058	1059	1060	1061	1062	1063	1064	1065	1066	1067	1068	1069	1070	1071	1072	1073	1074	1075	1076	1077	1078	1079	1080	1081	1082	1083	1084	1085	1086	1087	1088	1089	1090	1091	1092	1093	1094	1095	1096	1097	1098	1099	1100	1101	1102	1103	1104	1105	1106	1107	1108	1109	1110	1111	1112	1113	1114	1115	1116	1117	1118	1119	1120	1121	1122	1123	1124	1125	1126	1127	1128	1129	1130	1131	1132	1133	1134	1135	1136	1137	1138	1139	1140	1141	1142	1143	1144	1145	1146	1147	1148	1149	1150	1151	1152	1153	1154	1155	1156	1157	1158	1159	1160	1161	1162	1163	1164	1165	1166	1167	1168	1169	1170	1171	1172	1173	1174	1175	1176	1177	1178	1179	1180	1181	1182	1183	1184	1185	1186	1187	1188	1189	1190	1191	1192	1193	1194	1195	1196	1197	1198	1199	1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	1210	1211	1212	1213	1214	1215	1216	1217	1218	1219	1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233	1234	1235	1236	1237	1238	1239	1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254	1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296	1297	1298	1299	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	147
-----------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----

[illegible]

FIG. 4-27

FIG. 4-28

Consensus	127
> GAG207010.1311.39. translation	254
> G137383.100004.38. translation	355
> G51201000181.1.45. translation	425
> G33256...>_sec.099. translation	418
> G0210056.10255.03. translation	384
> G8101001804.1.46. translation	277
> G070954.10003.03. translation	145
> G073633.005083.en. translation	414
> G073687.00427.87. translation	379
> G172751.100152.03. translation	451
> G6180732.100096.09. translation	387
> G8480100518.1.90. translation	427
> G0307930.10760.03. translation	380
> G80531.000144.1.84. translation	462
> G0182743.10166.82. translation	384
> G144242.100005.02. translation	282
> G294697.100662.46. translation	72
> G173751.100019.66. translation	477
> G1320104953.1.938. translation	503
> G079258.100279.03. translation	389
> G315293.101066.03. translation	425
> G352297.100324.03. translation	453
> G0256843.10000.35. translation	387
> G0204223.10001.03. translation	457
> G0275562.10018.21. translation	344
> G18622.1001165.03. translation	388
> G2328048.100101.184. translation	322
> G2431753.1138.38. translation	450
> G2897632.100704.53. translation	460
> G0208636.10005.03. translation	372
> G0208347.10732.03. translation	443
> G46G010209070.30. translation	387
> G80231002963.1.12. translation	477
> G4241001493.81. translation	394
> G0257069.10008.03. translation	414
> G85C01000551.1.84. translation	415
> G05310136.0055.15. translation	381
> G02A01007129.1.30. translation	367
> G0268280.10165.09. translation	341
> G326531.100417.09. translation	361
> G0334866.10007.09. translation	258
> G540201003158.1.10. translation	295
> G4201000564.1.905. translation	361
> G172375.100049.39. translation	417
> G14001007529.1.76. translation	402
> G114919.100246.30. translation	234
> G80376415.0000.95. translation	402
> G0355344.10060.34. translation	288
> G480100014.1.867. translation	285
> G56...>_sec.8.312. translation	421
> G079224.100382.03. translation	404
> G193084.100123.95. translation	364
> G0256631.10002.02. translation	422
> G310914.100610.39. translation	418
> G0256671.10007.39. translation	435
> G272826.100077.71. translation	362
> G209656.100188.81. translation	385
> G401000099.1.012. translation	425
> G0315580.10165.03. translation	475
> G4010001779.1.08. translation	349
> G1001000153.1.413. translation	408
> G032627.1...>_sec. translation	315
> G20201001524.1.76. translation	339
> G2009941.1041.30. translation	362
> G0375514.10033.en. translation	304
> G177384.100014.03. translation	278
> G6G01003498.1.34. translation	276
> G0256830.10015.03. translation	301
> G...>_sec.133. translation	511
> G2181000343.1.646. translation	334
> G207433.100598.03. translation	411
> G3940101340.1.48. translation	
> G2076733.100643.03. translation	

FIG. 4-29

Consensus	R	125
> GAG207030.1311.39. translation	R	125
> G137383.100004.38. translation	R	255
> U517201000181.1.45. translation	R	335
> 333236...>_acc.699. translation	F	425
> A0210056.10755.03. translation	J	419
> N8101001804.1.44. translation	A1	365
> A0307954.10003.03. translation	R	277
> A077633.003052.en. translation	R	146
> A077667.00427.87. translation	Y	415
> 317275.100152.03. translation	F	374
> A6180732.10006.09. translation	F	452
> 0P480100518.1.90. translation	L	388
> A0307936.10760.03. translation	V	458
> 080551000144.1.88. translation	B	381
> A057671.10766.82.1. translation	B	463
> 0443682.100003.03. translation	R	395
> 0260697.100662.01. translation	R	282
> 0117371.100016.06. translation	R	71
> 0120104655.1.338. translation	A	479
> 0079228.100273.03. translation	T	504
> 0315293.101066.03. translation	T	380
> 0305207.100324.03. translation	W	425
> A0258843.10006.75. translation	K	284
> A0304023.10001.03. translation	R	389
> A0257952.100018.21. translation	R	527
> 310622.1001168.03. translation	K	544
> 02120648.100101.03. translation	B	383
> 02120648.100101.03. translation	B	323
> 02120648.100101.03. translation	R	452
> 02120648.100101.03. translation	R	462
> A0208036.100005.03. translation	N	373
> A0208036.100005.03. translation	R	444
> A0208036.100005.03. translation	R	397
> A0208036.100005.03. translation	S	396
> A0208036.100005.03. translation	S	428
> A0208036.100005.03. translation	B	371
> A0208036.100005.03. translation	B	395
> A0208036.100005.03. translation	C	415
> A0208036.100005.03. translation	Y	417
> A0208036.100005.03. translation	R	381
> A0208036.100005.03. translation	V	142
> A0208036.100005.03. translation	R	382
> A0208036.100005.03. translation	R	359
> A0208036.100005.03. translation	J	295
> A0208036.100005.03. translation	M	418
> A0208036.100005.03. translation	R	403
> A0208036.100005.03. translation	T	234
> A0208036.100005.03. translation	R	403
> A0208036.100005.03. translation	R	387
> A0208036.100005.03. translation	R	285
> A0208036.100005.03. translation	K	422
> A0208036.100005.03. translation	K	405
> A0208036.100005.03. translation	T	364
> A0208036.100005.03. translation	T	423
> A0208036.100005.03. translation	R	419
> A0208036.100005.03. translation	S	436
> A0208036.100005.03. translation	R	363
> A0208036.100005.03. translation	M	426
> A0208036.100005.03. translation	R	561
> A0208036.100005.03. translation	R	350
> A0208036.100005.03. translation	R	409
> A0208036.100005.03. translation	D	216
> A0208036.100005.03. translation	D	239
> A0208036.100005.03. translation	R	383
> A0208036.100005.03. translation	R	365
> A0208036.100005.03. translation	F	273
> A0208036.100005.03. translation	Y	277
> A0208036.100005.03. translation	R	502
> A0208036.100005.03. translation	R	512
> A0208036.100005.03. translation	R	334
> A0208036.100005.03. translation	R	412

FIG. 4-30

FIG. 4-31

Consensus

GC207030,10311,39, translation
U137383,100064,39, translation
U85720,1000181,25, translation
335256, ...isc,109, translation
a02110056,100259,65, translation
M61101,1001804,1,ase, translation
GC007954,1000033,60, translation
a0372633,605952,69, translation
a0376687,606247,87, translation
U172795,1001552,63, translation
U09010403181,1,90, translation
GC0070700,100160,61, translation
M60527,1001144,1,ase, translation
U015274,100166,82, translation
U0386697,1000662,82, translation
U137371,1000018,62, translation
U03101045551,1,639, translation
U0920236,100272,63, translation
U031295,1001062,63, translation
G565297,100234,63, translation
G565298,1000139,39, translation
a0209123,100011,63, translation
a0371956,1001185,63, translation
U16682,1001185,63, translation
U0209494,100101,63, translation
K42329468,1, ...194, translation
GC031753,11338,30, translation
CAE001050,1000006,51, translation
GC009630,10005,61, translation
a0205427,100032,61, translation
JAG0010100300,30, translation
M60240100206,112, translation
U092016013438,1,81, translation
a02571069,10004,81, translation
M60201000651,1,84, translation
G50313136,100518, translation
GC031401007129,1,30, translation
U0187878,100092,30, translation
G5058380,100169,69, translation
G229511,100147,69, translation
U0313866,10007,69, translation
M602010006110,10, translation
U040101005641,305, translation
U17170100259,1,36, translation
U0140919,100136,36, translation
G503756493,10020,35, translation
a0255414,100060,36, translation
U140010000141,867, translation
536, ...368,8,312, translation
U094726,100282,61, translation
U194944,100123,95, translation
GC056631,100062,61, translation
U0310141,100610,39, translation
U0356671,10007,39, translation
G223865,100077,71, translation
U09101000091,1,612, translation
U0314550,101165,63, translation
M601010015781,1,98, translation
A01010001031,1,413, translation
G2032427,1, ...ase, translation
U02071001534,176, translation
U0307544,100039,39, translation
GC059974,1294,39, translation
G50375534,100035,69, translation
U177384,100014,63, translation
M6001002481,1,34, translation
a0358629,100115,63, translation
U05, ...869,17,130, translation
U0110003631,1,446, translation
G2027433,100094,63, translation
U08007013246,1,48, translation
U0207433,100064,63, translation

Consensus	144
66227030_1311_33_translation	144
61317381_100604_36_translation	267
18575100101_1_51_translation	325
732256_>_loc.894_translation	438
60210058_10255_02_translation	411
60101001304_1_100_translation	277
60101001304_1_100_translation	162
6027553_102051_46_translation	279
6027553_102051_46_translation	485
6172375_100150_03_translation	401
60101001304_1_100_translation	467
6027553_102051_46_translation	807
6027553_102051_46_translation	415
6027553_102051_46_translation	262
6172375_100150_03_translation	84
6172375_100150_03_translation	691
6172375_100150_03_translation	320
6172375_100150_03_translation	438
6172375_100150_03_translation	433
6172375_100150_03_translation	404
6172375_100150_03_translation	311
6172375_100150_03_translation	244
6172375_100150_03_translation	405
6172375_100150_03_translation	805
6172375_100150_03_translation	236
6172375_100150_03_translation	947
6172375_100150_03_translation	470
6172375_100150_03_translation	386
6172375_100150_03_translation	459
6172375_100150_03_translation	397
6172375_100150_03_translation	310
6172375_100150_03_translation	401
6172375_100150_03_translation	264
6172375_100150_03_translation	469
6172375_100150_03_translation	427
6172375_100150_03_translation	428
6172375_100150_03_translation	381
6172375_100150_03_translation	381
6172375_100150_03_translation	132
6172375_100150_03_translation	398
6172375_100150_03_translation	373
6172375_100150_03_translation	309
6172375_100150_03_translation	389
6172375_100150_03_translation	429
6172375_100150_03_translation	422
6172375_100150_03_translation	341
6172375_100150_03_translation	414
6172375_100150_03_translation	420
6172375_100150_03_translation	285
6172375_100150_03_translation	842
6172375_100150_03_translation	421
6172375_100150_03_translation	377
6172375_100150_03_translation	435
6172375_100150_03_translation	430
6172375_100150_03_translation	452
6172375_100150_03_translation	375
6172375_100150_03_translation	400
6172375_100150_03_translation	427
6172375_100150_03_translation	576
6172375_100150_03_translation	363
6172375_100150_03_translation	425
6172375_100150_03_translation	328
6172375_100150_03_translation	239
6172375_100150_03_translation	359
6172375_100150_03_translation	398
6172375_100150_03_translation	391
6172375_100150_03_translation	277
6172375_100150_03_translation	514
6172375_100150_03_translation	527
6172375_100150_03_translation	314
6172375_100150_03_translation	429

FIG. 4-32

Consensus	146
147	148
149	150
151	152
153	154
155	156
157	158
159	160
161	162
163	164
165	166
167	168
169	170
171	172
173	174
175	176
177	178
179	180
181	182
183	184
185	186
187	188
189	190
191	192
193	194
195	196
197	198
199	200
201	202
203	204
205	206
207	208
209	210
211	212
213	214
215	216
217	218
219	220
221	222
223	224
225	226
227	228
229	230
231	232
233	234
235	236
237	238
239	240
241	242
243	244
245	246
247	248
249	250
251	252
253	254
255	256
257	258
259	260
261	262
263	264
265	266
267	268
269	270
271	272
273	274
275	276
277	278
279	280
281	282
283	284
285	286
287	288
289	290
291	292
293	294
295	296
297	298
299	300
301	302
303	304
305	306
307	308
309	310
311	312
313	314
315	316
317	318
319	320
321	322
323	324
325	326
327	328
329	330
331	332
333	334
335	336
337	338
339	340
341	342
343	344
345	346
347	348
349	350
351	352
353	354
355	356
357	358
359	360
361	362
363	364
365	366
367	368
369	370
371	372
373	374
375	376
377	378
379	380
381	382
383	384
385	386
387	388
389	390
391	392
393	394
395	396
397	398
399	400
401	402
403	404
405	406
407	408
409	410
411	412
413	414
415	416
417	418
419	420
421	422
423	424
425	426
427	428
429	430
431	432
433	434
435	436
437	438
439	440
441	442
443	444
445	446
447	448
449	450
451	452
453	454
455	456
457	458
459	460
461	462
463	464
465	466
467	468
469	470
471	472
473	474
475	476
477	478
479	480
481	482
483	484
485	486
487	488
489	490
491	492
493	494
495	496
497	498
499	500
501	502
503	504
505	506
507	508
509	510
511	512
513	514
515	516
517	518
519	520
521	522
523	524
525	526
527	528
529	530
531	532
533	534
535	536
537	538
539	540
541	542
543	544
545	546
547	548
549	550
551	552
553	554
555	556
557	558
559	560
561	562
563	564
565	566
567	568
569	570
571	572
573	574
575	576
577	578
579	580
581	582
583	584
585	586
587	588
589	590
591	592
593	594
595	596
597	598
599	600
601	602
603	604
605	606
607	608
609	610
611	612
613	614
615	616
617	618
619	620
621	622
623	624
625	626
627	628
629	630
631	632
633	634
635	636
637	638
639	640
641	642
643	644
645	646
647	648
649	650
651	652
653	654
655	656
657	658
659	660
661	662
663	664
665	666
667	668
669	670
671	672
673	674
675	676
677	678
679	680
681	682
683	684
685	686
687	688
689	690
691	692
693	694
695	696
697	698
699	700
701	702
703	704
705	706
707	708
709	710
711	712
713	714
715	716
717	718
719	720
721	722
723	724
725	726
727	728
729	730
731	732
733	734
735	736
737	738
739	740
741	742
743	744
745	746
747	748
749	750
751	752
753	754
755	756
757	758
759	760
761	762
763	764
765	766
767	768
769	770
771	772
773	774
775	776
777	778
779	780
781	782
783	784
785	786
787	788
789	790
791	792
793	794
795	796
797	798
799	800
801	802
803	804
805	806
807	808
809	810
811	812
813	814
815	816
817	818
819	820
821	822
823	824
825	826
827	828
829	830
831	832
833	834
835	836
837	838
839	840
841	842
843	844
845	846
847	848
849	850
851	852
853	854
855	856
857	858
859	860
861	862
863	864
865	866
867	868
869	870
871	872
873	874
875	876
877	878
879	880
881	882
883	884
885	886
887	888
889	890
891	892
893	894
895	896
897	898
899	900
901	902
903	904
905	906
907	908
909	910
911	912
913	914
915	916
917	918
919	920
921	922
923	924
925	926
927	928
929	930
931	932
933	934
935	936
937	938
939	940
941	942
943	944
945	946
947	948
949	950
951	952
953	954
955	956
957	958
959	960
961	962
963	964
965	966
967	968
969	970
971	972
973	974
975	976
977	978
979	980
981	982
983	984
985	986
987	988
989	990
991	992
993	994
995	996
997	998
999	1000

FIG. 4-33

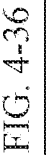
Consensus	
> 600305930.1811.79 translation	157
> 6137383.100204.39 translation	270
> 05226100181.1.55 translation	335
> 335236 --> 1cc.99 translation	469
> 60210956.10555.03 translation	432
> 60101001904.1.agg translation	411
> 603079354.10057.03 translation	277
> 6037653.303053.en translation	365
> 60376607.30427.07 translation	420
> 6172175.100152.03 translation	379
> 601801732.10008.09 translation	467
> 60180102005.18.1.01 translation	414
> 60307930.10092.01 translation	474
> 6040501000144.1.agg translation	410
> 60183791.10166.02 translation	482
> 6164342.100035.03 translation	415
> 6029697.100622.44 translation	282
> 6171717.100019.02 translation	89
> 61701001806.1.956 translation	495
> 6075935.100173.03 translation	325
> 6133255.100066.03 translation	386
> 607560470.10001.70 translation	456
> 60720123.10001.03 translation	305
> 60730594.100153.21 translation	316
> 6166232.1001185.09 translation	344
> 6200148.100151.03 translation	400
> 60356609.1 --> 139 translation	357
> 60101163.1136.98 translation	450
> 6205850100006.63 translation	420
> 60209608.10016.03 translation	391
> 60309342.10032.03 translation	462
> 60101002000.98 translation	297
> 60200100012.1.12 translation	215
> 054201010439.1.81 translation	405
> 60257069.10004.03 translation	307
> 050021000351.1.64 translation	409
> 60610136.10055.16 translation	425
> 600501002129.1.20 translation	694
> 6076878.10002.94 translation	381
> 6076880.10105.04 translation	396
> 6126511.100417.09 translation	142
> 6034486.10007.09 translation	377
> 530001000158.1.10 translation	401
> 05020100064.1.305 translation	314
> 0122275.100059.39 translation	403
> 05001007529.1.76 translation	654
> 014813.100256.39 translation	425
> 60276448.0000.95 translation	341
> 60153344.10000.29 translation	417
> 556 --> 1cc.9.8.512 translation	405
> 6073226.100382.33 translation	285
> 614044.100123.95 translation	444
> 60756831.10003.03 translation	424
> 6030914.100610.36 translation	382
> 60258871.10007.36 translation	439
> 6023826.100077.71 translation	432
> 6039656.100188.81 translation	454
> 60410100098.1.012 translation	386
> 6015350.10185.02 translation	400
> 60101001579.1.08 translation	442
> 60301000103.1.413 translation	591
> 6022627.1 --> aec translation	399
> 6021001534.1.76 translation	427
> 60009761.1041.36 translation	333
> 60276312.100035.46 translation	238
> 6137384.10014.03 translation	402
> 606001003498.1.04 translation	402
> 60756819.10015.03 translation	399
> 6101003815.1.846 translation	377
> 6207433.100604.31 translation	517
> 603001013340.1.48 translation	530
> 6267433.100644.33.5063605	314
> 6267433.100644.33.5063605	431

FIG. 4-34

Consensus	6V	S	H	R	152
	GI	H	A	R	175
	QI	N	L	S	325
	SV	K	I	D	424
	6V	S	H	R	412
	6V	S	H	R	170
	HK	N	I	K	279
	TI	S	V	K	472
	KL	DT	F	G	419
	GI	S	Y	R	980
	GI	N	H	R	415
	KL	S	Y	R	487
	SA	S	Y	R	319
	SA	S	Y	R	284
	SA	S	Y	R	94
	SA	S	Y	R	498
	SA	S	Y	R	528
	SA	S	Y	R	368
	TE	IK	F	K	442
	GI	N	Y	K	368
	GI	N	Y	K	412
	GA	K	Y	A	321
	GI	N	Y	K	244
	SV	N	Y	K	413
	GI	N	Y	K	413
	SV	N	Y	K	227
	GI	N	Y	K	455
	SV	N	Y	K	424
	SV	N	Y	K	356
	TA	R	Y	K	467
	HA	K	K	F	257
	HA	K	K	F	350
	TP	K	S	K	435
	TP	K	S	K	402
	DV	K	P	K	434
	SK	S	V	D	434
	ES	V	D	K	438
	6V	S	A	K	261
	SA	S	A	K	145
	V	N	H	K	405
	GI	S	Y	K	382
	KA	E	L	R	319
	PI	N	E	S	408
	SL	N	M	K	435
	TA	P	A	A	430
	CV	N	L	T	346
	VA	K	T	K	422
	AT	S	Y	K	410
	SA	S	LP	S	285
	SA	S	LP	S	450
	NA	K	A	Q	428
	SS	A	P	K	387
	KA	S	P	K	444
	ET	S	W	K	459
	YL	R	A	D	365
	DV	H	P	K	405
	DI	N	T	K	447
	MQ	A	R	A	366
	SV	S	A	K	374
	FM	S	Y	K	432
	FT	A	S	K	327
	GI	N	H	R	239
	SV	S	Y	K	407
	SV	S	Y	K	408
	SV	S	Y	K	399
	SV	S	Y	K	279
	GI	N	S	K	522
	GA	N	R	R	525
	EV	A	A	K	314
	EV	A	A	K	435

FIG. 4-35

Consensus	6V	S	H	R	152
	GI	H	A	R	175
	QI	N	L	S	325
	SV	K	I	D	424
	6V	S	H	R	412
	6V	S	H	R	170
	HK	N	I	K	279
	TI	S	V	K	472
	KL	DT	F	G	419
	GI	S	Y	R	980
	GI	N	H	R	415
	KL	S	Y	R	487
	SA	S	Y	R	319
	SA	S	Y	R	284
	SA	S	Y	R	94
	SA	S	Y	R	498
	SA	S	Y	R	528
	SA	S	Y	R	368
	TE	IK	F	K	442
	GI	N	Y	K	368
	GI	N	Y	K	412
	GA	K	Y	A	321
	GI	N	Y	K	244
	SV	N	Y	K	413
	GI	N	Y	K	413
	SV	N	Y	K	227
	GI	N	Y	K	455
	SV	N	Y	K	424
	SV	N	Y	K	356
	TA	R	Y	K	467
	HA	K	K	F	257
	HA	K	K	F	350
	TP	K	S	K	435
	TP	K	S	K	402
	DV	K	P	K	434
	SK	S	V	D	434
	ES	V	D	K	438
	6V	S	A	K	261
	SA	S	A	K	145
	V	N	H	K	405
	GI	S	Y	K	382
	KA	E	L	R	319
	PI	N	E	S	408
	SL	N	M	K	435
	TA	P	A	A	430
	CV	N	L	T	346
	VA	K	T	K	422
	AT	S	Y	K	410
	SA	S	LP	S	285
	SA	S	LP	S	450
	NA	K	A	Q	428
	SS	A	P	K	387
	KA	S	P	K	444
	ET	S	W	K	459
	YL	R	A	D	365
	DV	H	P	K	405
	DI	N	T	K	447
	MQ	A	R	A	366
	SV	S	A	K	374
	FM	S	Y	K	432
	FT	A	S	K	327
	GI	N	H	R	239
	SV	S	Y	K	407
	SV	S	Y	K	408
	SV	S	Y	K	399
	SV	S	Y	K	279
	GI	N	S	K	522
	GA	N	R	R	525
	EV	A	A	K	314
	EV	A	A	K	435



[illegible]

FIG. 4-37

[illegible]

FIG. 4-38

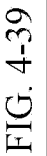
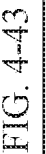


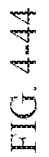
FIG. 4-40



FIG. 4-41

FIG. 4-42





[illegible]

FIG. 4-45

[illegible]

FIG. 4-46

[illegible]

FIG. 5

SEQ ID NOS: 2209-2223

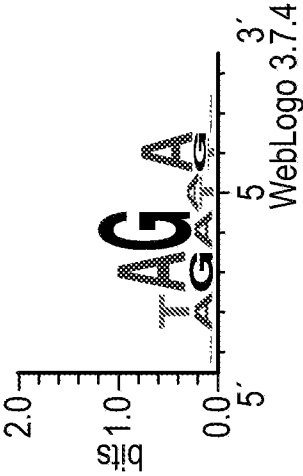


FIG. 6A

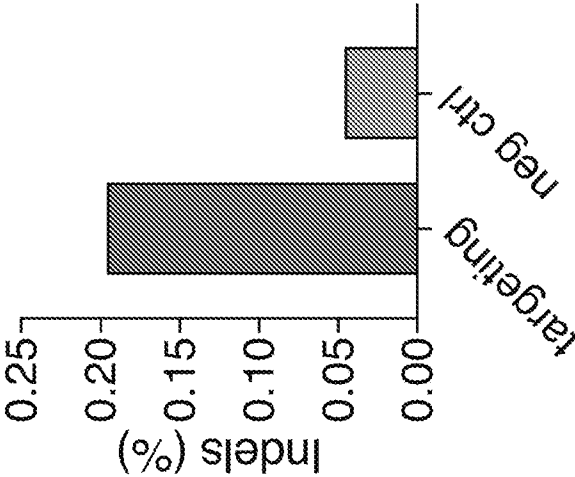


FIG. 6B

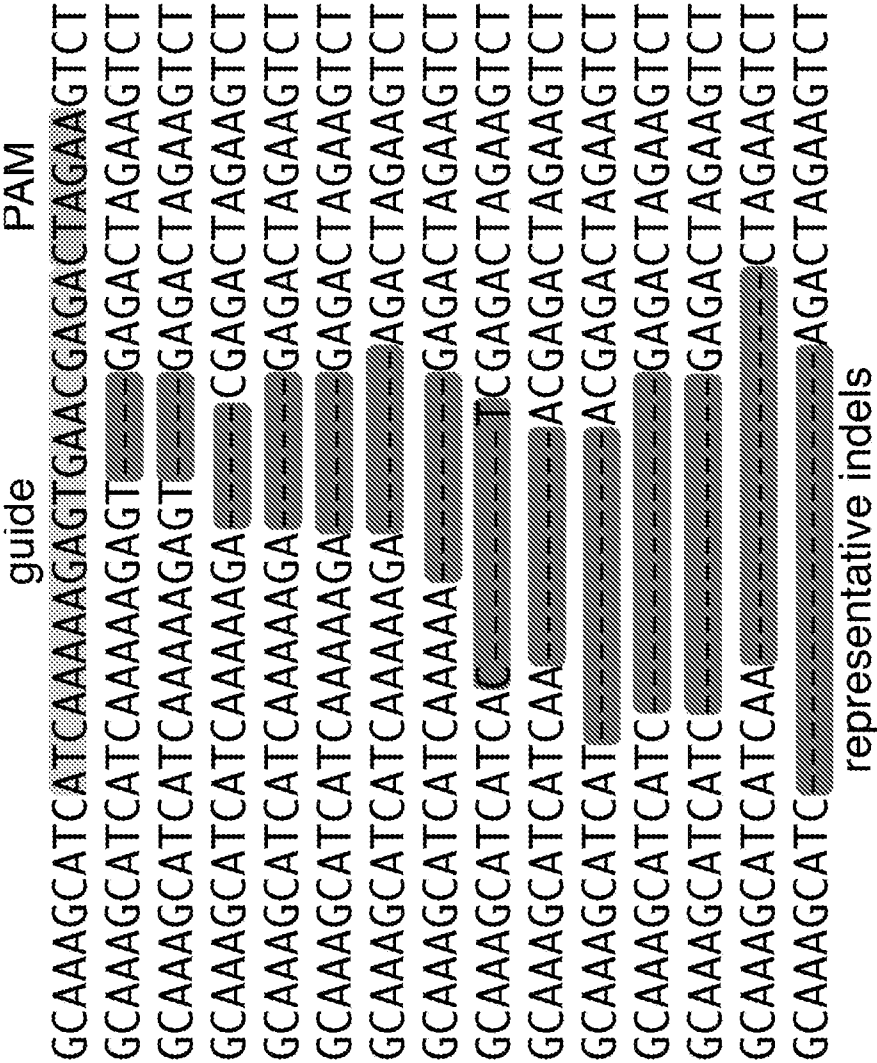


FIG. 6C

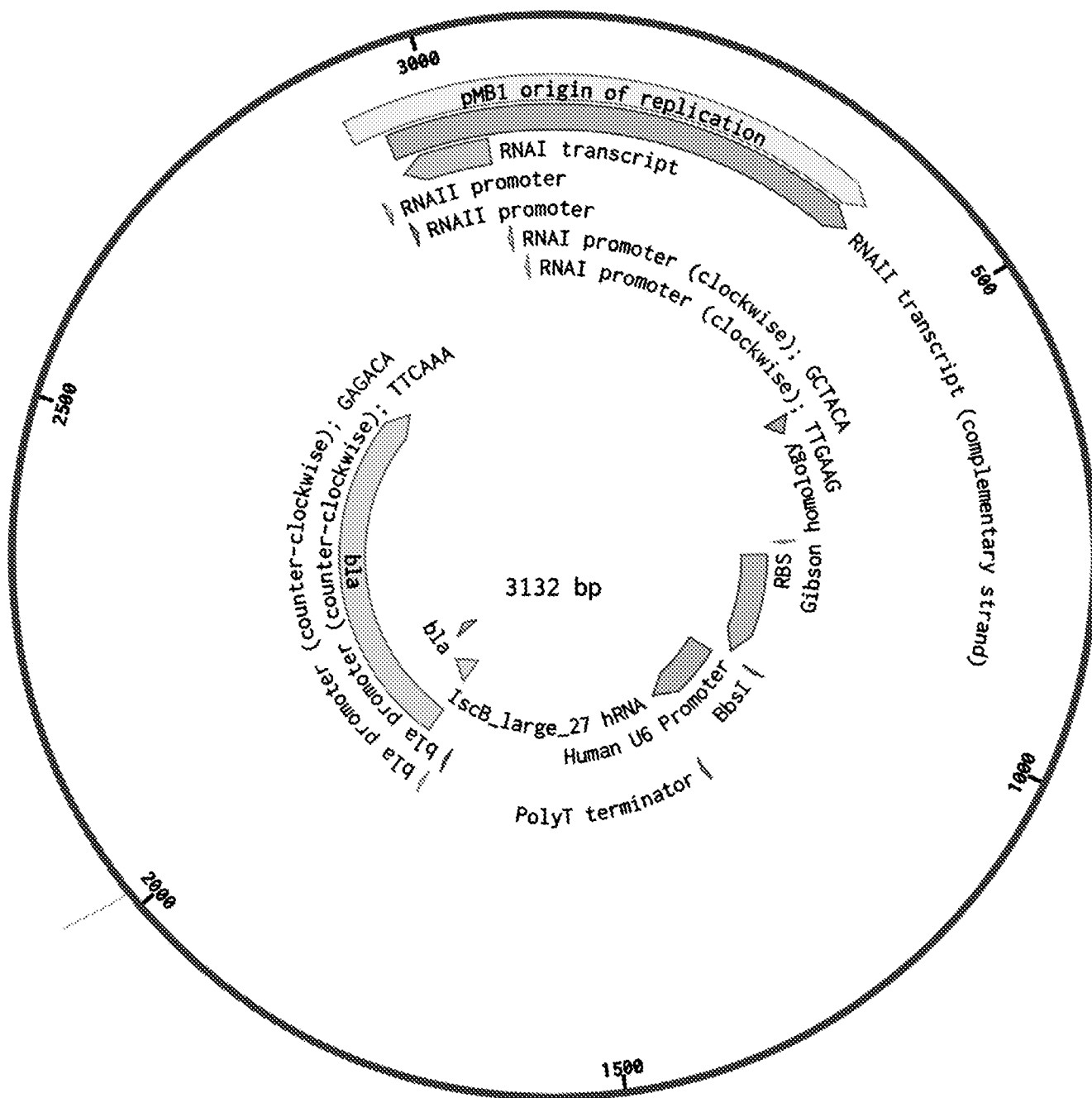
Protein sequence

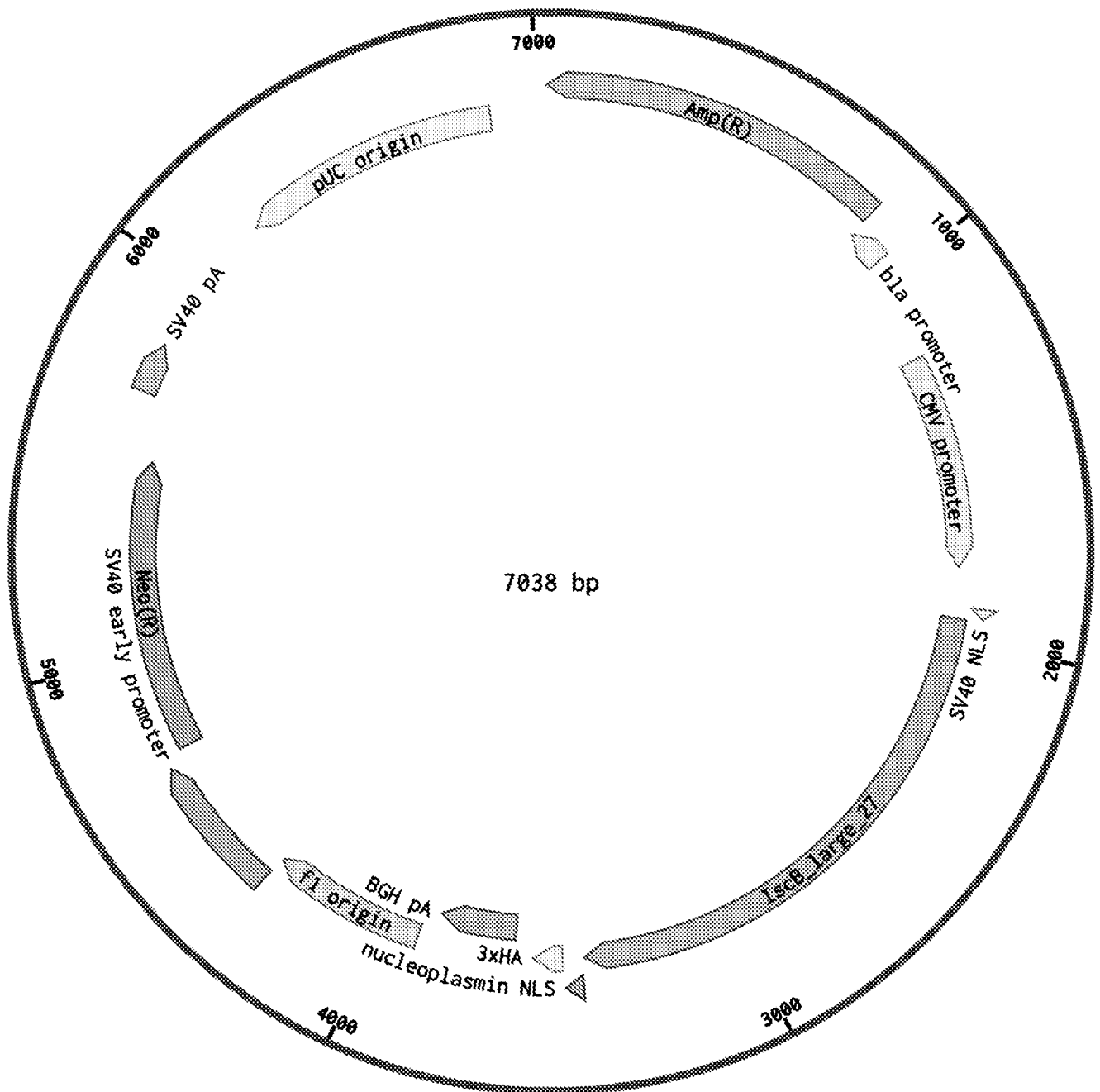
>OGEU01000025.1
 MMAVVVISKSGKPLMPTTRCGHVRIILLKEGKARVVERKPFITQLTYESAEEETQPLVLGIDP
 GRTNIGMSVVTESGESVFNAQIETRNKDVPKLMKDRKQYRMAHRRLKRRCKRRRRRAKAA
 GTAFEEGEKQRLLPGCFKPITCKSIRNKEARFNNRKRPPVGWLTPTANHLLVTHLNVVKKVQ
 KILPVAKVLELNRFSFMAMNNPKVQRWQYQRGPLYGKGSVEEAVSMQQDGHCLFCKHG
 IDHYHHVVP RRKNGSETLENRVGLCEEHHRLVHTDKEWEANLASKKSGMNNKKYHALSVL
 NQIIPYLADQLADMFPGNFCVTSGQDTYLFREEHGIPKDHLYDAYCIACSALTDAKKVSSPK
 GRPYMVHQFRRHRDRQACHKANLNRSYMMGGKLVATNRHKAMDQKTDLSLEEYRAAHSA
 DVSKLTVKHPSAQYKDMSRIMPGSILVSGEGKLFTLSRSEGRNKGQVNYFVSTEGIKYWA
 RKCQYLRNNGGLQIYV* SEQ ID NO: 2063

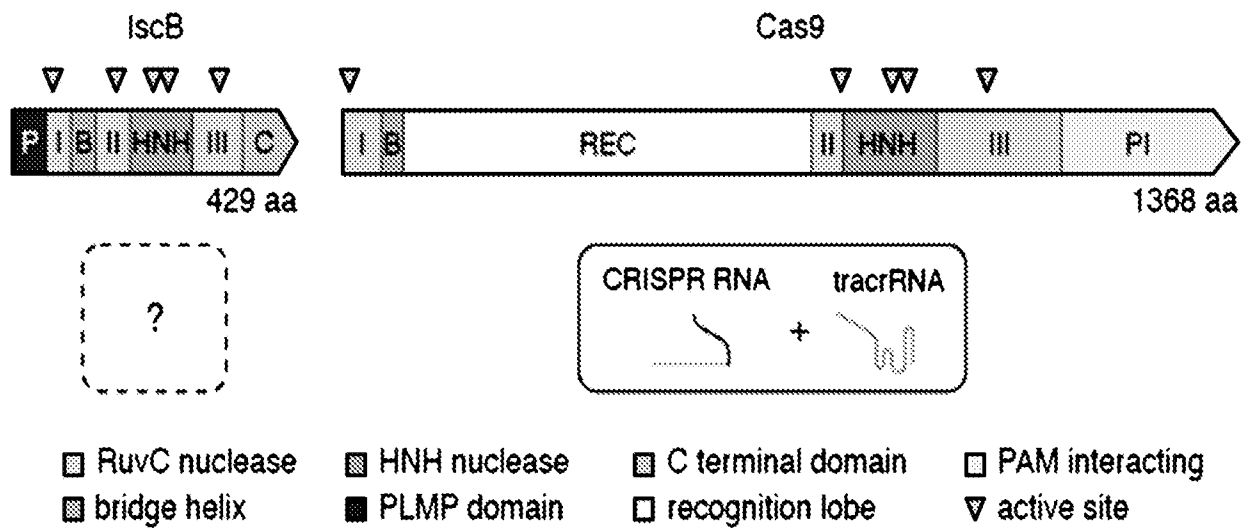
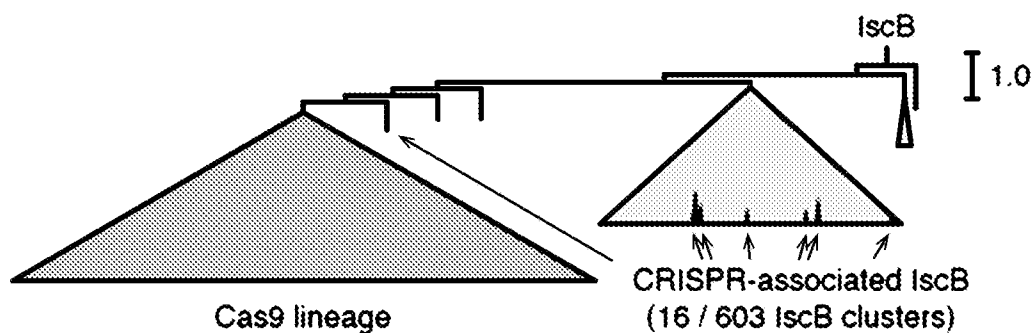
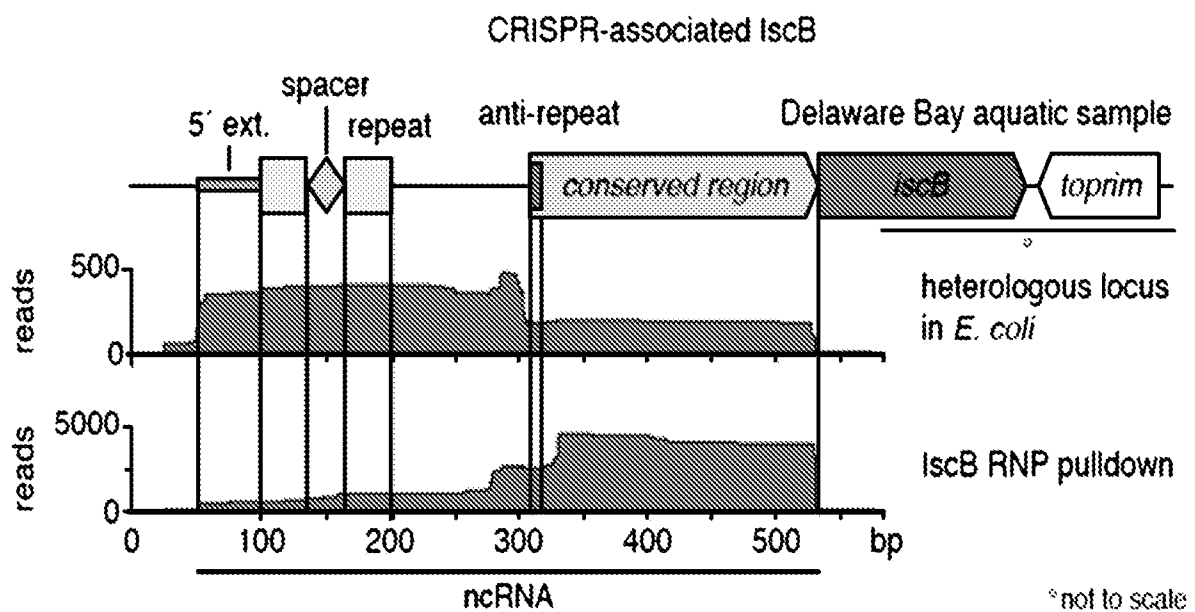
FIG. 7A**hRNA scaffold**

>OGEU01000025.1_hRNA
 GGCTCTTCCAACCTTATGGTTGCGACCGTAGGTTGAAAGAGCACAGGCTGAGACATTC
 GTAAGGCCGAAAGACCGGACGCCCTGGGATTTCGCCAGTCCCGGAACTGCATAG
 CGGATGCCAGTTGATGGAGCAATCTATCAGATAAGCCAGGGGGAACAATCACCTCTCT
 GTATCAGAGAGAGTTTACAAAAGGAGGAACGG SEQ ID NO: 2064

FIG. 7B

**FIG. 8A**

**FIG. 8B**

**FIG. 9A****FIG. 9B****FIG. 9C**

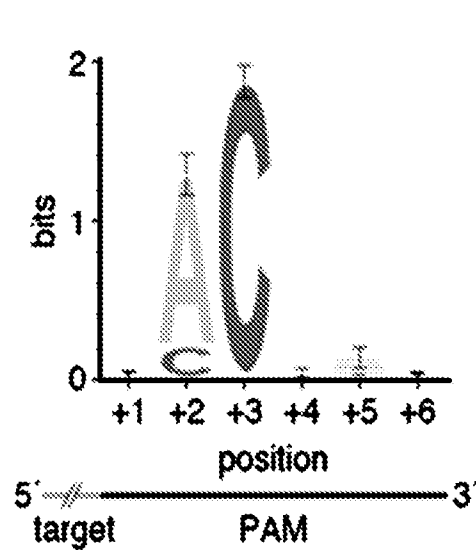


FIG. 9D

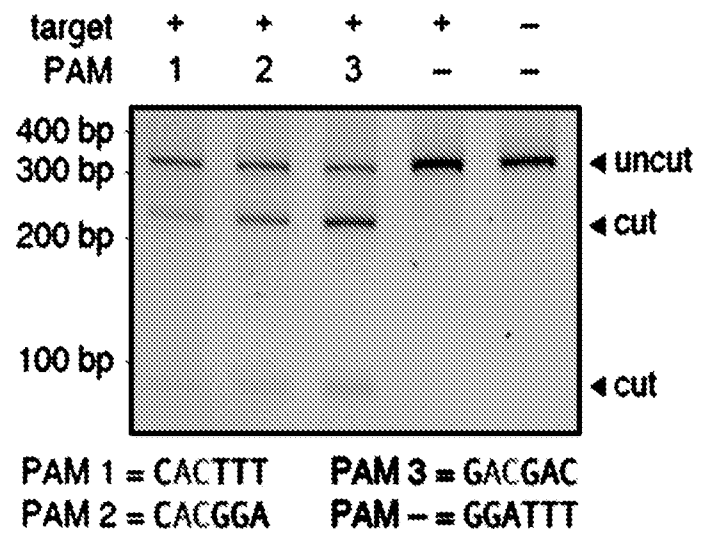


FIG. 9E

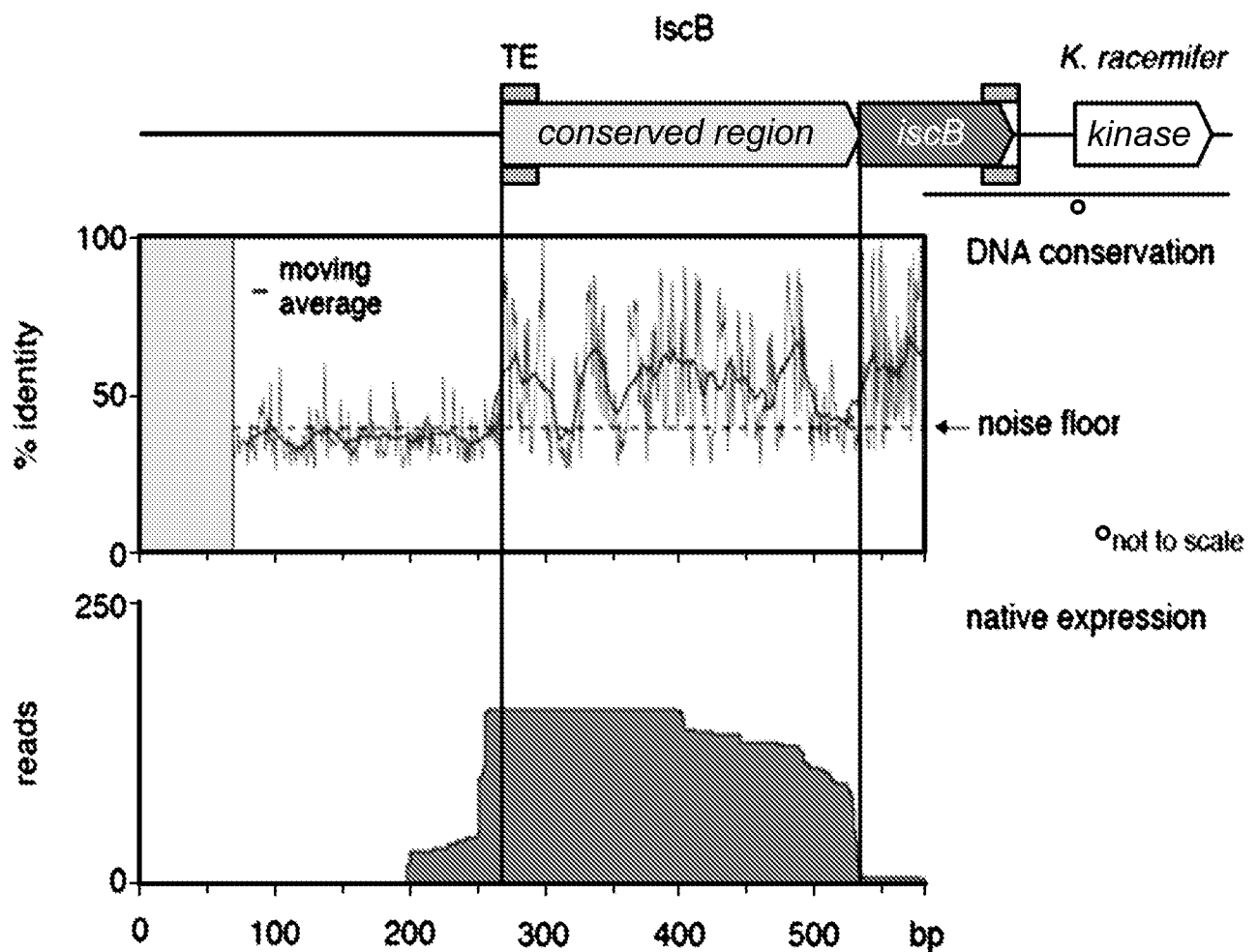


FIG. 9F

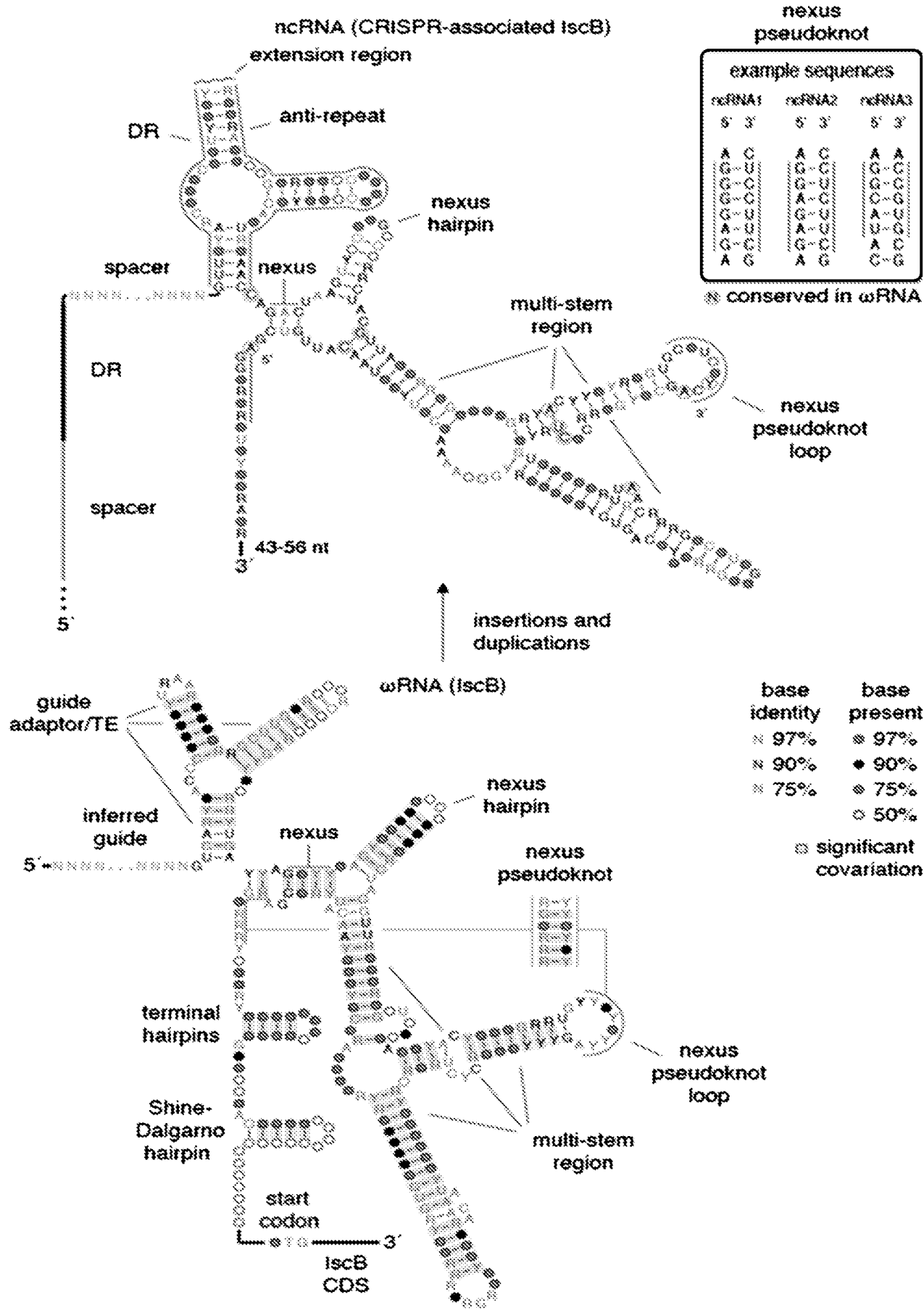


FIG. 9G

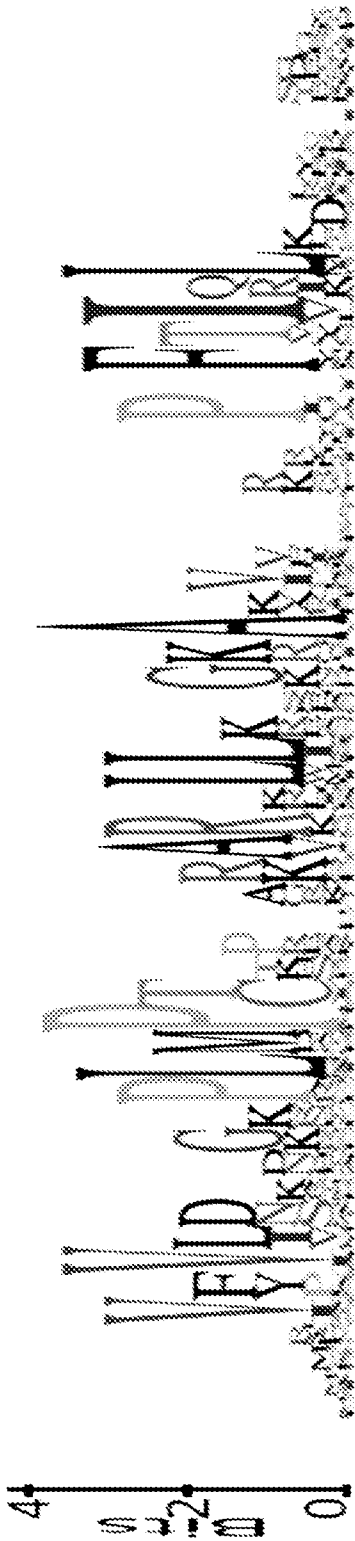


FIG. 10

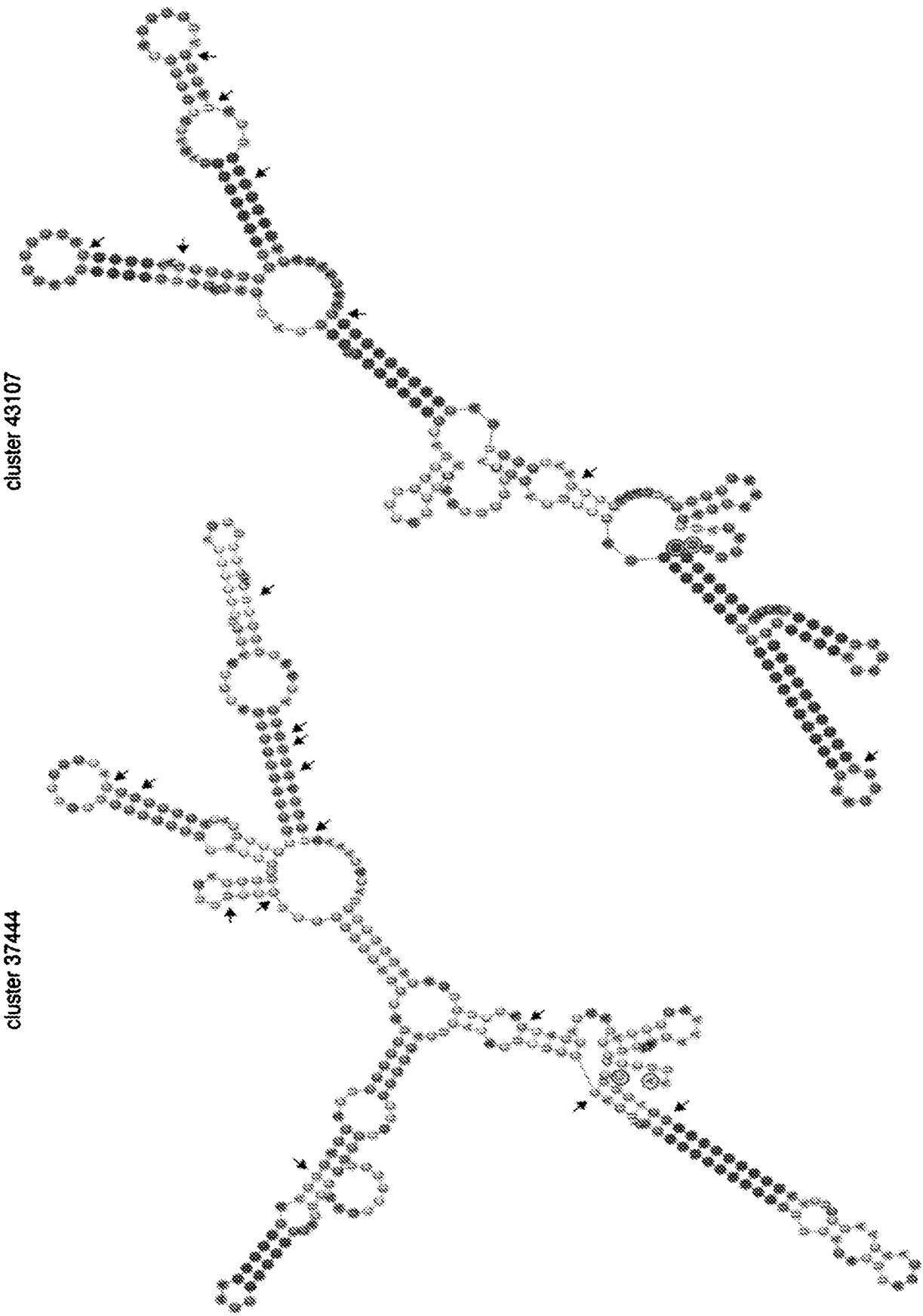


FIG. 11

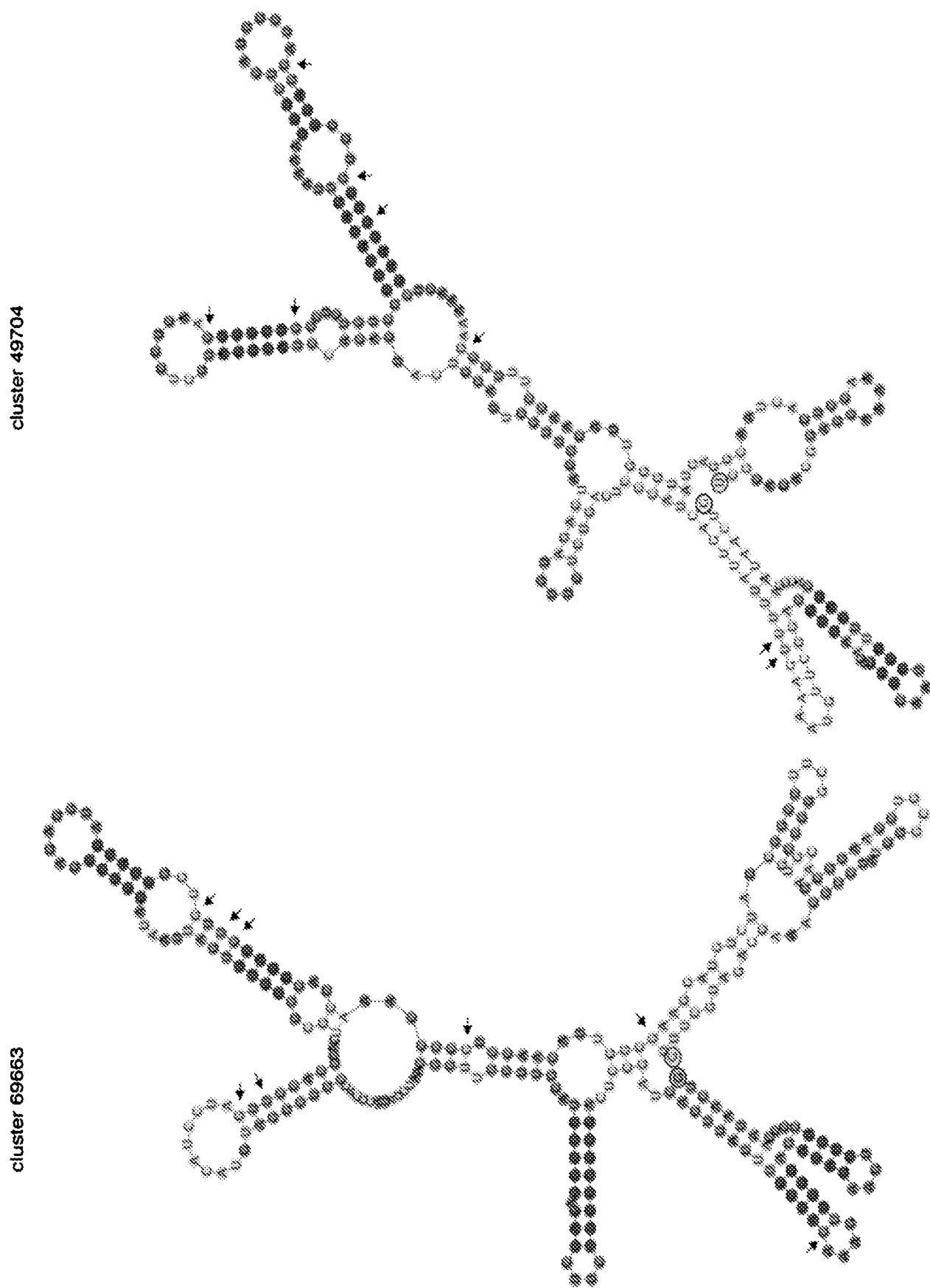


FIG. 11 (continued)

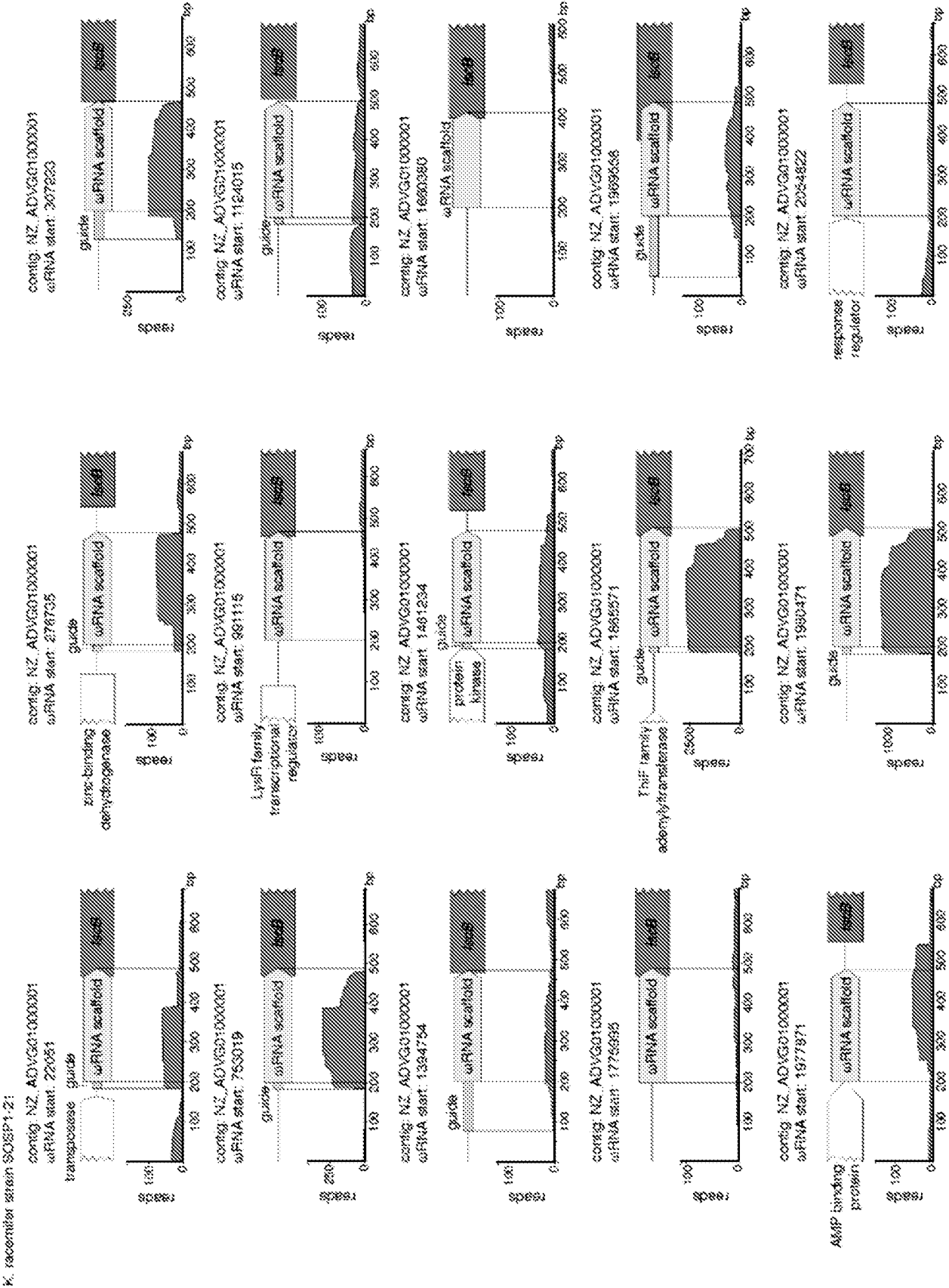


FIG. 12

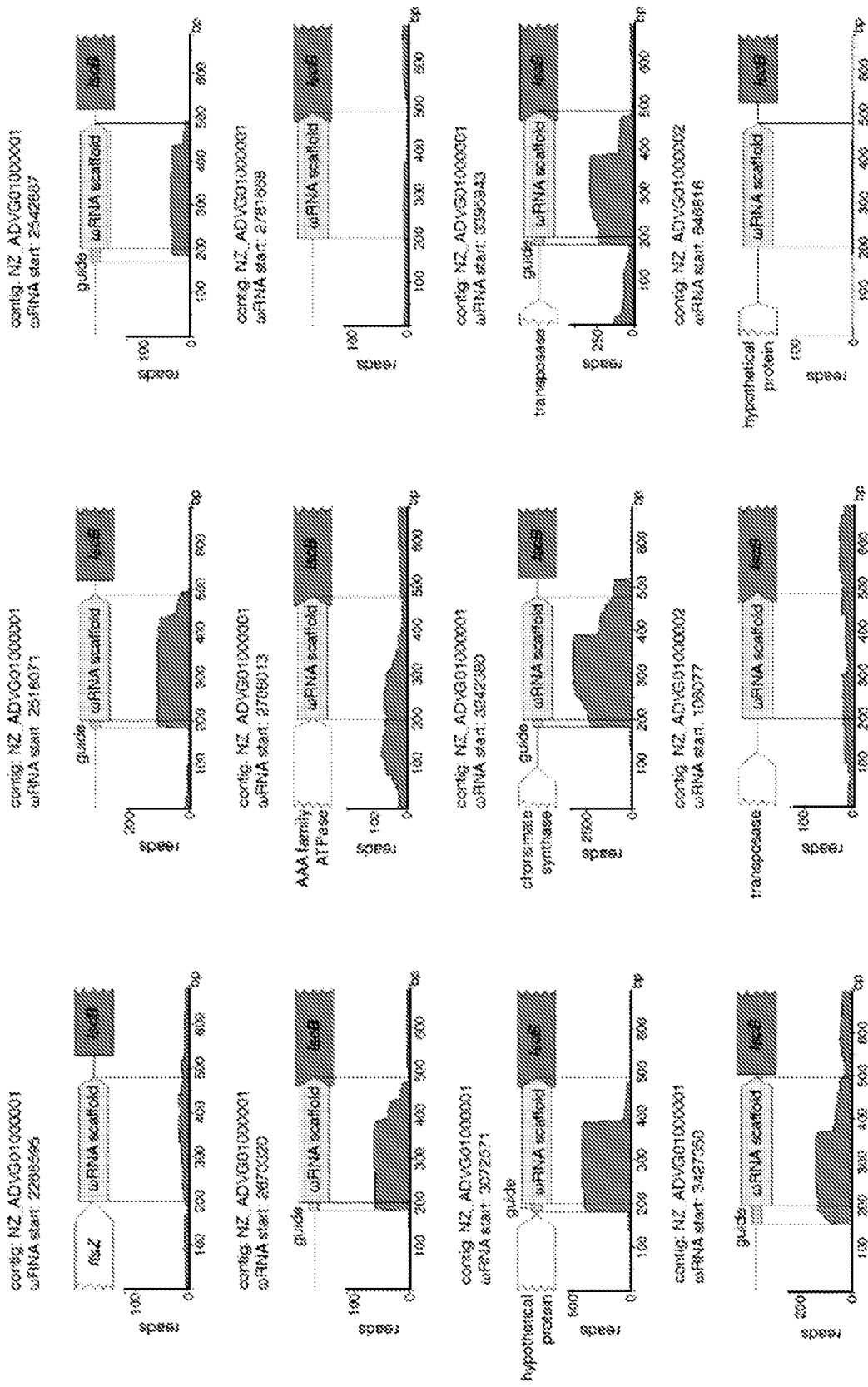


FIG. 12 (continued)

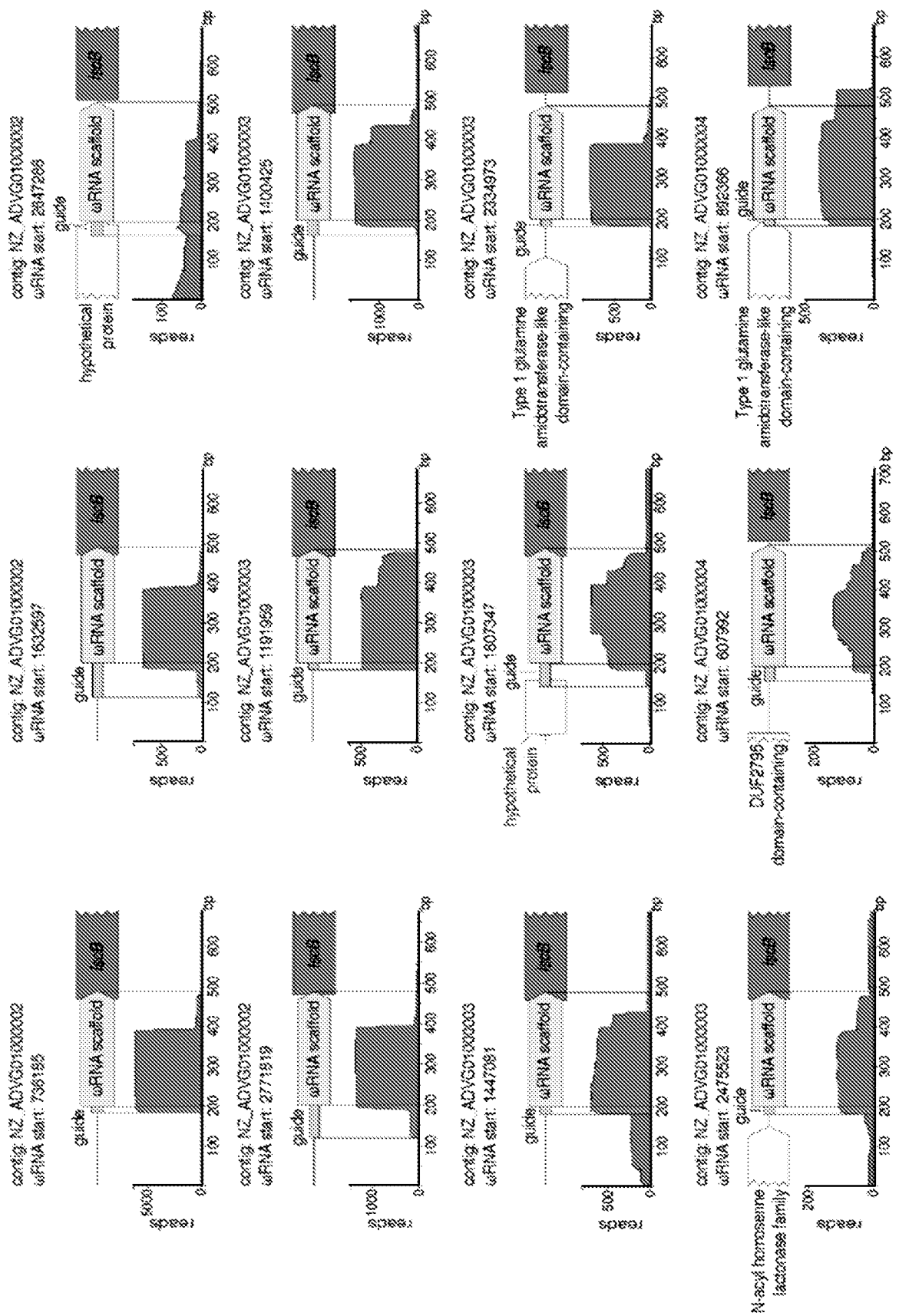


FIG. 12 (continued)

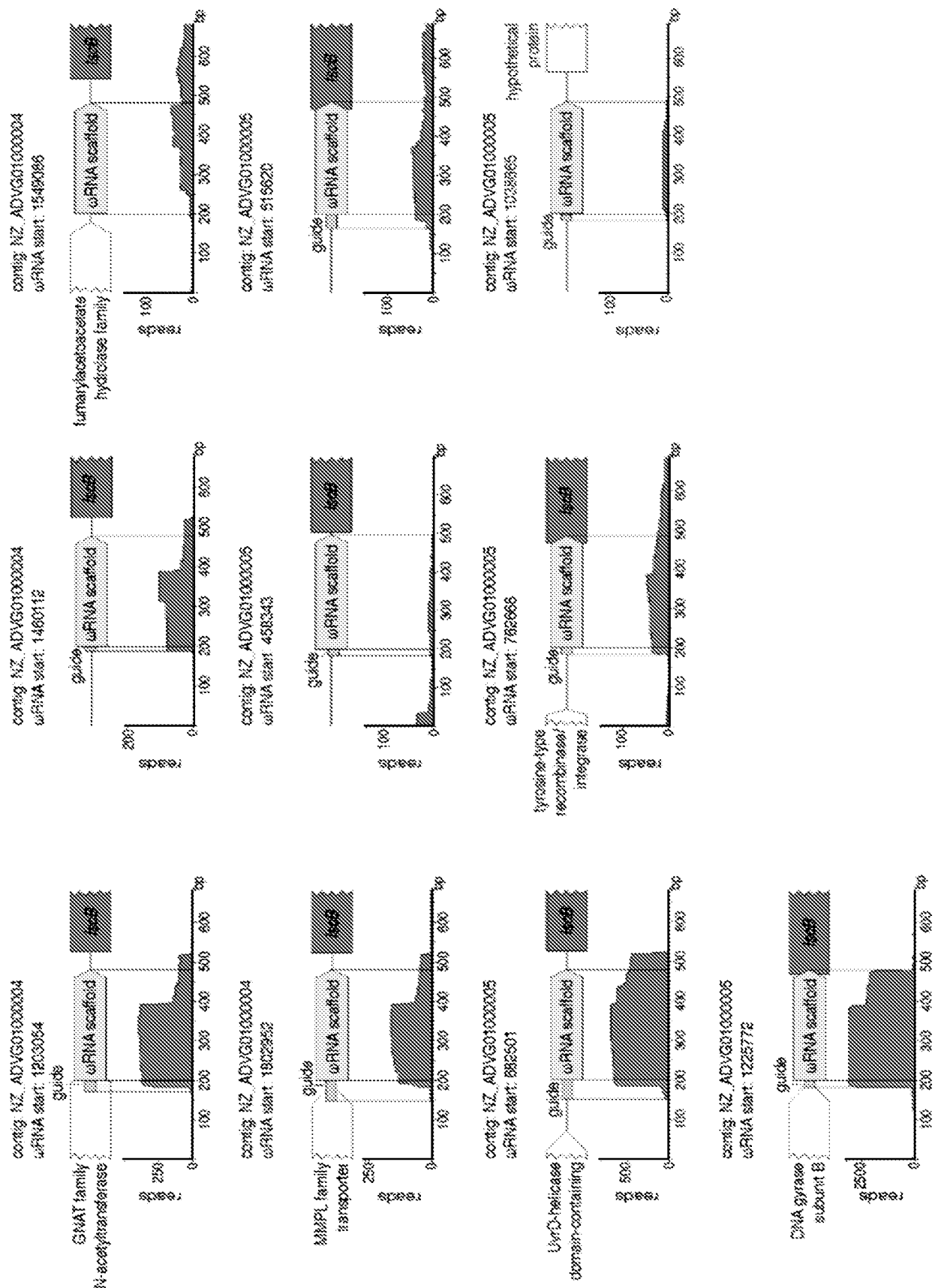


FIG. 12 (continued)

K. racemiter lsc8-1 RNP pull-down

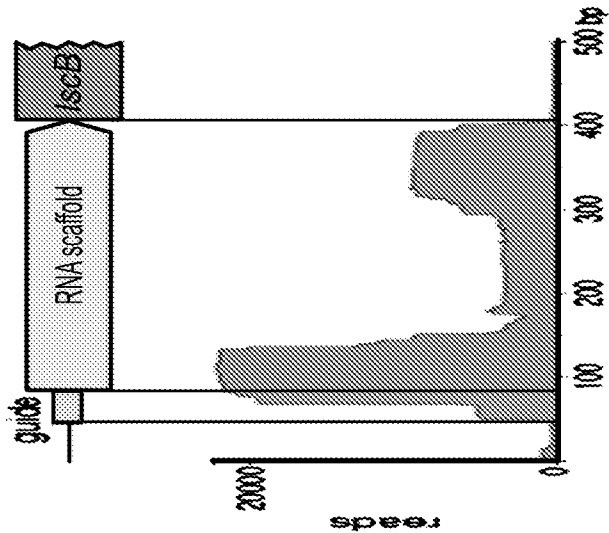


FIG. 13A

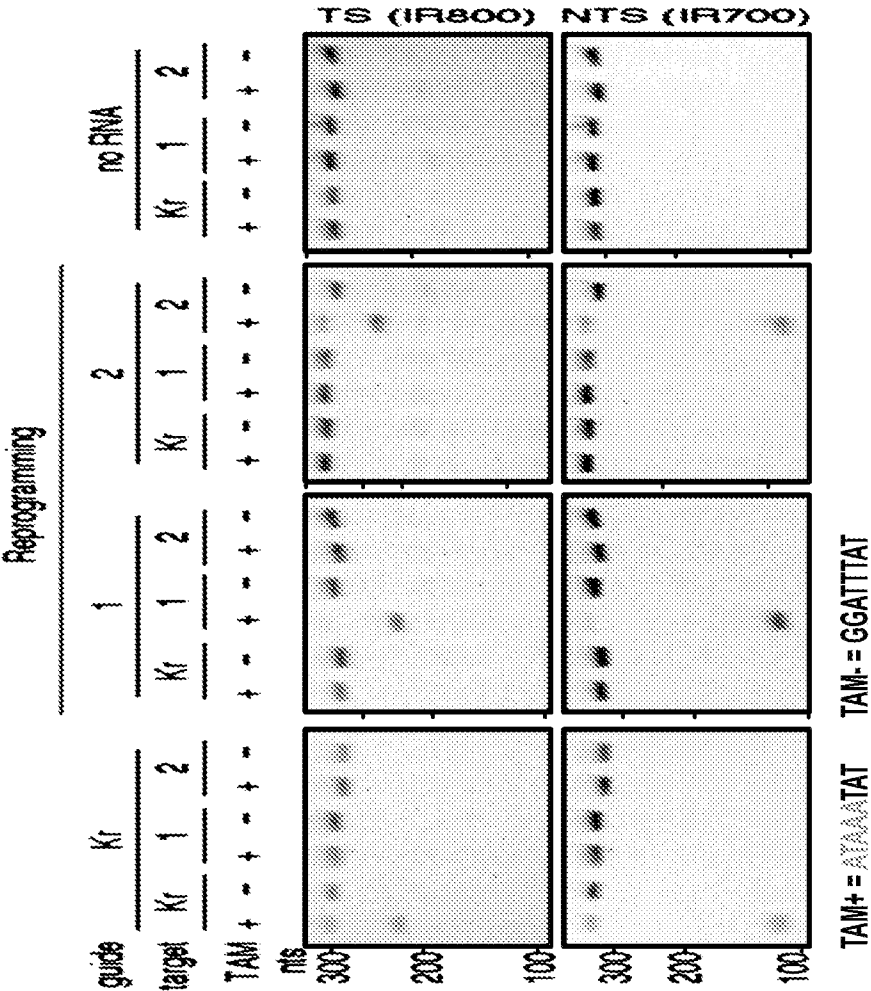


FIG. 13B

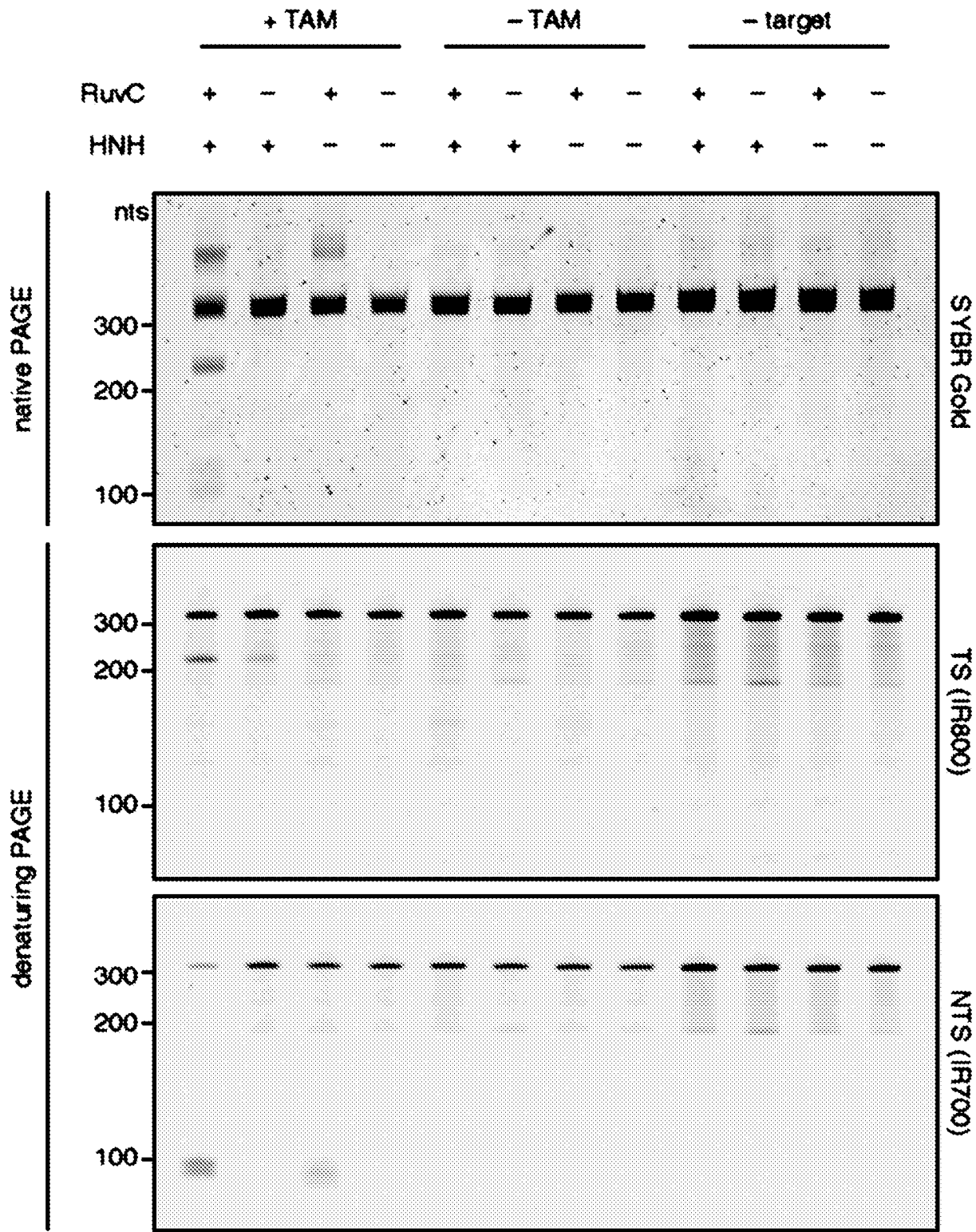


FIG. 13C

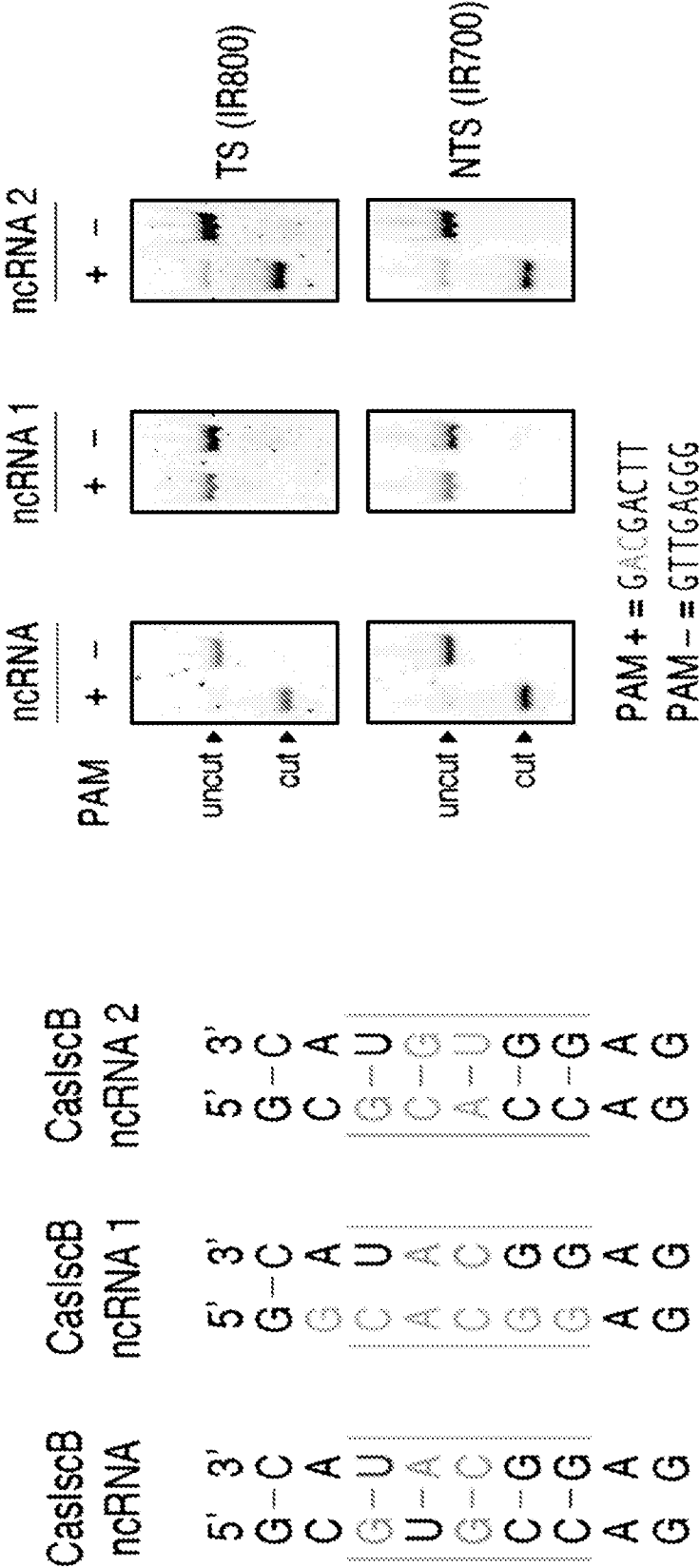


FIG. 14A

FIG. 14B

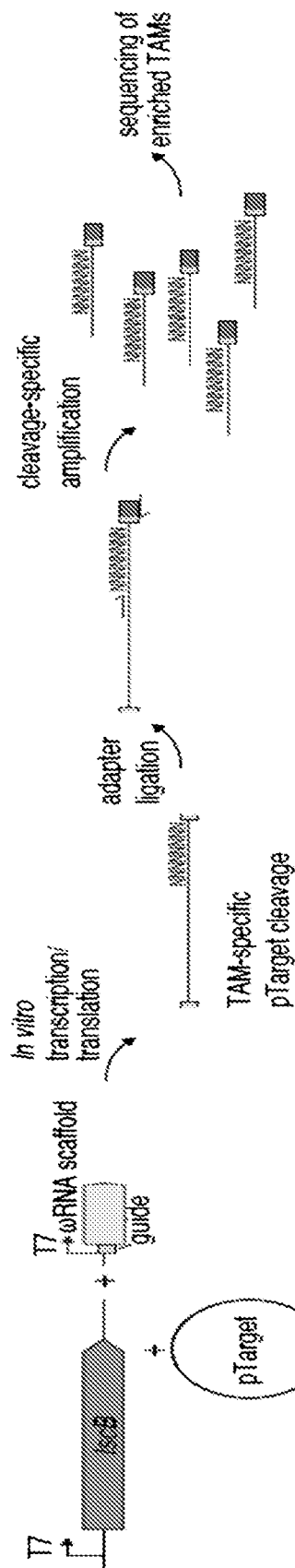


FIG. 15A

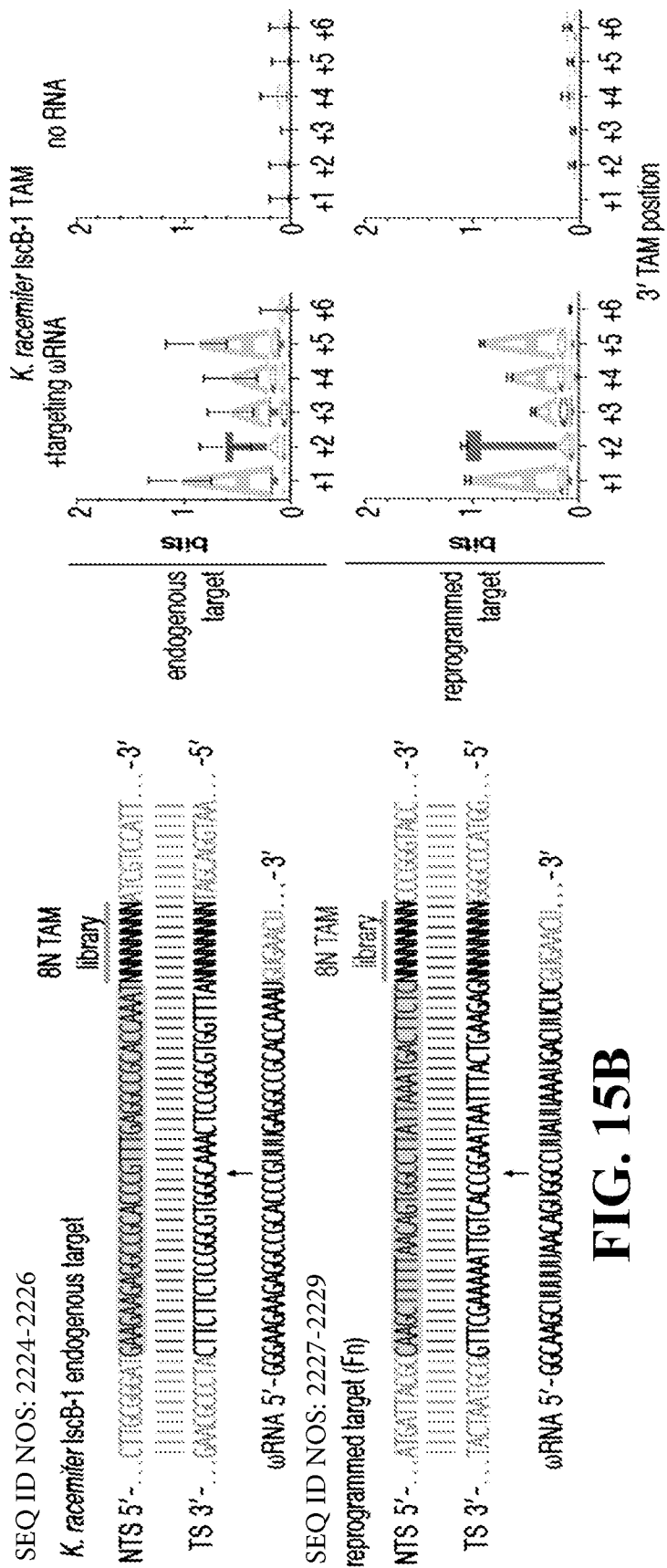


FIG. 15C

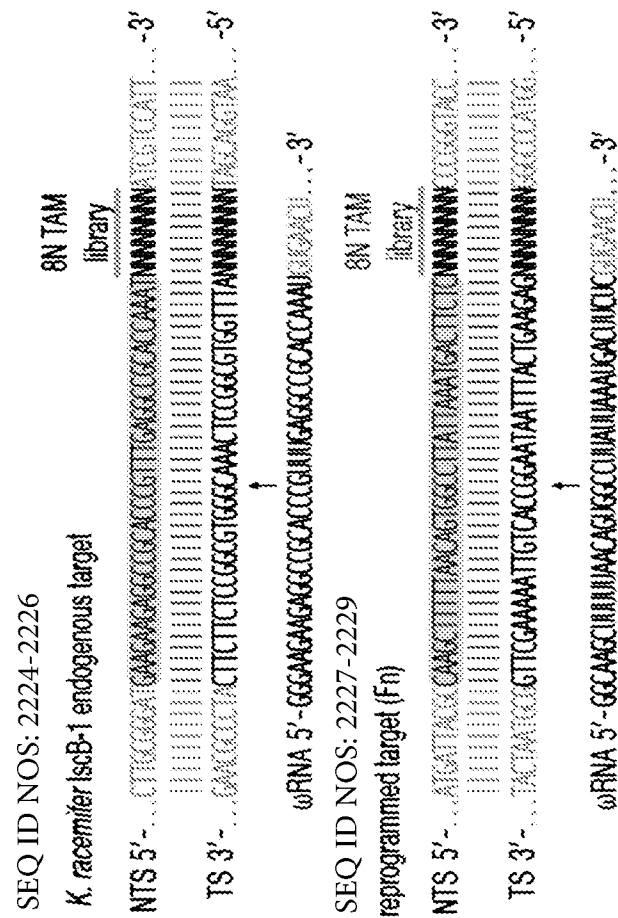
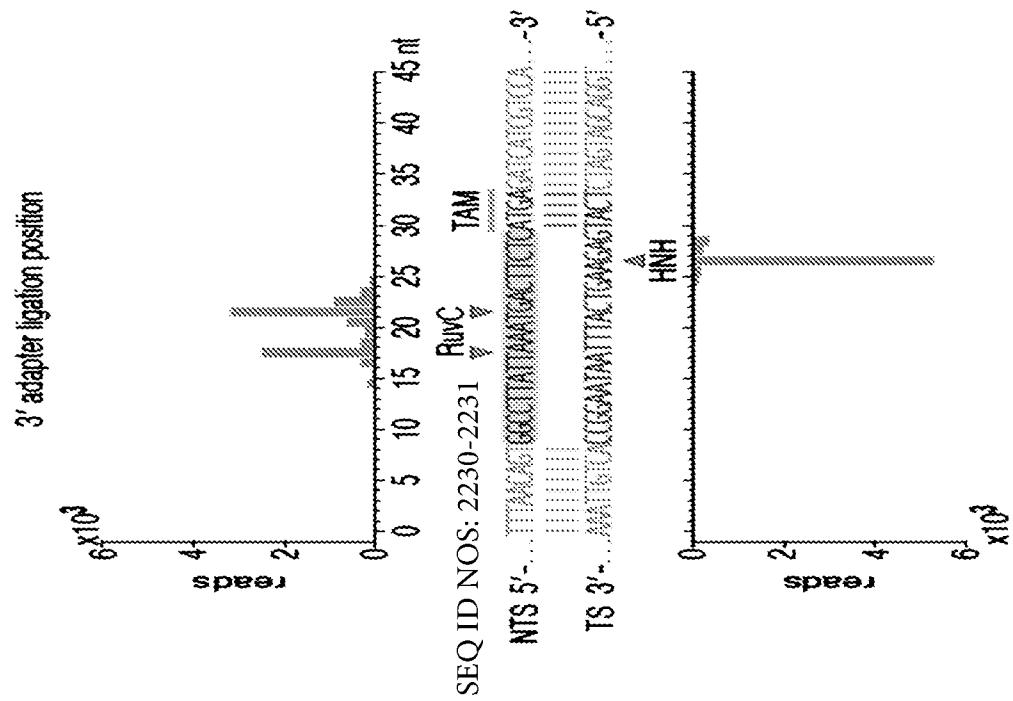
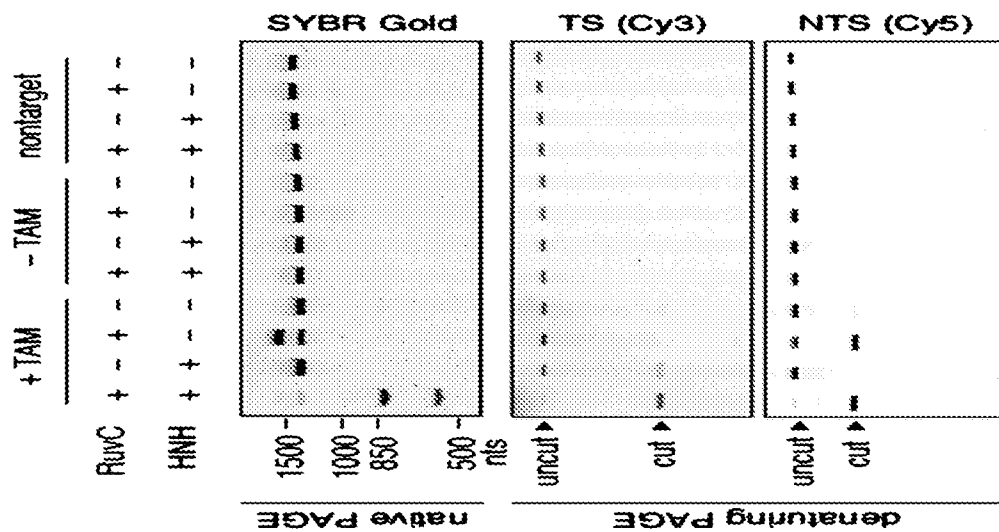
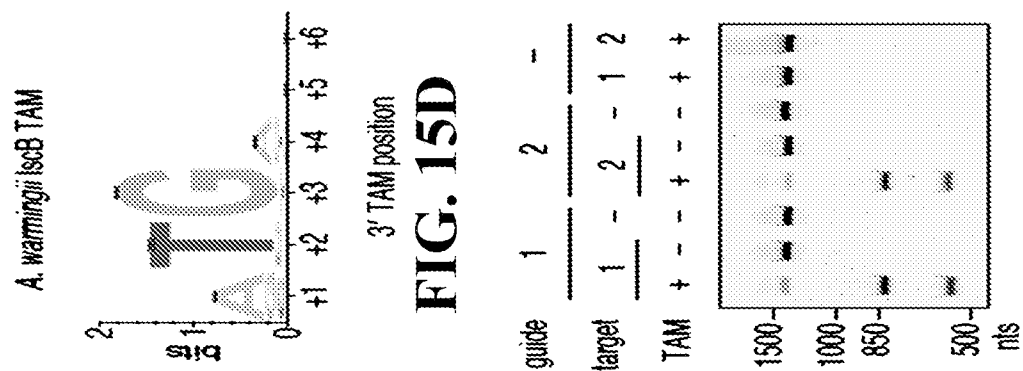


FIG. 15B



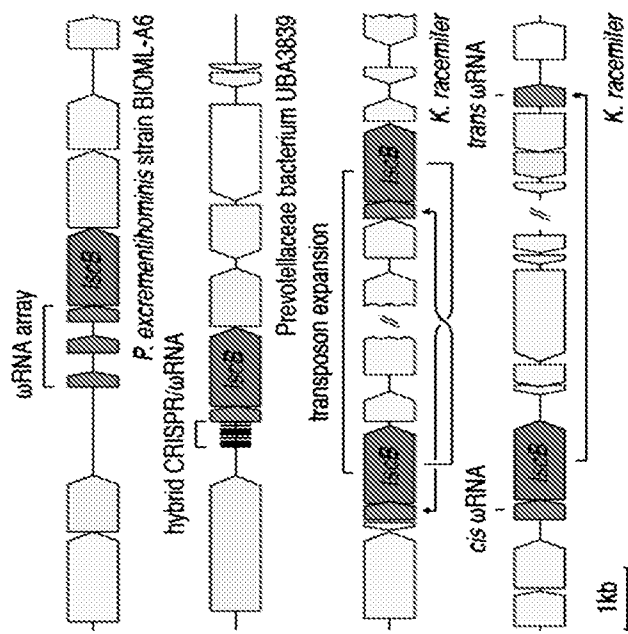


FIG. 16A

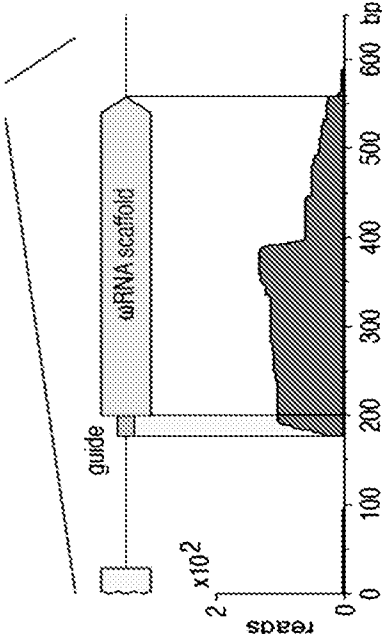


FIG. 16C

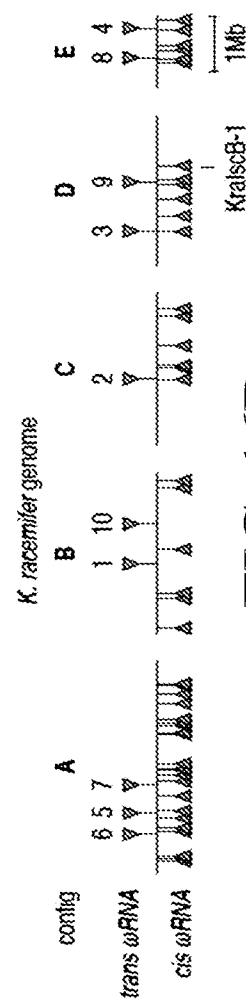


FIG. 16B

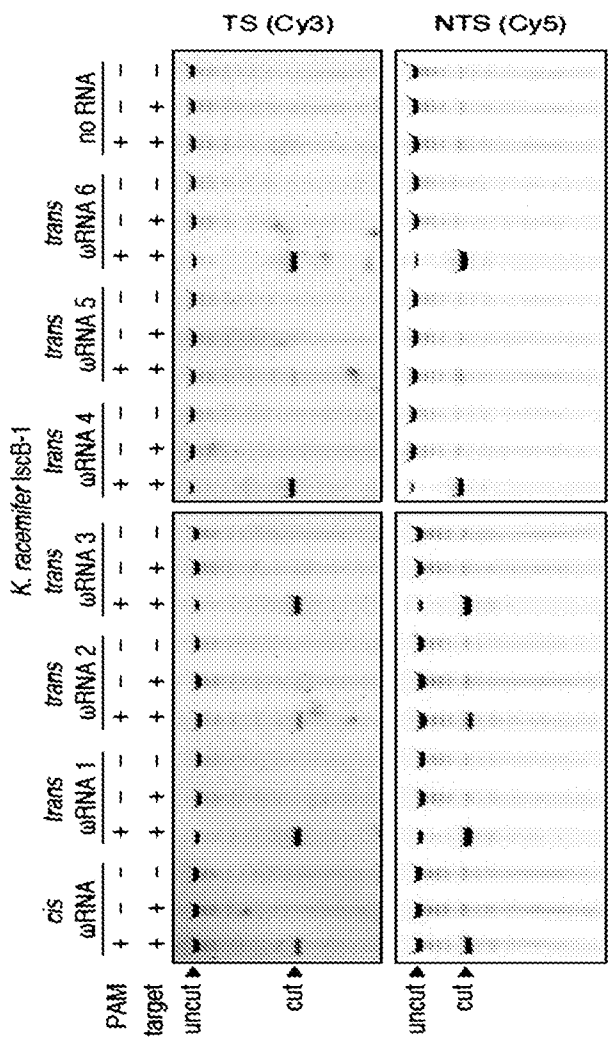


FIG. 16D

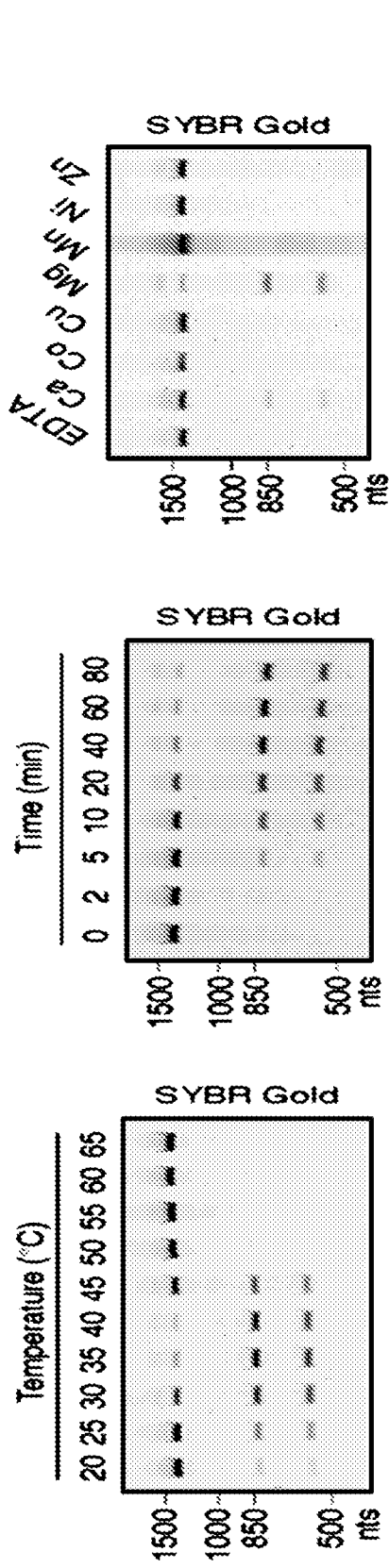


FIG. 17A

FIG. 17B

FIG. 17C

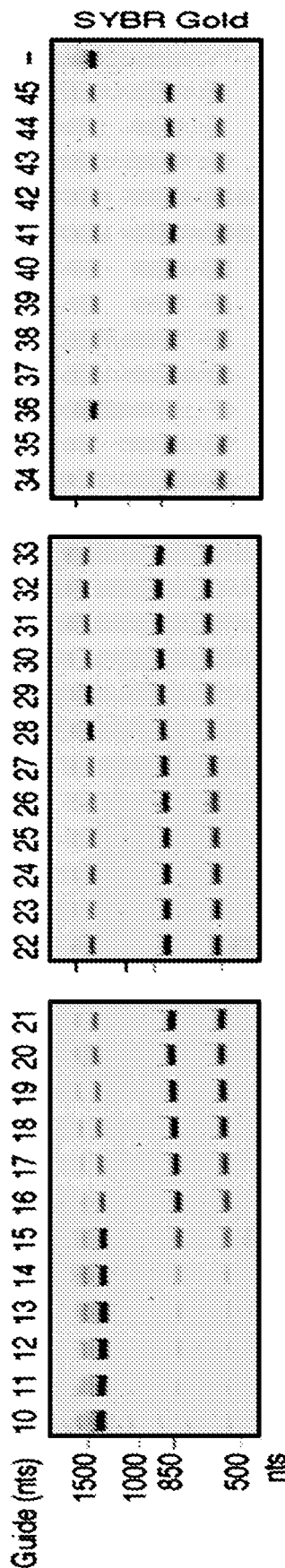


FIG. 17D

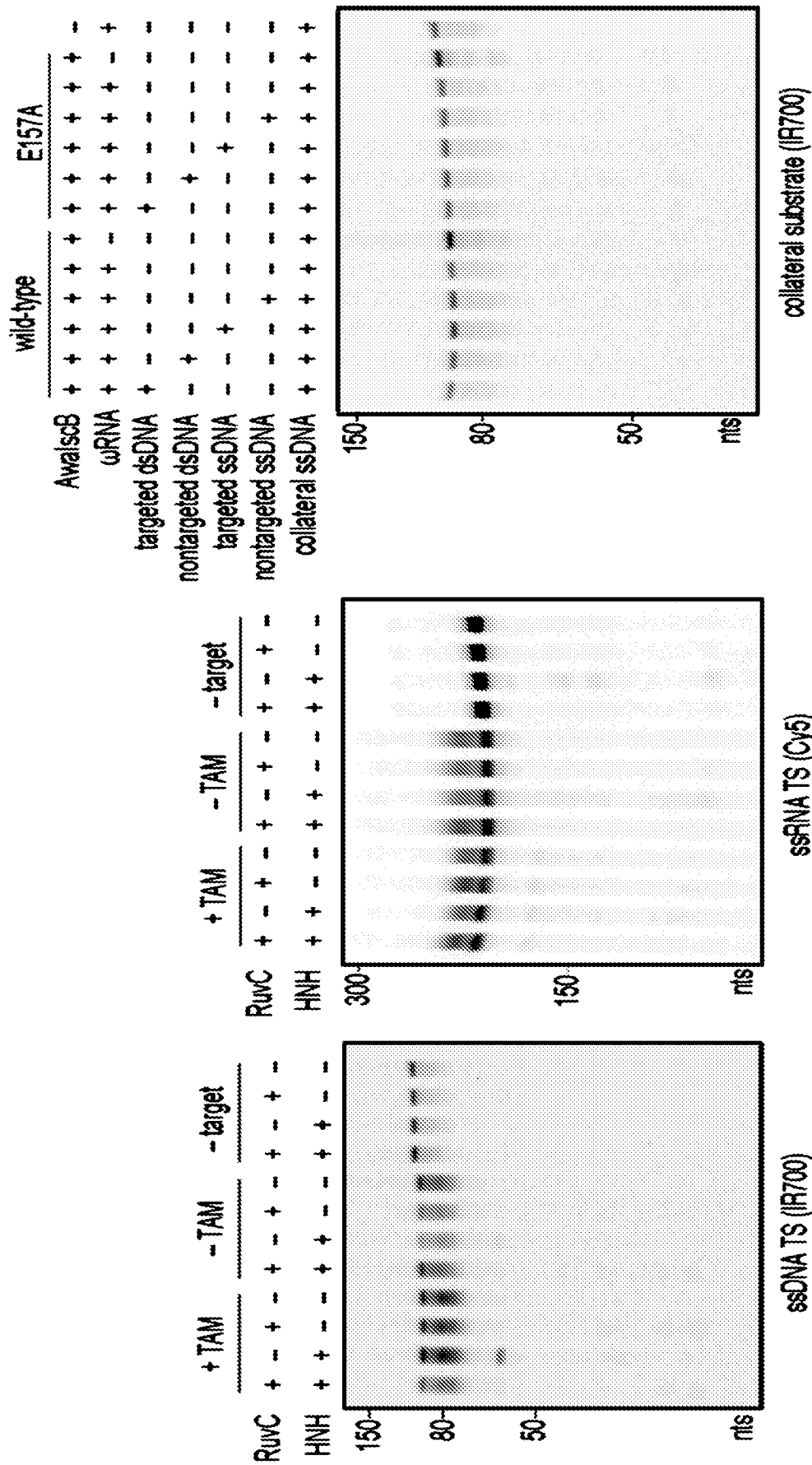


FIG. 17G

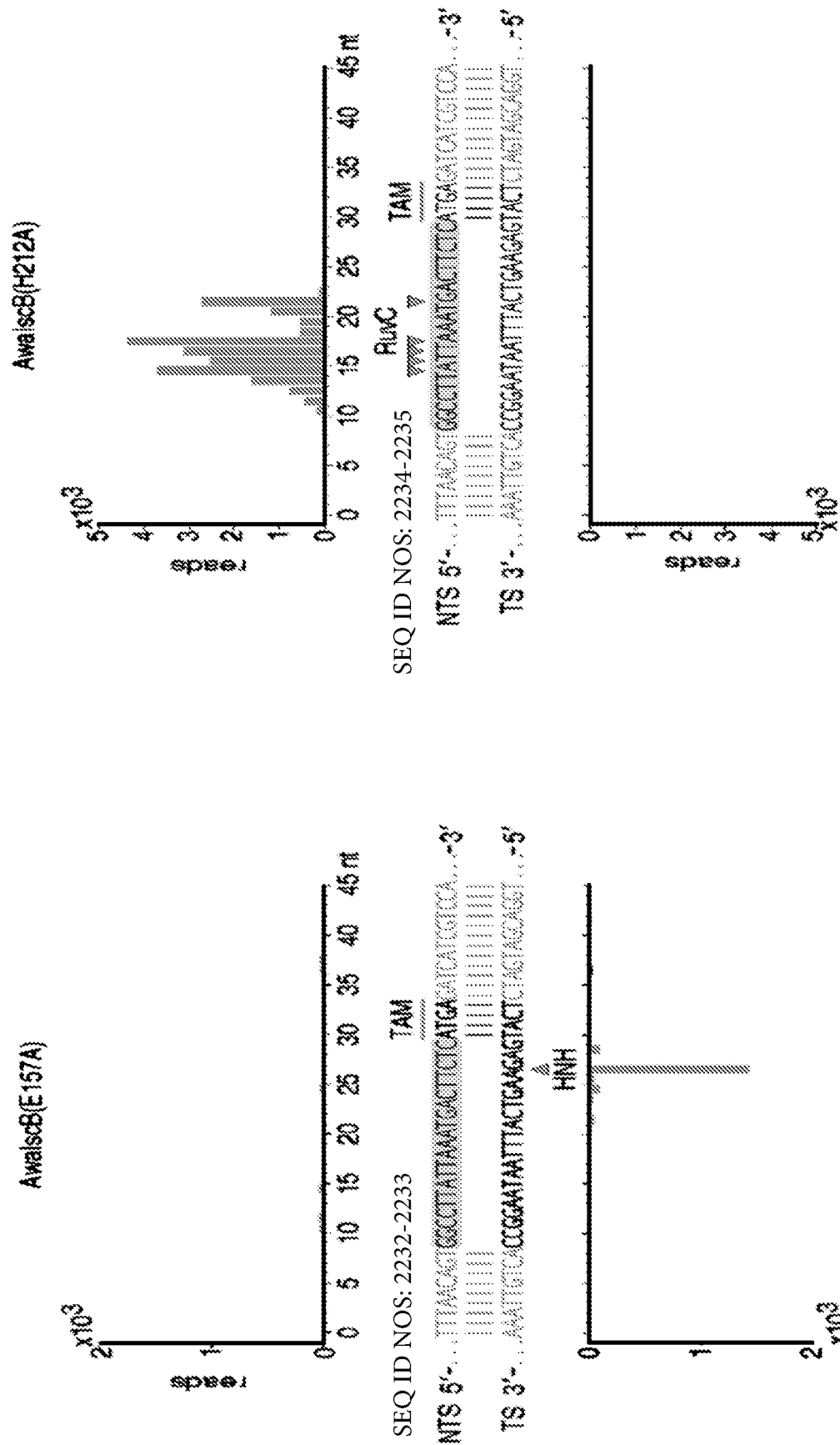


FIG. 18A

FIG. 18B

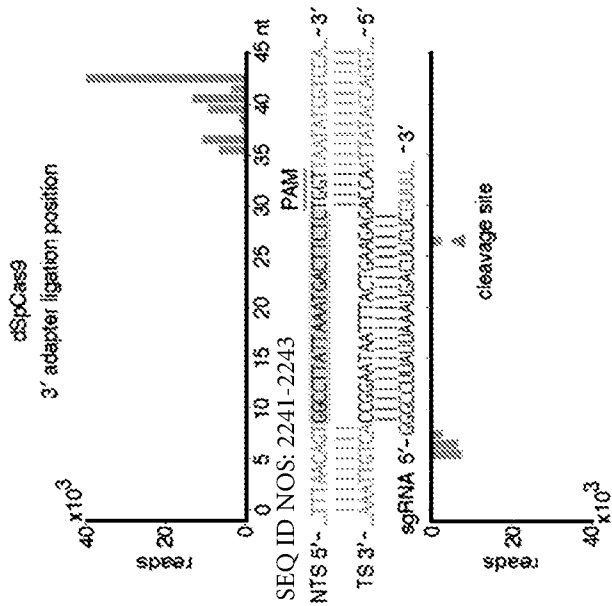


FIG. 19A

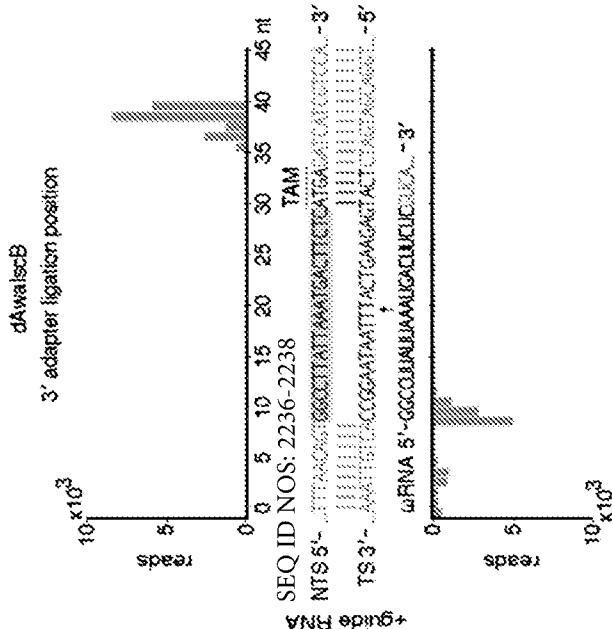


FIG. 19B

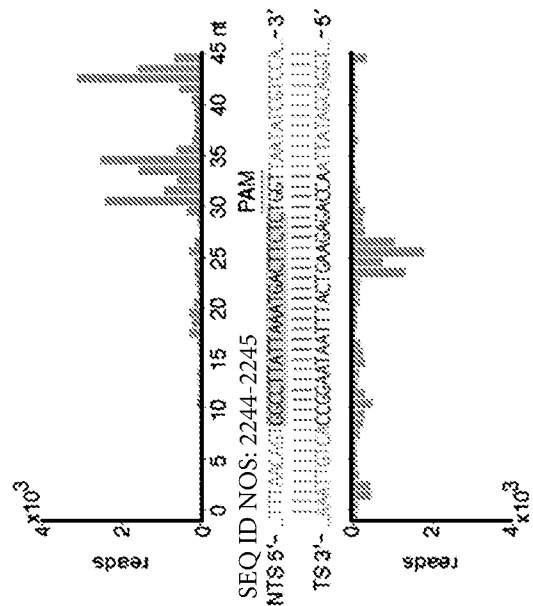


FIG. 19D

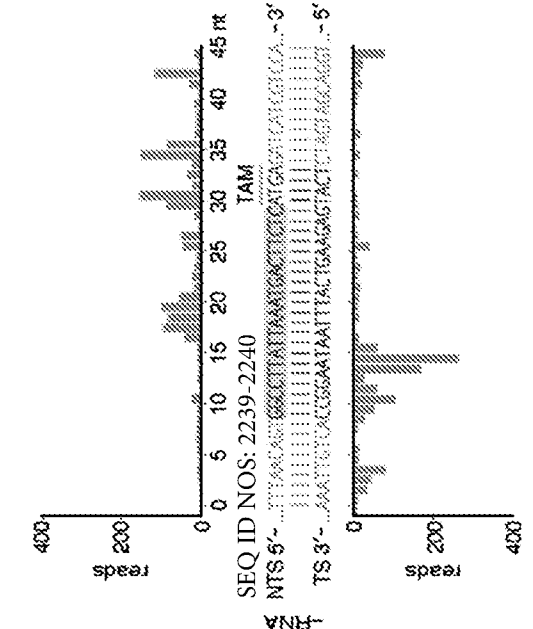


FIG. 19C

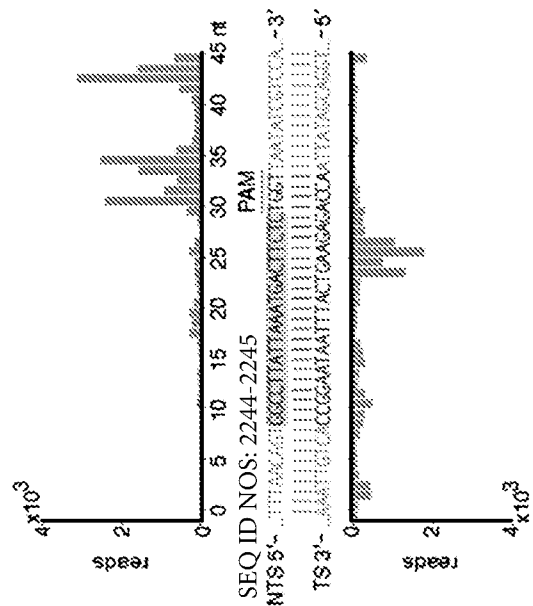


FIG. 19E

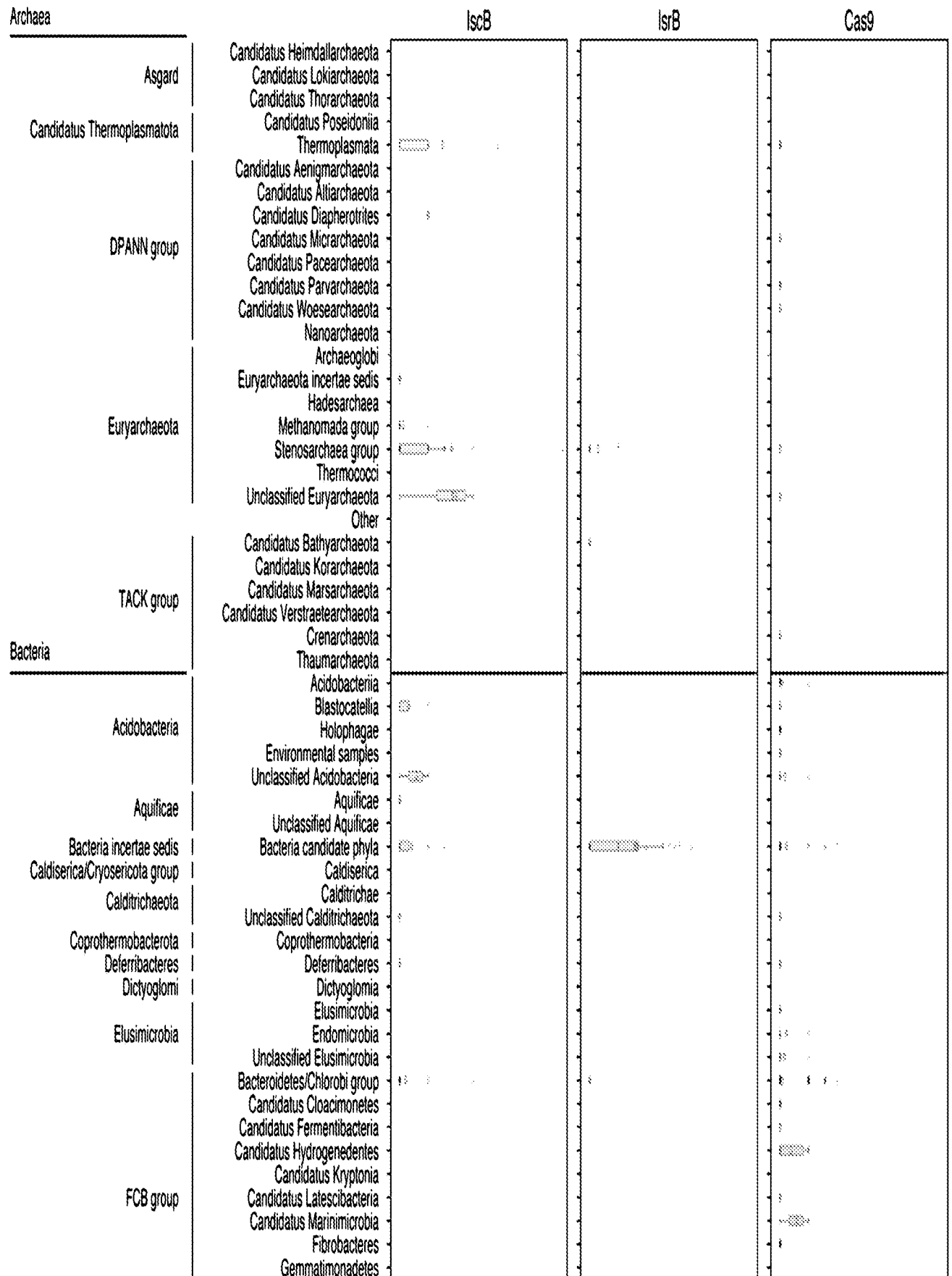


FIG. 20

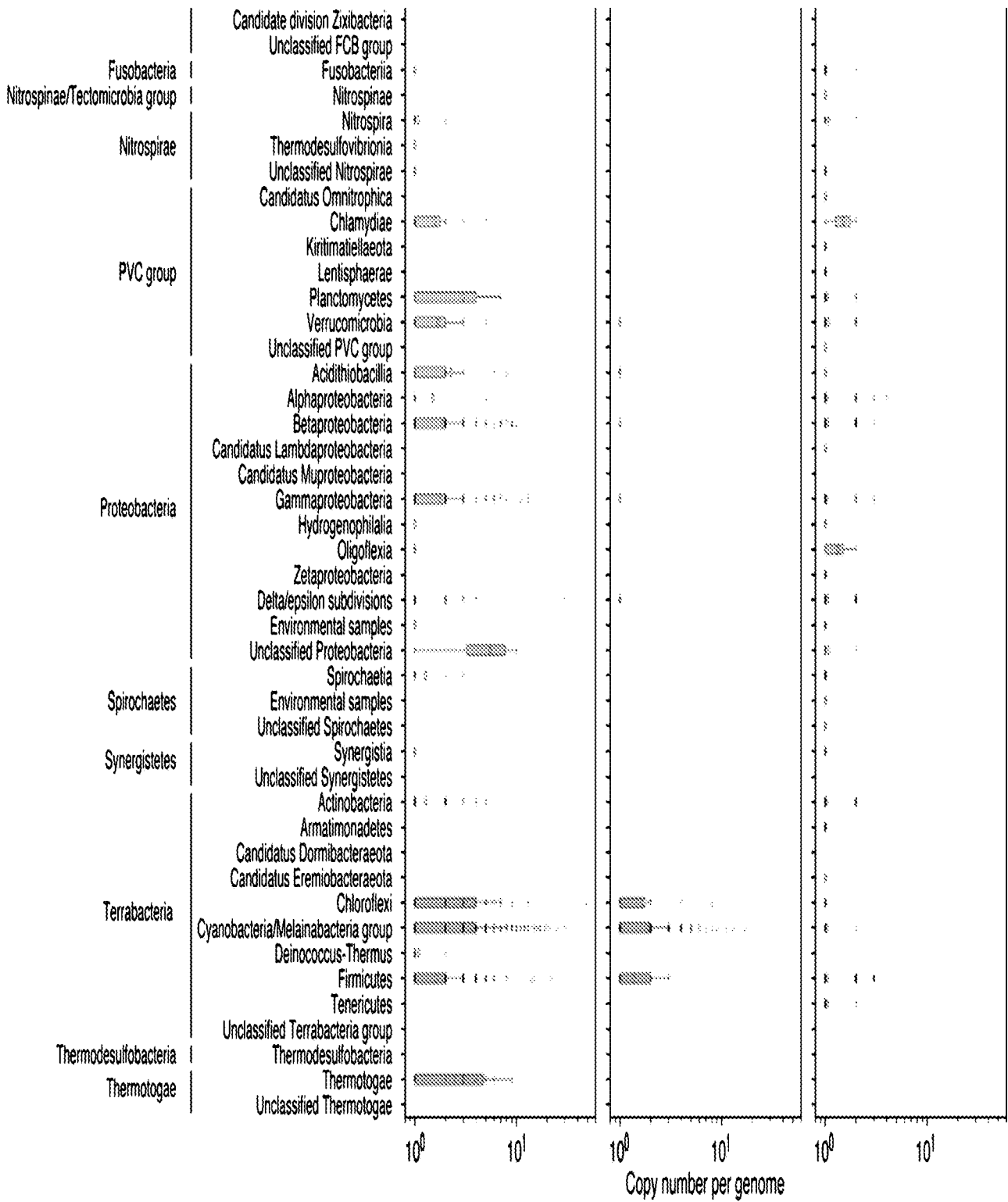


FIG. 20 (continued)

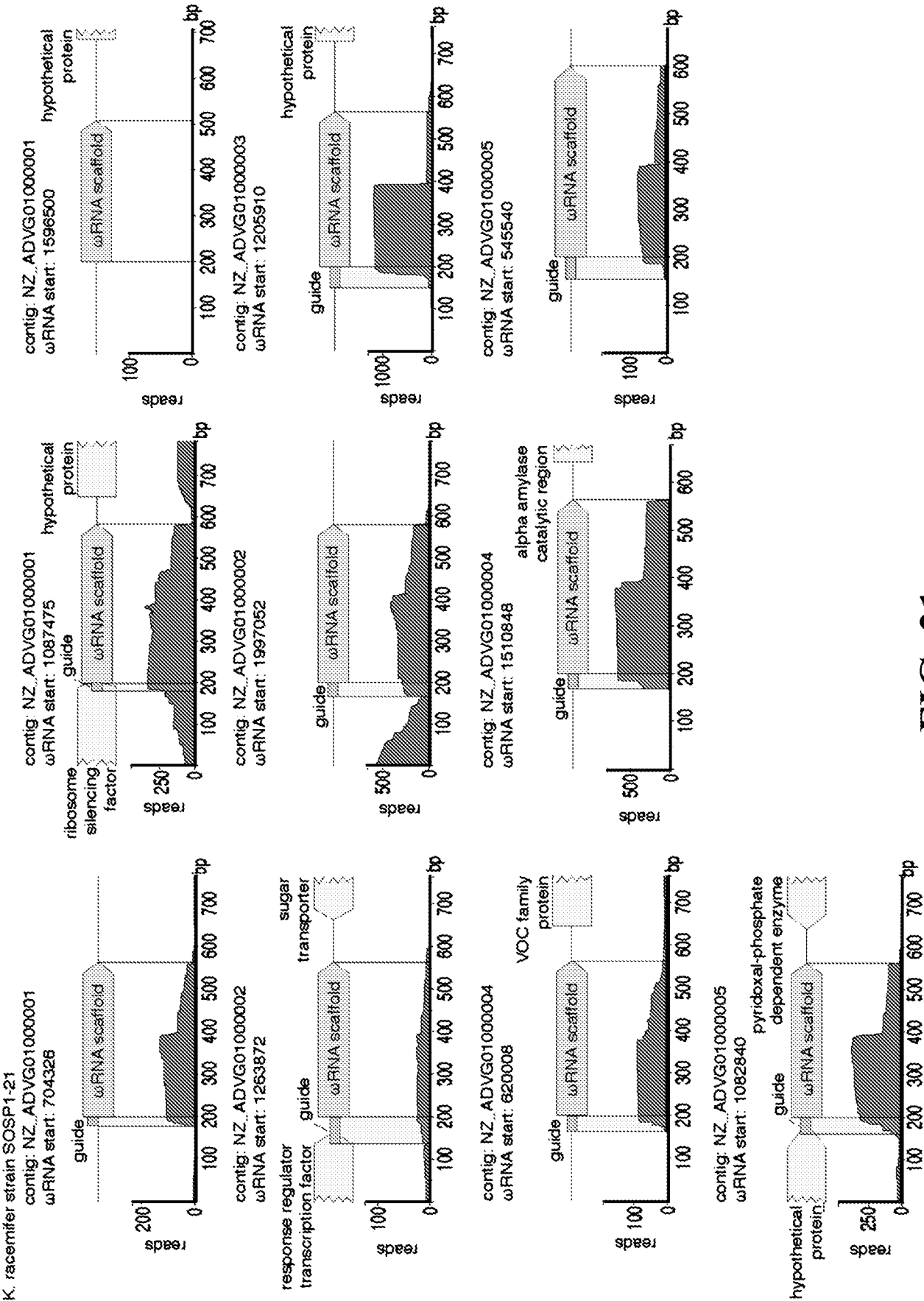
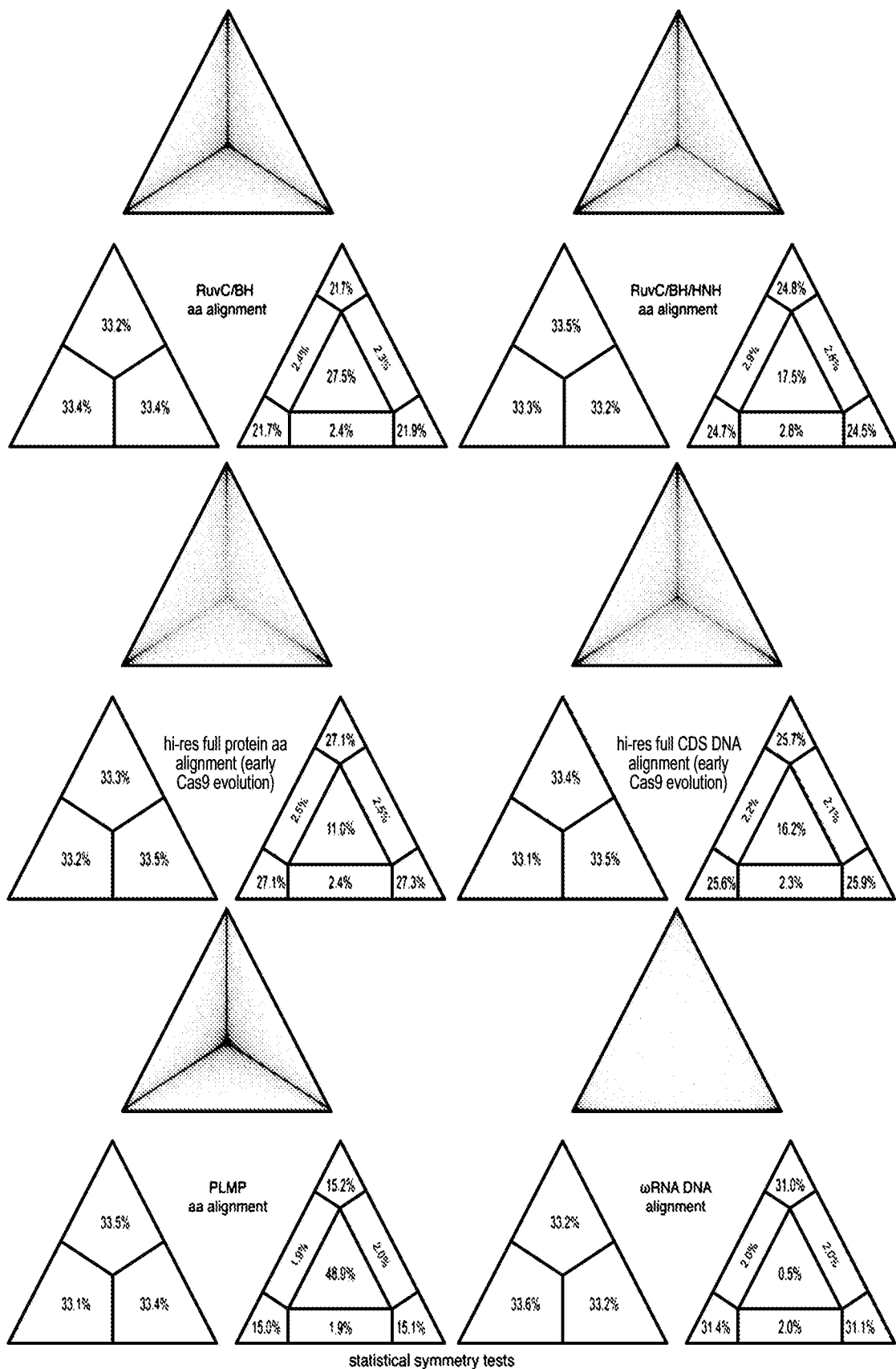


FIG. 21

**FIG. 22A**

alignment	significant pairs (sym)	non-significant pairs (sym)	sym P-val	significant pairs (mar)	non-significant pairs (mar)	mar P-val	significant pairs (int)	non-significant pairs (int)	int P-val
RuvC/BH aa	9554	1706324	0.22104	331597	1229041	0.0420983	1584	1559054	0.539489
RuvC/BH/HNH aa	45493	1415702	0.257994	520615	921462	0.000602025	1705	1440372	0.919923
RuvC/BH/HNH aa (no II-B)	43939	1351346	0.307609	498062	879306	0.0308016	1633	1375735	0.657229
RuvC/BH/HNH aa (early Cas9 + IscB)	9667	260813	0.396151	101356	162044	0.227067	293	263107	0.506638
hi-res full protein aa (early Cas9)	1566	6561	0.291625	5077	2983	0.835659	41	8019	0.181417
hi-res full CDS DNA (early Cas9)	7203	925	0	7396	731	0	616	7511	0.404016
PLMP aa	3	2100134	0.514766	1379	280476	0.411688	1647	280182	0.563905
ωRNA DNA	58169	194235	0.241292	71960	180443	0.508349	9161	243239	0.130791

FIG. 22B

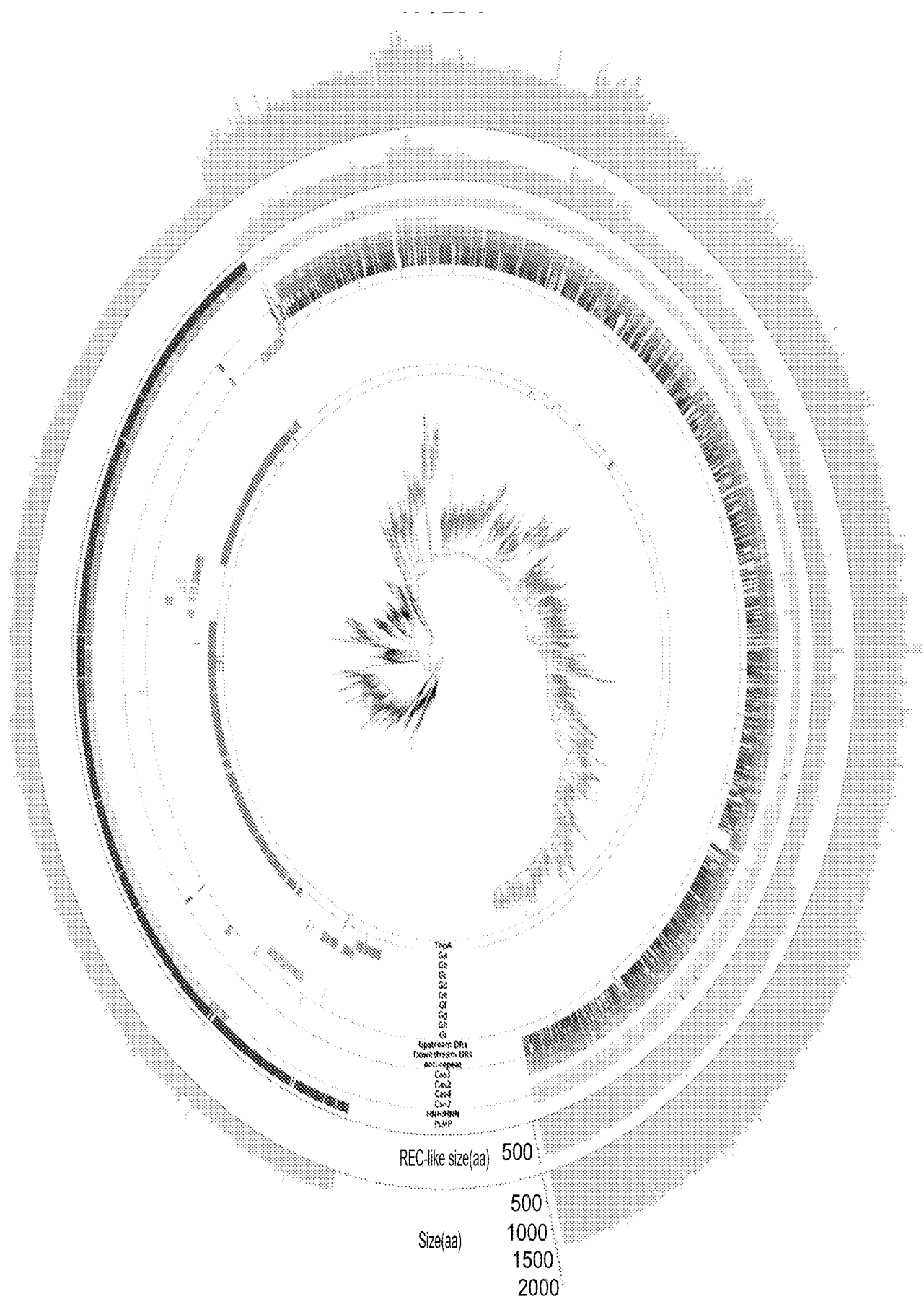


FIG. 23

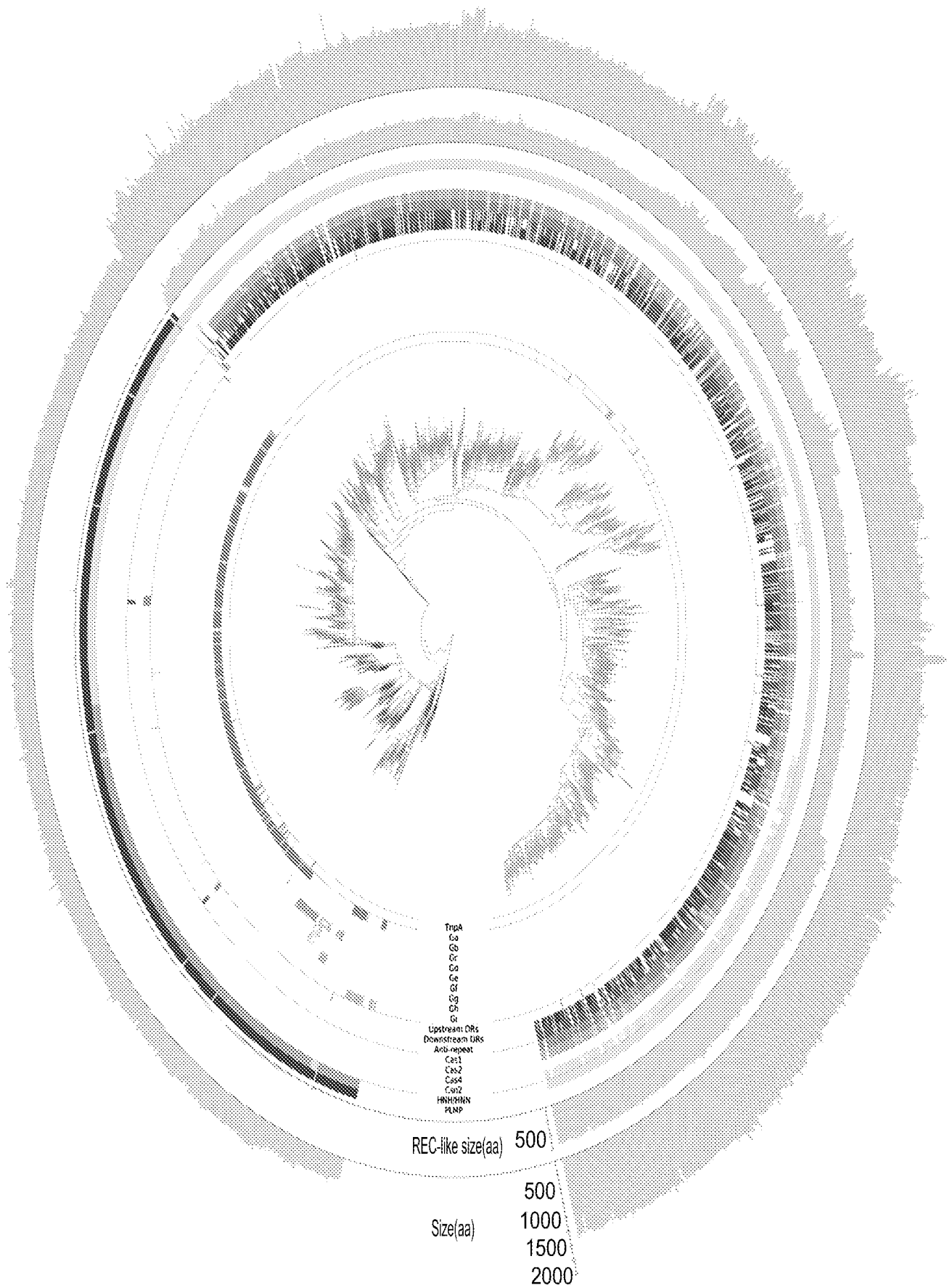


FIG. 24

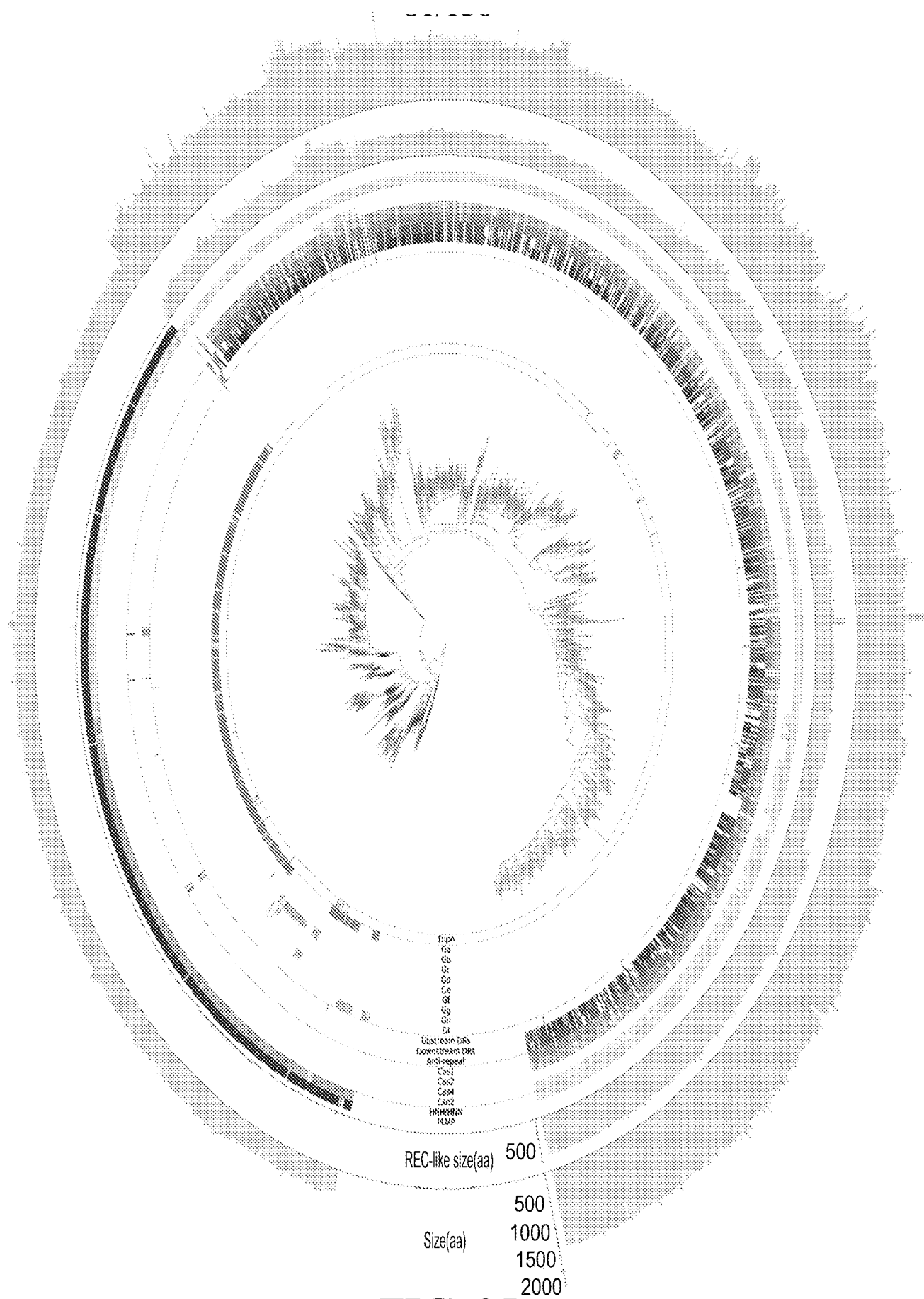


FIG. 25

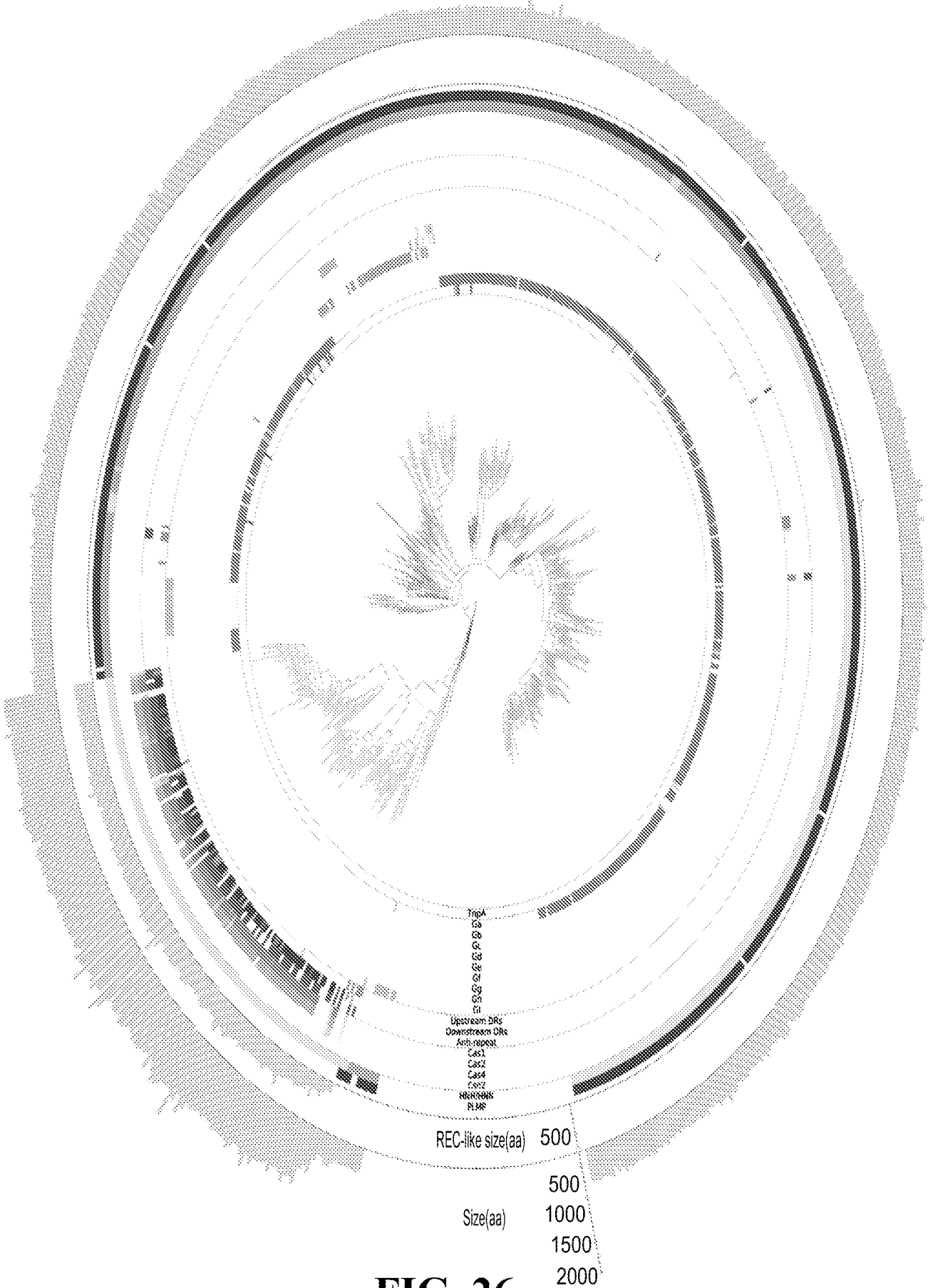


FIG. 26

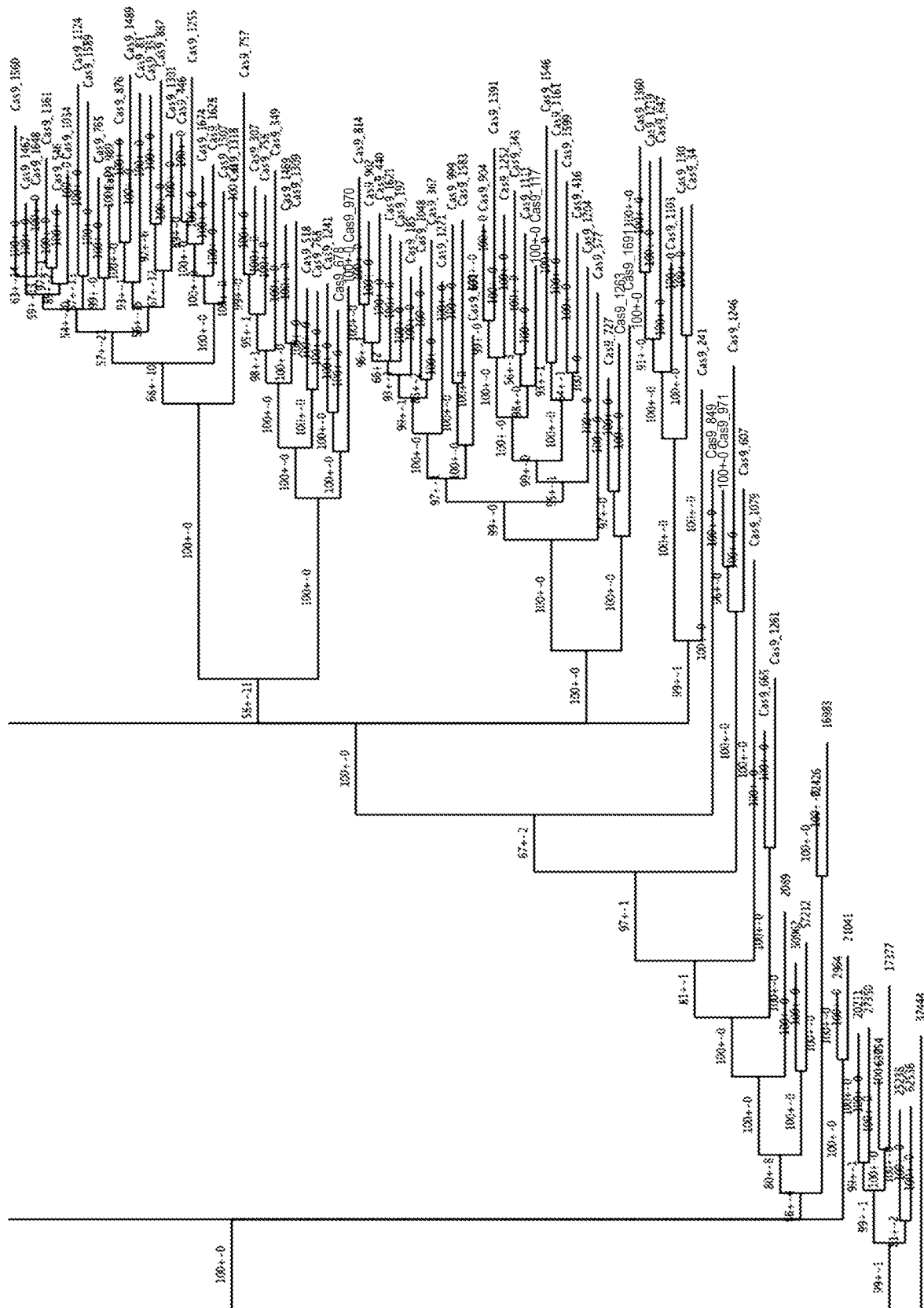


FIG. 27



0.4

FIG. 28

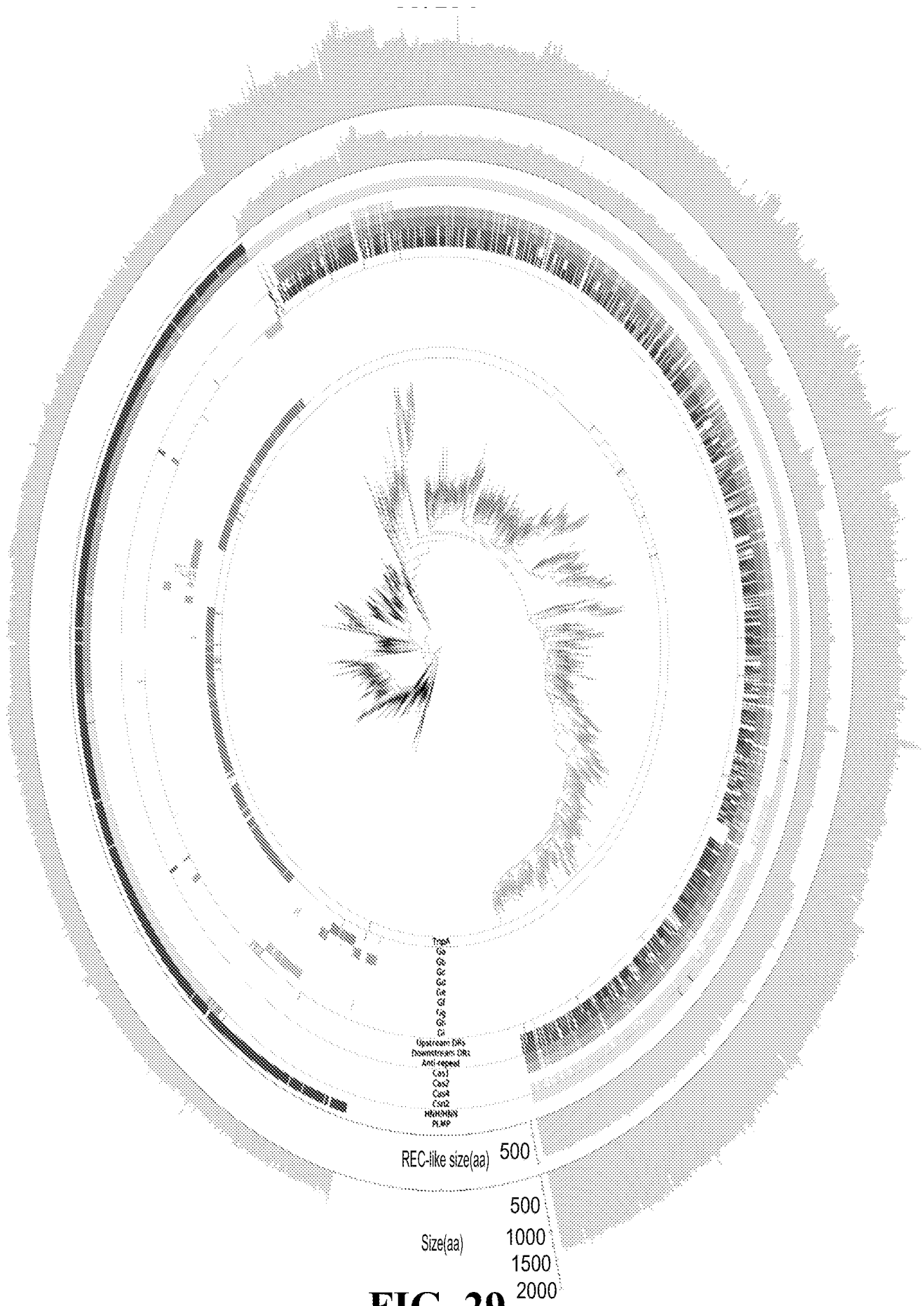


FIG. 29

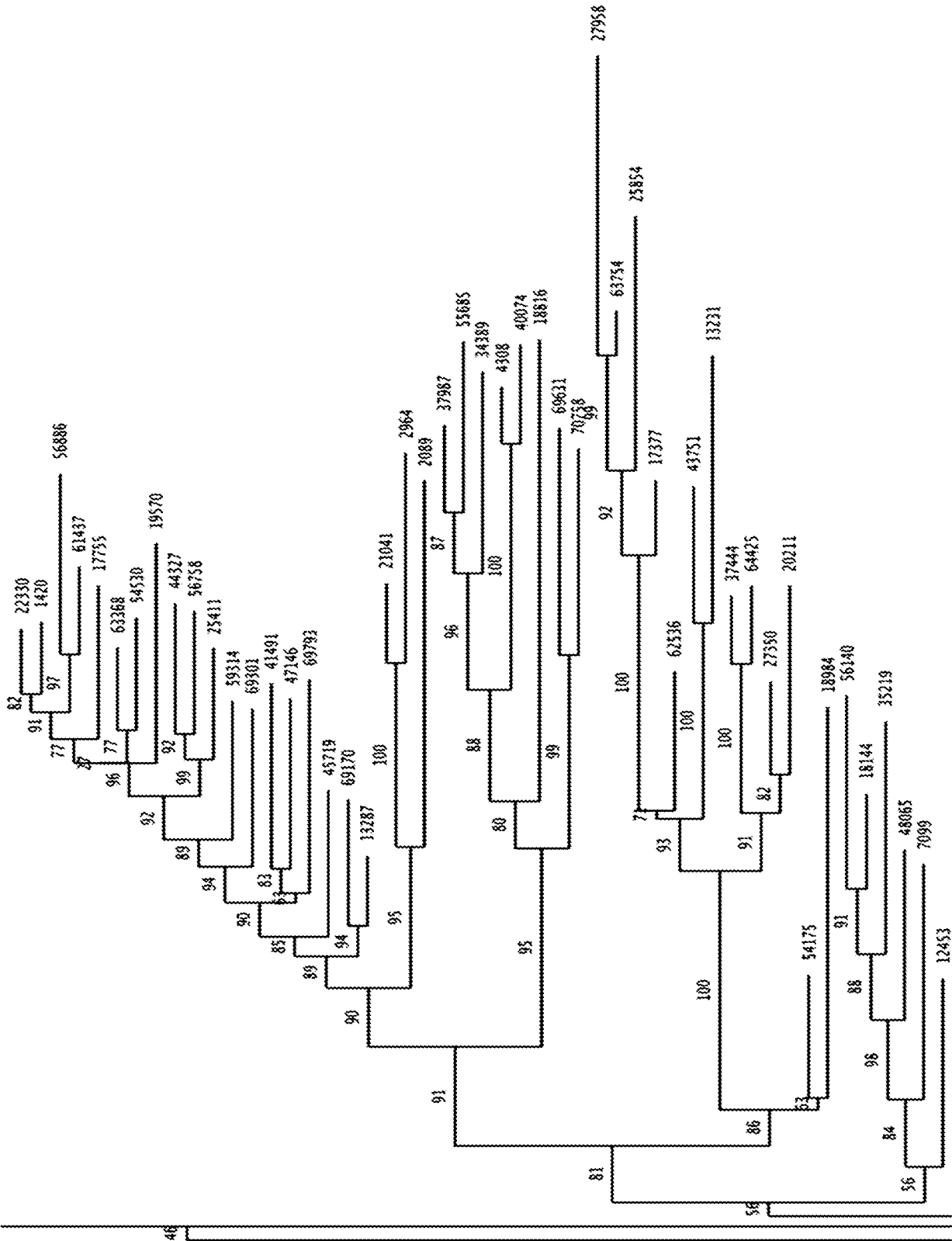


FIG. 30

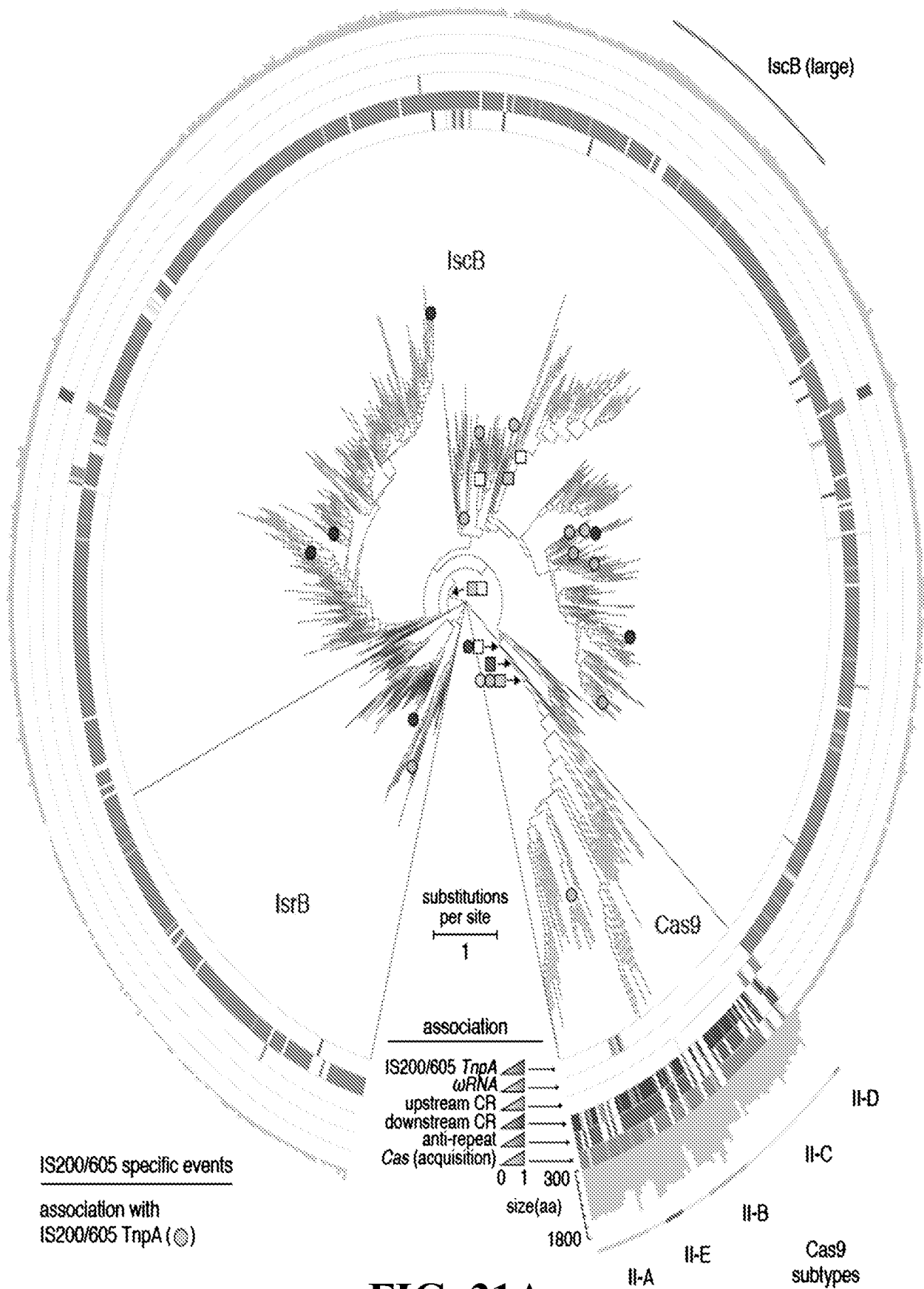


FIG. 31A

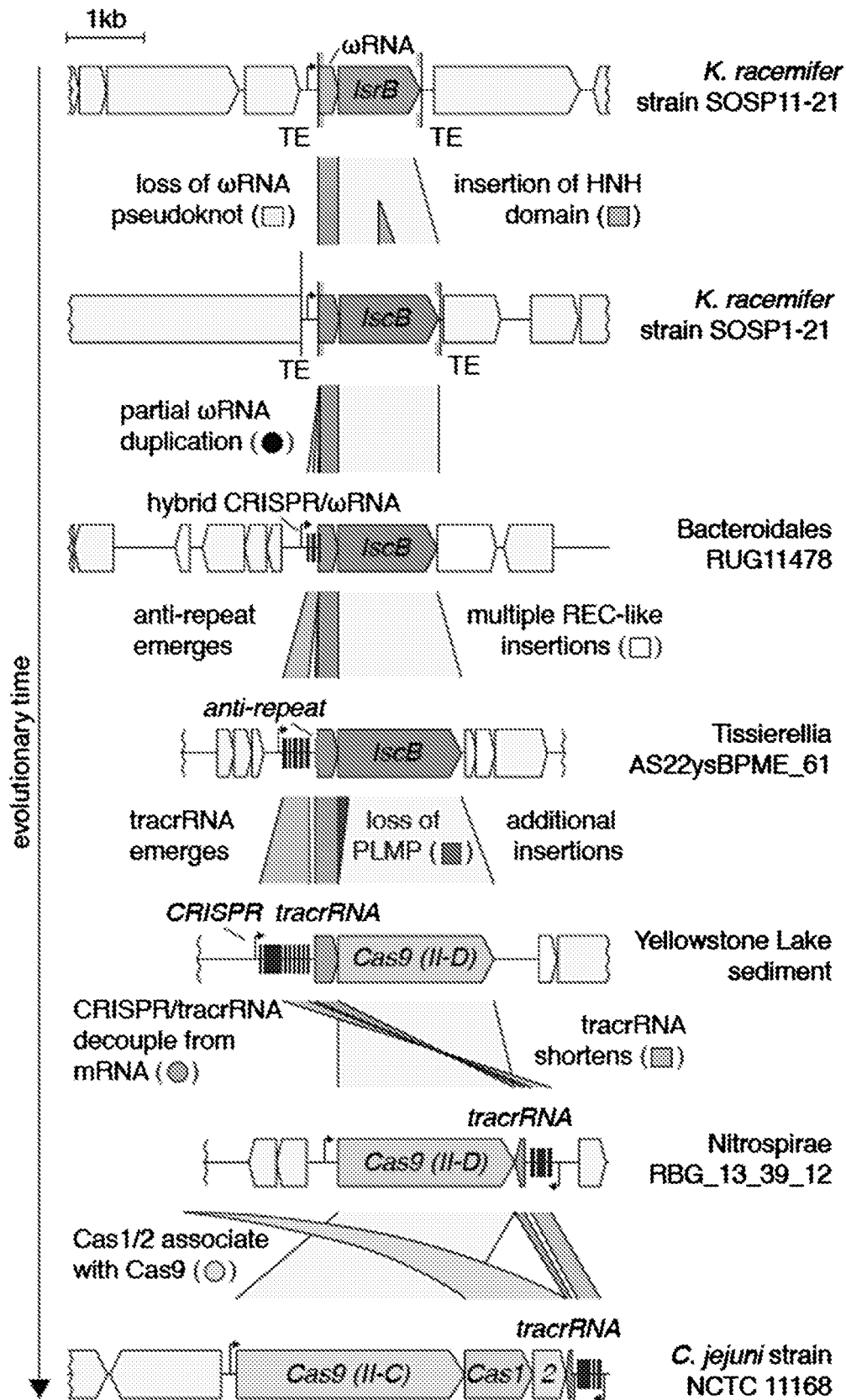


FIG. 31B

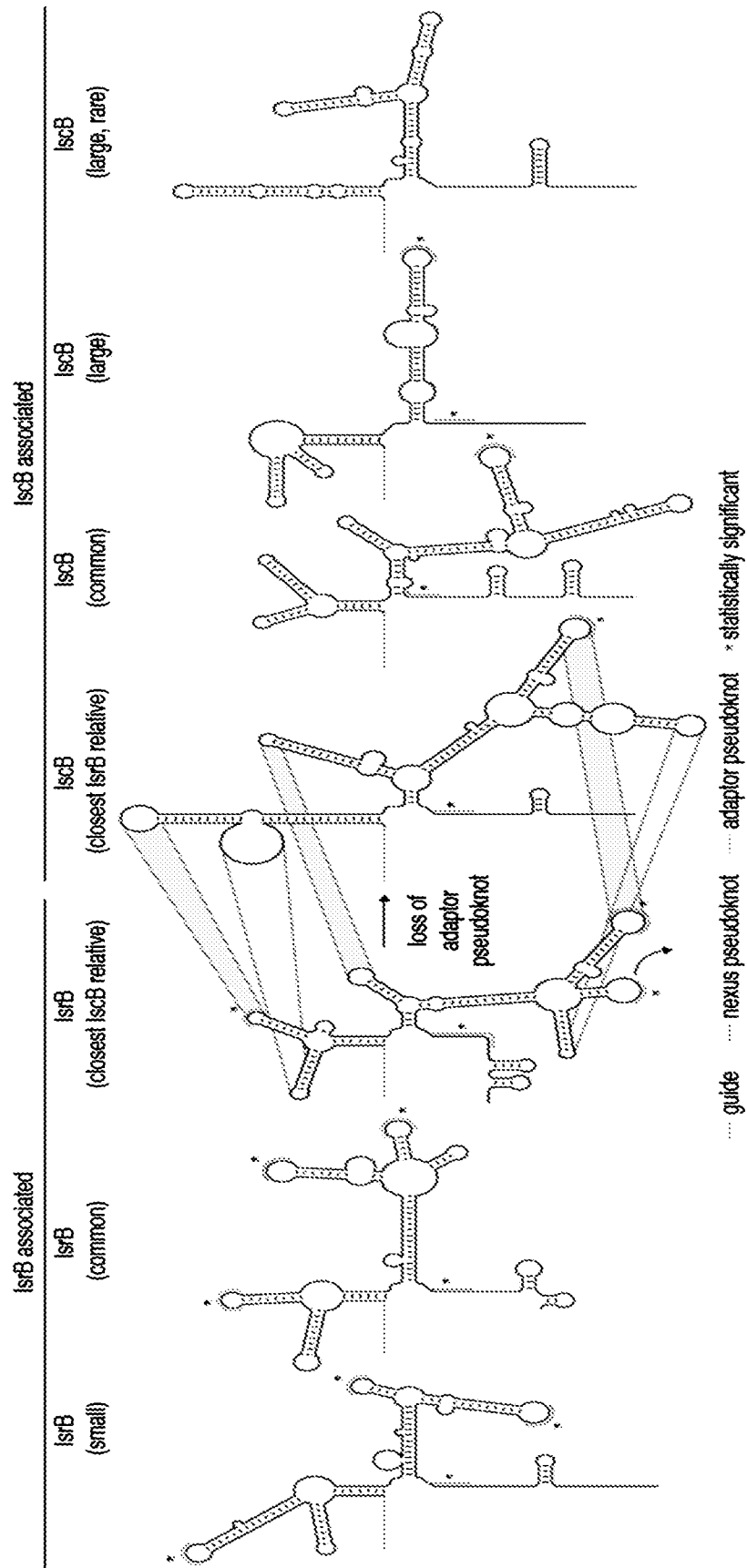


FIG. 31C

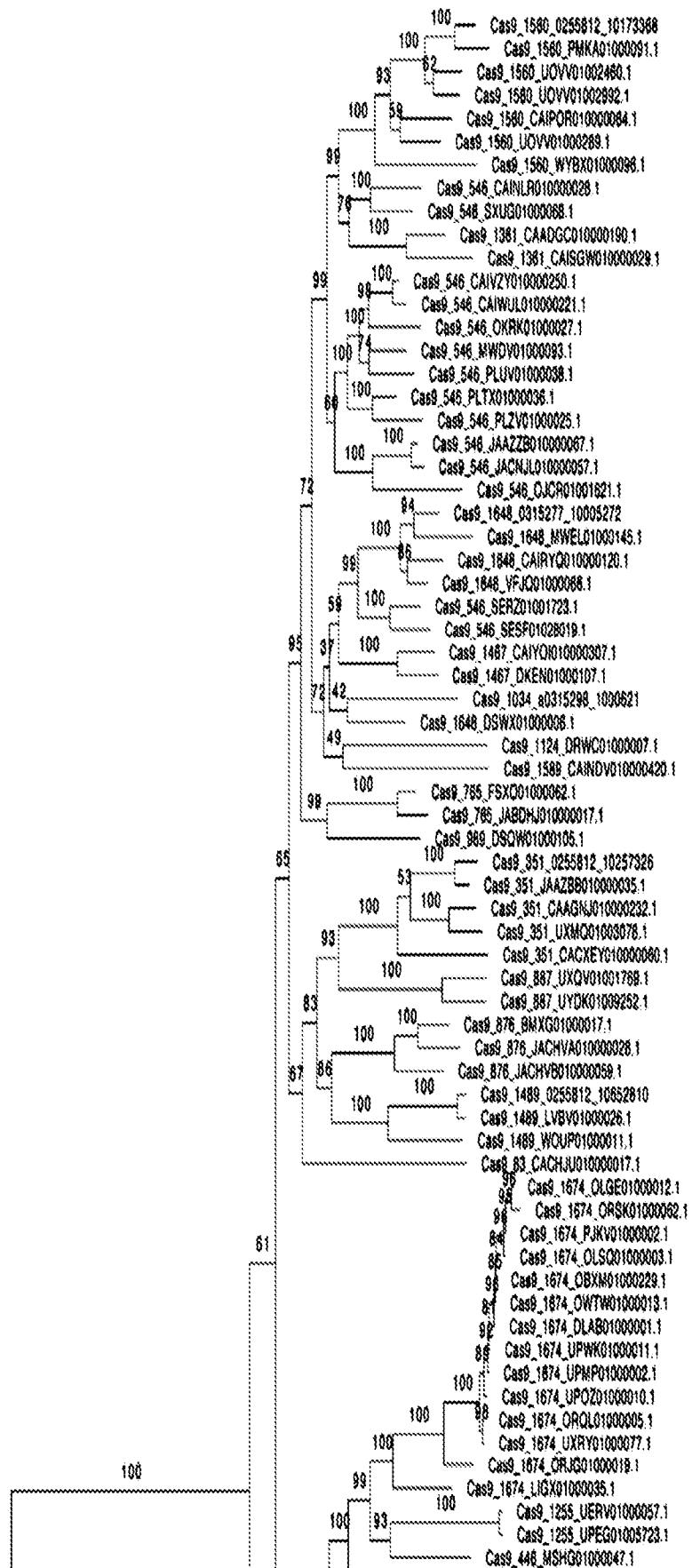
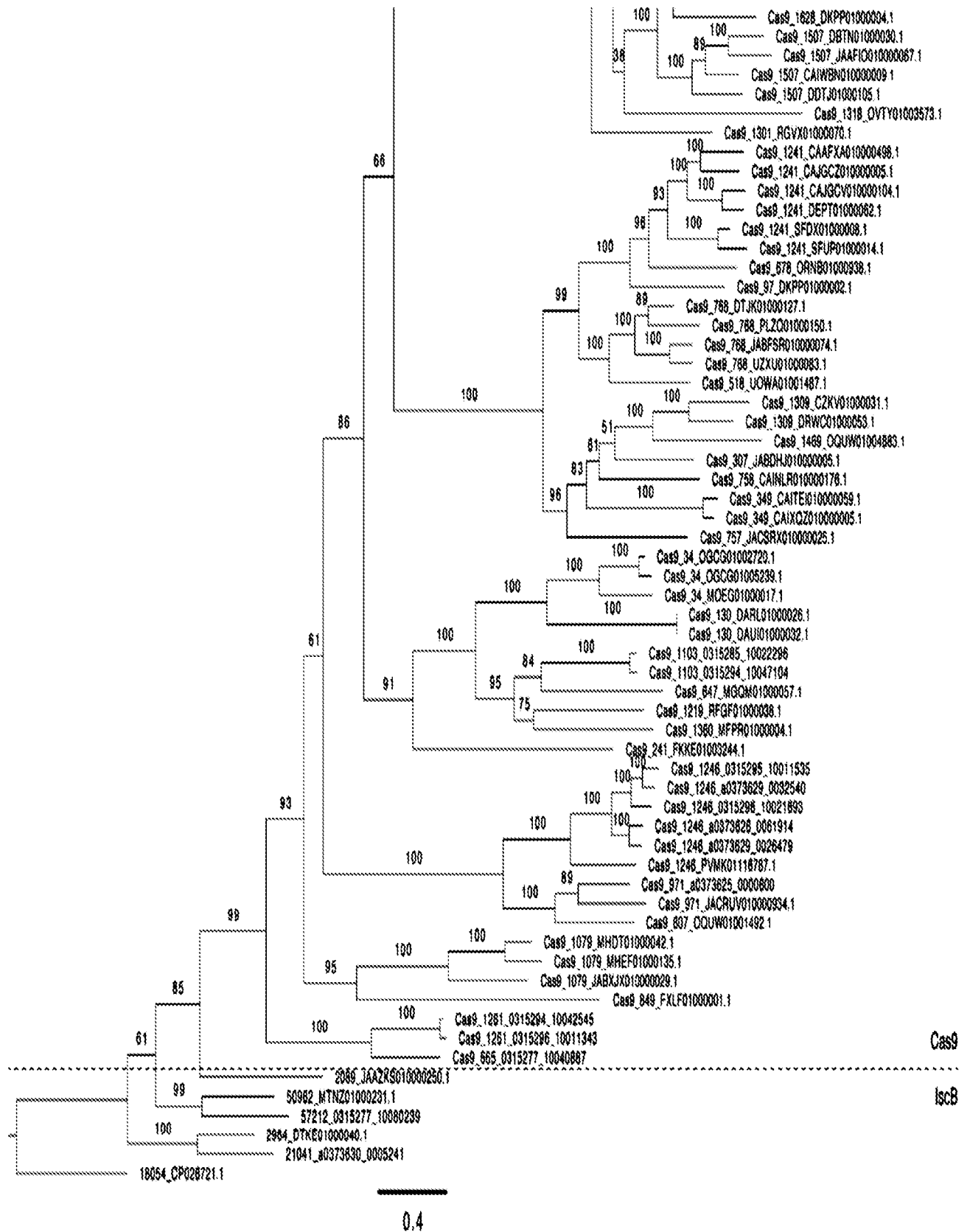


FIG. 32A

**FIG. 32A (continued)**

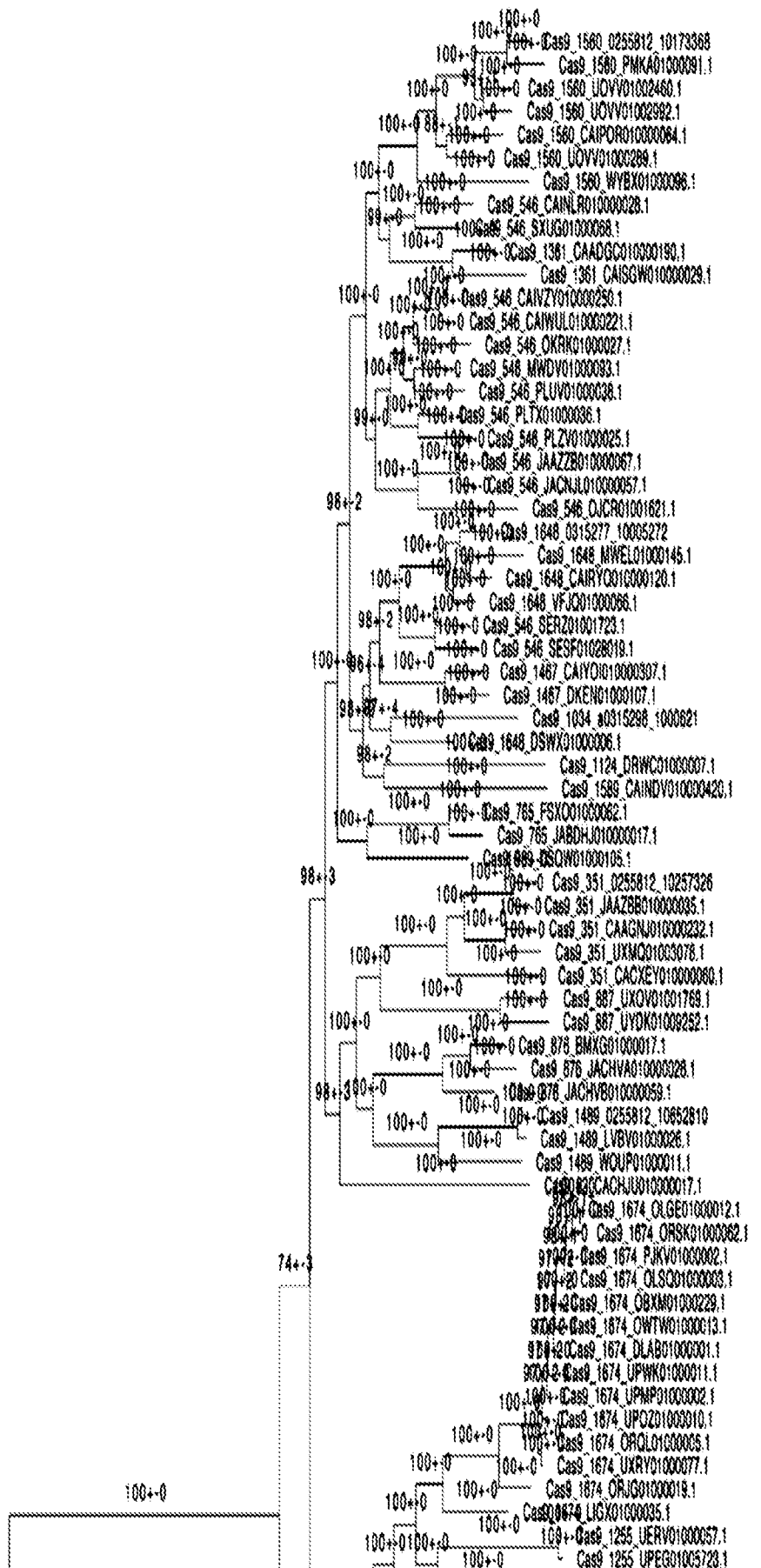


FIG. 32B

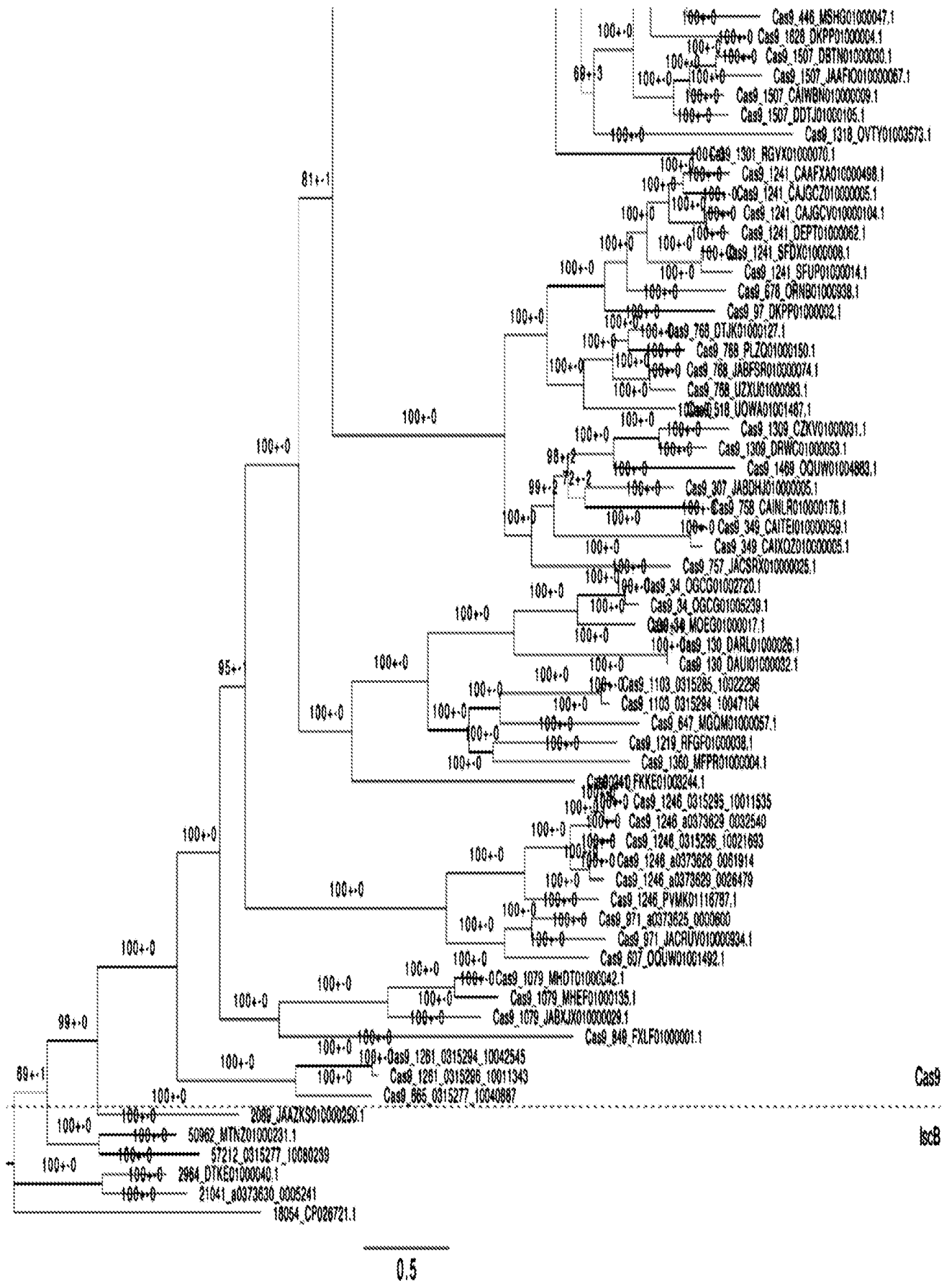


FIG. 32B (continued)

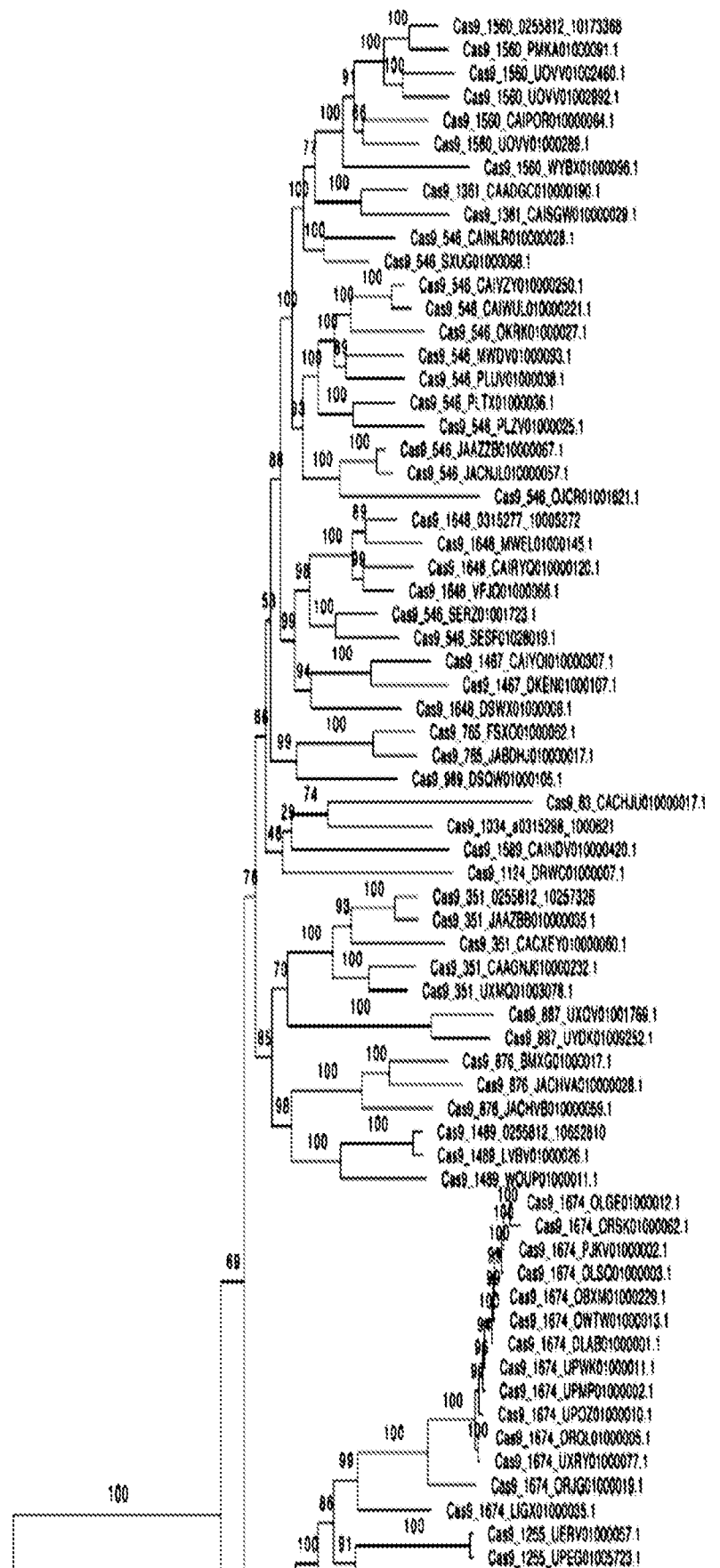
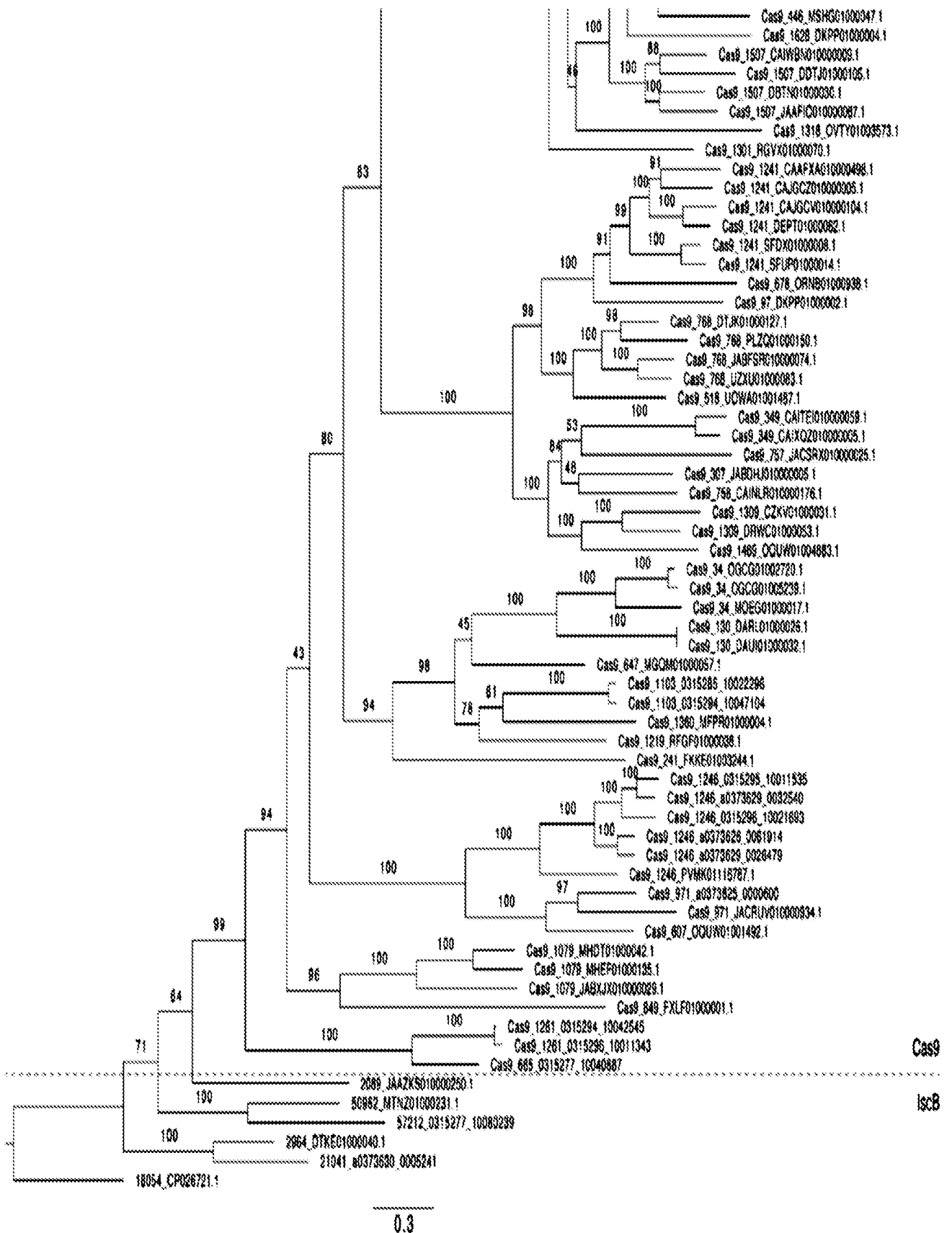


FIG. 33A

**FIG. 33A (continued)**

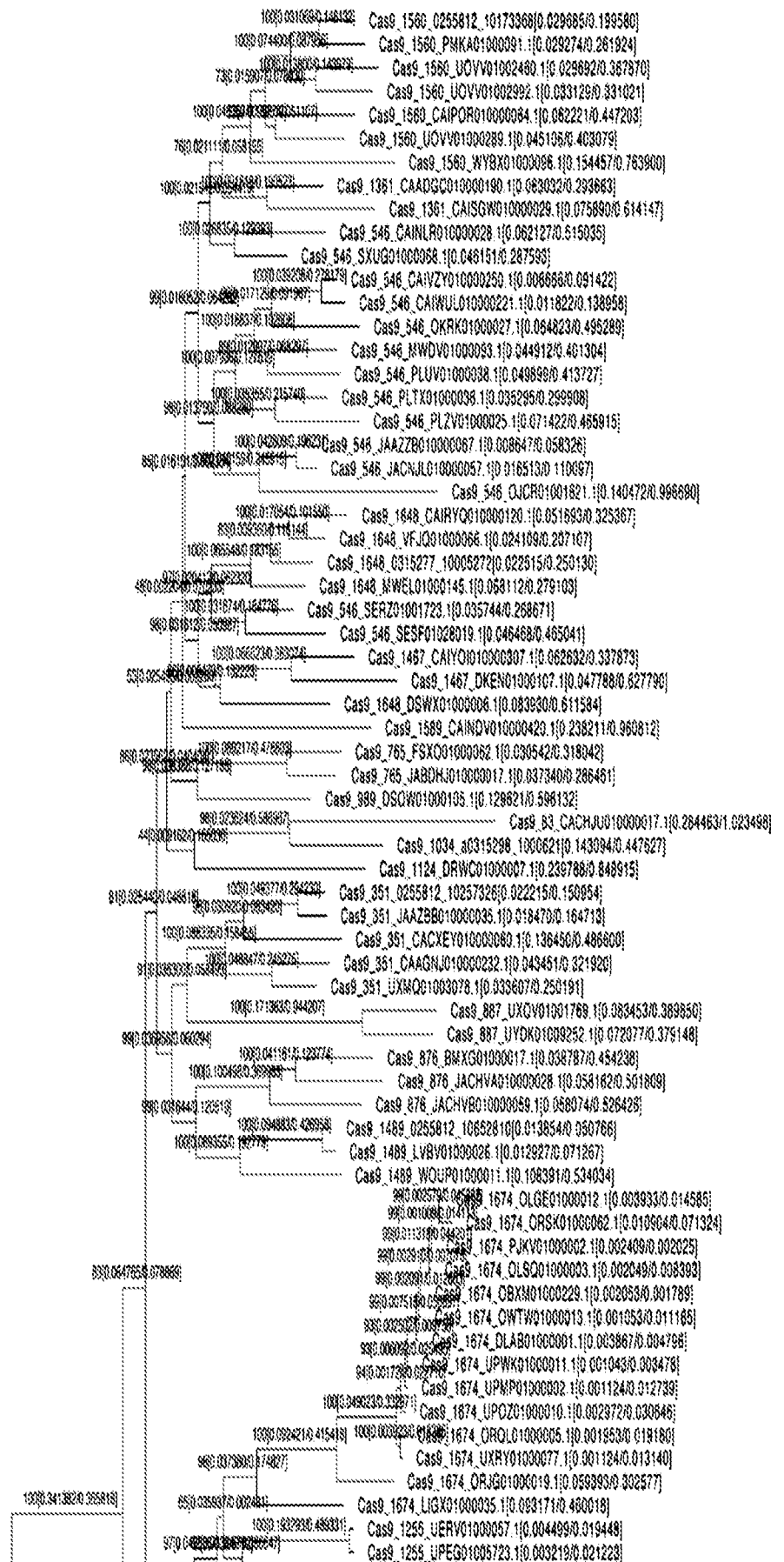


FIG. 33B

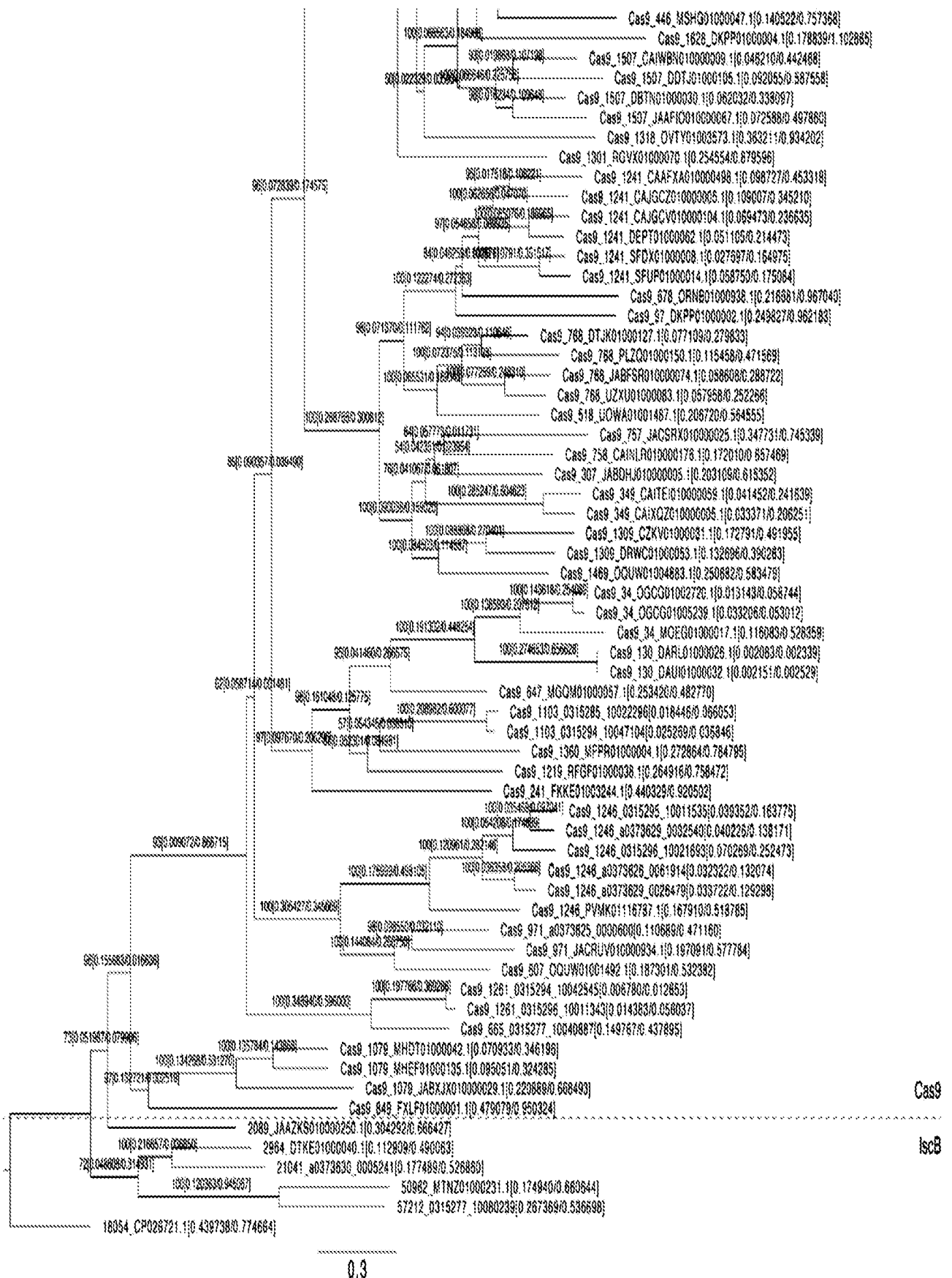


FIG. 33B (continued)

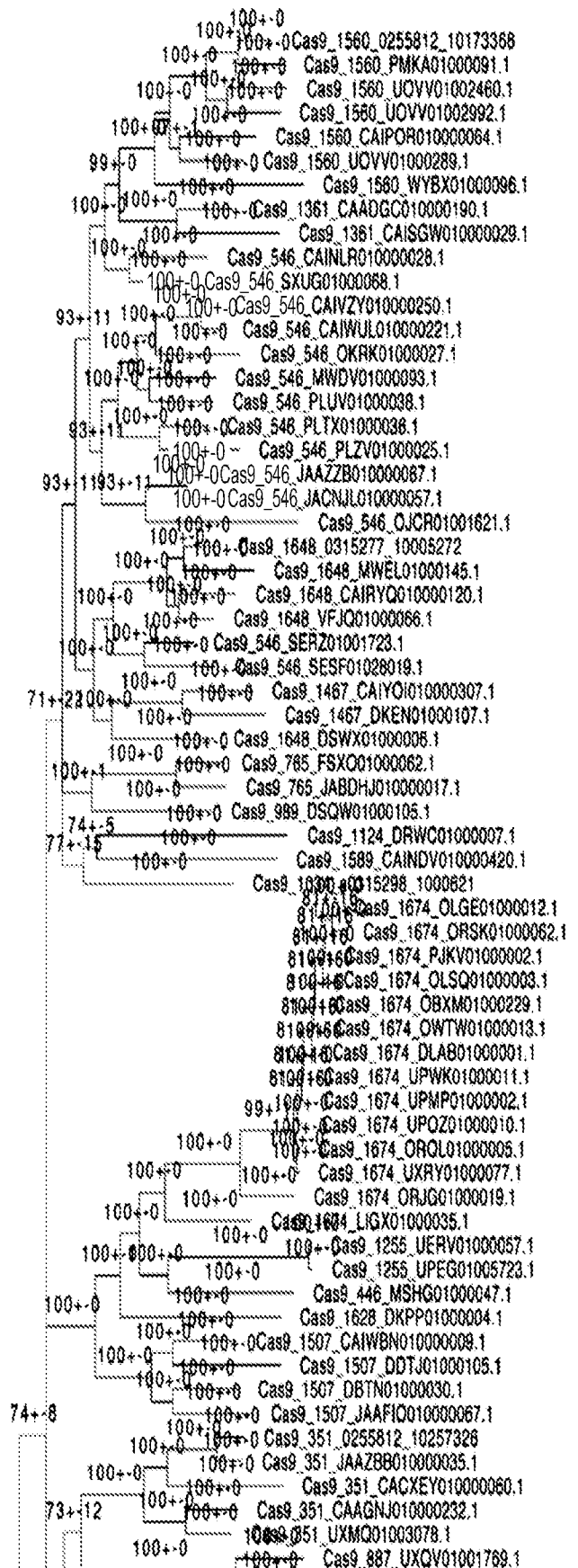
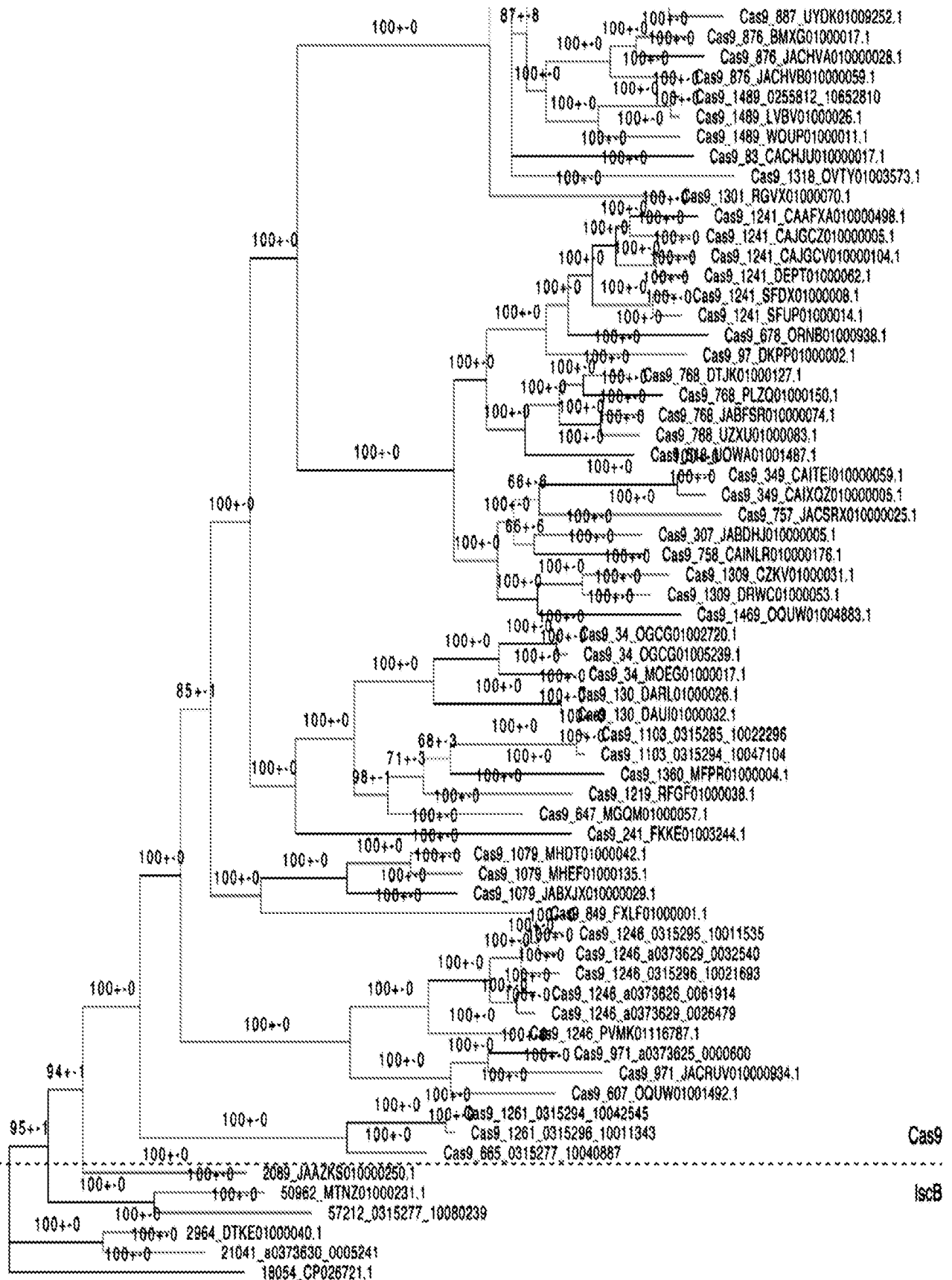


FIG. 33C



0.4

FIG. 33C (continued)

RuvC/BH/HNH
early IscB

LG
+ Gamma(4)

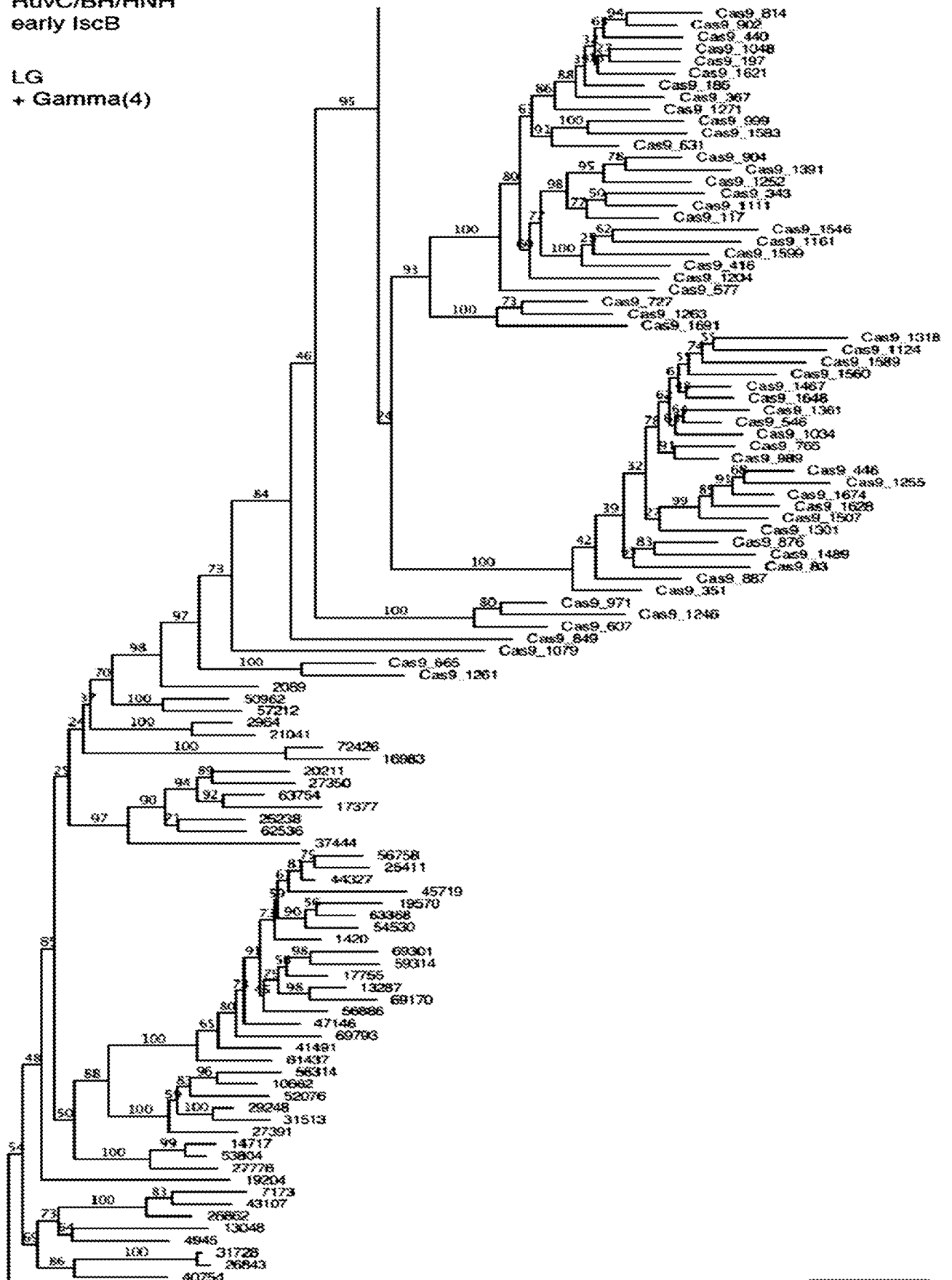


FIG. 34A

0.5

RuvC/BH/HNH
early IscB

LG

+ invariant sites
+ Gamma(4)

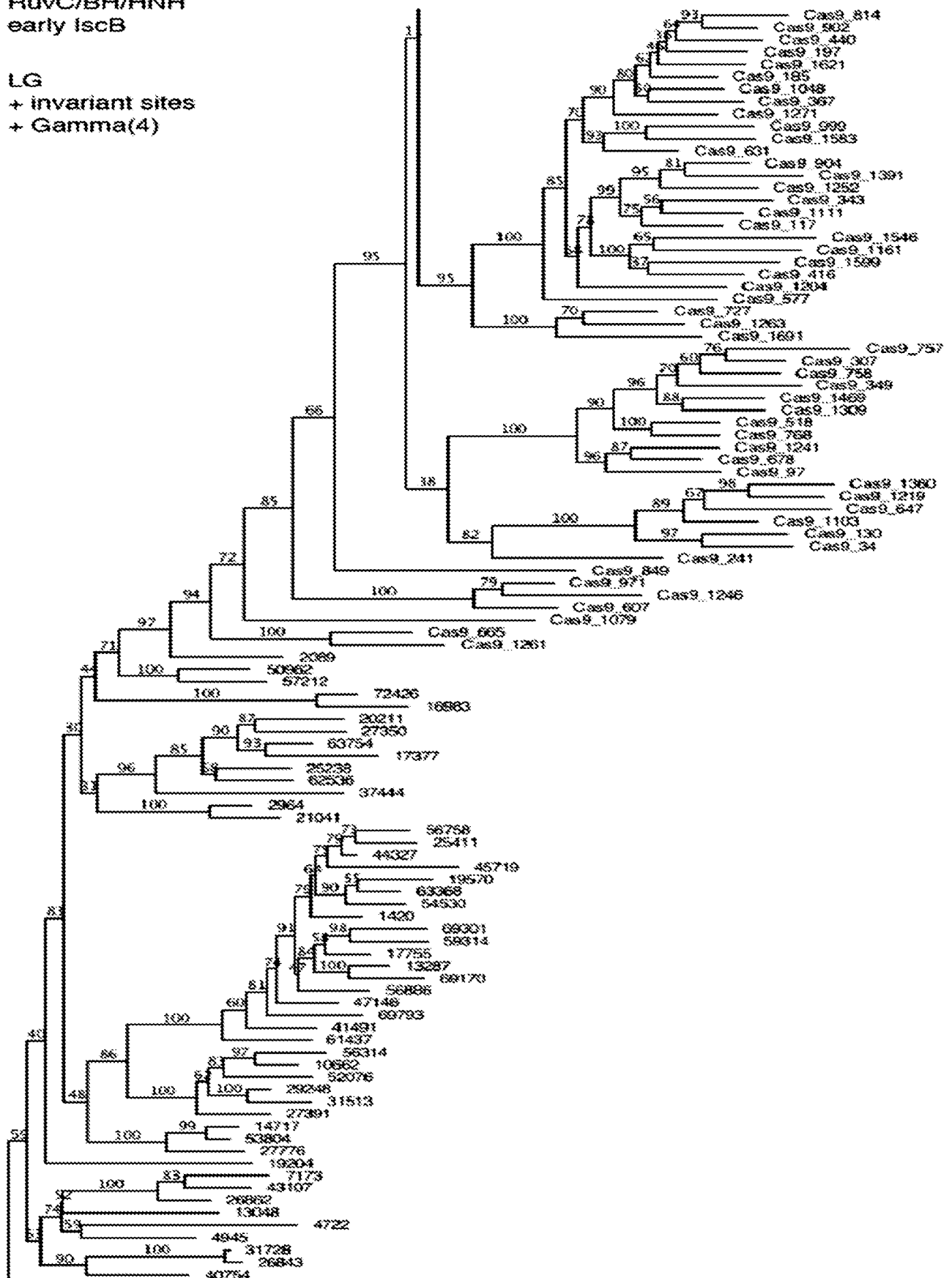


FIG. 34B

0.5

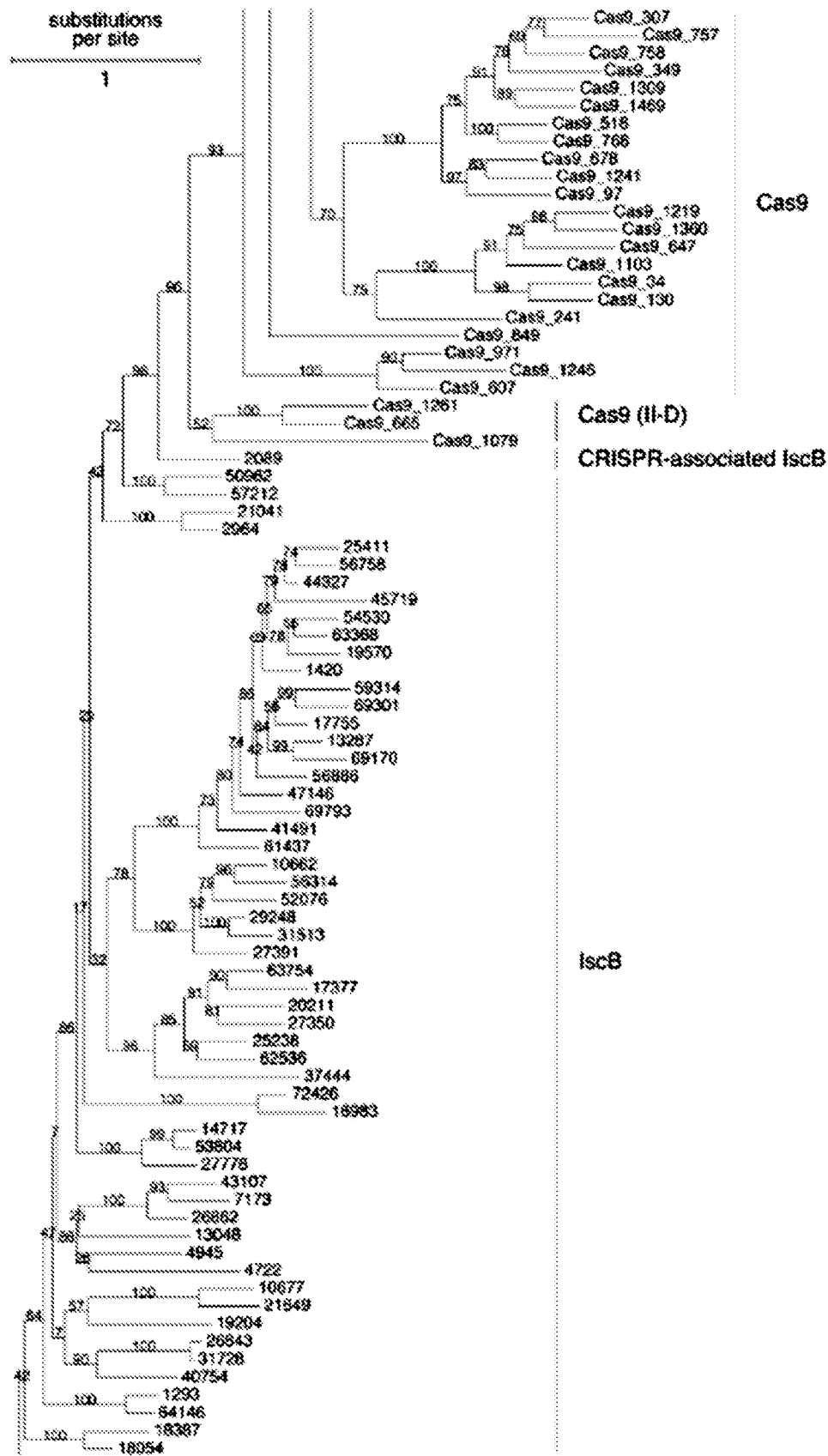


FIG. 35A

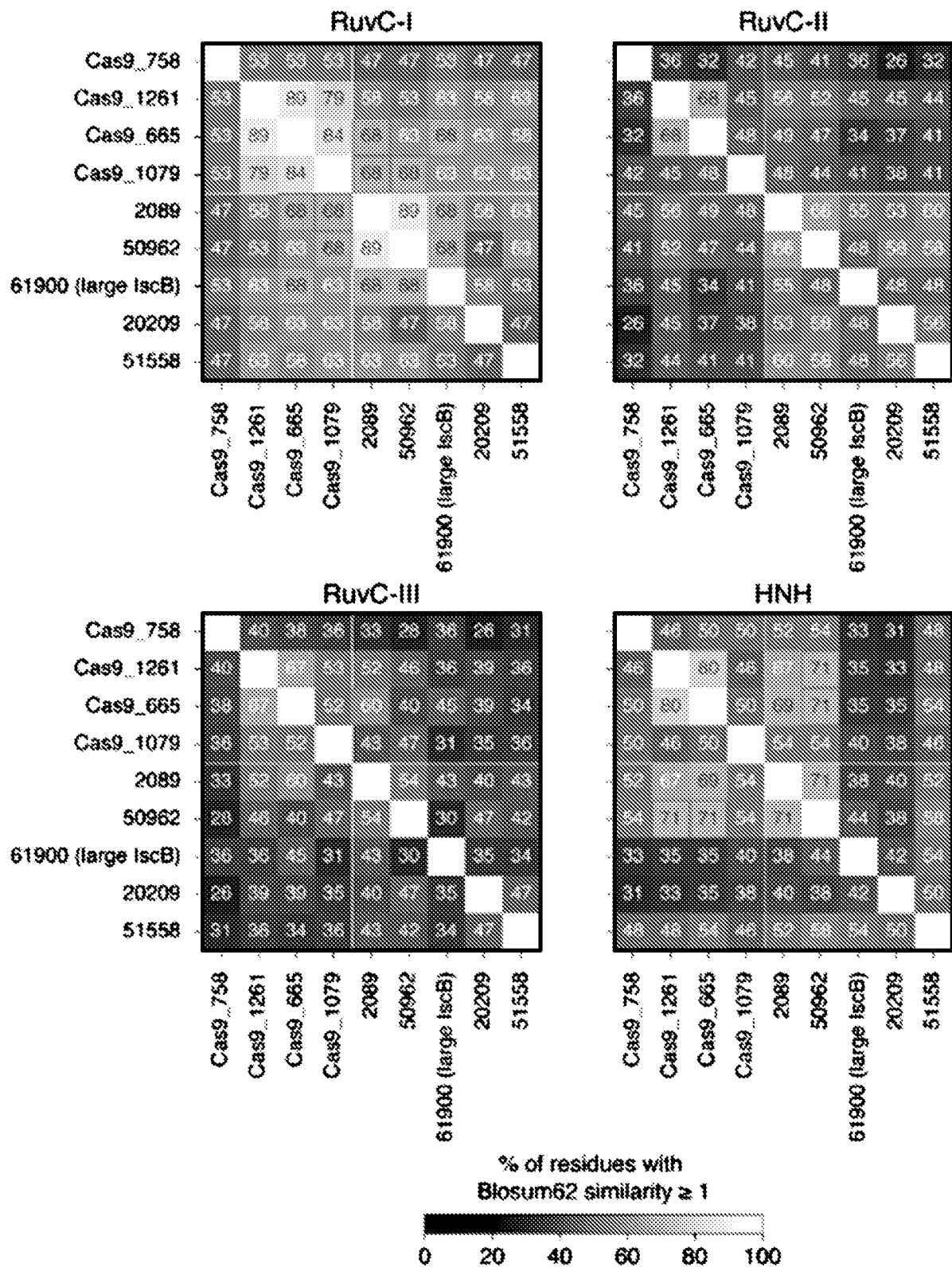


FIG. 35B

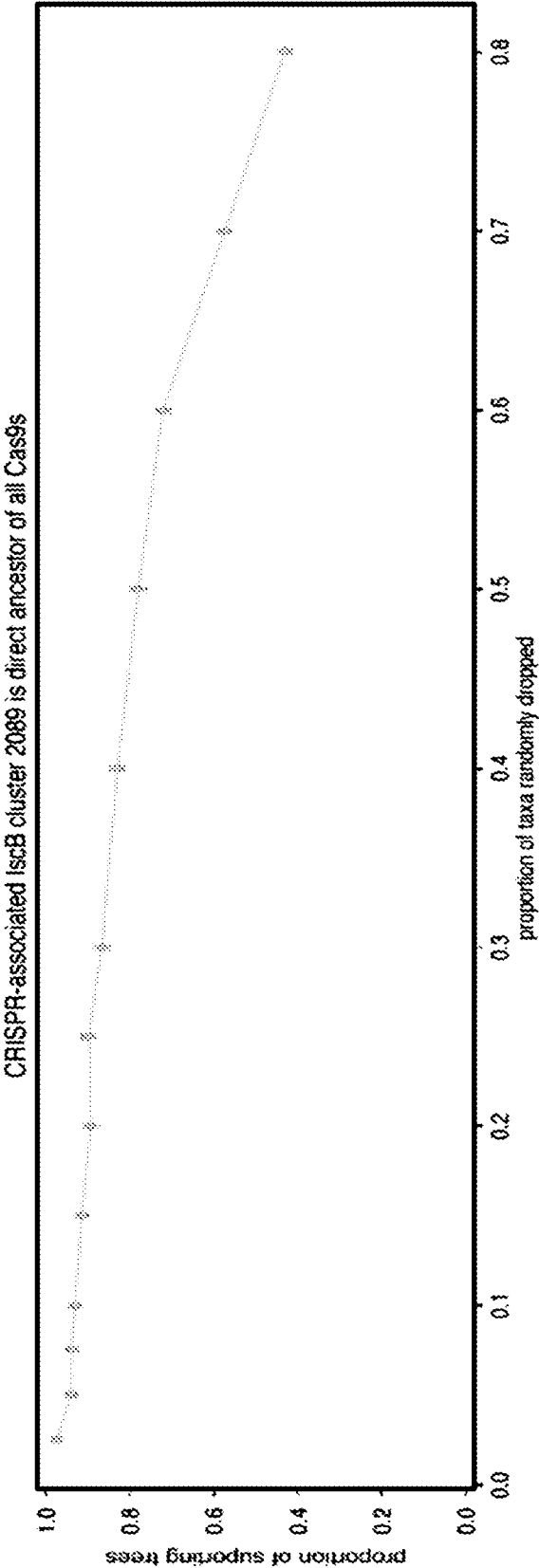


FIG. 35C

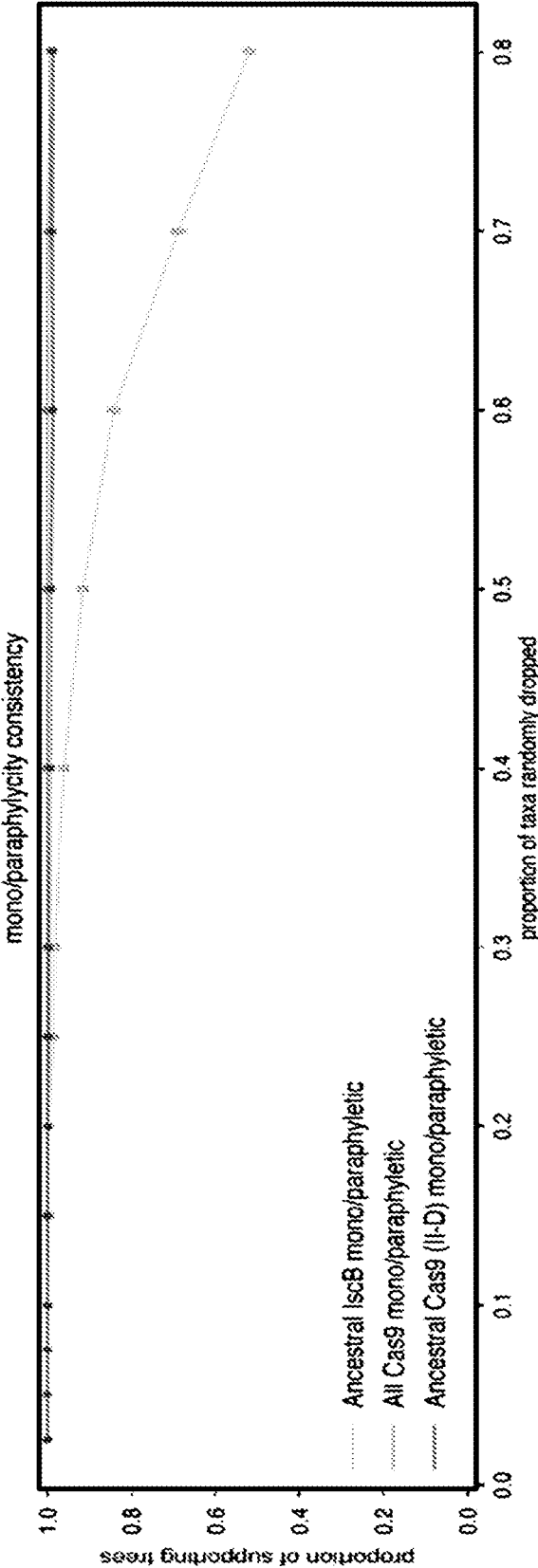


FIG. 35D

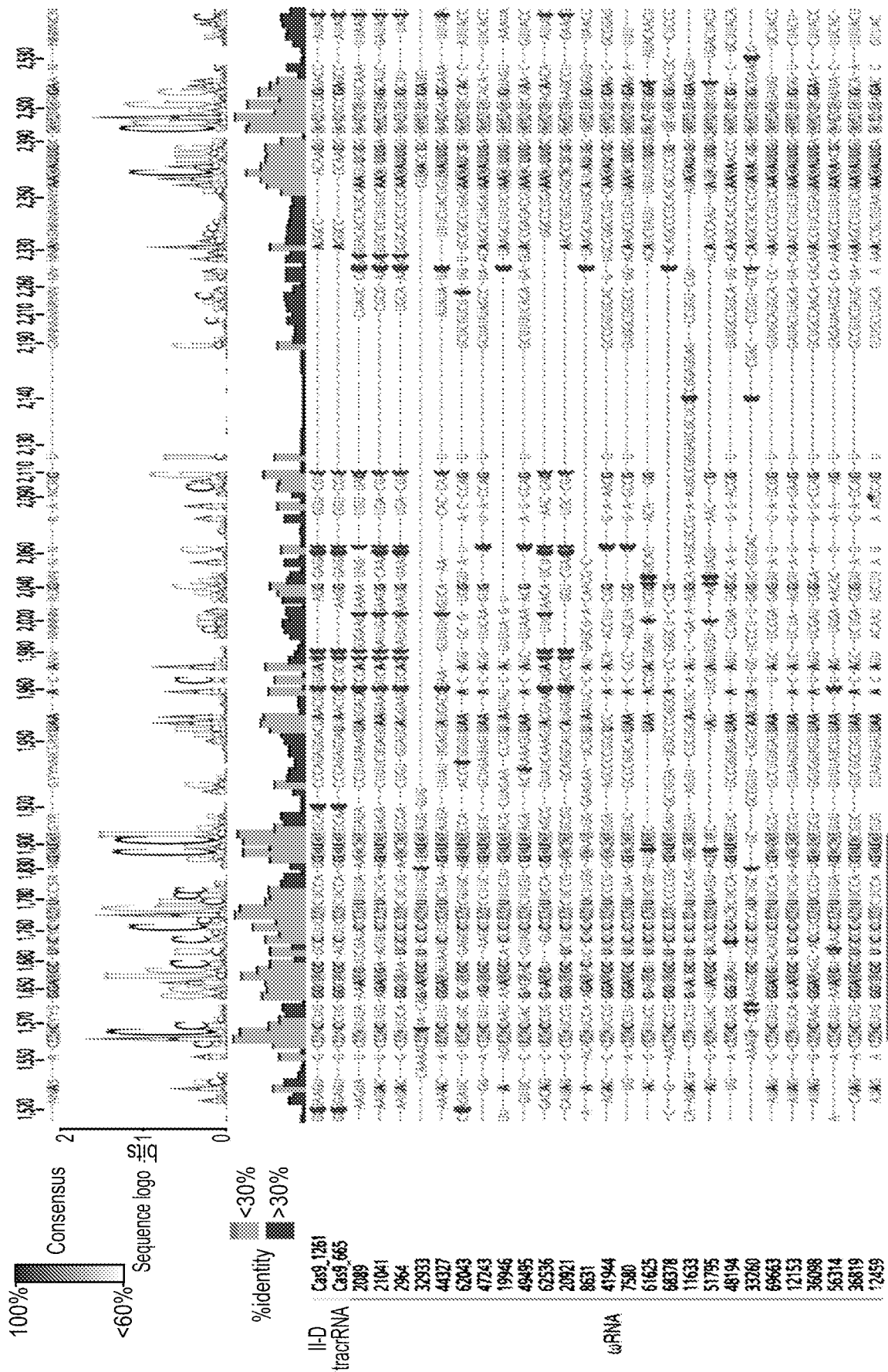


FIG. 36

nexus pseudoknot hairpin is conserved in both IscB/IscB and early Cas9 (Cas9_1261, Cas9_665)

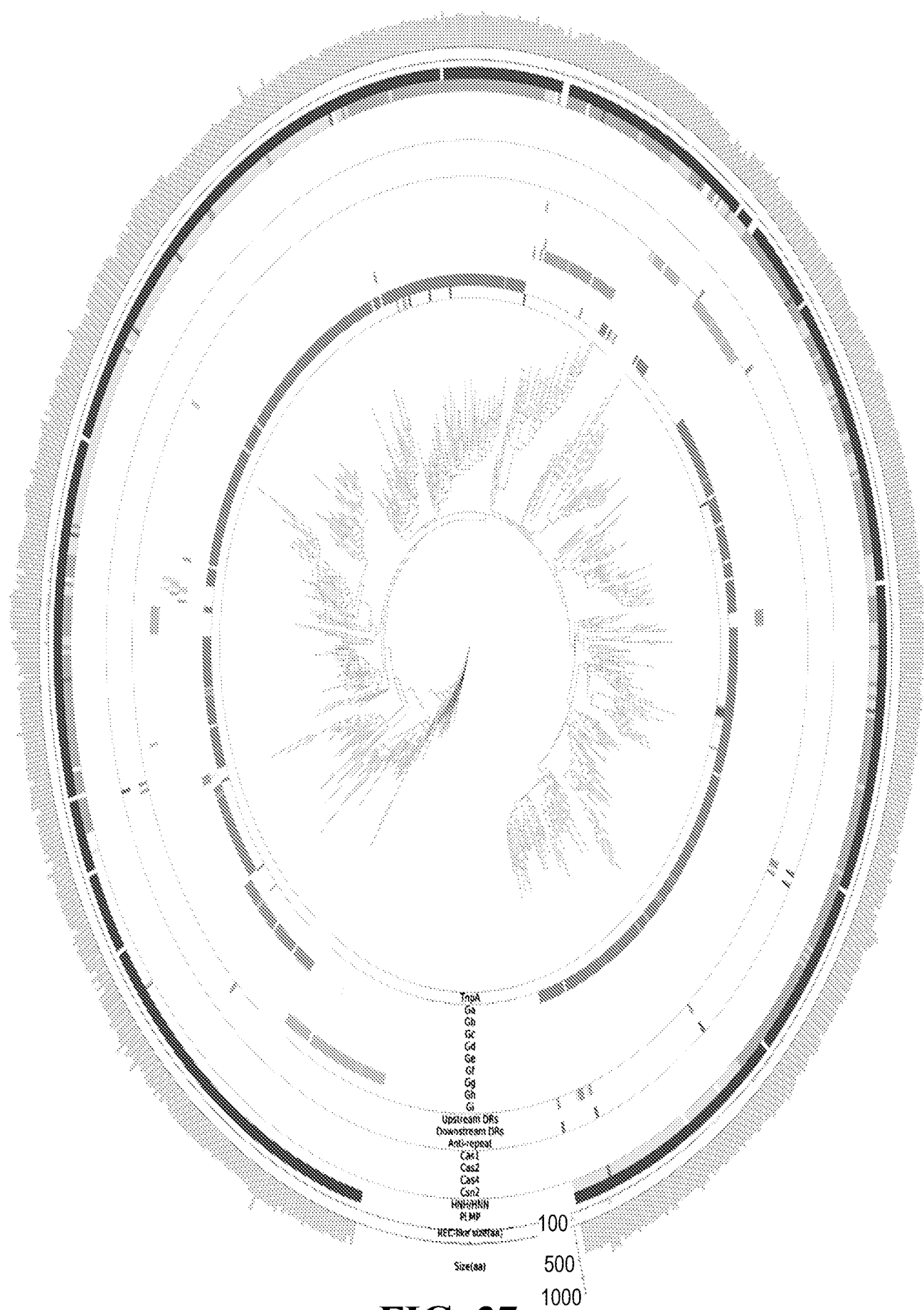


FIG. 37

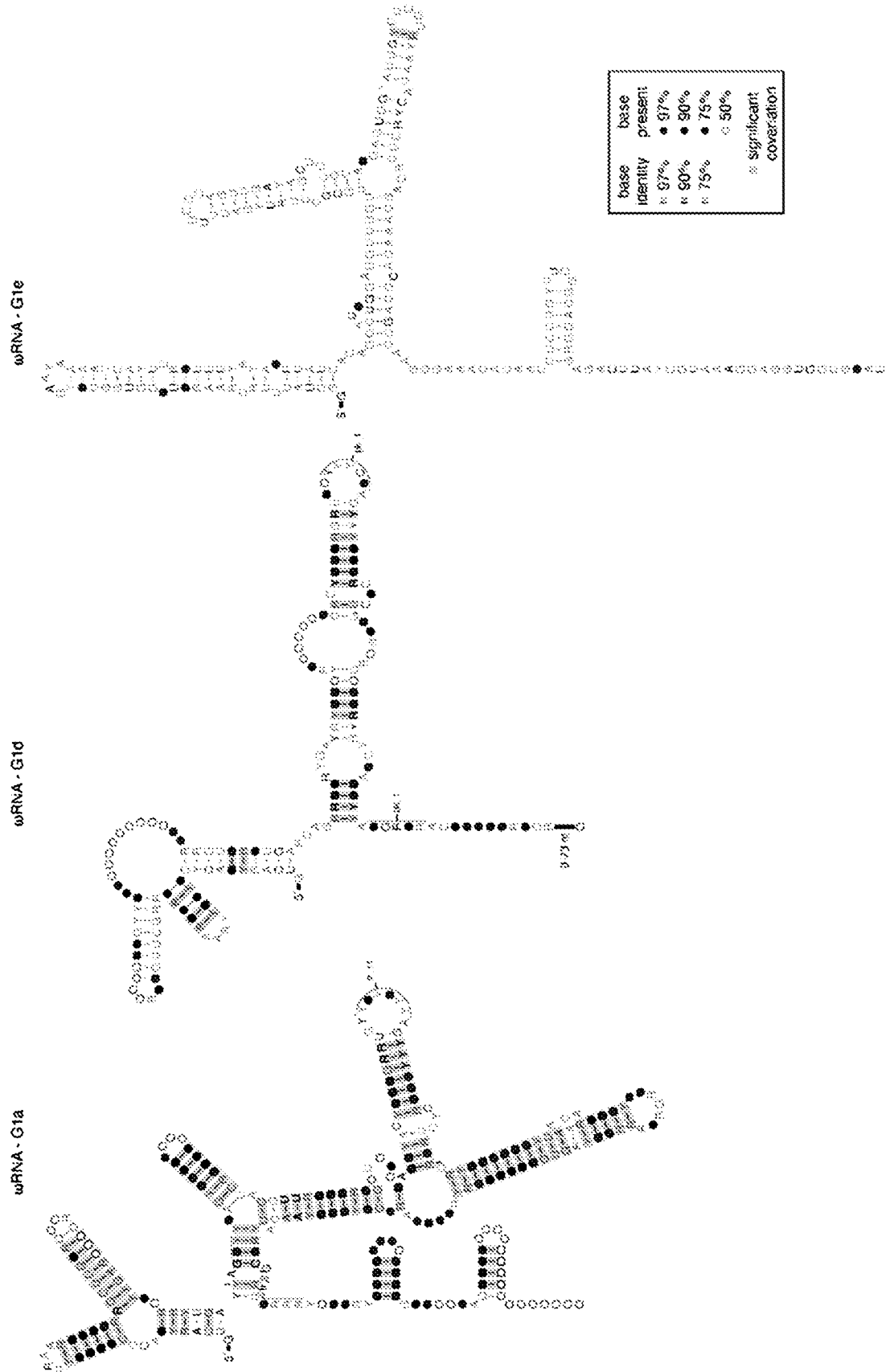


FIG. 38A

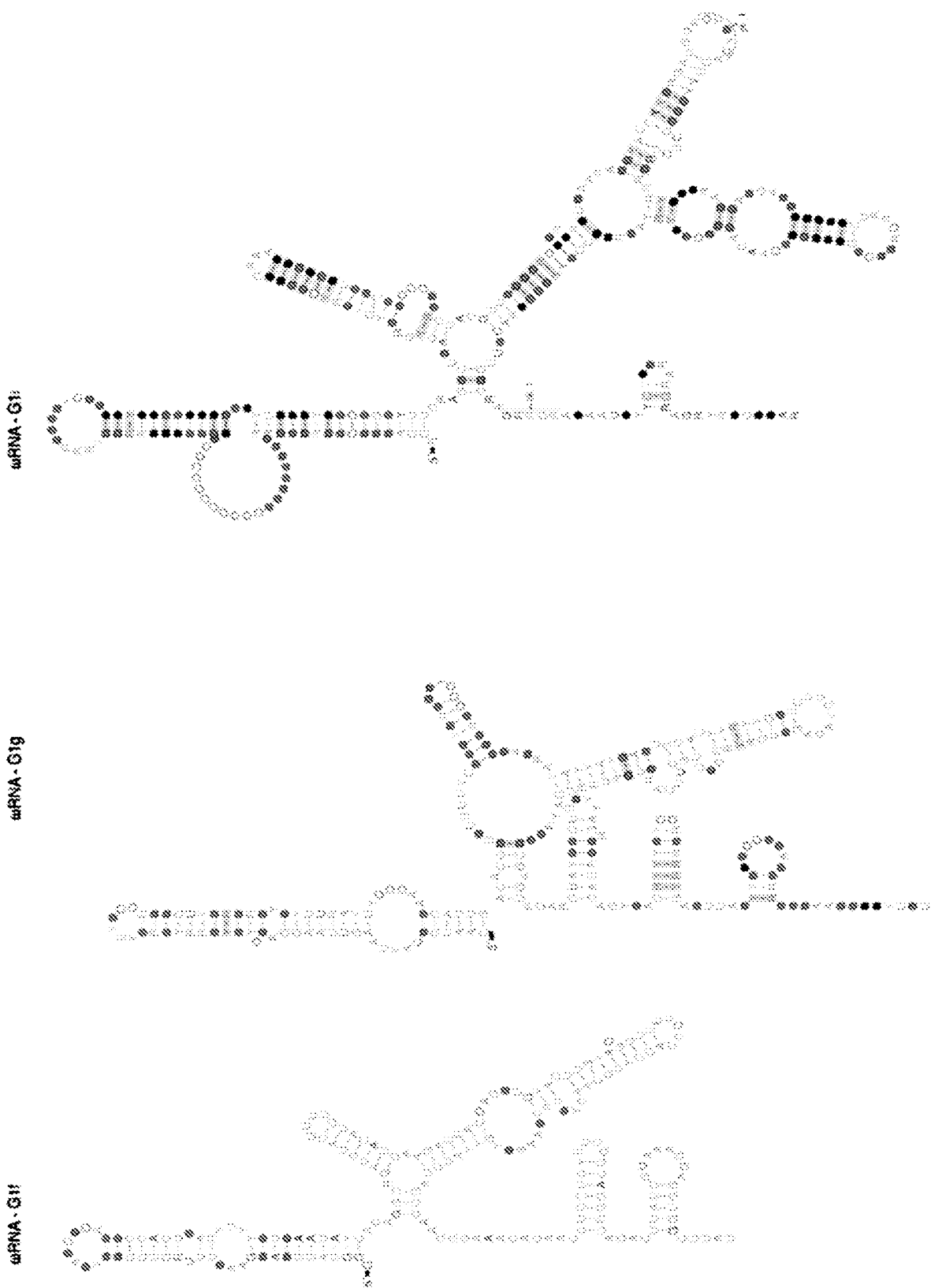
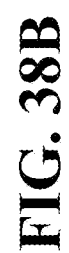


FIG. 38A (continued)



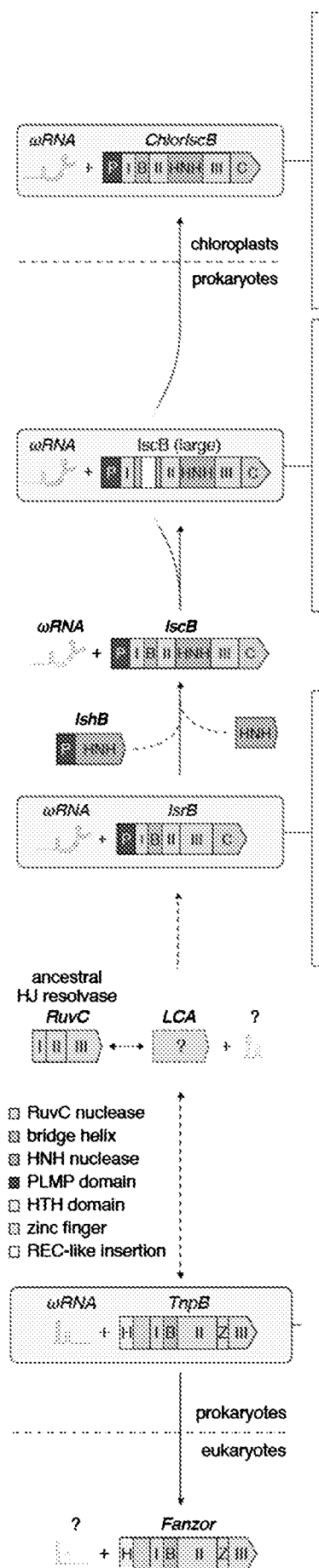


FIG. 39A

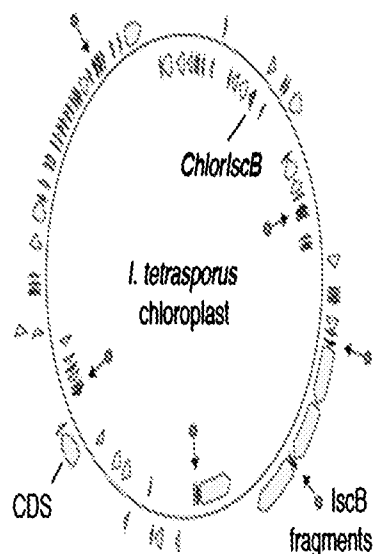


FIG. 39B

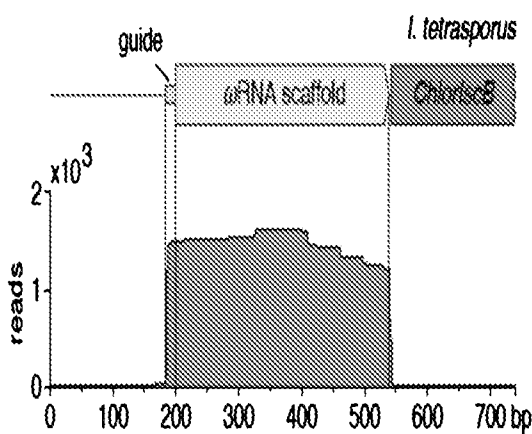


FIG. 39C

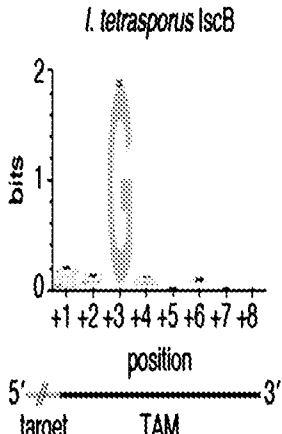


FIG. 39D

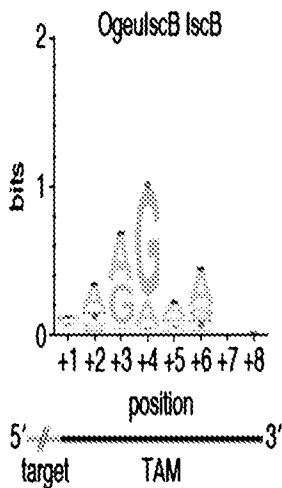


FIG. 39E

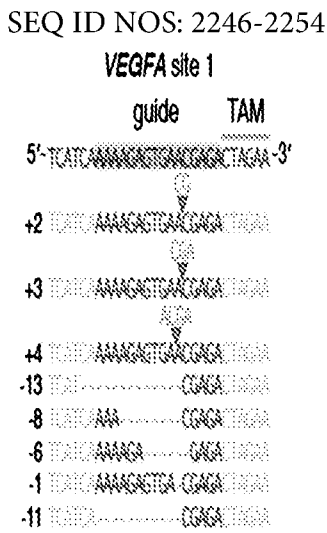


FIG. 39F

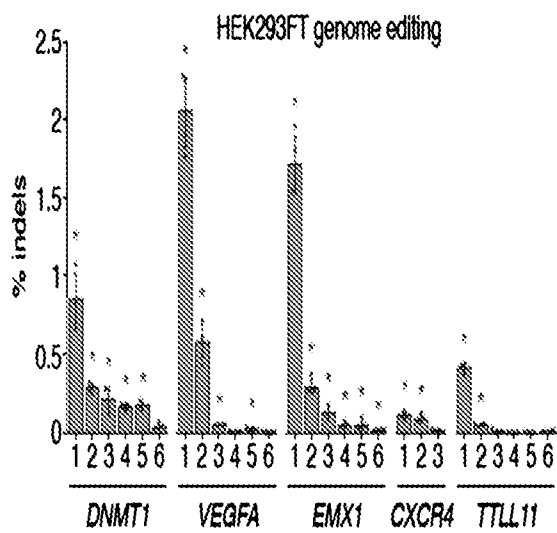


FIG. 39G

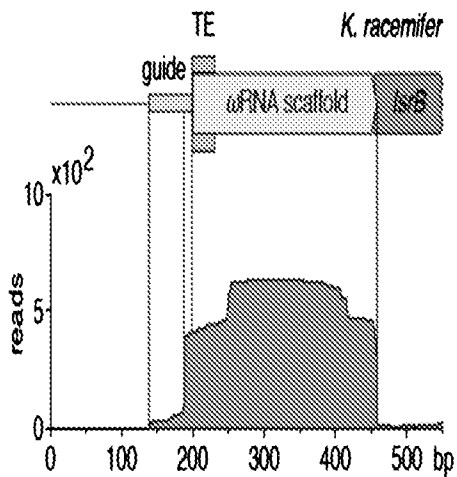


FIG. 39H

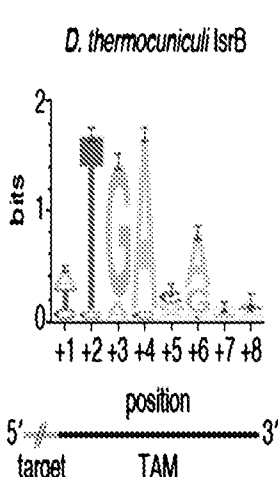


FIG. 39I

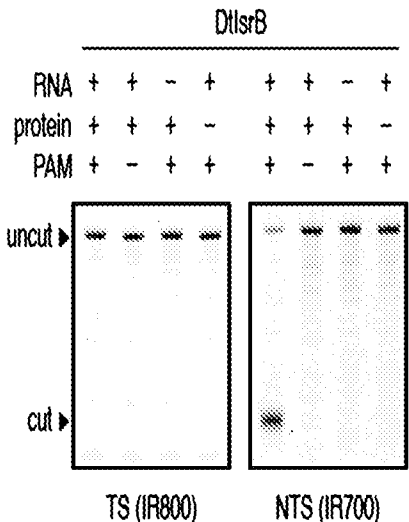


FIG. 39J

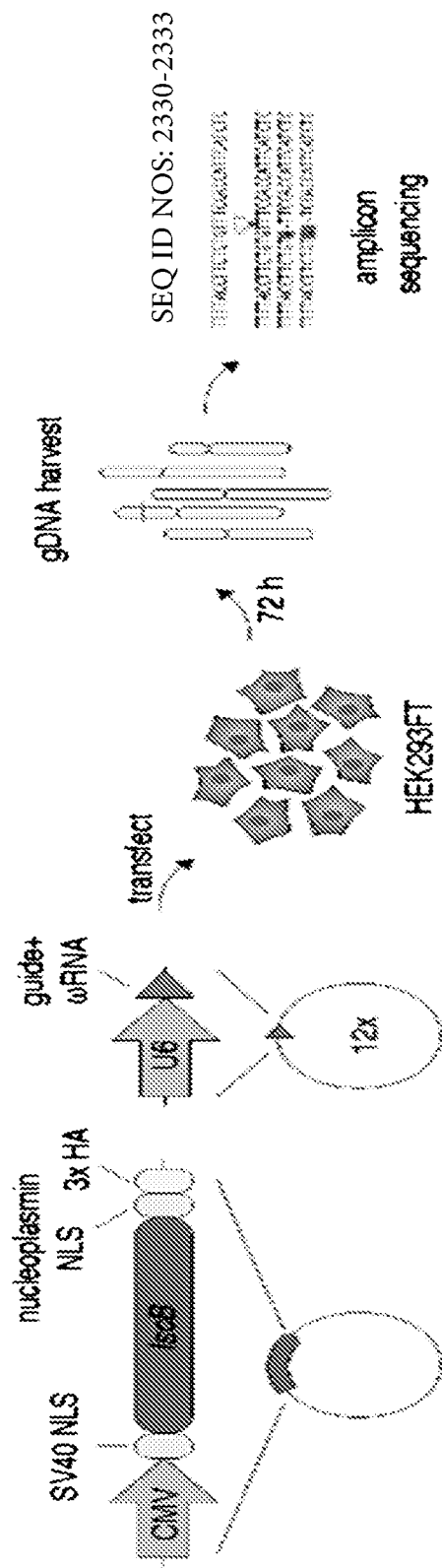


FIG. 40A

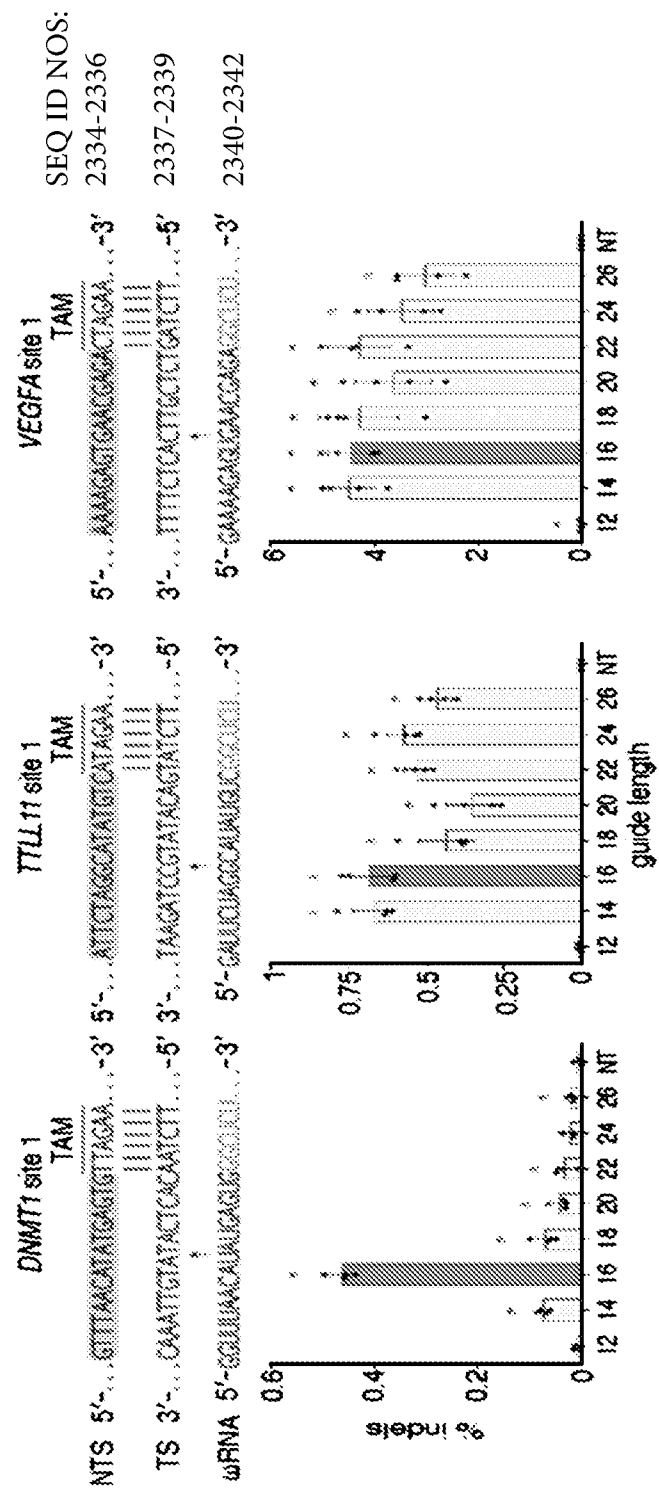


FIG. 40B

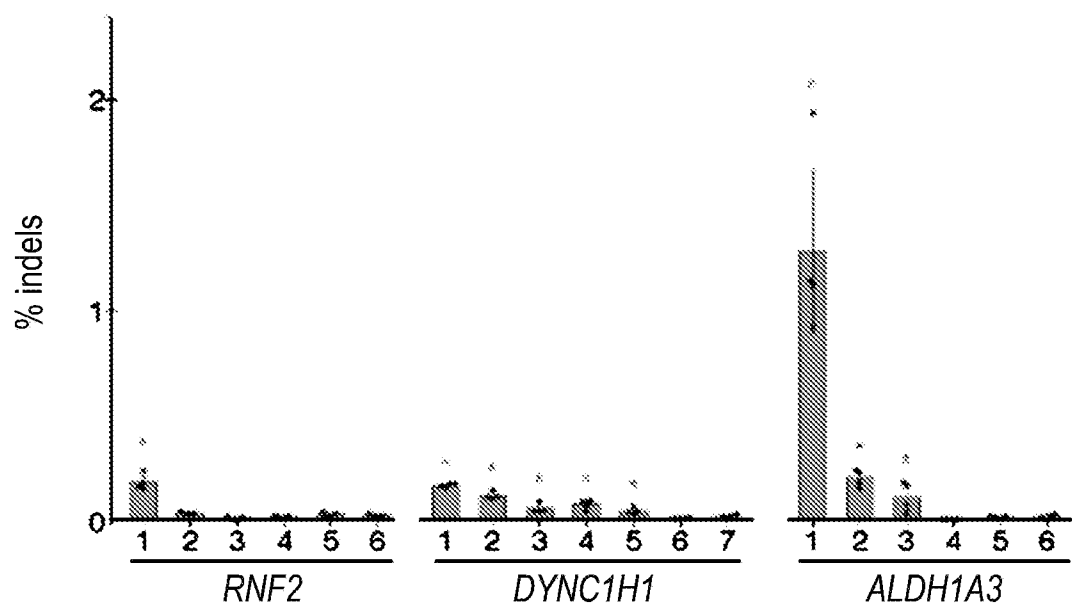


FIG. 40C

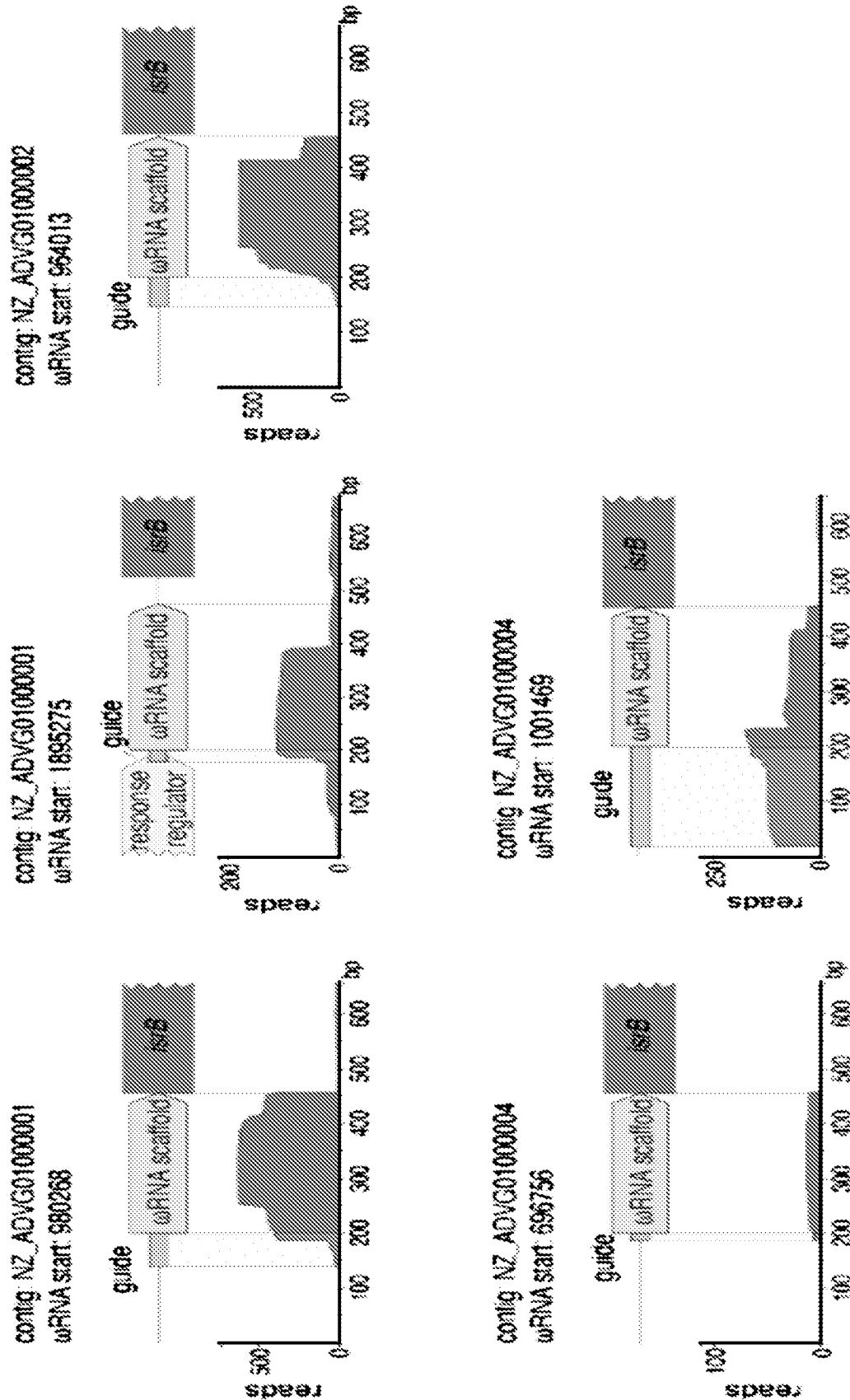


FIG. 41

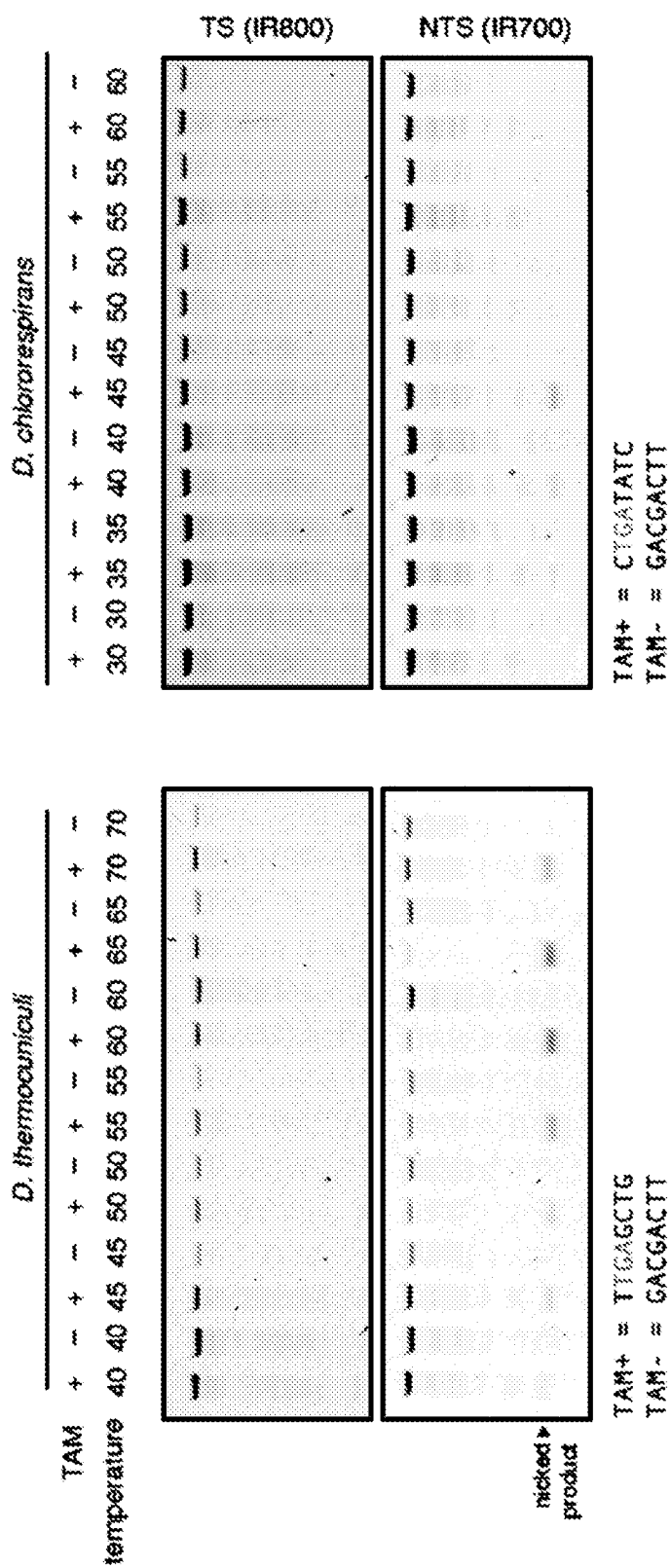


FIG. 42A

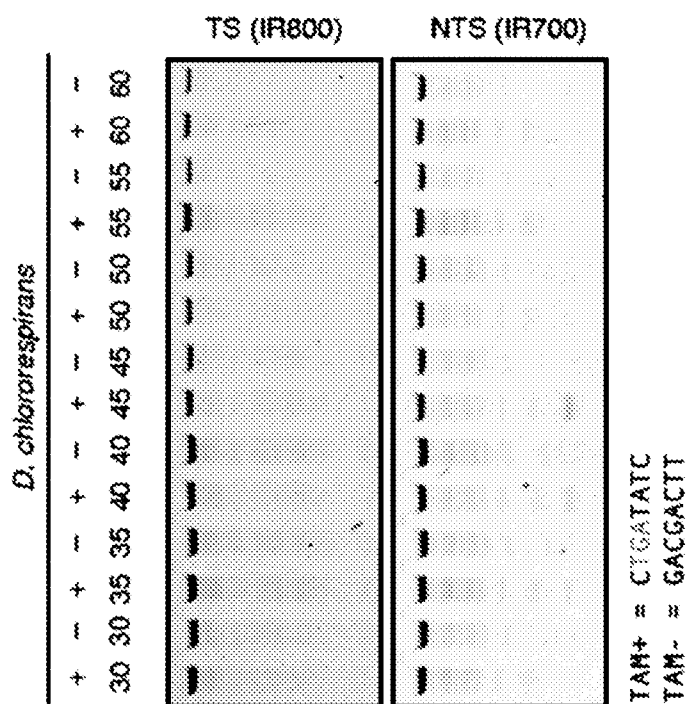


FIG. 42B

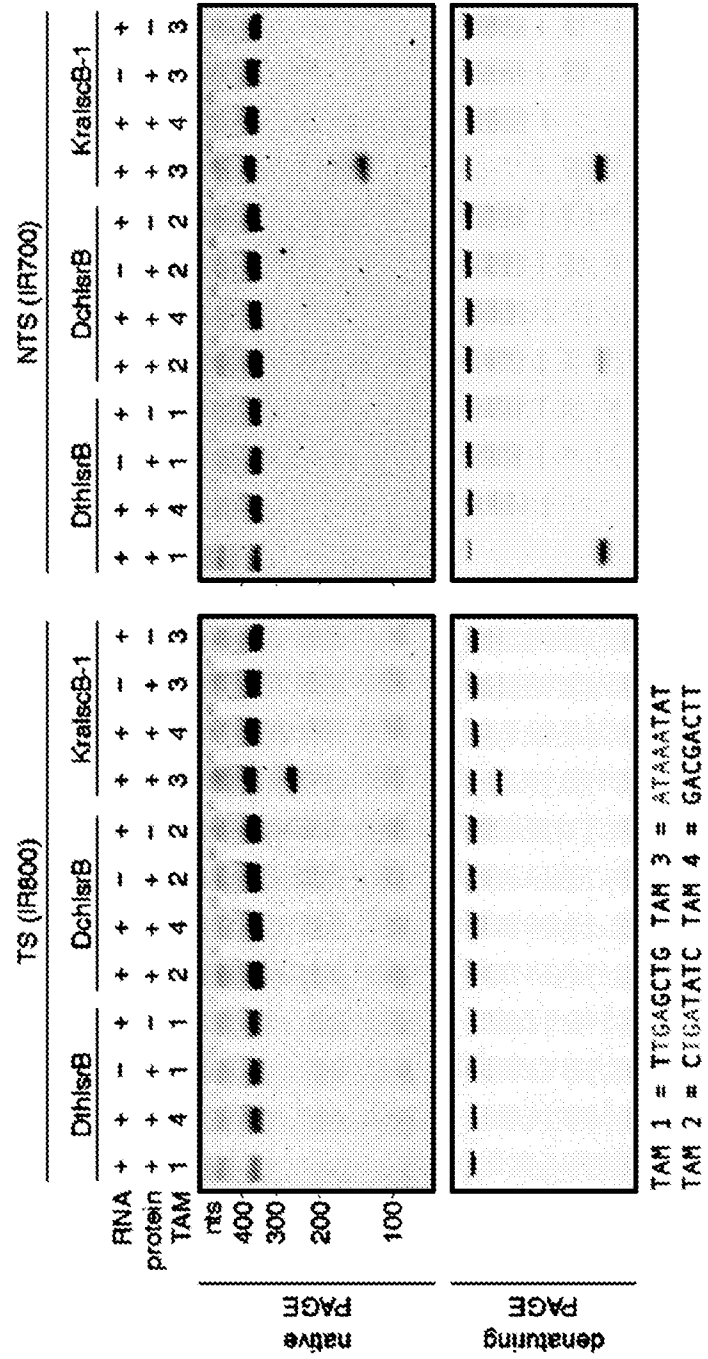


FIG. 42C

			<i>iscB</i>	<i>isrB</i>	<i>cas9</i>	# genomes total
Archaea	Asgard	Candidatus Heimdallarchaeota	0	0	0	15
		Candidatus Lokiarchaeota	0	0	0	44
		Candidatus Thorarchaeota	0	0	0	39
		Candidatus Poseidonia	0	0	0	454
	Candidatus Thermoplasmatota	Thermoplasmatia	2.5	0	1.7	355
		Candidatus Aenigmarchaeota	0	0	0	82
		Candidatus Altiaarchaeota	0	0	0	35
		Candidatus Diapherotrites	6.9	0	0	29
	DPANN group	Candidatus Micrarchaeota	0	0	1.5	136
		Candidatus Pacearchaeota	0	0	0	78
		Candidatus Parvarchaeota	0	0	0	18
		Candidatus Woesearchaeota	0	0	0.65	153
	Nanoarchaeota	Nanoarchaeota	0	0	0	50
		Archaeoglobi	0	0	0	75
		Euryarchaeota incertae sedis	7.5	0	0	40
		Hadesarchaea	0	0	0	29
	Euryarchaeota	Methanomada group	2.9	0	0	420
		Stenosarchaea group	2.9	0.69	0.077	1303
		Thermococci	0	0	0	104
		Unclassified Euryarchaeota	1.1	0	0.21	473
	TACK group	Other	0	0	0	14
		Candidatus Bathyarchaeota	0	0.58	0	347
		Candidatus Korarchaeota	0	0	0	30
		Candidatus Marsarchaeota	0	0	0	15
	Candidatus Verstraetearchaeota	Candidatus Verstraetearchaeota	0	0	0	34
		Crenarchaeota	0	0	0.21	474
		Thaumarchaeota	0	0	0	496
	Acidobacteria	Acidobacteria	0	0	4.2	529
		Blastocatellia	6.7	0	2.7	75
		Holophagae	0	0	26	197
		Environmental samples	0	0	3	66
	Aquificae	Unclassified Acidobacteria	0.34	0	1.5	582
		Aquificae	2.6	0	0	76
		Unclassified Aquificae	0	0	0	19
		Bacteria incertae sedis	0.24	0.6	1.2	5529
	Caldiserica/Cryosepicota group	Caldiserica	0	0	0	47
		Calditrichae	0	0	0	23
		Unclassified Calditrichae	3.4	0	3.4	29
		Coprothermobacterota	0	0	0	10
	Deferribacteres	Deferribacteres	3.7	0	11	27
		Dictyoglomi	0	0	0	12
		Elusimicrobia	0	0	5.6	72
		Endomicrobia	0	0	64	11
	Elusimicrobia	Unclassified Elusimicrobia	0	0	5.6	182
		Bacteroidetes/Chlorobi group	0.3	0.013	20	15748
		Candidatus Cloacimonetes	0	0	6.5	139
		Candidatus Fermentibacteria	0	0	6.7	15
	Candidatus Hydrogenedentes	Candidatus Hydrogenedentes	0	0	18	73
		Candidatus Kryptonia	0	0	0	34

FIG. 43

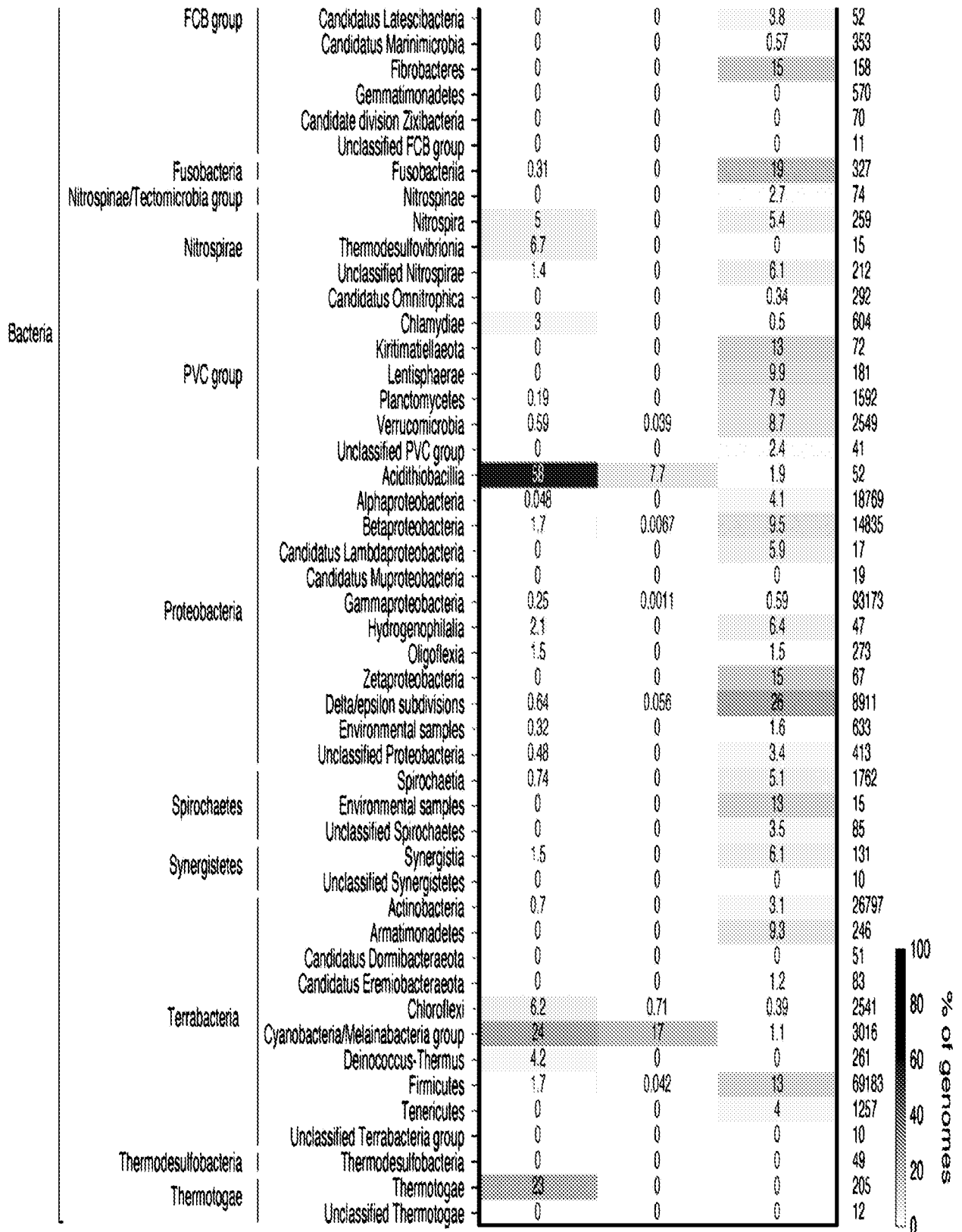


FIG. 43 (continued)

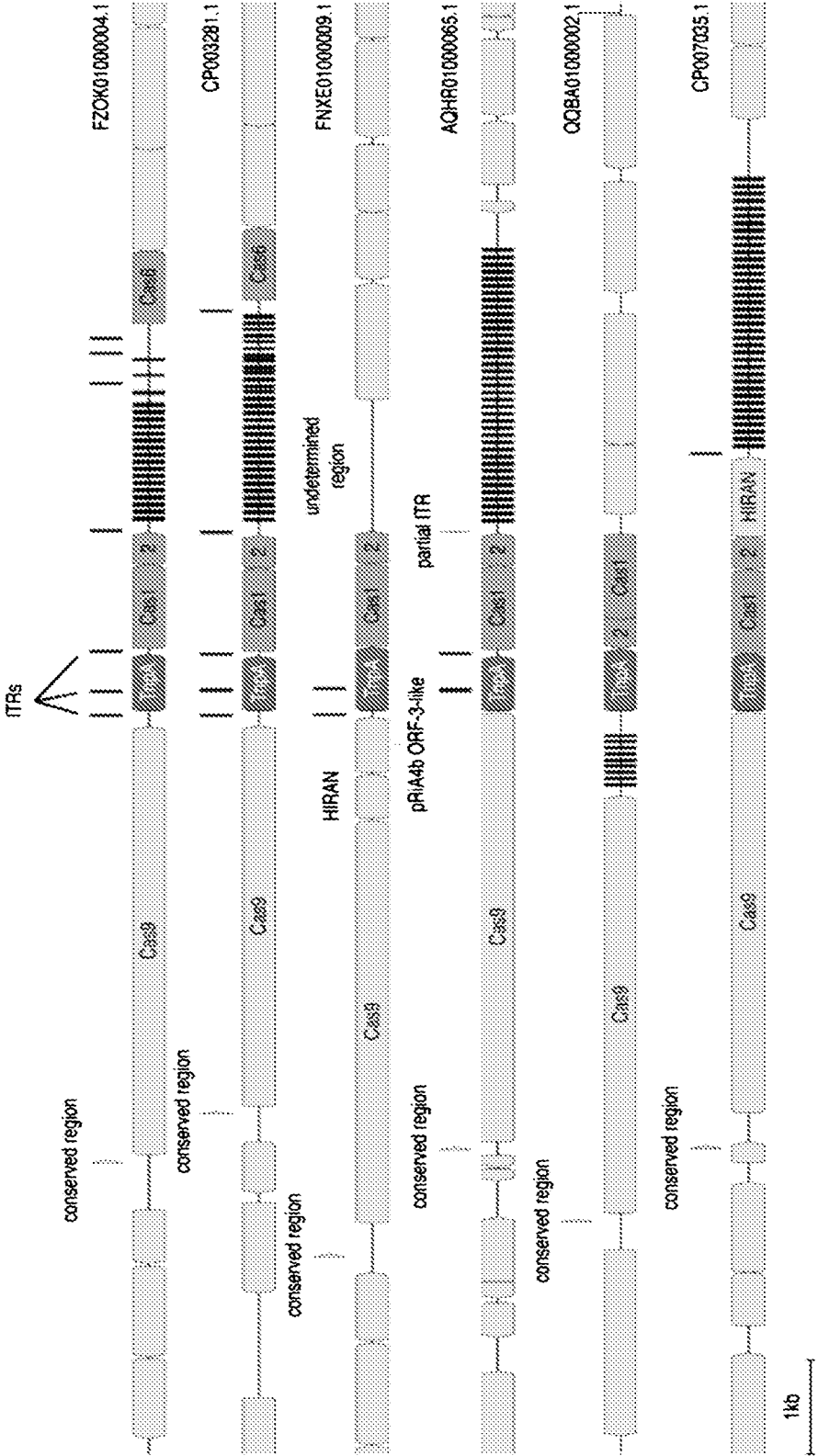


FIG. 44

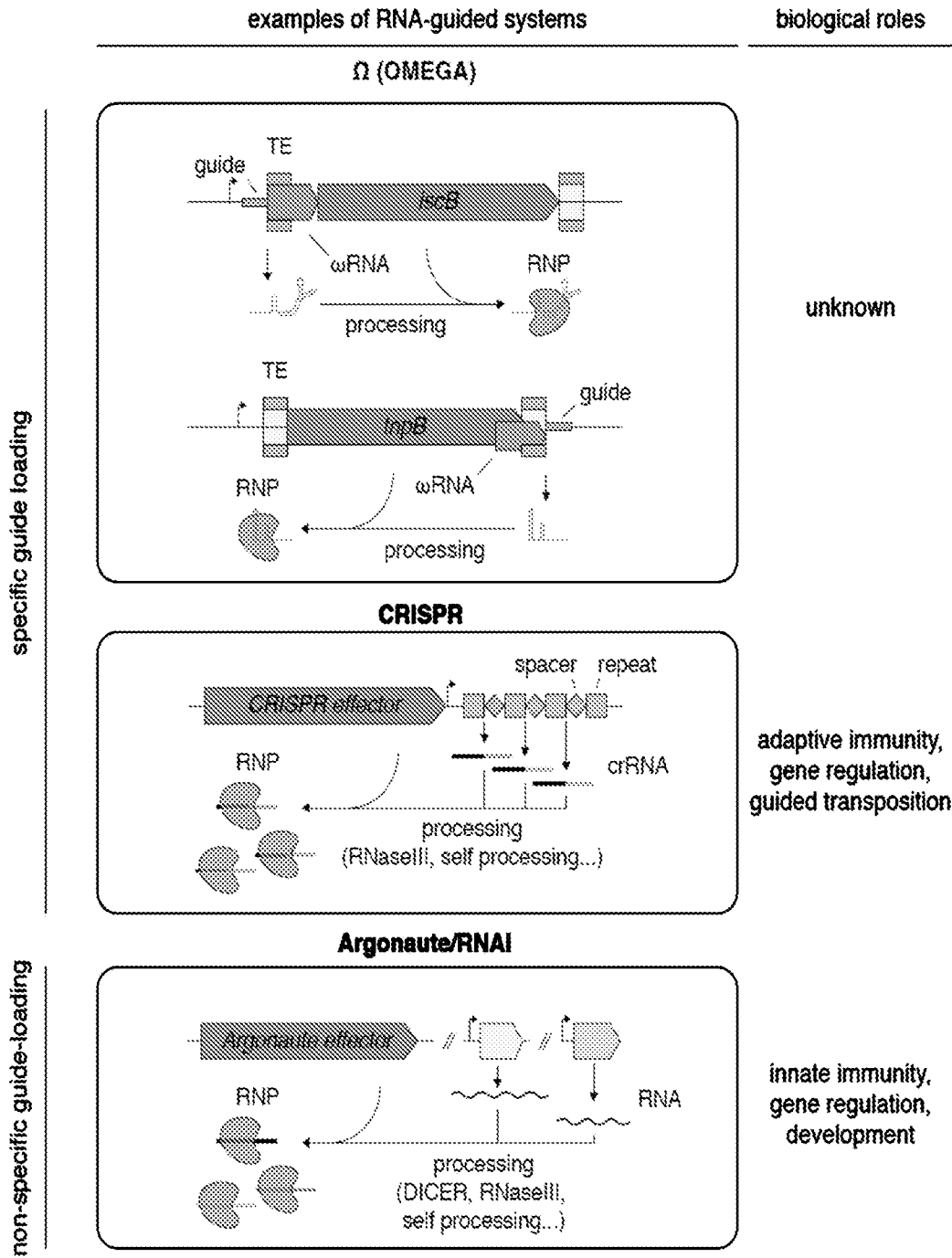


FIG. 45



FIG. 46A

SEQ ID NOS: 2343-2345

PSP1 target

8N TAM
library

NTS 5'-...TAGGCGGCTGATCCGTTTAGCTCAGCGAAAGTCGATNNNNNNNNATCGTCC...-3'

TS 3'-...ATCCCGCGACTACGCAATCGAGTCCGCTTTCAGCTANNNNNNNTAGCAGG...-5'

ncRNA 5'-CUGAUGCGUUAAGCUCAGCGAAAGUCGATGUUGUAG...-3'

SEQ ID NOS: 2346-2348

PSP2 target

8N TAM
library

NTS 5'-...TAGGCGGCTGGGTCGATTTCATCTCGAGTCTGGGAATC'NNNNNNNNATCGTCC...-3'

TS 3'-...ATCCCGCGCAECCACCTAAGTAGACCTCAGACCCTTAGNNNNNNNTAGCAGG...-5'

ncRNA 5'-GUGGGUGGAUUCAUUCUGGAGUCUGGGAUUCGUUGUAG...-3'

SEQ ID NOS: 2349-2351

PSP3 (Fn) target

8N TAM
library

NTS 5'-...CAAGCTTTTAAACAGTGGCCTTATTAAATGACTTCTC'NNNNNNNNATCGTCC...-3'

TS 3'-...GTTCCGAAAATTGTCACCGGAATAATTTACTGAAGAGNNNNNNNTAGCAGG...-5'

ncRNA 5'-UUUACAGUGGCCUUUAUUAAUACAUUCUGUUGUAG...-3'

FIG. 46B

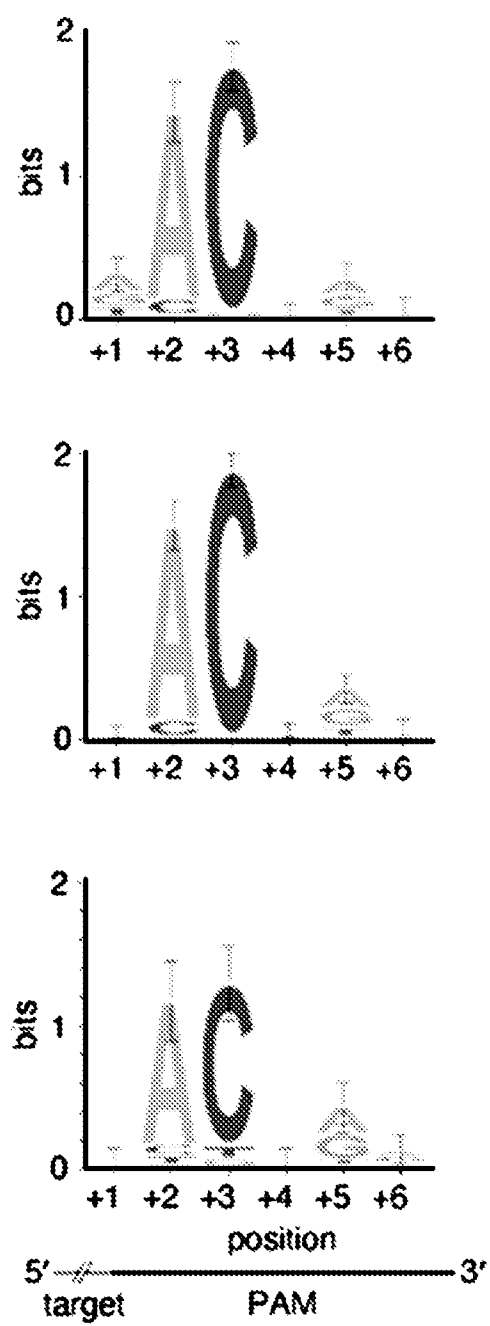


FIG. 46C

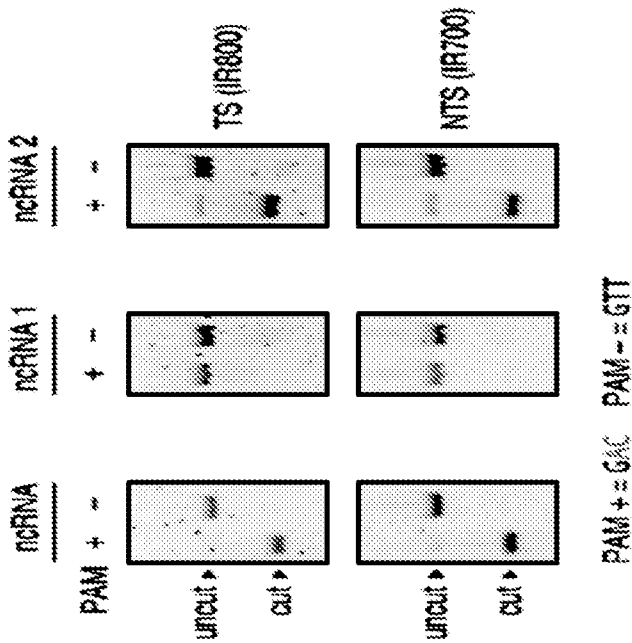
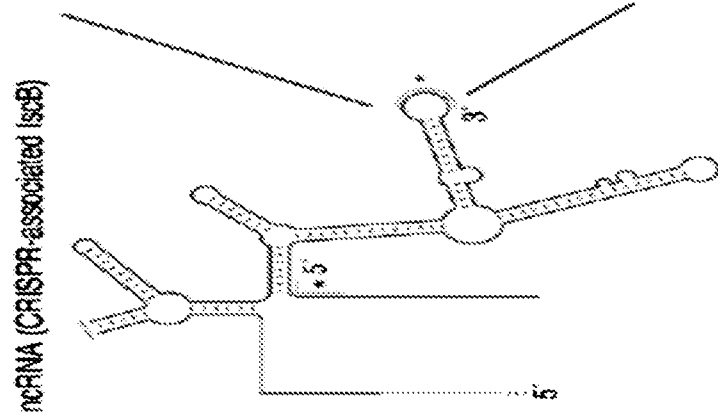


FIG. 47B

* nexus pseudoknot mutants			
wild-type ncRNA	ncRNA mutant 1	ncRNA mutant 2	
5' G C C G U G C C A G	5' G C C G U G C C A G	5' G C C G U G C C A G	
3' C A U A C C G A G	3' C A U A C C G A G	3' C A U A C C G A G	
mutated base relative to wild-type			

FIG. 47A



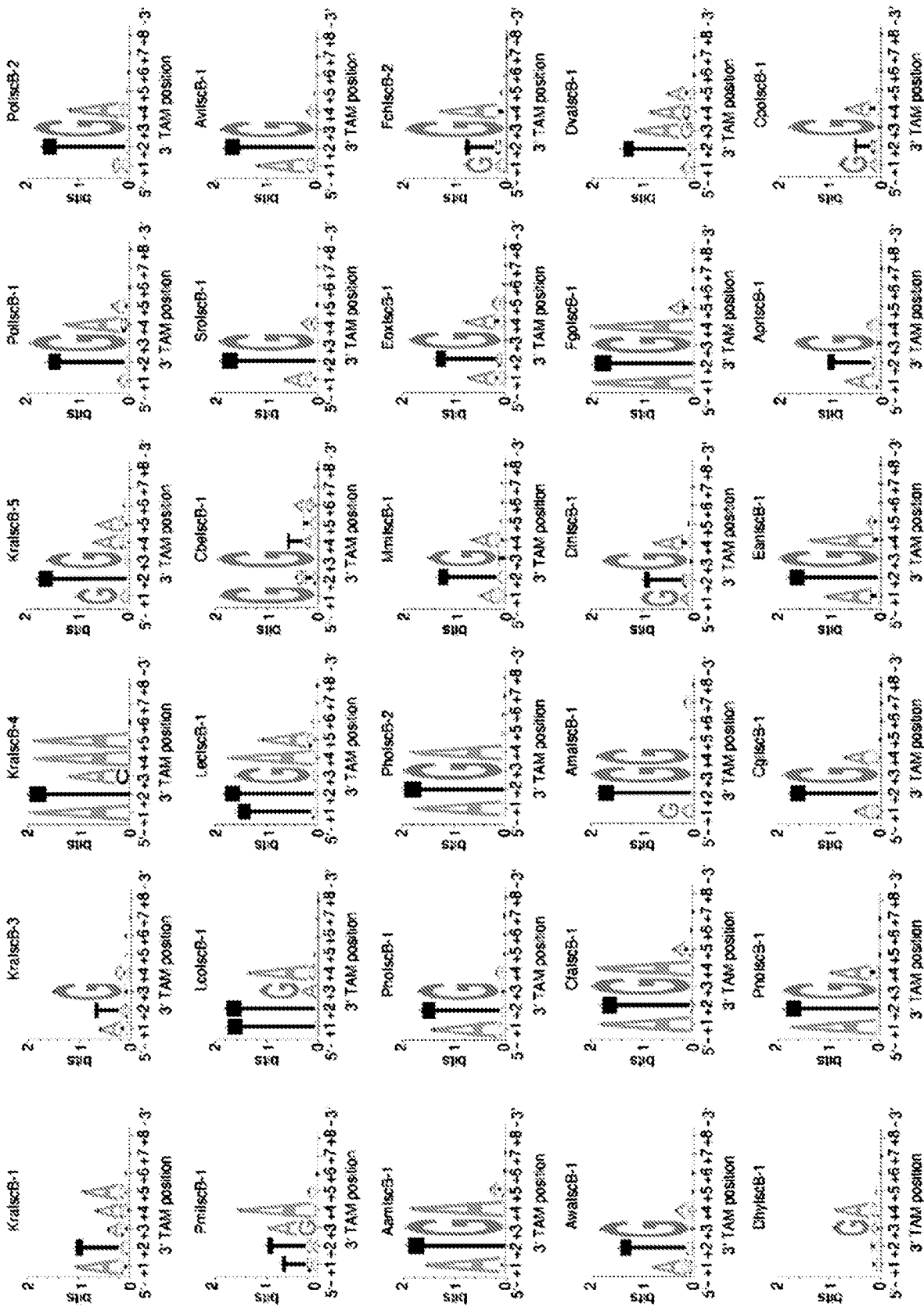


FIG. 48

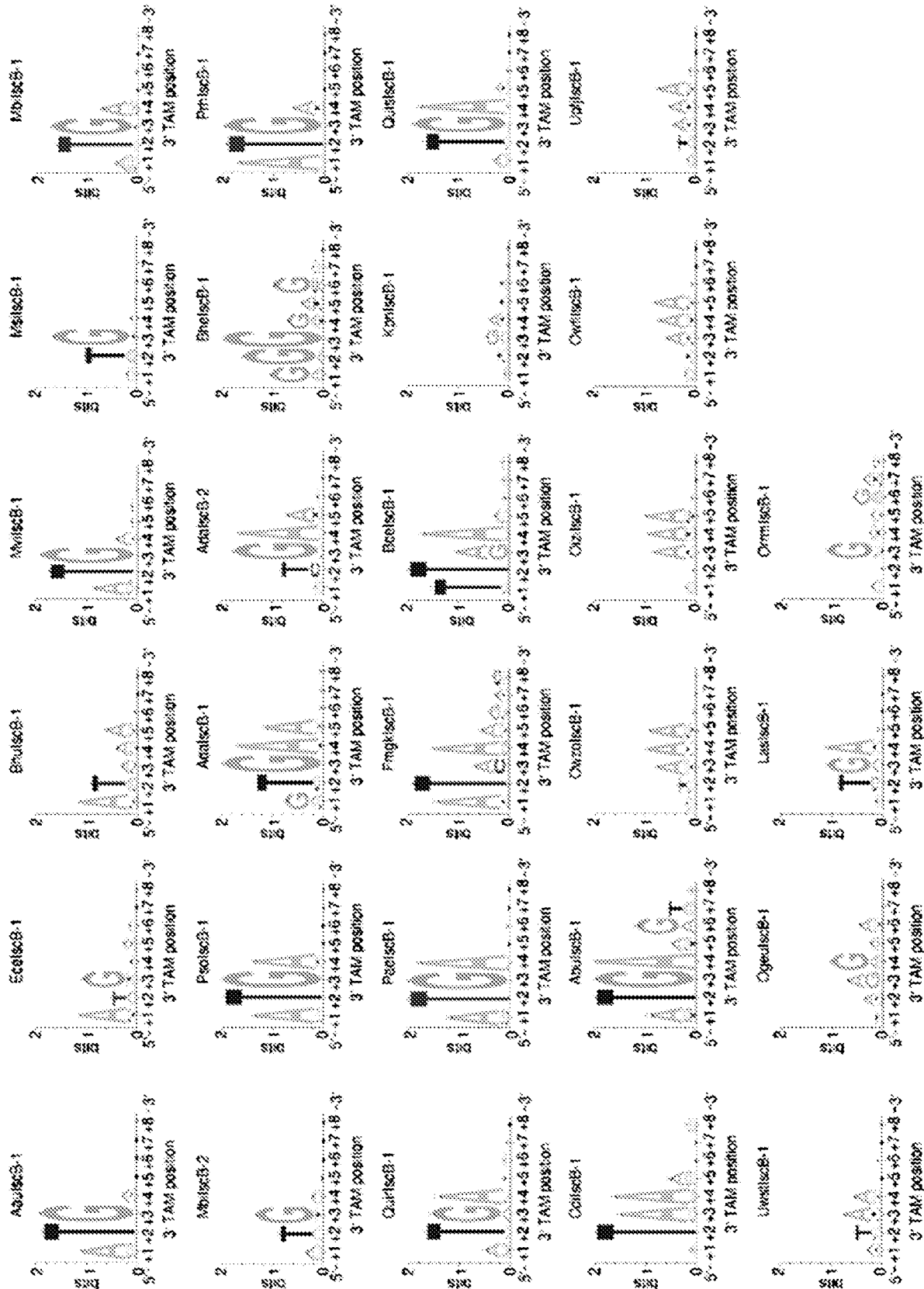


FIG. 48 (continued)

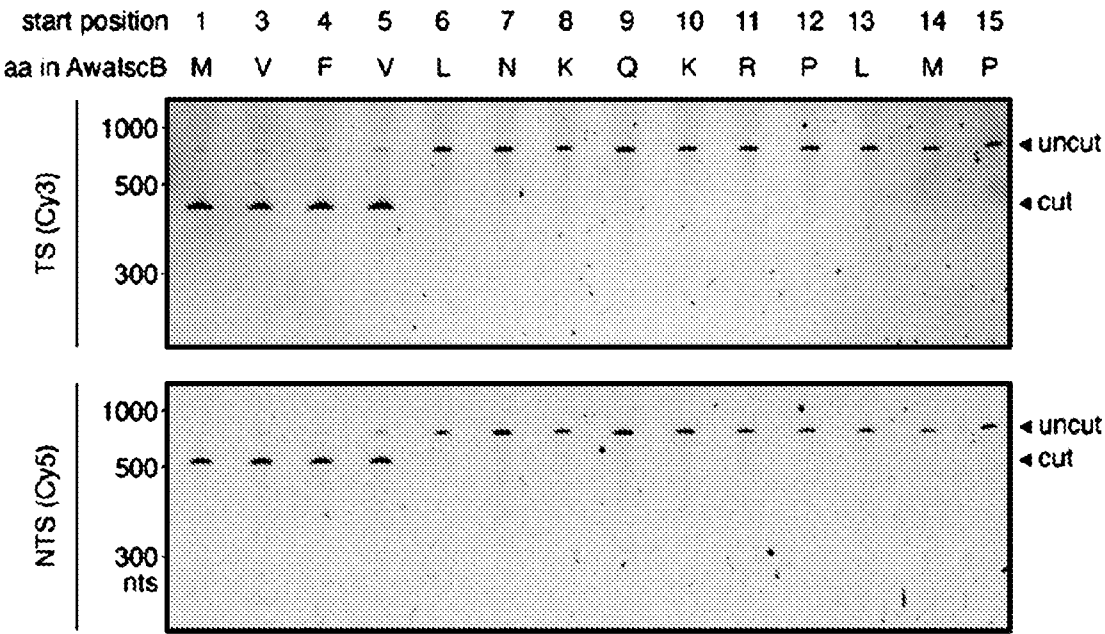


FIG. 49

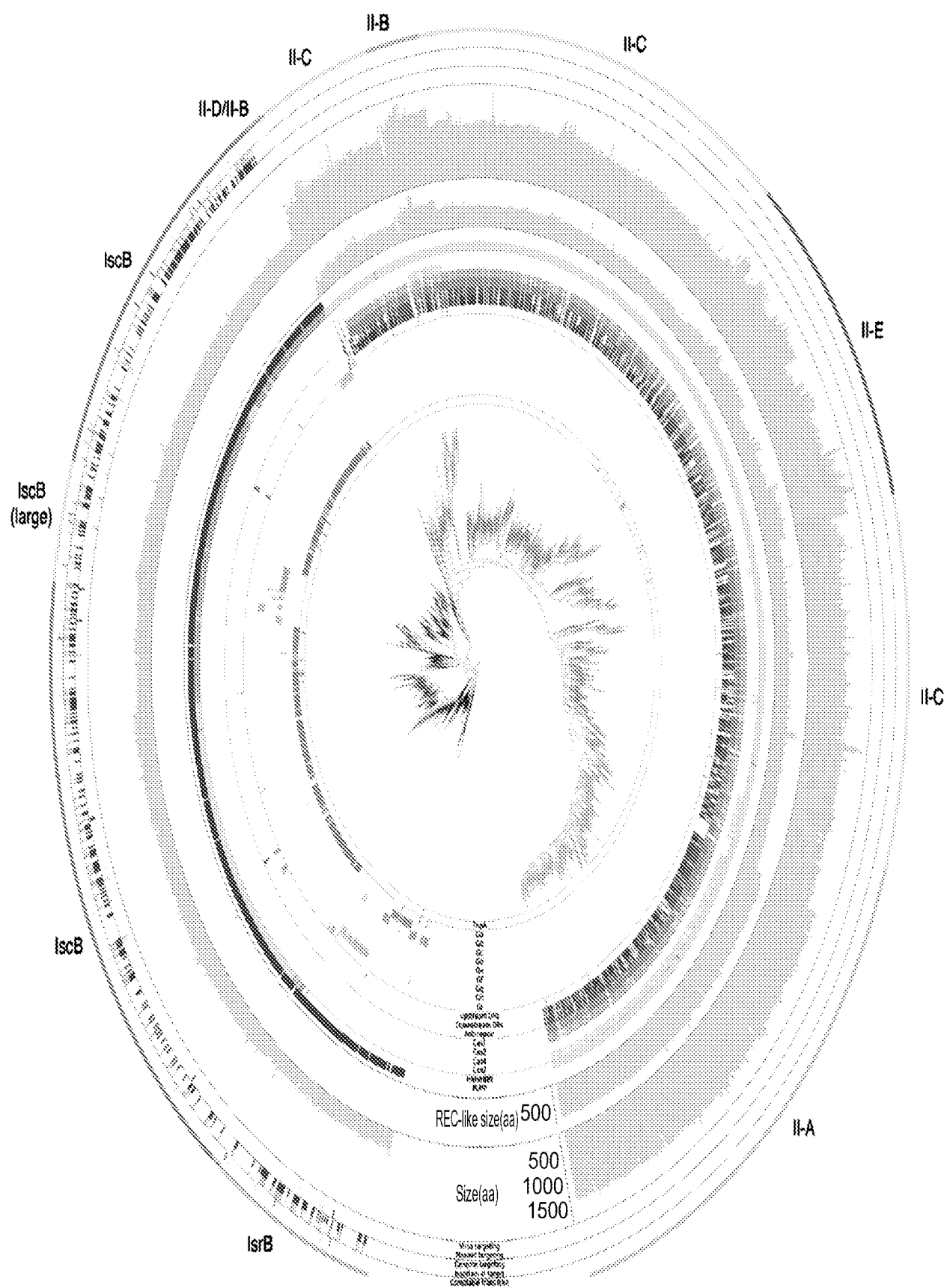


FIG. 50

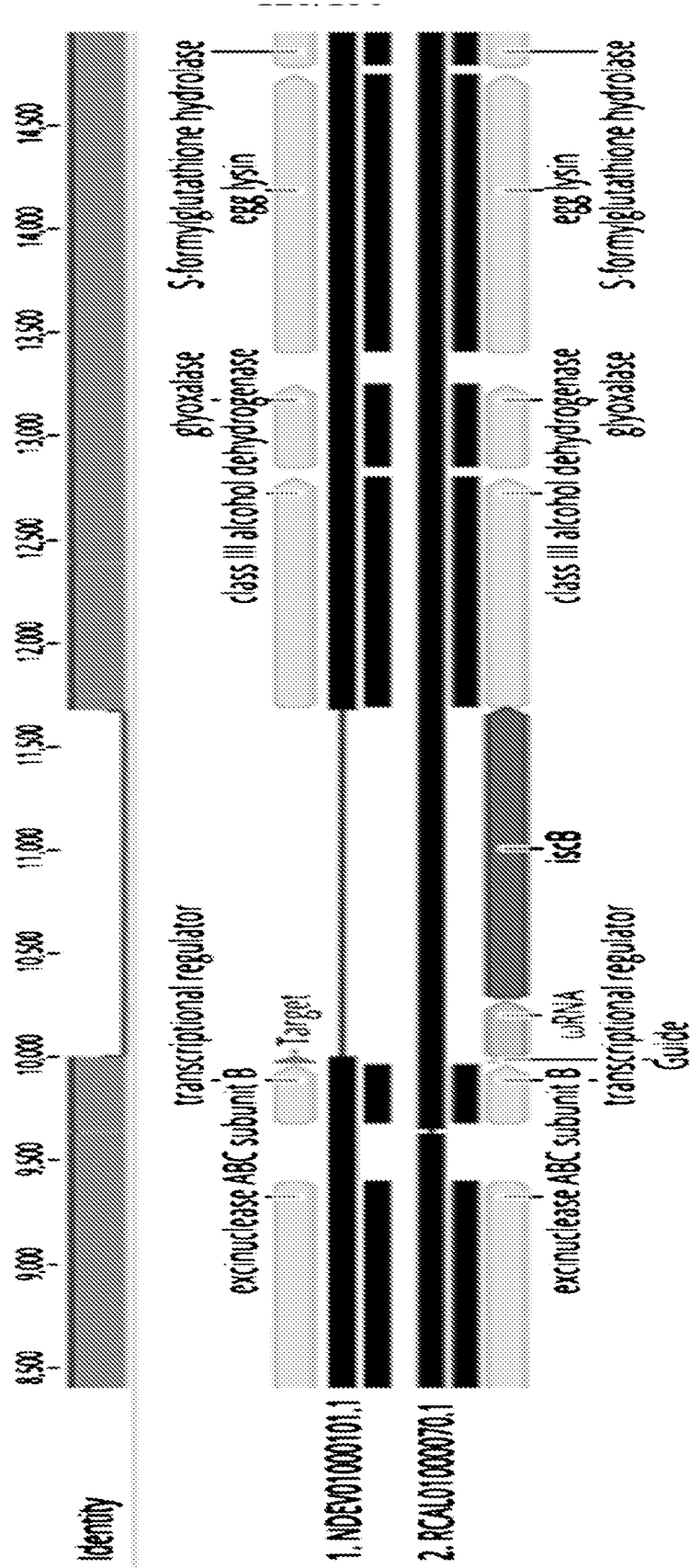


FIG. 51A

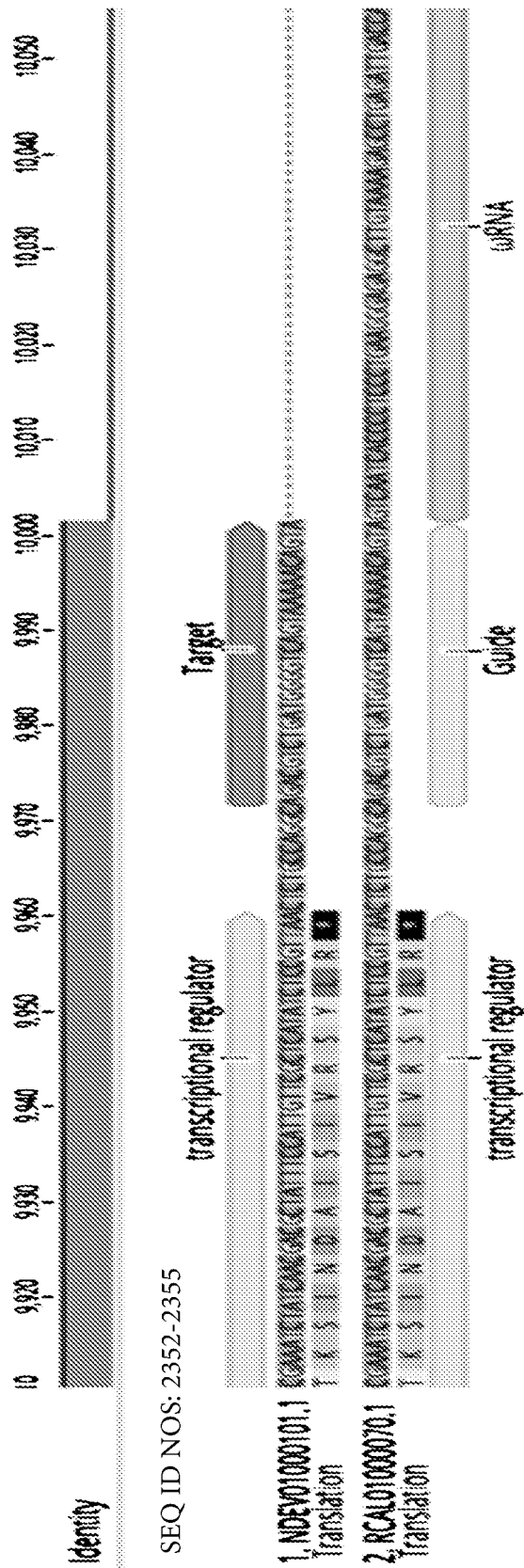


FIG. 51B

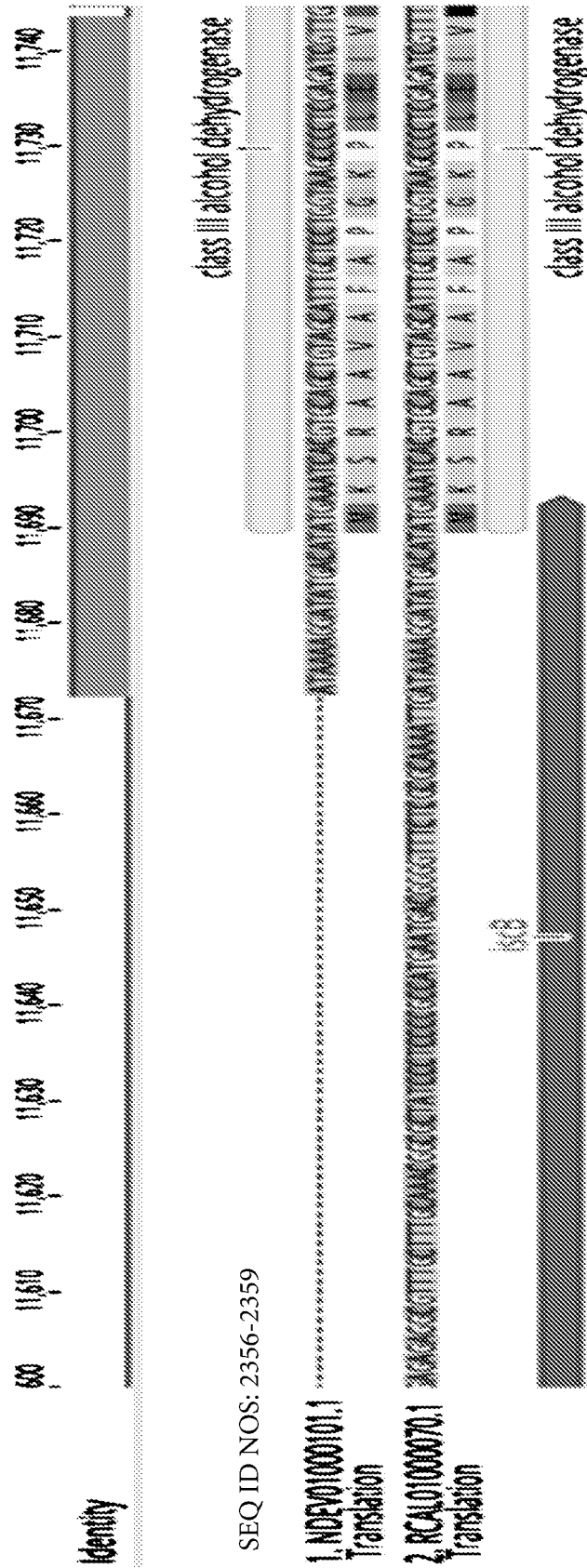


FIG. 51C

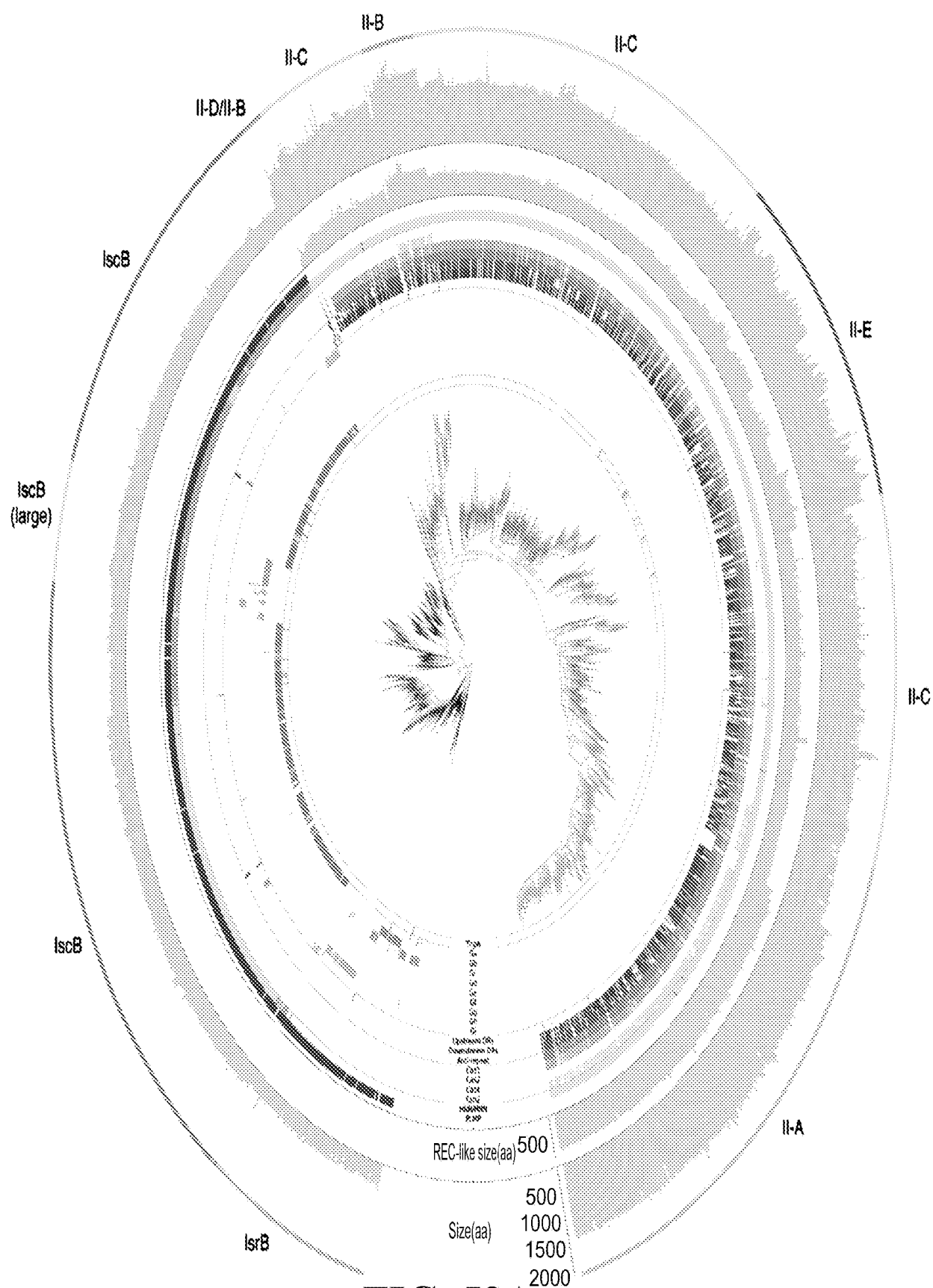


FIG. 52A

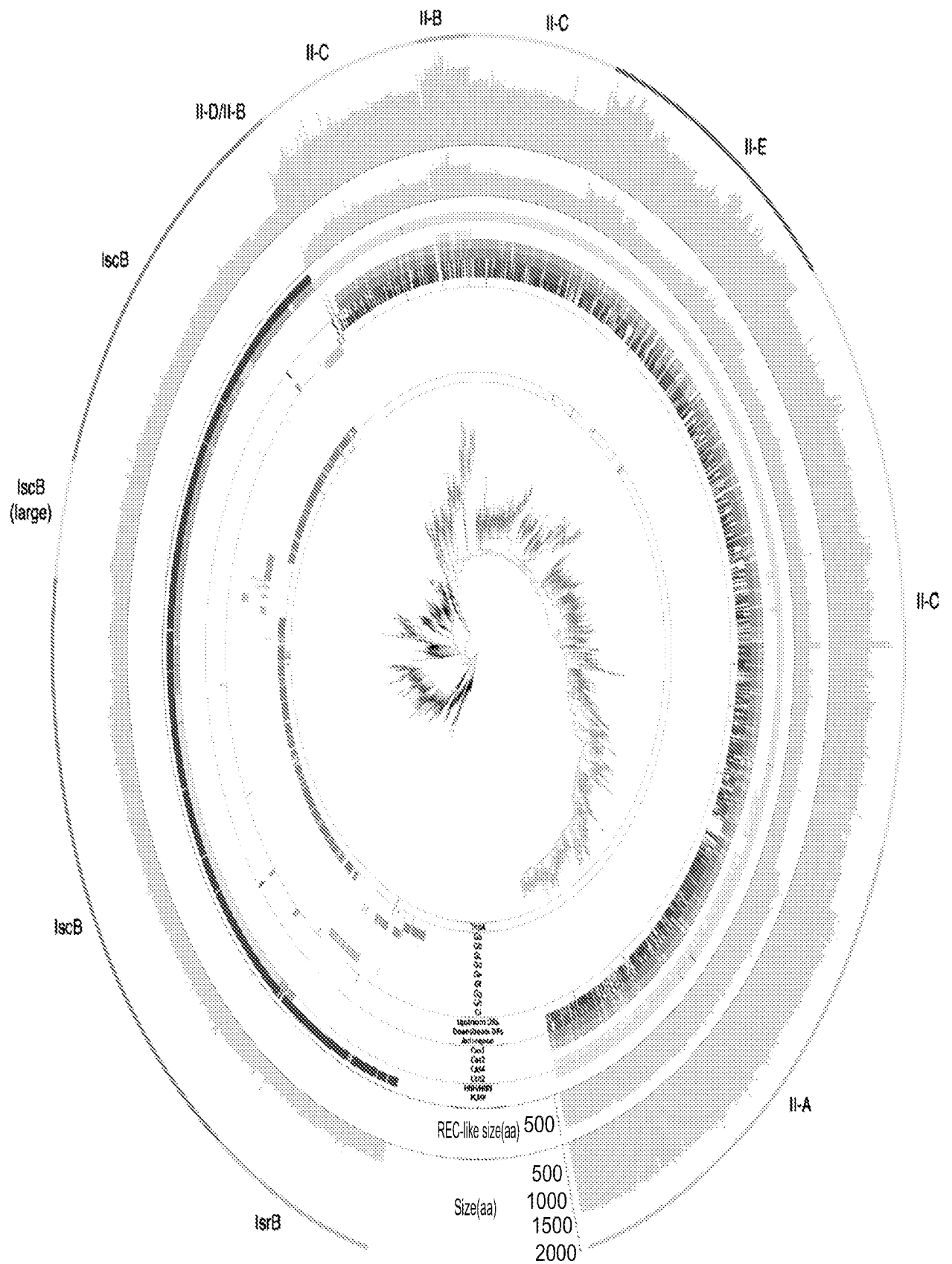


FIG. 52B

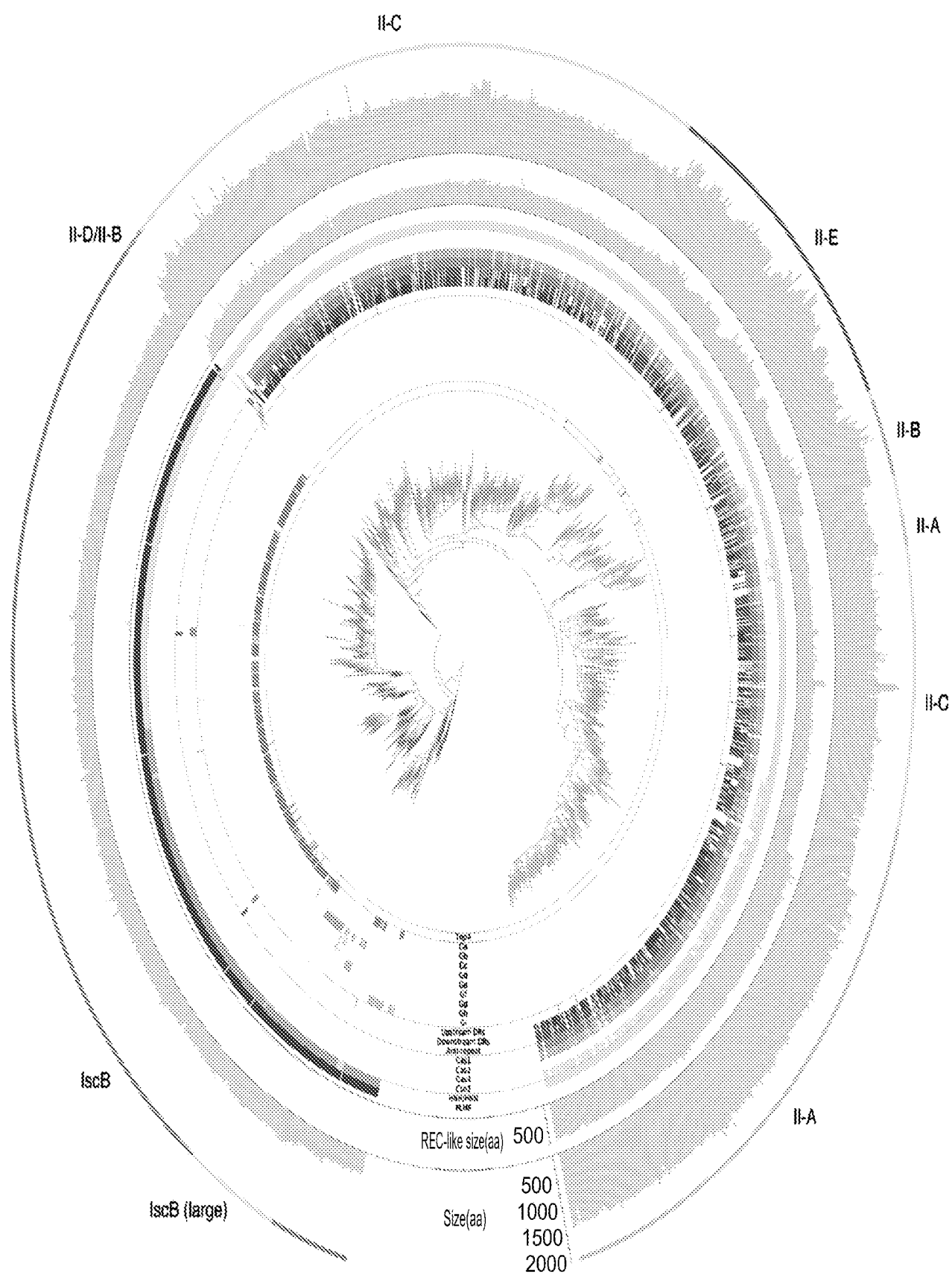


FIG. 53A

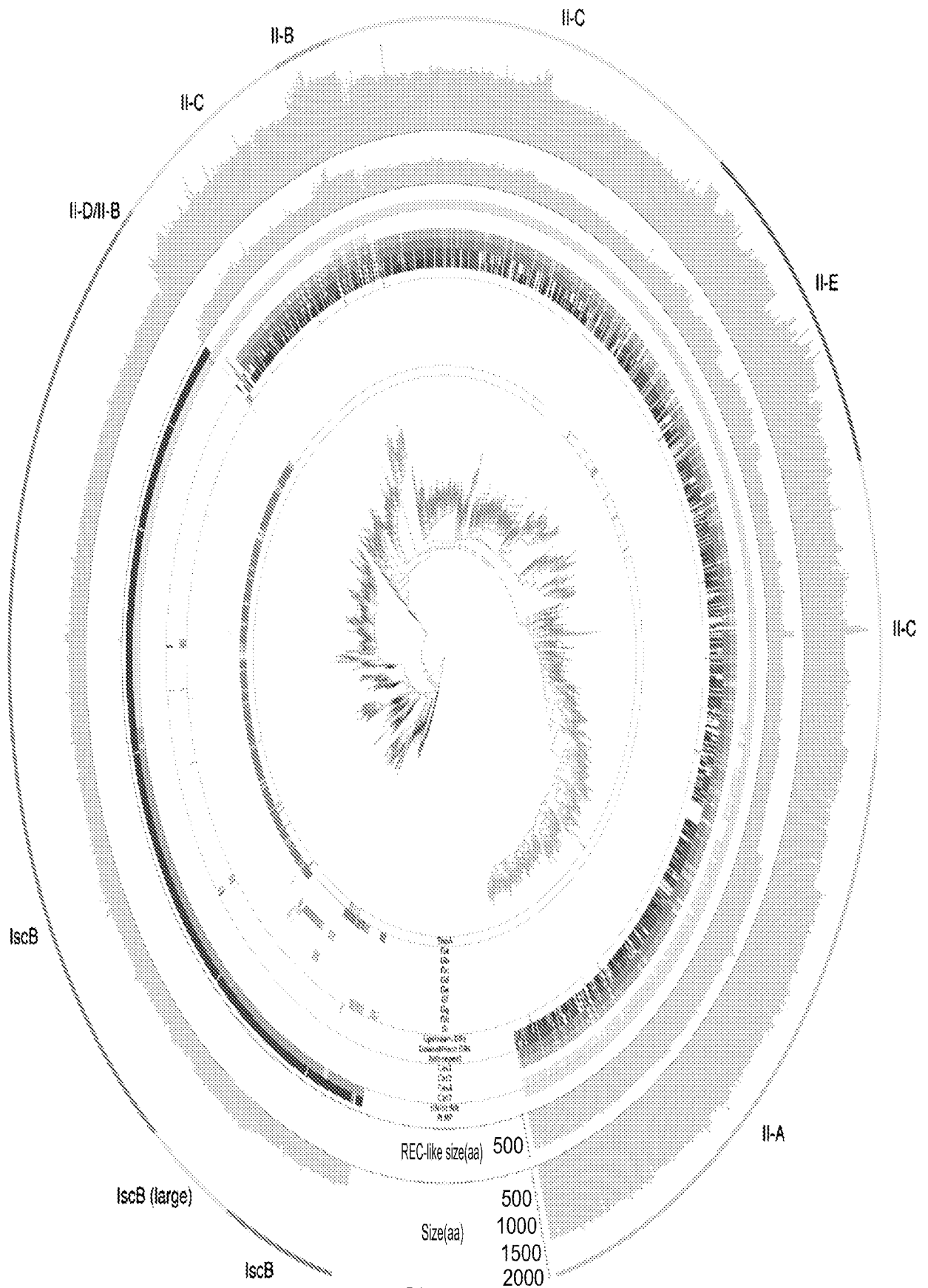


FIG. 53B

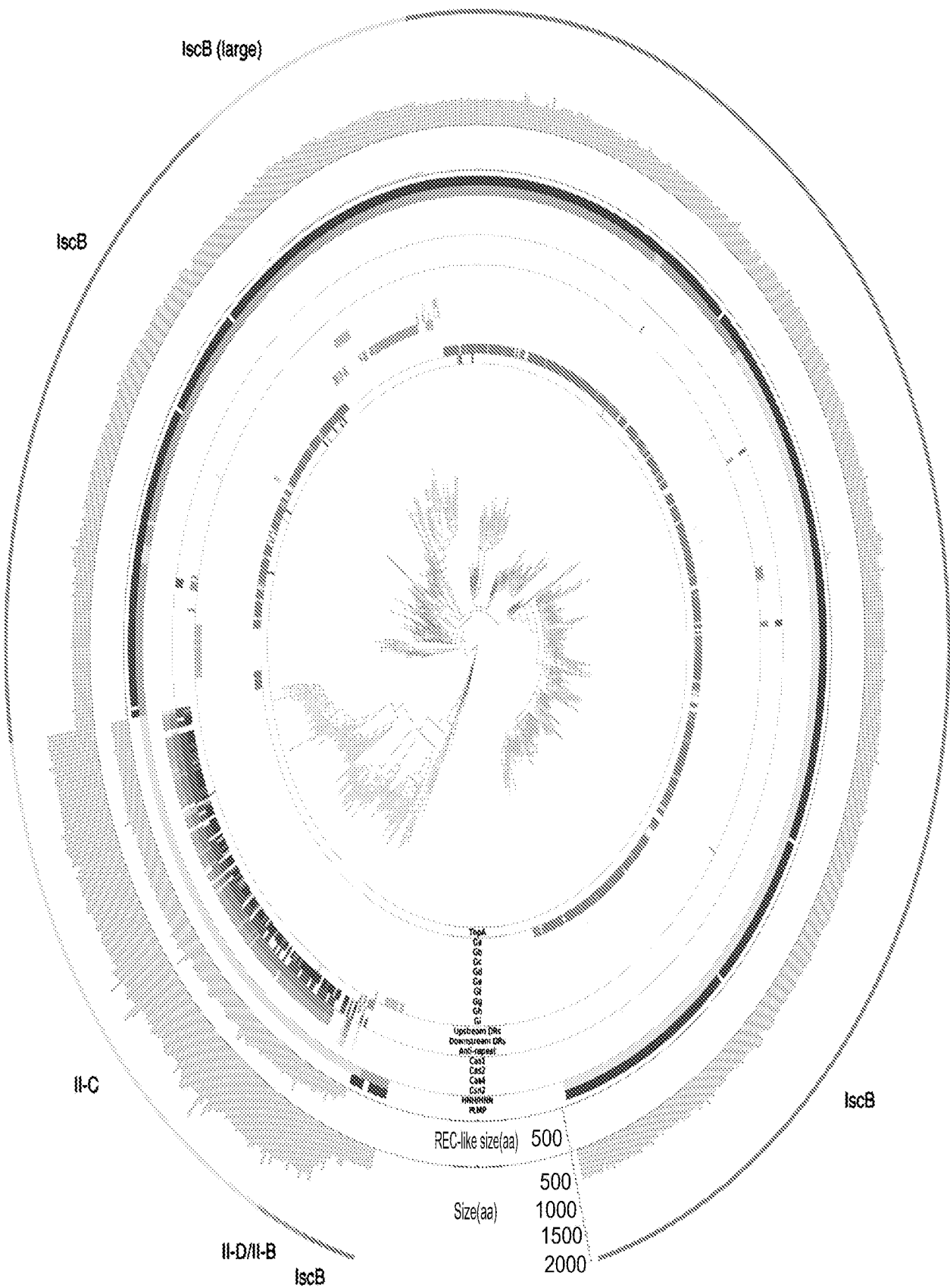
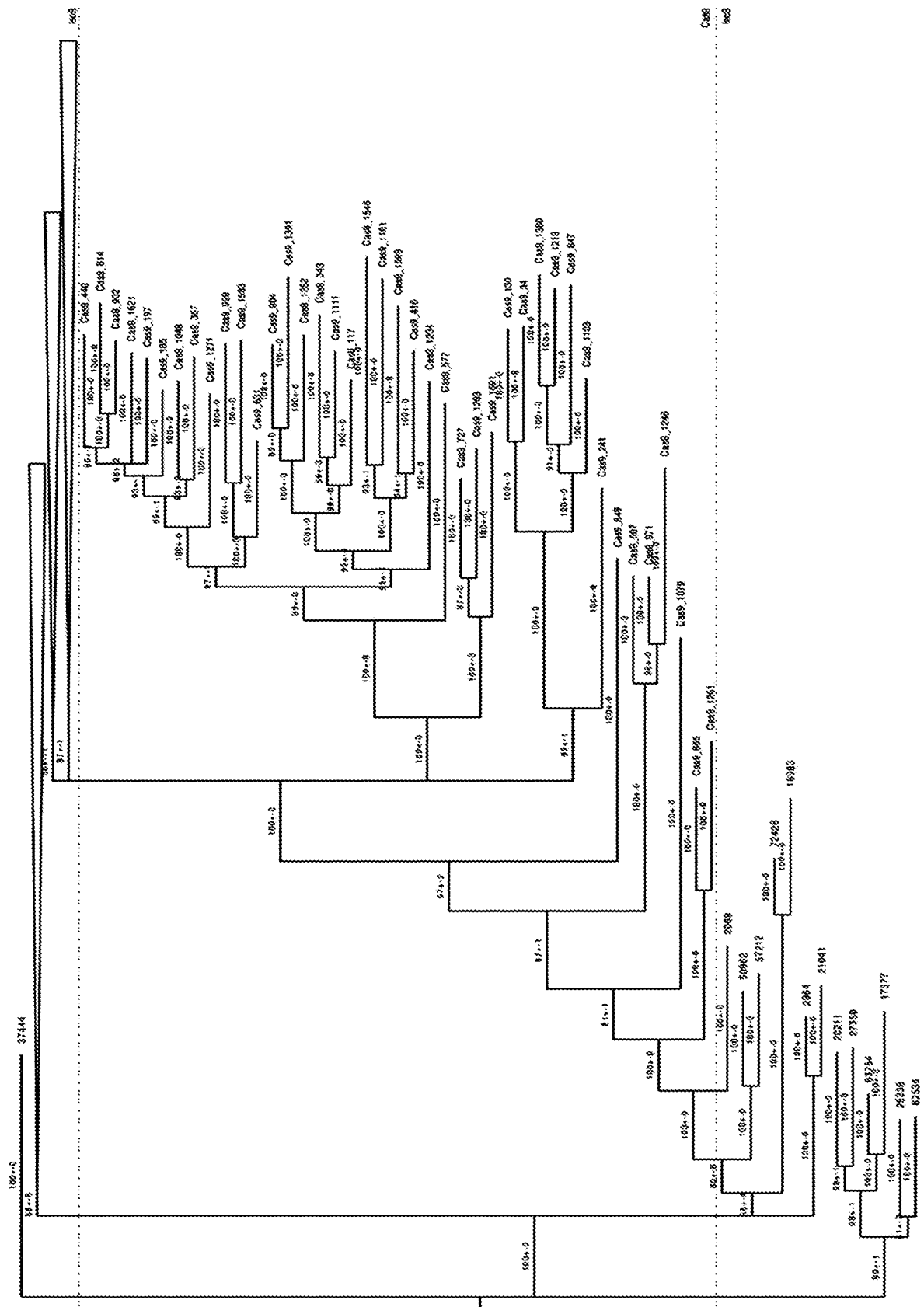


FIG. 54A



0.4
FIG. 54B

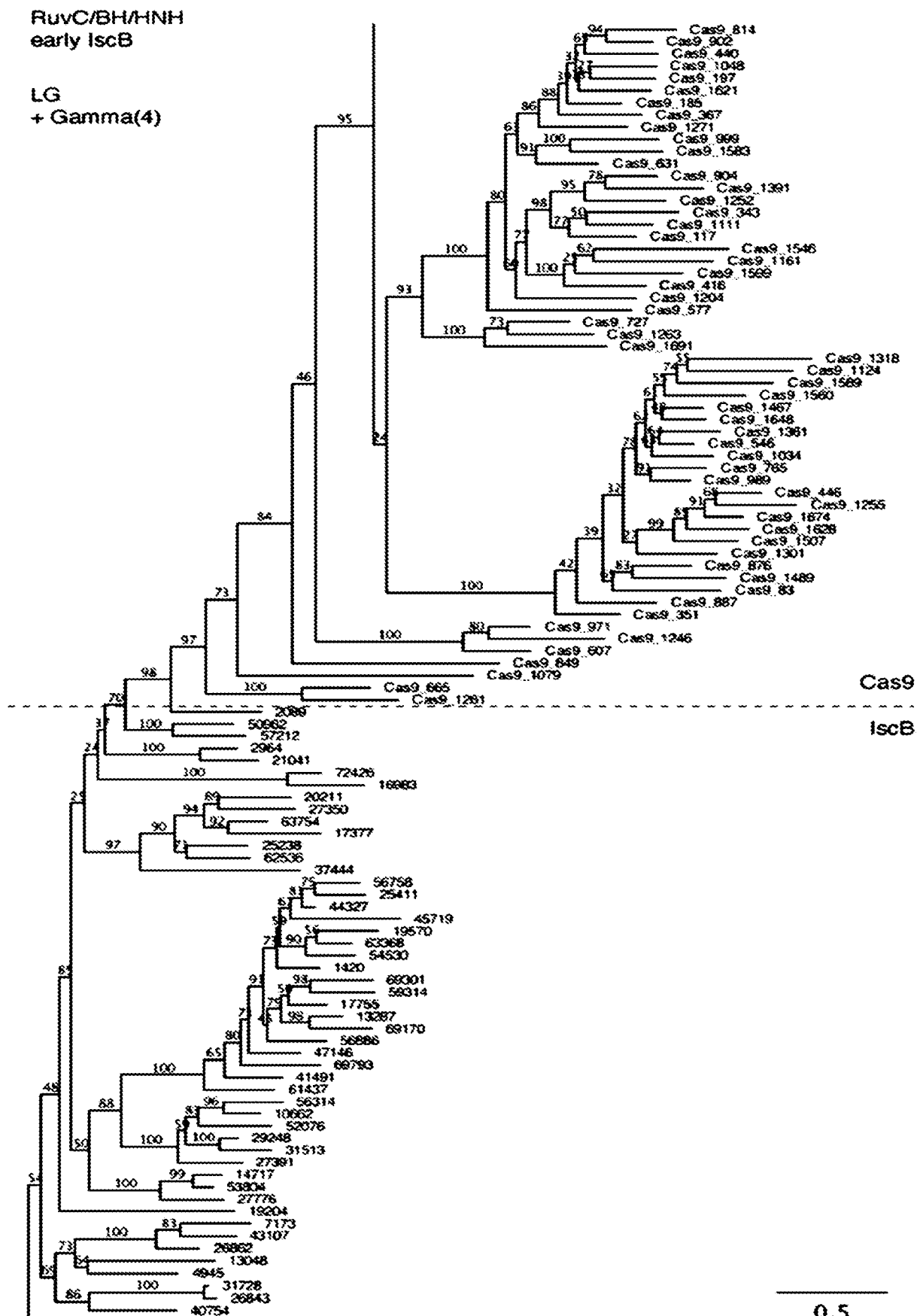
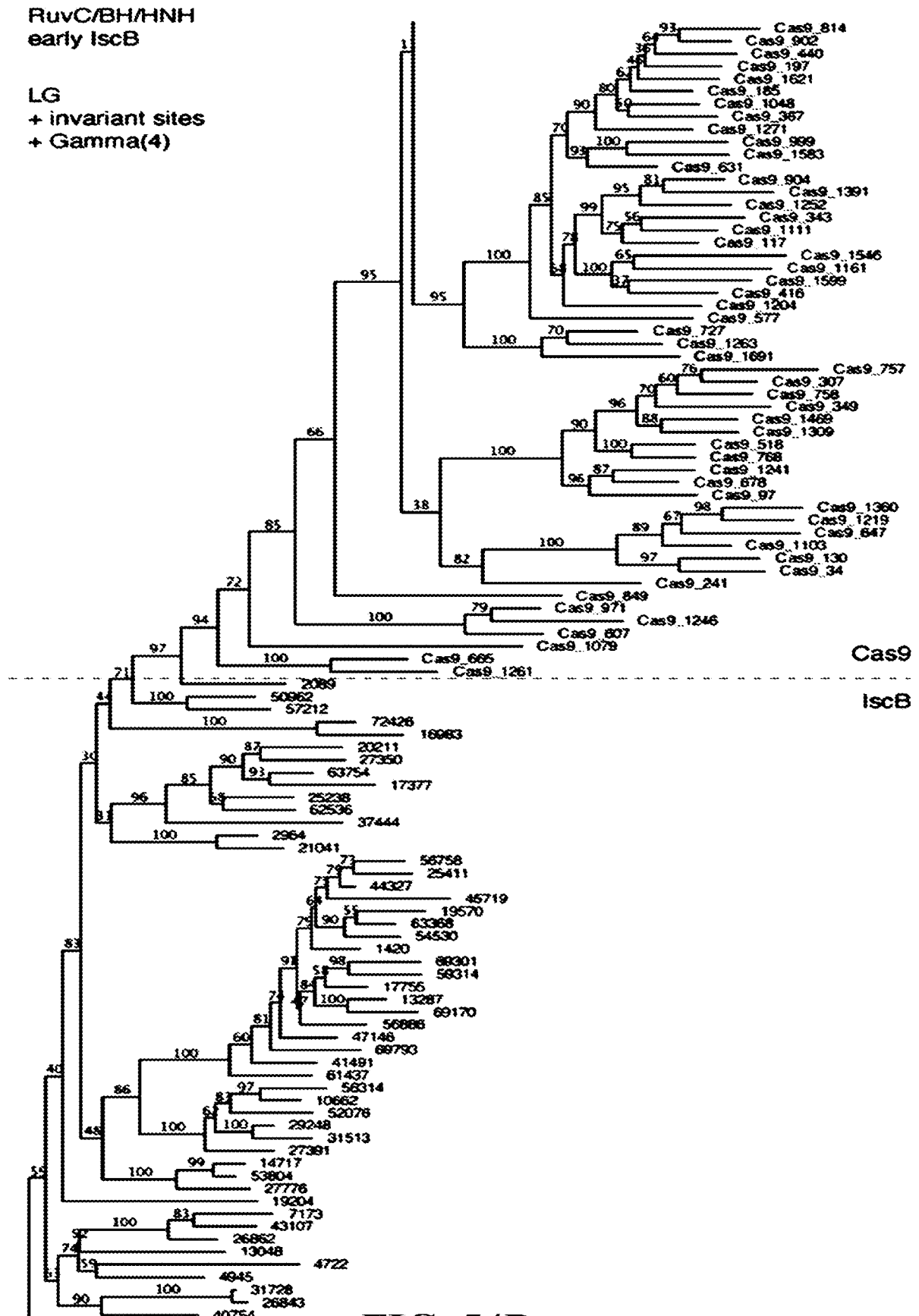


FIG. 54C



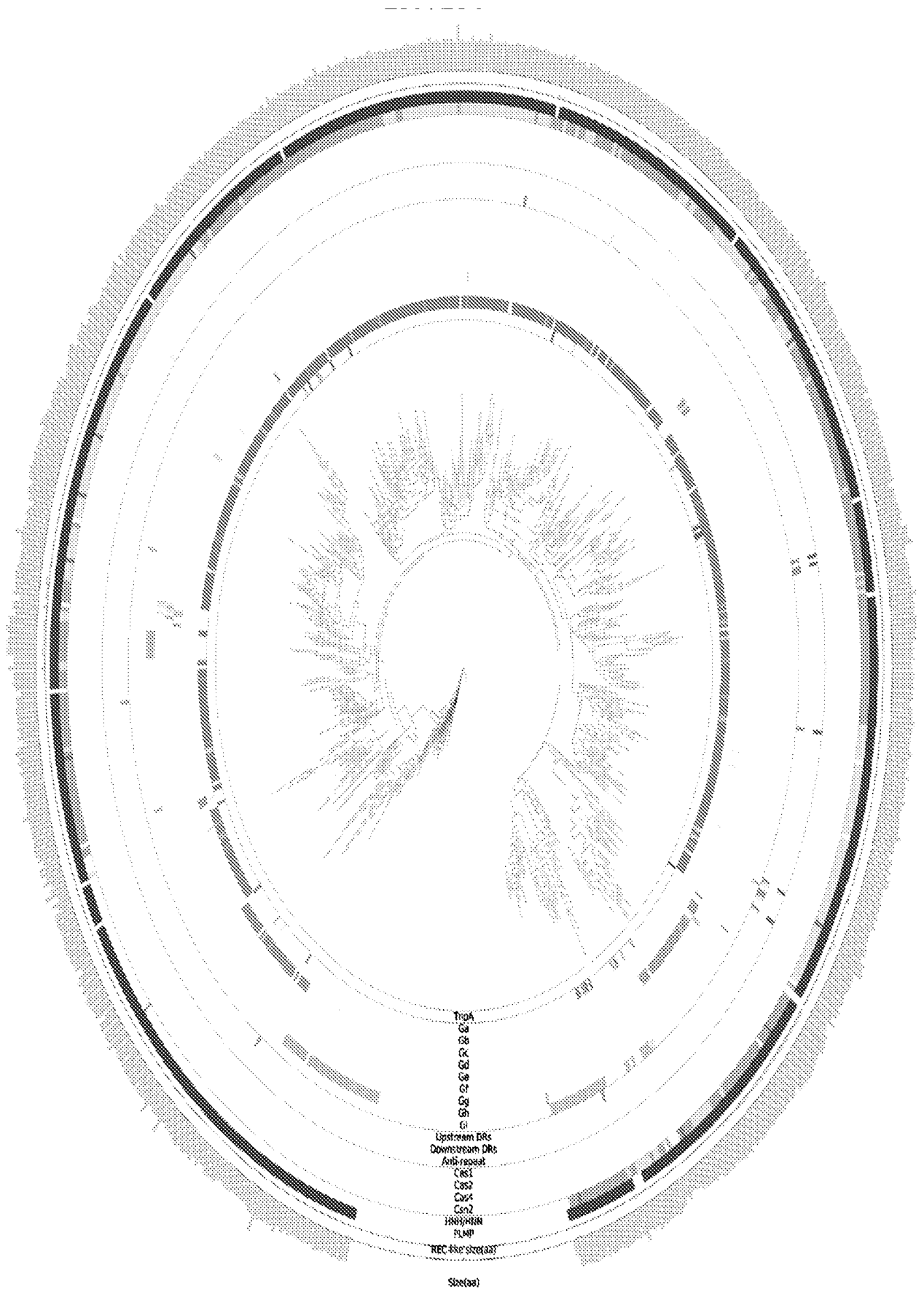
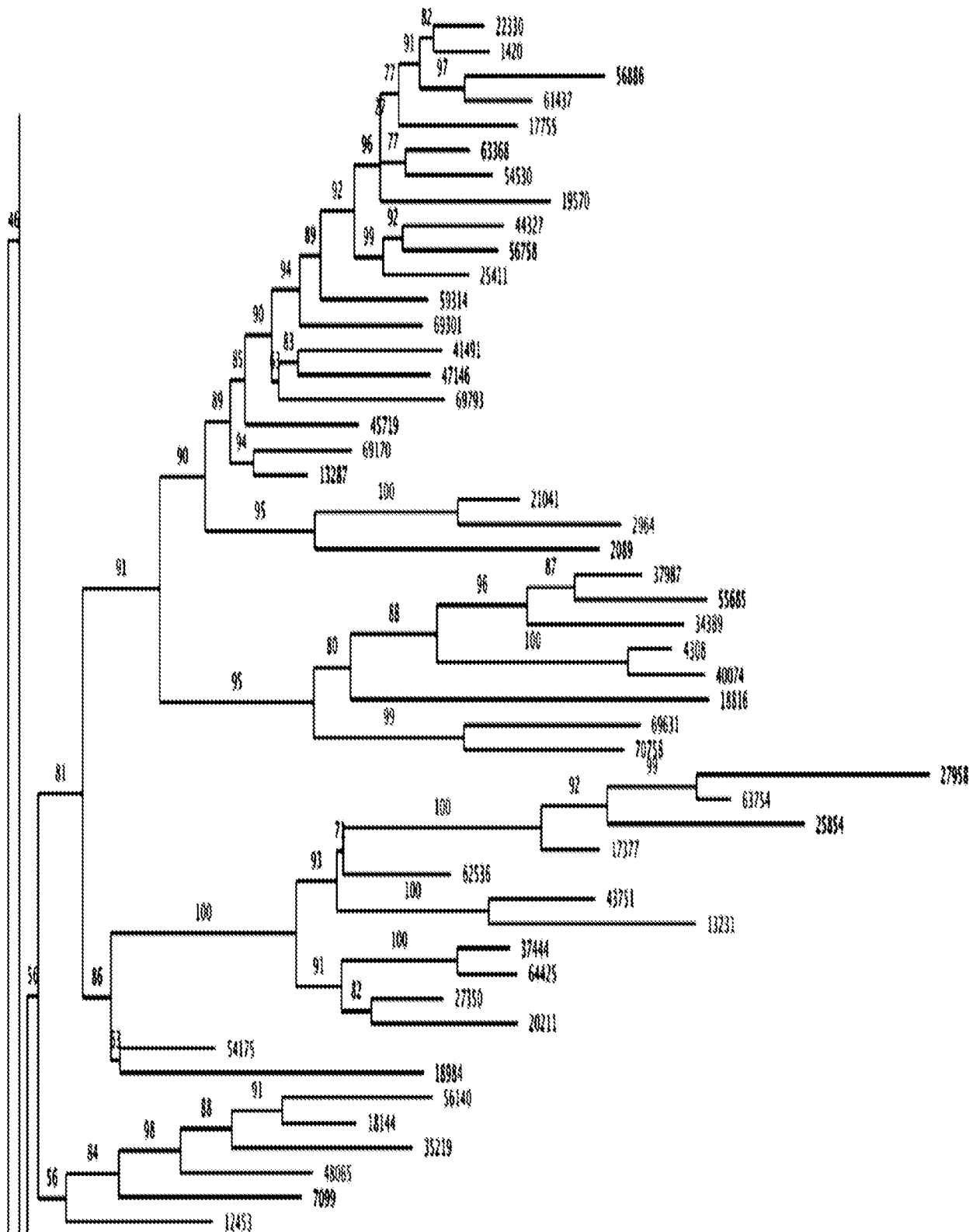


FIG. 55A

**FIG. 55B**

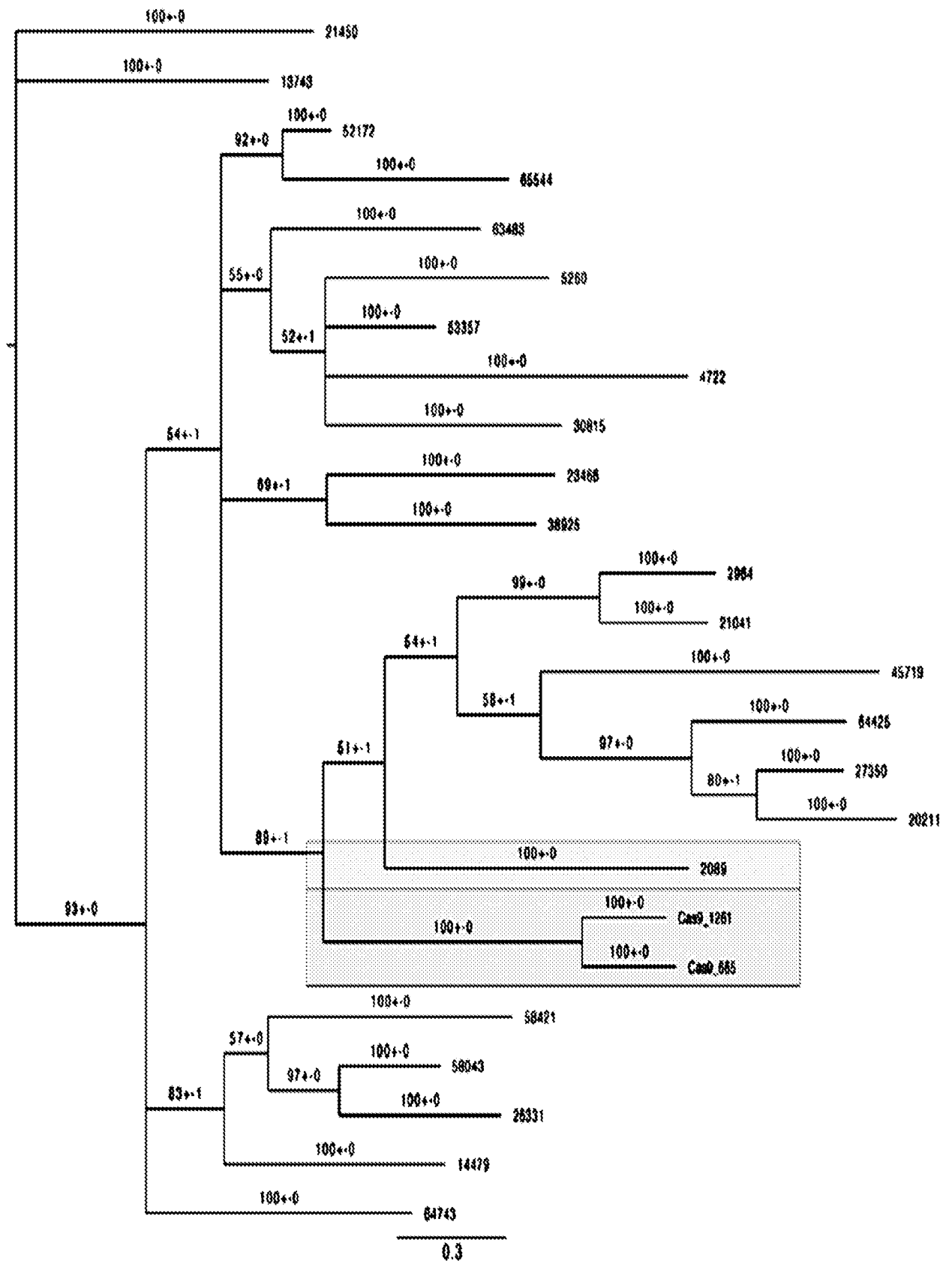


FIG. 55C

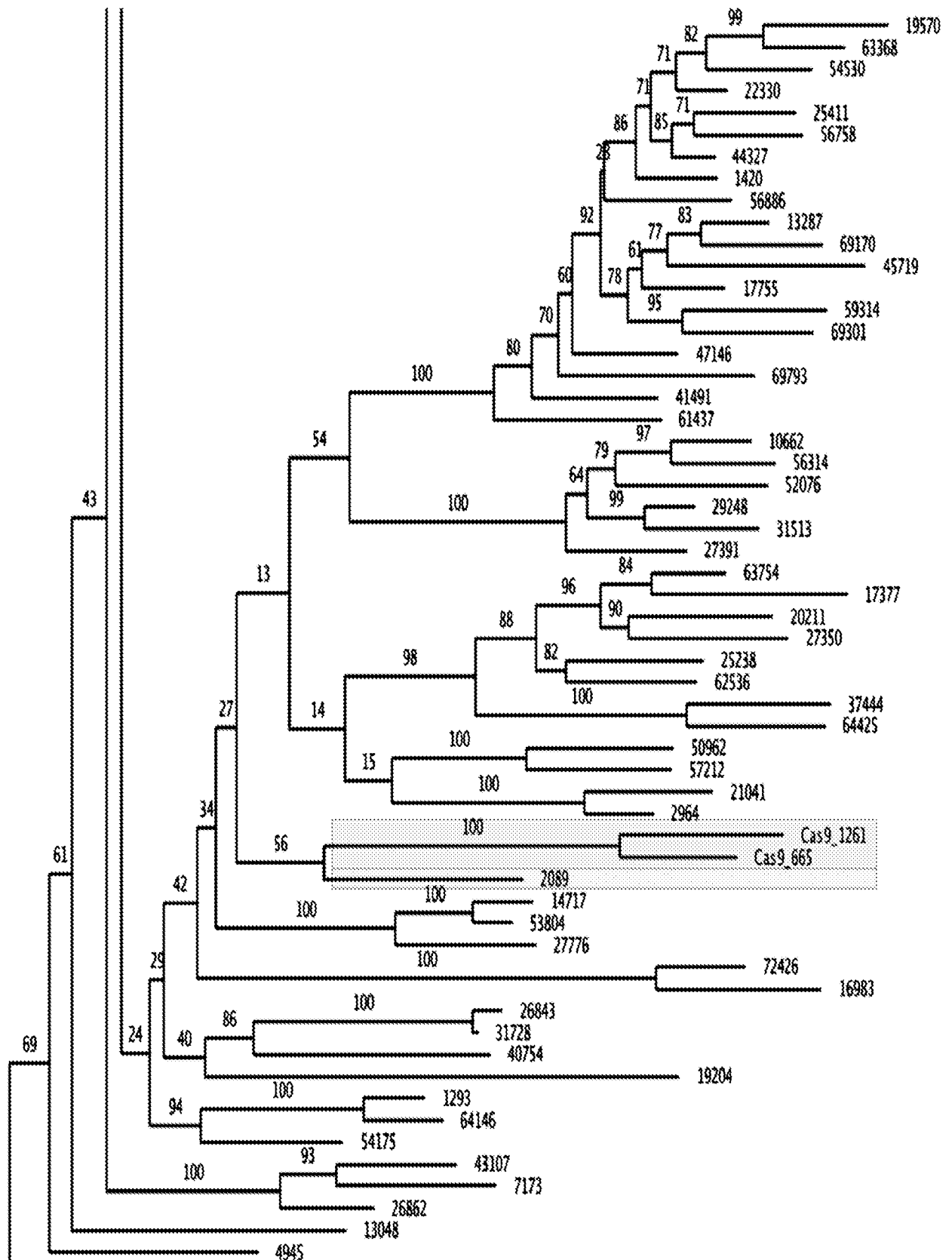


FIG. 56

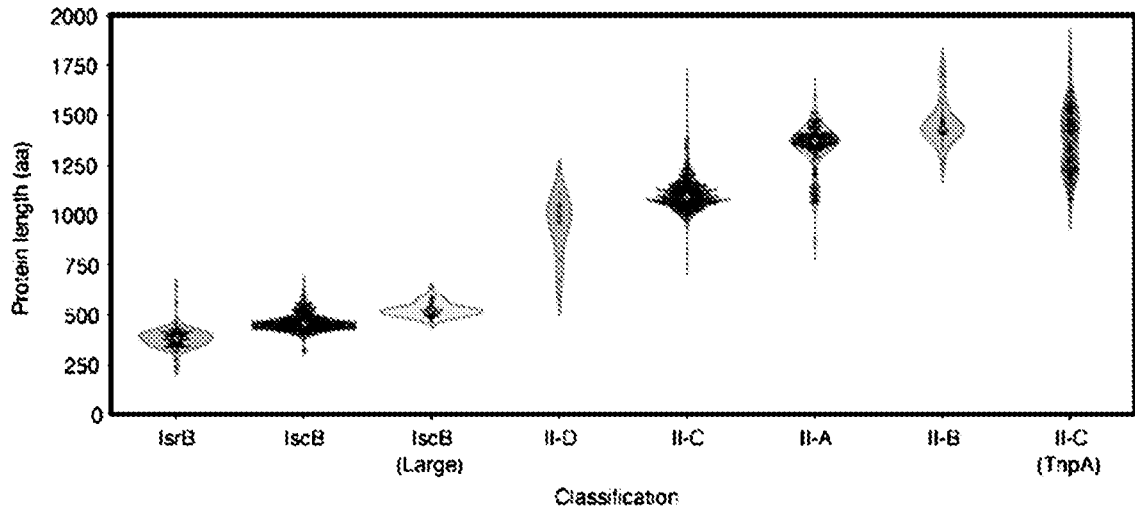


FIG. 57A

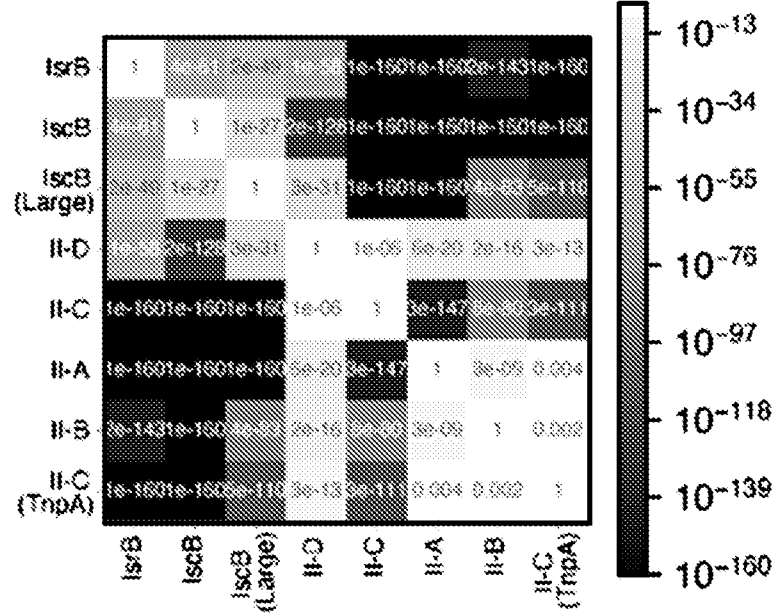


FIG. 57B

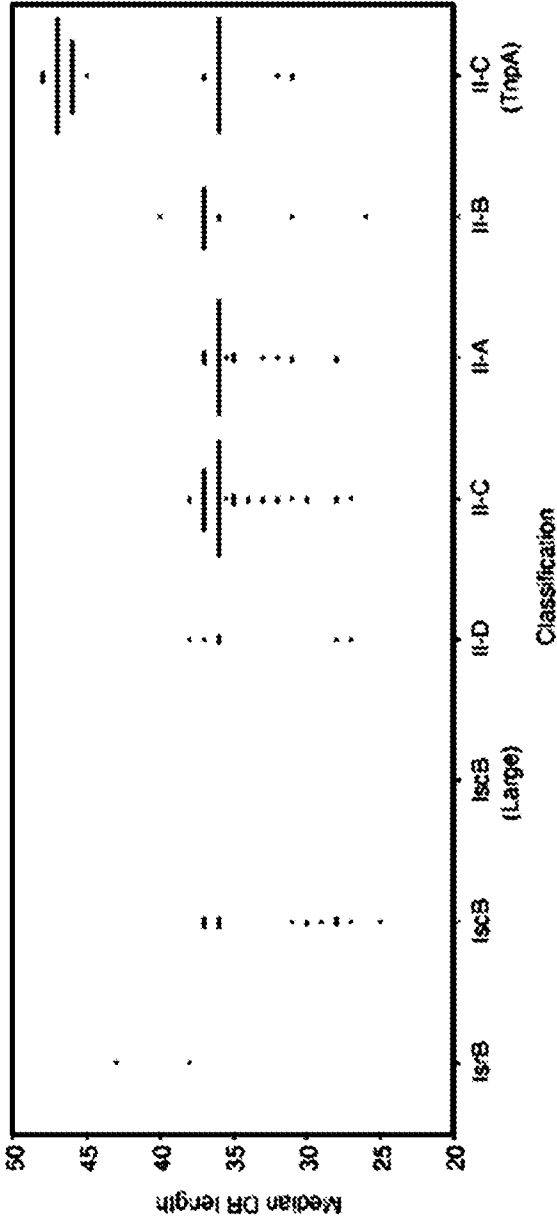


FIG. 57C

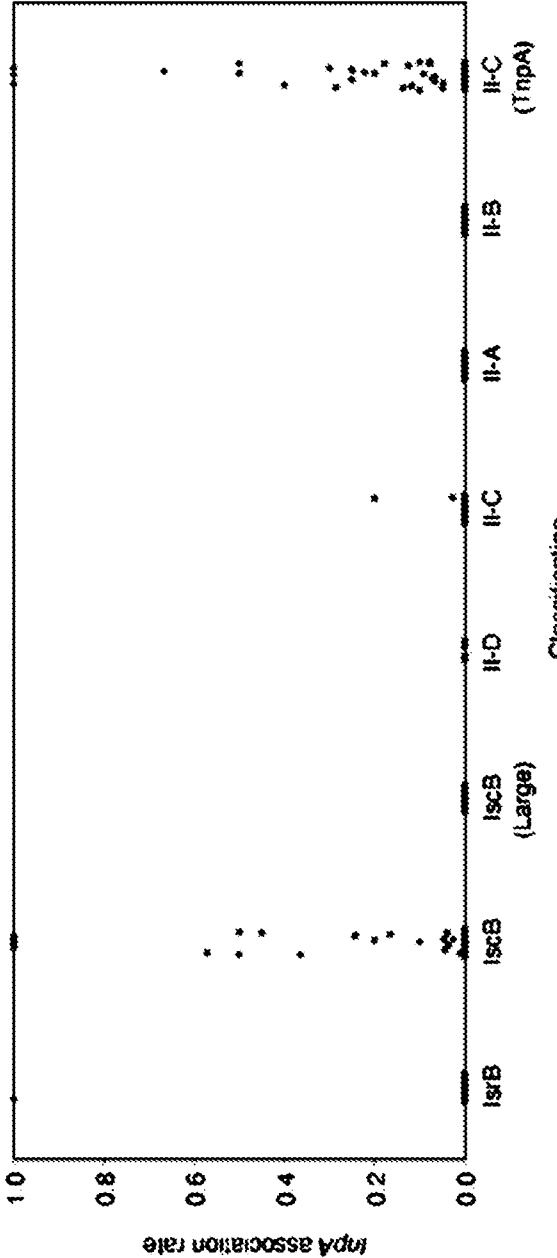
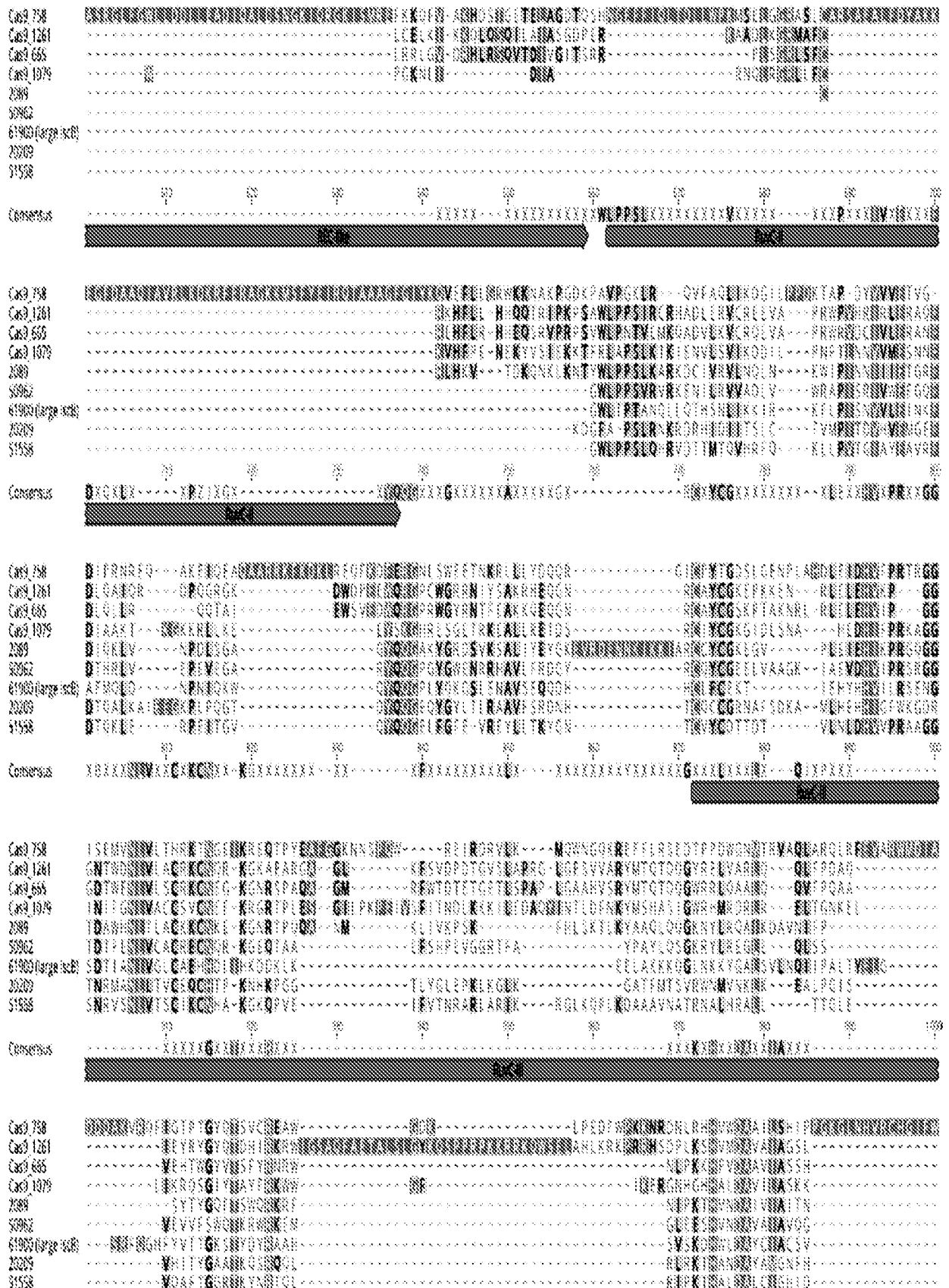
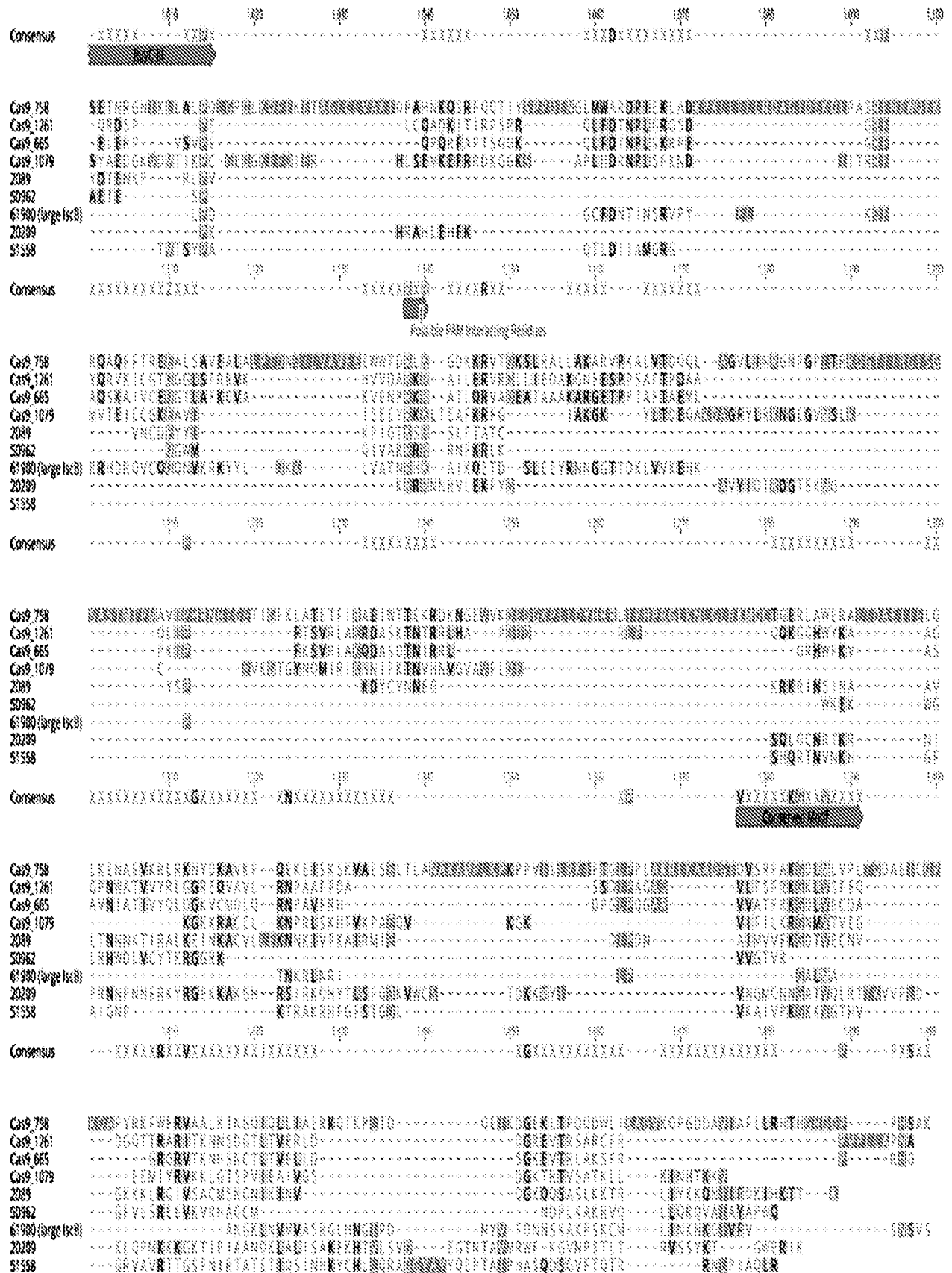


FIG. 57D

[illegible]

FIG. 58





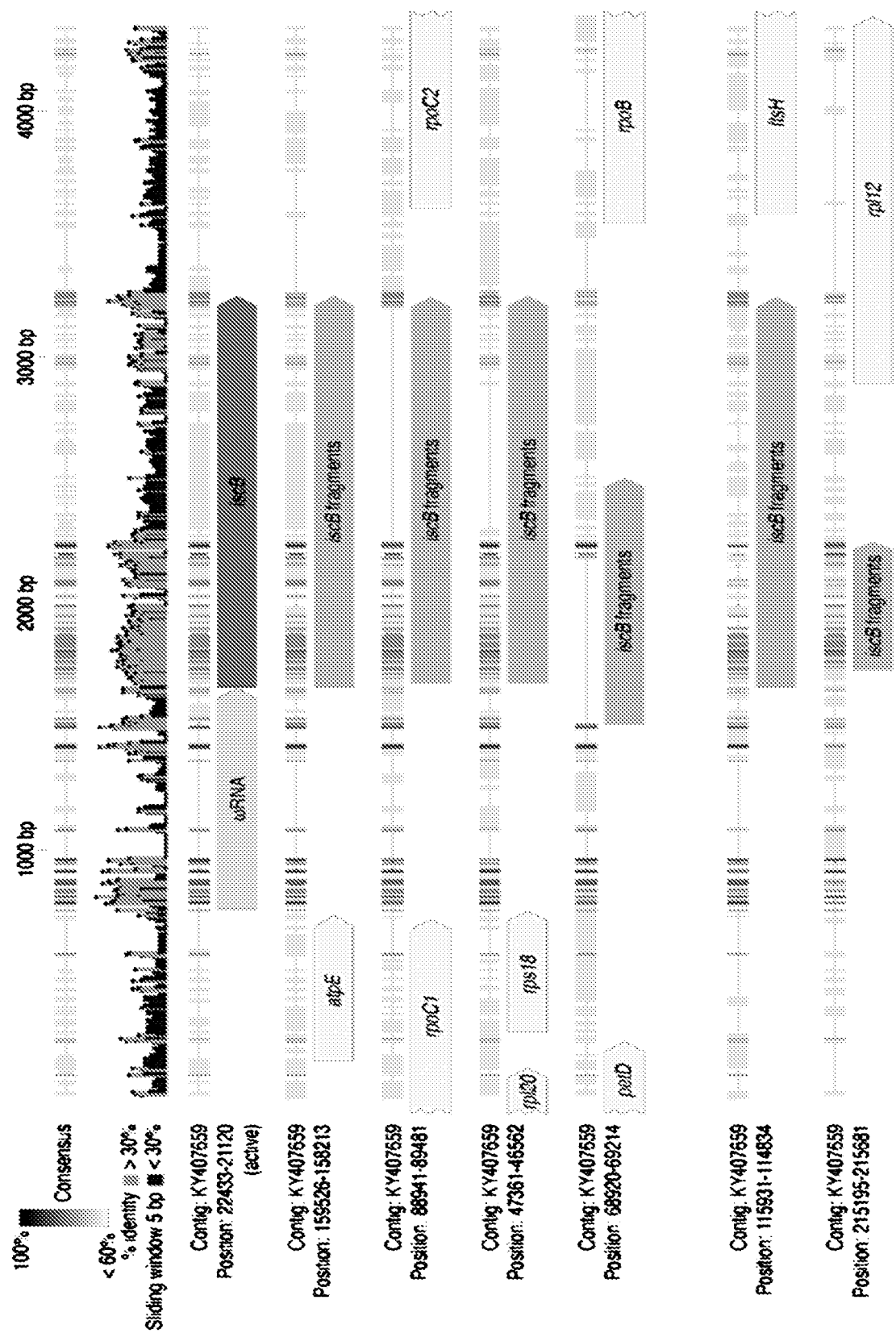


FIG. 59A

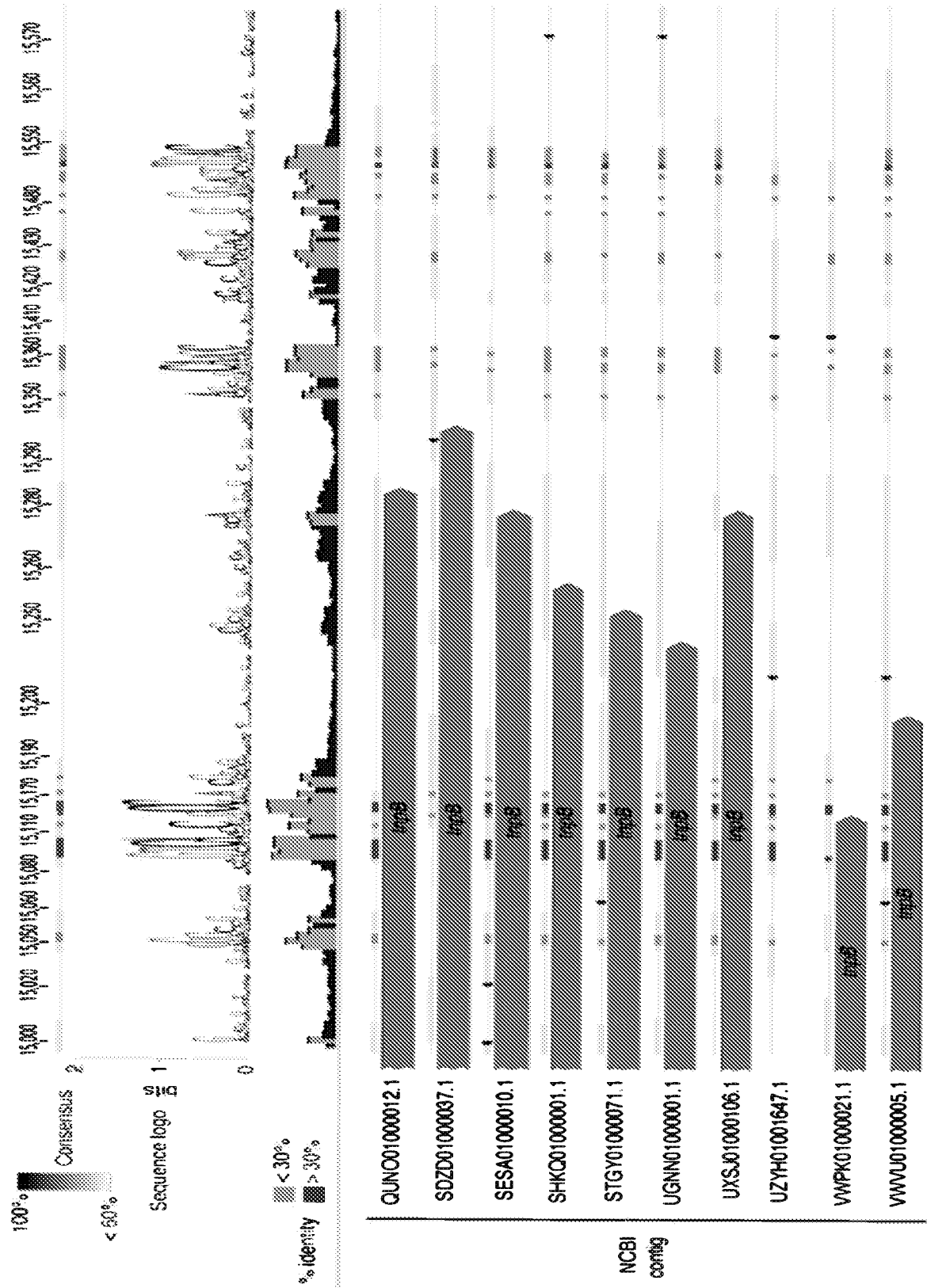


FIG. 60

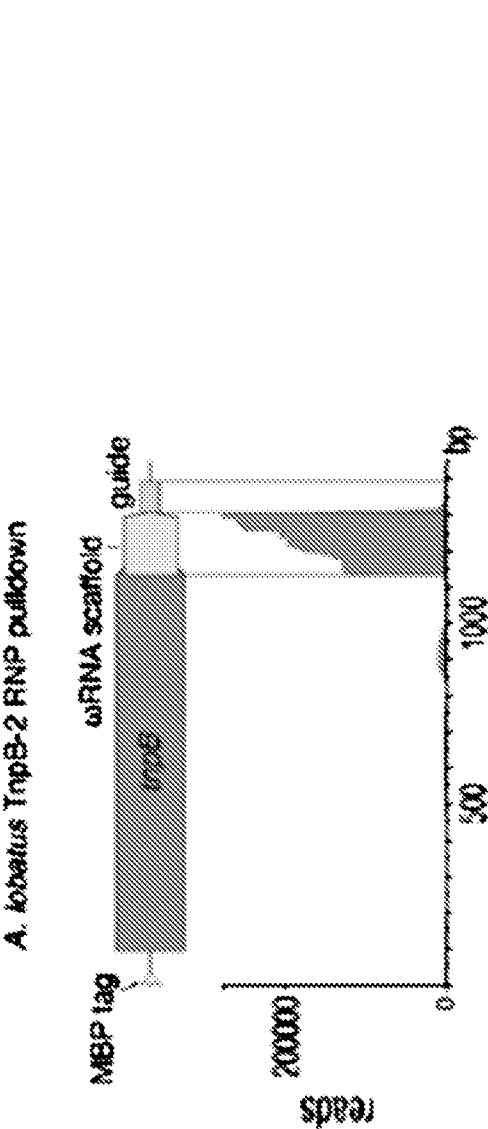


FIG. 61A

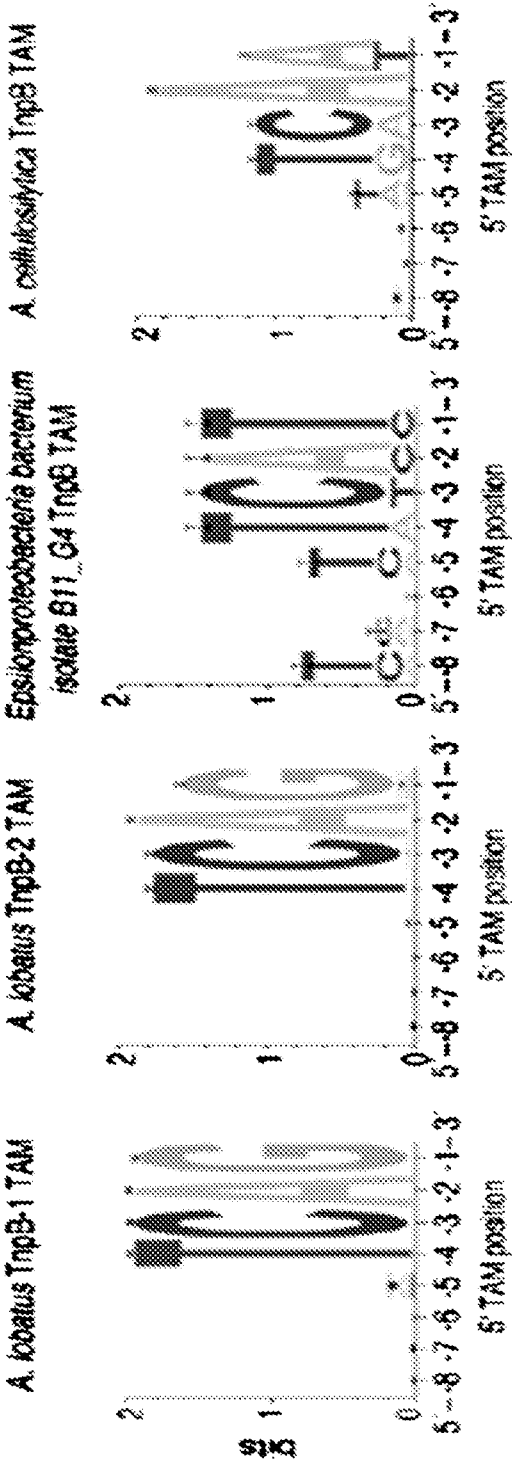


FIG. 61B

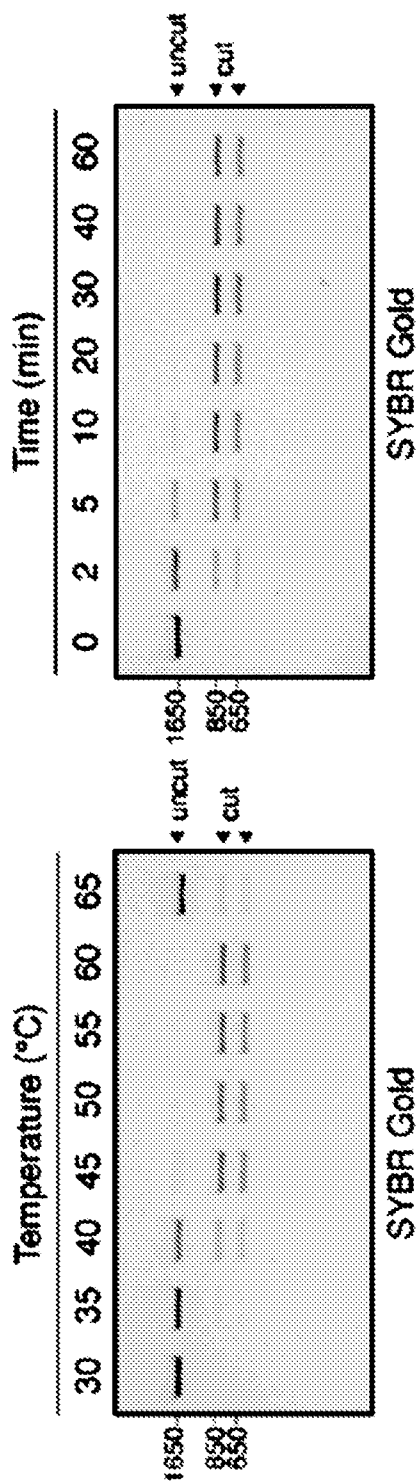


FIG. 61C

FIG. 61D

	ωRNA	+	+	+	-
target	+	+	+	+	-
TAM	+	-	-	-	+



FIG. 61E

FIG. 61F

Cy5

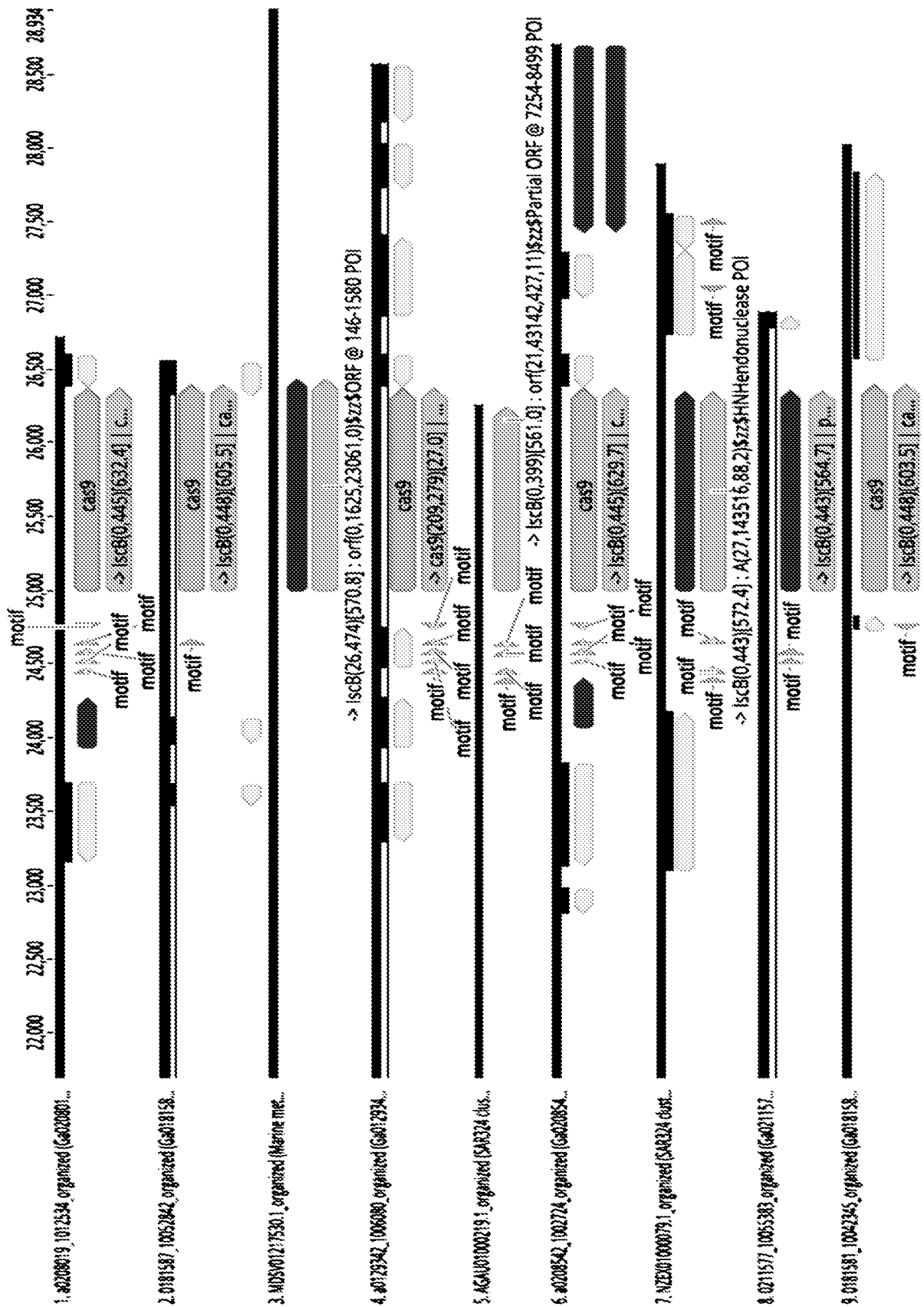
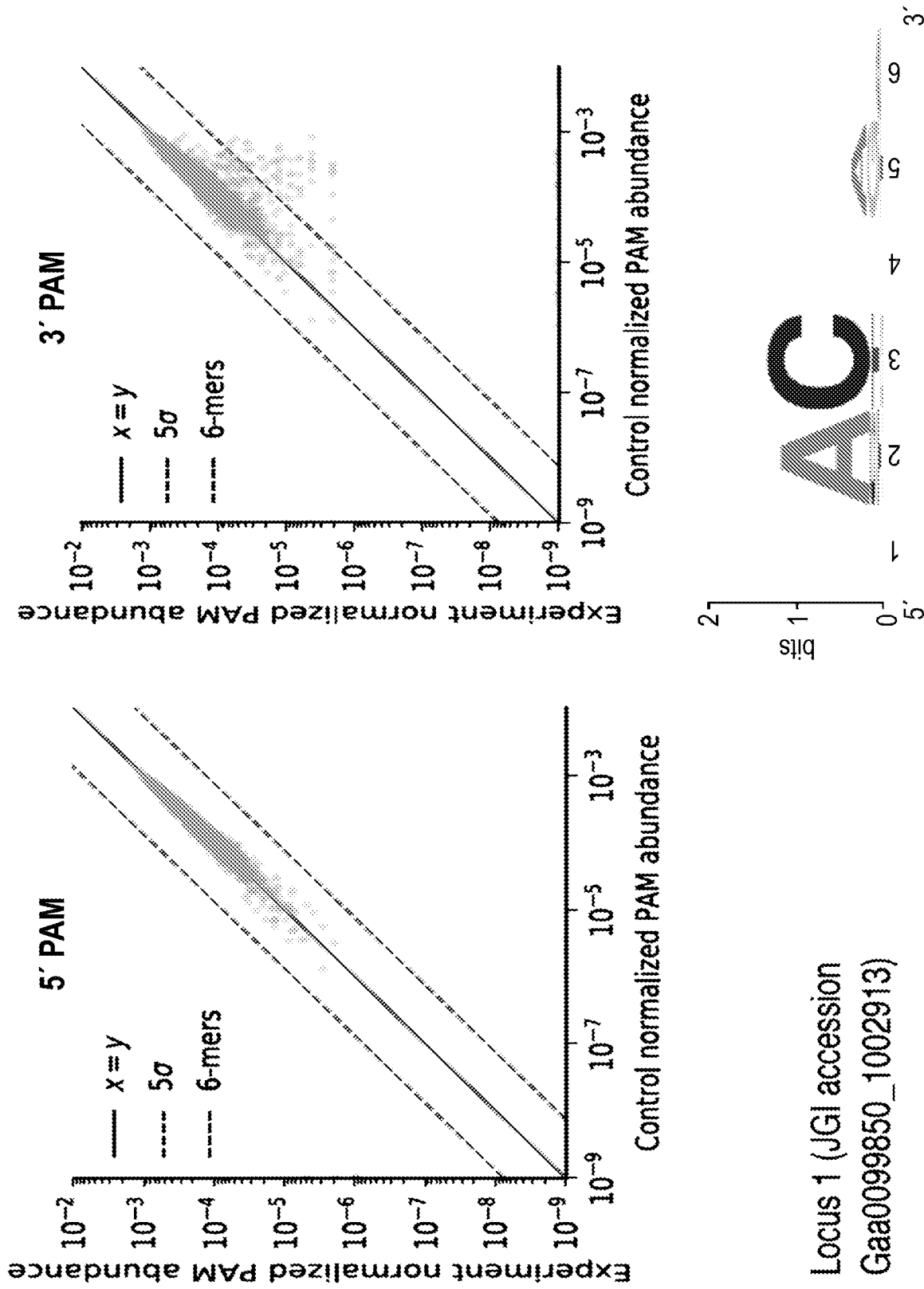


FIG. 62



Locus 1 (JGI accession
Gaa0099850_1002913)

FIG. 63A

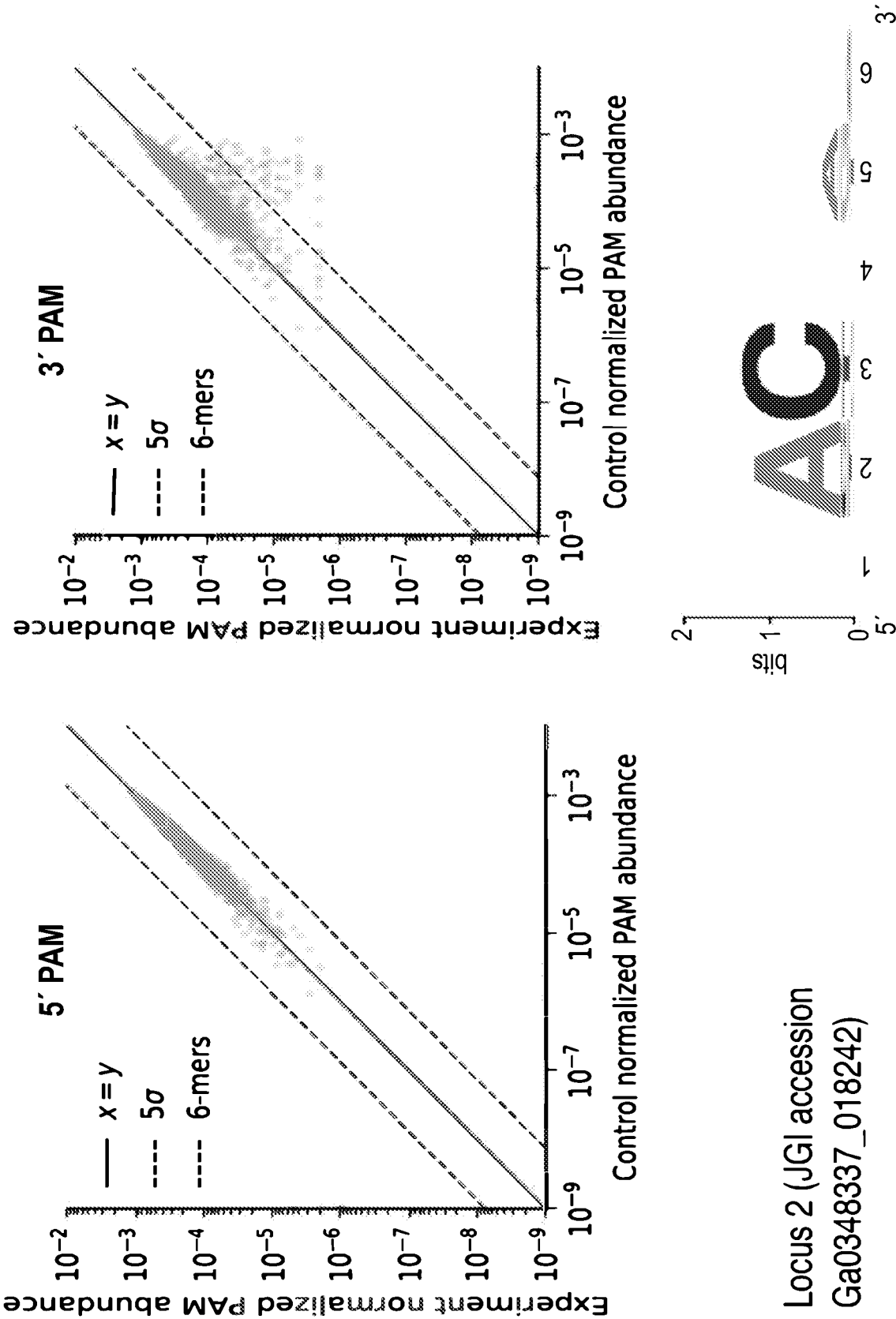
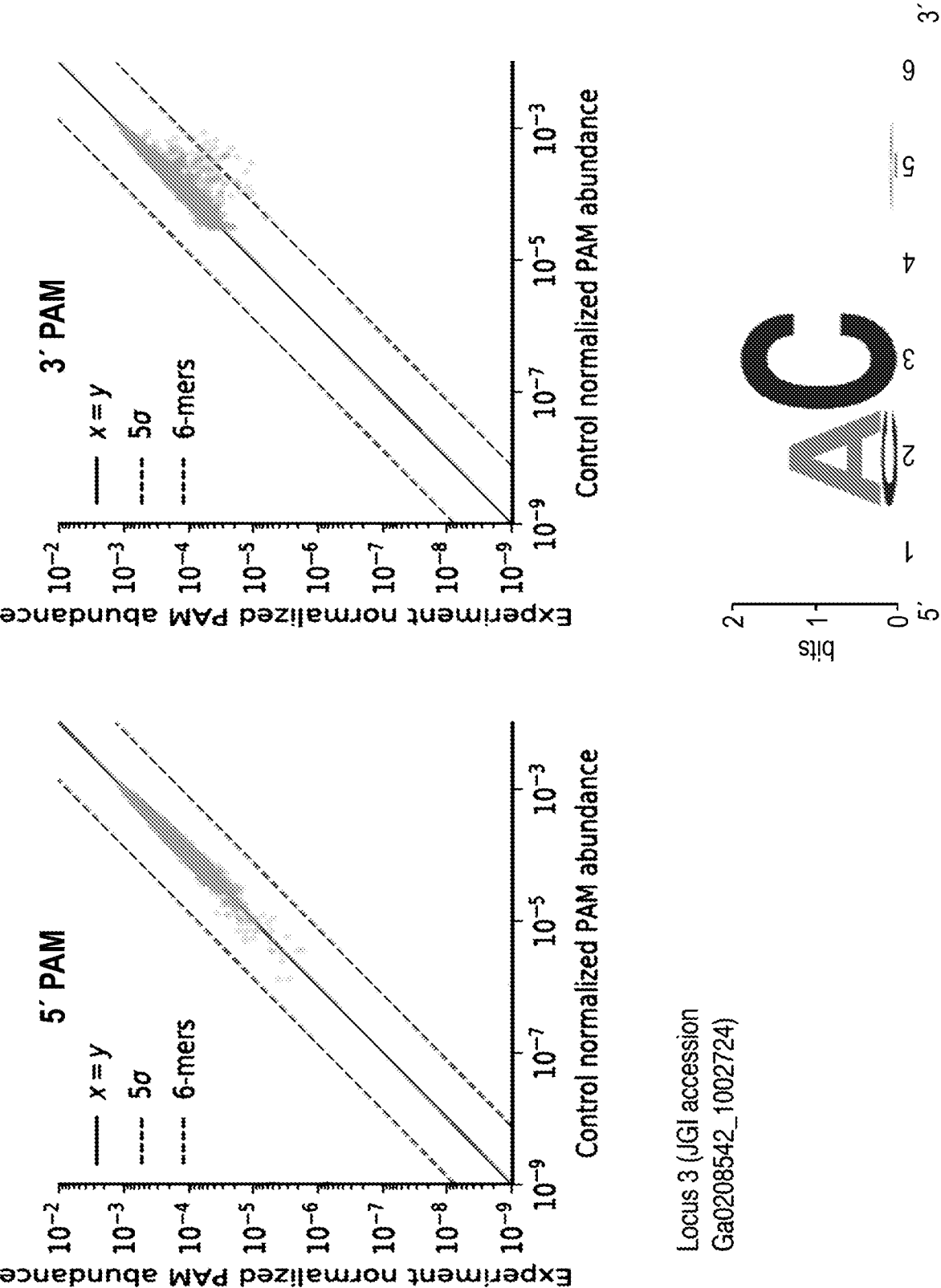


FIG. 63B



Locus 3 (JGI accession
Ga0208542_1002724)

FIG. 63C

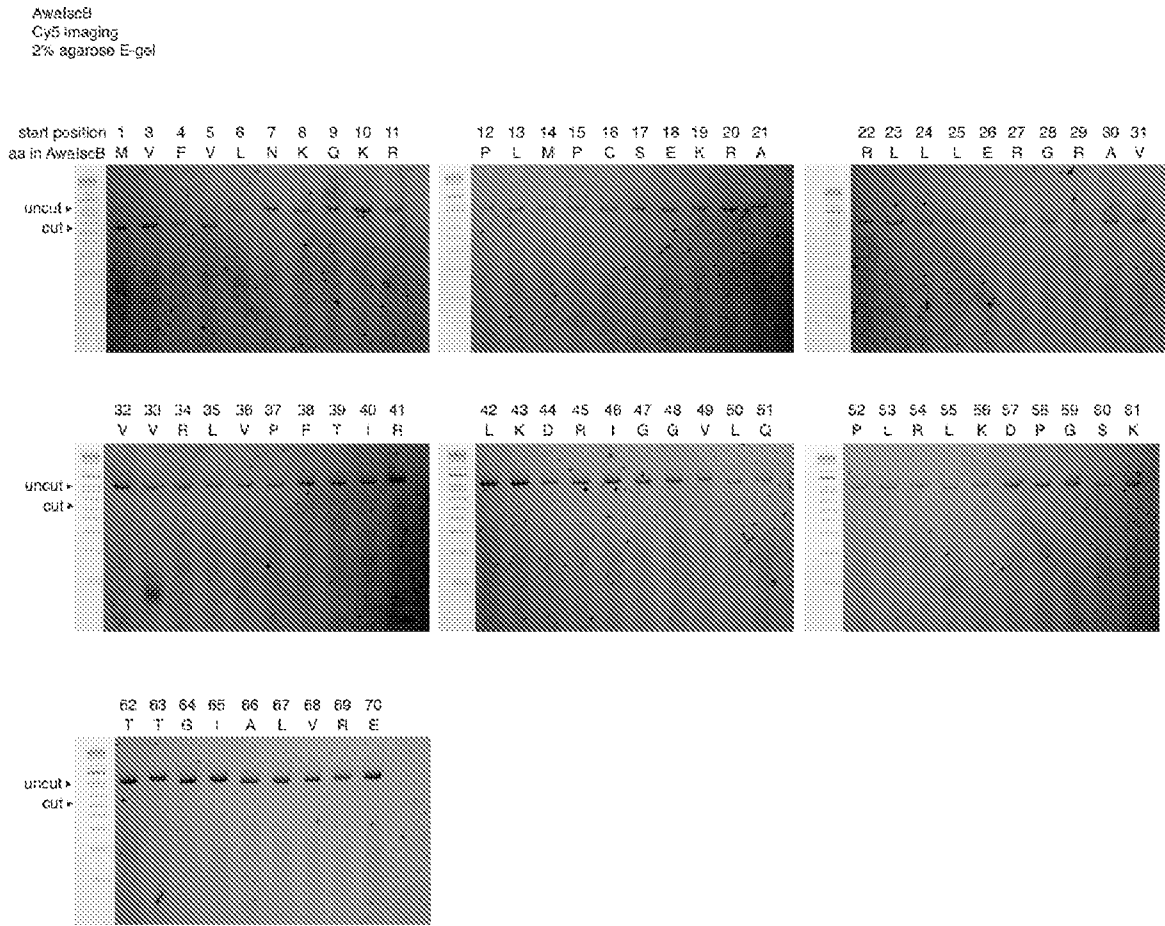


FIG. 64

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2021/056361

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - A61K 38/46; C12N 9/22; C12N 15/10; C12N 15/11; C12N 15/87 (2022.01)

CPC - C12N 9/22; C12N 15/102; C07K 2319/00; C07K 2319/80; C12N 2310/20 (2022.01)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

see Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

see Search History document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

see Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2020/191102 A1 (THE BROAD INSTITUTE INC. et al) 24 September 2020 (24.09.2020) entire document	1, 6, 16, 17, 22-29, 32
A	The UniProt Consortium, "UniProtKB - D6U4F6 (D6U4F6_9CHLR)" 10 August 2010 (10.08.2010), Pgs. 1-4, Retrieved from the Internet: <URL: https://www.uniprot.org/uniprot/D6U4F6 < ib 08 February 2022. entire document	1, 6, 16, 17, 22-29, 32
A	OSAKA et al. "GENBANK Sequence AP024096.1" 09 October 2020 (09.10.2020), Pgs. 1-2, Retrieved on 08 February 2022. Retrieved from the Internet: <URL: https://www.ncbi.nlm.nih.gov/nuccore/1915456248?sat=49&satkey=11026392 >. entire document	1, 6, 16, 17, 22-29, 32
A	WO 2020/191249 A1 (THE BROAD INSTITUTE INC. et al) 24 September 2020 (24.09.2020) entire document	1, 6, 16, 17, 22-29, 32
A	WO 2018/035250 A1 (THE BROAD INSTITUTE INC. et al) 22 February 2018 (22.02.2018) entire document	1, 6, 16, 17, 22-29, 32

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"D" document cited by the applicant in the international application

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

08 February 2022

Date of mailing of the international search report

FEB 24 2022

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents

P.O. Box 1450, Alexandria, VA 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Harry Kim

Telephone No. PCT Helpdesk: 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2021/056361

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

a. ☒ forming part of the international application as filed:

☒ in the form of an Annex C/ST.25 text file.

☐ on paper or in the form of an image file.

b. ☐ furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

c. ☐ furnished subsequent to the international filing date for the purposes of international search only:

☐ in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)).

☐ on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

SEQ ID NOs: 2041 and 2060 were searched.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2021/056361

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 8-13, 15, 18-21, 33-44, 47-53, 57-59, 62, 68-72
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See extra sheet(s).

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1, 6, 16, 17, 22-29, 32

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2021/056361

Continued from Box No. III Observations where unity of invention is lacking

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees need to be paid.

Group I+: claims 1-7, 14, 16, 17, 22-32, 45, 46, 54-56, 60, 61, and 63-67 are drawn to non-naturally occurring, engineered compositions comprising LscB polypeptides.

The first invention of Group I+ is restricted to a non-naturally occurring, engineered composition comprising a) an LscB polypeptide selected to be LscB SEQ ID NO:2041, and an omega RNA selected to be SEQ ID NO:2060. It is believed that claims 1, 6, 16, 17, 22-29, and 32 read on this first named invention and thus these claims will be searched without fee to the extent that they read on SEQ ID NOs: 2041 and 2060.

Applicant is invited to elect additional LscB polypeptides, omega RNAs, and their respective, corresponding, SEQ ID NOs to be searched in a specific combination by paying additional fee for each set of election. An exemplary election would be a non-naturally occurring, engineered composition comprising a) an LscB polypeptide selected to be LscB SEQ ID NO:2042, and an omega RNA selected to be SEQ ID NO:2064. Additional LscB polypeptides, omega RNAs, and their respective, corresponding, SEQ ID NOs will be searched upon the payment of additional fees. Applicants must specify the claims that read on any additional elected inventions. Applicants must further indicate, if applicable, the claims which read on the first named invention if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched/examined.

The inventions listed in Groups I+ do not relate to a single general inventive concept under PCT Rule 13.1, because under PCT Rule 13.2 they lack the same or corresponding special technical features for the following reasons:

The Groups I+ formulas do not share a significant structural element responsible for genome editing, requiring the selection of alternative LscB polypeptides and omega RNAs where "wherein the LscB protein shares at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with a LscB protein selected from Tables 2 and 3."

Additionally, even if Groups I+ were considered to share the technical features of a non-naturally occurring, engineered composition comprising a) an LscB polypeptide comprising a split RuvC nuclease domain comprising RuvC-I, RuvC-II, and RuvC-III subdomains, an HNH domain or both and b) an omega RNA molecule comprising a scaffold and a reprogrammable spacer sequence, the omega RNA molecule capable of forming a complex with the LscB polypeptide and directing the LscB polypeptide to a target polynucleotide; an engineered, non-naturally occurring composition comprising an LscB protein, wherein the LscB protein comprises an N-terminal X domain, a RuvC domain, a Bridge Helix domain, and a C-terminal Y domain. However, these shared technical features do not represent a contribution over the prior art.

Specifically, WO 2020/191102 A1 to The Broad Institute Inc. et al. discloses a non-naturally occurring, engineered composition (engineered composition comprising a Type VII CRISPR- Cas effector protein and one or more CRISPR system nucleic acid components, Para. [0015]) comprising a) an LscB polypeptide comprising a split RuvC nuclease domain comprising RuvC-I, RuvC-II, and RuvC-III subdomains, an HNH domain or both (said Cas split domains (RuvC and HNH domains in the case of Cas9) can be simultaneously or sequentially introduced into the cell such that said split Cas domain(s) process the target nucleic acid sequence in the algae cell, Para. [0808]; LscB Nucleases (which are HNH Nucleases with a split RuvC and a Bridge Helix domain), Para. [0025]); an engineered, non-naturally occurring composition (engineered composition comprising a Type VII CRISPR- Cas effector protein and one or more CRISPR system nucleic acid components, Para. [0015]) comprising an LscB protein (said Cas split domains (RuvC and HNH domains in the case of Cas9) can be simultaneously or sequentially introduced into the cell such that said split Cas domain(s) process the target nucleic acid sequence in the algae cell, Para. [0808]; LscB Nucleases (which are HNH Nucleases with a split RuvC and a Bridge Helix domain), Para. [0025]; LscB proteins, Para. [0027]), wherein the LscB protein comprises an N-terminal X domain (N-terminus of the deaminase domain, Para. [0146]), a RuvC domain, a Bridge Helix domain, and a C-terminal Y domain (split RuvC domain comprises structural domains linked in the following order: RuvC-I, Bridge Helix, RuvC-II, HNH, RuvC- III, Para. [0018]).

Further, WO 2020/191249 A1 to The Broad Institute Inc. et al. and b) an omega RNA molecule (a specialized guide RNA (i.e., PEGRNA) containing a spacer sequence that anneals to a complementary protospacer in the target DNA, Para. [0009]) comprising a scaffold and a reprogrammable spacer sequence, the omega RNA molecule capable of forming a complex with the polypeptide and directing the polypeptide to a target polynucleotide (the gRNA scaffold. Homopolymer tailing of the product with terminal transferase followed by Klenow extension and Sanger sequencing revealed that the full gRNA synthesis template was copied in addition to the terminal 3 nucleotides of the gRNA core. In the future, the flap terminus will be reprogrammed by modifying the terminal gRNA sequence 38,39. This result demonstrates that 3'-extended gRNAs can serve as efficient nuclease targeting guides and can template reverse transcription, Para. [1160]).

The inventions listed in Groups I+ therefore lack unity under Rule 13 because they do not share a same or corresponding special technical features.