(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 December 2019 (12.12.2019)

WIPO | PCT

(10) International Publication Number
WO 2019/236982 A1

(51) International Patent Classification:
C12N 9/22 (2006.01) *C12N 15/113* (2010.01)

(21) International Application Number:
PCT/US2019/036021

(22) International Filing Date:
07 June 2019 (07.06.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/682,271 08 June 2018 (08.06.2018) US

(71) Applicant: **LOCANA, INC.** [US/US]; 7770 Regents Road, San Diego, CA 92122 (US).

(72) Inventors: **NELLES, David A.**; 7770 Regents Road, San Diego, California 92122 (US). **BATRA, Ranjan**; 7770 Regents Road, San Diego, California 92122 (US). **YEO, Gene**; 7770 Regents Road, San Diego, California 92122 (US).

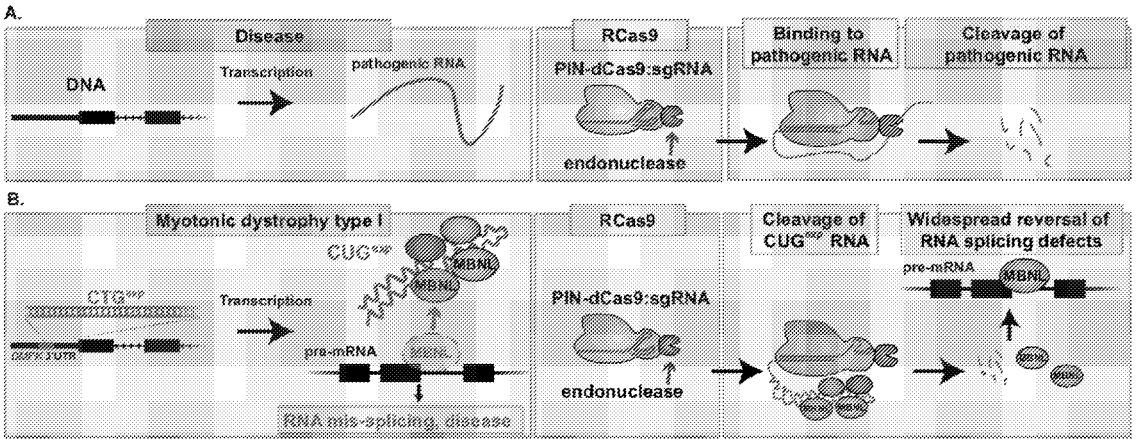
(74) Agent: **BROWN, Fraser** et al.; COOLEY LLP, 1299 Pennsylvania Avenue, NW Suite 700, Washington, District of Columbia 20004-2400 (US).

(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, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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,

(54) Title: RNA-TARGETING FUSION PROTEIN COMPOSITIONS AND METHODS FOR USE

FIGURE 1



(57) Abstract: Disclosed are compositions comprising: (a) a sequence comprising a guide RNA (gRNA) that specifically binds a target sequence within an RNA molecule and (b) a sequence encoding a fusion protein, the sequence comprising a sequence encoding a first RNA-binding polypeptide and a sequence encoding a second RNA-binding polypeptide, wherein neither the first RNA-binding polypeptide nor the second RNA-binding polypeptide comprises a significant DNA-nuclease activity, wherein the first RNA-binding polypeptide and the second RNA-binding polypeptide are not identical, and wherein the second RNA-binding polypeptide comprises an RNA-nuclease activity. Methods of making and methods of using compositions of the disclosure are also provided. For example, compositions of the disclosure may be used in the treatment of a disease or disorder in a subject. Exemplary disease or disorders of the disclosure include genetic and epigenetic diseases or disorders.

WO 2019/236982 A1

[Continued on next page]

TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

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))*

RNA-TARGETING FUSION PROTEIN COMPOSITIONS AND METHODS FOR USE

FIELD OF THE DISCLOSURE

[01] The disclosure is directed to molecular biology, and more, specifically, to compositions and methods for modifying expression and activity of RNA molecules.

RELATED APPLICATIONS

[02] This application claims priority to U.S. Patent Application No. 62/682,271, filed June 8, 2018, the contents of which are herein incorporated by reference in their entirety. The contents of U.S. Patent Application No. 62/682,276, filed June 8, 2018, are herein incorporated by reference in their entirety.

INCORPORATION OF SEQUENCE LISTING

[03] The contents of the text file named "LOCN_002_001WO_SeqList_ST25", which was created on June 6, 2019 and is 773 KB in size, are hereby incorporated by reference in their entirety.

BACKGROUND

[04] There has been a long-felt but unmet need in the art for a method of specifically binding target RNA molecules for modification of expression or activity of the RNA molecule or a protein encoded by the RNA molecule. The disclosure provides compositions and methods for specifically targeting RNA molecules in sequence-specific manner that further precludes modification of DNA sequences.

SUMMARY

[05] The disclosure provides a composition comprising (a) a sequence comprising a guide RNA (gRNA) that specifically binds a target sequence within an RNA molecule and (b) a sequence encoding a fusion protein, the sequence comprising a sequence encoding a first RNA-binding polypeptide and a sequence encoding a second RNA-binding polypeptide, wherein

neither the first RNA-binding polypeptide nor the second RNA-binding polypeptide comprises a significant DNA-nuclease activity, wherein the first RNA-binding polypeptide and the second RNA-binding polypeptide are not identical, and wherein the second RNA-binding polypeptide comprises an RNA-nuclease activity wherein the first RNA-binding polypeptide and the second RNA-binding polypeptide are not identical, and wherein the second RNA-binding polypeptide comprises an RNA-nuclease activity.

[06] The disclosure also provides a composition comprising a sequence encoding an RNA-guided target RNA-binding fusion protein comprising (a) a sequence encoding a first RNA-binding polypeptide or portion thereof; and (b) a sequence encoding a second RNA-binding polypeptide, wherein the first RNA-binding polypeptide binds a target RNA guided by a gRNA sequence, and wherein the second RNA-binding polypeptide comprises RNA-nuclease activity.

[07] The disclosure additionally provides a composition comprising a sequence encoding a target RNA-binding fusion protein comprising (a) a sequence encoding a first RNA-binding polypeptide or portion thereof; and (b) a sequence encoding a second RNA-binding polypeptide, wherein the first RNA-binding polypeptide binds a target RNA without a gRNA sequence, and wherein the second RNA-binding polypeptide comprises RNA-nuclease activity.

[08] In some embodiments of the compositions of the disclosure, the target sequence comprises at least one repeated sequence.

[09] In some embodiments of the compositions of the disclosure, the sequence comprising the gRNA further comprises a sequence encoding a promoter capable of expressing the gRNA in a eukaryotic cell.

[010] In some embodiments of the compositions of the disclosure, the eukaryotic cell is an animal cell. In some embodiments, the animal cell is a mammalian cell. In some embodiments, the animal cell is a human cell.

[011] In some embodiments of the compositions of the disclosure, the promoter is a constitutively active promoter. In some embodiments, the promoter sequence is isolated or derived from a promoter capable of driving expression of an RNA polymerase. In some embodiments, the promoter sequence is isolated or derived from a U6 promoter. In some embodiments, the promoter is a sequence isolated or derived from a promoter capable of driving expression of a transfer RNA (tRNA). In some embodiments, the promoter is isolated or derived

from an alanine tRNA promoter, an arginine tRNA promoter, an asparagine tRNA promoter, an aspartic acid tRNA promoter, a cysteine tRNA promoter, a glutamine tRNA promoter, a glutamic acid tRNA promoter, a glycine tRNA promoter, a histidine tRNA promoter, an isoleucine tRNA promoter, a leucine tRNA promoter, a lysine tRNA promoter, a methionine tRNA promoter, a phenylalanine tRNA promoter, a proline tRNA promoter, a serine tRNA promoter, a threonine tRNA promoter, a tryptophan tRNA promoter, a tyrosine tRNA promoter, or a valine tRNA promoter. In some embodiments, the promoter is isolated or derived from a valine tRNA promoter.

[012] In some embodiments of the compositions of the disclosure, the sequence comprising the gRNA further comprises a spacer sequence that specifically binds to the target RNA sequence. In some embodiments, the spacer sequence has at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 87%, 90%, 95%, 97%, 99% or any percentage in between of complementarity to the target RNA sequence. In some embodiments, the spacer sequence has 100% complementarity to the target RNA sequence. In some embodiments, the spacer sequence comprises or consists of 20 nucleotides. In some embodiments, the spacer sequence comprises or consists of 21 nucleotides. In some embodiments, the spacer sequence comprises or consists of the sequence UGGAGCGAGCAUCCCCCAAA (SEQ ID NO: 1), GUUUGGGGGAUGCUCGCUCCA (SEQ ID NO: 2), CCCUCACUGCUGGGGAGUCC (SEQ ID NO: 3), GGACUCCCCAGCAGUGAGGG (SEQ ID NO: 4), GCAACUGGAUCAAUUUGCUG (SEQ ID NO: 5), GCAGCAAAUUGAUCCAGUUGC (SEQ ID NO: 6), GCAUUCUUAUCUGGUCAGUGC (SEQ ID NO: 7), GCACUGACCAGAUAAGAAUG (SEQ ID NO: 8), GAGCAGCAGCAGCAGCAGCAG (SEQ ID NO: 9), GCAGGCAGGCAGGCAGGCAGG (SEQ ID NO: 10), GCCCCGGCCCCGGCCCCGGC (SEQ ID NO: 11) , or GCTGCTGCTGCTGCTGCTGC (SEQ ID NO: 12), GGGGCCGGGGCCGGGGCCGG (SEQ ID NO: 74), GGGCCGGGGCCGGGGCCGGG (SEQ ID NO: 75), GGCCGGGGCCGGGGCCGGGG (SEQ ID NO: 76), GCCGGGGCCGGGGCCGGGGC (SEQ ID NO: 77), CCGGGGCCGGGGCCGGGGCC (SEQ ID NO: 78), or CGGGGCCGGGGCCGGGGCCG (SEQ ID NO: 79).

[013] In some embodiments of the compositions of the disclosure, the sequence comprising the gRNA further comprises a spacer sequence that specifically binds to the target RNA

sequence. In some embodiments, the spacer sequence has at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 87%, 90%, 95%, 97%, 99% or any percentage in between of complementarity to the target RNA sequence.

[014] In some embodiments, the spacer sequence has 100% complementarity to the target RNA sequence. In some embodiments, the spacer sequence comprises or consists of 20 nucleotides. In some embodiments, the spacer sequence comprises or consists of 21 nucleotides. In some embodiments, the spacer sequence comprises or consists of the sequence GUGAUAAGUGGAAUGCCAUG (SEQ ID NO: 14), CUGGUGAACUUCCGAUAGUG (SEQ ID NO: 15), or GAGATATAGCCTGGTGGTTC (SEQ ID NO: 16).

[015] In some embodiments of the compositions of the disclosure, the sequence comprising the gRNA further comprises a spacer sequence that specifically binds to the target RNA sequence. In some embodiments, the spacer sequence has at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 87%, 90%, 95%, 97%, 99% or any percentage in between of complementarity to the target RNA sequence. In some embodiments, the spacer sequence has 100% complementarity to the target RNA sequence. In some embodiments, the spacer sequence comprises or consists of 20 nucleotides. In some embodiments, the spacer sequence comprises or consists of 21 nucleotides. In some embodiments, the spacer sequence comprises or consists of a sequence comprising at least 1, 2, 3, 4, 5, 6, or 7 repeats of the sequence CUG (SEQ ID NO: 18), CCUG (SEQ ID NO: 19), CAG (SEQ ID NO: 80), GGGGCC (SEQ ID NO: 81) or any combination thereof.

[016] In some embodiments of the compositions of the disclosure, the sequence comprising the gRNA further comprises a scaffold sequence that specifically binds to the first RNA binding protein. In some embodiments, the scaffold sequence comprises a stem-loop structure. In some embodiments, the scaffold sequence comprises or consists of 90 nucleotides. In some embodiments, the scaffold sequence comprises or consists of 93 nucleotides. In some embodiments, the scaffold sequence comprises or consists of the sequence GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 13). In some embodiments, the scaffold sequence comprises or consists of the sequence GGACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUU (SEQ ID NO: 17). In some embodiments, the scaffold

sequence comprises or consists of the sequence GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUU AUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 82) or GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGA AAAAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 83).

[017] In some embodiments of the compositions of the disclosure, the gRNA does not bind or does not selectively bind to a second sequence within the RNA molecule.

[018] In some embodiments of the compositions of the disclosure, an RNA genome or an RNA transcriptome comprises the RNA molecule.

[019] In some embodiments of the compositions of the disclosure, the first RNA binding protein comprises a CRISPR-Cas protein. In some embodiments, the CRISPR-Cas protein is a Type II CRISPR-Cas protein. In some embodiments, the first RNA binding protein comprises a Cas9 polypeptide or an RNA-binding portion thereof. In some embodiments, the CRISPR-Cas protein comprises a native RNA nuclease activity. In some embodiments, the native RNA nuclease activity is reduced or inhibited. In some embodiments, the native RNA nuclease activity is increased or induced. In some embodiments, the CRISPR-Cas protein comprises a native DNA nuclease activity and the native DNA nuclease activity is inhibited. In some embodiments, the CRISPR-Cas protein comprises a mutation. In some embodiments, a nuclease domain of the CRISPR-Cas protein comprises the mutation. In some embodiments, the mutation occurs in a nucleic acid encoding the CRISPR-Cas protein. In some embodiments, the mutation occurs in an amino acid encoding the CRISPR-Cas protein. In some embodiments, the mutation comprises a substitution, an insertion, a deletion, a frameshift, an inversion, or a transposition. In some embodiments, the mutation comprises a deletion of a nuclease domain, a binding site within the nuclease domain, an active site within the nuclease domain, or at least one essential amino acid residue within the nuclease domain.

[020] In some embodiments of the compositions of the disclosure, the first RNA binding protein comprises a CRISPR-Cas protein. In some embodiments, the CRISPR-Cas protein is a Type V CRISPR-Cas protein. In some embodiments, the first RNA binding protein comprises a Cpf1 polypeptide or an RNA-binding portion thereof. In some embodiments, the CRISPR-Cas protein comprises a native RNA nuclease activity. In some embodiments, the native RNA

nuclease activity is reduced or inhibited. In some embodiments, the native RNA nuclease activity is increased or induced. In some embodiments, the CRISPR-Cas protein comprises a native DNA nuclease activity and the native DNA nuclease activity is inhibited. In some embodiments, the CRISPR-Cas protein comprises a mutation. In some embodiments, a nuclease domain of the CRISPR-Cas protein comprises the mutation. In some embodiments, the mutation occurs in a nucleic acid encoding the CRISPR-Cas protein. In some embodiments, the mutation occurs in an amino acid encoding the CRISPR-Cas protein. In some embodiments, the mutation comprises a substitution, an insertion, a deletion, a frameshift, an inversion, or a transposition. In some embodiments, the mutation comprises a deletion of a nuclease domain, a binding site within the nuclease domain, an active site within the nuclease domain, or at least one essential amino acid residue within the nuclease domain.

[021] In some embodiments of the compositions of the disclosure, the first RNA binding protein comprises a CRISPR-Cas protein. In some embodiments, the CRISPR-Cas protein is a Type VI CRISPR-Cas protein. In some embodiments, the first RNA binding protein comprises a Cas13 polypeptide or an RNA-binding portion thereof. In some embodiments, the first RNA binding protein comprises a CasRx/Cas13d polypeptide or an RNA-binding portion thereof. In some embodiments, the CRISPR-Cas protein comprises a native RNA nuclease activity. In some embodiments, the native RNA nuclease activity is reduced or inhibited. In some embodiments, the native RNA nuclease activity is increased or induced. In some embodiments, the CRISPR-Cas protein comprises a native DNA nuclease activity and the native DNA nuclease activity is inhibited. In some embodiments, the CRISPR-Cas protein comprises a mutation. In some embodiments, a nuclease domain of the CRISPR-Cas protein comprises the mutation. In some embodiments, the mutation occurs in a nucleic acid encoding the CRISPR-Cas protein. In some embodiments, the mutation occurs in an amino acid encoding the CRISPR-Cas protein. In some embodiments, the mutation comprises a substitution, an insertion, a deletion, a frameshift, an inversion, or a transposition. In some embodiments, the mutation comprises a deletion of a nuclease domain, a binding site within the nuclease domain, an active site within the nuclease domain, or at least one essential amino acid residue within the nuclease domain.

[022] In some embodiments of the compositions of the disclosure, the first RNA binding protein comprises a Pumilio and FBF (PUF) protein or an RNA binding portion thereof. In some

embodiments, the first RNA binding protein comprises a Pumilio-based assembly (PUMBY) protein or an RNA binding portion thereof.

[023] In some embodiments of the compositions of the disclosure, the first RNA binding protein does not require multimerization for RNA-binding activity. In some embodiments, the first RNA binding protein is not a monomer of a multimer complex. In some embodiments, a multimer protein complex does not comprise the first RNA binding protein.

[024] In some embodiments of the compositions of the disclosure, the first RNA binding protein selectively binds to a target sequence within the RNA molecule. In some embodiments, the first RNA binding protein does not comprise an affinity for a second sequence within the RNA molecule. In some embodiments, the first RNA binding protein does not comprise a high affinity for or selectively bind a second sequence within the RNA molecule.

[025] In some embodiments of the compositions of the disclosure, an RNA genome or an RNA transcriptome comprises the RNA molecule.

[026] In some embodiments of the compositions of the disclosure, the first RNA binding protein comprises between 2 and 1300 amino acids, inclusive of the endpoints.

[027] In some embodiments of the compositions of the disclosure, the sequence encoding the first RNA binding protein further comprises a sequence encoding a nuclear localization signal (NLS), a nuclear export signal (NES) or tag. In some embodiments, the sequence encoding a nuclear localization signal (NLS) is positioned 3' to the sequence encoding the first RNA binding protein. In some embodiments, the first RNA binding protein comprises an NLS at a C-terminus of the protein.

[028] In some embodiments of the compositions of the disclosure, the sequence encoding the first RNA binding protein further comprises a first sequence encoding a first NLS and a second sequence encoding a second NLS. In some embodiments, the sequence encoding the first NLS or the second NLS is positioned 3' to the sequence encoding the first RNA binding protein. In some embodiments, the first RNA binding protein comprises the first NLS or the second NLS at a C-terminus of the protein.

[029] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a nuclease domain. In some embodiments, the second RNA

binding protein binds RNA in a manner in which it associates with RNA. In some embodiments, the second RNA binding protein associates with RNA in a manner in which it cleaves RNA.

[030] In some embodiments of the compositions of the disclosure, the sequence encoding the second RNA binding protein comprises or consists of an RNAse. In some embodiments, the second RNA binding protein comprises or consists of an RNAse1. In some embodiments, the RNAse1 comprises or consists of SEQ ID NO: 20. In some embodiments, the second RNA binding protein comprises or consists of an RNAse4. In some embodiments, the RNAse4 comprises or consists of SEQ ID NO: 21. In some embodiments, the second RNA binding protein comprises or consists of an RNAse6. In some embodiments, the RNAse6 comprises or consists of SEQ ID NO: 22. In some embodiments, the second RNA binding protein comprises or consists of an RNAse7. In some embodiments, the RNAse7 comprises or consists of SEQ ID NO: 23. In some embodiments, the second RNA binding protein comprises or consists of an RNAse8. In some embodiments, the RNAse8 protein comprises or consists of SEQ ID NO: 24. In some embodiments, the second RNA binding protein comprises or consists of an RNAse2. In some embodiments, the RNAse2 protein comprises or consists of SEQ ID NO: 25. In some embodiments, the second RNA binding protein comprises or consists of an RNAse6PL. In some embodiments, the RNAse6PL protein comprises or consists of SEQ ID NO: 26. In some embodiments, the second RNA binding protein comprises or consists of an RNAseL. In some embodiments the RNAseL protein comprises or consists of SEQ ID NO: 27. In some embodiments, the second RNA binding protein comprises or consists of an RNAseT2. In some embodiments, the RNAseT2 protein comprises or consists of SEQ ID NO: 28. In some embodiments, the second RNA binding protein comprises or consists of an RNAse11. In some embodiments, the RNAse11 protein comprises or consists of SEQ ID NO: 29. In some embodiments, the second RNA binding protein comprises or consists of an RNAseT2-like. In some embodiments, the RNAseT2-like protein comprises or consists of SEQ ID NO: 30.

[031] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a mutated RNAse. In some embodiments, the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(K41R)) polypeptide. In some embodiments, the Rnase1 (K41R) polypeptide comprises or consists of SEQ ID NO: 116. In some embodiments, the second RNA binding protein comprises or consists of a mutated

Rnase1 (Rnase1(K41R, D121E)) polypeptide. In some embodiments, the Rnase1 (Rnase1(K41R, D121E)) polypeptide comprises or consists of SEQ ID NO: 66. In some embodiments, the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(K41R, D121E, H119N)) polypeptide. In some embodiments, the Rnase1 (Rnase1(K41R, D121E, H119N)) polypeptide comprises or consists of SEQ ID NO: 118. In some embodiments, the second RNA binding protein comprises or consists of a mutated Rnase1. In some embodiments, the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(H119N)) polypeptide. In some embodiments, the Rnase1 (Rnase1(H119N)) polypeptide comprises or consists SEQ ID NO: 119. In some embodiments, the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(R39D, N67D, N88A, G89D, R91D, H119N)) polypeptide. In some embodiments, the Rnase1 (Rnase1(R39D, N67D, N88A, G89D, R91D, H119N)) polypeptide comprises or consists of SEQ ID NO: 120. In some embodiments, the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(R39D, N67D, N88A, G89D, R91D, H119N)) polypeptide. In some embodiments, the Rnase1 (Rnase1(R39D, N67D, N88A, G89D, R91D, H119N, K41R, D121E)) polypeptide comprises or consists of SEQ ID NO: 121. In some embodiments, the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(R39D, N67D, N88A, G89D, R91D, H119N)) polypeptide. In some embodiments, the Rnase1 (Rnase1(R39D, N67D, N88A, G89D, R91D)) polypeptide comprises or consists of SEQ ID NO: 122.

[032] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a NOB1 polypeptide. In some embodiments, the NOB1 polypeptide comprises or consists of SEQ ID NO: 31.

[033] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of an endonuclease. In some embodiments, the second RNA binding protein comprises or consists of an endonuclease V (ENDOV). In some embodiments, the ENDOV protein comprises or consists of SEQ ID NO: 32. In some embodiments, the second RNA binding protein comprises or consists of an endonuclease G (ENDOG). In some embodiments, the ENDOG protein comprises or consists of SEQ ID NO: 33. In some embodiments, the second RNA binding protein comprises or consists of an endonuclease D1 (ENDOD1). In some embodiments, the ENDOD1 protein comprises or consists of SEQ ID NO:

34. In some embodiments, the second RNA binding protein comprises or consists of a Human flap endonuclease-1 (hFEN1). In some embodiments, the hFEN1 protein comprises or consists of SEQ ID NO: 35. In some embodiments, the second RNA binding protein comprises or consists of a DNA repair endonuclease XPF (ERCC4) polypeptide. In some embodiments, the ERCC4 protein comprises or consists of SEQ ID NO: 64.

[034] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of an Endonuclease III-like protein 1 (NTHL) polypeptide. In some embodiments, the NTHL polypeptide comprises or consists of SEQ ID NO: 123.

[035] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a human Schlafen 14 (hSLFN14) polypeptide. In some embodiments, the hSLFN14 polypeptide comprises or consists of SEQ ID NO: 36.

[036] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a human beta-lactamase-like protein 2 (hLACTB2) polypeptide. In some embodiments, the hLACTB2 polypeptide comprises or consists of SEQ ID NO: 37.

[037] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of an apurinic/apyrimidinic (AP) endodeoxyribonuclease (APEX) polypeptide. In some embodiments, the second RNA binding protein comprises or consists of an apurinic/apyrimidinic (AP) endodeoxyribonuclease (APEX2) polypeptide. In some embodiments, the APEX2 polypeptide comprises or consists of SEQ ID NO: 38. In some embodiments, the APEX2 polypeptide comprises or consists of SEQ ID NO: 39. In some embodiments, the second RNA binding protein comprises or consists of an apurinic or apyrimidinic site lyase (APEX1) polypeptide. In some embodiments, the APEX1 polypeptide comprises or consists of SEQ ID NO: 125.

[038] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of an angiogenin (ANG) polypeptide. In some embodiments, the ANG polypeptide comprises or consists SEQ ID NO: 40.

[039] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a heat responsive protein 12 (HRSP12) polypeptide. In some embodiments, the HRSP12 polypeptide comprises or consists of SEQ ID NO: 41.

[040] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a Zinc Finger CCCH-Type Containing 12A (ZC3H12A) polypeptide. In some embodiments, the ZC3H12A polypeptide comprises or consists of SEQ ID NO: 42. In some embodiments, the ZC3H12A polypeptide comprises or consists of SEQ ID NO: 43.

[041] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a Reactive Intermediate Imine Deaminase A (RIDA) polypeptide. In some embodiments, the RIDA polypeptide comprises or consists of SEQ ID NO: 44.

[042] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a Phospholipase D Family Member 6 (PDL6) polypeptide. In some embodiments, the PDL6 polypeptide comprises or consists of SEQ ID NO: 126.

[043] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a mitochondrial ribonuclease P catalytic subunit (KIAA0391) polypeptide. In some embodiments, the KIAA0391 polypeptide comprises or consists of SEQ ID NO: 127.

[044] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of an argonaute 2 (AGO2) polypeptide.
In some embodiments of the compositions of the disclosure, the AGO2 polypeptide comprises or consists of SEQ ID NO: 128.

[045] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a mitochondrial nuclease EXOG (EXOG) polypeptide. In some embodiments, the EXOG polypeptide comprises or consists of SEQ ID NO: 129.

[046] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a Zinc Finger CCCH-Type Containing 12D (ZC3H12D) polypeptide. In some embodiments, the ZC3H12D polypeptide comprises or consists of SEQ ID NO: 130.

[047] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of an endoplasmic reticulum to nucleus signaling 2 (ERN2)

polypeptide. In some embodiments, the ERN2 polypeptide comprises or consists of SEQ ID NO: 131.

[048] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a pelota mRNA surveillance and ribosome rescue factor (PELO) polypeptide. In some embodiments, the PELO polypeptide comprises or consists of SEQ ID NO: 132.

[049] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a YBEY metallopeptidase (YBEY) polypeptide. In some embodiments, the YBEY polypeptide comprises or consists of SEQ ID NO: 133.

[050] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a cleavage and polyadenylation specific factor 4 like (CPSF4L) polypeptide. In some embodiments, the CPSF4L polypeptide comprises or consists of SEQ ID NO: 134.

[051] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of an hCG_2002731polypeptide. In some embodiments, the hCG_2002731 comprises or consists of SEQ ID NO: 135. In some embodiments, the hCG_2002731 polypeptide comprises or consists of SEQ ID NO: 136.

[052] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of an Excision Repair Cross-Complementation Group 1 (ERCC1) polypeptide. In some embodiments, the ERCC1 polypeptide comprises or consists of SEQ ID NO: 137.

[053] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a ras-related C3 botulinum toxin substrate 1 isoform (RAC1) polypeptide. In some embodiments, the RAC1 polypeptide comprises or consists of SEQ ID NO: 138.

[054] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a Ribonuclease A A1 (RAA1) polypeptide. In some embodiments, the RAA1 polypeptide comprises or consists of SEQ ID NO: 139.

[055] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a Ras Related Protein (RAB1) polypeptide. In some embodiments, the RAB1 polypeptide comprises or consists of SEQ ID NO: 140.

[056] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a DNA Replication Helicase/Nuclease 2 (DNA2) polypeptide. In some embodiments, the DNA2 polypeptide comprises or consists of SEQ ID NO: 141.

[057] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a FLJ35220 polypeptide. In some embodiments, the FLJ35220 polypeptide comprises or consists of SEQ ID NO: 142.

[058] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a FLJ13173 polypeptide. In some embodiments, the FLJ13173 polypeptide comprises or consists of SEQ ID NO: 143.

[059] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of Teneurin Transmembrane Protein (TENM) polypeptide. In some embodiments, the second RNA binding protein comprises or consists of Teneurin Transmembrane Protein 1 (TENM1) polypeptide. In some embodiments, the TENM1 polypeptide comprises or consists of SEQ ID NO: 144. In some embodiments, the second RNA binding protein comprises or consists of Teneurin Transmembrane Protein 2 (TENM2) polypeptide. In some embodiments, the TENM2 polypeptide comprises or consists of SEQ ID NO: 145.

[060] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a Ribonuclease Kappa (RNAseK) polypeptide. In some embodiments, the RNAseK polypeptide comprises or consists of SEQ ID NO: 204.

[061] In some embodiments, the fusion proteins of the disclosure are used in methods for treating a subject in need thereof, the methods comprising contacting a target RNA with a fusion protein or the sequence encoding the fusion protein.

BRIEF DESCRIPTION OF THE DRAWINGS

[062] The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

[063] Figure 1A-B is a schematic diagram of an exemplary embodiment of a composition of the disclosure. (A) An RNA-targeting Cas9 system fused to an endonuclease targets and cleaves a disease-causing RNA. (B) Depicts an application of (A) in the context of myotonic dystrophy type 1, wherein an RNA-targeting Cas9 system fused to an endonuclease targets and cleaves a repetitive RNA composed of repeating CUG units. In the absence of the RNA-targeting Cas9 system, the repetitive RNA composed of repeating CUG units binds to a splicing factor MBNL and causes pathology via dysfunctional RNA splicing. Cleavage of this repetitive RNA ameliorates disease.

[064] Figure 2 is a schematic diagram depicting an exemplary modular therapeutic platform for treating genetic disease by targeting RNA molecules.

[065] Figure 3A-B is a pair of schematic diagrams depicting (A) a "high expression" control system (also referred to as "pos control") comprising a two plasmid system comprising a cytomegalovirus promoter driving expression of the RNA endonuclease/Cas9 fusion and (B) a "low expression" control system (also referred to as "P13") comprising a single plasmid system comprising a lower-expression promoter (pEFS) driving expression of the RNA endonuclease/Cas9 fusion.

[066] Figure 4A is a pair of schematic diagrams depicting an exemplary RNA Endonuclease-C. jejuni Cas9 fusion protein (left) and a vector comprising an exemplary RNA Endonuclease-S. pyogenes Cas9 fusion protein (right)

[067] Figure 4B is a graph depicting the ability of a variety of fusion proteins comprising either C. jejuni Cas9 or S. pyogenes Cas9, as shown in Figure 4A, to cleave repetitive RNA molecules.

[068] Figure 5A is a pair of schematic diagrams depicting an exemplary RNA Endonuclease-C. jejuni Cas9 fusion protein (left) and a vector comprising an exemplary RNA Endonuclease-S. pyogenes Cas9 fusion protein (right)

[069] Figure 5B is a graph depicting the ability of a variety of fusion proteins comprising either C. jejuni Cas9 or S. pyogenes Cas9, as shown in Figure 5A, to cleave mRNA molecules encoding a luciferase protein.

[070] Figure 6 is a table providing a key to the endonucleases shown in Figures 4B, 5B, and 9.

[071] Figure 7A is a schematic diagram depicting an exemplary RNA Endonuclease-C. jejuni Cas9 fusion protein.

[072] Figure 7B is a graph depicting changes in expression levels of Zika NS5 in the presence of both E43 and E67 CjeCas9-endonuclease fusions with sgRNAs containing the various NS5-targeting spacer sequences as indicated in Table 2. Zika NS5 expression is displayed as fold change relative to the endonuclease loaded with an sgRNA containing a control (Lambda) spacer sequence.

[073] Figure 8A is a fluorescence microscopy image of cells transfected with CjeCas9-endonuclease fusions loaded with an sgRNA containing a Zika NS5-targeting spacer sequence.

[074] Figure 8B is a graph depicting changes of expression of Zika NS5 in the presence of CjeCas9-endonuclease fusions loaded with the appropriate Zika NS5-targeting sgRNA as compared to a CjeCas9-endonuclease fusions loaded with a non-Zika NS5 targeting sgRNA.

[075] Figure 9 is a graph depicting the cleavage efficiencies of a variety of exemplary fusion proteins (SpyCas9 fused to the annotated endonuclease).

DETAILED DESCRIPTION

[076] The disclosure provides an RNA-guided fusion protein that selectively binds and, optionally, cleaves RNA molecules. The disclosure provides vectors, compositions and cells comprising the RNA-guided fusion protein. The disclosure provides methods of using the RNA-guided fusion protein, vectors, compositions and cells of the disclosure to treat a disease or disorder.

Guide RNA

[077] The terms guide RNA (gRNA) and single guide RNA (sgRNA) are used interchangeably throughout the disclosure.

[078] Guide RNAs (gRNAs) of the disclosure may comprise of a spacer sequence and a scaffolding sequence. In some embodiments, a guide RNA is a single guide RNA (sgRNA)

comprising a contiguous spacer sequence and scaffolding sequence. In some embodiments, the spacer sequence and the scaffolding sequence are not contiguous. In some embodiments, a scaffold sequence comprises a "direct repeat" (DR) sequence. DR sequences refer to the repetitive sequences in the CRISPR locus (naturally-occurring in a bacterial genome or plasmid) that are interspersed with the spacer sequences. It is well known that one would be able to infer the DR sequence of a corresponding Cas protein if the sequence of the associated CRISPR locus is known. In some embodiments, a sequence encoding a guide RNA or single guide RNA of the disclosure comprises or consists of a spacer sequence and a scaffolding sequence, that are separated by a linker sequence. In some embodiments, the linker sequence may comprise or consist of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50 or any number of nucleotides in between. In some embodiments, the linker sequence may comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50 or any number of nucleotides in between.

[079] Guide RNAs (gRNAs) of the disclosure may comprise non-naturally occurring nucleotides. In some embodiments, a guide RNA of the disclosure or a sequence encoding the guide RNA comprises or consists of modified or synthetic RNA nucleotides. Exemplary modified RNA nucleotides include, but are not limited to, pseudouridine (Ψ), dihydrouridine (D), inosine (I), and 7-methylguanosine (m7G), hypoxanthine, xanthine, xanthosine, 7-methylguanine, 5, 6-Dihydrouracil, 5-methylcytosine, 5-methylcytidine, 5-hydropxymethylcytosine, isoguanine, and isocytosine.

[080] Guide RNAs (gRNAs) of the disclosure may bind modified RNA within a target sequence. Within a target sequence, guide RNAs (gRNAs) of the disclosure may bind modified RNA. Exemplary epigenetically or post-transcriptionally modified RNA include, but are not limited to, 2'-O-Methylation (2'-OMe) (2'-O-methylation occurs on the oxygen of the free 2'-OH of the ribose moiety), N6-methyladenosine (m6A), and 5-methylcytosine (m5C).

[081] In some embodiments of the compositions of the disclosure, a guide RNA of the disclosure comprises at least one sequence encoding a non-coding C/D box small nucleolar RNA (snoRNA) sequence. In some embodiments, the snoRNA sequence comprises at least one sequence that is complementary to the target RNA, wherein the target sequence of the RNA molecule comprises at least one 2'-OMe. In some embodiments, the snoRNA sequence comprises at least one sequence that is complementary to the target RNA, wherein the at least

one sequence that is complementary to the target RNA comprises a box C motif (RUGAUGA) and a box D motif (CUGA).

[082] Spacer sequences of the disclosure bind to the target sequence of an RNA molecule. Spacer sequences of the disclosure may comprise a CRISPR RNA (crRNA). Spacer sequences of the disclosure comprise or consist of a sequence having sufficient complementarity to a target sequence of an RNA molecule to bind selectively to the target sequence. Upon binding to a target sequence of an RNA molecule, the spacer sequence may guide one or more of a scaffolding sequence and a fusion protein to the RNA molecule. In some embodiments, a sequence having sufficient complementarity to a target sequence of an RNA molecule to bind selectively to the target sequence has at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96, 97%, 98%, 99%, or any percentage identity in between to the target sequence. In some embodiments, a sequence having sufficient complementarity to a target sequence of an RNA molecule to bind selectively to the target sequence has 100% identity the target sequence.

[083] Scaffolding sequences of the disclosure bind the first RNA-binding polypeptide of the disclosure. Scaffolding sequences of the disclosure may comprise a trans acting RNA (tracrRNA). Scaffolding sequences of the disclosure comprise or consist of a sequence having sufficient complementarity to a target sequence of an RNA molecule to bind selectively to the target sequence. Upon binding to a target sequence of an RNA molecule, the scaffolding sequence may guide a fusion protein to the RNA molecule. In some embodiments, a sequence having sufficient complementarity to a target sequence of an RNA molecule to bind selectively to the target sequence has at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96, 97%, 98%, 99%, or any percentage identity in between to the target sequence. In some embodiments, a sequence having sufficient complementarity to a target sequence of an RNA molecule to bind selectively to the target sequence has 100% identity the target sequence. Alternatively, or in addition, in some embodiments, scaffolding sequences of the disclosure comprise or consist of a sequence that binds to a first RNA binding protein or a second RNA binding protein of a fusion protein of the disclosure. In some embodiments, scaffolding sequences of the disclosure comprise a secondary structure or a tertiary structure. Exemplary secondary structures include, but are not limited to, a helix, a stem loop, a bulge, a tetraloop and a pseudoknot. Exemplary tertiary structures include, but are not limited to, an A-form of a helix,

a B-form of a helix, and a Z-form of a helix. Exemplary tertiary structures include, but are not limited to, a twisted or helicized stem loop. Exemplary tertiary structures include, but are not limited to, a twisted or helicized pseudoknot. In some embodiments, scaffolding sequences of the disclosure comprise at least one secondary structure or at least one tertiary structure. In some embodiments, scaffolding sequences of the disclosure comprise one or more secondary structure(s) or one or more tertiary structure(s).

[084] In some embodiments of the compositions of the disclosure, a guide RNA or a portion thereof selectively binds to a tetraloop motif in an RNA molecule of the disclosure. In some embodiments, a target sequence of an RNA molecule comprises a tetraloop motif. In some embodiments, the tetraloop motif is a "GRNA" motif comprising or consisting of one or more of the sequences of GAAA, GUGA, GCAA or GAGA.

[085] In some embodiments of the compositions of the disclosure, a guide RNA or a portion thereof that binds to a target sequence of an RNA molecule hybridizes to the target sequence of the RNA molecule. In some embodiments, a guide RNA or a portion thereof that binds to a first RNA binding protein or to a second RNA binding protein covalently binds to the first RNA binding protein or to the second RNA binding protein. In some embodiments, a guide RNA or a portion thereof that binds to a first RNA binding protein or to a second RNA binding protein non-covalently binds to the first RNA binding protein or to the second RNA binding protein.

[086] In some embodiments of the compositions of the disclosure, a guide RNA or a portion thereof comprises or consists of between 10 and 100 nucleotides, inclusive of the endpoints. In some embodiments, a spacer sequence of the disclosure comprises or consists of between 10 and 30 nucleotides, inclusive of the endpoints. In some embodiments, a spacer sequence of the disclosure comprises or consists of 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 or 30 nucleotides. In some embodiments, the spacer sequence of the disclosure comprises or consists of 20 nucleotides. In some embodiments, the spacer sequence of the disclosure comprises or consists of 21 nucleotides. In some embodiments, a scaffold sequence of the disclosure comprises or consists of between 10 and 100 nucleotides, inclusive of the endpoints. In some embodiments, a scaffold sequence of the disclosure comprises or consists of 30, 35, 40, 45, 50, 55, 60, 65, 70, 76, 80, 87, 90, 95, 100 or any number of nucleotides in between. In some embodiments, the scaffold sequence of the disclosure comprises or consists of between 85 and

95 nucleotides, inclusive of the endpoints. In some embodiments, the scaffold sequence of the disclosure comprises or consists of 85 nucleotides. In some embodiments, the scaffold sequence of the disclosure comprises or consists of 90 nucleotides. In some embodiments, the scaffold sequence of the disclosure comprises or consists of 93 nucleotides.

[087] In some embodiments of the compositions of the disclosure, a guide RNA or a portion thereof does not comprise a nuclear localization sequence (NLS).

[088] In some embodiments of the compositions of the disclosure, a guide RNA or a portion thereof does not comprise a sequence complementary to a protospacer adjacent motif (PAM).

[089] Therapeutic or pharmaceutical compositions of the disclosure do not comprise a PAMmer oligonucleotide. In other embodiments, optionally, non-therapeutic or non-pharmaceutical compositions may comprise a PAMmer oligonucleotide. The term "PAMmer" refers to an oligonucleotide comprising a PAM sequence that is capable of interacting with a guide nucleotide sequence-programmable RNA binding protein. Non-limiting examples of PAMmers are described in O'Connell *et al. Nature* 516, pages 263–266 (2014), incorporated herein by reference. A PAM sequence refers to a protospacer adjacent motif comprising about 2 to about 10 nucleotides. PAM sequences are specific to the guide nucleotide sequence-programmable RNA binding protein with which they interact and are known in the art. For example, *Streptococcus pyogenes* PAM has the sequence 5'-NGG-3', where "N" is any nucleobase followed by two guanine ("G") nucleobases. Cas9 of *Francisella novicida* recognizes the canonical PAM sequence 5'-NGG-3', but has been engineered to recognize the PAM 5'-YG-3' (where "Y" is a pyrimidine), thus adding to the range of possible Cas9 targets. The Cpf1 nuclease of *Francisella novicida* recognizes the PAM 5'-TTTN-3' or 5'-YTN-3'.

[090] In some embodiments of the compositions of the disclosure, a guide RNA or a portion thereof comprises a sequence complementary to a protospacer flanking sequence (PFS). In some embodiments, including those wherein a guide RNA or a portion thereof comprises a sequence complementary to a PFS, the first RNA binding protein may comprise a sequence isolated or derived from a Cas13 protein. In some embodiments, including those wherein a guide RNA or a portion thereof comprises a sequence complementary to a PFS, the first RNA binding protein may comprise a sequence encoding a Cas13 protein or an RNA-binding portion thereof. In some

embodiments, the guide RNA or a portion thereof does not comprise a sequence complementary to a PFS.

[091] In some embodiments of the compositions of the disclosure, guide RNA sequence of the disclosure comprises a promoter sequence to drive expression of the guide RNA. In some embodiments, a vector comprising a guide RNA sequence of the disclosure comprises a promoter sequence to drive expression of the guide RNA. In some embodiments, the promoter to drive expression of the guide RNA is a constitutive promoter. In some embodiments, the promoter sequence is an inducible promoter. In some embodiments, the promoter is a sequence is a tissue-specific and/or cell-type specific promoter. In some embodiments, the promoter is a hybrid or a recombinant promoter. In some embodiments, the promoter is a promoter capable of expressing the guide RNA in a mammalian cell. In some embodiments, the promoter is a promoter capable of expressing the guide RNA in a human cell. In some embodiments, the promoter is a promoter capable of expressing the guide RNA and restricting the guide RNA to the nucleus of the cell. In some embodiments, the promoter is a human RNA polymerase promoter or a sequence isolated or derived from a sequence encoding a human RNA polymerase promoter. In some embodiments, the promoter is a U6 promoter or a sequence isolated or derived from a sequence encoding a U6 promoter. In some embodiments, the promoter is a human tRNA promoter or a sequence isolated or derived from a sequence encoding a human tRNA promoter. In some embodiments, the promoter is a human valine tRNA promoter or a sequence isolated or derived from a sequence encoding a human valine tRNA promoter.

[092] In some embodiments of the compositions of the disclosure, a promoter to drive expression of the guide RNA further comprises a regulatory element. In some embodiments, a vector comprising a promoter sequence to drive expression of the guide RNA further comprises a regulatory element. In some embodiments, a regulatory element enhances expression of the guide RNA. Exemplary regulatory elements include, but are not limited to, an enhancer element, an intron, an exon, or a combination thereof.

[093] In some embodiments of the compositions of the disclosure, a vector of the disclosure comprises one or more of a sequence encoding a guide RNA, a promoter sequence to drive expression of the guide RNA and a sequence encoding a regulatory element. In some

embodiments of the compositions of the disclosure, the vector further comprises a sequence encoding a fusion protein of the disclosure.

Fusion Proteins

[094] Fusion proteins of the disclosure comprise a first RNA binding protein and a second RNA binding protein. In some embodiments, along a sequence encoding the fusion protein, the sequence encoding the first RNA binding protein is positioned 5' of the sequence encoding the second RNA binding protein. In some embodiments, along a sequence encoding the fusion protein, the sequence encoding the first RNA binding protein is positioned 3' of the sequence encoding the second RNA binding protein.

[095] In some embodiments of the compositions of the disclosure, the sequence encoding the first RNA binding protein comprises a sequence isolated or derived from a protein capable of binding an RNA molecule. In some embodiments, the sequence encoding the first RNA binding protein comprises a sequence isolated or derived from a protein capable of selectively binding an RNA molecule and not binding a DNA molecule, a mammalian DNA molecule or any DNA molecule. In some embodiments, the sequence encoding the first RNA binding protein comprises a sequence isolated or derived from a protein capable of binding an RNA molecule and inducing a break in the RNA molecule. In some embodiments, the sequence encoding the first RNA binding protein comprises a sequence isolated or derived from a protein capable of binding an RNA molecule, inducing a break in the RNA molecule, and not binding a DNA molecule, a mammalian DNA molecule or any DNA molecule. In some embodiments, the sequence encoding the first RNA binding protein comprises a sequence isolated or derived from a protein capable of binding an RNA molecule, inducing a break in the RNA molecule, and neither binding nor inducing a break in a DNA molecule, a mammalian DNA molecule or any DNA molecule.

[096] In some embodiments of the compositions of the disclosure, the sequence encoding the first RNA binding protein comprises a sequence isolated or derived from a protein with no DNA nuclease activity.

[097] In some embodiments of the compositions of the disclosure, the sequence encoding the first RNA binding protein comprises a sequence isolated or derived from a protein having DNA nuclease activity, wherein the DNA nuclease activity does not induce a break in a DNA

molecule, a mammalian DNA molecule or any DNA molecule when a composition of the disclosure is contacted to an RNA molecule or introduced into a cell or into a subject of the disclosure.

[098] In some embodiments of the compositions of the disclosure, the sequence encoding the first RNA binding protein comprises a sequence isolated or derived from a protein having DNA nuclease activity, wherein the DNA nuclease activity is inactivated and wherein the DNA nuclease activity does not induce a break in a DNA molecule, a mammalian DNA molecule or any DNA molecule when a composition of the disclosure is contacted to an RNA molecule or introduced into a cell or into a subject of the disclosure. In some embodiments, the sequence encoding the first RNA binding protein comprises a mutation that inactivates or decreases the DNA nuclease activity to a level at which the DNA nuclease activity does not induce a break in a DNA molecule, a mammalian DNA molecule or any DNA molecule when a composition of the disclosure is contacted to an RNA molecule or introduced into a cell or into a subject of the disclosure. In some embodiments, the sequence encoding the first RNA binding protein comprises a mutation that inactivates or decreases the DNA nuclease activity and the mutation comprises one or more of a substitution, inversion, transposition, insertion, deletion, or any combination thereof to a nucleic acid sequence or amino acid sequence encoding the first RNA binding protein or a nuclease domain thereof.

[099] In some embodiments of the compositions of the disclosure, the sequence encoding the first RNA binding protein of an RNA-guided fusion protein disclosed herein comprises a sequence isolated or derived from a CRISPR Cas protein. In some embodiments, the CRISPR Cas protein comprises a Type II CRISPR Cas protein. In some embodiments, the Type II CRISPR Cas protein comprises a Cas9 protein. Exemplary Cas9 proteins of the disclosure may be isolated or derived from any species, including, but not limited to, a bacteria or an archaea. Exemplary Cas9 proteins of the disclosure may be isolated or derived from any species, including, but not limited to, *Streptococcus pyogenes*, *Haloferax mediteranii*, *Mycobacterium tuberculosis*, *Francisella tularensis subsp. novicida*, *Pasteurella multocida*, *Neisseria meningitidis*, *Campylobacter jejune*, *Streptococcus thermophilus*, *Campylobacter lari CF89-12*, *Mycoplasma gallisepticum str. F*, *Nitratifractor salsuginis str. DSM 16511*, *Parvibaculum lavamentivorans*, *Roseburia intestinalis*, *Neisseria cinerea*, *a Gluconacetobacter diazotrophicus*,

an Azospirillum B510, a Sphaerochaeta globus str. Buddy, Flavobacterium columnare, Fluviicola taffensis, Bacteroides coprophilus, Mycoplasma mobile, Lactobacillus farciminis, Streptococcus pasteurianus, Lactobacillus johnsonii, Staphylococcus pseudintermedius, Filifactor alocis, Treponema denticola, Legionella pneumophila str. Paris, Sutterella wadsworthensis, Corynebacter diphtherias, Streptococcus aureus, and Francisella novicida.

[0100] Exemplary wild type *S. pyogenes* Cas9 proteins of the disclosure may comprise or consist of the amino acid sequence:

```
   1 MDKKYSIGLD IGTNSVGWAV ITDEYKVPSK KFKVLGNTDR HSIKKNLIGA LLFDSGETAE
  61 ATRLKRTARR RYTRRKNRIC YLQEIFSNEM AKVDDSFFHR LEESFLVEED KKHERHPIFG
 121 NIVDEVAYHE KYPTIYHLRK KLVDSTDKAD LRLIYLALAH MIKFRGHFLI EGDLNPDNSD
 181 VDKLFIQLVQ TYNQLFEENP INASGVDAKA ILSARLSKSR RLENLIAQLP GEKKNGLFGN
 241 LIALSLGLTP NFKSNFDLAE DAKLQLSKDT YDDDLDNLLA QIGDQYADLF LAAKNLSDAI
 301 LLSDILRVNT EITKAPLSAS MIKRYDEHHQ DLTLLKALVR QQLPEKYKEI FFDQSKNGYA
 361 GYIDGGASQE EFYKFIKPIL EKMDGTEELL VKLNREDLLR KQRTFDNGSI PHQIHLGELH
 421 AILRRQEDFY PFLKDNREKI EKILTFRIPY YVGPLARGNS RFAWMTRKSE ETITPWNFEE
 481 VVDKGASAQS FIERMTNFDK NLPNEKVLPK HSLLYEYFTV YNELTKVKYV TEGMRKPAFL
 541 SGEQKKAIVD LLFKTNRKVT VKQLKEDYFK KIECFDSVEI SGVEDRFNAS LGTYHDLLKI
 601 IKDKDFLDNE ENEDILEDIV LTLTLFEDRE MIEERLKTYA HLFDDKVMKQ LKRRRYTGWG
 661 RLSRKLINGI RDKQSGKTIL DFLKSDGFAN RNFMQLIHDD SLTFKEDIQK AQVSGQGDSL
 721 HEHIANLAGS PAIKKGILQT VKVVDELVKV MGRHKPENIV IEMARENQTT QKGQKNSRER
 781 MKRIEEGIKE LGSQILKEHP VENTQLQNEK LYLYYLQNGR DMYVDQELDI NRLSDYDVDH
 841 IVPQSFLKDD SIDNKVLTRS DKNRGKSDNV PSEEVVKKMK NYWRQLLNAK LITQRKFDNL
 901 TKAERGGLSE LDKAGFIKRQ LVETRQITKH VAQILDSRMN TKYDENDKLI REVKVITLKS
 961 KLVSDFRKDF QFYKVREINN YHHAHDAYLN AVVGTALIKK YPKLESEFVY GDYKVYDVRK
1021 MIAKSEQEIG KATAKYFFYS NIMNFFKTEI TLANGEIRKR PLIETNGETG EIVWDKGRDF
1081 ATVRKVLSMP QVNIVKKTEV QTGGFSKESI LPKRNSDKLI ARKKDWDPKK YGGFDSPTVA
1141 YSVLVVAKVE KGKSKKLKSV KELLGITIME RSSFEKNPID FLEAKGYKEV KKDLIIKLPK
1201 YSLFELENGR KRMLASAGEL QKGNELALPS KYVNFLYLAS HYEKLKGSPE DNEQKQLFVE
1261 QHKHYLDEII EQISEFSKRV ILADANLDKV LSAYNKHRDK PIREQAENII HLFTLTNLGA
1321 PAAFKYFDTT IDRKRYTSTK EVLDATLIHQ SITGLYETRI DLSQLGGD (SEQ ID NO: 147).
```

[0101] Nuclease inactivated *S. pyogenes* Cas9 proteins may comprise a substitution of an Alanine (A) for an Aspartic Acid (D) at position 10 and an alanine (A) for a Histidine (H) at position 840. Exemplary nuclease inactivated *S. pyogenes* Cas9 proteins of the disclosure may comprise or consist of the amino acid sequence (D10A and H840A bolded and underlined):

```
   1 MDKKYSIGLA IGTNSVGWAV ITDEYKVPSK KFKVLGNTDR HSIKKNLIGA LLFDSGETAE
  61 ATRLKRTARR RYTRRKNRIC YLQEIFSNEM AKVDDSFFHR LEESFLVEED KKHERHPIFG
 121 NIVDEVAYHE KYPTIYHLRK KLVDSTDKAD LRLIYLALAH MIKFRGHFLI EGDLNPDNSD
 181 VDKLFIQLVQ TYNQLFEENP INASGVDAKA ILSARLSKSR RLENLIAQLP GEKKNGLFGN
 241 LIALSLGLTP NFKSNFDLAE DAKLQLSKDT YDDDLDNLLA QIGDQYADLF LAAKNLSDAI
 301 LLSDILRVNT EITKAPLSAS MIKRYDEHHQ DLTLLKALVR QQLPEKYKEI FFDQSKNGYA
 361 GYIDGGASQE EFYKFIKPIL EKMDGTEELL VKLNREDLLR KQRTFDNGSI PHQIHLGELH
 421 AILRRQEDFY PFLKDNREKI EKILTFRIPY YVGPLARGNS RFAWMTRKSE ETITPWNFEE
 481 VVDKGASAQS FIERMTNFDK NLPNEKVLPK HSLLYEYFTV YNELTKVKYV TEGMRKPAFL
 541 SGEQKKAIVD LLFKTNRKVT VKQLKEDYFK KIECFDSVEI SGVEDRFNAS LGTYHDLLKI
 601 IKDKDFLDNE ENEDILEDIV LTLTLFEDRE MIEERLKTYA HLFDDKVMKQ LKRRRYTGWG
 661 RLSRKLINGI RDKQSGKTIL DFLKSDGFAN RNFMQLIHDD SLTFKEDIQK AQVSGQGDSL
 721 HEHIANLAGS PAIKKGILQT VKVVDELVKV MGRHKPENIV IEMARENQTT QKGQKNSRER
 781 MKRIEEGIKE LGSQILKEHP VENTQLQNEK LYLYYLQNGR DMYVDQELDI NRLSDYDVDA
 841 IVPQSFLKDD SIDNKVLTRS DKNRGKSDNV PSEEVVKKMK NYWRQLLNAK LITQRKFDNL
 901 TKAERGGLSE LDKAGFIKRQ LVETRQITKH VAQILDSRMN TKYDENDKLI REVKVITLKS
 961 KLVSDFRKDF QFYKVREINN YHHAHDAYLN AVVGTALIKK YPKLESEFVY GDYKVYDVRK
1021 MIAKSEQEIG KATAKYFFYS NIMNFFKTEI TLANGEIRKR PLIETNGETG EIVWDKGRDF
1081 ATVRKVLSMP QVNIVKKTEV QTGGFSKESI LPKRNSDKLI ARKKDWDPKK YGGFDSPTVA
1141 YSVLVVAKVE KGKSKKLKSV KELLGITIME RSSFEKNPID FLEAKGYKEV KKDLIIKLPK
1201 YSLFELENGR KRMLASAGEL QKGNELALPS KYVNFLYLAS HYEKLKGSPE DNEQKQLFVE
1261 QHKHYLDEII EQISEFSKRV ILADANLDKV LSAYNKHRDK PIREQAENII HLFTLTNLGA
1321 PAAFKYFDTT IDRKRYTSTK EVLDATLIHQ SITGLYETRI DLSQLGGD (SEQ ID NO: 148).
```

[0102] Nuclease inactivated *S. pyogenes* Cas9 proteins may comprise deletion of a RuvC nuclease domain or a portion thereof, an HNH domain, a DNAse active site, a ββα-metal fold or a portion thereof comprising a DNAse active site or any combination thereof.

[0103] Other exemplary Cas9 proteins or portions thereof may comprise or consist of the following amino acid sequences.

[0104] In some embodiments the Cas9 protein can be *S. pyogenes* Cas9 and may comprise or consist of the amino acid sequence:

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRR
YTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV
DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSAR
LSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQY
ADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNG

YAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYP
FLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLP
NEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE
CFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQ
GDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEG
IKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT
RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQIT
KHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALI
KKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETG
EIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA
YSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKR
MLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILAD
ANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE
TRIDLSQLGGD (SEQ ID NO: 149)

[0105] In some embodiments the Cas9 protein can be *S. aureus* Cas9 and may comprise or consist of the amino acid sequence:

MKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKKLLF
DYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSK
ALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEG
PGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQII
ENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIY
QSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLS
QQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIE
EIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEE
NSKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYA
TRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKEWKKLD
KAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKD
DKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGN
YLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLD
VIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYL
ENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG (SEQ ID NO:150)

[0106] In some embodiments the Cas9 protein can be *S. thermophiles* CRISPR1 Cas9 and may comprise or consist of the amino acid sequence:

MSDLVLGLDIGIGSVGVGILNKVTGEIIHKNSRIFPAAQAENNLVRRTNRQGRRLARRKKHRRVRLNRLFEE

SGLITDFTKISINLNPYQLRVKGLTDELSNEELFIALKNMVKHRGISYLDDASDDGNSSVGDYAQIVKENSK
QLETKTPGQIQLERYQTYGQLRGDFTVEKDGKKHRLINVFPTSAYRSEALRILQTQQEFNPQITDEFINRYLE
ILTGKRKYYHGPGNEKSRTDYGRYRTSGETLDNIFGILIGKCTFYPDEFRAAKASYTAQEFNLLNDLNNLTV
PTETKKLSKEQKNQIINYVKNEKAMGPAKLFKYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLE
TLDIEQMDRETLDKLAYVLTLNTEREGIQEALEHEFADGSFSQKQVDELVQFRKANSSIFGKGWHNFSVKL
MMELIPELYETSEEQMTILTRLGKQKTTSSSNKTKYIDEKLLTEEIYNPVVAKSVRQAIKIVNAAIKEYGDFD
NIVIEMARETNEDDEKKAIQKIQKANKDEKDAAMLKAANQYNGKAELPHSVFHGHKQLATKIRLWHQQG
ERCLYTGKTISIHDLINNSNQFEVDHILPLSITFDDSLANKVLVYATANQEKGQRTPYQALDSMDDAWSFRE
LKAFVRESKTLSNKKKEYLLTEEDISKFDVRKKFIERNLVDTRYASRVVLNALQEHFRAHKIDTKVSVVRG
QFTSQLRRHWGIEKTRDTYHHHAVDALIIAASSQLNLWKKQKNTLVSYSEDQLLDIETGELISDDEYKESVF
KAPYQHFVDTLKSKEFEDSILFSYQVDSKFNRKISDATIYATRQAKVGKDKADETYVLGKIKDIYTQDGYD
AFMKIYKKDKSKFLMYRHDPQTFEKVIEPILENYPNKQINDKGKEVPCNPFLKYKEEHGYIRKYSKKGNGP
EIKSLKYYDSKLGNHIDITPKDSNNKVVLQSVSPWRADVYFNKTTGKYEILGLKYADLQFDKGTGTYKISQ
EKYNDIKKKEGVDSDSEFKFTLYKNDLLLVKDTETKEQQLFRFLSRTMPKQKHYVELKPYDKQKFEGGEA
LIKVLGNVANSGQCKKGLGKSNISIYKVRTDVLGNQHIIKNEGDKPKLDF (SEQ ID NO: 151)

[0107] In some embodiments the Cas9 protein can be *N. meningitidis* Cas9 and may comprise or consist of the amino acid sequence:

MAAFKPNPINYILGLDIGIASVGWAMVEIDEDENPICLIDLGVRVFERAEVPKTGDSLAMARRLARSVRRLT
RRRAHRLLRARRLLKREGVLQAADFDENGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLS
QRKNEGETADKELGALLKGVADNAHALQTGDFRTPAELALNKFEKESGHIRNQRGDYSHTFSRKDLQAEL
ILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKL
NNLRILEQGSERPLTDTERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAY
HAISRALEKEGLKDKKSPLNLSPELQDEIGTAFSLFKTDEDITGRLKDRIQPEILEALLKHISFDKFVQISLKAL
RRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVVLRALSQARKVINGVVRRYGSPAR
IHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCLYSGKE
INLGRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETS
RFPRSKKQRILLQKFDEDGFKERNLNDTRYVNRFLCQFVADRMRLTGKGKKRVFASNGQITNLLRGFWGL
RKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGEVLHQKTHFPQPWEFFAQ
EVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSGQGHMETVKSAKR
LDEGVSVLRVPLTQLKLKDLEKMVNREREPKLYEALKARLEAHKDDPAKAFAEPFYKYDKAGNRTQQVK
AVRVEQVQKTGVWVRNHNGIADNATMVRVDVFEKGDKYYLVPIYSWQVAKGILPDRAVVQGKDEEDW
QLIDDSFNFKFSLHPNDLVEVITKKARMFGYFASCHRGTGNINIRIHDLDHKIGKNGILEGIGVKTALSFQKY
QIDELGKEIRPCRLKKRPPVR (SEQ ID NO: 152)

[0108] In some embodiments the Cas9 protein can be *Parvibaculum. lavamentivorans* Cas9 and may comprise or consist of the amino acid sequence:

MERIFGFDIGTTSIGFSVIDYSSTQSAGNIQRLGVRIFPEARDPDGTPLNQQRRQKRMMRRQLRRRRIRRKAL
NETLHEAGFLPAYGSADWPVVMADEPYELRRRGLEEGLSAYEFGRAIYHLAQHRHFKGRELEESDTPDPD
VDDEKEAANERAATLKALKNEQTTLGAWLARRPPSDRKRGIHAHRNVVAEEFERLWEVQSKFHPALKSE
EMRARISDTIFAQRPVFWRKNTLGECRFMPGEPLCPKGSWLSQQRRMLEKLNNLAIAGGNARPLDAEERD
AILSKLQQQASMSWPGVRSALKALYKQRGEPGAEKSLKFNLELGGESKLLGNALEAKLADMFGPDWPAH
PRKQEIRHAVHERLWAADYGETPDKKRVIILSEKDRKAHREAAANSFVADFGITGEQAAQLQALKLPTGW
EPYSIPALNLFLAELEKGERFGALVNGPDWEGWRRTNFPHRNQPTGEILDKLPSPASKEERERISQLRNPTV
VRTQNELRKVVNNLIGLYGKPDRIRIEVGRDVGKSKREREEIQSGIRRNEKQRKKATEDLIKNGIANPSRDD
VEKWILWKEGQERCPYTGDQIGFNALFREGRYEVEHIWPRSRSFDNSPRNKTLCRKDVNIEKGNRMPFEAF
GHDEDRWSAIQIRLQGMVSAKGGTGMSPGKVKRFLAKTMPEDFAARQLNDTRYAAKQILAQLKRLWPD
MGPEAPVKVEAVTGQVTAQLRKLWTLNNILADDGEKTRADHRHHAIDALTVACTHPGMTNKLSRYWQL
RDDPRAEKPALTPPWDTIRADAEKAVSEIVVSHRVRKKVSGPLHKETTYGDTGTDIKTKSGTYRQFVTRKK
IESLSKGELDEIRDPRIKEIVAAHVAGRGGDPKKAFPPYPCVSPGGPEIRKVRLTSKQQLNLMAQTGNGYAD
LGSNHHIAIYRLPDGKADFEIVSLFDASRRLAQRNPIVQRTRADGASFVMSLAAGEAIMIPEGSKKGIWIVQ
GVWASGQVVLERDTDADHSTTTRPMPNPILKDDAKKVSIDPIGRVRPSND (SEQ ID NO: 153)

[0109] In some embodiments the Cas9 protein can be *Corynebacter diphtheria* Cas9 and may comprise or consist of the amino acid sequence:

MKYHVGIDVGTFSVGLAAIEVDDAGMPIKTLSLVSHIHDSGLDPDEIKSAVTRLASSGIARRTRRLYRRKRR
RLQQLDKFIQRQGWPVIELEDYSDPLYPWKVRAELAASYIADEKERGEKLSVALRHIARHRGWRNPYAKV
SSLYLPDGPSDAFKAIREEIKRASGQPVPETATVGQMVTLCELGTLKLRGEGGVLSARLQQSDYAREIQEIC
RMQEIGQELYRKIIDVVFAAESPKGSASSRVGKDPLQPGKNRALKASDAFQRYRIAALIGNLRVRVDGEKRI
LSVEEKNLVFDHLVNLTPKKEPEWVTIAEILGIDRGQLIGTATMTDDGERAGARPPTHDTNRSIVNSRIAPL
VDWWKTASALEQHAMVKALSNAEVDDFDSPEGAKVQAFFADLDDDVHAKLDSLHLPVGRAAYSEDTLV
RLTRRMLSDGVDLYTARLQEFGIEPSWTPPTPRIGEPVGNPAVDRVLKTVSRWLESATKTWGAPERVIIEHV
REGFVTEKRAREMDGDMRRRAARNAKLFQEMQEKLNVQGKPSRADLWRYQSVQRQNCQCAYCGSPITF
SNSEMDHIVPRAGQGSTNTRENLVAVCHRCNQSKGNTPFAIWAKNTSIEGVSVKEAVERTRHWVTDTGM
RSTDFKKFTKAVVERFQRATMDEEIDARSMESVAWMANELRSRVAQHFASHGTTVRVYRGSLTAEARRA
SGISGKLKFFDGVGKSRLDRRHHAIDAAVIAFTSDYVAETLAVRSNLKQSQAHRQEAPQWREFTGKDAEH
RAAWRVWCQKMEKLSALLTEDLRDDRVVVMSNVRLRLGNGSAHKETIGKLSKVKLSSQLSVSDIDKASS
EALWCALTREPGFDPKEGLPANPERHIRVNGTHVYAGDNIGLFPVSAGSIALRGGYAELGSSFHHARVYKI
TSGKKPAFAMLRVYTIDLLPYRNQDLFSVELKPQTMSMRQAEKKLRDALATGNAEYLGWLVVDDELVVD
TSKIATDQVKAVEAELGTIRRWRVDGFFSPSKLRLRPLQMSKEGIKKESAPELSKIIDRPGWLPAVNKLFSD
GNVTVVRRDSLGRVRLESTAHLPVTWKVQ (SEQ ID NO: 154)

[0110] In some embodiments the Cas9 protein can be *Streptococcus pasteurianus* Cas9 and may comprise or consist of the amino acid sequence:

MTNGKILGLDIGIASVGVGIIEAKTGKVVHANSRLFSAANAENNAERRGFRGSRRLNRRKKHRVKRVRDLF
EKYGIVTDFRNLNLNPYELRVKGLTEQLKNEELFAALRTISKRRGISYLDDAEDDSTGSTDYAKSIDENRRL
LKNKTPGQIQLERLEKYGQLRGNFTVYDENGEAHRLINVFSTSDYEKEARKILETQADYNKKITAEFIDDYV
EILTQKRKYYHGPGNEKSRTDYGRFRTDGTTLENIFGILIGKCNFYPDEYRASKASYTAQEYNFLNDLNNLK
VSTETGKLSTEQKESLVEFAKNTATLGPAKLLKEIAKILDCKVDEIKGYREDDKGKPDLHTFEPYRKLKFNL
ESINIDDLSREVIDKLADILTLNTEREGIEDAIKRNLPNQFTEEQISEIIKVRKSQSTAFNKGWHSFSAKLMNE
LIPELYATSDEQMTILTRLEKFKVNKKSSKNTKTIDEKEVTDEIYNPVVAKSVRQTIKIINAAVKKYGDFDKI
VIEMPRDKNADDEKKFIDKRNKENKKEKDDALKRAAYLYNSSDKLPDEVFHGNKQLETKIRLWYQQGER
CLYSGKPISIQELVHNSNNFEIDHILPLSLSFDDSLANKVLVYAWTNQEKGQKTPYQVIDSMDAAWSFREM
KDYVLKQKGLGKKKRDYLLTTENIDKIEVKKKFIERNLVDTRYASRVVLNSLQSALRELGKDTKVSVVRG
QFTSQLRRKWKIDKSRETYHHHAVDALIIAASSQLKLWEKQDNPMFVDYGKNQVVDKQTGEILSVSDDEY
KELVFQPPYQGFVNTISSKGFEDEILFSYQVDSKYNRKVSDATIYSTRKAKIGKDKKEETYVLGKIKDIYSQ
NGFDTFIKKYNKDKTQFLMYQKDSLTWENVIEVILRDYPTTKKSEDGKNDVKCNPFEEYRRENGLICKYSK
KGKGTPIKSLKYYDKKLGNCIDITPEESRNKVILQSINPWRADVYFNPETLKYELMGLKYSDLSFEKGTGNY
HISQEKYDAIKEKEGIGKKSEFKFTLYRNDLILIKDIASGEQEIYRFLSRTMPNVNHYVELKPYDKEKFDNVQ
ELVEALGEADKVGRCIKGLNKPNISIYKVRTDVLGNKYFVKKKGDKPKLDFKNNKK (SEQ ID NO: 155)

[0111] In some embodiments the Cas9 protein can be *Neisseria cinerea* Cas9 and may comprise or consist of the amino acid sequence:

MAAFKPNPMNYILGLDIGIASVGWAIVEIDEEENPIRLIDLGVRVFERAEVPKTGDSLAAARRLARSVRRLT
RRRAHRLLRARRLLKREGVLQAADFDENGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLS
QRKNEGETADKELGALLKGVADNTHALQTGDFRTPAELALNKFEKESGHIRNQRGDYSHTFNRKDLQAEL
NLLFEKQKEFGNPHVSDGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPTEPKAAKNTYTAERFVWLTK
LNNLRILEQGSERPLTDTERATLMDEPYRKSKLTYAQARKLLDLDDTAFFKGLRYGKDNAEASTLMEMKA
YHAISRALEKEGLKDKKSPLNLSPELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDKFVQISLK
ALRRIVPLMEQGNRYDEACTEIYGDHYGKKNTEEKIYLPPIPADEIRNPVVLRALSQARKVINGVVRRYGSP
ARIHIETAREVGKSFKDRKEIEKRQEENRKDREKSAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCLYSG
KEINLGRLNEKGYVEIDHALPFSRTWDDSFNNKVLALGSENQNKGNQTPYEYFNGKDNSREWQEFKARVE
TSRFPRSKKQRILLQKFDEDGFKERNLNDTRYINRFLCQFVADHMLLTGKGKRRVFASNGQITNLLRGFWG
LRKVRAENDRHHALDAVVVACSTIAMQQKITRFVRYKEMNAFDGKTIDKETGEVLHQKAHFPQPWEFFA
QEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHKYVTPLFISRAPNRKMSGQGHMETVKSAK
RLDEGISVLRVPLTQLKLKDLEKMVNREREPKLYEALKARLEAHKDDPAKAFAEPFYKYDKAGNRTQQV
KAVRVEQVQKTGVWVHNHNGIADNATIVRVDVFEKGGKYYLVPIYSWQVAKGILPDRAVVQGKDEEDW
TVMDDSFEFKFVLYANDLIKLTAKKNEFLGYFVSLNRATGAIDIRTHDTDSTKGKNGIFQSVGVKTALSFQ
KYQIDELGKEIRPCRLKKRPPVR (SEQ ID NO: 156)

[0112] In some embodiments the Cas9 protein can be *Campylobacter lari* Cas9 and may comprise or consist of the amino acid sequence:

MRILGFDIGINSIGWAFVENDELKDCGVRIFTKAENPKNKESLALPRRNARSSRRRLKRRKARLIAIKRILAK
ELKLNYKDYVAADGELPKAYEGSLASVYELRYKALTQNLETKDLARVILHIAKHRGYMNKNEKKSNDAK
KGKILSALKNNALKLENYQSVGEYFYKEFFQKYKKNTKNFIKIRNTKDNYNNCVLSSDLEKELKLILEKQK
EFGYNYSEDFINEILKVAFFQRPLKDFSHLVGACTFFEEEKRACKNSYSAWEFVALTKIINEIKSLEKISGEIV
PTQTINEVLNLILDKGSITYKKFRSCINLHESISFKSLKYDKENAENAKLIDFRKLVEFKKALGVHSLSRQEL
DQISTHITLIKDNVKLKTVLEKYNLSNEQINNLLEIEFNDYINLSFKALGMILPLMREGKRYDEACEIANLKP
KTVDEKKDFLPAFCDSIFAHELSNPVVNRAISEYRKVLNALLKKYGKVHKIHLELARDVGLSKKAREKIEK
EQKENQAVNAWALKECENIGLKASAKNILKLKLWKEQKEICIYSGNKISIEHLKDEKALEVDHIYPYSRSFD
DSFINKVLVFTKENQEKLNKTPFEAFGKNIEKWSKIQTLAQNLPYKKKNKILDENFKDKQQEDFISRNLNDT
RYIATLIAKYTKEYLNFLLLSENENANLKSGEKGSKIHVQTISGMLTSVLRHTWGFDKKDRNNHLHHALDA
IIVAYSTNSIIKAFSDFRKNQELLKARFYAKELTSDNYKHQVKFFEPFKSFREKILSKIDEIFVSKPPRKRARR
ALHKDTFHSENKIIDKCSYNSKEGLQIALSCGRVRKIGTKYVENDTIVRVDIFKKQNKFYAIPIYAMDFALGI
LPNKIVITGKDKNNNPKQWQTIDESYEFCFSLYKNDLILLQKKNMQEPEFAYYNDFSISTSSICVEKHDNKF
ENLTSNQKLLFSNAKEGSVKVESLGIQNLKVFEKYIITPLGDKIKADFQPRENISLKTSKKYGLR (SEQ ID NO: 157)

[0113] In some embodiments the Cas9 protein can be *T. denticola* Cas9 and may comprise or consist of the amino acid sequence:

MKKEIKDYFLGLDVGTGSVGWAVTDTDYKLLKANRKDLWGMRCFETAETAEVRRLHRGARRRIERRKK
RIKLLQELFSQEIAKTDEGFFQRMKESPFYAEDKTILQENTLFNDKDFADKTYHKAYPTINHLIKAWIENKV
KPDPRLLYLACHNIIKKRGHFLFEGDFDSENQFDTSIQALFEYLREDMEVDIDADSQKVKEILKDSSLKNSE
KQSRLNKILGLKPSDKQKKAITNLISGNKINFADLYDNPDLKDAEKNSISFSKDDFDALSDDLASILGDSFEL
LLKAKAVYNCSVLSKVIGDEQYLSFAKVKIYEKHKTDLTKLKNVIKKHFPKDYKKVFGYNKNEKNNNNY
SGYVGVCKTKSKKLIINNSVNQEDFYKFLKTILSAKSEIKEVNDILTEIETGTFLPKQISKSNAEIPYQLRKME
LEKILSNAEKHFSFLKQKDEKGLSHSEKIIMLLTFKIPYYIGPINDNHKKFFPDRCWVVKKEKSPSGKTTPWN
FFDHIDKEKTAEAFITSRTNFCTYLVGESVLPKSSLLYSEYTVLNEINNLQIIIDGKNICDIKLKQKIYEDLFKK
YKKITQKQISTFIKHEGICNKTDEVIILGIDKECTSSLKSYIELKNIFGKQVDEISTKNMLEEIIRWATIYDEGE
GKTILKTKIKAEYGKYCSDEQIKKILNLKFSGWGRLSRKFLETVTSEMPGFSEPVNIITAMRETQNNLMELLS
SEFTFTENIKKINSGFEDAEKQFSYDGLVKPLFLSPSVKKMLWQTLKLVKEISHITQAPPKKIFIEMAKGAEL
EPARTKTRLKILQDLYNNCKNDADAFSSEIKDLSGKIENEDNLRLRSDKLYLYYTQLGKCMYCGKPIEIGH
VFDTSNYDIDHIYPQSKIKDDSISNRVLVCSSCNKNKEDKYPLKSEIQSKQRGFWNFLQRNNFISLEKLNRLT
RATPISDDETAKFIARQLVETRQATKVAAKVLEKMFPETKIVYSKAETVSMFRNKFDIVKCREINDFHHAH
DAYLNIVVGNVYNTKFTNNPWNFIKEKRDNPKIADTYNYYKVFDYDVKRNNITAWEKGKTIITVKDMLKR
NTPIYTRQAACKKGELFNQTIMKKGLGQHPLKKEGPFSNISKYGGYNKVSAAYYTLIEYEEKGNKIRSLETI

PLYLVKDIQKDQDVLKSYLTDLLGKKEFKILVPKIKINSLLKINGFPCHITGKTNDSFLLRPAVQFCCSNNEV
LYFKKIIRFSEIRSQREKIGKTISPYEDLSFRSYIKENLWKKTKNDEIGEKEFYDLLQKKNLEIYDMLLTKHKD
TIYKKRPNSATIDILVKGKEKFKSLIIENQFEVILEILKLFSATRNVSDLQHIGGSKYSGVAKIGNKISSLDNCI
LIYQSITGIFEKRIDLLKV (SEQ ID NO: 158)

[0114] In some embodiments the Cas9 protein can be *S. mutans* Cas9 and may comprise or consist of the amino acid sequence:

MKKPYSIGLDIGTNSVGWAVVTDDYKVPAKKMKVLGNTDKSHIEKNLLGALLFDSGNTAEDRRLKRTAR
RRYTRRRNRILYLQEIFSEEMGKVDDSFFHRLEDSFLVTEDKRGERHPIFGNLEEEVKYHENFPTIYHLRQYL
ADNPEKVDLRLVYLALAHIIKFRGHFLIEGKFDTRNNDVQRLFQEFLAVYDNTFENSSLQEQNVQVEEILTD
KISKSAKKDRVLKLFPNEKSNGRFAEFLKLIVGNQADFKKHFELEEKAPLQFSKDTYEEELEVLLAQIGDNY
AELFLSAKKLYDSILLSGILTVTDVGTKAPLSASMIQRYNEHQMDLAQLKQFIRQKLSDKYNEVFSDVSKD
GYAGYIDGKTNQEAFYKYLKGLLNKIEGSGYFLDKIEREDFLRKQRTFDNGSIPHQIHLQEMRAIIRRQAEF
YPFLADNQDRIEKLLTFRIPYYVGPLARGKSDFAWLSRKSADKITPWNFDEIVDKESSAEAFINRMTNYDLY
LPNQKVLPKHSLLYEKFTVYNELTKVKYKTEQGKTAFFDANMKQEIFDGVFKVYRKVTKDKLMDFLEKE
FDEFRIVDLTGLDKENKVFNASYGTYHDLCKILDKDFLDNSKNEKILEDIVLTLTLFEDREMIRKRLENYSD
LLTKEQVKKLERRHYTGWGRLSAELIHGIRNKESRKTILDYLIDDGNSNRNFMQLINDDALSFKEEIAKAQV
IGETDNLNQVVSDIAGSPAIKKGILQSLKIVDELVKIMGHQPENIVVEMARENQFTNQGRRNSQQRLKGLTD
SIKEFGSQILKEHPVENSQLQNDRLFLYYLQNGRDMYTGEELDIDYLSQYDIDHIIPQAFIKDNSIDNRVLTSS
KENRGKSDDVPSKDVVRKMKSYWSKLLSAKLITQRKFDNLTKAERGGLTDDDKAGFIKRQLVETRQITKH
VARILDERFNTETDENNKKIRQVKIVTLKSNLVSNFRKEFELYKVREINDYHHAHDAYLNAVIGKALLGVY
PQLEPEFVYGDYPHFHGHKENKATAKKFFYSNIMNFFKKDDVRTDKNGEIIWKKDEHISNIKKVLSYPQVN
IVKKVEEQTGGFSKESILPKGNSDKLIPRKTKKFYWDTKKYGGFDSPIVAYSILVIADIEKGKSKKLKTVKAL
VGVTIMEKMTFERDPVAFLERKGYRNVQEENIIKLPKYSLFKLENGRKRLLASARELQKGNEIVLPNHLGT
LLYHAKNIHKVDEPKHLDYVDKHKDEFKELLDVVSNFSKKYTLAEGNLEKIKELYAQNNGEDLKELASSFI
NLLTFTAIGAPATFKFFDKNIDRKRYTSTTEILNATLIHQSITGLYETRIDLNKLGGD (SEQ ID NO: 159)

[0115] In some embodiments the Cas9 protein can be *S. thermophilus* CRISPR 3 Cas9 and may comprise or consist of the amino acid sequence:

MTKPYSIGLDIGTNSVGWAVTTDNYKVPSKKMKVLGNTSKKYIKKNLLGVLLFDSGITAEGRRLKRTARR
RYTRRRNRILYLQEIFSTEMATLDDAFFQRLDDSFLVPDDKRDSKYPIFGNLVEEKAYHDEFPTIYHLRKYL
ADSTKKADLRLVYLALAHMIKYRGHFLIEGEFNSKNNDIQKNFQDFLDTYNAIFESDLSLENSKQLEEIVKD
KISKLEKKDRILKLFPGEKNSGIFSEFLKLIVGNQADFRKCFNLDEKASLHFSKESYDEDLETLLGYIGDDYS
DVFLKAKKLYDAILLSGFLTVTDNETEAPLSSAMIKRYNEHKEDLALLKEYIRNISLKTYNEVFKDDTKNG
YAGYIDGKTNQEDFYVYLKKLLAEFEGADYFLEKIDREDFLRKQRTFDNGSIPYQIHLQEMRAILDKQAKF
YPFLAKNKERIEKILTFRIPYYVGPLARGNSDFAWSIRKRNEKITPWNFEDVIDKESSAEAFINRMTSFDLYL
PEEKVLPKHSLLYETFNVYNELTKVRFIAESMRDYQFLDSKQKKDIVRLYFKDKRKVTDKDIIEYLHAIYGY

DGIELKGIEKQFNSSLSTYHDLLNIINDKEFLDDSSNEAIIEEIIHTLTIFEDREMIKQRLSKFENIFDKSVLKKL
SRRHYTGWGKLSAKLINGIRDEKSGNTILDYLIDDGISNRNFMQLIHDDALSFKKKIQKAQIIGDEDKGNIKE
VVKSLPGSPAIKKGILQSIKIVDELVKVMGGRKPESIVVEMARENQYTNQGKSNSQQRLKRLEKSLKELGS
KILKENIPAKLSKIDNNALQNDRLYLYYLQNGKDMYTGDDLDIDRLSNYDIDHIIPQAFLKDNSIDNKVLVS
SASNRGKSDDVPSLEVVKKRKTFWYQLLKSKLISQRKFDNLTKAERGGLSPEDKAGFIQRQLVETRQITKH
VARLLDEKFNNKKDENNRAVRTVKIITLKSTLVSQFRKDFELYKVREINDFHHAHDAYLNAVVASALLKK
YPKLEPEFVYGDYPKYNSFRERKSATEKVYFYSNIMNIFKKSISLADGRVIERPLIEVNEETGESVWNKESDL
ATVRRVLSYPQVNVVKKVEEQNHGLDRGKPKGLFNANLSSKPKPNSNENLVGAKEYLDPKKYGGYAGIS
NSFTVLVKGTIEKGAKKKITNVLEFQGISILDRINYRKDKLNFLLEKGYKDIELIIELPKYSLFELSDGSRRML
ASILSTNNKRGEIHKGNQIFLSQKFVKLLYHAKRISNTINENHRKYVENHKKEFEELFYYILEFNENYVGAK
KNGKLLNSAFQSWQNHSIDELCSSFIGPTGSERKGLFELTSRGSAADFEFLGVKIPRYRDYTPSSLLKDATLI
HQSVTGLYETRIDLAKLGEG (SEQ ID NO: 160)

[0116] In some embodiments the Cas9 protein can be *C. jejuni* Cas9 and may comprise or consist of the amino acid sequence:

MARILAFDIGISSIGWAFSENDELKDCGVRIFTKVENPKTGESLALPRRLARSARKRLARRKARLNHLKHLI
ANEFKLNYEDYQSFDESLAKAYKGSLISPYELRFRALNELLSKQDFARVILHIAKRRGYDDIKNSDDKEKG
AILKAIKQNEEKLANYQSVGEYLYKEYFQKFKENSKEFTNVRNKKESYERCIAQSFLKDELKLIFKKQREFG
FSFSKKFEEEVLSVAFYKRALKDFSHLVGNCSFFTDEKRAPKNSPLAFMFVALTRIINLLNNLKNTEGILYTK
DDLNALLNEVLKNGTLTYKQTKKLLGLSDDYEFKGEKGTYFIEFKKYKEFIKALGEHNLSQDDLNEIAKDI
TLIKDEIKLKKALAKYDLNQNQIDSLSKLEFKDHLNISFKALKLVTPLMLEGKKYDEACNELNLKVAINED
KKDFLPAFNETYYKDEVTNPVVLRAIKEYRKVLNALLKKYGKVHKINIELAREVGKNHSQRAKIEKEQNE
NYKAKKDAELECEKLGLKINSKNILKLRLFKEQKEFCAYSGEKIKISDLQDEKMLEIDHIYPYSRSFDDSYM
NKVLVFTKQNQEKLNQTPFEAFGNDSAKWQKIEVLAKNLPTKKQKRILDKNYKDKEQKNFKDRNLNDTR
YIARLVLNYTKDYLDFLPLSDDENTKLNDTQKGSKVHVEAKSGMLTSALRHTWGFSAKDRNNHLHHAID
AVIIAYANNSIVKAFSDFKKEQESNSAELYAKKISELDYKNKRKFFEPFSGFRQKVLDKIDEIFVSKPERKKP
SGALHEETFRKEEEFYQSYGGKEGVLKALELGKIRKVNGKIVKNGDMFRVDIFKHKKTNKFYAVPIYTMD
FALKVLPNKAVARSKKGEIKDWILMDENYEFCFSLYKDSLILIQTKDMQEPEFVYYNAFTSSTVSLIVSKHD
NKFETLSKNQKILFKNANEKEVIAKSIGIQNLKVFEKYIVSALGEVTKAEFRQREDFKK (SEQ ID NO: 161)

[0117] In some embodiments the Cas9 protein can be *P. multocida* Cas9 and may comprise or consist of the amino acid sequence:

MQTTNLSYILGLDLGIASVGWAVVEINENEDPIGLIDVGVRIFERAEVPKTGESLALSRRLARSTRRLIRRRA
HRLLLAKRFLKREGILSTIDLEKGLPNQAWELRVAGLERRLSAIEWGAVLLHLIKHRGYLSKRKNESQTNN
KELGALLSGVAQNHQLLQSDDYRTPAELALKKFAKEEGHIRNQRGAYTHTFNRLDLLAELNLLFAQQHQF
GNPHCKEHIQQYMTELLMWQKPALSGEAILKMLGKCTHEKNEFKAAKHTYSAERFVWLTKLNNLRILED
GAERALNEEERQLLINHPYEKSKLTYAQVRKLLGLSEQAIFKHLRYSKENAESATFMELKAWHAIRKALEN

QGLKDTWQDLAKKPDLLDEIGTAFSLYKTDEDIQQYLTNKVPNSVINALLVSLNFDKFIELSLKSLRKILPL
MEQGKRYDQACREIYGHHYGEANQKTSQLLPAIPAQEIRNPVVLRTLSQARKVINAIIRQYGSPARVHIETG
RELGKSFKERREIQKQQEDNRTKRESAVQKFKELFSDFSSEPKSKDILKFRLYEQQHGKCLYSGKEINIHRL
NEKGYVEIDHALPFSRTWDDSFNNKVLVLASENQNKGNQTPYEWLQGKINSERWKNFVALVLGSQCSAA
KKQRLLTQVIDDNKFIDRNLNDTRYIARFLSNYIQENLLLVGKNKKNVFTPNGQITALLRSRWGLIKARENN
NRHHALDAIVVACATPSMQQKITRFIRFKEVHPYKIENRYEMVDQESGEIISPHFPEPWAYFRQEVNIRVFD
NHPDTVLKEMLPDRPQANHQFVQPLFVSRAPTRKMSGQGHMETIKSAKRLAEGISVLRIPLTQLKPNLLEN
MVNKEREPALYAGLKARLAEFNQDPAKAFATPFYKQGGQQVKAIRVEQVQKSGVLVRENNGVADNASIV
RTDVFIKNNKFFLVPIYTWQVAKGILPNKAIVAHKNEDEWEEMDEGAKFKFSLFPNDLVELKTKKEYFFGY
YIGLDRATGNISLKEHDGEISKGKDGVYRVGVKLALSFEKYQVDELGKNRQICRPQQRQPVR (SEQ ID NO:
162)

[0118] In some embodiments the Cas9 protein can be *F. novicida* Cas9 and may comprise or consist of the amino acid sequence:

MNFKILPIAIDLGVKNTGVFSAFYQKGTSLERLDNKNGKVYELSKDSYTLLMNNRTARRHQRRGIDRKQL
VKRLFKLIWTEQLNLEWDKDTQQAISFLFNRRGFSFITDGYSPEYLNIVPEQVKAILMDIFDDYNGEDDLDS
YLKLATEQESKISEIYNKLMQKILEFKLMKLCTDIKDDKVSTKTLKEITSYEFELLADYLANYSESLKTQKFS
YTDKQGNLKELSYYHHDKYNIQEFLKRHATINDRILDTLLTDDLDIWNFNFEKFDFDKNEEKLQNQEDKD
HIQAHLHHFVFAVNKIKSEMASGGRHRSQYFQEITNVLDENNHQEGYLKNFCENLHNKKYSNLSVKNLVN
LIGNLSNLELKPLRKYFNDKIHAKADHWDEQKFTETYCHWILGEWRVGVKDQDKKDGAKYSYKDLCNEL
KQKVTKAGLVDFLLELDPCRTIPPYLDNNNRKPPKCQSLILNPKFLDNQYPNWQQYLQELKKLQSIQNYLD
SFETDLKVLKSSKDQPYFVEYKSSNQQIASGQRDYKDLDARILQFIFDRVKASDELLLNEIYFQAKKLKQKA
SSELEKLESSKKLDEVIANSQLSQILKSQHTNGIFEQGTFLHLVCKYYKQRQRARDSRLYIMPEYRYDKKLH
KYNNTGRFDDDNQLLTYCNHKPRQKRYQLLNDLAGVLQVSPNFLKDKIGSDDDLFISKWLVEHIRGFKKA
CEDSLKIQKDNRGLLNHKINIARNTKGKCEKEIFNLICKIEGSEDKKGNYKHGLAYELGVLLFGEPNEASKP
EFDRKIKKFNSIYSFAQIQQIAFAERKGNANTCAVCSADNAHRMQQIKITEPVEDNKDKIILSAKAQRLPAIP
TRIVDGAVKKMATILAKNIVDDNWQNIKQVLSAKHQLHIPIITESNAFEFEPALADVKGKSLKDRRKKALE
RISPENIFKDKNNRIKEFAKGISAYSGANLTDGDFDGAKEELDHIIPRSHKKYGTLNDEANLICVTRGDNKN
KGNRIFCLRDLADNYKLKQFETTDDLEIEKKIADTIWDANKKDFKFGNYRSFINLTPQEQKAFRHALFLADE
NPIKQAVIRAINNRNRTFVNGTQRYFAEVLANNIYLRAKKENLNTDKISFDYFGIPTIGNGRGIAEIRQLYEK
VDSDIQAYAKGDKPQASYSHLIDAMLAFCIAADEHRNDGSIGLEIDKNYSLYPLDKNTGEVFTKDIFSQIKIT
DNEFSDKKLVRKKAIEGFNTHRQMTRDGIYAENYLPILIHKELNEVRKGYTWKNSEEIKIFKGKKYDIQQL
NNLVYCLKFVDKPISIDIQISTLEELRNILTTNNIAATAEYYYINLKTQKLHEYYIENYNTALGYKKYSKEME
FLRSLAYRSERVKIKSIDDVKQVLDKDSNFIIGKITLPFKKEWQRLYREWQNTTIKDDYEFLKSFFNVKSITK
LHKKVRKDFSLPISTNEGKFLVKRKTWDNNFIYQILNDSDSRADGTKPFIPAFDISKNEIVEAIIDSFTSKNIF
WLPKNIELQKVDNKNIFAIDTSKWFEVETPSDLRDIGIATIQYKIDNNSRPKVRVKLDYVIDDDSKINYFMN
HSLLKSRYPDKVLEILKQSTIIEFESSGFNKTIKEMLGMKLAGIYNETSNN (SEQ ID NO: 163)

[0119] In some embodiments the Cas9 protein can be *Lactobacillus buchneri* Cas9 and may comprise or consist of the amino acid sequence:

MKVNNYHIGLDIGTSSIGWVAIGKDGKPLRVKGKTAIGARLFQEGNPAADRRMFRTTRRRLSRRKWRLKL
LEEIFDPYITPVDSTFFARLKQSNLSPKDSRKEFKGSMLFPDLTDMQYHKNYPTIYHLRHALMTQDKKFDIR
MVYLAIHHIVKYRGNFLNSTPVDSFKASKVDFVDQFKKLNELYAAINPEESFKINLANSEDIGHQFLDPSIRK
FDKKKQIPKIVPVMMNDKVTDRLNGKIASEIIHAILGYKAKLDVVLQCTPVDSKPWALKFDDEDIDAKLEK
ILPEMDENQQSIVAILQNLYSQVTLNQIVPNGMSLSESMIEKYNDHHDHLKLYKKLIDQLADPKKKAVLKK
AYSQYVGDDGKVIEQAEFWSSVKKNLDDSELSKQIMDLIDAEKFMPKQRTSQNGVIPHQLHQRELDEIIEH
QSKYYPWLVEINPNKHDLHLAKYKIEQLVAFRVPYYVGPMITPKDQAESAETVFSWMERKGTETGQITPW
NFDEKVDRKASANRFIKRMTTKDTYLIGEDVLPDESLLYEKFKVLNELNMVRVNGKLLKVADKQAIFQDL
FENYKHVSVKKLQNYIKAKTGLPSDPEISGLSDPEHFNNSLGTYNDFKKLFGSKVDEPDLQDDFEKIVEWST
VFEDKKILREKLNEITWLSDQQKDVLESSRYQGWGRLSKKLLTGIVNDQGERIIDKLWNTNKNFMQIQSDD
DFAKRIHEANADQMQAVDVEDVLADAYTSPQNKKAIRQVVKVVDDIQKAMGGVAPKYISIEFTRSEDRNP
RRTISRQRQLENTLKDTAKSLAKSINPELLSELDNAAKSKKGLTDRLYLYFTQLGKDIYTGEPINIDELNKYD
IDHILPQAFIKDNSLDNRVLVLTAVNNGKSDNVPLRMFGAKMGHFWKQLAEAGLISKRKLKNLQTDPDTIS
KYAMHGFIRRQLVETSQVIKLVANILGDKYRNDDTKIIEITARMNHQMRDEFGFIKNREINDYHHAFDAYL
TAFLGRYLYHRYIKLRPYFVYGDFKKFREDKVTMRNFNFLHDLTDDTQEKIADAETGEVIWDRENSIQQLK
DVYHYKFMLISHEVYTLRGAMFNQTVYPASDAGKRKLIPVKADRPVNVYGGYSGSADAYMAIVRIHNKK
GDKYRVVGVPMRALDRLDAAKNVSDADFDRALKDVLAPQLTKTKKSRKTGEITQVIEDFEIVLGKVMYR
QLMIDGDKKFMLGSSTYQYNAKQLVLSDQSVKTLASKGRLDPLQESMDYNNVYTEILDKVNQYFSLYDM
NKFRHKLNLGFSKFISFPNHNVLDGNTKVSSGKREILQEILNGLHANPTFGNLKDVGITTPFGQLQQPNGILL
SDETKIRYQSPTGLFERTVSLKDL (SEQ ID NO: 164)

[0120] In some embodiments the Cas9 protein can be *Listeria innocua* Cas9 and may comprise or consist of the amino acid sequence:

MKKPYTIGLDIGTNSVGWAVLTDQYDLVKRKMKIAGDSEKKQIKKNFWGVRLFDEGQTAADRRMARTA
RRRIERRRNRISYLQGIFAEEMSKTDANFFCRLSDSFYVDNEKRNSRHPFFATIEEEVEYHKNYPTIYHLREE
LVNSSEKADLRLVYLALAHIIKYRGNFLIEGALDTQNTSVDGIYKQFIQTYNQVFASGIEDGSLKKLEDNKD
VAKILVEKVTRKEKLERILKLYPGEKSAGMFAQFISLIVGSKGNFQKPFDLIEKSDIECAKDSYEEDLESLLA
LIGDEYAELFVAAKNAYSAVVLSSIITVAETETNAKLSASMIERFDTHEEDLGELKAFIKLHLPKHYEEIFSN
TEKHGYAGYIDGKTKQADFYKYMKMTLENIEGADYFIAKIEKENFLRKQRTFDNGAIPHQLHLEELEAILH
QQAKYYPFLKENYDKIKSLVTFRIPYFVGPLANGQSEFAWLTRKADGEIRPWNIEEKVDFGKSAVDFIEKM
TNKDTYLPKENVLPKHSLCYQKYLVYNELTKVRYINDQGKTSYFSGQEKEQIFNDLFKQKRKVKKKDLEL
FLRNMSHVESPTIEGLEDSFNSSYSTYHDLLKVGIKQEILDNPVNTEMLENIVKILTVFEDKRMIKEQLQQFS
DVLDGVVLKKLERRHYTGWGRLSAKLLMGIRDKQSHLTILDYLMNDDGLNRNLMQLINDSNLSFKSIIEK
EQVTTADKDIQSIVADLAGSPAIKKGILQSLKIVDELVSVMGYPPQTIVVEMARENQTTGKGKNNSRPRYKS

LEKAIKEFGSQILKEHPTDNQELRNNRLYLYYLQNGKDMYTGQDLDIHNLSNYDIDHIVPQSFITDNSIDNL
VLTSSAGNREKGDDVPPLEIVRKRKVFWEKLYQGNLMSKRKFDYLTKAERGGLTEADKARFIHRQLVETR
QITKNVANILHQRFNYEKDDHGNTMKQVRIVTLKSALVSQFRKQFQLYKVRDVNDYHHAHDAYLNGVV
ANTLLKVYPQLEPEFVYGDYHQFDWFKANKATAKKQFYTNIMLFFAQKDRIIDENGEILWDKKYLDTVKK
VMSYRQMNIVKKTEIQKGEFSKATIKPKGNSSKLIPRKTNWDPMKYGGLDSPNMAYAVVIEYAKGKNKLV
FEKKIIRVTIMERKAFEKDEKAFLEEQGYRQPKVLAKLPKYTLYECEEGRRRMLASANEAQKGNQQVLPN
HLVTLLHHAANCEVSDGKSLDYIESNREMFAELLAHVSEFAKRYTLAEANLNKINQLFEQNKEGDIKAIAQ
SFVDLMAFNAMGAPASFKFFETTIERKRYNNLKELLNSTIIYQSITGLYESRKRLDD (SEQ ID NO: 165)

[0121] In some embodiments the Cas9 protein can be *L. pneumophilia* Cas9 and may comprise or consist of the amino acid sequence:

MESSQILSPIGIDLGGKFTGVCLSHLEAFAELPNHANTKYSVILIDHNNFQLSQAQRRATRHRVRNKKRNQF
VKRVALQLFQHILSRDLNAKEETALCHYLNNRGYTYVDTDLDEYIKDETTINLLKELLPSESEHNFIDWFLQ
KMQSSEFRKILVSKVEEKKDDKELKNAVKNIKNFITGFEKNSVEGHRHRKVYFENIKSDITKDNQLDSIKKK
IPSVCLSNLLGHLSNLQWKNLHRYLAKNPKQFDEQTFGNEFLRMLKNFRHLKGSQESLAVRNLIQQLEQSQ
DYISILEKTPPEITIPPYEARTNTGMEKDQSLLLNPEKLNNLYPNWRNLIPGIIDAHPFLEKDLEHTKLRDRKR
IISPSKQDEKRDSYILQRYLDLNKKIDKFKIKKQLSFLGQGKQLPANLIETQKEMETHFNSSLVSVLIQIASAY
NKEREDAAQGIWFDNAFSLCELSNINPPRKQKILPLLVGAILSEDFINNKDKWAKFKIFWNTHKIGRTSLKS
KCKEIEEARKNSGNAFKIDYEEALNHPEHSNNKALIKIIQTIPDIIQAIQSHLGHNDSQALIYHNPFSLSQLYTI
LETKRDGFHKNCVAVTCENYWRSQKTEIDPEISYASRLPADSVRPFDGVLARMMQRLAYEIAMAKWEQIK
HIPDNSSLLIPIYLEQNRFEFEESFKKIKGSSSDKTLEQAIEKQNIQWEEKFQRIINASMNICPYKGASIGGQGE
IDHIYPRSLSKKHFGVIFNSEVNLIYCSSQGNREKKEEHYLLEHLSPLYLKHQFGTDNVSDIKNFISQNVANI
KKYISFHLLTPEQQKAARHALFLDYDDEAFKTITKFLMSQQKARVNGTQKFLGKQIMEFLSTLADSKQLQL
EFSIKQITAEEVHDHRELLSKQEPKLVKSRQQSFPSHAIDATLTMSIGLKEFPQFSQELDNSWFINHLMPDEV
HLNPVRSKEKYNKPNISSTPLFKDSLYAERFIPVWVKGETFAIGFSEKDLFEIKPSNKEKLFTLLKTYSTKNP
GESLQELQAKSKAKWLYFPINKTLALEFLHHYFHKEIVTPDDTTVCHFINSLRYYTKKESITVKILKEPMPVL
SVKFESSKKNVLGSFKHTIALPATKDWERLFNHPNFLALKANPAPNPKEFNEFIRKYFLSDNNPNSDIPNNG
HNIKPQKHKAVRKVFSLPVIPGNAGTMMRIRRKDNKGQPLYQLQTIDDTPSMGIQINEDRLVKQEVLMDA
YKTRNLSTIDGINNSEGQAYATFDNWLTLPVSTFKPEIIKLEMKPHSKTRRYIRITQSLADFIKTIDEALMIKP
SDSIDDPLNMPNEIVCKNKLFGNELKPRDGKMKIVSTGKIVTYEFESDSTPQWIQTLYVTQLKKQP (SEQ ID
NO: 166)

[0122] In some embodiments the Cas9 protein can be *N. lactamica* Cas9 and may comprise or consist of the amino acid sequence:

MAAFKPNPMNYILGLDIGIASVGWAMVEVDEEENPIRLIDLGVRVFERAEVPKTGDSLAMARRLARSVRRL
TRRRAHRLLRARRLLKREGVLQDADFDENGLVKSLPNTPWQLRAAALDRKLTCLEWSAVLLHLVKHRGY
LSQRKNEGETADKELGALLKGVADNAHALQTGDFRTPAELALNKFEKESGHIRNQRGDYSHTFSRKDLQA

ELNLLFEKQKEFGNPHVSDGLKEDIETLLMAQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWL
TKLNNLRILEQGSERPLTDTERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEM
KAYHAISRALEKEGLKDKKSPLNLSTELQDEIGTAFSLFKTDKDITGRLKDRVQPEILEALLKHISFDKFVQIS
LKALRRIVPLMEQGKRYDEACAEIYGDHYCKKNAEEKIYLPPIPADEIRNPVVLRALSQARKVINCVVRRY
GSPARIHIETARAEVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCL
YSGKEINLVRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKA
RVETSRFPRSKKQRILLQKFDEEGFKERNLNDTRYVNRFLCQFVADHILLTGKGKRRVFASNGQITNLLRGF
WGLRKVRTENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGEVLHQKAHFPQPWE
FFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSGQGHMETVK
SAKRLDEGISVLRVPLTQLKLKGLEKMVNREREPKLYDALKAQLETHKDDPAKAFAEPFYKYDKAGSRTQ
QVKAVRIEQVQKTGVWVRNHNGIADNATMVRVDVFEKGGKYYLVPIYSWQVAKGILPDRAVVAFKDEE
DWTVMDDSFEFRFVLYANDLIKLTAKKNEFLGYFVSLNRATGAIDIRTHDTDSTKGKNGIFQSVGVKTALS
FQKNQIDELGKEIRPCRLKKRPPVR (SEQ ID NO: 167)

[0123] In some embodiments the Cas9 protein can be *N. meningitides* Cas9 and may comprise or consist of the amino acid sequence:

MAAFKPNPINYILGLDIGIASVGWAMVEIDEDENPICLIDLGVRVFERAEVPKTGDSLAMARRLARSVRRLT
RRRAHRLLRARRLLKREGVLQAADFDENGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLS
QRKNEGETADKELGALLKGVADNAHALQTGDFRTPAELALNKFEKESGHIRNQRGDYSHTFSRKDLQAEL
ILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKL
NNLRILEQGSERPLTDTERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAY
HAISRALEKEGLKDKKSPLNLSPELQDEIGTAFSLFKTDEDITGRLKDRIQPEILEALLKHISFDKFVQISLKAL
RRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVVLRALSQARKVINGVVRRYGSPAR
IHIETARAEVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCLYSGKE
INLGRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETS
RFPRSKKQRILLQKFDEDGFKERNLNDTRYVNRFLCQFVADRMRLTGKGKKRVFASNGQITNLLRGFWGL
RKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGEVLHQKTHFPQPWEFFAQ
EVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSGQGHMETVKSAKR
LDEGVSVLRVPLTQLKLKDLEKMVNREREPKLYEALKARLEAHKDDPAKAFAEPFYKYDKAGNRTQQVK
AVRVEQVQKTGVWVRNHNGIADNATMVRVDVFEKGDKYYLVPIYSWQVAKGILPDRAVVQGKDEEDW
QLIDDSFNFKFSLHPNDLVEVITKKARMFGYFASCHRGTGNINIRIHDLDHKIGKNGILEGIGVKTALSFQKY
QIDELGKEIRPCRLKKRPPVR (SEQ ID NO: 168)

[0124] In some embodiments the Cas9 protein can be *B. longum* Cas9 and may comprise or consist of the amino acid sequence:

MLSRQLLGASHLARPVSYSYNVQDNDVHCSYGERCFMRGKRYRIGIDVGLNSVGLAAVEVSDENSPVRLL
NAQSVIHDGGVDPQKNKEAITRKNMSGVARRTRRMRRRKRERLHKLDMLLGKFGYPVIEPESLDKPFEEW

HVRAELATRYIEDDELRRESISIALRHMARHRGWRNPYRQVDSLISDNPYSKQYGELKEKAKAYNDDATA
AEEESTPAQLVVAMLDAGYAEAPRLRWRTGSKKPDAEGYLPVRLMQEDNANELKQIFRVQRVPADEWKP
LFRSVFYAVSPKGSAEQRVGQDPLAPEQARALKASLAFQEYRIANVITNLRIKDASAELRKLTVDEKQSIYD
QLVSPSSEDITWSDLCDFLGFKRSQLKGVGSLTEDGEERISSRPPRLTSVQRIYESDNKIRKPLVAWWKSAS
DNEHEAMIRLLSNTVDIDKVREDVAYASAIEFIDGLDDDALTKLDSVDLPSGRAAYSVETLQKLTRQMLTT
DDDLHEARKTLFNVTDSWRPPADPIGEPLGNPSVDRVLKNVNRYLMNCQQRWGNPVSVNIEHVRSSFSSV
AFARKDKREYEKNNEKRSIFRSSLSEQLRADEQMEKVRESDLRRLEAIQRQNGQCLYCGRTITFRTCEMDH
IVPRKGVGSTNTRTNFAAVCAECNRMKSNTPFAIWARSEDAQTRGVSLAEAKKRVTMFTFNPKSYAPREV
KAFKQAVIARLQQTEDDAAIDNRSIESVAWMADELHRRIDWYFNAKQYVNSASIDDAEAETMKTTVSVFQ
GRVTASARRAAGIEGKIHFIGQQSKTRLDRRHHAVDASVIAMMNTAAAQTLMERESLRESQRLIGLMPGER
SWKEYPYEGTSRYESFHLWLDNMDVLLELLNDALDNDRIAVMQSQRYVLGNSIAHDATIHPLEKVPLGSA
MSADLIRRASTPALWCALTRLPDYDEKEGLPEDSHREIRVHDTRYSADDEMGFFASQAAQIAVQEGSADIG
SAIHHARVYRCWKTNAKGVRKYFYGMIRVFQTDLLRACHDDLFTVPLPPQSISMRYGEPRVVQALQSGNA
QYLGSLVVGDEIEMDFSSLDVDGQIGEYLQFFSQFSGGNLAWKHWVVDGFFNQTQLRIRPRYLAAEGLAK
AFSDDVVPDGVQKIVTKQGWLPPVNTASKTAVRIVRRNAFGEPRLSSAHHMPCSWQWRHE (SEQ ID NO:
169)

[0125] In some embodiments the Cas9 protein can be *A. muciniphila* Cas9 and may comprise or consist of the amino acid sequence:

MSRSLTFSFDIGYASIGWAVIASASHDDADPSVCGCGTVLFPKDDCQAFKRREYRRLRRNIRSRRVRIERIG
RLLVQAQIITPEMKETSGHPAPFYLASEALKGHRTLAPIELWHVLRWYAHNRGYDNNASWSNSLSEDGGN
GEDTERVKHAQDLMDKHGTATMAETICRELKLEEGKADAPMEVSTPAYKNLNTAFPRLIVEKEVRRILELS
APLIPGLTAEIIELIAQHHPLTTEQRGVLLQHGIKLARRYRGSLLFGQLIPRFDNRIISRCPVTWAQVYEAELK
KGNSEQSARERAEKLSKVPTANCPEFYEYRMARILCNIRADGEPLSAEIRRELMNQARQEGKLTKASLEKAI
SSRLGKETETNVSNYFTLHPDSEEALYLNPAVEVLQRSGIGQILSPSVYRIAANRLRRGKSVTPNYLLNLLKS
RGESGEALEKKIEKESKKKEADYADTPLKPKYATGRAPYARTVLKKVVEEILDGEDPTRPARGEAHPDGEL
KAHDGCLYCLLDTDSSVNQHQKERRLDTMTNNHLVRHRMLILDRLLKDLIQDFADGQKDRISRVCVEVG
KELTTFSAMDSKKIQRELTLRQKSHTDAVNRLKRKLPGKALSANLIRKCRIAMDMNWTCPFTGATYGDHE
LENLELEHIVPHSFRQSNALSSLVLTWPGVNRMKGQRTGYDFVEQEQENPVPDKPNLHICSLNNYRELVEK
LDDKKGHEDDRRRKKKRKALLMVRGLSHKHQSQNHEAMKEIGMTEGMMTQSSHLMKLACKSIKTSLPD
AHIDMIPGAVTAEVRKAWDVFGVFKELCPEAADPDSGKILKENLRSLTHLHHALDACVLGLIPYIIPAHHN
GLLRRVLAMRRIPEKLIPQVRPVANQRHYVLNDDGRMMLRDLSASLKENIREQLMEQRVIQHVPADMGG
ALLKETMQRVLSVDGSGEDAMVSLSKKKDGKKEKNQVKASKLVGVFPEGPSKLKALKAAIEIDGNYGVA
LDPKPVVIRHIKVFKRIMALKEQNGGKPVRILKKGMLIHLTSSKDPKHAGVWRIESIQDSKGGVKLDLQRA
HCAVPKNKTHECNWREVDLISLLKKYQMKRYPTSYTGTPR (SEQ ID NO: 170)

[0126] In some embodiments the Cas9 protein can be *O. laneus* Cas9 and may comprise or consist of the amino acid sequence:

METTLGIDLGTNSIGLALVDQEEHQILYSGVRIFPEGINKDTIGLGEKEESRNATRRAKRQMRRQYFRKKLR
KAKLLELLIAYDMCPLKPEDVRRWKNWDKQQKSTVRQFPDTPAFREWLKQNPYELRKQAVTEDVTRPEL
GRILYQMIQRRGFLSSRKGKEEGKIFTGKDRMVGIDETRKNLQKQTLGAYLYDIAPKNGEKYRFRTERVRA
RYTLRDMYIREFEIIWQRQAGHLGLAHEQATRKKNIFLEGSATNVRNSKLITHLQAKYGRGHVLIEDTRITV
TFQLPLKEVLGGKIEIEEEQLKFKSNESVLFWQRPLRSQKSLLSKCVFEGRNFYDPVHQKWIIAGPTPAPLSH
PEFEEFRAYQFINNIIYGKNEHLTAIQREAVFELMCTESKDFNFEKIPKHLKLFEKFNFDDTTKVPACTTISQL
RKLFPHPVWEEKREEIWHCFYFYDDNTLLFEKLQKDYALQTNDLEKIKKIRLSESYGNVSLKAIRRINPYLK
KGYAYSTAVLLGGIRNSFGKRFEYFKEYEPEIEKAVCRILKEKNAEGEVIRKIKDYLVHNRFGFAKNDRAFQ
KLYHHSQAITTQAQKERLPETGNLRNPIVQQGLNELRRTVNKLLATCREKYGPSFKFDHIHVEMGRELRSS
KTEREKQSRQIRENEKKNEAAKVKLAEYGLKAYRDNIQKYLLYKEIEEKGGTVCCPYTGKTLNISHTLGSD
NSVQIEHIIPYSISLDDSLANKTLCDATFNREKGELTPYDFYQKDPSPEKWGASSWEEIEDRAFRLLPYAKAQ
RFIRRKPQESNEFISRQLNDTRYISKKAVEYLSAICSDVKAFPGQLTAELRHLWGLNNILQSAPDITFPLPVSA
TENHREYYVITNEQNEVIRLFPKQGETPRTEKGELLLTGEVERKVFRCKGMQEFQTDVSDGKYWRRIKLSS
SVTWSPLFAPKPISADGQIVLKGRIEKGVFVCNQLKQKLKTGLPDGSYWISLPVISQTFKEGESVNNSKLTSQ
QVQLFGRVREGIFRCHNYQCPASGADGNFWCTLDTDTAQPAFTPIKNAPPGVGGGQIILTGDVDDKGIFHA
DDDLHYELPASLPKGKYYGIFTVESCDPTLIPIELSAPKTSKGENLIEGNIWVDEHTGEVRFDPKKNREDQR
HHAIDAIVIALSSQSLFQRLSTYNARRENKKRGLDSTEHFPSPWPGFAQDVRQSVVPLLVSYKQNPKTLCKI
SKTLYKDGKKIHSCGNAVRGQLHKETVYGQRTAPGATEKSYHIRKDIRELKTSKHIGKVVDITIRQMLLKH
LQENYHIDITQEFNIPSNAFFKEGVYRIFLPNKHGEPVPIKKIRMKEELGNAERLKDNINQYVNPRNNHHVMI
YQDADGNLKEEIVSFWSVIERQNQGQPIYQLPREGRNIVSILQINDTFLIGLKEEEPEVYRNDLSTLSKHLYR
VQKLSGMYYTFRHHLASTLNNEREEFRIQSLEAWKRANPVKVQIDEIGRITFLNGPLC (SEQ ID NO: 171).

[0127]

[0128] In some embodiments of the compositions of the disclosure, the sequence encoding the first RNA binding protein comprises a sequence isolated or derived from a CRISPR Cas protein or portion thereof. In some embodiments, the CRISPR Cas protein comprises a Type V CRISPR Cas protein. In some embodiments, the Type V CRISPR Cas protein comprises a Cpf1 protein. Exemplary Cpf1 proteins of the disclosure may be isolated or derived from any species, including, but not limited to, a bacteria or an archaea. Exemplary Cpf1 proteins of the disclosure may be isolated or derived from any species, including, but not limited to, *Francisella tularensis subsp. novicida, Acidaminococcus sp. BV3L6* and *Lachnospiraceae bacterium sp. ND2006*. Exemplary Cpf1 proteins of the disclosure may be nuclease inactivated.

[0129] Exemplary wild type *Francisella tularensis subsp. Novicida* Cpf1 (FnCpf1) proteins of the disclosure may comprise or consist of the amino acid sequence:

```
   1 MSIYQEFVNK YSLSKTLRFE LIPQGKTLEN IKARGLILDD EKRAKDYKKA KQIIDKYHQF
  61 FIEEILSSVC ISEDLLQNYS DVYFKLKKSD DDNLQKDFKS AKDTIKKQIS EYIKDSEKFK
 121 NLFNQNLIDA KKGQESDLIL WLKQSKDNGI ELFKANSDIT DIDEALEIIK SFKGWTTYFK
 181 GFHENRKNVY SSNDIPTSII YRIVDDNLPK FLENKAKYES LKDKAPEAIN YEQIKKDLAE
 241 ELTFDIDYKT SEVNQRVFSL DEVFEIANFN NYLNQSGITK FNTIIGGKFV NGENTKRKGI
 301 NEYINLYSQQ INDKTLKKYK MSVLFKQILS DTESKSFVID KLEDDSDVVT TMQSFYEQIA
 361 AFKTVEEKSI KETLSLLFDD LKAQKLDLSK IYFKNDKSLT DLSQQVFDDY SVIGTAVLEY
 421 ITQQIAPKNL DNPSKKEQEL IAKKTEKAKY LSLETIKLAL EEFNKHRDID KQCRFEEILA
 481 NFAAIPMIFD EIAQNKDNLA QISIKYQNQG KKDLLQASAE DDVKAIKDLL DQTNNLLHKL
 541 KIFHISQSED KANILDKDEH FYLVFEECYF ELANIVPLYN KIRNYITQKP YSDEKFKLNF
 601 ENSTLANGWD KNKEPDNTAI LFIKDDKYYL GVMNKKNNKI FDDKAIKENK GEGYKKIVYK
 661 LLPGANKMLP KVFFSAKSIK FYNPSEDILR IRNHSTHTKN GSPQKGYEKF EFNIEDCRKF
 721 IDFYKQSISK HPEWKDFGFR FSDTQRYNSI DEFYREVENQ GYKLTFENIS ESYIDSVVNQ
 781 GKLYLFQIYN KDFSAYSKGR PNLHTLYWKA LFDERNLQDV VYKLNGEAEL FYRKQSIPKK
 841 ITHPAKEAIA NKNKDNPKKE SVFEYDLIKD KRFTEDKFFF HCPITINFKS SGANKFNDEI
 901 NLLLKEKAND VHILSIDRGE RHLAYYTLVD GKGNIIKQDT FNIIGNDRMK TNYHDKLAAI
 961 EKDRDSARKD WKKINNIKEM KEGYLSQVVH EIAKLVIEYN AIVVFEDLNF GFKRGRFKVE
1021 KQVYQKLEKM LIEKLNYLVF KDNEFDKTGG VLRAYQLTAP FETFKKMGKQ TGIIYYVPAG
1081 FTSKICPVTG FVNQLYPKYE SVSKSQEFFS KFDKICYNLD KGYFEFSFDY KNFGDKAAKG
1141 KWTIASFGSR LINFRNSDKN HNWDTREVYP TKELEKLLKD YSIEYGHGEC IKAAICGESD
1201 KKFFAKLTSV LNTILQMRNS KTGTELDYLI SPVADVNGNF FDSRQAPKNM PQDADANGAY
1261 HIGLKGLMLL GRIKNNQEGK KLNLVIKNEE YFEFVQNRNN (SEQ ID NO: 172).
```

[0130] Exemplary wild type *Lachnospiraceae bacterium sp. ND2006* Cpf1 (LbCpf1) proteins of the disclosure may comprise or consist of the amino acid sequence:

```
   1 AASKLEKFTN CYSLSKTLRF KAIPVGKTQE NIDNKRLLVE DEKRAEDYKG VKKLLDRYYL
  61 SFINDVLHSI KLKNLNNYIS LFRKKTRTEK ENKELENLEI NLRKEIAKAF KGAAGYKSLF
 121 KKDIIETILP EAADDKDEIA LVNSFNGFTT AFTGFFDNRE NMFSEEAKST SIAFRCINEN
 181 LTRYISNMDI FEKVDAIFDK HEVQEIKEKI LNSDYDVEDF FEGEFFNFVL TQEGIDVYNA
 241 IIGGFVTESG EKIKGLNEYI NLYNAKTKQA LPKFKPLYKQ VLSDRESLSF YGEGYTSDEE
 301 VLEVFRNTLN KNSEIFSSIK KLEKLFKNFD EYSSAGIFVK NGPAISTISK DIFGEWNLIR
 361 DKWNAEYDDI HLKKKAVVTE KYEDDRRKSF KKIGSFSLEQ LQEYADADLS VVEKLKEIII
 421 QKVDEIYKVY GSSEKLFDAD FVLEKSLKKN DAVVAIMKDL LDSVKSFENY IKAFFGEGKE
 481 TNRDESFYGD FVLAYDILLK VDHIYDAIRN YVTQKPYSKD KFKLYFQNPQ FMGGWDKDKE
 541 TDYRATILRY GSKYYLAIMD KKYAKCLQKI DKDDVNGNYE KINYKLLPGP NKMLPKVFFS
 601 KKWMAYYNPS EDIQKIYKNG TFKKGDMFNL NDCHKLIDFF KDSISRYPKW SNAYDFNFSE
 661 TEKYKDIAGF YREVEEQGYK VSFESASKKE VDKLVEEGKL YMFQIYNKDF SDKSHGTPNL
 721 HTMYFKLLFD ENNHGQIRLS GGAELFMRRA SLKKEELVVH PANSPIANKN PDNPKKTTTL
 781 SYDVYKDKRF SEDQYELHIP IAINKCPKNI FKINTEVRVL LKHDDNPYVI GIDRGERNLL
 841 YIVVVDGKGN IVEQYSLNEI INNFNGIRIK TDYHSLLDKK EKERFEARQN WTSIENIKEL
 901 KAGYISQVVH KICELVEKYD AVIALEDLNS GFKNSRVKVE KQVYQKFEKM LIDKLNYMVD
 961 KKSNPCATGG ALKGYQITNK FESFKSMSTQ NGFIFYIPAW LTSKIDPSTG FVNLLKTKYT
1021 SIADSKKFIS SFDRIMYVPE EDLFEFALDY KNFSRTDADY IKKWKLYSYG NRIRIFAAAK
1081 KNNVFAWEEV CLTSAYKELF NKYGINYQQG DIRALLCEQS DKAFYSSFMA LMSLMLQMRN
1141 SITGRTDVDF LISPVKNSDG IFYDSRNYEA QENAILPKNA DANGAYNIAR KVLWAIGQFK
1201 KAEDEKLDKV KIAISNKEWL EYAQTSVK (SEQ ID NO: 173).
```

[0131] Exemplary wild type *Acidaminococcus sp. BV3L6* Cpf1 (AsCpf1) proteins of the disclosure may comprise or consist of the amino acid sequence:

```
   1 MTQFEGFTNL YQVSKTLRFE LIPQGKTLKH IQEQGFIEED KARNDHYKEL KPIIDRIYKT
  61 YADQCLQLVQ LDWENLSAAI DSYRKEKTEE TRNALIEEQA TYRNAIHDYF IGRTDNLTDA
 121 INKRHAEIYK GLFKAELFNG KVLKQLGTVT TTEHENALLR SFDKFTTYFS GFYENRKNVF
 181 SAEDISTAIP HRIVQDNFPK FKENCHIFTR LITAVPSLRE HFENVKKAIG IFVSTSIEEV
 241 FSFPFYNQLL TQTQIDLYNQ LLGGISREAG TEKIKGLNEV LNLAIQKNDE TAHIIASLPH
 301 RFIPLFKQIL SDRNTLSFIL EEFKSDEEVI QSFCKYKTLL RNENVLETAE ALFNELNSID
 361 LTHIFISHKK LETISSALCD HWDTLRNALY ERRISELTGK ITKSAKEKVQ RSLKHEDINL
 421 QEIISAAGKE LSEAFKQKTS EILSHAHAAL DQPLPTTLKK QEEKEILKSQ LDSLLGLYHL
 481 LDWFAVDESN EVDPEFSARL TGIKLEMEPS LSFYNKARNY ATKKPYSVEK FKLNFQMPTL
 541 ASGWDVNKEK NNGAILFVKN GLYYLGIMPK QKGRYKALSF EPTEKTSEGF DKMYYDYFPD
 601 AAKMIPKCST QLKAVTAHFQ THTTPILLSN NFIEPLEITK EIYDLNNPEK EPKKFQTAYA
 661 KKTGDQKGYR EALCKWIDFT RDFLSKYTKT TSIDLSSLRP SSQYKDLGEY YAELNPLLYH
 721 ISFQRIAEKE IMDAVETGKL YLFQIYNKDF AKGHHGKPNL HTLYWTGLFS PENLAKTSIK
 781 LNGQAELFYR PKSRMKRMAH RLGEKMLNKK LKDQKTPIPD TLYQELYDYV NHRLSHDLSD
 841 EARALLPNVI TKEVSHEIIK DRRFTSDKFF FHVPITLNYQ AANSPSKFNQ RVNAYLKEHP
 901 ETPIIGIDRG ERNLIYITVI DSTGKILEQR SLNTIQQFDY QKKLDNREKE RVAARQAWSV
 961 VGTIKDLKQG YLSQVIHEIV DLMIHYQAVV VLENLNFGFK SKRTGIAEKA VYQQFEKMLI
1021 DKLNCLVLKD YPAEKVGGVL NPYQLTDQFT SFAKMGTQSG FLFYVPAPYT SKIDPLTGFV
1081 DPFVWKTIKN HESRKHFLEG FDFLHYDVKT GDFILHFKMN RNLSFQRGLP GFMPAWDIVF
1141 EKNETQFDAK GTPFIAGKRI VPVIENHRFT GRYRDLYPAN ELIALLEEKG IVFRDGSNIL
1201 PKLLENDDSH AIDTMVALIR SVLQMRNSNA ATGEDYINSP VRDLNGVCFD SRFQNPEWPM
1261 DADANGAYHI ALKGQLLLNH LKESKDLKLQ NGISNQDWLA YIQELRN (SEQ ID NO:
174).
```

[0132] In some embodiments of the compositions of the disclosure, the sequence encoding the first RNA binding protein comprises a sequence isolated or derived from a CRISPR Cas protein. In some embodiments, the CRISPR Cas protein comprises a Type VI CRISPR Cas protein or portion thereof. In some embodiments, the Type VI CRISPR Cas protein comprises a Cas13 protein or portion thereof. Exemplary Cas13 proteins of the disclosure may be isolated or derived from any species, including, but not limited to, a bacteria or an archaea. Exemplary Cas13 proteins of the disclosure may be isolated or derived from any species, including, but not limited to, *Leptotrichia wadei, Listeria seeligeri serovar 1/2b (strain ATCC 35967 / DSM 20751 / CIP 100100 / SLCC 3954), Lachnospiraceae bacterium, Clostridium aminophilum DSM 10710, Carnobacterium gallinarum DSM 4847, Paludibacter propionicigenes WB4, Listeria weihenstephanensis FSL R9-0317, Listeria weihenstephanensis FSL R9-0317, bacterium FSL M6-0635 (Listeria newyorkensis), Leptotrichia wadei F0279, Rhodobacter capsulatus SB 1003, Rhodobacter capsulatus R121, Rhodobacter capsulatus DE442* and *Corynebacterium ulcerans.* Exemplary Cas13 proteins of the disclosure may be DNA nuclease inactivated. Exemplary Cas13 proteins of the disclosure include, but are not limited to, Cas13a, Cas13b, Cas13c, Cas13d

and orthologs thereof. Exemplary Cas13b proteins of the disclosure include, but are not limited to, subtypes 1 and 2 referred to herein as Csx27 and Csx28, respectively.

[0133] Exemplary Cas13a proteins include, but are not limited to:

Cas13a number	**Cas13a abbreviation**	**Organism name**	**Accession number**	**Direct Repeat sequence**
Cas13a1	LshCas13a	Leptotrichia shahii	WP_018451595.1	CCACCCCAATATCGAAGGGGACTAAAAC (SEQ ID NO: 175)
Cas13a2	LwaCas13a	Leptotrichia wadei	WP_021746774.1	GATTTAGACTACCCCAAAAACGAAGGGGACTAAAAC (SEQ ID NO: 176)
Cas13a3	LseCas13a	Listeria seeligeri	WP_012985477.1	GTAAGAGACTACCTCTATATGAAAGAGGACTAAAAC (SEQ ID NO: 177)
Cas13a4	LbmCas13a	Lachnospiraceae bacterium MA2020	WP_044921188.1	GTATTGAGAAAAGCCAGATATAGTTGGCAATAGAC (SEQ ID NO: 178)
Cas13a5	LbnCas13a	Lachnospiraceae bacterium NK4A179	WP_022785443.1	GTTGATGAGAAGAGCCCAAGATAGAGGGCAATAAC (SEQ ID NO: 179)
Cas13a6	CamCas13a	[Clostridium] aminophilum DSM 10710	WP_031473346.1	GTCTATTGCCCTCTATATCGGGCTGTTCTCCAAAC (SEQ ID NO: 180)
Cas13a7	CgaCas13a	Carnobacterium gallinarum DSM 4847	WP_034560163.1	ATTAAAGACTACCTCTAAATGTAAGAGGACTATAAC (SEQ ID NO: 181)
Cas13a8	Cga2Cas13a	Carnobacterium gallinarum DSM 4847	WP_034563842.1	AATATAAACTACCTCTAAATGTAAGAGGACTATAAC (SEQ ID NO: 182)
Cas13a9	Pprcas13a	Paludibacter propionicigenes WB4	WP_013443710.1	CTTGTGGATTATCCCAAAATTGAAGGGAACTACAAC (SEQ ID NO: 183)
Cas13a10	LweCas13a	Listeria weihenstephanensis FSL R9-0317	WP_036059185.1	GATTTAGAGTACCTCAAAATAGAAGAGGTCTAAAAC (SEQ ID NO: 184)
Cas13a11	LbfCas13a	Listeriaceae bacterium FSL M6-0635 (Listeria newyorkensis)	WP_036091002.1	GATTTAGAGTACCTCAAAACAAAAGAGGACTAAAAC (SEQ ID NO: 185)

Cas13a12	Lwa2cas13a	Leptotrichia wadei F0279	WP_021746774.1	GATATAGATAACCCCAAAAACGAAGGGATCTAAAAC (SEQ ID NO: 186)
Cas13a13	RcsCas13a	Rhodobacter capsulatus SB 1003	WP_013067728.1	GCCTCACATCACCGCCAAGACGACGGCGGACTGAAC (SEQ ID NO: 187)
Cas13a14	RcrCas13a	Rhodobacter capsulatus R121	WP_023911507.1	GCCTCACATCACCGCCAAGACGACGGCGGACTGAAC (SEQ ID NO: 188)
Cas13a15	RcdCas13a	Rhodobacter capsulatus DE442	WP_023911507.1	GCCTCACATCACCGCCAAGACGACGGCGGACTGAAC (SEQ ID NO: 189)

[0134] Exemplary wild type Cas13a proteins of the disclosure may comprise or consist of the amino acid sequence:

```
   1 MGNLFGHKRW YEVRDKKDFK IKRKVKVKRN YDGNKYILNI NENNNKEKID NNKFIRKYIN
  61 YKKNDNILKE FTRKFHAGNI LFKLKGKEGI IRIENNDDFL ETEEVVLYIE AYGKSEKLKA
 121 LGITKKKIID EAIRQGITKD DKKIEIKRQE NEEEIEIDIR DEYTNKTLND CSIILRIIEN
 181 DELETKKSIY EIFKNINMSL YKIIEKIIEN ETEKVFENRY YEEHLREKLL KDDKIDVILT
 241 NFMEIREKIK SNLEILGFVK FYLNVGGDKK KSKNKKMLVE KILNINVDLT VEDIADFVIK
 301 ELEFWNITKR IEKVKKVNNE FLEKRRNRTY IKSYVLLDKH EKFKIERENK KDKIVKFFVE
 361 NIKNNSIKEK IEKILAEFKI DELIKKLEKE LKKGNCDTEI FGIFKKHYKV NFDSKKFSKK
 421 SDEEKELYKI IYRYLKGRIE KILVNEQKVR LKKMEKIEIE KILNESILSE KILKRVKQYT
 481 LEHIMYLGKL RHNDIDMTTV NTDDFSRLHA KEELDLELIT FFASTNMELN KIFSRENINN
 541 DENIDFFGGD REKNYVLDKK ILNSKIKIIR DLDFIDNKNN ITNNFIRKFT KIGTNERNRI
 601 LHAISKERDL QGTQDDYNKV INIIQNLKIS DEEVSKALNL DVVFKDKKNI ITKINDIKIS
 661 EENNNDIKYL PSFSKVLPEI LNLYRNNPKN EPFDTIETEK IVLNALIYVN KELYKKLILE
 721 DDLEENESKN IFLQELKKTL GNIDEIDENI IENYYKNAQI SASKGNNKAI KKYQKKVIEC
 781 YIGYLRKNYE ELFDFSDFKM NIQEIKKQIK DINDNKTYER ITVKTSDKTI VINDDFEYII
 841 SIFALLNSNA VINKIRNRFF ATSVWLNTSE YQNIIDILDE IMQLNTLRNE CITENWNLNL
 901 EEFIQKMKEI EKDFDDFKIQ TKKEIFNNYY EDIKNNILTE FKDDINGCDV LEKKLEKIVI
 961 FDDETKFEID KKSNILQDEQ RKLSNINKKD LKKKVDQYIK DKDQEIKSKI LCRIIFNSDF
1021 LKKYKKEIDN LIEDMESENE NKFQEIYYPK ERKNELYIYK KNLFLNIGNP NFDKIYGLIS
1081 NDIKMADAKF LFNIDGKNIR KNKISEIDAI LKNLNDKLNG YSKEYKEKYI KKLKENDDFF
1141 AKNIQNKNYK SFEKDYNRVS EYKKIRDLVE FNYLNKIESY LIDINWKLAI QMARFERDMH
1201 YIVNGLRELG IIKLSGYNTG ISRAYPKRNG SDGFYTTTAY YKFFDEESYK KFEKICYGFG
1261 IDLSENSEIN KPENESIRNY ISHFYIVRNP FADYSIAEQI DRVSNLLSYS TRYNNSTYAS
1321 VFEVFKKDVN LDYDELKKKF KLIGNNDILE RLMKPKKVSV LELESYNSDY IKNLIIELLT
1381 KIENTNDTL (SEQ ID NO: 190).
```

[0135] Exemplary Cas13b proteins include, but are not limited to:

Species	Cas13b Accession	Cas13b Size (aa)
Paludibacter propionicigenes WB4	WP_013446107.1	1155
Prevotella sp. P5-60	WP_044074780.1	1091
Prevotella sp. P4-76	WP_044072147.1	1091
Prevotella sp. P5-125	WP_044065294.1	1091
Prevotella sp. P5-119	WP_042518169.1	1091
Capnocytophaga canimorsus Cc5	WP_013997271.1	1200
Phaeodactylibacter xiamenensis	WP_044218239.1	1132

Porphyromonas gingivalis W83	WP_005873511.1	1136
Porphyromonas gingivalis F0570	WP_021665475.1	1136
Porphyromonas gingivalis ATCC 33277	WP_012458151.1	1136
Porphyromonas gingivalis F0185	ERJ81987.1	1136
Porphyromonas gingivalis F0185	WP_021677657.1	1136
Porphyromonas gingivalis SJD2	WP_023846767.1	1136
Porphyromonas gingivalis F0568	ERJ65637.1	1136
Porphyromonas gingivalis W4087	ERJ87335.1	1136
Porphyromonas gingivalis W4087	WP_021680012.1	1136
Porphyromonas gingivalis F0568	WP_021663197.1	1136
Porphyromonas gingivalis	WP_061156637.1	1136
Porphyromonas gulae	WP_039445055.1	1136
Bacteroides pyogenes F0041	ERI81700.1	1116
Bacteroides pyogenes JCM 10003	WP_034542281.1	1116
Alistipes sp. ZOR0009	WP_047447901.1	954
Flavobacterium branchiophilum FL-15	WP_014084666.1	1151
Prevotella sp. MA2016	WP_036929175.1	1323
Myroides odoratimimus CCUG 10230	EHO06562.1	1160
Myroides odoratimimus CCUG 3837	EKB06014.1	1158
Myroides odoratimimus CCUG 3837	WP_006265509.1	1158
Myroides odoratimimus CCUG 12901	WP_006261414.1	1158
Myroides odoratimimus CCUG 12901	EHO08761.1	1158
Myroides odoratimimus (NZ_CP013690.1)	WP_058700060.1	1160
Bergeyella zoohelcum ATCC 43767	EKB54193.1	1225
Capnocytophaga cynodegmi	WP_041989581.1	1219
Bergeyella zoohelcum ATCC 43767	WP_002664492.1	1225
Flavobacterium sp. 316	WP_045968377.1	1156
Psychroflexus torquis ATCC 700755	WP_015024765.1	1146
Flavobacterium columnare ATCC 49512	WP_014165541.1	1180
Flavobacterium columnare	WP_060381855.1	1214
Flavobacterium columnare	WP_063744070.1	1214
Flavobacterium columnare	WP_065213424.1	1215
Chryseobacterium sp. YR477	WP_047431796.1	1146
Riemerella anatipestifer ATCC 11845 = DSM 15868	WP_004919755.1	1096
Riemerella anatipestifer RA-CH-2	WP_015345620.1	949
Riemerella anatipestifer	WP_049354263.1	949
Riemerella anatipestifer	WP_061710138.1	951
Riemerella anatipestifer	WP_064970887.1	1096
Prevotella saccharolytica F0055	EKY00089.1	1151
Prevotella saccharolytica JCM 17484	WP_051522484.1	1152
Prevotella buccae ATCC 33574	EFU31981.1	1128
Prevotella buccae ATCC 33574	WP_004343973.1	1128
Prevotella buccae D17	WP_004343581.1	1128
Prevotella sp. MSX73	WP_007412163.1	1128
Prevotella pallens ATCC 700821	EGQ18444.1	1126
Prevotella pallens ATCC 700821	WP_006044833.1	1126

Prevotella intermedia ATCC 25611 = DSM 20706	WP_036860899.1	1127
Prevotella intermedia	WP_061868553.1	1121
Prevotella intermedia 17	AFJ07523.1	1135
Prevotella intermedia	WP_050955369.1	1133
Prevotella intermedia	BAU18623.1	1134
Prevotella intermedia ZT	KJJ86756.1	1126
Prevotella aurantiaca JCM 15754	WP_025000926.1	1125
Prevotella pleuritidis F0068	WP_021584635.1	1140
Prevotella pleuritidis JCM 14110	WP_036931485.1	1117
Prevotella falsenii DSM 22864 = JCM 15124	WP_036884929.1	1134
Porphyromonas gulae	WP_039418912.1	1176
Porphyromonas sp. COT-052 OH4946	WP_039428968.1	1176
Porphyromonas gulae	WP_039442171.1	1175
Porphyromonas gulae	WP_039431778.1	1176
Porphyromonas gulae	WP_046201018.1	1176
Porphyromonas gulae	WP_039434803.1	1176
Porphyromonas gulae	WP_039419792.1	1120
Porphyromonas gulae	WP_039426176.1	1120
Porphyromonas gulae	WP_039437199.1	1120
Porphyromonas gingivalis TDC60	WP_013816155.1	1120
Porphyromonas gingivalis ATCC 33277	WP_012458414.1	1120
Porphyromonas gingivalis A7A1-28	WP_058019250.1	1176
Porphyromonas gingivalis JCVI SC001	EOA10535.1	1176
Porphyromonas gingivalis W50	WP_005874195.1	1176
Porphyromonas gingivalis	WP_052912312.1	1176
Porphyromonas gingivalis AJW4	WP_053444417.1	1120
Porphyromonas gingivalis	WP_039417390.1	1120
Porphyromonas gingivalis	WP_061156470.1	1120

[0136] Exemplary wild type *Bergeyella zoohelcum* ATCC 43767 Cas13b (BzCas13b) proteins of the disclosure may comprise or consist of the amino acid sequence:

```
  1 menktslgnn iyynpfkpqd ksyfagyfna amentdsvfr elgkrlkgke ytsenffdai
 61 fkenislvey eryvkllsdy fpmarlldkk evpikerken fkknfkgiik avrdlrnfyt
121 hkehgeveit deifgvldem lkstvltvkk kkvktdktke ilkksiekql dilcqkkley
181 lrdtarkiee krrnqrerge kelvapfkys dkrddliaai yndafdvyid kkkdslkess
241 kakyntksdp qqeegdlkip iskngvvfll slfltkqeih afkskiagfk atvideatvs
301 eatvshgkns icfmatheif shlaykklkr kvrtaeinyg eaenaeqlsv yaketlmmqm
361 ldelskvpdv vyqnlsedvq ktfiedwney lkenngdvgt meeeqvihpv irkryedkfn
421 yfairfldef aqfptlrfqv hlgnylhdsr pkenlisdrr ikekitvfgr lselehkkal
481 fikntetned rehyweifpn pnydfpkeni svndkdfpia gsildrekqp vagkigikvk
541 llnqqyvsev dkavkahqlk qrkaskpsiq niieeivpin esnpkeaivf ggqptaylsm
601 ndihsilyef fdkwekkkek lekkgekelr keigkelekk ivgkiqaqiq qiidkdtnak
661 ilkpyqdgns taidkeklik dlkqeqnilq klkdeqtvre keyndfiayq dknreinkvr
721 drnhkqylkd nlkrkypeap arkevlyyre kgkvavwlan dikrfmptdf knewkgeqhs
781 llqkslayye qckeelknll pekvfqhlpf klggyfqqky lyqfytcyld krleyisglv
841 qqaenfksen kvfkkvenec fkflkkqnyt hkeldarvqs ilgypifler gfmdekptii
901 kgktfkgnea lfadwfryyk eyqnfqtfyd tenyplvele kkqadrkrkt kiyqqkkndv
961 ftllmakhif ksvfkqdsid qfsledlyqs reerlgnqer arqtgerntn yiwnktvdlk
```

```
1021 lcdgkitven vklknvgdfi kyeydqrvqa flkyeeniew qaflikeske eenypyvver
1081 eieqyekvrr eellkevhli eeyilekvkd keilkkgdnq nfkyyilngl lkqlknedve
1141 sykvfnlnte pedvninqlk qeatdleqka fvltyirnkf ahnqlpkkef wdycqekygk
1201      ektyaey faevfkkeke alik (SEQ ID NO: 191).
```

[0137] In some embodiments of the compositions of the disclosure, the sequence encoding the first RNA binding protein comprises a sequence isolated or derived from a CasRX/Cas13d protein. CasRX/Cas13d is an effector of the type VI-D CRISPR-Cas systems. In some embodiments, the CasRX/Cas13d protein is an RNA-guided RNA endonuclease enzyme that can cut or bind RNA. In some embodiments, the CasRX/Cas13d protein can include one or more higher eukaryotes and prokaryotes nucleotide-binding (HEPN) domains. In some embodiments, the CasRX/Cas13d protein can include either a wild-type or mutated HEPN domain. In some embodiments, the CasRX/Cas13d protein includes a mutated HEPN domain that cannot cut RNA but can process guide RNA. In some embodiments, the CasRX/Cas13d protein does not require a protospacer flanking sequence. Also see WO Publication No. WO2019/040664 & US2019/0062724, which is incorporated herein by reference in its entirety, for further examples and sequences of CasRX/Cas13d protein, without limitation, specific reference is made to

[0138] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence:

CasRX/Cas13d Gut_metagenome_contig6049000251:

```
LYLTSFGKGN AAVIEQKIEP ENGYRVTGMQ ITPSITVNKA TDESVRFRVK RKIAQKDEFI        60
ADNPMHEGRH RIEPSAGSDM LGLKTKLEKY YFGKEFDDNL HIQIIYNILD IEKILAVYST       120
NITA                                                                     124
```

(SEQ ID NO: 54).

[0139] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence:

CasRX/Cas13d Gut_metagenome_contig546000275:

```
MDSYRPKLYK LIDFCIFKHY HEYTEISEKN VDTLRAAVSE EQKESFYADE AKRLWGIFDK        60
QFLGFCKKIN VWVNGSHEKE ILGYIDKDAY RKKSDVSYFS KFLYAMSFFL DGKEINDLLT       120
TLINKFDNIA SFISTAKELD AEIDRILEKK LDPVTGKPLK GKNSFRNFIA NNVIENKRFI       180
YVIKFCNPKN VLKLVKNTKV TEFVLKRMPE SQIDRYYSSC IDTEKNPSVD KKISDLAEMI       240
KKIAFDDFRN VRQKTRTREE SLEKERFKAV IGLYLTVVYL LIKNLVNVNS RYVMAFHCLE       300
RDAKLYGINI GKNYIELTED LCRENENSRS AYLARNKRLR DCVKQNIDNA KNMKSKEK          358
```

(SEQ ID NO: 57).

[0140] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence:

CasRX/Cas13d Gut_metagenome_contig4114000374:

```
DTKINPQTWL YQLENTPDLD NEYRDTLDHF FDERFNEINE HFVTQNATNL CIMKEVFPDE        60
DFKSIADLYY DFIVVKSYKN IGFSIKKLRE KMLELPEAKR VTSTEMDSVR SKLYKLIDFC       120
```

```
IFKHYHEKPE TVEMIVSMLR AYTSEDMKE                                    149
```

(SEQ ID NO: 61).

[0141] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Gut_metagenome_contig721000619:

```
KEGSTMAKNE KKKSTAKALG LKSSFVVNND IYMTSFGKGN KAVLEKKITE NTIENKSDTT    60
YFDVINRDPK GFTLEGRRIA DMTAFSNDPK YHVNVVNGKF LEDQLGARSE LEKKVFGRTF   120
DDNVHIQLIH NILDIEKIMA QYVSDIVYLL HNTIKRDMND DIMGYISIRN SFDDFCHPER   180
IPDRKAKDNL QKQHDIFFDE ILKCGRLAYF GNAFFEDGSD NKEIAKLKRY KEIYHIIALM   240
GSLRQSYFHG ENSDKNFQGP TWAYTLESNL TGKYKEFKDT LDKTFDERYE MISKDFGSTN   300
MVNLQILEEL LKMLYGNVSP                                              320
```

(SEQ ID NO: 67).

[0142] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Gut_metagenome_contig2002000411:

```
EKQNKAKYQA IISLYLMVMY QIVKNMIYVN SRYVIAFHCL ERDSNQLLGR FNSRDASMYN    60
KLTQKFITDK YLNDGAQGCS KKVGNYLSHN ITCCSDELRK EYRNQVDHFA VVRMIGKYAA   120
DIGKFSTWFE LYHYVMQRII FDKRNPLSET ERTYKQLIAK HHTYCKDLVK ALNTPFGYNL   180
ARYKNLSIGE LFDRNNYNAK TKET                                         204
```

(SEQ ID NO: 69).

[0143] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Gut_metagenome_contig13552000311:

```
LIDFLIYDLY YNRKPARIEE IVDKLRESVN DEEKESIYSA ETKYVYEALG KVLVRSLKKY    60
LNGATIRDLK NRYDAKTANR IWDISEHSKS GHVNCFCKLI YMMTLMLDGK EINDLLTTLV   120
NKFDNIASFI DVMDELGLEH SFTDNYKMFA DSKAICLDLQ FINSFARMSK IDDEKSKRQL   180
FRDALVVLDI GDKNEDWIEK YLTSDIFKRD ENGNKIDGEK RDFRNFIANN VIKSARFKYL   240
VKYSSADGMI KLKKNEKLIS FVLEQLPETQ IDRYYESCGL DCAVADRKVR IEKLTGLIRD   300
MRFDNFRGVN YSNDACKKDK QAKAKYQAII SLYLMVLYQI VKNMIYVNSR YVIAFHCLER   360
DLLFFNIELD NSYQYSNCNE LTEKFIKDKY MKEGALGFNM KAGRYLTKNI GNCSNELRKI   420
YRNQVDHFAV VRKIGNYAAD IASVGSWFE                                     449
```

(SEQ ID NO: 71).

[0144] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Gut_metagenome_contig10037000527:

```
YMDQNFANSD AWAIHVYRNK IQHLDAVRHA DMYIGDIREF HSWFELYHYI IQRRIIDQYA    60
YESTPGSSRD GSAIIDEERL NPATRRYFRL ITTYKT                             96
```

(SEQ ID NO: 72).

[0145] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Gut_metagenome_contig238000329:

```
RYDKDRSKIY TMMDFVIYRY YIDNNNDSID FINKLRSSID EKSKEKLYNE EANRLWNKLK    60
```

```
EYMLYIKEFN GKLASRTPDR DGNISEFVES LPKIHRLLPR GQKISNFSKL MYLLTMFLDG       120
KEINDLLTTL INKFENIQGF LDIMPEINVN AKFEPEYVFF NKSHEIAGEL KLIKGFAQMG       180
EPAATLKLEM TADAIKILGT EKEDAELIKL AESLFKDENG KLLGNKQHGM RNFIGNNVIK       240
SKRFHYLIRY GDPAHLHKIA TNKNVVRFVL GRIADMQKKQ GQKGKNQIDR YYEVCVGNKD       300
IKKTIEEKID ALTDIIVNMN YDQFEKKKAV IENQNRGKTF EEKNKYKRDN AEREKFKKII       360
SLYLTVIYHI LKNIVNVNSR YILGFHCLER DKQLYIEKYN KDKLDGFVAL TKFCLGDEER       420
YEDLKAKAQA SIQALETANP KLYAKYMNYS DEEKKEEFKK QLNRERVKNA RNAYLKNIKN       480
YIMIRLQLRD QTDSSGYLCG EFRDKVAHLE VARHAHEYI                              519
```

(SEQ ID NO: 73).

[0146] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Gut_metagenome_contig2643000492:

```
NGEIVSLAEK EAFSAKIADK NIGCKIENKQ FRHPKGYDVI ADNPIYKGSP RQDMLGLKET        60
LEKRYFSPSD SIDNVRVQVA HNILDIEKIL AEYITNAVYS FDNIAGFGKD IIGDDFSPVY       120
TYDKFEKSDR YEYFKNLLNN SRLGYYGQAF FECDDSKENK KKKDAIKCYN IIALLSGLRH       180
W                                                                       181
```

(SEQ ID NO: 84).

[0147] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Gut_metagenome_contig874000057:

```
MSKNKESYAK GMGLKSALVS GSKVYMTSFE GGNDAKLEKV VENSEIVSLA EKESFSAEIF        60
KKNIGCKIEN KKFKHPKRYD VIADNPLYKG SVRQDMLGLK ETLEKRYFNS ADGTDNVCIQ       120
VIHNILDIEK ILAEYITNAV YSFDNIAGFG EDIIGMGGFK PIYTYKQFKE PDKYNKKFDD       180
ILNNSRLGYY GKAFFEKNDL KHNPNKKKRD KNPYILKYDN ECYYIIALLS GLRHWNIHSH       240
AKDDLVSYRW LYNLDSILNR EYISTLNYLY DDIADELTES FSKNSSANVN YIAETLNIDP       300
SEFAQQYFRF SIMKEQKNMG FNVSKLREIM LDRKELSDIR DNHRVFDSIR SKLYTMMDFV       360
IYRYYIEEAA KTEAENRNLP ENEKKISEKD FFVINLRGSF DENQKEKLYI EEAKRLWEKL       420
KDIMLKIKEF RGEKVKEYKK                                                   440
```

(SEQ ID NO: 85).

[0148] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Gut_metagenome_contig4781000489:

```
LDKQLDYEYI RTLNYMFNDI ADELTRTFSK NSAANVNYIA ETLNIDPNKF AEQYFRFSIM        60
KEQKNLGFNL TKLRESMLDR RELSDIRDNH NVFDSIRPKL YTMMDFVIYK HYIDEAKKTE       120
AENKSLPDDR KNLSEKD                                                      137
```

(SEQ ID NO: 86).

[0149] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Gut_metagenome_contig12144000352:

```
RMGEPVANTK RVMMIDAVKI LGTDLSDDEL KEMADSFFKD SDGNLLKKGK HGMRNFITNN        60
VIKNKRFHYL IRYGDPAHLH EIAKNEA                                           87
```

(SEQ ID NO: 87).

[0150] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Gut_metagenome_contig5590000448:

```
VHNNEEKDLI KYTWLYNLDK YLDAEYITTL NYMYNDIGDE LTDSFSKNSA ANINYIAETL        60
GIDPKTFAEQ YFRFSIMKEQ KNLGFNLTKL REVMLDRKDM SEIRENHNDF DSIRAKVYTM       120
MDFVIYRYYI EEAAKVNAAN KSLPDNEKSL SEKDIFVISL RGSFNEDQKD RLYYDEAQRL       180
WSKVGKLMLK IKKFRGKDTR KYKNMGTPRI RRLIPEGRDI STFSKLMYAL TMFLDGKEIN       240
DLLTTLINKF DNIQSFLKVM PLIGVNAKFA EEYSFFNNSE KIADELRLIK SFARMGEPVA       300
DARRAMYIDA IRILGTDLSD DELKALADSF SLDENGNKLG KGKHGMRNFI INNVITNKRF       360
HYLIRYGNPV HLHEIAKNEA VVKFVLGRIA DIQKKQGQNG KNQIDRYYET CIGK             414
```

(SEQ ID NO: 88).

[0151] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Gut_metagenome_contig525000349:

```
MSKKENRKSY VKGLGLKSTL VSDSKVYLTT FADGSNAKLE KCVENNKIIC ISNDKEAFAA        60
SIANKNVGYK IKNDEKFRHP KGYDIISNNP LLHNNSVQQD MLGLKNVLEK RYFGKSSGGD       120
NNLCIQIIHN IIDIEKILSE YIPNVVYAFN NIAGFKDEHN NIIDIIGTQT YNSSYTYADF       180
SKDKSDKKYI EFQKLLKNKR LGYWGKAFFT GQGNNAKVRQ ENQCFHIIAL LISLRNWATH       240
SNELDKHTKR TWLYKLDDTN ILNAEYVKTL NYLYDTIADE LTKSFSKNGA VNVNYLAKKY       300
NIKDDLPGFS EQYFRFSIMK EQKNLGFNIS KLRENMLDFK DMSVI                        345
```

(SEQ ID NO: 89).

[0152] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Gut_metagenome_contig7229000302:

```
KKISSLTKFC LGESDEKKLK ALAKKSLEEL KTTNSKLYEN YIKYSDERKA EEAKRQINRE        60
RAKTAMNAHL RNTKWNDIMY GQLKDLADSK SRICSEFRNK AAHLEVARYA HMYINDISEV       120
KSYFRLYHYI MQRRIIDVIE NNPKAKYEGK VKVYFEDVKK NKKYNKNLLK LMCVPFGYCI       180
PRFKNLSIEQ MFDMNETDNS DKKKEK                                            206
```

(SEQ ID NO: 90).

[0153] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Gut_metagenome_contig3227000343:

```
IGDISEVNSY FQLYHYIMQR ILIDKIGSKT TGKAKEYFDS VIVNKKYDDR LLKLLCSPLG        60
YCLTRYKDLS IEALFDMNEA AKYDKLNKER KNKKK                                   95
```

(SEQ ID NO: 91).

[0154] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Gut_metagenome_contig7030000469:

```
SIRSKLYTMM DFVIYRYYIE ESAKAAAENK PSESDSFVIR LRGSFNENQK EELYIEEAER        60
LWKKFGEIML KIKEFRGEKV KEYKKEVPRI ERILPHGKDI SAFSKLMYML SMFLD             115
```

(SEQ ID NO: 92).

[0155] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d gut_metagenome_P17E0k2120140920, c87000043:

```
MYFSKMIYML TYFLDGKEIN DLLTTLISKF DNIKEFLKIM KSSAVDVECE LTAGYKLFND        60
SQRITNELFI VKNIASMRKP AASAKLTMFR DALTILGIDD KITDDRISEI LKLKEKGKGI       120
HGLRNFITNN VIESSRFVYL IKYANAQKIR EVAKNEKVVM FVLGGIPDTQ IERYYKSCVE       180
FPDMNSSLEA KRSELARMIK NISFDDFKNV KQQAKGRENV AKERAKAVIG LYLT             234
```

(SEQ ID NO: 93).

[0156] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Metagenomic hit (no protein accession): contig emb|OBVH01003037.1, human gut metagenome sequence (also found in WGS contigs emb|OBXZ01000094.1| and emb|OBJF01000033.1|):

```
MAKKKRITAK ERKQNHRELL MKKADSNAEK EKAKKPVVEN KPDTAISKDN TPKPNKEIKK        60
SKAKLAGVKW VIKANDDVAY ISSFGKGNNS VLEKRIMGDV SSNVNKDSHM YVNPKYTKKN       120
YEIKNGFSSG SSLVTYPNKP DKNSGMDALC LKPYFEKDFF GHIFTDNMHI QAIYNIFDIE       180
KILAKHITNI IYTVNSFDRN YNQSGNDTIG FGLNYRVPYS EYGGGKDSNG EPKNQSKWEK       240
RDNFIKFYNE SKPHLGYYEN IFYDHGEPIS EEKFYNYLNI LNFIRNNTFH YKDDDIELYS       300
ENYSEEFVFI NCLNKFVKNK FKNVNKNFIS NEKNNLYIIL NAYGKDTENV EVVKKYSKEL       360
YKLSVLKTNK NLGVNVKKLR ESAIEYGYCP LPYDKEKEVA KLSSVKHKLY KTYDFVITHY       420
LNSNDKLLLE IVETLRLSKN DDEKENVYKK YAEKLFKADD VINPIKAISK LFARKGNKLF       480
KEKIIIKKEY IEDVSIDKNI YDFTKVIFFM TCFLDGKEIN DLLTNIISKL QVIEDHNNVI       540
KFISNNKDAV YKDYSDKYAI FRNAGKIATE LEAIKSIARM ENKIENAPQE PLLKDALLSL       600
GVSDDTKVLE NTYNKYFDSK EKTDKQSQKV STFLMNNVIN NNRFKYVIKY INPADINGLA       660
KNRYLVKFVL SKIPEEQIDS YYKLFSNEEE PGCEEKIKLL TKKISKLNFQ TLFENNKIPN       720
VEKEKKKAII TLYFTIVYIL VKNLVNINGL YTLALYFVER DGYFYKDICG KKDKKKSYND       780
VDYLLLPEIF SGSKYREETK NLKLPKEKDR DIMKKYLPND KDREKYNKFF TAYRNNIVHL       840
NIIAKLSELT KNIDKDINSY FDIYHYCTQR VMFNYCKEKN DVVLAKMKDL AHIKSDCNEF       900
SSKHTYPFSS AVLRFMNLPF AYNVPRFKNL SYKKFFDKQ                              939
```

(SEQ ID NO: 94).

[0157] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Metagenomic hit (no protein accession): contig tpg|DJXD01000002.1| (uncultivated Ruminococcus assembly, UBA7013, from sheep gut metagenome):

```
MKKQKSKKTV SKTSGLKEAL SVQGTVIMTS FGKGNMANLS YKIPSSQKPQ NLNSSAGLKN        60
VEVSGKKIKF QGRHPKIATT DNPLFKPQPG MDLLCLKDKL EMHYFGKTFD DNIHIQLIYQ       120
ILDIEKILAV HVNNIVFTLD NVLHPQKEEL TEDFIGAGGW RINLDYQTLR GQTNKYDRFK       180
NYIKRKELLY FGEAFYHENE RRYEEDIFAI LTLLSALRQF CFHSDLSSDE SDHVNSFWLY       240
QLEDQLSDEF KETLSILWEE VTERIDSEFL KTNTVNLHIL CHVFPKESKE TIVRAYYEFL       300
IKKSFKNMGF SIKKLREIML EQSDLKSFKE DKYNSVRAKL YKLFDFIITY YYDHHAFEKE       360
ALVSSLRSSL TEENKEEIYI KTARTLASAL GADFKKAAAD VNAKNIRDYQ KKANDYRISF       420
```

```
EDIKIGNTGI GYFSELIYML TLLLDGKEIN DLLTTLINKF DNIISFIDIL KKLNLEFKFK       480
PEYADFFNMT NCRYTLEELR VINSIARMQK PSADARKIMY RDALRILGMD NRPDEEIDRE       540
LERTMPVGAD GKFIKGKQGF RNFIASNVIE SSRFHYLVRY NNPHKTRTLV KNPNVVKFVL       600
EGIPETQIKR YFDVCKGQEI PPTSDKSAQI DVLARIISSV DYKIFEDVPQ SAKINKDDPS       660
RNFSDALKKQ RYQAIVSLYL TVMYLITKNL VYVNSRYVIA FHCLERDAFL HGVTLPKMNK       720
KIVYSQLTTH LLTDKNYTTY GHLKNQKGHR KWYVLVKNNL QNSDITAVSS FRNIVAHISV       780
VRNSNEYISG IGELHSYFEL YHYLVQSMIA KNNWYDTSHQ PKTAEYLNNL KKHHTYCKDF       840
VKAYCIPFGY VVPRYKNLTI NELFDRNNPN PEPKEEV                                877
```

(SEQ ID NO: 95).

[0158] An exemplary direct repeat sequence of CasRX/Cas13d Metagenomic hit (no protein accession): contig tpg|DJXD01000002.1| (uncultivated Ruminococcus assembly, UBA7013, from sheep gut metagenome) (SEQ ID NO: 95) comprises or consists of the nucleic acid sequence:

CasRX/Cas13d DR:

```
caactacaac cccgtaaaaa tacggggttc tgaaac                                  36
```

[0159] (SEQ ID NO: 96).

[0160] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence:

CasRX/Cas13d Metagenomic hit (no protein accession): contig OGZC01000639.1 (human gut metagenome assembly):

```
MKKKNIRATR EALKAQKIKK SQENEALKKQ KLAEEAAQKR REELEKKNLA QWEETSAEGR        60
RSRVKAVGVK SVFVVGDDLY LATFGNGNET VLEKKITPDG KITTFPEEET FTAKLKFAQT       120
EPTVATSIGI SNGRIVLPEI SVDNPLHTTM QKNTIKRSAG EDILQLKDVL ENRYFDRSFN       180
DDLHIRLIYN ILDIEKILAE YTTNAVFAID NVSGCSDDFL SNFSTRNQWD EFQNPEQHRE       240
HFGNKDNVIC SVKKQQDLFF NFFKNNRIGY FGKAFFHAES ERKIVKKTEK EVYHILTLIG       300
SLRQWITHST EGGISRLWLY QLEDALSREY QETMNNCYNS TIYGLQKDFE KTNAPNLNFL       360
AEILGKNASE LAEPYFRFII TKEYKNLGFS IKTLREMLLD QPDLQEIREN HNVYDSIRSK       420
LYKMIDFVLV YAYSNERKSK ADALASNLRS AITEDAKKRI YQNEADQLWT SYQELFKRIR       480
GFKGAQVKEY SSKNMPIPIQ KQIQNILKPA EQVTYFTKLM YLLTMFLDGK EINDLLTTLI       540
NKFDNISSLL KTMEQLELQT TFKEDYTFFQ QSSRLCKEIT QLKSFARMGN PISNLKEVMM       600
VDAIQILGTE KSEQELQSMA CFFFRDKNGK KLNTGEHGMR NFIGNNVISN TRFQYLIRYG       660
NPQKLHTLSQ NETVVRFVLS RIAKNQRVQG MNGKNQIDRY YETCGGTNSW SVSEEEKINF       720
LCKILTNMSY DQFQDVKQSG AEITAEEKRK KERYKAIISL YLTVLYQLIK NLVNINARYI       780
IAFHCLERDA ILYSSKFNTS INLKKRYTAL TEMILGYETD EKARRKDTRT VYEKAEAAKN       840
RHLKNVKWNC KTRENLENAD KNAIVAFRNI VAHLWIIRDA DRFITGMGAM KRYFDCYHYL       900
LQRELGYILE KSNQGSEYTK KSLEKVQQYH SYCKDFLHML CLPFAYCIPR YKNLSIAELF       960
DRHEPEAEPK EEASSVNNSQ FITT                                              984
```

(SEQ ID NO: 97).

[0161] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence:

CasRX/Cas13d Metagenomic hit (no protein accession): contig emb|OHBM01000764.1 (human gut metagenome assembly):

```
XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX        60
XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX       120
XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX       180
XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXHPLQKRYR YLTSTNLKSF       240
ETYKNNLVNK KKFDLDRVKK IPQLAYFGSA FYNTPEDTSA KITKTKIKSN EEIYYTFMLL       300
STARNFSAHY LDRNRAKSSD AEDFDGTSVI MYNLDNEELY KKLYNKKVHM ALTGMKKVLD       360
ANFNKKVEHL NNSFIKNSAK DFVILCEVLG IKSRDEKTKF VKDYYDFVVR KNYKHLGFSV       420
KELRELLFAN HDSNKYIKEF DKISNKKFDS VRSRLNRLAD YIIYDYYNKN NAKVSDLVKY       480
LRAAADDEQK KKIYLNESIN LVKSGILERI KKILPKLNGK IIGNMQPDST ITASMLHNTG       540
KDWHPISENA HYFTKWIYTL TLFMDGKEIN DLVTTLINKF DNIASFIEVL KSQSVCTHFS       600
EERKMFIDSA EICSELSAMN SFARMEAPGA SSKRAMFVEA ARILGDNRSK EELEEYFDTL       660
FDKSASKKEK GFRNFIRNNV VDSNRFKYLT RYTDTSSVKA FSNNKALVKF AIKDIPQEQI       720
LRYYNSCFGA SERYYNDGMS DKLVEAIGKI NLMQFNGVIQ QADRNMLPEE KKKANAQKEK       780
YKSIIRLYLT VCYLFFKNLV YVNSRYYSAF YNLEKDRSLF EINGELKPTG KFDEGHYTGL       840
VKLFIDNGWI NPRASAYLTV NLANSDETAI RTFRNTAEHL EALRNADKYL NDLKQFDSYF       900
EIYHYITQRN IKEKCEMLKE QTVKYNNDLL KYHGYSKDFV KALCVPFGYN LPRFKNLSID       960
ALFDKNDKRE KLKKGFED                                                     978
```

(SEQ ID NO: 98).

[0162] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence:

CasRX/Cas13d Metagenomic hit (no protein accession): contig emb|OHCP01000044.1 (human gut metagenome assembly):

```
MAKKITAKQK REEKERLNKQ KWAKNDSVII VPETKEEIKT GEIQDNNRKR SRQKSQAKAM        60
GLKAVLSFDN KIAIASFVSS KNAKSSHIER ITDKEGTTIS VNSKMFESSV NKRDINIEKR       120
ITIEEPQQDG TIKKEEKGVK STTCNPYFKV GGKDYIGIKE IAEEHFFGRA FPNENLRVQI       180
AYNIFDVQKI LGTFVNNIIY SFYNLSRDEV QSDNDVIGML YSISDYDRQK ETETFLQAKS       240
LLKQTEAYYA YFDDVFKKNK KPDKNKEGDN SKQYQENLRH NFNILRVLSF LRQICMHAEV       300
HVSDDEGCTR TQNYTDSLEA LFNISKAFGK KMPELKTLID NIYSKGINAI NDEFVKNGKN       360
NLYILSKVYP NEKREVLLRE YYNFVVCKEG SNIGISTRKL KETMIAQNMP SLKEENTYRN       420
KLYTVMNFIL VRELKNCATI REQMIKELRA NMDEEEGRDR IYSKYAKEIY LYVKDKLKLM       480
LNVFKEEAEG IIIPGKEDPV KFSHGKLDKK EIESFCLTTK NTEDITKVIY FLCKFLDGKE       540
INELCCAMMN KLDGISDLIE TAKQCGEDVE FVDQFKCLSK CATMSNQIRI VKNISRMKKE       600
MTIDNDTIFL DALELLGRKI EKYQKDKNGD YVKDEKGKKV YTKDYNNFQD MFFEGKNHRV       660
RNFVSNNVIK SKWFSYVVRY NKPAECQALM RNSKLVKFAL DELPDSQIEK YYISVFGEKS       720
SSSNEEMRRE LLKKLCDFSV RGFLDEIVLL SEDEMKQKDK FSEKEKKKSL IRLYLTIVYL       780
ITKSMVKINT RFSIACATYE RDYILLCQSE KAERAWEKGA TAFALTRKFL NHDKPTFEQY       840
YTREREISAM PQEKRKELRK ENDQLLKKTH YSKHAYCYIV DNVNNLTGAV ANDNGRGLPC       900
LSEKNDNANL FLEMRNKIVH LNVVHDMVKY INEIKNITSY YAFFCYVLQR MIIGNNSNEQ       960
NKFKAKYSKT LQEFGTYSKD LMWVLNLPFA YNLPRYKNLS NEQLFYDEEE RMEKIVGRKN      1020
```

DSR 1023 (SEQ ID NO: 99).

[0163] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Metagenomic hit (no protein accession): contig emb|OGDF01008514.1| (human gut metagenome assembly):

```
MTETKPKRED IAKTPAAKSR SKAAGLKSTF AVNGSVLLTS FGRGNDAVPE KLITEKAVSE        60
INTVKPRFSV EKPATSYSSS FGIKSHISAT ADNPLAGRAP VGEDAIHAKE VLEQRVFGKT       120
FSDDNIHIQL IYNILDIRKI LSTYANNVVF TINSMRRLDE YDREQDYLGY LYTGNSYERL       180
LDIADKYAVD GEDWRNTAAG ISNDFEKKQF QTINGFWDLL DMIEPYMCYF SEAFFCETTV       240
KDPDSGRIVP CLEQRSDGDI YNILRILSIV RQTCMHDNAS MRTVMFTLGQ NSVRDRKNGF       300
DELAELLDYL YDEKIDIVNR DFLRNQKNNI ELLSRIYGSS ADSPERDRLV QNFYDFRVLS       360
QDKNLGFSIK KLREKLLDSP ALSVVRSKKY DTMRSKIYSL IDFMIYRKFS ENHVAVDDFV       420
EELRSLLTED EKESAYSRWA ETLINDGFAQ EILVKLLPQT DPAVIGKIKG KKLLNDSIAG       480
IKLKKDASFF TKIINVLCMF QDGKEINELV SSLVNKFANI QSFVDVMRSQ GIDSGFTADY       540
AMFAESGRIS RELHILKGIA RMQHSIAGLG DVKIYGSDDK FHGVSRRVYT DAAYILGFGE       600
RSEDNDGYVD DYVSSKLLGG ADKNLRNFIT NNVIKNRRFL YTVRYMNPKR AKKLVQNDAL       660
VVLALSGIPE TQIDRYYKSC IEKRSFNPDL NEKIAALSEM ITTLKIDDFE DVKQNPEKNA       720
NYEAKKNQRI SKERYKACIG LYLTVLYLIC KNLVKINARY SIAIGCLERD TQLHGVDFKG       780
AAYMTRDVFI AKGWINPKKP TVKSIKEQYA FLTPYIFTTY RNMIAHLAAV TNAYKYIPQM       840
DRFKSWFHLY HTVIQHSLIQ QYEYDRDYGR KGAPVVSERV LQLLEQCREH SNYSRDLLHI       900
LNLPFGYNLP RYLNLSSEKY FDANAI                                            926
```

[0164] (SEQ ID NO: 100).

[0165] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Metagenomic hit (no protein accession): contig emb|OGPN01002610.1 (human gut metagenome assembly):

```
MAKKITAKQK REEKERLNKQ KWAKQDTPVV PKSKTEEKPV AASDDKLLKT TQVKKVQTKS        60
KAKAMGLKTV LSFDDKIAIA SFVNDKKTKL PHIERITDKS GTTIHENARM FDSSVDEQNV       120
NIEKRMTIEE KQNDGTFKKD EKDVKATICN PYFKTCGKDY IGIKDVAEKY FFGKTFPNEN       180
LRVQIAYNVF DIQKILGTYV NNIIYSFYNL RRDGKSDVDI IGSLYAFADF DNQLKDKPAF       240
REAKDLLKNT EAYFSYFGDV FKKSKKGKKD ENNEDYEKNL RHNFNVLRVL SFLRQICTHA       300
YVKCTGGAKN NGDSTKVEAE SLDALFNITE YFAKTAPELS KTINEIYKEG IDRINNDFVT       360
NGKNNLYILS KVYPDMQRNE LVKKYYQFVV CKEGNNVGIN TRKLKESIIS QHPWITTPQD       420
NNKANDYESC RHKLYTIMCF ILVAELDAHE SIRDNMVAEL RANMDGDDGR DAIYEKYAKD       480
IYHIVKDKLL AMQKVFDEEL VPVKVEGKND PQQFTHGKLG KKEIESFCLS DKNTSDIAKV       540
VYFLCNFLDG KEINELCCAM MNKFDGIGDL IDTAKQCGEE VKFIEEFACL SNCRKITNDI       600
RVAKSISKMK NKVNIDNDII YLDAIELLGR KIEKYQKDEN GKILLGTDGK RLYTQEYKYF       660
NDMFFNAGNH KVRNFIANNV MQSKWFFYVV RYNKPAECQI IMRNKTLVKF TLDDLPDMQI       720
QRYYSSVFGD NNMPAVDEMR KRLLDKINQF SVRGFLDELD EIVLMSDEES KRNKSSEKEQ       780
KKSLIRLYLT IAYLITKSMV KINTRFSIAC AMYERDYALL CQSEMKGGPW DGGAQALAVT       840
```

```
RKFLNHDREV FDRYCAREAE IARLPSEERK PLRKANDKLL KQTHYTNHSY TYIVNNLNSF      900
TDIDYCAKDV GLPAPNDKND NASILGEMRN DIAHLNIVHD MVKYIEELKD ISSYYAFYCY      960
VLQRRLVGKD PNCQNKFKAK YAKELNDYGT YNKNLMWMLN LPFAYNLPRY KNLSSEFLFY     1020
DMEYNKKDDE                                                            1030
```

(SEQ ID NO: 101).

[0166] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Metagenomic hit (no protein accession): from contig emb|OBLI01020244 and emb|OBLI01038679 (from pig gut metagenome):

```
MAKKITAKQR REERERQNKQ KWAKKQADAT AVFECEADIK PADSKDEDCT NIYIKREKKK       60
TQAKAMGLKT VLGFDNKIAI ASFMSSKDSK SSHIERITDP NGKTIREDVR MFDSNVDECS      120
INLEKRMTVE ERQKDGTIKK DEKDVKSTIC NPYSNECGKD YIGIKSVAEE LFFGRTFPND      180
NLRVQIAYNI FDIQKILGTY INNIIYSFYN LSRDESQSDN DVIGTLYMLK DFDGQKETDT      240
FRQARALLER TEAYYSYFDN VFKKIDKNKK KSDDCKRERN EILRYNFNVL RVLSFLRQIC      300
AHAQVKISNE HDREKGGGLV DSLDALFNIS RFFDAVAPEL NEVINSVYSK GIDDINDNFV      360
KNGKNNFYIL SKIYPEVARE DLLREYYYFV VSKEGNNIGI STKKLKEAII VQDMSYIKSE      420
DYDTYRNKLY TVLCFILVKE LNERTTIREQ MVADLRANMN GDIGREDIYS KYAKIIYAQV      480
KPRFDTMKSA FEEEAKDVIV PDKKKPVKFS HGKLDKNEIE RFCITSANTD SVAKIIYFLC      540
KFLDGKEINE LCCAMMNKLD GINDLIETAE QCGAKVEFVD KFSVLSNCET ISDQIRIVKS      600
ISKMKKEIAI DNDTIFLDAL ELLGRKIDKY KKDATGKYLK DENGKYLYSK EYDDFQYMFF      660
KDSHRVRNFI SNSVIKSKWF SYIVRYNQPS ECRAIMKNKT LVKFALDELP DLQIQRYFVA      720
LYGDEDLPSY GEMRKILLKK LHDFSIKGFL DEIVLLSDLD MESQDKYCEK EQKKSLFRLY      780
LTIAYLITKS MVKINTRFSI ACATYERDYA LLCASNKQER AWSSGATALA LTRRFLNQDK      840
LIFEKHYARE GEISKLPKEE RKAMRKVNDQ LLKRTHFSKH SYCYIVDNVN RLTGGECRTD      900
KRVLPVLNEK NDNAGILLDF RKTIAHLNVV HKMVDYVDEI KGITSYYAFF CYVLQRMLVG      960
NNLNEKNAIK EKYSATVKSF GTYSKDFMWL INLPFAYNLP RYKNLSNEQL FYDEEERNET     1020
EEQIDRL                                                               1027
```

(SEQ ID NO: 102).

[0167] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Metagenomic hit (no protein accession): contig OIZX01000427.1:

```
MAKKKTARQ LREEMQQQRK QAIQKQQEQR QEKAAAARET AAPEQPAAAP VPKRQRKSLA       60
KAAGLKSNFI LDPQRRTTVM TAFGQGSTAI LEKQIVDRAI SDLQPVQQFQ VEPASAAKYR      120
LKNSRVRFPN VTADDPLYRR KDGGFVPGMD ALRRKNVLEQ RFFGKSFADN IHIQMIYSIL      180
DIHKILAAAS GHIVHLLNIV NGSKDRDFIG MLAAHVLYNE LNEEAKRSIA DFCKSPRLIY      240
YSAAFYETLD NGKSERRSNE DIFNILALMT CLRNFSSHHS IAIKVKDYSA AGLYNLRRLG      300
PDMKKMLDTF YTEAFIQLNQ SFQDHNTTNL TCLFDILNIS DSARQKQLAE EFYRYVVFKE      360
QKNLGFSVRK LREEMLLLPD AAVIADKRYD TCRSKLYNLM DFLILRVYRT GRADRCDKLP      420
EALRAALTDE EKAVVYHKEA LSLWNEMRTL ILDGLLPQMT PENLSRLSGQ KRKGELSLDD      480
AMLKECLYEP GPVPEDAAPE EANAEYFCRM IYLATLFMDG KEINTLLTTL ISKFENIAAF      540
LQTMEQLNIE AELGPEYAMF TRSRAVAEQL RVINSFALMK KPQVNAKQQL YRAAVTLLGT      600
EDPDGVTDEM LCIDPVTGKM LPPNQRHHGD TGLRNFIANN VVESRRFQYL IRYSDPAQLH      660
```

```
QLASNKKLVR FVLSSIPDTQ INRYYETCGQ TRLAGRAAKV EFLTDMIAAI RFDQFRDVNQ       720
KERGANTQKE RYKAMLGLYQ TVLYLAVKNL VNINARYVMA FHCVERDMFL YDGELTDPKG       780
ESVSAFLAVN GKKGVQPQYL LLTQLFIRRD YLKRSACEQI QHNMENISDR LLREYRNAVA       840
HLNVIAHLAD YSADMREITS YYGLYHYLMQ RHLFKRHAWQ IRQPERPTEE EQKLIEQEQK       900
QLAWEKALFD KTLQYHSYNK DLVKALNAPF GYNLARYKNL SIEPLFSKEA APAAEIKATH       960
A                                                                       961
```

(SEQ ID NO: 103).

[0168] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence:

CasRX/Cas13d Metagenomic hit (no protein accession): contig OCTW011587266.1:

```
MKQNDRENNN KIKKSAAKAV GVKSLARLSD GSTVVSSFGK GAAAELESLI TGGEIRKLSD        60
KAILEITDDT QNKNAYNVKS SRIPNLTART DKLSDKSGMD DLGFKRELEL EVFGQCFDDS       120
IHIQIAHAVF DIQKSLAAVI PNVLYTLNNL DRSYSTDNTS DKKDIIGNTL NYQHSYESFN       180
VEKRGEFTEY YNAAKDRFSY FPDILCVLEK VNGKDRYQPK SEKDAFNVLS SVNMLRNSLF       240
HFAPKSNDGK ARIAVFKNQF DSDFSHITST VNKIYSAKIA GVNENFLNNE GNNLYIILKA       300
TNWDIKKIVP QLYRFSVLKS DKNMGFNMRK LREFAVESKN IDLSRLNDKF LTNNRKKLYK       360
VIDFIIYYHL NKVLKDSFVD DFVAALRASQ SEEEKEKLYA QYSERLFADE GLKSAIKKAV       420
DMISDTKSNI FKMKTPLDKA LIENIKVNSD ASDFCKLIYV FTRFLDGKEI NILLNSLIKK       480
FQDIHSFNTT VKKLSENNLI INADYVDDYS LFEQSGTVAR ELMLIKSISK MDFGLDNINL       540
SFMYDDALRT LGVSDENLPE VKREYFGKTK NLSAYIRNNV LENRRFKYVI KYIHPSDVQK       600
IACNKAIAGF VLNRMPDTQI KRYYDSLINK GATDIQAQAK ALLDCITGIS FDAIKDDKHL       660
HKSKEKSPQR SADRERKKAM LTLYYTIVYI FVKQMLHINS LYTIGFFYLE RDQRFIYSRA       720
KKENKNPSKN SYLNDFRSVT AYFIPSEIMK RIEKNENKGF LEDFEALWNS CGKTSRLRKE       780
DVLLYARYIS PDHALKNYKM ILNSYRNKIA HINVIMSAGK YTGGIKRMDS YFSVFQHLVQ       840
CDILSNPNNK GKCFESESLK PLLLDMKFDG TDEKLYSKRL TRALNIPFGY NVPRYKNLTF       900
EKIYLKSSIN E                                                            911
```

(SEQ ID NO: 104).

[0169] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence:

CasRX/Cas13d Metagenomic hit (no protein accession): contig emb|OGNF01009141.1:

```
MADIDKKKSS AKAAGLKSTF VLENNKLLMT SFGNGNKAVI EKIIDEKVDS INEPEVFSVT        60
PCDKKFELQP AKRGLAADSL VDNPLKSKKT AGDDAIHSRK FLERQFFDGN TFNDNIHIQL       120
IYNILDIEKI LSVHVNDIVY SVNNILSRGE GMEYNDYIGT LNLKSFETYK NNLVNKKKFD       180
LDRVKKIPQL AYFGSAFYNT PEDTSAKITK TKIKSNEEIY YTFMLLSTAR NFSAHYLDRN       240
RAKSSDAEDF DGTSVIMYNL DNEELYKKLY NKKVHMALTG MKKVLDANFN KKVEHLNNSF       300
IKNSAKDFVI LCEVLGIKSR DEKTKFVKDY YDFVVRKNYK HLGFSVKELR ELLFANHDSN       360
KYIKEFDKIS NKKFDSVRSR LNRLADYIIY DYYNKNNAKV SDLVKYLRAA ADDEQKKKIY       420
LNESINLVKS GILERIKKIL PKLNGKIIGN MQPDSTITAS MLHNTGKDWH PISENAHYFT       480
KWIYTLTLFM DGKEINDLVT TLINKFDNIA SFIEVLKSQS VCTHFSEERK MFIDSAEICS       540
ELSAMNSFAR MEAPGASSKR AMFVEAARIL GDNRSKEELE EYFDTLFDKS ASKKEKGFRN       600
FIRNNVVDSN RFKYLTRYTD TSSVKAFSNN KALVKFAIKD IPQEQILRYY NSCFGASERY       660
YNDGMSDKLV EAIGKINLMQ FNGVIQQADR NMLPEEKKKA NAQKEKYKSI IRLYLTVCYL       720
```

```
FFKNLVYVNS RYYSAFYNLE KDRSLFEING ELKPTGKFDE GHYTGLVKLF IDNGWINPRA       780
SAYLTVNLAN SDETAIRTFR NTAEHLEALR NADKYLNDLK QFDSYFEIYH YITQRNIKEK       840
CEMLKEQTVK YNNDLLKYHG YSKDFVKALC VPFGYNLPRF KNLSIDALFD KNDKREKLKK       900
GFED                                                                    904
```

(SEQ ID NO: 105).

[0170] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Metagenomic hit (no protein accession): contig emb|OIEN01002196.1:

```
MERQKRKMKS KSKMAGVKSV FVIGDELLMT SFGDGDDAVL EKDIDENGVV NDCRNPAAYD        60
AVYGTDSIRV KKTNNNIRAK VNNPLAKSNI RSEESALFRT RVNEYKREQK DKYETLFFGK       120
TFDDNIHIQL ISKILDIEKT FSVVIGNIVY AINNLSLEQS IDRPIDIFGD KNTQGISLRE       180
DNDYLKTMLP RCEYLFHNIL NSDSDNNSKM NYNKVNKGKE EKDNRNNENI EKLKKALEVI       240
KIIRVDSFHG VDGIKGDQKF PRSKYNLAVN YNEEIQKTIS EPFNRKVEEV QQDFYRNSCV       300
NIDFLKEIMY GSNYTDRGSD SLECSYFNFA ILKQNKNMGF SITSIRECLL DLYELNFESM       360
QNLRPRANSF CDFLIYDYYC KNESERANLV DCLRSAASEE EKKNIYFQTA ERVKEKFRNA       420
FNRISRFDAS YIKNSREKNL SGGSSLPKYS FIEGFTKRSK KINDNDEKNA DLFCNMLYYL       480
AQFLDGKEIN IFLTSIHNIF QNIDSFLKVM KEKGMECKFQ KDFKMFSHAG HVAKKIEIVI       540
SLAKMKKTLD FYNAQALKDA VTILGVSKKH QYLDMNSYLD FYMFDNRSGA TGKNAGKDHN       600
LRNFLVSNVI RSRKFNYLSR YSNLAEVKKL AQNPSLVQFV LSRIEPSLIC RYYESSQGIS       660
SEGITIDEQI KKLTGIIVDM NIDSFENINN GEIGMRYSKA TPQSIERRNQ MRVCVGLYLN       720
VLYQIEKNLM NVNARYVLAF AFAERDALML NFTLEECKKN KKRSSGGFSF IEMTQFFIDK       780
KLFKVATEAI KKNVLKYNGN PESLNHIPGE YICKNMEGYH ENTVRNFRNM VAHLTAVARV       840
PLYISEVTQI DSYYALYHYC MQMNILQGIE QSGKILDNIK LKNALENARV HRTYSKDAVK       900
YLCLPFAYNI SRYKALTIKD LFDWTEYSCK KDE                                    933
```

(SEQ ID NO: 106).

[0171] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence: CasRX/Cas13d Metagenomic hit (no protein accession): contig e-k87_11092736:

```
MKRQKTFAKR IGIKSTVAYG QGKYAITTFG KGSKAEIAVR SADPPEETLP TESDATLSIH        60
AKFAKAGRDG REFKCGDVDE TRIHTSRSEY ESLISNPAES PREDYLGLKG TLERKFFGDE       120
YPKDNLRIQI IYSILDIQKI LGLYVEDILH FVDGLQDEPE DLVGLGLGDE KMQKLLSKAL       180
PYMGFFGSTD VFKVTKKREE RAAADEHNAK VFRALGAIRQ KLAHFKWKES LAIFGANANM       240
PIRFFQGATG GRQLWNDVIA PLWKKRIERV RKSFLSNSAK NLWVLYQVFK DDTDEKKKAR       300
ARQYYHFSVL KEGKNLGFNL TKTREYFLDK FFPIFHSSAP DVKRKVDTFR SKFYAILDFI       360
IYEASVSVAN SGQMGKVAPW KGAIDNALVK LREAPDEEAK EKIYNVLAAS IRNDSLFLRL       420
KSACDKFGAE QNRPVFPNEL RNNRDIRNVR SEWLEATQDV DAAAFVQLIA FLCNFLEGKE       480
INELVTALIK KFEGIQALID LLRNLEGVDS IRFENEFALF NDDKGNMAGR IARQLRLLAS       540
VGKMKPDMTD AKRVLYKSAL EILGAPPDEV SDEWLAENIL LDKSNNDYQK AKKTVNPFRN       600
YIAKNVITSR SFYYLVRYAK PTAVRKLMSN PKIVRYVLKR LPEKQVASYY SAIWTQSESN       660
SNEMVKLIEM IDRLTTEIAG FSFAVLKDKK DSIVSASRES RAVNLEVERL KKLTTLYMSI       720
AYIAVKSLVK VNARYFIAYS ALERDLYFFN EKYGEEFRLH FIPYELNGKT CQFEYLAILK       780
YYLARDEETL KRKCEICEEI KVGCEKHKKN ANPPYEYDQE WIDKKKALNS ERKACERRLH       840
FSTHWAQYAT KRDENMAKHP QKWYDILASH YDELLALQAT GWLATQARND AEHLNPVNEF       900
```

```
DVYIEDLRRY PEGTPKNKDY HIGSYFEIYH YIRQRAYLEE VLAKRKEYRD SGSFTDEQLD       960
KLQKILDDIR ARGSYDKNLL KLEYLPFAYN LPRYKNLTTE ALFDDDSVSG KKRVAEWRER      1020
EKTREAEREQ RRQR                                                        1034
```

(SEQ ID NO: 107).

An exemplary direct repeat sequence of CasRX/Cas13d Metagenomic hit (no protein accession): contig e-k87_11092736 (SEQ ID NO: 107) comprises or consists of the nucleic acid sequence:

CasRX/Cas13d Direct repeat 1: `gtgagaagtc tccttatggg gagatgctac`

(SEQ ID NO: 108).

[0172] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence:

CasRX/Cas13d Ga0129306_1000735:

```
MQKQREQQTV TDESERKKKP LKSGAKAAGL KSVFVLSEGK ELLTSFGRGN EAVPEKRVTG        60
GTIANARTDN KEAFSAALQN KRFEVFGRTA GSSDDPLAVS RAPGQDLIGA KTALEERYFG       120
RAFADNIHMQ VIYAIQDINK ILAVHANNIV YTLNNLDREA DPETDDFIGS GYLTLKNTFE       180
TYCDPAALNE REREKVTVSK QHFDAFMQNP RLAYYGNAFF RKLSKAERLA RGREIFDKES       240
PERRQEILGS RGKNKSVDDE IRALAPEWVK REERDVYSEL VLMSELRQSC FHGQQKNSAR       300
IFRLDNDLGP GVDGARELLD RLYAEKINDL RSFDKTSASS NFRLLFNAYH ADNEKKKELA       360
QEFYRFSVLK VSKNTGFSIR TLREKIIEDH AAQYRDKIYD SMRKKLFSTF DFFLWRFYEE       420
REDEAEELRA CLRAARSDEE KEQIYAEAAA SCWPSVKPFV ESVAATLCDV VKGRTKLNKL       480
KLSADESTLV RNAIDGVRIS PRASYFTKLI YLMTLFLDGK EINDLLTTLI HAFENIDSFL       540
SVLGSERLER TFDANYRIFA DSGVIAQELR AVNSFARMTT EPFNSKLVMF EDAAQLFGMS       600
GGLVEHAEEL REYLDNKMLD KTKLRLLPDG KVDTGFRNFI ISNVTESRRF RYLVRYCEPR       660
AVRDYMSCRP LIRLTLRDMP DTILRRYYEQ SVGAATVDRE RILDTLADKL LSLRFTDFEN       720
VNQRANAERN REKQKMMGII SLYLNVAYQI VKNLVYVNAR YTMAYHCAER DTELLLNAAG       780
EGNLLRRDRS WPARLHLPRR ALARRRDRVE VMERDVARGP EAYNRDEWLG LVRTLRREKR       840
VCDNLHNNYA YLCGADAEPG DASLSLLFVY RNKAAHLSVL NKGGRLSGDL KEAKSWFYVY       900
HFLMQRVLEE EFRNTQALPE RLRELLMMAE RYRGCSKDLI KVLNLTFAYN LPRYKNLSID       960
GRFDKNHPDP SDE                                                          973
```

(SEQ ID NO: 109).

[0173] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence:

CasRX/Cas13d Ga0129317_1008067:

```
MKKQKKSLVK AAGLKSAFVV GDSVYLTSFG KGNAARLDTK INPDNSTERY VSDSEKHTLK        60
INSITDTELR LSGPFPKQAE AKNPTHKKDN EQKNTRQDML GLKSTLEKFY FGSTFDDNIH       120
IQIIHNIQDI AKILAAHSNN AGYALDNMLA YQGVEFSDMI GYMGTSRTFD NYDPNHKNNK       180
DFFRFLKLPR LGYFGSAFYS QKGKDFEKRS DEEVYNICAL MGQIRQCCFH GKQEKYQLKW       240
LYNFHNFKSN KPFLDTLDKH FDEMIDRINK NFIKNNTPDL IILSGLYPDM AKKELVRLFY       300
DFTTVKEYKN MGFSVKKLRE KMLESEEASD FRDKDYDSVR RKLYKLMDFC IYYLYYSDSE       360
RNENLVSRLR ESLTDENKDI IYSKEAKIVW NELRKKFSTI LDNVKGSNIK KLENVKEKFI       420
SEDEFDDIKL DIDISYFSKL MYVMCYFLDG KEINDLLTTL VSKFDNIGSI IEAATQIGIN       480
IEFIDDFKFF DRSKDISVEL NIIRNFARMQ APVPNAKRAM QEDAIRILGG SEEDIFSILD       540
```

```
DMTGYDKSGK KLAQSKKGFR NFIINNVVES SRFKYIVRYS NPQKIRKLAN NSVVVGFVLG       600
KLPDAQIESY FNSCLPNRVY STPDKARESL RDMLHNISFN DFADVKQDDR RATPEEKVEK       660
ERYKAIIGLY LTVMYHLVKN LVYVNSRYVM AFHCLERDAM HYDVSLDNYR DLIRHLISEG       720
DSSCNHFISH NRRMRDCIEE NVKNSEQLIF GKEDAVIRFR NNVAHLSAIR NANEYIGDIR       780
EITSYFALYH YLMQRKLIDD CKVNDTAHKY FEQLTKYKTY VMDMVKALCS PFGYNLPRFK       840
NLSIEGKFDM HESK                                                        854
```

(SEQ ID NO: 110).

[0174] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence:

CasRX/Cas13d Ga0224415_10048792:

```
MSKKENRKSY VKGLGLKSTL VSDSKVYLTT FADGSNAKLE KCVENNKIIC ISNDKEAFAA        60
SIANKNVGYK IKNDEKFRHP KGYDIISNNP LLHNNSVQQD MLGLKNVLEK RYFGKSSGGD       120
NNLCIQIIHN IIDIEKILSE YIPNVVYAFN NIAGFKDEHN NIIDIIGTQT YNSSYTYADF       180
SKDKSDKKYI EFQKLLKNKR LGYWGKAFFT GQGNNAKVRQ ENQCFHIIAL LISLRNWATH       240
SNELDKHTKR TWLYKLDDTN ILNAEYVKTL NYLYDTIADE LTKSFSKNGA VNVNYLAKKY       300
NIKDDLPGFS EQYFRFSIMK EQKNLGFNIS KLRENMLDFK DMSVIRDDHN RYDKDRSKIY       360
TMMDFVIYRY YIDNNNDSID FINKLRSSID EKSKEKLYNE EANRLWNKLK EYMLYIKEFN       420
GKLASRTPDR DGNISEFVES LPKIHRLLPR GQKISNFSKL MYLLTMFLDG KEINDLLTTL       480
INKFENIQGF LDIMPEINVN AKFEPEYVFF NKSHEIAGEL KLIKGFAQMG EPAATLKLEM       540
TADAIKILGT EKEDAELIKL AESLFKDENG KLLGNKQHGM RNFIGNNVIK SKRFHYLIRY       600
GDPAHLHKIA TNKNVVRFVL GRIADMQKKQ GQKGKNQIDR YYEVCVGNKD IKKTIEEKID       660
ALTDIIVNMN YDQFEKKKAV IENQNRGKTF EEKNKYKRDN AEREKFKKII SLYLTVIYHI       720
LKNIVNVNSR YILGFHCLER DKQLYIEKYN KDKLDGFVAL TKFCLGDEER FEDLKAKAQA       780
SIQALETANP KLYAKYMNYS DEEKKEEFKK QLNRERVKNA RNAYLKNIKN YIMIRLQLRD       840
QTDSSGYLCG EFRDKVAHLE VARHAHEYIG NIKEVNSYFQ LYHYIMQCRL YDVLKNNTKA       900
EAMVKGKAKE YFEALEKEGT YNDKLLKIAC VPFGYCIPRY KNLSMEELFD MNEEKKFKKK       960
APENT                                                                  965
```

(SEQ ID NO: 111).

[0175] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence

CasRX/Cas13d 160582958 _gene49834:

```
MKNSVTFKLI QAQENKEAAR KKAKDIAEQA RIAKRNGVVK KEENRINRIQ IEIQTQKKSN        60
TQNAYHLKSL AKAAGVKSVF AIGNDLLMTG FGPGNDATIE KRVFQNRAIE TLSSPEQYSA       120
EFQNKQFKIK GNIKVLNHST QKMEEIQTEL QDNYNRPHFD LLGCKNVLEQ KYFGRTFSDN       180
IHVQIAYNIM DIEKLLTPYI NNIIYTLNEL MRDNSKDDFF GCDSHFSVAY LYDELKAGYS       240
DRLKTKPNLS KNIDRIWNNF CNYMNSDSGN TEARLAYFGE LFYKPKETGD AKSDYKTHLS       300
NNQKEEWELK SDKEVYNIFA ILCDLRHFCT HGESITPSGK PFPYNLEKNL FPEAKQVLNS       360
LFEEKAESLG AEAFGKTAGK TDVSILLKVF EKEQASQKEQ QALLKEYYDF KVQKTYKNMG       420
FSIKKLREAI MEIPDAAKFK DDLYSSLRHK LYGLFDFILV KHFLDTSDSE NLQNNDIFRQ       480
LRACRCEEEK DQVYRSIAVK VWEKVKKKEL NMFKQVVVIP SLSKDELKQM EMTKNTELLS       540
SIETISTQAS LFSEMIFMMT YLLDGKEINL LCTSLIEKFE NIASFNEVLK SPQIGYETKY       600
TEGYAFFKNA DKTAKELRQV NNMARMTKPL GGVNTKCVMY NEAAKILGAK PMSKAELESV       660
```

```
FNLDNHDYTY SPSGKKIPNK NFRNFIINNV ITSRRFLYLI RYGNPEKIRK IAINPSIISF       720
VLKQIPDEQI KRYYPPCIGK RTDDVTLMRD ELGKMLQSVN FEQFSRVNNK QNAKQNPNGE       780
KARLQACVRL YLTVPYLFIK NMVNINARYV LAFHCLERDH ALCFNSRKLN DDSYNEMANK       840
FQMVRKAKKE QYEKEYKCKK QETGTAHTKK IEKLNQQIAY IDKDIKNMHS YTCRNYRNLV       900
AHLNVVSKLQ NYVSELPNDY QITSYFSFYH YCMQLGLMEK VSSKNIPLVE SLKNEANDAQ       960
SYSAKKTLEY FDLIEKNRTY CKDFLKALNA PFSYNLPRFK NLSIEALFDK NIVYEQADLK      1020
KE                                                                     1022
```

(SEQ ID NO: 112).

[0176] An exemplary direct repeat sequence of CasRX/Cas13d proteins may comprise or consist of the sequence

CasRX/Cas13d 160582958 _gene49834 (SEQ ID NO: 112) comprises or consists of the nucleic acid sequence: CasRX/Cas13d DR:

```
gaactacacc cctctgttct tgtaggggtc taacac                                  36
```

(SEQ ID NO: 113).

[0177] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence:

CasRX/Cas13d 250twins_35838_GL0110300:

```
MGNKQRVSAQ KRRENAKLCN QQKARQAESQ RDKIKNMNVE KMKNINTNDI KHTKTTAKKL        60
GLKSTIIADK KIILTSFINE QSSKTANIEK VAGFKGDTID TISYTPRMFR SEINPGEIVI       120
SKGDDLSEFA NPANFPIGRD YVKIRSALEK QYFGKEFPED NLHVQIAYNV ADIKKILSVY       180
INNIIYMFYN LARSEEYDIF YNSQSENSGR DCDVIGSLYY QASYRNQDAN RFEKDGKKKA       240
IDSLLDDTRA YYTYFDGLFS VPKREDDGKI KESEKEKAKD QNFDVLRLLS VGRQLTFHSD       300
KSNNEAYLFD LSKLTRAAQD ENRRQDIQSL LNILNSTCRS NLEGVNGDFV KHAKNNLYVL       360
NQLYPSLKAN DLIGEYYNFI VKKENRNIGI RLITVRELII EHNYTNLKDS KYDTYRNKIY       420
TVLNFILFRE IQENSIAIKN FREKLRSTEK AEQPALYQAF ANKIYPMVQA KFAKAIDLFE       480
EQYKTKFKSE FKGGISIENM QQQNILLQTE NIDYFSKYVL FLTKFLDGKE INELLCALIN       540
KFDNIADLLD ISKQIGTPVV FCADYESLND AAKIAENIRL IKNIAHLRPA IQEAQSSKDN       600
ADAAGTPATL LIDAYNMLNT DIQLVYGEAA YEELRKDLFE RKNGTKYNKK GKKVDVYDHK       660
FRNFLINNVI KSKWFFYIAK YVKPADCAKM MSNKKMIEFA LRDLPETQIK RYYYTITGNE       720
ALGDAESLKG VIIEQLHAFS IKNTLLSIKN MGEGEYKIQQ IGSSKEKLKA IVNLYLTVAY       780
LLTKSLVKVN IRFSIAFGCL ERDLVLQKKS EKKFDAIINE ILLEDDKIRK ECDKERAQAK       840
TLPRELAQER FAQIKRRESG CYFKSYHVYD YLSKNSNEFK QNHIDFAVTS YRNNVEHLNV       900
VHCMTKYFSE VKDVKSYYGV YCYIMQRMLC DELIIKNQDK PDVRQTFEEY NRLLKDHGTY       960
SKNLMWLLNF PFAYNLARYK NLSNEDLFNA KNNDQKSK                               998
```

(SEQ ID NO: 114).

[0178] Exemplary CasRX/Cas13d proteins may comprise or consist of the sequence:

CasRX/Cas13d 250twins_36050_GL0158985:

```
MKKKHQSAAE KRQVKKLKNQ EKAQKYASEP SPLQSDTAGV ECSQKKTVVS HIASSKTLAK        60
AMGLKSTLVM GDKLVITSFA ASKAVGGAGY KSANIEKITD LQGRVIEEHE RMFSADVGEK       120
```

```
NIELSKNDCH TNVNNPVVTN IGKDYIGLKS RLEQEFFGKT FENDNLHVQL AYNILDIKKI       180
LGTYVNNIIY IFYNLNRAGT GRDERMYDDL IGTLYAYKPM EAQQTYLLKG DKDMRRFEEV       240
KQLLQNTSAY VVYYGTLFEK VKAKSKKEQR AKEAEIDACT AHNYDVLRLL SLMRQLCMHS       300
VAGTAFKLAE SALFNIEDVL SADLKEILDE AFSGAVNKLN DGFVQHSGNN LYVLQQLYPN       360
ETIERIAEKY YRLTVRKEDL NMGVNIKKLR ELIVGQYFPE VLDKEYDLSK NGDSVVTYRS       420
KIYTVMNYIL LYYLEDHDSS RESMVEALRQ NREGDEGKEE IYRQFAKKVW NGVSGLFGVC       480
LNLFKTEKRN KFRSKVALPD VSGAAYMLSS ENIDYFVKML FFVCKFLDGK EINELLCALI       540
NKFDNIADIL DAAAQCGSSV WFVDSYRFFE RSRRISAQIR IVKNIASKDF KKSKKDSDES       600
YPEQLYLDAL ALLGDVISKY KQNRDGSVVI DDQGNAVLTE QYKRFRYEFF EEIKRDESGG       660
IKYKKSGKPE YNHQRRNFIL NNVLKSKWFF YVVKYNRPSS CRELMKNKEI LRFVLRDIPD       720
SQVRRYFKAV QGEEAYASAE AMRTRLVDAL SQFSVTACLD EVGGMTDKEF ASQRAVDSKE       780
KLRAIIRLYL TVAYLITKSM VKVNTRFSIA FSVLERDYYL LIDGKKKSSD YTGEDMLALT       840
RKFVGEDAGL YREWKEKNAE AKDKYFDKAE RKKVLRQNDK MIRKMHFTPH SLNYVQKNLE       900
SVQSNGLAAV IKEYRNAVAH LNIINRLDEY IGSARADSYY SLYCYCLQMY LSKNFSVGYL       960
INVQKQLEEH HTYMKDLMWL LNIPFAYNLA RYKNLSNEKL FYDEEAAAEK ADKAENERGE      1020
```

(SEQ ID NO: 115).

[0179] Yan et al. (2018) *Mol Cell.* 70(2):327-339 (doi: 10.1016/j.molcel.2018.02.2018) and Konermann et al. (2018) *Cell* 173(3):665-676 (doi: 10.1016/j.cell/2018.02.033) have described CasRX/Cas13d proteins and both of which are incorporated by reference herein in their entireties. Also see WO Publication Nos. WO2018/183703 (CasM) and WO2019/006471 (Cas13d), which are incorporated herein by reference in their entirety.

[0180] Exemplary wild type Cas13d proteins of the disclosure may comprise or consist of the amino acid sequence:

[0181] Cas13d (Ruminococcus flavefaciens XPD3002) sequence:

```
  1 IEKKKSFAKG MGVKSTLVSG SKVYMTTFAE GSDARLEKIV EGDSIRSVNE GEAFSAEMAD
 61 KNAGYKIGNA KFSHPKGYAV VANNPLYTGP VQQDMLGLKE TLEKRYFGES ADGNDNICIQ
121 VIHNILDIEK ILAEYITNAA YAVNNISGLD KDIIGFGKFS TVYTYDEFKD PEHHRAAFNN
181 NDKLINAIKA QYDEFDNFLD NPRLGYFGQA FFSKEGRNYI INYGNECYDI LALLSGLAHW
241 VVANNEEESR ISRTWLYNLD KNLDNEYIST LNYLYDRITN ELTNSFSKNS AANVNYIAET
301 LGINPAEFAE QYFRFSIMKE QKNLGFNITK LREVMLDRKD MSEIRKNHKV FDSIRTKVYT
361 MMDFVIYRYY IEEDAKVAAA NKSLPDNEKS LSEKDIFVIN LRGSFNDDQK DALYYDEANR
421 IWRKLENIMH NIKEFRGNKT REYKKKDAPR LPRILPAGRD VSAFSKLMYA LTMFLDGKEI
481 NDLLTTLINK FDNIQSFLKV MPLIGVNAKF VEEYAFFKDS AKIADELRLI KSFARMGEPI
541 ADARRAMYID AIRILGTNLS YDELKALADT FSLDENGNKL KKGKHGMRNF IINNVISNKR
601 FHYLIRYGDP AHLHEIAKNE AVVKFVLGRI ADIQKKQGQN GKNQIDRYYE TCIGKDKGKS
661 VSEKVDALTK IITGMNYDQF DKKRSVIEDT GRENAEREKF KKIISLYLTV IYHILKNIVN
721 INARYVIGFH CVERDAQLYK EKGYDINLKK LEEKGFSSVT KLCAGIDETA PDKRKDVEKE
781 MAERAKESID SLESANPKLY ANYIKYSDEK KAEEFTRQIN REKAKTALNA YLRNTKWNVI
841 IREDLLRIDN KTCTLFANKA VALEVARYVH AYINDIAEVN SYFQLYHYIM QRIIMNERYE
901 KSSGKVSEYF DAVNDEKKYN DRLLKLLCVP FGYCIPRFKN LSIEALFDRN EAAKFDKEKK
961 KVSGNS (SEQ ID NO: 45).
```

[0182] Exemplary wild type Cas13d proteins of the disclosure may comprise or consist of the amino acid sequence:

[0183] Cas13d (contig e-k87_11092736):

MKRQKTFAKRIGIKSTVAYGQGKYAITTFGKGSKAEIAVRSADPPEETLPTESDATLSIHAKFA
KAGRDGREFKCGDVDETRIHTSRSEYESLISNPAESPREDYLGLKGTLERKFFGDEYPKDNLRI
QIIYSILDIQKILGLYVEDILHFVDGLQDEPEDLVGLGLGDEKMQKLLSKALPYMGFFGSTDVF
KVTKKREERAAADEHNAKVFRALGAIRQKLAHFKWKESLAIFGANANMPIRFFQGATGGRQLWN
DVIAPLWKKRIERVRKSFLSNSAKNLWVLYQVFKDDTDEKKKARARQYYHFSVLKEGKNLGFNL
TKTREYFLDKFFPIFHSSAPDVKRKVDTFRSKFYAILDFIIYEASVSVANSGQMGKVAPWKGAI
DNALVKLREAPDEEAKEKIYNVLAASIRNDSLFLRLKSACDKFGAEQNRPVFPNELRNNRDIRN
VRSEWLEATQDVDAAAFVQLIAFLCNFLEGKEINELVTALIKKFEGIQALIDLLRNLEGVDSIR
FENEFALFNDDKGNMAGRIARQLRLLASVGKMKPDMTDAKRVLYKSALEILGAPPDEVSDEWLA
ENILLDKSNNDYQKAKKTVNPFRNYIAKNVITSRSFYYLVRYAKPTAVRKLMSNPKIVRYVLKR
LPEKQVASYYSAIWTQSESNSNEMVKLIEMIDRLTTEIAGFSFAVLKDKKDSIVSASRESRAVN
LEVERLKKLTTLYMSIAYIAVKSLVKVNARYFIAYSALERDLYFFNEKYGEEFRLHFIPYELNG
KTCQFEYLAILKYYLARDEETLKRKCEICEEIKVGCEKHKKNANPPYEYDQEWIDKKKALNSER
KACERRLHFSTHWAQYATKRDENMAKHPQKWYDILASHYDELLALQATGWLATQARNDAEHLNP
VNEFDVYIEDLRRYPEGTPKNKDYHIGSYFEIYHYIRQRAYLEEVLAKRKEYRDSGSFTDEQLD
KLQKILDDIRARGSYDKNLLKLEYLPFAYNLPRYKNLTTEALFDDDSVSGKKRVAEWREREKTR
EAEREQRRQR (SEQ ID NO: 46).

[0184] An exemplary direct repeat sequence of Cas13d (contig e-k87_11092736) (SEQ ID NO: 46) comprises or consists of the nucleic acid sequence:Cas13d (contig e-k87_11092736) Direct Repeat Sequence): GTGAGAAGTCTCCTTATGGGGAGATGCTAC (SEQ ID NO: 47).

[0185] Exemplary wild type Cas13d proteins of the disclosure may comprise or consist of the amino acid sequence:

[0186] Cas13d (160582958_gene49834):

MKNSVTFKLIQAQENKEAARKKAKDIAEQARIAKRNGVVKKEENRINRIQIEIQTQKKSNTQNA
YHLKSLAKAAGVKSVFAIGNDLLMTGFGPGNDATIEKRVFQNRAIETLSSPEQYSAEFQNKQFK
IKGNIKVLNHSTQKMEEIQTELQDNYNRPHFDLLGCKNVLEQKYFGRTFSDNIHVQIAYNIMDI
EKLLTPYINNIIYTLNELMRDNSKDDFFGCDSHFSVAYLYDELKAGYSDRLKTKPNLSKNIDRI

WNNFCNYMNSDSGNTEARLAYFGELFYKPKETGDAKSDYKTHLSNNQKEEWELKSDKEVYNIFA
ILCDLRHFCTHGESITPSGKPFPYNLEKNLFPEAKQVLNSLFEEKAESLGAEAFGKTAGKTDVS
ILLKVFEKEQASQKEQQALLKEYYDFKVQKTYKNMGFSIKKLREAIMEIPDAAKFKDDLYSSLR
HKLYGLFDFILVKHFLDTSDSENLQNNDIFRQLRACRCEEEKDQVYRSIAVKVWEKVKKKELNM
FKQVVVIPSLSKDELKQMEMTKNTELLSSIETISTQASLFSEMIFMMTYLLDGKEINLLCTSLI
EKFENIASFNEVLKSPQIGYETKYTEGYAFFKNADKTAKELRQVNNMARMTKPLGGVNTKCVMY
NEAAKILGAKPMSKAELESVFNLDNHDYTYSPSGKKIPNKNFRNFIINNVITSRRFLYLIRYGN
PEKIRKIAINPSIISFVLKQIPDEQIKRYYPPCIGKRTDDVTLMRDELGKMLQSVNFEQFSRVN
NKQNAKQNPNGEKARLQACVRLYLTVPYLFIKNMVNINARYVLAFHCLERDHALCFNSRKLNDD
SYNEMANKFQMVRKAKKEQYEKEYKCKKQETGTAHTKKIEKLNQQIAYIDKDIKNMHSYTCRNY
RNLVAHLNVVSKLQNYVSELPNDYQITSYFSFYHYCMQLGLMEKVSSKNIPLVESLKNEANDAQ
SYSAKKTLEYFDLIEKNRTYCKDFLKALNAPFSYNLPRFKNLSIEALFDKNIVYEQADLKKE
(SEQ ID NO: 48).

[0187] An exemplary direct repeat sequence of Cas13d (160582958_gene49834) (SEQ ID NO: 48) comprises or consists of the nucleic acid sequence:

[0188] Cas13d (160582958_gene49834) Direct Repeat Sequence:

GAACTACACCCCTCTGTTCTTGTAGGGGTCTAACAC (SEQ ID NO: 49).

[0189] Exemplary wild type Cas13d proteins of the disclosure may comprise or consist of the amino acid sequence:

[0190] Cas13d (contig tpg|DJXD01000002.1| ; uncultivated *Ruminococcus* assembly, UBA7013, from sheep gut metagenome):

MKKQKSKKTVSKTSGLKEALSVQGTVIMTSFGKGNMANLSYKIPSSQKPQNLNSSAGLKNVEVS
GKKIKFQGRHPKIATTDNPLFKPQPGMDLLCLKDKLEMHYFGKTFDDNIHIQLIYQILDIEKIL
AVHVNNIVFTLDNVLHPQKEELTEDFIGAGGWRINLDYQTLRGQTNKYDRFKNYIKRKELLYFG
EAFYHENERRYEEDIFAILTLLSALRQFCFHSDLSSDESDHVNSFWLYQLEDQLSDEFKETLSI
LWEEVTERIDSEFLKTNTVNLHILCHVFPKESKETIVRAYYEFLIKKSFKNMGFSIKKLREIML
EQSDLKSFKEDKYNSVRAKLYKLFDFIITYYYDHHAFEKEALVSSLRSSLTEENKEEIYIKTAR
TLASALGADFKKAAADVNAKNIRDYQKKANDYRISFEDIKIGNTGIGYFSELIYMLTLLLDGKE
INDLLTTLINKFDNIISFIDILKKLNLEFKFKPEYADFFNMTNCRYTLEELRVINSIARMQKPS
ADARKIMYRDALRILGMDNRPDEEIDRELERTMPVGADGKFIKGKQGFRNFIASNVIESSRFHY

LVRYNNPHKTRTLVKNPNVVKFVLEGIPETQIKRYFDVCKGQEIPPTSDKSAQIDVLARIISSV
DYKIFEDVPQSAKINKDDPSRNFSDALKKQRYQAIVSLYLTVMYLITKNLVYVNSRYVIAFHCL
ERDAFLHGVTLPKMNKKIVYSQLTTHLLTDKNYTTYGHLKNQKGHRKWYVLVKNNLQNSDITAV
SSFRNIVAHISVVRNSNEYISGIGELHSYFELYHYLVQSMIAKNNWYDTSHQPKTAEYLNNLKK
HHTYCKDFVKAYCIPFGYVVPRYKNLTINELFDRNNPNPEPKEEV (SEQ ID NO: 50).

[0191] An exemplary direct repeat sequence of Cas13d (contig tpg|DJXD01000002.1| ; uncultivated *Ruminococcus* assembly, UBA7013, from sheep gut metagenome) (SEQ ID NO: 50) comprises or consists of the nucleic acid sequence:Cas13d (contig tpg|DJXD01000002.1| ; uncultivated *Ruminococcus* assembly, UBA7013, from sheep gut metagenome) Direct Repeat Sequence: CAACTACAACCCCGTAAAAATACGGGGTTCTGAAAC (SEQ ID NO: 51).

gRNA Target Sequences

[0192] In some embodiments of the compositions of the disclosure, a target sequence of an RNA molecule comprises a sequence motif corresponding to the first RNA binding protein and/or the second RNA binding protein.

[0193] In some embodiments of the compositions and methods of the disclosure, the sequence motif is a signature of a disease or disorder.

[0194] A sequence motif of the disclosure may be isolated or derived from a sequence of foreign or exogenous sequence found in a genomic sequence, and therefore translated into an mRNA molecule of the disclosure or a sequence of foreign or exogenous sequence found in an RNA sequence of the disclosure.

[0195] A sequence motif of the disclosure may comprise or consist of a mutation in an endogenous sequence that causes a disease or disorder. The mutation may comprise or consist of a sequence substitution, inversion, deletion, insertion, transposition, or any combination thereof.

[0196] A sequence motif of the disclosure may comprise or consist of a repeated sequence. In some embodiments, the repeated sequence may be associated with a microsatellite instability (MSI). MSI at one or more loci results from impaired DNA mismatch repair mechanisms of a cell of the disclosure. A hypervariable sequence of DNA may be transcribed into an mRNA of the disclosure comprising a target sequence comprising or consisting of the hypervariable sequence.

[0197] A sequence motif of the disclosure may comprise or consist of a biomarker. The biomarker may indicate a risk of developing a disease or disorder. The biomarker may indicate a healthy gene (low or no determinable risk of developing a disease or disorder. The biomarker may indicate an edited gene. Exemplary biomarkers include, but are not limited to, single nucleotide polymorphisms (SNPs), sequence variations or mutations, epigenetic marks, splice acceptor sites, exogenous sequences, heterologous sequences, and any combination thereof.

[0198] A sequence motif of the disclosure may comprise or consist of a secondary, tertiary or quaternary structure. The secondary, tertiary or quaternary structure may be endogenous or naturally occurring. The secondary, tertiary or quaternary structure may be induced or non-naturally occurring. The secondary, tertiary or quaternary structure may be encoded by an endogenous, exogenous, or heterologous sequence.

[0199] In some embodiments of the compositions and methods of the disclosure, a target sequence of an RNA molecule comprises or consists of between 2 and 100 nucleotides or nucleic acid bases, inclusive of the endpoints. In some embodiments, the target sequence of an RNA molecule comprises or consists of between 2 and 50 nucleotides or nucleic acid bases, inclusive of the endpoints. In some embodiments, the target sequence of an RNA molecule comprises or consists of between 2 and 20 nucleotides or nucleic acid bases, inclusive of the endpoints.

[0200] In some embodiments of the compositions and methods of the disclosure, a target sequence of an RNA molecule is continuous. In some embodiments, the target sequence of an RNA molecule is discontinuous. For example, the target sequence of an RNA molecule may comprise or consist of one or more nucleotides or nucleic acid bases that are not contiguous because one or more intermittent nucleotides are positioned in between the nucleotides of the target sequence.

[0201] In some embodiments of the compositions and methods of the disclosure, a target sequence of an RNA molecule is naturally occurring. In some embodiments, the target sequence of an RNA molecule is non-naturally occurring. Exemplary non-naturally occurring target sequences may comprise or consist of sequence variations or mutations, chimeric sequences, exogenous sequences, heterologous sequences, chimeric sequences, recombinant sequences, sequences comprising a modified or synthetic nucleotide or any combination thereof.

[0202] In some embodiments of the compositions and methods of the disclosure, a target sequence of an RNA molecule binds to a guide RNA of the disclosure.

[0203] In some embodiments of the compositions and methods of the disclosure, a target sequence of an RNA molecule binds to a first RNA binding protein of the disclosure.

[0204] In some embodiments of the compositions and methods of the disclosure, a target sequence of an RNA molecule binds to a second RNA binding protein of the disclosure.

RNA Molecules

[0205] In some embodiments of the compositions and methods of the disclosure, an RNA molecule of the disclosure comprises a target sequence. In some embodiments, the RNA molecule of the disclosure comprises at least one target sequence. In some embodiments, the RNA molecule of the disclosure comprises one or more target sequence(s). In some embodiments, the RNA molecule of the disclosure comprises two or more target sequences.

[0206] In some embodiments of the compositions and methods of the disclosure, an RNA molecule of the disclosure is a naturally occurring RNA molecule. In some embodiments, the RNA molecule of the disclosure is a non-naturally occurring molecule. Exemplary non-naturally occurring RNA molecules may comprise or consist of sequence variations or mutations, chimeric sequences, exogenous sequences, heterologous sequences, chimeric sequences, recombinant sequences, sequences comprising a modified or synthetic nucleotide or any combination thereof.

[0207] In some embodiments of the compositions and methods of the disclosure, an RNA molecule of the disclosure comprises or consists of a sequence isolated or derived from a virus.

[0208] In some embodiments of the compositions and methods of the disclosure, an RNA molecule of the disclosure comprises or consists of a sequence isolated or derived from a prokaryotic organism. In some embodiments, an RNA molecule of the disclosure comprises or consists of a sequence isolated or derived from a species or strain of archaea or a species or strain of bacteria.

[0209] In some embodiments of the compositions and methods of the disclosure, the RNA molecule of the disclosure comprises or consists of a sequence isolated or derived from a eukaryotic organism. In some embodiments, an RNA molecule of the disclosure comprises or consists of a sequence isolated or derived from a species of protozoa, parasite, protist, algae,

fungi, yeast, amoeba, worm, microorganism, invertebrate, vertebrate, insect, rodent, mouse, rat, mammal, or a primate. In some embodiments, an RNA molecule of the disclosure comprises or consists of a sequence isolated or derived from a human.

[0210] In some embodiments of the compositions and methods of the disclosure, the RNA molecule of the disclosure comprises or consists of a sequence derived from a coding sequence from a genome of an organism or a virus. In some embodiments, the RNA molecule of the disclosure comprises or consists of a primary RNA transcript, a precursor messenger RNA (pre-mRNA) or messenger RNA (mRNA). In some embodiments, the RNA molecule of the disclosure comprises or consists of a gene product that has not been processed (e.g. a transcript). In some embodiments, the RNA molecule of the disclosure comprises or consists of a gene product that has been subject to post-transcriptional processing (e.g. a transcript comprising a 5'cap and a 3' polyadenylation signal). In some embodiments, the RNA molecule of the disclosure comprises or consists of a gene product that has been subject to alternative splicing (e.g. a splice variant). In some embodiments, the RNA molecule of the disclosure comprises or consists of a gene product that has been subject to removal of non-coding and/or intronic sequences (e.g. a messenger RNA (mRNA)).

[0211] In some embodiments of the compositions and methods of the disclosure, the RNA molecule of the disclosure comprises or consists of a sequence derived from a non-coding sequence (e.g. a non-coding RNA (ncRNA)). In some embodiments, the RNA molecule of the disclosure comprises or consists of a ribosomal RNA. In some embodiments, the RNA molecule of the disclosure comprises or consists of a small ncRNA molecule. Exemplary small RNA molecules of the disclosure include, but are not limited to, microRNAs (miRNAs), small interfering (siRNAs), piwi-interacting RNAs (piRNAs), small nucleolar RNAs (snoRNAs), small nuclear RNAs (snRNAs), extracellular or exosomal RNAs (exRNAs), and small Cajal body-specific RNAs (scaRNAs). In some embodiments, the RNA molecule of the disclosure comprises or consists of a long ncRNA molecule. Exemplary long RNA molecules of the disclosure include, but are not limited to, X-inactive specific transcript (Xist) and HOX transcript antisense RNA (HOTAIR).

[0212] In some embodiments of the compositions and methods of the disclosure, the RNA molecule of the disclosure contacted by a composition of the disclosure in an intracellular space.

In some embodiments, the RNA molecule of the disclosure contacted by a composition of the disclosure in a cytosolic space. In some embodiments, the RNA molecule of the disclosure contacted by a composition of the disclosure in a nucleus. In some embodiments, the RNA molecule of the disclosure contacted by a composition of the disclosure in a vesicle, membrane-bound compartment of a cell, or an organelle.

[0213] In some embodiments of the compositions and methods of the disclosure, the RNA molecule of the disclosure contacted by a composition of the disclosure in an extracellular space. In some embodiments, the RNA molecule of the disclosure contacted by a composition of the disclosure in an exosome. In some embodiments, the RNA molecule of the disclosure contacted by a composition of the disclosure in a liposome, a polymersome, a micelle or a nanoparticle. In some embodiments, the RNA molecule of the disclosure contacted by a composition of the disclosure in an extracellular matrix. In some embodiments, the RNA molecule of the disclosure contacted by a composition of the disclosure in a droplet. In some embodiments, the RNA molecule of the disclosure contacted by a composition of the disclosure in a microfluidic droplet.

[0214] In some embodiments of the compositions and methods of the disclosure, a RNA molecule of the disclosure comprises or consists of a single-stranded sequence. In some embodiments, the RNA molecule of the disclosure comprises or consists of a double-stranded sequence. In some embodiments, the double-stranded sequence comprises two RNA molecules. In some embodiments, the double-stranded sequence comprises one RNA molecule and one DNA molecule. In some embodiments, including those wherein the double-stranded sequence comprises one RNA molecule and one DNA molecule, compositions of the disclosure selectively bind and, optionally, selectively cut the RNA molecule.

RNA-Binding Endonucleases

[0215] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a nuclease domain. In some embodiments, the second RNA binding protein binds RNA in a manner in which it associates with RNA. In some embodiments, the second RNA binding protein associates with RNA in a manner in which it cleaves RNA.

[0216] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of an RNAse.

[0217] In some embodiments, the second RNA binding protein comprises or consists of an RNAse1. In some embodiments, the RNAse1 protein comprises or consists of: KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGLCKPVNTFVHEPLVDVQNV CFQEKVTCKNGQGNCYKSNSSMHITDCRLTNGSRYPNCAYRTSPKERHIIVACEGSPYV PVHFDASVEDST (SEQ ID NO: 20).

[0218] In some embodiments, the second RNA binding protein comprises or consists of an RNAse4. In some embodiments, the RNAse4 protein comprises or consists of: QDGMYQRFLRQHVHPEETGGSDRYCDLMMQRRKMTLYHCKRFNTFIHEDIWNIRSICS TTNIQCKNGKMNCHEGVVKVTDCRDTGSSRAPNCRYRAIASTRRVVIACEGNPQVPVH FDG (SEQ ID NO: 21).

[0219] In some embodiments, the second RNA binding protein comprises or consists of an RNAse6. In some embodiments, the RNAse6 protein comprises or consists of: WPKRLTKAHWFEIQHIQPSPLQCNRAMSGINNYTQHCKHQNTFLHDSFQNVAAVCDLL SIVCKNRRHNCHQSSKPVNMTDCRLTSGKYPQCRYSAAAQYKFFIVACDPPQKSDPPYK LVPVHLDSIL (SEQ ID NO: 22).

[0220] In some embodiments, the second RNA binding protein comprises or consists of an RNAse7. In some embodiments, the RNAse7 protein comprises or consists of: APARAGFCPLLLLLLLGLWVAEIPVSAKPKGMTSSQWFKIQHMQPSPQACNSAMKNINK HTKRCKDLNTFLHEPFSSVAATCQTPKIACKNGDKNCHQSHGPVSLTMCKLTSGKYPNC RYKEKRQNKSYVVACKPPQKKDSQQFHLVPVHLDRVL (SEQ ID NO: 23).

[0221] In some embodiments, the second RNA binding protein comprises or consists of an RNAse8. In some embodiments, the RNAse8 protein comprises or consists of: TSSQWFKTQHVQPSPQACNSAMSIINKYTERCKDLNTFLHEPFSSVAITCQTPNIACKNSC KNCHQSHGPMSLTMGELTSGKYPNCRYKEKHLNTPYIVACDPPQQGDPGYPLVPVHLD KVV (SEQ ID NO: 24).

[0222] In some embodiments, the second RNA binding protein comprises or consists of an RNAse2. In some embodiments, the RNAse2 protein comprises or consists of: KPPQFTWAQWFETQHINMTSQQCTNAMQVINNYQRRCKNQNTFLLTTFANVVNVCGN PNMTCPSNKTRKNCHHSGSQVPLIHCNLTTPSPQNISNCRYAQTPANMFYIVACDNRDQ RRDPPQYPVVPVHLDRII (SEQ ID NO: 25).

[0223] In some embodiments, the second RNA binding protein comprises or consists of an RNAse6PL. In some embodiments, the RNAse6PL protein comprises or consists of: DKRLRDNHEWKKLIMVQHWPETVCEKIQNDCRDPPDYWTIHGLWPDKSEGCNRSWPF NLEEIKKNWMEITDSSLPSPSMGPAPPRWMRSTPRRSTLAEAWNSTGSWTSTGGCALPP AALPSGDLCCRPSLTAGSRGVGVDLTALHQLLHVHYSATGIIPEECSEPTKPFQIILHHDH TEWVQSIGMPIWGTISSSESAIGKNEESQPACAVLSHDS (SEQ ID NO: 26).

[0224] In some embodiments, the second RNA binding protein comprises or consists of an RNAseL. In some embodiments, the RNAseL protein comprises or consists of: AAVEDNHLLIKAVQNEDVDLVQQLLEGGANVNFQEEEGGWTPLHNAVQMSREDIVEL LLRHGADPVLRKKNGATPFILAAIAGSVKdLLKLFLSKGADVNECDFYGFTAFMEAAVY GKVKALKFLYKRGANVNLRRKTKEDQERLRKGGATALMDAAEKGHVEVLKILLDEM GADVNACDNMGRNALIHALLSSDDSDVEAITHLLLDHGADVNVRGERGKTPLILAVEK KHLGLVQRLLEQEHIEINDTDSDGKTALLLAVELKLKKIAELLCKRGASTDCGDLVMTA RRNYDHSLVKVLLSHGAKEDFHPPAEDWKPQSSHWGAALKDLHRIYRPMIGKLKFFID EKYKIADTSEGGIYLGFYEKQEVAVKTFCEGSPRAQREVSCLQSSRENSHLVTFYGSESH RGHLFVCVTLCEQTLEACLDVHRGEDVENEEDEFARNVLSSIFKAVQELHLSCGYTHQD LQPQNILIDSKKAAHLADFDKSIKWAGDPQEVKRDLEDLGRLVLYVVKKGSISFEDLKA QSNEEVVQLSPDEETKDLIHRLFHPGEHVRDCLSDLLGHPFFWTWESRYRTLRNVGNES DIKTRKSESEILRLLQPGPSEHSKSFDKWTTKINECVMKKMNKFYEKRGNFYQNTVGDL LKFIRNLGEHIDEEKHKKMKLKIGDPSLYFQKTFPDLVIYVYTKLQNTEYRKHFPQTHSP NKPQCDGAGGASGLASPGC (SEQ ID NO: 27).

[0225] In some embodiments, the second RNA binding protein comprises or consists of an RNAseT2. In some embodiments, the RNAseT2 protein comprises or consists of: VQHWPETVCEKIQNDCRDPPDYWTIHGLWPDKSEGCNRSWPFNLEEIKDLLPEMRAYW PDVIHSFPNRSRFWKHEWEKHGTCAAQVDALNSQKKYFGRSLELYRELDLNSVLLKLGI KPSINYYQVADFKDALARVYGVIPKIQCLPPSQDEEVQTIGQIELCLTKQDQQLQNCTEP GEQPSPKQEVWLANGAAESRGLRVCEDGPVFYPPPKKTKH (SEQ ID NO: 28).

[0226] In some embodiments, the second RNA binding protein comprises or consists of an RNAse11. In some embodiments, the RNAse11 protein comprises or consists of: EASESTMKIIKEEFTDEEMQYDMAKSGQEKQTIEILMNPILLVKNTSLSMSKDDMSSTLL

TFRSLHYNDPKGNSSGNDKECCNDMTVWRKVSEANGSCKWSNNFIRSSTEVMRRVHR APSCKFVQNPGISCCESLELENTVCQFTTGKQFPRCQYHSVTSLEKILTVLTGHSLMSWL VCGSKL (SEQ ID NO: 29).

[0227] In some embodiments, the second RNA binding protein comprises or consists of an RNAseT2-like. In some embodiments, the RNAseT2-like protein comprises or consists of: XLGGADKRLRDNHEWKKLIMVQHWPETVCEKIQNDCRDPPDYWTIHGLWPDKSEGCN RSWPFNLEEIKDLLPEMRAYWPDVIHSFPNRSRFWKHEWEKHGTCAAQVDALNSQKKY FGRSLELYRELDLNSVLLKLGIKPSINYYQTTEEDLNLDVEPTTEDTAEEVTIHVLLHSAL FGEIGPRRW (SEQ ID NO: 30).

[0228] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a mutated RNAse.

[0229] In some embodiments, the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(K41R)) polypeptide. In some embodiments, the Rnase1(K41R) polypeptide comprises or consists of: KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGRCRPVNTFVHEPLVDVQNV CFQEKVTCKNGQGNCYKSNSSMHITDCRLTNGSRYPNCAYRTSPKERHIIVACEGSPYV PVHFDASVEDST (SEQ ID NO: 116).

[0230] In some embodiments, the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(K41R, D121E)) polypeptide. In some embodiments, the Rnase1 (Rnase1(K41R, D121E)) polypeptide comprises or consists of: KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGRCRPVNTFVHEPLVDVQNV CFQEKVTCKNGQGNCYKSNSSMHITDCRLTNGSRYPNCAYRTSPKERHIIVACEGSPYV PVHFEASVEDST (SEQ ID NO: 117).

[0231] In some embodiments, the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(K41R, D121E, H119N)) polypeptide. In some embodiments, the Rnase1 (Rnase1(K41R, D121E, H119N)) polypeptide comprises or consists of: KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGRCRPVNTFVHEPLVDVQNV CFQEKVTCKNGQGNCYKSNSSMHITDCRLTNGSRYPNCAYRTSPKERHIIVACEGSPYV PVNFEASVEDST (SEQ ID NO: 118).

[0232] In some embodiments, the second RNA binding protein comprises or consists of a mutated Rnase1. In some embodiments, the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(H119N)) polypeptide. In some embodiments, the Rnase1 (Rnase1(H119N)) polypeptide comprises or consists of:
KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGRCKPVNTFVHEPLVDVQNV
CFQEKVTCKNGQGNCYKSNSSMHITDCRLTNGSRYPNCAYRTSPKERHIIVACEGSPYV
PVNFDASVEDST (SEQ ID NO: 119).

[0233] In some embodiments, the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(R39D, N67D, N88A, G89D, R91D, H119N)) polypeptide. In some embodiments, the Rnase1 (Rnase1(R39D, N67D, N88A, G89D, R91D, H119N)) polypeptide comprises or consists of:
KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGDCKPVNTFVHEPLVDVQNV
CFQEKVTCKDGQGNCYKSNSSMHITDCRLTADSDYPNCAYRTSPKERHIIVACEGSPYV
PVNFDASVEDST (SEQ ID NO: 120). In some embodiments, the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(R39D, N67D, N88A, G89D, R91D, H119N)) polypeptide. In some embodiments, the Rnase1 (Rnase1(R39D, N67D, N88A, G89D, R91D, H119N, K41R, D121E)) polypeptide comprises or consists of:
KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGDCRPVNTFVHEPLVDVQNV
CFQEKVTCKDGQGNCYKSNSSMHITDCRLTADSDYPNCAYRTSPKERHIIVACEGSPYV
PVNFEASVEDST (SEQ ID NO: 121).

In some embodiments, the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(R39D, N67D, N88A, G89D, R91D, H119N)) polypeptide. In some embodiments, the Rnase1 (Rnase1(R39D, N67D, N88A, G89D, R91D)) polypeptide comprises or consists of:
KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGDCKPVNTFVHEPLVDVQNV
CFQEKVTCKDGQGNCYKSNSSMHITDCRLTADSDYPNCAYRTSPKERHIIVACEGSPYV
PVHFDASVEDST (SEQ ID NO: 122).

In some embodiments, the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1 (R39D, N67D, N88A, G89D, R91D, H119N, K41R, D121E)) polypeptide that comprises or consists of:

KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGDCRPVNTFVHEPLVDVQNV CFQEKVTCKDGQGNCYKSNSSMHITDCRLTADSDYPNCAYRTSPKERHIIVACEGSPYV PVNFEASVEDST (SEQ ID NO: 208).

[0234] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a NOB1 polypeptide. In some embodiments, the NOB1 polypeptide comprises or consists of: APVEHVVADAGAFLRHAALQDIGKNIYTIREVVTEIRDKATRRRLAVLPYELRFKEPLPE YVRLVTEFSKKTGDYPSLSATDIQVLALTYQLEAEFVGVSHLKQEPQKVKVSSSIQHPET PLHISGFHLPYKPKPPQETEKGHSACEPENLEFSSFMFWRNPLPNIDHELQELLIDRGEDV PSEEEEEEENGFEDRKDDSDDDGGGWITPSNIKQIQQELEQCDVPEDVRVGCLTTDFAM QNVLLQMGLHVLAVNGMLIREARSYILRCHGCFKTTSDMSRVFCSHCGNKTLKKVSVT V (SEQ ID NO: 31).

[0235] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of an endonuclease. In some embodiments, the second RNA binding protein comprises or consists of an endonuclease V (ENDOV). In some embodiments, the ENDOV protein comprises or consists of: AFSGLQRVGGVDVSFVKGDSVRACASLVVLSFPELEVVYEESRMVSLTAPYVSGFLAFR EVPFLLELVQQLREKEPGLMPQVLLVDGNGVLHHRGFGVACHLGVLTDLPCVGVAKKL LQVDGLENNALHKEKIRLLQTRGDSFPLLGDSGTVLGMALRSHDRSTRPLYISVGHRMS LEAAVRLTCCCCRFRIPEPVRQADICSREHIRKS (SEQ ID NO: 32).

[0236] In some embodiments, the second RNA binding protein comprises or consists of an endonuclease G (ENDOG). In some embodiments, the ENDOG protein comprises or consists of: AELPPVPGGPRGPGELAKYGLPGLAQLKSRESYVLCYDPRTRGALWVVEQLRPERLRG DGDRRECDFREDDSVHAYHRATNADYRGSGFDRGHLAAAANHRWSQKAMDDTFYLS NVAPQVPHLNQNAWNNLEKYSRSLTRSYQNVYVCTGPLFLPRTEADGKSYVKYQVIGK NHVAVPTHFFKVLILEAAGGQIELRTYVMPNAPVDEAIPLERFLVPIESIERASGLLFVPNI LARAGSLKAITAGSK (SEQ ID NO: 33).

[0237] In some embodiments, the second RNA binding protein comprises or consists of an endonuclease D1 (ENDOD1). In some embodiments, the ENDOD1 protein comprises or

consists of:

RLVGEEEAGFGECDKFFYAGTPPAGLAADSHVKICQRAEGAERFATLYSTRDRIPVYSA
FRAPRPAPGGAEQRWLVEPQIDDPNSNLEEAINEAEAITSVNSLGSKQALNTDYLDSDYQ
RGQLYPFSLSSDVQVATFTLTNSAPMTQSFQERWYVNLHSLMDRALTPQCGSGEDLYIL
TGTVPSDYRVKDKVAVPEFVWLAACCAVPGGGWAMGFVKHTRDSDIIEDVMVKDLQ
KLLPFNPQLFQNNCGETEQDTEKMKKILEVVNQIQDEERMVQSQKSSSPLSSTRSKRSTL
LPPEASEGSSSFLGKLMGFIATPFIKLFQLIYYLVVAILKNIVYFLWCVTKQVINGIESCLY
RLGSATISYFMAIGEELVSIPWKVLKVVAKVIRALLRILCCLLKAICRVLSIPVRVLVDVA
TFPVYTMGAIPIVCKDIALGLGGTVSLLFDTAFGTLGGLFQVVFSVCKRIGYKVTFDNSG
EL (SEQ ID NO: 34).

[0238] In some embodiments, the second RNA binding protein comprises or consists of a Human flap endonuclease-1 (hFEN1). In some embodiments, the hFEN1 polypeptide comprises or consists of:

MGIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSIYQFLIAVRQGGDVLQNEEGET
TSHLMGMFYRTIRMMENGIKPVYVFDGKPPQLKSGELAKRSERRAEAEKQLQQAQAAG
AEQEVEKFTKRLVKVTKQHNDECKHLLSLMGIPYLDAPSEAEASCAALVKAGKVYAAA
TEDMDCLTFGSPVLMRHLTASEAKKLPIQEFHLSRILQELGLNQEQFVDLCILLGSDYCE
SIRGIGPKRAVDLIQKHKSIEEIVRRLDPNKYPVPENWLHKEAHQLFLEPEVLDPESVELK
WSEPNEEELIKFMCGEKQFSEERIRSGVKRLSKSRQGSTQGRLDDFFKVTGSLSSAKRKE
PEPKGSTKKKAKTGAAGKFKRGK (SEQ ID NO: 35).

[0239] In some embodiments, the second RNA binding protein comprises or consists of a DNA repair endonuclease XPF (ERCC4) polypeptide. In some embodiments, the ERCC4 polypeptide comprises or consists of:

MESGQPARRIAMAPLLEYERQLVLELLDTDGLVVCARGLGADRLLYHFLQLHCHPACL
VLVLNTQPAEEEYFINQLKIEGVEHLPRRVTNEITSNSRYEVYTQGGVIFATSRILVVDFL
TDRIPSDLITGILVYRAHRIIESCQEAFILRLFRQKNKRGFIKAFTDNAVAFDTGFCHVERV
MRNLFVRKLYLWPRFHVAVNSFLEQHKPEVVEIHVSMTPTMLAIQTAILDILNACLKEL
KCHNPSLEVEDLSLENAIGKPFDKTIRHYLDPLWHQLGAKTKSLVQDLKILRTLLQYLSQ
YDCVTFLNLLESLRATEKAFGQNSGWLFLDSSTSMFINARARVVYHLPDAKMSKKEKISE
KMEIKEGEGILWG (SEQ ID NO: 124).

[0240] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of an Endonuclease III-like protein 1 (NTHL) polypeptide. In some embodiments, the NTHL polypeptide comprises or consists of:
CSPQESGMTALSARMLTRSRSLGPGAGPRGCREEPGPLRRREAAAEARKSHSPVKRPRK
AQRLRVAYEGSDSEKGEGAEPLKVPVWEPQDWQQQLVNIRAMRNKKDAPVDHLGTEH
CYDSSAPPKVRRYQVLLSLMLSSQTKDQVTAGAMQRLRARGLTVDSILQTDDATLGKLI
YPVGFWRSKVKYIKQTSAILQQHYGGDIPASVAELVALPGVGPKMAHLAMAVAWGTV
SGIAVDTHVHRIANRLRWTKKATKSPEETRAALEEWLPRELWHEINGLLVGFGQQTCLP
VHPRCHACLNQALCPAAQGL (SEQ ID NO: 123).

[0241] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a human Schlafen 14 (hSLFN14) polypeptide. In some embodiments, the hSLFN14 polypeptide comprises or consists of:
ESTHVEFKRFTTKKVIPRIKEMLPHYVSAFANTQGGYVLIGVDDKSKEVVGCKWEKVNP
DLLKKEIENCIEKLPTFHFCCEKPKVNFTTKILNVYQKDVLDGYVCVIQVEPFCCVVFAE
APDSWIMKDNSVTRLTAEQWVVMMLDTQSAPPSLVTDYNSCLISSASSARKSPGYPIKV
HKFKEALQ (SEQ ID NO: 36).

[0242] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a human beta-lactamase-like protein 2 (hLACTB2) polypeptide. In some embodiments, the hLACTB2 polypeptide comprises or consists of:
TLQGTNTYLVGTGPRRILIDTGEPAIPEYISCLKQALTEFNTAIQEIVVTHWHRDHSGGIG
DICKSINNDTTYCIKKLPRNPQREEIIGNGEQQYVYLKDGDVIKTEGATLRVLYTPGHTD
DHMALLLEEENAIFSGDCILGEGTTVFEDLYDYMNSLKELLKIKADIIYPGHGPVIHNAE
AKIQQYISHRNIREQQILTLFRENFEKSFTVMELVKIIYKNTPENLHEMAKHNLLLHLKKL
EKEGKIFSNTDPDKKWKAHL (SEQ ID NO: 37).

[0243] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of an apurinic/apyrimidinic (AP) endodeoxyribonuclease (APEX) polypeptide. In some embodiments, the second RNA binding protein comprises or consists of an apurinic/apyrimidinic (AP) endodeoxyribonuclease (APEX2) polypeptide. In some embodiments, the APEX2 polypeptide comprises or consists of:
MLRVVSWNINGIRRPLQGVANQEPSNCAAVAVGRILDELDADIVCLQETKVTRDALTEP

LAIVEGYNSYFSFSRNRSGYSGVATFCKDNATPVAAEEGLSGLFATQNGDVGCYGNMD
EFTQEELRALDSEGRALLTQHKIRTWEGKEKTLTLINVYCPHADPGRPERLVFKMRFYR
LLQIRAEALLAAGSHVIILGDLNTAHRPIDHWDAVNLECFEEDPGRKWMDSLLSNLGCQ
SASHVGPFIDSYRCFQPKQEGAFTCWSAVTGARHLNYGSRLDYVLGDRTLVIDTFQASF
LLPEVMGSDHCPVGAVLSVSSVPAKQCPPLCTRFLPEFAGTQLKILRFLVPLEQSPVLEQ
STLQHNNQTRVQTCQNKAQVRSTRPQPSQVGSSRGQKNLKSYFQPSPSCPQASPDIELPS
LPLMSALMTPKTPEEKAVAKVVKGQAKTSEAKDEKELRTSFWKSVLAGPLRTPLCGGH
REPCVMRTVKKPGPNLGRRFYMCARPRGPPTDPSSRCNFFLWSRPS (SEQ ID NO: 38).

[0244] In some embodiments, the APEX2 polypeptide comprises or consists of:
MLRVVSWNINGIRRPLQGVANQEPSNCAAVAVGRILDELDADIVCLQETKVTRDALTEP
LAIVEGYNSYFSFSRNRSGYSGVATFCKDNATPVAAEEGLSGLFATQNGDVGCYGNMD
EFTQEELRALDSEGRALLTQHKIRTWEGKEKTLTLINVYCPHADPGRPERLVFKMRFYR
LLQIRAEALLAAGSHVIILGDLNTAHRPIDHWDAVNLECFEEDPGRKWMDSLLSNLGCQ
SASHVGPFIDSYRCFQPKQEGAFTCWSAVTGARHLNYGSRLDYVLGDRTLVIDTFQASF
LLPEVMGSDHCPVGAVLSVSSVPAKQCPPLCTRFLPEFAGTQLKILRFLVPLEQSP (SEQ ID NO: 39).

[0245] In some embodiments, the second RNA binding protein comprises or consists of an apurinic or apyrimidinic site lyase (APEX1) polypeptide. In some embodiments, the APEX1 polypeptide comprises or consists of:
PKRGKKGAVAEDGDELRTEPEAKKSKTAAKKNDKEAAGEGPALYEDPPDQKTSPSGKP
ATLKICSWNVDGLRAWIKKKGLDWVKEEAPDILCLQETKCSENKLPAELQELPGLSHQ
YWSAPSDKEGYSGVGLLSRQCPLKVSYGIGDEEHDQEGRVIVAEFDSFVLVTAYVPNAG
RGLVRLEYRQRWDEAFRKFLKGLASRKPLVLCGDLNVAHEEIDLRNPKGNKKNAGFTP
QERQGFGELLQAVPLADSFRHLYPNTPYAYTFWTYMMNARSKNVGWRLDYFLLS
(SEQ ID NO: 125).

[0246] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of an angiogenin (ANG) polypeptide. In some embodiments, the ANG polypeptide comprises or consists of:
QDNSRYTHFLTQHYDAKPQGRDDRYCESIMRRRGLTSPCKDINTFIHGNKRSIKAICENK

NGNPHRENLRISKSSFQVTTCKLHGGSPWPPCQYRATAGFRNVVVACENGLPVHLDQSI FRRP (SEQ ID NO: 40).

[0247] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a heat responsive protein 12 (HRSP12) polypeptide. In some embodiments, the HRSP12 polypeptide comprises or consists of: SSLIRRVISTAKAPGAIGPYSQAVLVDRTIYISGQIGMDPSSGQLVSGGVAEEAKQALKN MGEILKAAGCDFTNVVKTTVLLADINDFNTVNEIYKQYFKSNFPARAAYQVAALPKGS RIEIEAVAIQGPLTTASL (SEQ ID NO: 41).

[0248] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a Zinc Finger CCCH-Type Containing 12A (ZC3H12A) polypeptide. In some embodiments, the ZC3H12A polypeptide comprises or consists of: GGGTPKAPNLEPPLPEEEKEGSDLRPVVIDGSNVAMSHGNKEVFSCRGILLAVNWFLER GHTDITVFVPSWRKEQPRPDVPITDQHILRELEKKKILVFTPSRRVGGKRVVCYDDRFIV KLAYESDGIVVSNDTYRDLQGERQEWKRFIEERLLMYSFVNDKFMPPDDPLGRHGPSLD NFLRKKPLTLE (SEQ ID NO: 42).

[0249] In some embodiments, the ZC3H12A polypeptide comprises or consists of: SGPCGEKPVLEASPTMSLWEFEDSHSRQGTPRPGQELAAEEASALELQMKVDFFRKLGY SSTEIHSVLQKLGVQADTNTVLGELVKHGTATERERQTSPDPCPQLPLVPRGGGTPKAP NLEPPLPEEEKEGSDLRPVVIDGSNVAMSHGNKEVFSCRGILLAVNWFLERGHTDITVFV PSWRKEQPRPDVPITDQHILRELEKKKILVFTPSRRVGGKRVVCYDDRFIVKLAYESDGI VVSNDTYRDLQGERQEWKRFIEERLLMYSFVNDKFMPPDDPLGRHGPSLDNFLRKKPL TLEHRKQPCPYGRKCTYGIKCRFFHPERPSCPQRSVADELRANALLSPPRAPSKDKNGRR PSPSSQSSSLLTESEQCSLDGKKLGAQASPGSRQEGLTQTYAPSGRSLAPSGGSGSSFGPT DWLPQTLDSLPYVSQDCLDSGIGSLESQMSELWGVRGGGPGEPGPPRAPYTGYSPYGSE LPATAAFSAFGRAMGAGHFSVPADYPPAPPAFPPREYWSEPYPLPPPTSVLQEPPVQSPG AGRSPWGRAGSLAKEQASVYTKLCGVFPPHLVEAVMGRFPQLLDPQQLAAEILSYKSQ HPSE (SEQ ID NO: 43).

[0250] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a Reactive Intermediate Imine Deaminase A (RIDA) polypeptide. In some embodiments, the RIDA polypeptidecomprises or consists of:

SSLIRRVISTAKAPGAIGPYSQAVLVDRTIYISGQIGMDPSSGQLVSGGVAEEAKQALKN
MGEILKAAGCDFTNVVKTTVLLADINDFNTVNEIYKQYFKSNFPARAAYQVAALPKGS
RIEIEAVAIQGPLTTASL (SEQ ID NO: 44).

[0251] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a Phospholipase D Family Member 6 (PDL6) polypeptide. In some embodiments, the PDL6 polypeptide comprises or consists of:
EALFFPSQVTCTEALLRAPGAELAELPEGCPCGLPHGESALSRLLRALLAARASLDLCLF
AFSSPQLGRAVQLLHQRGVRVRVVTDCDYMALNGSQIGLLRKAGIQVRHDQDPGYMH
HKFAIVDKRVLITGSLNWTTQAIQNNRENVLITEDDEYVRLFLEEFERIWEQFNPTKYTF
FPPKKSHGSCAPPVSRAGGRLLSWHRTCGTSSESQT (SEQ ID NO: 126).

[0252] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a mitochondrial ribonuclease P catalytic subunit (KIAA0391) polypeptide. In some embodiments, the KIAA0391 polypeptide comprises or consists of:
KARYKTLEPRGYSLLIRGLIHSDRWREALLLLEDIKKVITPSKKNYNDCIQGALLHQDVN
TAWNLYQELLGHDIVPMLETLKAFFDFGKDIKDDNYSNKLLDILSYLRNNQLYPGESFA
HSIKTWFESVPGKQWKGQFTTVRKSGQCSGCGKTIESIQLSPEEYECLKGKIMRDVIDGG
DQYRKTTPQELKRFENFIKSRPPFDVVIDGLNVAKMFPKVRESQLLLNVVSQLAKRNLR
LLVLGRKHMLRRSSQWSRDEMEEVQKQASCFFADDISEDDPFLLYATLHSGNHCRFITR
DLMRDHKACLPDAKTQRLFFKWQQGHQLAIVNRFPGSKLTFQRILSYDTVVQTTGDSW
HIPYDEDLVERCSCEVPTKWLCLHQKT (SEQ ID NO: 127).

[0253] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of an argonaute 2 (AGO2) polypeptide.
In some embodiments of the compositions of the disclosure, the AGO2 polypeptide comprises or consists of:
SVEPMFRHLKNTYAGLQLVVVILPGKTPVYAEVKRVGDTVLGMATQCVQMKNVQRTT
PQTLSNLCLKINVKLGGVNNILLPQGRPPVFQQPVIFLGADVTHPPAGDGKKPSIAAVVG
SMDAHPNRYCATVRVQQHRQEIIQDLAAMVRELLIQFYKSTRFKPTRIIFYRDGVSEGQF
QQVLHHELLAIREACIKLEKDYQPGITFIVVQKRHHTRLFCTDKNERVGKSGNIPAGTTV
DTKITHPTEFDFYLCSHAGIQGTSRPSHYHVLWDDNRFSSDELQILTYQLCHTYVRCTRS

VSIPAPAYYAHLVAFRARYHLVDKEHDSAEGSHTSGQSNGRDHQALAKAVQVHQDTL RTMYFA (SEQ ID NO: 128).

[0254] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a mitochondrial nuclease EXOG (EXOG) polypeptide. In some embodiments, the EXOG polypeptide comprises or consists of: QGAEGALTGKQPDGSAEKAVLEQFGFPLTGTEARCYTNHALSYDQAKRVPRWVLEHIS KSKIMGDADRKHCKFKPDPNIPPTFSAFNEDYVGSGWSRGHMAPAGNNKFSSKAMAET FYLSNIVPQDFDNNSGYWNRIEMYCRELTERFEDVWVVSGPLTLPQTRGDGKKIVSYQV IGEDNVAVPSHLYKVILARRSSVSTEPLALGAFVVPNEAIGFQPQLTEFQVSLQDLEKLSG LVFFPHLDRTSDIRNICSVDTCKLLDFQEFTLYLSTRKIEGARSVLRLEKIMENLKNAEIEP DDYFMSRYEKKLEELKAKEQSGTQIRKPS (SEQ ID NO: 129).

[0255] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a Zinc Finger CCCH-Type Containing 12D (ZC3H12D) polypeptide. In some embodiments, the ZC3H12D polypeptide comprises or consists of: EHPSKMEFFQKLGYDREDVLRVLGKLGEGALVNDVLQELIRTGSRPGALEHPAAPRLVP RGSCGVPDSAQRGPGTALEEDFRTLASSLRPIVIDGSNVAMSHGNKETFSCRGIKLAVD WFRDRGHTYIKVFVPSWRKDPPRADTPIREQHVLAELERQAVLVYTPSRKVHGKRLVC YDDRYIVKVAYEQDGVIVSNDNYRDLQSENPEWKWFIEQRLLMFSFVNDRFMPPDDPL GRHGPSLSNFLSRKPKPPEPSWQHCPYGKKCTYGIKCKFYHPERPHHAQLAVADELRAK TGARPGAGAEEQRPPRAPGGSAGARAAPREPFAHSLPPARGSPDLAALRGSFSRLAFSD DLGPLGPPLPVPACSLTPRLGGPDWVSAGGRVPGPLSLPSPESQFSPGDLPPPPGLQLQPR GEHRPRDLHGDLLSPRRPPDDPWARPPRSDRFPGRSVWAEPAWGDGATGGLSVYATED DEGDARARARIALYSVFPRDQVDRVMAAFPELSDLARLILLVQRCQSAGAPLGKP (SEQ ID NO: 130).

[0256] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of an endoplasmic reticulum to nucleus signaling 2 (ERN2) polypeptide. In some embodiments, the ERN2 polypeptide comprises or consists of: RQQQPQVVEKQQETPLAPADFAHISQDAQSLHSGASRRSQKRLQSPSKQAQPLDDPEAE QLTVVGKISFNPKDVLGRGAGGTFVFRGQFEGRAVAVKRLLRECFGLVRREVQLLQES DRHPNVLRYFCTERGPQFHYIALELCRASLQEYVENPDLDRGGLEPEVVLQQLMSGLAH

LHSLHIVHRDLKPGNILITGPDSQGLGRVVLSDFGLCKKLPAGRCSFSLHSGIPGTEGWM APELLQLLPPDSPTSAVDIFSAGCVFYYVLSGGSHPFGDSLYRQANILTGAPCLAHLEEEV HDKVVARDLVGAMLSPLPQPRPSAPQVLAHPFFWSRAKQLQFFQDVSDWLEKESEQEP LVRALEAGGCAVVRDNWHEHISMPLQTDLRKFRSYKGTSVRDLLRAVRNKKHHYREL PVEVRQALGQVPDGFVQYFTNRFPRLLLHTHRAMRSCASESLFLPYYPPDSEARRPCPG ATGR (SEQ ID NO: 131).

[0257] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a pelota mRNA surveillance and ribosome rescue factor (PELO) polypeptide. In some embodiments, the PELO polypeptide comprises or consists of: KLVRKNIEKDNAGQVTLVPEEPEDMWHTYNLVQVGDSLRASTIRKVQTESSTGSVGSN RVRTTLTLCVEAIDFDSQACQLRVKGTNIQENEYVKMGAYHTIELEPNRQFTLAKKQW DSVVLERIEQACDPAWSADVAAVVMQEGLAHICLVTPSMTLTRAKVEVNIPRKRKGNC SQHDRALERFYEQVVQAIQRHIHFDVVKCILVASPGFVREQFCDYLFQQAVKTDNKLLL ENRSKFLQVHASSGHKYSLKEALCDPTVASRLSDTKAAGEVKALDDFYKMLQHEPDRA FYGLKQVEKANEAMAIDTLLISDELFRHQDVATRSRYVRLVDSVKENAGTVRIFSSLHV SGEQLSQLTGVAAILRFPVPELSDQEGDSSSEED (SEQ ID NO: 132).

[0258] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a YBEY metallopeptidase (YBEY) polypeptide. In some embodiments, the YBEY polypeptide comprises or consists of: SLVIRNLQRVIPIRRAPLRSKIEIVRRILGVQKFDLGIICVDNKNIQHINRIYRDRNVPTDVL SFPFHEHLKAGEFPQPDFPDDYNLGDIFLGVEYIFHQCKENEDYNDVLTVTATHGLCHLL GFTHGTEAEWQQMFQKEKAVLDELGRRTGTRLQPLTRGLFGGS (SEQ ID NO: 133).

[0259] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a cleavage and polyadenylation specific factor 4 like (CPSF4L) polypeptide. In some embodiments, the CPSF4L polypeptide comprises or consists of: QEVIAGLERFTFAFEKDVEMQKGTGLLPFQGMDKSASAVCNFFTKGLCEKGKLCPFRH DRGEKMVVCKHWLRGLCKKGDHCKFLHQYDLTRMPECYFYSKFGDCSNKECSFLHVK PAFKSQDCPWYDQGFCKDGPLCKYRHVPRIMCLNYLVGFCPEGPKCQFAQKIREFKLLP GSKI (SEQ ID NO: 134).

[0260] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of an hCG_2002731 polypeptide. In some embodiments, the hCG_2002731 polypeptide comprises or consists of:
KLVRKNIEKDNAGQVTLVPEEPEDMWHTYNLVQVGDSLRASTIRKVQTESSTGSVGSN
RVRTTLTLCVEAIDFDSQACQLRVKGTNIQENEYVKMGAYHTIELEPNRQFTLAKKQW
DSVVLERIEQACDPAWSADVAAVVMQEGLAHICLVTPSMTLTRAKVEVNIPRKRKGNC
SQHDRALERFYEQVVQAIQRHIHFDVVKCILVASPGFVREQFCDYMFQQAVKTDNKLLL
ENRSKFLQVHASSGHKYSLKEALCDPTVASRLSDTKAAGEVKALDDFYKMLQHEPDRA
FYGLKQVEKANEAMAIDTLLISDELFRHQDVATRSRYVRLVDSVKENAGTVRIFSSLHV
SGEQLSQLTGVAAILRFPVPELSDQEGDSSSEED (SEQ ID NO: 135).

[0261] In some embodiments, the hCG_2002731 polypeptide comprises or consists of:
DPAWSADVAAVVMQEGLAHICLVTPSMTLTRAKVEVNIPRKRKGNCSQHDRALERFYE
QVVQAIQRHIHFDVVKCILVASPGFVREQFCDYMFQQAVKTDNKLLLENRSKFLQVHAS
SGHKYSLKEALCDPTVASRLSDTKAAGEVKALDDFYKMLQHEPDRAFYGLKQVEKAN
EAMAIDTLLISDELFRHQDVATRSRYVRLVDSVKENAGTVRIFSSLHVSGEQLSQLTGVA
AILRFPVPELSDQEGDSSSEED (SEQ ID NO: 136).

[0262] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of an Excision Repair Cross-Complementation Group 1 (ERCC1) polypeptide. In some embodiments, the ERCC1 polypeptide comprises or consists of:
MDPGKDKEGVPQPSGPPARKKFVIPLDEDEVPPGVRGNPVLKFVRNVPWEFGDVIPDYV
LGQSTCALFLSLRYHNLHPDYIHGRLQSLGKNFALRVLLVQVDVKDPQQALKELAKMC
ILADCTLILAWSPEEAGRYLETYKAYEQKPADLLMEKLEQDFVSRVTECLTTVKSVNKT
DSQTLLTTFGSLEQLIAASREDLALCPGLGPQK (SEQ ID NO: 137).

[0263] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a ras-related C3 botulinum toxin substrate 1 isoform (RAC1) polypeptide. In some embodiments, the RAC1 polypeptide comprises or consists of:
KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGRCKPVNTFVHEPLVDVQNV
CFQEKVTCKNGQGNCYKSNSSMHITDCRLTNGSRYPNCAYRTSPKERHIIVACEGSPYV
PVHFDASVEDST (SEQ ID NO: 138).

[0264] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a Ribonuclease A A1 (RAA1) polypeptide. In some embodiments, the RAA1 polypeptide comprises or consists of: QDNSRYTHFLTQHYDAKPQGRDDRYCESIMRRRGLTSPCKDINTFIHGNKRSIKAICENK NGNPHRENLRISKSSFQVTTCKLHGGSPWPPCQYRATAGFRNVVVACENGLPVHLDQSI FRRP (SEQ ID NO: 139).

[0265] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a Ras Related Protein (RAB1) polypeptide. In some embodiments, the RAB1 polypeptide comprises or consists of: GLGLVQPSYGQDGMYQRFLRQHVHPEETGGSDRYCNLMMQRRKMTLYHCKRFNTFIH EDIWNIRSICSTTNIQCKNGKMNCHEGVVKVTDCRDTGSSRAPNCRYRAIASTRRVVIAC EGNPQVPVHFDG (SEQ ID NO: 140).

[0266] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a DNA Replication Helicase/Nuclease 2 (DNA2) polypeptide. In some embodiments, the DNA2 polypeptide comprises or consists of: XSAVDNILLKLAKFKIGFLRLGQIQKVHPAIQQFTEQEICRSKSIKSLALLEELYNSQLIVA TTCMGINHPIFSRKIFDFCIVDEASQISQPICLGPLFFSRRFVLVGDHQQLPPLVLNREARA LGMSESLFKRLEQNKSAVVQLTVQYRMNSKIMSLSNKLTYEGKLECGSDKVANAVINL RHFKDVKLELEFYADYSDNPWLMGVFEPNNPVCFLNTDKVPAPEQVEKGGVSNVTEA KLIVFLTSIFVKAGCSPSDIGIIAPYRQQLKIINDLLARSIGMVEVNTVDKYQGRDKSIVLV SFVRSNKDGTVGELLKDWRRLNVAITRAKHKLILLGCVPSLNCYPPLEKLLNHLNSEKLI SFFFCIWSHLIALL (SEQ ID NO: 141).

[0267] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a FLJ35220 polypeptide. In some embodiments, the FLJ35220 polypeptide comprises or consists of: MALRSHDRSTRPLYISVGHRMSLEAAVRLTCCCCRFRIPEPVRQADICSREHIRKSLGLP GPPTPRSPKAQRPVACPKGDSGESSALC (SEQ ID NO: 142).

[0268] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a FLJ13173 polypeptide. In some embodiments, the FLJ13173 polypeptide comprises or consists of:

CYTNHALSYDQAKRVPRWVLEHISKSKIMGDADRKHCKFKPDPNIPPTFSAFNEDYVGS GWSRGHMAPAGNNKFSSKAMAETFYLSNIVPQDFDNNSGYWNRIEMYCRELTERFEDV WVVSGPLTLPQTRGDGKKIVSYQVIGEDNVAVPSHLYKVILARRSSVSTEPLALGAFVV PNEAIGFQPQLTEFQVSLQDLEKLSGLVFFPHLDRT (SEQ ID NO: 143).

[0269] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of Teneurin Transmembrane Protein (TENM) polypeptide. In some embodiments, the second RNA binding protein comprises or consists of Teneurin Transmembrane Protein 1 (TENM1) polypeptide. In some embodiments, the TENM1 polypeptide comprises or consists of:

VTVSQMTSVLNGKTRRFADIQLQHGALCFNIRYGTTVEEEKNHVLEIARQRAVAQAWT KEQRRLQEGEEGIRAWTEGEKQQLLSTGRVQGYDGYFVLSVEQYLELSDSANNIHFMR QSEIGRR (SEQ ID NO: 144).

In some embodiments, the second RNA binding protein comprises or consists of Teneurin Transmembrane Protein 2 (TENM2) polypeptide. In some embodiments, the TENM2 polypeptide comprises or consists of:

TVSQPTLLVNGKTRRFTNIEFQYSTLLLSIRYGLTPDTLDEEKARVLDQARQRALGTAW AKEQQKARDGREGSRLWTEGEKQQLLSTGRVQGYEGYYVLPVEQYPELADSSSNIQFL RQNEMGKR (SEQ ID NO: 145).

In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a Ribonuclease Kappa (RNAseK) polypeptide. In some embodiments, the RNAseK polypeptide comprises or consists of:

MGWLRPGPRPLCPPARASWAFSHRFPSPLAPRRSPTPFFMASLLCCGPKLAACGIVLSA WGVIMLIMLGIFFNVHSAVLIEDVPFTEKDFENGPQNIYNLYEQVSYNCFIAAGLYLLLG GFSFCQVRLNKRKEYMVR (SEQ ID NO: 204).

[0270] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a transcription activator-like effector nuclease (TALEN) polypeptide or a nuclease domain thereof. In some embodiments, the TALEN polypeptide comprises or consists of:

```
  1 MRIGKSSGWL NESVSLEYEH VSPPTRPRDT RRRPRAAGDG GLAHLHRRLA VGYAEDTPRT
 61 EARSPAPRRP LPVAPASAPP APSLVPEPPM PVSLPAVSSP RFSAGSSAAI TDPFPSLPPT
121 PVLYAMAREL EALSDATWQP AVPLPAEPPT DARRGNTVFD EASASSPVIA SACPQAFASP
181 PRAPRSARAR RARTGGDAWP APTFLSRPSS SRIGRDVFGK LVALGYSREQ IRKLKQESLS
```

```
 241 EIAKYHTTLT GQGFTHADIC RISRRRQSLR VVARNYPELA AALPELTRAH IVDIARQRSG
 301 DLALQALLPV ATALTAAPLR LSASQIATVA QYGERPAIQA LYRLRRKLTR APLHLTPQQV
 361 VAIASNTGGK RALEAVCVQL PVLRAAPYRL STEQVVAIAS NKGGKQALEA VKAHLLDLLG
 421 APYVLDTEQV VAIASHNGGK QALEAVKADL LDLRGAPYAL STEQVVAIAS HNGGKQALEA
 481 VKADLLELRG APYALSTEQV VAIASHNGGK QALEAVKAHL LDLRGVPYAL STEQVVAIAS
 541 HNGGKQALEA VKAQLLDLRG APYALSTAQV VAIASNGGGK QALEGIGEQL LKLRTAPYGL
 601 STEQVVAIAS HDGGKQALEA VGAQLVALRA APYALSTEQV VAIASNKGGK QALEAVKAQL
 661 LELRGAPYAL STAQVVAIAS HDGGNQALEA VGTQLVALRA APYALSTEQV VAIASHDGGK
 721 QALEAVGAQL VALRAAPYAL NTEQVVAIAS SHGGKQALEA VRALFPDLRA APYALSTAQL
 781 VAIASNPGGK QALEAVRALF RELRAAPYAL STEQVVAIAS NHGGKQALEA VRALFRGLRA
 841 APYGLSTAQV VAIASSNGGK QALEAVWALL PVLRATPYDL NTAQIVAIAS HDGGKPALEA
 901 VWAKLPVLRG APYALSTAQV VAIACISGQQ ALEAIEAHMP TLRQASHSLS PERVAAIACI
 961 GGRSAVEAVR QGLPVKAIRR IRREKAPVAG PPPASLGPTP QELVAVLHFF RAHQQPRQAF
1021 VDALAAFQAT RPALLRLLSS VGVTEIEALG GTIPDATERW QRLLGRLGFR PATGAAAPSP
1081 DSLQGFAQSL ERTLGSPGMA GQSACSPHRK RPAETAIAPR SIRRSPNNAG QPSEPWPDQL
1141 AWLQRRKRTA RSHIRADSAA SVPANLHLGT RAQFTPDRLR AEPGPIMQAH TSPASVSFGS
1201 HVAFEPGLPD PGTPTSADLA SFEAEPFGVG PLDFHLDWLL QILET(SEQ ID NO: 205).
```

In some embodiments, the TALEN polypeptide comprises or consists of:

```
   1 mdpirsrtps parellpgpq pdrvqptadr ggappaggpl dglparrtms rtrlpsppap
  61 spafsagsfs dllrqfdpsl ldtslldsmp avgtphtaaa paecdevqsg lraaddpppt
 121 vrvavtaarp prakpaprrr aaqpsdaspa aqvdlrtlgy sqqqqekikp kvgstvaqhh
 181 ealvghgfth ahivalsrhp aalgtvavky qdmiaalpea thedivgvgk qwsgaralea
 241 lltvagelrg pplqldtgql vkiakrggvt aveavhasrn altgaplnlt paqvvaiasn
 301 nggkqaletv qrllpvlcqa hgltpaqvva iashdggkqa letmqrllpv lcqahglppd
 361 qvvaiasnig gkqaletvqr llpvlcqahg ltpdqvvaia shgggkqale tvqrllpvlc
 421 qahgltpdqv vaiashdggk qaletvqrll pvlcqahglt pdqvvaiasn gggkqaletv
 481 qrllpvlcqa hgltpdqvva iasnggkqal etvqrllpvl cqahgltpdq vvaiashdgg
 541 kqaletvqrl lpvlcqthgl tpaqvvaias hdggkqalet vqqllpvlcq ahgltpdqvv
 601 aiasniggkq alatvqrllp vlcqahgltp dqvvaiasng ggkqaletvq rllpvlcqah
 661 gltpdqvvai asngggkqal etvqrllpvl cqahgltqvq vvaiasnigg kqaletvqrl
 721 lpvlcqahgl tpaqvvaias hdggkqalet vqrllpvlcq ahgltpdqvv aiasngggkq
 781 aletvqrllp vlcqahgltq eqvvaiasnn ggkqaletvq rllpvlcqah gltpdqvvai
 841 asngggkqal etvqrllpvl cqahgltpaq vvaiasnigg kqaletvqrl lpvlcqdhgl
 901 tlaqvvaias niggkqalet vqrllpvlcq ahgltqdqvv aiasniggkq aletvqrllp
 961 vlcqdhgltp dqvvaiasni ggkqaletvq rllpvlcqdh gltldqvvai asnggkqale
1021 tvqrllpvlc qdhgltpdqv vaiasnsggk qaletvqrll pvlcqdhglt pnqvvaiasn
1081 ggkqalesiv aqlsrpdpal aaltndhlva laclggrpam davkkglpha pelirrvnrr
1141 igertshrva dyaqvvrvle ffqchshpay afdeamtqfg msrnglvqlf rrvgvtelea
1201 rggtlppasq rwdrilqasg mkrakpspts aqtpdqaslh afadslerdl dapspmhegd
1261 qtgassrkrs rsdravtgps aqhsfevrvp eqrdalhlpl swrvkrprtr iggglpdpgt
1321 piaadlaass tvmweqdaap fagaaddfpa fneeelawlm ellpqsgsvg gti (SEQ ID
NO: 206).
```

[0271] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists a zinc finger nuclease polypeptide or a nuclease domain thereof. In some embodiments, the second RNA binding protein comprises or consists of a ZNF638 polypeptide or a nuclease domain thereof. In some embodiments, the ZNF638 polypeptide polypeptide comprises or consists of:

```
   1 MSRPRFNPRG DFPLQRPRAP NPSGMRPPGP FMRPGSMGLP RFYPAGRARG IPHRFAGHES
  61 YQNMGPQRMN VQVTQHRTDP RLTKEKLDFH EAQQKKGKPH GSRWDDEPHI SASVAVKQSS
```

```
 121 VTQVTEQSPK VQSRYTKESA SSILASFGLS NEDLEELSRY PDEQLTPENM PLILRDIRMR
 181 KMGRRLPNLP SQSRNKETLG SEAVSSNVID YGHASKYGYT EDPLEVRIYD PEIPTDEVEN
 241 EFQSQQNISA SVPNPNVICN SMFPVEDVFR QMDFPGESSN NRSFFSVESG TKMSGLHISG
 301 GQSVLEPIKS VNQSINQTVS QTMSQSLIPP SMNQQPFSSE LISSVSQQER IPHEPVINSS
 361 NVHVGSRGSK KNYQSQADIP IRSPFGIVKA SWLPKFSHAD AQKMKRLPTP SMMNDYYAAS
 421 PRIFPHLCSL CNVECSHLKD WIQHQNTSTH IESCRQLRQQ YPDWNPEILP SRRNEGNRKE
 481 NETPRRRSHS PSPRRSRRSS SSHRFRRSRS PMHYMYRPRS RSPRICHRFI SRYRSRSRSR
 541 SPYRIRNPFR GSPKCFRSVS PERMSRRSVR SSDRKKALED VVQRSGHGTE FNKQKHLEAA
 601 DKGHSPAQKP KTSSGTKPSV KPTSATKSDS NLGGHSIRCK SKNLEDDTLS ECKQVSDKAV
 661 SLQRKLRKEQ SLHYGSVLLI TELPEDGCTE EDVRKLFQPF GKVNDVLIVP YRKEAYLEME
 721 FKEAITAIMK YIETTPLTIK GKSVKICVPG KKKAQNKEVK KKTLESKKVS ASTLKRDADA
 781 SKAVEIVTST SAAKTGQAKA SVAKVNKSTG KSASSVKSVV TVAVKGNKAS IKTAKSGGKK
 841 SLEAKKTGNV KNKDSNKPVT IPENSEIKTS IEVKATENCA KEAISDAALE ATENEPLNKE
 901 TEEMCVMLVS NLPNKGYSVE EVYDLAKPFG GLKDILILSS HKKAYIEINR KAAESMVKFY
 961 TCFPVLMDGN QLSISMAPEN MNIKDEEAIF ITLVKENDPE ANIDTIYDRF VHLDNLPEDG
1021 LQCVLCVGLQ FGKVDHHVFI SNRNKAILQL DSPESAQSMY SFLKQNPQNI GDHMLTCSLS
1081 PKIDLPEVQI EHDPELEKES PGLKNSPIDE SEVQTATDSP SVKPNELEEE STPSIQTETL
1141 VQQEEPCEEE AEKATCDSDF AVETLELETQ GEEVKEEIPL VASASVSIEQ FTENAEECAL
1201 NQQMFNSDLE KKGAEIINPK TALLPSDSVF AEERNLKGIL EESPSEAEDF ISGITQTMVE
1261 AVAEVEKNET VSEILPSTCI VTLVPGIPTG DEKTVDKKNI SEKKGNMDEK EEKEFNTKET
1321 RMDLQIGTEK AEKNEGRMDA EKVEKMAAMK EKPAENTLFK AYPNKGVGQA NKPDETSKTS
1381 ILAVSDVSSS KPSIKAVIVS SPKAKATVSK TENQKSFPKS VPRDQINAEK KLSAKEFGLL
1441 KPTSARSGLA ESSSKFKPTQ SSLTRGGSGR ISALQGKLSK LDYRDITKQS QETEARPSIM
1501 KRDDSNNKTL AEQNTKNPKS TTGRSSKSKE EPLFPFNLDE FVTVDEVIEE VNPSQAKQNP
1561 LKGKRKETLK NVPFSELNLK KKKGKTSTPR GVEGELSFVT LDEIGEEEDA AAHLAQALVT
1621 VDEVIDEEEL NMEEMVKNSN SLFTLDELID QDDCISHSEP KDVTVLSVAE EQDLLKQERL
1681 VTVDEIGEVE ELPLNESADI TFATLNTKGN EGDTVRDSIG FISSQVPEDP STLVTVDEIQ
1741 DDSSDLHLVT LDEVTEEDED SLADFNNLKE ELNFVTVDEV GEEEDGDNDL KVELAQSKND
1801 HPTDKKGNRK KRAVDTKKTK LESLSQVGPV NENVMEEDLK TMIERHLTAK TPTKRVRIGK
1861 TLPSEKAVVT EPAKGEEAFQ MSEVDEESGL KDSEPERKRK KTEDSSSGKS VASDVPEELD
1921 FLVPKAGFFC PICSLFYSGE KAMTNHCKST RHKQNTEKFM AKQRKEKEQN EAEERSSR
```

(SEQ ID NO: 207).

[0272] In some embodiments of the compositions of the disclosure, the second RNA binding protein comprises or consists of a PIN domain derived from the human SMG6 protein, also commonly known as telomerase-binding protein EST1A isoform 3, NCBI Reference Sequence: NP_001243756.1. In some embodiments, the PIN from hSMG6 is used herein in the form of a Cas fusion protein and as an internal control, for example, and without limitation, see Figure 9, which shows PIN-dSauCas9, PIN-dSauCas9dHNH, PIN-dSPCas9, and dcjeCas9-PIN.

[0273] In some embodiments of the compositions of the disclosure, the composition further comprises (a) a sequence comprising a gRNA that specifically binds within an RNA molecule and (b) a sequence encoding a nuclease. In some embodiments, a nuclease comprises a sequence isolated or derived from a CRISPR/Cas protein. In some embodiments, the CRISPR/Cas protein is isolated or derived from any one of a type I, a type IA, a type IB, a type IC, a type ID, a type IE, a type IF, a type IU, a type III, a type IIIA, a type IIIB, a type IIIC, a type IIID, a type IV, a

type IVA, a type IVB, a type II, a type IIA, a type IIB, a type IIC, a type V, or a type VI CRISPR/Cas protein .In some embodiments, a nuclease comprises a sequence isolated or derived from a TALEN or a nuclease domain thereof. In some embodiments, a nuclease comprises a sequence isolated or derived from a zinc finger nuclease or a nuclease domain thereof.

Fusion Proteins

[0274] In some embodiments of the compositions and methods of the disclosure, the composition comprises a sequence encoding a target RNA-binding fusion protein comprising (a) a sequence encoding a first RNA-binding polypeptide or portion thereof; and (b) a sequence encoding a second RNA-binding polypeptide, wherein the first RNA-biding polypeptide binds a target RNA, and wherein the second RNA-binding polypeptide comprises RNA-nuclease activity.

[0275] In some embodiments, a target RNA-binding fusion protein is an RNA-guided target RNA-binding fusion protein. RNA-guided target RNA-binding fusion proteins comprise at least one RNA-binding polypeptide which corresponds to a gRNA which guides the RNA-binding polypeptide to target RNA. RNA-guided target RNA-binding fusion proteins include without limitation, RNA-binding polypeptides which are CRISPR/Cas-based RNA-binding polypeptides or portions thereof.

[0276] In some embodiments, a target RNA-binding fusion protein is not an RNA-guided target RNA-binding fusion protein and as such comprises at least one RNA-binding polypeptide which is capable of binding a target RNA without a corresponding gRNA sequence. Such non-guided RNA-binding polypeptides include, without limitation, at least one RNA-binding protein or RNA-binding portion thereof which is a PUF (Pumilio and FBF homology family). This type RNA-binding polypeptide can be used in place of a gRNA-guided RNA binding protein such as CRISPR/Cas. The unique RNA recognition mode of PUF proteins (named for Drosophila Pumilio and C. elegans fem-3 binding factor) that are involved in mediating mRNA stability and translation are well known in the art. The PUF domain of human Pumilio1, also known in the art, binds tightly to cognate RNA sequences and its specificity can be modified. It contains eight PUF repeats that recognize eight consecutive RNA bases with each repeat recognizing a single base. Since two amino acid side chains in each repeat recognize the Watson-Crick edge of the

corresponding base and determine the specificity of that repeat, a PUF domain can be designed to specifically bind most 8-nt RNA. *Wang et al., Nat Methods. 2009; 6(11): 825-830.* See also WO2012/068627 which is incorporated by reference herein in its entirety.

[0277] In some embodiments of the non-guided RNA-binding fusion proteins of the disclosure, the fusion protein comprises at least one RNA-binding protein or RNA-binding portion thereof which is a PUMBY (Pumilio-based assembly) protein. RNA-binding protein PumHD (Pumilio homology domain, a member of the PUF family), which has been widely used in native and modified form for targeting RNA, has been engineered to yield a set of four canonical protein modules, each of which targets one RNA base. These modules (*i.e.*, Pumby, for Pumilio-based assembly) can be concatenated in chains of varying composition and length, to bind desired target RNAs. The specificity of such Pumby–RNA interactions is high, with undetectable binding of a Pumby chain to RNA sequences that bear three or more mismatches from the target sequence. *Katarzyna et al., PNAS, 2016; 113(19): E2579-E2588.* See also US 2016/0238593 which is incorporated by reference herein in its entirety.

[0278] In some embodiments of the compositions of the disclosure, the first RNA binding protein comprises a Pumilio and FBF (PUF) protein. In some embodiments, the first RNA binding protein comprises a Pumilio-based assembly (PUMBY) protein. In some embodiments, a PUF1 protein of the disclosure comprises or consists of the amino acid sequence of

```
MDKSKQMNIN NLSNIPEVID PGITIPIYEE EYENNGESNS QLQQQPQKLG SYRSRAGKFS        60
NTLSNLLPSI SAKLHHSKKN SHGKNGAEFS SSNNSSQSTV ASKTPRASPS RSKMMESSID       120
GVTMDRPGSL TPPQDMEKLV HFPDSSNNFL IPAPRGSSDS FNLPHQISRT RNNTMSSQIT       180
SISSIAPKPR TSSGIWSSNA SANDPMQQHL LQQLQPTTSN NTTNSNTLND YSTKTAYFDN       240
MVSTSGSQMA DNKMNTNNLA IPNSVWSNTR QRSQSNASSI YTDAPLYEQP ARASISSHYT       300
IPTQESPLIA DEIDPQSINW VTMDPTVPSI NQISNLLPTN TISISNVFPL QHQQPQLNNA       360
INLTSTSLAT LCSKYGEVIS ARTLRNLNMA LVEFSSVESA VKALDSLQGK EVSMIGAPSK       420
ISFAKILPMH QQPPQFLLNS QGLPLGLENN NLQPQPLLQE QLFNGAVTFQ QQGNVSIPVF       480
NQQSQQSQHQ NHSSGSAGFS NVLHGYNNNN SMHGNNNNSA NEKEQCPFPL PPPNVNEKED       540
LLREIIELFE ANSDEYQINS LIKKSLNHKG TSDTQNFGPL PEPLSGREFD PPKLRELRKS       600
IDSNAFSDLE IEQLAIAMLD ELPELSSDYL GNTIVQKLFE HSSDIIKDIM LRKTSKYLTS       660
MGVHKNGTWA CQKMITMAHT PRQIMQVTQG VKDYCTPLIN DQFGNYVIQC VLKFGFPWNQ       720
FIFESIIANF WVIVQNRYGA RAVRACLEAH DIVTPEQSIV LSAMIVTYAE YLSTNSNGAL       780
LVTWFLDTSV LPNRHSILAP RLTKRIVELC GHRLASLTIL KVLNYRGDDN ARKIILDSLF       840
GNVNAHDSSP PKELTKLLCE TNYGPTFVHK VLAMPLLEDD LRAHIIKQVR KVLTDSTQIQ       900
PSRRLLEEVG LASPSSTHNK TKQQQQQHHN SSISHMFATP DTSGQHMRGL SVSSVKSGGS       960
KHTTMNTTTT NGSSASTLSP GQPLNANSNS SMGYFSYPGV FPVSGFSGNA SNGYAMNNDD      1020
LSSQFDMLNF NNGTRLSLPQ LSLTNHNNTT MELVNNVGSS QPHTNNNNNN NNTNYNDDNT      1080
VFETLTLHSA N                                                           1091
(SEQ ID NO: 209).
```

In some embodiments, a PUF3 protein of the disclosure comprises or consists of the amino acid sequence of

```
  1 MEMNMDMDMD MELASIVSSL SALSHSNNNG GQAAAAGIVN GGAAGSQQIG GFRRSSFTTA
 61 NEVDSEILLL HGSSESSPIF KKTALSVGTA PPFSTNSKKF FGNGGNYYQY RSTDTASLSS
121 ASYNNYHTHH TAANLGKNNK VNHLLGQYSA SIAGPVYYNG NDNNNSGGEG FFEKFGKSLI
181 DGTRELESQD RPDAVNTQSQ FISKSVSNAS LDTQNTFEQN VESDKNFNKL NRNTTNSGSL
241 YHSSSNSGSS ASLESENAHY PKRNIWNVAN TPVFRPSNNP AAVGATNVAL PNQQDGPANN
301 NFPPYMNGFP PNQFHQGPHY QNFPNYLIGS PSNFISQMIS VQIPANEDTE DSNGKKKKKA
361 NRPSSVSSPS SPPNNSPFPF AYPNPMMFMP PPPLSAPQQQ QQQQQQQQQE DQQQQQQQEN
421 PYIYYPTPNP IPVKMPKDEK TFKKRNNKNH PANNSNNANK QANPYLENSI PTKNTSKKNA
481 SSKSNESTAN NHKSHSHSHP HSQSLQQQQQ TYHRSPLLEQ LRNSSSDKNS NSNMSLKDIF
541 GHSLEFCKDQ HGSRFIQREL ATSPASEKEV IFNEIRDDAI ELSNDVFGNY VIQKFFEFGS
601 KIQKNTLVDQ FKGNMKQLSL QMYACRVIQK ALEYIDSNQR IELVLELSDS VLQMIKDQNG
661 NHVIQKAIET IPIEKLPFIL SSLTGHIYHL STHSYGCRVI QRLLEFGSSE DQESILNELK
721 DFIPYLIQDQ YGNYVIQYVL QQDQFTNKEM VDIKQEIIET VANNVVEYSK HKFASNVVEK
781 SILYGSKNQK DLIISKILPR DKNHALNLED DSPMILMIKD QFANYVIQKL VNVSEGEGKK
841 LIVIAIRAYL DKLNKSNSLG NRHLASVEKL AALVENAEV (SEQ ID NO: 210).
```

In some embodiments, a PUF4 protein of the disclosure comprises or consists of the amino acid sequence of

```
  1 MSTKGLKEEI DDVPSVDPVV SETVNSALEQ LQLDDPEENA TSNAFANKVS QDSQFANGPP
 61 SQMFPHPQMM GGMGFMPYSQ MMQVPHNPCP FFPPPDFNDP TAPLSSSPLN AGGPPMLFKN
121 DSLPFQMLSS GAAVATQGGQ NLNPLINDNS MKVLPIASAD PLWTHSNVPG SASVAIEETT
181 ATLQESLPSK GRESNNKASS FRRQTFHALS PTDLINAANN VTLSKDFQSD MQNFSKAKKP
241 SVGANNTAKT RTQSISFDNT PSSTSFIPPT NSVSEKLSDF KIETSKEDLI NKTAPAKKES
301 PTTYGAAYPY GGPLLQPNPI MPGHPHNISS PIYGIRSPFP NSYEMGAQFQ PFSPILNPTS
361 HSLNANSPIP LTQSPIHLAP VLNPSSNSVA FSDMKNDGGK PTTDNDKAGP NVRMDLINPN
421 LGPSMQPFHI LPPQQNTPPP PWLYSTPPPF NAMVPPHLLA QNHMPLMNSA NNKHHGRNNN
481 SMSSHNDNDN IGNSNYNNKD TGRSNVGKMK NMKNSYHGYY NNNNNNNNNN NNNNNSNATN
541 SNSAEKQRKI EESSRFADAV LDQYIGSIHS LCKDQHGCRF LQKQLDILGS KAADAIFEET
601 KDYTVELMTD SFGNYLIQKL LEEVTTEQRI VLTKISSPHF VEISLNPHGT RALQKLIECI
661 KTDEEAQIVV DSLRPYTVQL SKDLNGNHVI QKCLQRLKPE NFQFIFDAIS DSCIDIATHR
721 HGCCVLQRCL DHGTTEQCDN LCDKLLALVD KLTLDPFGNY VVQYIITKEA EKNKYDYTHK
781 IVHLLKPRAI ELSIHKFGSN VIEKILKTAI VSEPMILEIL NNGGETGIQS LLNDSYGNYV
841 LQTALDISHK QNDYLYKRLS EIVAPLLVGP IRNTPHGKRI IGMLHLDS (SEQ ID NO:
211).
```

In some embodiments, a PUF5 protein of the disclosure comprises or consists of the amino acid sequence of

```
  1 MSDSTGRINS KASDSSSISD HQTADLSIFN GSFDGGAFSS SNIPLFNFMG TGNQRFQYSP
 61 HPFAKSSDPC RLAALTPSTP KGPLNLTPAD FGLADFSVGN ESFADFTANN TSFVGNVQSN
121 VRSTRLLPAW AVDNSGNIRD DLTLQDVVSN GSLIDFAMDR TGVKFLERHF PEDHDNEMHF
181 VLFDKLTEQG AVFTSLCRSA AGNFIIQKFV EHATLDEQER LVRKMCDNGL IEMCLDKFAC
241 RVVQMSIQKF DVSIAMKLVE KISSLDFLPL CTDQCAIHVL QKVVKLLPIS AWSFFVKFLC
301 RDDNLMTVCQ DKYGCRLVQQ TIDKLSDNPK LHCFNTRLQL LHGLMTSVAR NCFRLSSNEF
361 ANYVVQYVIK SSGVMEMYRD TIIEKCLLRN ILSMSQDKYA SHVVEGAFLF APPLLLSEMM
421 DEIFDGYVKD QETNRDALDI LLFHQYGNYV VQQMISICIS ALLGKEERKM VASEMRLYAK
481 WFDRIKNRVN RHSGRLERFS SGKKIIESLQ KLNVPMTMTN EPMPYWAMPT PLMDISAHFM
541 NKLNFQKNSV FDE (SEQ ID NO: 212).
```

In some embodiments, a PUF6 protein of the disclosure comprises or consists of the amino acid sequence of

```
  1 MTPNRRSTDS YNMLGASFDF DPDFSLLSNK THKNKNPKPP VKLLPYRHGS NTTSSDLDNY
```

```
 61 IFNSGSGSSD DETPPPAAPI FISLEEVLLN GLLIDFAIDP SGVKFLEANY PLDSEDQIRK
121 AVFEKLTEST TLFVGLCHSR NGNFIVQKLV ELATPAEQRE LLRQMIDGGL LVMCKDKFAC
181 RVVQLALQKF DHSNVFQLIQ ELSTFDLAAM CTDQISIHVI QRVVKQLPVD MWTFFVHFLS
241 SGDSLMAVCQ DKYGCRLVQQ VIDRLAENPK LPCFKFRIQL LHSLMTCIVR NCYRLSSNEF
301 ANYVIQYVIK SSGIMEMYRD TIIDKCLLRN LLSMSQDKYA SHVIEGAFLF APPALLHEMM
361 EEIFSGYVKD VELNRDALDI LLFHQYGNYV VQQMISICTA ALIGKEERQL PPAILLLYSG
421 WYEKMKQRVL QHASRLERFS SGKKIIDSVM RHGVPTAAAI NAQAAPSLME LTAQFDAMFP
481 SFLAR (SEQ ID NO: 213).
```

In some embodiments, a PUF7 protein of the disclosure comprises or consists of the amino acid sequence of

```
  1 MTPNRRSTDS YNMLGASFDF DPDFSLLLSNK THKNKNPKPP VKLLPYRHGS NTTSSDSDSY
 61 IFNSGSGSSD AETPAPVAPI FISLEDVLLN GQLIDFAIDP SGVKFLEANY PLDSEDQIRK
121 AVFEKFTEST TLFVGLCHSR NGNFIVQKLV ELATPAEQRE LLRQMIDGGL LAMCKDKFAC
181 RVVQLALQKF DHSNVFQLIQ ELSTFDLAAM CTDQISIHVI QRVVKQLPVD MWTFFVHFLS
241 SGDSLMAVCQ DKYGCRLVQQ VIDRLAENPK LPCFKFRIQL LHSLMTCIVR NCYRLSSNEF
301 ANYVIQYVIK SSGIMEMYRD TIIDKCLLRN LLSMSQDKYA SHVIEGAFLF APPALLHEMM
361 EEIFSGYVKD VESNRDALDI LLFHQYGNYV VQQMISICTA ALIGKEEREL PPAILLLYSG
421 WYEKMKQRVL QHASRLERFS SGKKIIDSVM RHGVPTAAAV NAQAAPSLME LTAQFDAMFP
481 SFLAR (SEQ ID NO: 214).
```

In some embodiments, a PUF8 protein of the disclosure comprises or consists of the amino acid sequence of

```
  1 MSRPISIGNT CTFDPSASPI ESLGRSIGAQ KIVDSVCGSP IRSYGRHIST NPKNERLPDT
 61 PEFQFATYMH QGGKVIGQNT LHMFGTPPSC YCAQENIPIS SNVGHVLSTI NNNYMNHQYN
121 GSNMFSNQMT QMLQAQAYND LQMHQAHSQS IRVPVQPSAT GIFSNPYREP TTTDDLLTRY
181 RANPAMMKNL KLSDIRGALL KFAKDQVGSR FIQQELASSK DRFEKDSIFD EVVSNADELV
241 DDIFGNYVVQ KFFEYGEERH WARLVDAIID RVPEYAFQMY ACRVLQKALE KINEPLQIKI
301 LSQIRHVIHR CMKDQNGNHV VQKAIEKVSP QYVQFIVDTL LESSNTIYEM SVDPYGCRVV
361 QRCLEHCSPS QTKPVIGQIH KRFDEIANNQ YGNYVVQHVI EHGSEEDRMV IVTRVSNNLF
421 EFATHKYSSN VIEKCLEQGA VYHKSMIVGA ACHHQEGSVP IVVQMMKDQY ANYVVQKMFD
481 QVTSEQRREL ILTVRPHIPV LRQFPHGKHI LAKLEKYFQK PAVMSYPYQD MQGSH (SEQ
ID NO: 215).
```

In some embodiments, a PUF9 protein of the disclosure comprises or consists of the amino acid sequence of

```
  1 MADPNWAYAP PTNYYADHSI AKPIMISGGH PSQDQGHSPK SESFGQSVTT AFNGMVDNLV
 61 GSPSSSVQQR NYFTTTPFPI SRSPNDRNDD KIMGNGSYGV PIPIPQDGVP QGTPDFQMTP
121 FLQQGGHLIG GSPNGPVQVS GNWYSGGAGI FSTMQQADPS NGMPGMAAEF VNNENGMPGP
181 NGMHQQAMIS GSPPFPYQNM MNLTTSFGAM GLGPQQIQQR DPQMFQQPIL HEPIQGMAQN
241 GFGQQVFFTQ MQNQQHPQGQ AQQQLQQLAQ QHQQQQNSQQ FFGQGPNGMG NGGVMNDWSQ
301 RSFGMPQQQA QQNGLPPNFS QNPPRRRGPE DPNGQTPKTL QDIKNNVIEF AKDQHGSRFI
361 QQKLERASLR DKAAIFTPVL ENAEELMTDV FGNYVIQKFF EFGNNEQRNQ LVGTIRGNVM
421 KLALQMYGCR VIQKALEYVE EKYQHEILGE MEGQVLKCVK DQNGNHVIQK VIERVEPERL
481 QFIIDAFTKN NSDNVYTLSV HPYGCRVIQR VLEYCNEEQK QPVLDALQIH LKQLVLDQYG
541 NYVIQHVIEH GSPSDKEQIV QDVISDDLLK FAQHKFASNV IEKCLTFGGH AERNLIIDKV
601 CGDPNDPSPP LLQMMKDPFA NYVVQKMLDV ADPQHRKKIT LTIKPHIATL RKYNFGKHIL
661 LKLEKYFAKQ APANSSNSSS NDQIYEHSPF DIPLGADFSN HPF (SEQ ID NO:
216).
```

[0279] In some embodiments of the compositions of the disclosure, at least one of the RNA-binding proteins or RNA-binding portions thereof is a PPR protein. PPR proteins (proteins with pentatricopeptide repeat (PPR) motifs derived from plants) are nuclear-encoded and exclusively

controlled at the RNA level organelles (chloroplasts and mitochondria), cutting, translation, splicing, RNA editing, genes specifically acting on RNA stability. PPR proteins are typically a motif of 35 amino acids and have a structure in which a PPR motif is about 10 contiguous amino acids. The combination of PPR motifs can be used for sequence-selective binding to RNA. PPR proteins are often comprised of PPR motifs of about 10 repeat domains. PPR domains or RNA-binding domains may be configured to be catalytically inactive. WO 2013/058404 incorporated herein by reference in its entirety.

[0280] In some embodiments, the fusion protein disclosed herein comprises a linker between the at least two RNA-binding polypeptides. In some embodiments, the linker is a peptide linker. In some embodiments, the peptide linker comprises one or more repeats of the tri-peptide GGS. In other embodiments, the linker is a non-peptide linker. In some embodiments, the non-peptide linker comprises polyethylene glycol (PEG), polypropylene glycol (PPG), co-poly(ethylene/propylene) glycol, polyoxyethylene (POE), polyurethane, polyphosphazene, polysaccharides, dextran, polyvinyl alcohol, polyvinylpyrrolidones, polyvinyl ethyl ether, polyacryl amide, polyacrylate, polycyanoacrylates, lipid polymers, chitins, hyaluronic acid, heparin, or an alkyl linker.

[0281] In some embodiments, the at least one RNA-binding protein does not require multimerization for RNA-binding activity. In some embodiments, the at least one RNA-binding protein is not a monomer of a multimer complex. In some embodiments, a multimer protein complex does not comprise the RNA binding protein. In some embodiments, the at least one of RNA-binding protein selectively binds to a target sequence within the RNA molecule. In some embodiments, the at least one RNA-binding protein does not comprise an affinity for a second sequence within the RNA molecule. In some embodiments, the at least one RNA-binding protein does not comprise a high affinity for or selectively bind a second sequence within the RNA molecule. In some embodiments, the at least one RNA-binding protein comprises between 2 and 1300 amino acids, inclusive of the endpoints.

[0282] In some embodiments, the at least one RNA-binding protein of the fusion proteins disclosed herein further comprises a sequence encoding a nuclear localization signal (NLS). In some embodiments, a nuclear localization signal (NLS) is positioned 3' to the RNA binding protein. In some embodiments, the at least one RNA-binding protein comprises an NLS at a C-

terminus of the protein. In some embodiments, the at least one RNA-binding protein further comprises a first sequence encoding a first NLS and a second sequence encoding a second NLS. In some embodiments, the first NLS or the second NLS is positioned 3' to the RNA-binding protein. In some embodiments, the at least one RNA-binding protein comprises the first NLS or the second NLS at a C-terminus of the protein. In some embodiments, the at least one RNA-binding protein further comprises an NES (nuclear export signal) or other peptide tag or secretory signal.

[0283] In some embodiments, a fusion protein disclosed herein comprises the at least one RNA-binding protein as a first RNA-binding protein together with a second RNA-binding protein comprising or consisting of a nuclease domain.

[0284] In some embodiments, the second RNA-binding polypeptide is operably configured to the first RNA-binding polypeptide at the C-terminus of the first RNA-binding polypeptide. In some embodiments, the second RNA-binding polypeptide is operably configured to the first RNA-binding polypeptide at the N-terminus of the first RNA-binding polypeptide. For example, one such exemplary fusion protein is E99 which is configured so that RNAse1(R39D, N67D, N88A, G89D, R19D, H119N, K41R) is located at the N-terminus of SpyCas9 whereas another exemplary fusion protein, E100, is configured so that RNAse1(R39D, N67D, N88A, G89D, R19D, H119N, K41R) is located at the C-terminus of SpyCas9. See Figure 6.

Vectors

[0285] In some embodiments of the compositions and methods of the disclosure, a vector comprises a guide RNA of the disclosure. In some embodiments, the vector comprises at least one guide RNA of the disclosure. In some embodiments, the vector comprises one or more guide RNA(s) of the disclosure. In some embodiments, the vector comprises two or more guide RNAs of the disclosure. In some embodiments, the vector further comprises a fusion protein of the disclosure. In some embodiments, the fusion protein comprises a first RNA binding protein and a second RNA binding protein.

[0286] In some embodiments of the compositions and methods of the disclosure, a first vector comprises a guide RNA of the disclosure and a second vector comprises a fusion protein of the disclosure. In some embodiments, the first vector comprises at least one guide RNA of the disclosure. In some embodiments, the first vector comprises one or more guide RNA(s) of the

disclosure. In some embodiments, the first vector comprises two or more guide RNA(s) of the disclosure. In some embodiments, the fusion protein comprises a first RNA binding protein and a second RNA binding protein. In some embodiments, the first vector and the second vector are identical. In some embodiments, the first vector and the second vector are not identical.

[0287] In some embodiments of the compositions and methods of the disclosure, the vector is or comprises a component of a "2-component RNA targeting system" comprising (a) nucleic acid sequence encoding a RNA-targeted fusion protein of the disclosure; and (b) a single guide RNA (sgRNA) sequence comprising: on its 5' end, an RNA sequence (or spacer sequence) that hybridizes to or binds to a target RNA sequence; and on its 3' end, an RNA sequence (or scaffold sequence) capable of binding to or associating with the CRISPR/Cas protein of the fusion protein; and wherein the 2-component RNA targeting system recognizes and alters the target RNA in a cell in the absence of a PAMmer. In some embodiments, the sequences of the 2-component system are in a single vector. In some embodiments, the spacer sequence of the 2-component system targets a repeat sequence selected from the group consisting of CUG, CCUG, CAG, and GGGGCC.

[0288] In some embodiments of the compositions and methods of the disclosure, a vector of the disclosure is a viral vector. In some embodiments, the viral vector comprises a sequence isolated or derived from a retrovirus. In some embodiments, the viral vector comprises a sequence isolated or derived from a lentivirus. In some embodiments, the viral vector comprises a sequence isolated or derived from an adenovirus. In some embodiments, the viral vector comprises a sequence isolated or derived from an adeno-associated virus (AAV). In some embodiments, the viral vector is replication incompetent. In some embodiments, the viral vector is isolated or recombinant. In some embodiments, the viral vector is self-complementary.

[0289] In some embodiments of the compositions and methods of the disclosure, the viral vector comprises a sequence isolated or derived from an adeno-associated virus (AAV). In some embodiments, the viral vector comprises an inverted terminal repeat sequence or a capsid sequence that is isolated or derived from an AAV of serotype AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11 or AAV12. In some embodiments, the viral vector is replication incompetent. In some embodiments, the viral vector is isolated or recombinant (rAAV). In some embodiments, the viral vector is self-complementary (scAAV).

[0290] In some embodiments of the compositions and methods of the disclosure, a vector of the disclosure is a non-viral vector. In some embodiments, the vector comprises or consists of a nanoparticle, a micelle, a liposome or lipoplex, a polymersome, a polyplex or a dendrimer. In some embodiments, the vector is an expression vector or recombinant expression system. As used herein, the term "recombinant expression system" refers to a genetic construct for the expression of certain genetic material formed by recombination.

[0291] In some embodiments of the compositions and methods of the disclosure, an expression vector, viral vector or non-viral vector provided herein, includes without limitation, an expression control element. An "expression control element" as used herein refers to any sequence that regulates the expression of a coding sequence, such as a gene. Exemplary expression control elements include but are not limited to promoters, enhancers, microRNAs, post-transcriptional regulatory elements, polyadenylation signal sequences, and introns. Expression control elements may be constitutive, inducible, repressible, or tissue-specific, for example. A "promoter" is a control sequence that is a region of a polynucleotide sequence at which initiation and rate of transcription are controlled. It may contain genetic elements at which regulatory proteins and molecules may bind such as RNA polymerase and other transcription factors. In some embodiments, expression control by a promoter is tissue-specific. Non-limiting exemplary promoters include CMV, CBA, CAG, Cbh, EF-1a, PGK, UBC, GUSB, UCOE, hAAT, TBG, Desmin, MCK, C5-12, NSE, Synapsin, PDGF, MecP2, CaMKII, mGluR2, NFL, NFH, nβ2, PPE, ENK, EAAT2, GFAP, MBP, and U6 promoters. An "enhancer" is a region of DNA that can be bound by activating proteins to increase the likelihood or frequency of transcription. Non-limiting exemplary enhancers and posttranscriptional regulatory elements include the CMV enhancer and WPRE.

[0292] In some embodiments of the compositions and methods of the disclosure, an expression vector, viral vector or non-viral vector provided herein, includes without limitation, vector elements such as an IRES or 2A peptide sites for configuration of "multicistronic" or "polycistronic" or "bicistronic" or tricistronic" constructs, i.e., having double or triple or multiple coding areas or exons, and as such will have the capability to express from mRNA two or more proteins from a single construct. Multicistronic vectors simultaneously express two or more separate proteins from the same mRNA. The two strategies most widely used for

constructing multicistronic configurations are through the use of an IRES or a 2A self-cleaving site. An "IRES" refers to an internal ribosome entry site or portion thereof of viral, prokaryotic, or eukaryotic origin which are used within polycistronic vector constructs. In some embodiments, an IRES is an RNA element that allows for translation initiation in a cap-independent manner. The term "self-cleaving peptides" or "sequences encoding self-cleaving peptides" or "2A self-cleaving site" refer to linking sequences which are used within vector constructs to incorporate sites to promote ribosomal skipping and thus to generate two polypeptides from a single promoter, such self-cleaving peptides include without limitation, T2A, and P2A peptides or sequences encoding the self-cleaving peptides.

[0293] In some embodiments, the vector is a viral vector. In some embodiments, the vector is an adenoviral vector, an adeno-associated viral (AAV) vector, or a lentiviral vector. In some embodiments, the vector is a retroviral vector, an adenoviral/retroviral chimera vector, a herpes simplex viral I or II vector, a parvoviral vector, a reticuloendotheliosis viral vector, a polioviral vector, a papillomaviral vector, a vaccinia viral vector, or any hybrid or chimeric vector incorporating favorable aspects of two or more viral vectors. In some embodiments, the vector further comprises one or more expression control elements operably linked to the polynucleotide. In some embodiments, the vector further comprises one or more selectable markers. In some embodiments, the AAV vector has low toxicity. In some embodiments, the AAV vector does not incorporate into the host genome, thereby having a low probability of causing insertional mutagenesis. In some embodiments, the AAV vector can encode a range of total polynucleotides from 4.5 kb to 4.75 kb. In some embodiments, exemplary AAV vectors that may be used in any of the herein described compositions, systems, methods, and kits can include an AAV1 vector, a modified AAV1 vector, an AAV2 vector, a modified AAV2 vector, an AAV3 vector, a modified AAV3 vector, an AAV4 vector, a modified AAV4 vector, an AAV5 vector, a modified AAV5 vector, an AAV6 vector, a modified AAV6 vector, an AAV7 vector, a modified AAV7 vector, an AAV8 vector, an AAV9 vector, an AAV.rh10 vector, a modified AAV.rh10 vector, an AAV.rh32/33 vector, a modified AAV.rh32/33 vector, an AAV.rh43 vector, a modified AAV.rh43 vector, an AAV.rh64R1 vector, and a modified AAV.rh64R1 vector and any combinations or equivalents thereof. In some embodiments, the lentiviral vector is an integrase-competent lentiviral vector (ICLV). In some embodiments, the

lentiviral vector can refer to the transgene plasmid vector as well as the transgene plasmid vector in conjunction with related plasmids (e.g., a packaging plasmid, a rev expressing plasmid, an envelope plasmid) as well as a lentiviral-based particle capable of introducing exogenous nucleic acid into a cell through a viral or viral-like entry mechanism. Lentiviral vectors are well-known in the art (see, e.g., Trono D. (2002) Lentiviral vectors, New York: Spring-Verlag Berlin Heidelberg and Durand et al. (2011) *Viruses* 3(2):132-159 doi: 10.3390/v3020132). In some embodiments, exemplary lentiviral vectors that may be used in any of the herein described compositions, systems, methods, and kits can include a human immunodeficiency virus (HIV) 1 vector, a modified human immunodeficiency virus (HIV) 1 vector, a human immunodeficiency virus (HIV) 2 vector, a modified human immunodeficiency virus (HIV) 2 vector, a sooty mangabey simian immunodeficiency virus (SIV_{SM}) vector, a modified sooty mangabey simian immunodeficiency virus (SIV_{SM}) vector, a African green monkey simian immunodeficiency virus (SIV_{AGM}) vector, a modified African green monkey simian immunodeficiency virus (SIV_{AGM}) vector, an equine infectious anemia virus (EIAV) vector, a modified equine infectious anemia virus (EIAV) vector, a feline immunodeficiency virus (FIV) vector, a modified feline immunodeficiency virus (FIV) vector, a Visna/maedi virus (VNV/VMV) vector, a modified Visna/maedi virus (VNV/VMV) vector, a caprine arthritis-encephalitis virus (CAEV) vector, a modified caprine arthritis-encephalitis virus (CAEV) vector, a bovine immunodeficiency virus (BIV), or a modified bovine immunodeficiency virus (BIV).

Nucleic Acids

[0294] Provided herein are the nucleic acid sequences encoding the fusion proteins disclosed herein for use in gene transfer and expression techniques described herein. It should be understood, although not always explicitly stated that the sequences provided herein can be used to provide the expression product as well as substantially identical sequences that produce a protein that has the same biological properties. These "biologically equivalent" or "biologically active" or "equivalent" polypeptides are encoded by equivalent polynucleotides as described herein. They may possess at least 60%, or alternatively, at least 65%, or alternatively, at least 70%, or alternatively, at least 75%, or alternatively, at least 80%, or alternatively at least 85%, or alternatively at least 90%, or alternatively at least 95% or alternatively at least 98%, identical primary amino acid sequence to the reference polypeptide when compared using sequence

identity methods run under default conditions. Specific polypeptide sequences are provided as examples of particular embodiments. Modifications to the sequences to amino acids with alternate amino acids that have similar charge. Additionally, an equivalent polynucleotide is one that hybridizes under stringent conditions to the reference polynucleotide or its complement or in reference to a polypeptide, a polypeptide encoded by a polynucleotide that hybridizes to the reference encoding polynucleotide under stringent conditions or its complementary strand. Alternatively, an equivalent polypeptide or protein is one that is expressed from an equivalent polynucleotide.

[0295] The nucleic acid sequences (e.g., polynucleotide sequences) disclosed herein may be codon-optimized which is a technique well known in the art. In some embodiments disclosed herein, exemplary Cas sequences, such as e.g., SEQ ID NO: 46 (Cas13d), are codon optimized for expression in human cells. Codon optimization refers to the fact that different cells differ in their usage of particular codons. This codon bias corresponds to a bias in the relative abundance of particular tRNAs in the cell type. By altering the codons in the sequence to match with the relative abundance of corresponding tRNAs, it is possible to increase expression. It is also possible to decrease expression by deliberately choosing codons for which the corresponding tRNAs are known to be rare in a particular cell type. Codon usage tables are known in the art for mammalian cells, as well as for a variety of other organisms. Based on the genetic code, nucleic acid sequences coding for, e.g., a Cas protein, can be generated. In some embodiments, such a sequence is optimized for expression in a host or target cell, such as a host cell used to express the Cas protein or a cell in which the disclosed methods are practiced (such as in a mammalian cell, e.g., a human cell). Codon preferences and codon usage tables for a particular species can be used to engineer isolated nucleic acid molecules encoding a Cas protein (such as one encoding a protein having at least 80%, at least 85%, at least 90%, at least 92%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to its corresponding wild-type protein) that takes advantage of the codon usage preferences of that particular species. For example, the Cas proteins disclosed herein can be designed to have codons that are preferentially used by a particular organism of interest. In one example, an Cas nucleic acid sequence is optimized for expression in human cells, such as one having at least 70%, at least 80%, at least 85%, at least 90%, at least 92%, at least 95%, at least 98%, or at least

99% sequence identity to its corresponding wild-type or originating nucleic acid sequence. In some embodiments, an isolated nucleic acid molecule encoding at least one Cas protein (which can be part of a vector) includes at least one Cas protein coding sequence that is codon optimized for expression in a eukaryotic cell, or at least one Cas protein coding sequence codon optimized for expression in a human cell. In one embodiment, such a codon optimized Cas coding sequence has at least 80%, at least 85%, at least 90%, at least 92%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to its corresponding wild-type or originating sequence. In another embodiment, a eukaryotic cell codon optimized nucleic acid sequence encodes a Cas protein having at least 85%, at least 90%, at least 92%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to its corresponding wild-type or originating protein. In another embodiment, a variety of clones containing functionally equivalent nucleic acids may be routinely generated, such as nucleic acids which differ in sequence but which encode the same Cas protein sequence. Silent mutations in the coding sequence result from the degeneracy (i.e., redundancy) of the genetic code, whereby more than one codon can encode the same amino acid residue. Thus, for example, leucine can be encoded by CTT, CTC, CTA, CTG, TTA, or TTG; serine can be encoded by TCT, TCC, TCA, TCG, AGT, or AGC; asparagine can be encoded by AAT or AAC; aspartic acid can be encoded by GAT or GAC; cysteine can be encoded by TGT or TGC; alanine can be encoded by GCT, GCC, GCA, or GCG; glutamine can be encoded by CAA or CAG; tyrosine can be encoded by TAT or TAC; and isoleucine can be encoded by ATT, ATC, or ATA. Tables showing the standard genetic code can be found in various sources (see, for example, Stryer, 1988, Biochemistry, 3.sup.rd Edition, W.H. 5 Freeman and Co., NY).

[0296] "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, Hoogstein 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 a PC reaction, or the enzymatic cleavage of a polynucleotide by a ribozyme.

[0297] Examples of stringent hybridization conditions include: incubation temperatures of about 25°C to about 37°C; hybridization buffer concentrations of about 6x SSC to about 10x SSC; formamide concentrations of about 0% to about 25%; and wash solutions from about 4x SSC to about 8x SSC. Examples of moderate hybridization conditions include: incubation temperatures of about 40°C to about 50°C; buffer concentrations of about 9x SSC to about 2x SSC; formamide concentrations of about 30% to about 50%; and wash solutions of about 5x SSC to about 2x SSC. Examples of high stringency conditions include: incubation temperatures of about 55°C to about 68°C; buffer concentrations of about lx SSC to about 0.1x SSC; formamide concentrations of about 55% to about 75%; and wash solutions of about lx SSC, 0.1x SSC, or deionized water. In general, hybridization incubation times are from 5 minutes to 24 hours, with 1, 2, or more washing steps, and wash incubation times are about 1, 2, or 15 minutes. SSC is 0.15 M NaCl and 15 mM citrate buffer. It is understood that equivalents of SSC using other buffer systems can be employed.

[0298] "Homology" or "identity" or "similarity" refers to sequence similarity between two peptides or between two nucleic acid molecules. Homology can be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same base or amino acid, then the molecules are homologous at that position. A degree of homology between sequences is a function of the number of matching or homologous positions shared by the sequences. An "unrelated" or "non-homologous" sequence shares less than 40% identity, or alternatively less than 25% identity, with one of the sequences of the present invention.

Cells

[0299] In some embodiments of the compositions and methods of the disclosure, a cell of the disclosure is a prokaryotic cell.

[0300] In some embodiments of the compositions and methods of the disclosure, a cell of the disclosure is a eukaryotic cell. In some embodiments, the cell is a mammalian cell. In some embodiments, the cell is a bovine, murine, feline, equine, porcine, canine, simian, or human cell. In some embodiments, the cell is a non-human mammalian cell such as a non-human primate cell.

[0301] In some embodiments, a cell of the disclosure is a somatic cell. In some embodiments, a cell of the disclosure is a germline cell. In some embodiments, a germline cell of the disclosure is not a human cell.

[0302] In some embodiments of the compositions and methods of the disclosure, a cell of the disclosure is a stem cell. In some embodiments, a cell of the disclosure is an embryonic stem cell. In some embodiments, an embryonic stem cell of the disclosure is not a human cell. In some embodiments, a cell of the disclosure is a multipotent stem cell or a pluripotent stem cell. In some embodiments, a cell of the disclosure is an adult stem cell. In some embodiments, a cell of the disclosure is an induced pluripotent stem cell (iPSC). In some embodiments, a cell of the disclosure is a hematopoietic stem cell (HSC).

[0303] In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is an immune cell. In some embodiments, an immune cell of the disclosure is a lymphocyte. In some embodiments, an immune cell of the disclosure is a T lymphocyte (also referred to herein as a T-cell). Exemplary T-cells of the disclosure include, but are not limited to, naïve T cells, effector T cells, helper T cells, memory T cells, regulatory T cells (Tregs) and Gamma delta T cells. In some embodiments, an immune cell of the disclosure is a B lymphocyte. In some embodiments, an immune cell of the disclosure is a natural killer cell. In some embodiments, an immune cell of the disclosure is an antigen-presenting cell.

[0304] In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is a muscle cell. In some embodiments, a muscle cell of the disclosure is a myoblast or a myocyte. In some embodiments, a muscle cell of the disclosure is a cardiac muscle cell, skeletal muscle cell or smooth muscle cell. In some embodiments, a muscle cell of the disclosure is a striated cell.

[0305] In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is an epithelial cell. In some embodiments, an epithelial cell of the disclosure forms a squamous cell epithelium, a cuboidal cell epithelium, a columnar cell epithelium, a stratified cell epithelium, a pseudostratified columnar cell epithelium or a transitional cell epithelium. In some embodiments, an epithelial cell of the disclosure forms a gland including, but not limited to, a pineal gland, a thymus gland, a pituitary gland, a thyroid gland, an adrenal gland, an apocrine gland, a holocrine gland, a merocrine gland, a serous gland,

a mucous gland and a sebaceous gland. In some embodiments, an epithelial cell of the disclosure contacts an outer surface of an organ including, but not limited to, a lung, a spleen, a stomach, a pancreas, a bladder, an intestine, a kidney, a gallbladder, a liver, a larynx or a pharynx. In some embodiments, an epithelial cell of the disclosure contacts an outer surface of a blood vessel or a vein.

[0306] In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is a neuronal cell. In some embodiments, a neuron cell of the disclosure is a neuron of the central nervous system. In some embodiments, a neuron cell of the disclosure is a neuron of the brain or the spinal cord. In some embodiments, a neuron cell of the disclosure is a neuron of the retina. In some embodiments, a neuron cell of the disclosure is a neuron of a cranial nerve or an optic nerve. In some embodiments, a neuron cell of the disclosure is a neuron of the peripheral nervous system. In some embodiments, a neuron cell of the disclosure is a neuroglial or a glial cell. In some embodiments, a glial of the disclosure is a glial cell of the central nervous system including, but not limited to, oligodendrocytes, astrocytes, ependymal cells, and microglia. In some embodiments, a glial of the disclosure is a glial cell of the peripheral nervous system including, but not limited to, Schwann cells and satellite cells.

[0307] In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is a primary cell.

[0308] In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is a cultured cell.

[0309] In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is in vivo, in vitro, ex vivo or in situ.

[0310] In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is autologous or allogeneic.

Methods of Use

[0311] The disclosure provides a method of modifying level of expression of an RNA molecule of the disclosure or a protein encoded by the RNA molecule comprising contacting the composition and the RNA molecule under conditions suitable for binding of one or more of the guide RNA or the fusion protein (or a portion thereof) to the RNA molecule.

[0312] The disclosure provides a method of modifying an activity of a protein encoded by an RNA molecule comprising contacting the composition and the RNA molecule under conditions suitable for binding of one or more of the guide RNA or the fusion protein (or a portion thereof) to the RNA molecule.

[0313] The disclosure provides a method of modifying level of expression of an RNA molecule of the disclosure or a protein encoded by the RNA molecule comprising contacting the composition and a cell comprising the RNA molecule under conditions suitable for binding of one or more of the guide RNA or the fusion protein (or a portion thereof) to the RNA molecule. In some embodiments, the cell is in vivo, in vitro, ex vivo or in situ. In some embodiments, the composition comprises a vector comprising composition comprising a guide RNA of the disclosure and a fusion protein of the disclosure. In some embodiments, the vector is an AAV.

[0314] The disclosure provides a method of modifying an activity of a protein encoded by an RNA molecule comprising contacting the composition and a cell comprising the RNA molecule under conditions suitable for binding of one or more of the guide RNA or the fusion protein (or a portion thereof) to the RNA molecule. In some embodiments, the cell is in vivo, in vitro, ex vivo or in situ. In some embodiments, the composition comprises a vector comprising composition comprising a guide RNA or a single guide RNA of the disclosure and a fusion protein of the disclosure. In some embodiments, the vector is an AAV.

[0315] The disclosure provides a method of modifying level of expression of an RNA molecule of the disclosure or a protein encoded by the RNA molecule comprising contacting the composition and the RNA molecule under conditions suitable for RNA nuclease activity wherein the fusion protein induces a break in the RNA molecule.

[0316] The disclosure provides a method of modifying an activity of a protein encoded by an RNA molecule comprising contacting the composition and the RNA molecule under conditions suitable for RNA nuclease activity wherein the fusion protein induces a break in the RNA molecule.

[0317] The disclosure provides a method of modifying a level of expression of an RNA molecule of the disclosure or a protein encoded by the RNA molecule comprising contacting the composition and a cell comprising the RNA molecule under conditions suitable for RNA nuclease activity wherein the fusion protein induces a break in the RNA molecule. In some

embodiments, the cell is in vivo, in vitro, ex vivo or in situ. In some embodiments, the composition comprises a vector comprising composition comprising a guide RNA of the disclosure and a fusion protein of the disclosure. In some embodiments, the vector is an AAV.

[0318] The disclosure provides a method of modifying an activity of a protein encoded by an RNA molecule comprising contacting the composition and a cell comprising the RNA molecule under conditions suitable for RNA nuclease activity wherein the fusion protein induces a break in the RNA molecule. In some embodiments, the cell is in vivo, in vitro, ex vivo or in situ. In some embodiments, the composition comprises a vector comprising composition comprising a guide RNA or a single guide RNA of the disclosure and a fusion protein of the disclosure. In some embodiments, the vector is an AAV.

[0319] The disclosure provides a method of treating a disease or disorder comprising administering to a subject a therapeutically effective amount of a composition of the disclosure.

[0320] The disclosure provides a method of treating a disease or disorder comprising administering to a subject a therapeutically effective amount of a composition of the disclosure, wherein the composition comprises a vector comprising composition comprising a guide RNA of the disclosure and a fusion protein of the disclosure and wherein the composition modifies a level of expression of an RNA molecule of the disclosure or a protein encoded by the RNA molecule.

[0321] The disclosure provides a method of treating a disease or disorder comprising administering to a subject a therapeutically effective amount of a composition of the disclosure, wherein the composition comprises a vector comprising composition comprising a guide RNA of the disclosure and a fusion protein of the disclosure and wherein the composition modifies an activity of a protein encoded by an RNA molecule.

[0322] In some embodiments of the compositions and methods of the disclosure, a disease or disorder of the disclosure includes, but is not limited to, a genetic disease or disorder. In some embodiments, the genetic disease or disorder is a single-gene disease or disorder. In some embodiments, the single-gene disease or disorder is an autosomal dominant disease or disorder, an autosomal recessive disease or disorder, an X-chromosome linked (X-linked) disease or disorder, an X-linked dominant disease or disorder, an X-linked recessive disease or disorder, a Y-linked disease or disorder or a mitochondrial disease or disorder. In some embodiments, the

genetic disease or disorder is a multiple-gene disease or disorder. In some embodiments, the genetic disease or disorder is a multiple-gene disease or disorder. In some embodiments, the single-gene disease or disorder is an autosomal dominant disease or disorder including, but not limited to, Huntington's disease, neurofibromatosis type 1, neurofibromatosis type 2, Marfan syndrome, hereditary nonpolyposis colorectal cancer, hereditary multiple exostoses, Von Willebrand disease, and acute intermittent porphyria. In some embodiments, the single-gene disease or disorder is an autosomal recessive disease or disorder including, but not limited to, Albinism, Medium-chain acyl-CoA dehydrogenase deficiency, cystic fibrosis, sickle-cell disease, Tay-Sachs disease, Niemann-Pick disease, spinal muscular atrophy, and Roberts syndrome. In some embodiments, the single-gene disease or disorder is X-linked disease or disorder including, but not limited to, muscular dystrophy, Duchenne muscular dystrophy, Hemophilia, Adrenoleukodystrophy (ALD), Rett syndrome, and Hemophilia A. In some embodiments, the single-gene disease or disorder is a mitochondrial disorder including, but not limited to, Leber's hereditary optic neuropathy.

[0323] In some embodiments of the compositions and methods of the disclosure, a disease or disorder of the disclosure includes, but is not limited to, an immune disease or disorder. In some embodiments, the immune disease or disorder is an immunodeficiency disease or disorder including, but not limited to, B-cell deficiency, T-cell deficiency, neutropenia, asplenia, complement deficiency, acquired immunodeficiency syndrome (AIDS) and immunodeficiency due to medical intervention (immunosuppression as an intended or adverse effect of a medical therapy). In some embodiments, the immune disease or disorder is an autoimmune disease or disorder including, but not limited to, Achalasia, Addison’s disease, Adult Still's disease, Agammaglobulinemia, Alopecia areata, Amyloidosis, Anti-GBM/Anti-TBM nephritis, Antiphospholipid syndrome, Autoimmune angioedema, Autoimmune dysautonomia, Autoimmune encephalomyelitis, Autoimmune hepatitis, Autoimmune inner ear disease (AIED), Autoimmune myocarditis, Autoimmune oophoritis, Autoimmune orchitis, Autoimmune pancreatitis, Autoimmune retinopathy, Autoimmune urticaria, Axonal & neuronal neuropathy (AMAN), Baló disease, Behcet’s disease, Benign mucosal pemphigoid, Bullous pemphigoid, Castleman disease (CD), Celiac disease, Chagas disease, Chronic inflammatory demyelinating polyneuropathy (CIDP), Chronic recurrent multifocal osteomyelitis (CRMO), Churg-Strauss

Syndrome (CSS) or Eosinophilic Granulomatosis (EGPA), Cicatricial pemphigoid, Cogan's syndrome, Cold agglutinin disease, Congenital heart block, Coxsackie myocarditis, CREST syndrome, Crohn's disease, Dermatitis herpetiformis, Dermatomyositis, Devic's disease (neuromyelitis optica), Discoid lupus, Dressler's syndrome, Endometriosis, Eosinophilic esophagitis (EoE), Eosinophilic fasciitis, Erythema nodosum, Essential mixed cryoglobulinemia, Evans syndrome, Fibromyalgia, Fibrosing alveolitis, Giant cell arteritis (temporal arteritis), Giant cell myocarditis, Glomerulonephritis, Goodpasture's syndrome, Granulomatosis with Polyangiitis, Graves' disease, Guillain-Barre syndrome, Hashimoto's thyroiditis, Hemolytic anemia, Henoch-Schonlein purpura (HSP), Herpes gestationis or pemphigoid gestationis (PG), Hidradenitis Suppurativa (HS) (Acne Inversa), Hypogammalglobulinemia, IgA Nephropathy, IgG4-related sclerosing disease, Immune thrombocytopenic purpura (ITP), Inclusion body myositis (IBM), Interstitial cystitis (IC), Juvenile arthritis, Juvenile diabetes (Type 1 diabetes), Juvenile myositis (JM), Kawasaki disease, Lambert-Eaton syndrome, Leukocytoclastic vasculitis, Lichen planus, Lichen sclerosus, Ligneous conjunctivitis, Linear IgA disease (LAD), Lupus, Lyme disease chronic, Meniere's disease, Microscopic polyangiitis (MPA), Mixed connective tissue disease (MCTD), Mooren's ulcer, Mucha-Habermann disease, Multifocal Motor Neuropathy (MMN) or MMNCB, Multiple sclerosis, Myasthenia gravis, Myositis, Narcolepsy, Neonatal Lupus, Neuromyelitis optica, Neutropenia, Ocular cicatricial pemphigoid, Optic neuritis, Palindromic rheumatism (PR), PANDAS, Paraneoplastic cerebellar degeneration (PCD), Paroxysmal nocturnal hemoglobinuria (PNH), Parry Romberg syndrome, Pars planitis (peripheral uveitis), Parsonnage-Turner syndrome, Pemphigus, Peripheral neuropathy, Perivenous encephalomyelitis, Pernicious anemia (PA), POEMS syndrome, Polyarteritis nodosa, Polyglandular syndromes type I, II, III, Polymyalgia rheumatica, Polymyositis, Postmyocardial infarction syndrome, Postpericardiotomy syndrome, Primary biliary cirrhosis, Primary sclerosing cholangitis, Progesterone dermatitis, Psoriasis, Psoriatic arthritis, Pure red cell aplasia (PRCA), Pyoderma gangrenosum, Raynaud's phenomenon, Reactive Arthritis, Reflex sympathetic dystrophy, Relapsing polychondritis, Restless legs syndrome (RLS), Retroperitoneal fibrosis, Rheumatic fever, Rheumatoid arthritis, Sarcoidosis, Schmidt syndrome, Scleritis, Scleroderma, Sjögren's syndrome, Sperm & testicular autoimmunity, Stiff person syndrome (SPS), Subacute bacterial endocarditis (SBE), Susac's syndrome, Sympathetic ophthalmia (SO), Takayasu's

arteritis, Temporal arteritis/Giant cell arteritis, Thrombocytopenic purpura (TTP), Tolosa-Hunt syndrome (THS), Transverse myelitis, Type 1 diabetes, Ulcerative colitis (UC), Undifferentiated connective tissue disease (UCTD), Uveitis, Vasculitis, Vitiligo, Vogt-Koyanagi-Harada Disease, or Wegener's granulomatosis.

[0324] In some embodiments of the compositions and methods of the disclosure, a disease or disorder of the disclosure includes, but is not limited to, an inflammatory disease or disorder.

[0325] In some embodiments of the compositions and methods of the disclosure, a disease or disorder of the disclosure includes, but is not limited to, a metabolic disease or disorder.

[0326] In some embodiments of the compositions and methods of the disclosure, a disease or disorder of the disclosure includes, but is not limited to, a degenerative or a progressive disease or disorder. In some embodiments, the degenerative or a progressive disease or disorder includes, but is not limited to, amyotrophic lateral sclerosis (ALS), Huntington's disease, Alzheimer's disease, and aging.

[0327] In some embodiments of the compositions and methods of the disclosure, a disease or disorder of the disclosure includes, but is not limited to, an infectious disease or disorder.

[0328] In some embodiments of the compositions and methods of the disclosure, a disease or disorder of the disclosure includes, but is not limited to, a pediatric or a developmental disease or disorder.

[0329] In some embodiments of the compositions and methods of the disclosure, a disease or disorder of the disclosure includes, but is not limited to, a cardiovascular disease or disorder.

[0330] In some embodiments of the compositions and methods of the disclosure, a disease or disorder of the disclosure includes, but is not limited to, a proliferative disease or disorder. In some embodiments, the proliferative disease or disorder is a cancer. In some embodiments, the cancer includes, but is not limited to, Acute Lymphoblastic Leukemia (ALL), Acute Myeloid Leukemia (AML), Adrenocortical Carcinoma, AIDS-Related Cancers, Kaposi Sarcoma (Soft Tissue Sarcoma), AIDS-Related Lymphoma (Lymphoma), Primary CNS Lymphoma (Lymphoma), Anal Cancer, Appendix Cancer, Gastrointestinal Carcinoid Tumors, Astrocytomas, Atypical Teratoid/Rhabdoid Tumor, Central Nervous System (Brain Cancer), Basal Cell Carcinoma, Bile Duct Cancer, Bladder Cancer, Bone Cancer, Ewing Sarcoma, Osteosarcoma, Malignant Fibrous Histiocytoma, Brain Tumors, Breast Cancer, Burkitt

Lymphoma, Carcinoid Tumor, Carcinoma, Cardiac (Heart) Tumors, Embryonal Tumors, Germ Cell Tumor, Primary CNS Lymphoma, Cervical Cancer, Cholangiocarcinoma, Chordoma, Chronic Lymphocytic Leukemia (CLL), Chronic Myelogenous Leukemia (CML), Chronic Myeloproliferative Neoplasms, Colorectal Cancer , Craniopharyngioma, Cutaneous T-Cell Lymphoma, Ductal Carcinoma In Situ, Embryonal Tumors, Endometrial Cancer (Uterine Cancer), Ependymoma, Esophageal Cancer, Esthesioneuroblastoma (Head and Neck Cancer), Ewing Sarcoma (Bone Cancer), Extracranial Germ Cell Tumor, Extragonadal Germ Cell Tumor, Eye Cancer, Childhood Intraocular Melanoma, Intraocular Melanoma, Retinoblastoma, Fallopian Tube Cancer, Fibrous Histiocytoma of Bone, Malignant, and Osteosarcoma, Gallbladder Cancer, Gastric (Stomach) Cancer, Gastrointestinal Carcinoid Tumor, Gastrointestinal Stromal Tumors (GIST) (Soft Tissue Sarcoma), Childhood Gastrointestinal Stromal Tumors, Germ Cell Tumors, Childhood Extracranial Germ Cell Tumors, Extragonadal Germ Cell Tumors, Ovarian Germ Cell Tumors, Testicular Cancer, Gestational Trophoblastic Disease, Hairy Cell Leukemia, Head and Neck Cancer, Heart Tumors, Hepatocellular (Liver) Cancer, Histiocytosis, Hodgkin Lymphoma, Hypopharyngeal Cancer (Head and Neck Cancer), Intraocular Melanoma, Islet Cell Tumors, Pancreatic Neuroendocrine Tumors, Kaposi Sarcoma (Soft Tissue Sarcoma), Kidney (Renal Cell) Cancer, Langerhans Cell Histiocytosis, Laryngeal Cancer (Head and Neck Cancer), Leukemia, Lip and Oral Cavity Cancer (Head and Neck Cancer), Liver Cancer, Lung Cancer (Non-Small Cell and Small Cell), Childhood Lung Cancer, Lymphoma, Male Breast Cancer, Malignant Fibrous Histiocytoma of Bone and Osteosarcoma, Melanoma, Merkel Cell Carcinoma (Skin Cancer), Mesothelioma, Metastatic Squamous Neck Cancer with Occult Primary (Head and Neck Cancer), Midline Tract Carcinoma With NUT Gene Changes, Mouth Cancer (Head and Neck Cancer), Multiple Endocrine Neoplasia Syndromes, Multiple Myeloma/Plasma Cell Neoplasms, Mycosis Fungoides (Lymphoma), Myelodysplastic Syndromes, Myelodysplastic/Myeloproliferative Neoplasms, Nasal Cavity and Paranasal Sinus Cancer (Head and Neck Cancer), Nasopharyngeal Cancer (Head and Neck Cancer), Neuroblastoma, Non-Hodgkin Lymphoma, Non-Small Cell Lung Cancer, Oral Cancer, Lip and Oral Cavity Cancer and Oropharyngeal Cancer, Osteosarcoma and Malignant Fibrous Histiocytoma of Bone, Ovarian Cancer, Pancreatic Cancer, Pancreatic Neuroendocrine Tumors (Islet Cell Tumors), Papillomatosis, Paraganglioma, Parathyroid Cancer, Penile Cancer,

Pharyngeal Cancer (Head and Neck Cancer), Pheochromocytoma , Plasma Cell Neoplasm/Multiple Myeloma, Pleuropulmonary Blastoma, Pregnancy and Breast Cancer, Primary Central Nervous System (CNS) Lymphoma, Primary Peritoneal Cancer, Prostate Cancer, Rectal Cancer, Recurrent Cancer, Renal Cell (Kidney) Cancer, Retinoblastoma, Rhabdomyosarcoma, Childhood (Soft Tissue Sarcoma), Salivary Gland Cancer (Head and Neck Cancer), Sarcoma, Childhood Rhabdomyosarcoma (Soft Tissue Sarcoma), Childhood Vascular Tumors (Soft Tissue Sarcoma), Ewing Sarcoma (Bone Cancer), Kaposi Sarcoma (Soft Tissue Sarcoma), Osteosarcoma (Bone Cancer), Uterine Sarcoma, Sézary Syndrome, Lymphoma, Skin Cancer, Small Cell Lung Cancer, Small Intestine Cancer, Soft Tissue Sarcoma, Squamous Cell Carcinoma of the Skin, Squamous Neck Cancer, Stomach (Gastric) Cancer, T-Cell Lymphoma, Testicular Cancer, Throat Cancer (Head and Neck Cancer), Nasopharyngeal Cancer, Oropharyngeal Cancer, Hypopharyngeal Cancer, Thymoma and Thymic Carcinoma , Thyroid Cancer, Transitional Cell Cancer of the Renal Pelvis and Ureter, Renal Cell Cancer, Urethral Cancer, Uterine Sarcoma, Vaginal Cancer, Vascular Tumors (Soft Tissue Sarcoma), Vulvar Cancer, Wilms Tumor and Other Childhood Kidney Tumors.

[0331] In some embodiments of the methods of the disclosure, a subject of the disclosure has been diagnosed with the disease or disorder. In some embodiments, the subject of the disclosure presents at least one sign or symptom of the disease or disorder. In some embodiments, the subject has a biomarker predictive of a risk of developing the disease or disorder. In some embodiments, the biomarker is a genetic mutation.

[0332] In some embodiments of the methods of the disclosure, a subject of the disclosure is female. In some embodiments of the methods of the disclosure, a subject of the disclosure is male. In some embodiments, a subject of the disclosure has two XX or XY chromosomes. In some embodiments, a subject of the disclosure has two XX or XY chromosomes and a third chromosome, either an X or a Y.

[0333] In some embodiments of the methods of the disclosure, a subject of the disclosure is a neonate, an infant, a child, an adult, a senior adult, or an elderly adult. In some embodiments of the methods of the disclosure, a subject of the disclosure is at least 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 or 31 days old. In some embodiments of the methods of the disclosure, a subject of the disclosure is at least 1, 2, 3, 4, 5,

6, 7, 8, 9, 10, 11 or 12 months old. In some embodiments of the methods of the disclosure, a subject of the disclosure is at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 or any number of years or partial years in between of age.

[0334] In some embodiments of the methods of the disclosure, a subject of the disclosure is a mammal. In some embodiments, a subject of the disclosure is a non-human mammal.

[0335] In some embodiments of the methods of the disclosure, a subject of the disclosure is a human.

[0336] In some embodiments of the methods of the disclosure, a therapeutically effective amount comprises a single dose of a composition of the disclosure. In some embodiments, a therapeutically effective amount comprises a therapeutically effective amount comprises at least one dose of a composition of the disclosure. In some embodiments, a therapeutically effective amount comprises a therapeutically effective amount comprises one or more dose(s) of a composition of the disclosure.

[0337] In some embodiments of the methods of the disclosure, a therapeutically effective amount eliminates a sign or symptom of the disease or disorder. In some embodiments, a therapeutically effective amount reduces a severity of a sign or symptom of the disease or disorder.

[0338] In some embodiments of the methods of the disclosure, a therapeutically effective amount eliminates the disease or disorder.

[0339] In some embodiments of the methods of the disclosure, a therapeutically effective amount prevents an onset of a disease or disorder. In some embodiments, a therapeutically effective amount delays the onset of a disease or disorder. In some embodiments, a therapeutically effective amount reduces the severity of a sign or symptom of the disease or disorder. In some embodiments, a therapeutically effective amount improves a prognosis for the subject.

[0340] In some embodiments of the methods of the disclosure, a composition of the disclosure is administered to the subject systemically. In some embodiments, the composition of the disclosure is administered to the subject by an intravenous route. In some embodiments, the composition of the disclosure is administered to the subject by an injection or an infusion.

[0341] In some embodiments of the methods of the disclosure, a composition of the disclosure is administered to the subject locally. In some embodiments, the composition of the disclosure is administered to the subject by an intraosseous, intraocular, intracerebrospinal or intraspinal route. In some embodiments, the composition of the disclosure is administered directly to the cerebral spinal fluid of the central nervous system. In some embodiments, the composition of the disclosure is administered directly to a tissue or fluid of the eye and does not have bioavailability outside of ocular structures. In some embodiments, the composition of the disclosure is administered to the subject by an injection or an infusion.

[0342] In some embodiments, the compositions comprising the RNA-binding fusion proteins disclosed herein are formulated as pharmaceutical compositions. Briefly, pharmaceutical compositions for use as disclosed herein may comprise a fusion protein(s) or a polynucleotide encoding the fusion protein(s), optionally comprised in an AAV, which is optionally also immune orthogonal, in combination with one or more pharmaceutically or physiologically acceptable carriers, diluents or excipients. Such compositions may comprise buffers such as neutral buffered saline, phosphate buffered saline and the like; carbohydrates such as glucose, mannose, sucrose or dextrans, mannitol; proteins; polypeptides or amino acids such as glycine; antioxidants; chelating agents such as EDTA or glutathione; adjuvants (e.g., aluminum hydroxide); and preservatives. Compositions of the disclosure may be formulated for oral, intravenous, topical, enteral, intraocular, and/or parenteral administration. In certain embodiments, the compositions of the present disclosure are formulated for intravenous administration.

Example Embodiments:

[0343] Embodiment 1. A composition comprising:

(a) a sequence comprising a guide RNA (gRNA) that specifically binds a target sequence within an RNA molecule and

(b) a sequence encoding a fusion protein, the sequence comprising a sequence encoding a first RNA-binding polypeptide and a sequence encoding a second RNA-binding polypeptide,

wherein neither the first RNA-binding polypeptide nor the second RNA-binding polypeptide comprises a significant DNA-nuclease activity,

wherein the first RNA-binding polypeptide and the second RNA-binding polypeptide are not identical, and

wherein the second RNA-binding polypeptide comprises an RNA-nuclease activity;

or

a composition comprising nucleic acid sequence encoding a fusion protein, the fusion protein comprising a first RNA-binding polypeptide and a second RNA-binding polypeptide, wherein the first RNA-binding polypeptide is not a guided RNA-binding polypeptide, wherein the first RNA-binding polypeptide and the second RNA-binding polypeptide are not identical, and wherein the second RNA-binding polypeptide comprises an RNA-nuclease activity.

Embodiment 2. The composition of embodiment 1, wherein the target sequence comprises at least one repeated sequence.

Embodiment 3. The composition of embodiment 1 or 2, wherein the sequence comprising the gRNA comprises a promoter capable of expressing the gRNA in a eukaryotic cell.

Embodiment 4. The composition of embodiment 3, wherein the eukaryotic cell is an animal cell.

Embodiment 5. The composition of embodiment 4, wherein the animal cell is a mammalian cell.

Embodiment 6. The composition of embodiment 5, wherein the animal cell is a human cell.

Embodiment 7. The composition of any one of embodiments 1-6, wherein the promoter is a constitutively active promoter.

Embodiment 8. The composition of any one of embodiments 1-7, wherein the promoter is isolated or derived from a promoter capable of driving expression of an RNA polymerase.

Embodiment 9. The composition of embodiment 8, wherein the promoter is isolated or derived from a U6 promoter.

Embodiment 10. The composition of any one of embodiments 1-7, wherein the promoter is isolated or derived from a promoter capable of driving expression of a transfer RNA (tRNA).

Embodiment 11. The composition of embodiment 10, wherein the promoter is isolated or derived from an alanine tRNA promoter, an arginine tRNA promoter, an asparagine tRNA promoter, an aspartic acid tRNA promoter, a cysteine tRNA promoter, a glutamine tRNA promoter, a glutamic acid tRNA promoter, a glycine tRNA promoter, a histidine tRNA promoter, an isoleucine tRNA promoter, a leucine tRNA promoter, a lysine tRNA promoter, a methionine tRNA promoter, a phenylalanine tRNA promoter, a proline tRNA promoter, a serine tRNA promoter, a threonine tRNA promoter, a tryptophan tRNA promoter, a tyrosine tRNA promoter, or a valine tRNA promoter.

Embodiment 12. The composition of embodiment 10, wherein the promoter is isolated or derived from a valine tRNA promoter.

Embodiment 13. The composition of any one of embodiments 1-12, wherein the sequence comprising the gRNA comprises a spacer sequence that specifically binds to the target RNA sequence.

Embodiment 14. The composition of embodiment 13, wherein the spacer sequence has at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 87%, 90%, 95%, 97%, 99% or any percentage in between of complementarity to the target RNA sequence.

Embodiment 15. The composition of embodiment 13, wherein the spacer sequence has 100% complementarity to the target RNA sequence.

Embodiment 16. The composition of any one of embodiments 13-15, wherein the spacer sequence comprises or consists of 20 nucleotides.

Embodiment 17. The composition of any one of embodiments 13-15, wherein the spacer sequence comprises or consists of 21 nucleotides.

Embodiment 18. The composition of embodiment 17, wherein the spacer sequence comprises the sequence UGGAGCGAGCAUCCCCCAAA (SEQ ID NO: 1), GUUUGGGGGAUGCUCGCUCCA (SEQ ID NO: 2), CCCUCACUGCUGGGGAGUCC (SEQ ID NO: 3), GGACUCCCCAGCAGUGAGGG (SEQ ID NO: 4), GCAACUGGAUCAAUUUGCUG (SEQ ID NO: 5), GCAGCAAAUUGAUCCAGUUGC (SEQ ID NO: 6), GCAUUCUUAUCUGGUCAGUGC (SEQ ID NO: 7), GCACUGACCAGAUAAGAAUG (SEQ ID NO: 8), GAGCAGCAGCAGCAGCAGCAG (SEQ ID NO: 9), GCAGGCAGGCAGGCAGGCAGG (SEQ ID NO: 10), GCCCCGGCCCCGGCCCCGGC (SEQ ID NO: 11) , or GCTGCTGCTGCTGCTGCTGC (SEQ ID NO: 12), GGGGCCGGGGCCGGGGCCGG (SEQ ID NO: 74), GGGCCGGGGCCGGGGCCGGG (SEQ ID NO: 75), GGCCGGGGCCGGGGCCGGGG (SEQ ID NO: 76), GCCGGGGCCGGGGCCGGGGC (SEQ ID NO: 77), CCGGGGCCGGGGCCGGGGCC (SEQ ID NO: 78), CGGGGCCGGGGCCGGGGCCG (SEQ ID NO: 79).

Embodiment 19. The composition of any one of embodiments 1-18, wherein the sequence comprising the gRNA comprises a scaffold sequence that specifically binds to the first RNA binding protein.

Embodiment 20. The composition of embodiment 19, wherein the scaffold sequence comprises a stem-loop structure.

Embodiment 21. The composition of embodiment 19 or 20, wherein the scaffold sequence comprises or consists of 90 nucleotides.

Embodiment 22. The composition of embodiment 19 or 20, wherein the scaffold sequence comprises or consists of 93 nucleotides.

Embodiment 23. The composition of embodiment 22, wherein the scaffold sequence comprises the sequence GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUU AUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 13).

Embodiment 24. The composition of embodiment 16, wherein the spacer sequence comprises the sequence GUGAUAAGUGGAAUGCCAUG (SEQ ID NO: 14), CUGGUGAACUUCCGAUAGUG (SEQ ID NO: 15), or GAGATATAGCCTGGTGGTTC (SEQ ID NO: 16).

Embodiment 25. The composition of embodiment 19 or 24, wherein the scaffold sequence comprises a step-loop structure.

Embodiment 26. The composition of embodiment 25, wherein the scaffold sequence comprises or consists of 85 nucleotides.

Embodiment 27. The composition of embodiment 26, wherein the scaffold sequence comprises the sequence GGACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG CACCGAGUCGGUGCUUUUU (SEQ ID NO: 17).

Embodiment 28. The composition of embodiment 16, wherein the spacer sequence comprises the sequence at least 1, 2, 3, 4, 5, 6, or 7 repeats of the sequence CUG (SEQ ID NO: 18), CCUG (SEQ ID NO: 19), CAG (SEQ ID NO: 80), GGGGCC (SEQ ID NO: 81) or any combination thereof.

Embodiment 29. The composition of embodiment 28, wherein the sequence comprising the gRNA comprises a scaffold sequence that specifically binds to the first RNA binding protein.

Embodiment 30. The composition of embodiment 29, wherein the scaffold sequence comprises a stem-loop structure.

Embodiment 31. The composition of embodiment 29 or 30, wherein the scaffold sequence comprises or consists of 90 nucleotides.

Embodiment 32. The composition of embodiment 30 or 31, wherein the scaffold sequence comprises or consists of 93 nucleotides.

Embodiment 33. The composition of embodiment 32, wherein the scaffold sequence comprises the sequence GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 82) or GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 83).

Embodiment 34. The composition of any one of embodiments 1-33, wherein the gRNA does not bind or does not selectively bind to a second sequence within the RNA molecule.

Embodiment 35. The composition of embodiment 34, wherein an RNA genome or an RNA transcriptome comprises the RNA molecule.

Embodiment 36. The composition of any one of embodiments 1-35, wherein the first RNA binding protein comprises a CRISPR-Cas protein.

Embodiment 37. The composition of embodiment 36, wherein the CRISPR-Cas protein is a Type II CRISPR-Cas protein.

Embodiment 38. The composition of embodiment 37, wherein the first RNA binding protein comprises a Cas9 polypeptide or an RNA-binding portion thereof.

Embodiment 39. The composition of embodiment 36, wherein the CRISPR-Cas protein is a Type V CRISPR-Cas protein.

Embodiment 40. The composition of embodiment 39, wherein the first RNA binding protein comprises a Cpf1 polypeptide or an RNA-binding portion thereof.

Embodiment 41. The composition of embodiment 36, wherein the CRISPR-Cas protein is a Type VI CRISPR-Cas protein.

Embodiment 42. The composition of embodiment 41, wherein the first RNA binding protein comprises a Cas13 polypeptide or an RNA-binding portion thereof.

Embodiment 43. The composition of any one of embodiments 36-42, wherein the CRISPR-Cas protein comprises a native RNA nuclease activity.

Embodiment 44. The composition of embodiment 43, wherein the native RNA nuclease activity is reduced or inhibited.

Embodiment 45. The composition of embodiment 43, wherein the native RNA nuclease activity is increased or induced.

Embodiment 46. The composition of any one of embodiments 36-45, wherein the CRISPR-Cas protein comprises a native DNA nuclease activity and wherein the native DNA nuclease activity is inhibited.

Embodiment 47. The composition of embodiment 46, wherein the CRISPR-Cas protein comprises a mutation.

Embodiment 48. The composition of embodiment 47, wherein a nuclease domain of the CRISPR-Cas protein comprises the mutation.

Embodiment 49. The composition of embodiment 47, wherein the mutation occurs in a nucleic acid encoding the CRISPR-Cas protein.

Embodiment 50. The composition of embodiment 47, wherein the mutation occurs in an amino acid encoding the CRISPR-Cas protein.

Embodiment 51. The composition of any one of embodiments 47-50, wherein the mutation comprises a substitution, an insertion, a deletion, a frameshift, an inversion, or a transposition.

Embodiment 52. The composition of any one of embodiments 47-50, wherein the mutation comprises a deletion of a nuclease domain, a binding site within the nuclease domain, an active site within the nuclease domain, or at least one essential amino acid residue within the nuclease domain.

Embodiment 53. The composition of any one of embodiments 1-35, wherein the first RNA binding protein comprises a Pumilio and FBF (PUF) protein.

Embodiment 54. The composition of embodiment 53, wherein the first RNA binding protein comprises a Pumilio-based assembly (PUMBY) protein.

Embodiment 55. The composition of any one of embodiments 1-54, wherein the first RNA binding protein does not require multimerization for RNA-binding activity.

Embodiment 56. The composition of embodiment 55, wherein the first RNA binding protein is not a monomer of a multimer complex

Embodiment 57. The composition of embodiment 55, wherein a multimer protein complex does not comprise the first RNA binding protein.

Embodiment 58. The composition of any one of embodiments 1-57, wherein the first RNA binding protein selectively binds to a target sequence within the RNA molecule.

Embodiment 59. The composition of embodiment 58, wherein the first RNA binding protein does not comprise an affinity for a second sequence within the RNA molecule.

Embodiment 60. The composition of embodiment 58 or 59, wherein the first RNA binding protein does not comprise a high affinity for or selectively bind a second sequence within the RNA molecule.

Embodiment 61. The composition of embodiment 60, wherein an RNA genome or an RNA transcriptome comprises the RNA molecule.

Embodiment 62. The composition of any one of embodiments 1-61, wherein the first RNA binding protein comprises between 2 and 1300 amino acids, inclusive of the endpoints.

Embodiment 63. The composition of any one of embodiments 1-62, wherein the sequence encoding the first RNA binding protein further comprises a sequence encoding a nuclear localization signal (NLS).

Embodiment 64. The composition of embodiment 63, wherein the sequence encoding a nuclear localization signal (NLS) is positioned 3' to the sequence encoding the first RNA binding protein.

Embodiment 65. The composition of embodiment 64, wherein the first RNA binding protein comprises an NLS at a C-terminus of the protein.

Embodiment 66. The composition of any one of embodiments 1-62, wherein the sequence encoding the first RNA binding protein further comprises a first sequence encoding a first NLS and a second sequence encoding a second NLS.

Embodiment 67. The composition of embodiment 66, wherein the sequence encoding the first NLS or the second NLS is positioned 3' to the sequence encoding the first RNA binding protein.

Embodiment 68. The composition of embodiment 67, wherein the first RNA binding protein comprises the first NLS or the second NLS at a C-terminus of the protein.

Embodiment 69. The composition of any one of embodiments 1-68, wherein the second RNA binding protein comprises or consists of a nuclease domain.

Embodiment 70. The composition of embodiment 69, wherein the sequence encoding the second RNA binding protein comprises or consists of an RNAse.

Embodiment 71. The composition of embodiment 70, wherein the second RNA binding protein comprises or consists of an RNAse1.

Embodiment 72. The composition of embodiment 71, wherein the RNAse1 protein comprises or consists of SEQ ID NO: 20.

Embodiment 73. The composition of embodiment 72, wherein the second RNA binding protein comprises or consists of an RNAse4.

Embodiment 74. The composition of embodiment 73, wherein the RNAse4 protein comprises or consists of: (SEQ ID NO: 21.

Embodiment 75. The composition of embodiment 70, wherein the second RNA binding protein comprises or consists of an RNAse6.

Embodiment 76. The composition of embodiment 75, wherein the RNAse6 protein comprises or consists of SEQ ID NO: 22.

Embodiment 77. The composition of embodiment 70, wherein the second RNA binding protein comprises or consists of an RNAse7.

Embodiment 78. The composition of embodiment 77, wherein the RNAse7 protein comprises or consists of SEQ ID NO: 23.

Embodiment 79. The composition of embodiment 70, wherein the second RNA binding protein comprises or consists of an RNAse8.

Embodiment 80. The composition of embodiment 79, wherein the RNAse8 protein comprises or consists of SEQ ID NO: 24.

Embodiment 81. The composition of embodiment 70, wherein the second RNA binding protein comprises or consists of an RNAse2.

Embodiment 82. The composition of embodiment 81, wherein the RNAse2 protein comprises or consists of SEQ ID NO: 25.

Embodiment 83. The composition of embodiment 70, wherein the second RNA binding protein comprises or consists of an RNAse6PL.

Embodiment 84. The composition of embodiment 83, wherein the RNAse6PL protein comprises or consists of SEQ ID NO: 26.

Embodiment 85. The composition of embodiment 70, wherein the second RNA binding protein comprises or consists of an RNAseL.

Embodiment 86. The composition of embodiment 85, wherein the RNAseL protein comprises or consists of SEQ ID NO: 27.

Embodiment 87. The composition of embodiment 70, wherein the second RNA binding protein comprises or consists of an RNAseT2.

Embodiment 88. The composition of embodiment 87, wherein the RNAseT2 protein comprises or consists of SEQ ID NO: 28.

Embodiment 89. The composition of embodiment 70, wherein the second RNA binding protein comprises or consists of an RNAse11.

Embodiment 90. The composition of embodiment 89, wherein the RNAse11 comprises or consists of SEQ ID NO: 29.

Embodiment 91. The composition of embodiment 70, wherein the second RNA binding protein comprises or consists of an RNAseT2-like.

Embodiment 92. The composition of embodiment 91, wherein the RNAseT2-like protein comprises or consists of SEQ ID NO: 30.

Embodiment 93. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a NOB1 polypeptide.

Embodiment 94. The composition of embodiment 93, wherein the NOB1 polypeptide comprises or consists of SEQ ID NO: 31.

Embodiment 95. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of an endonuclease.

Embodiment 96. The composition of embodiment 95, wherein the second RNA binding protein comprises or consists of an endonuclease V (ENDOV).

Embodiment 97. The composition of embodiment 96, wherein the ENDOV protein comprises or consists of SEQ ID NO: 32.

Embodiment 98. The composition of embodiment 95, wherein the second RNA binding protein comprises or consists of an endonuclease G (ENDOG).

Embodiment 99. The composition of embodiment 98, wherein the ENDOG protein comprises or consists of SEQ ID NO: 33.

Embodiment 100. The composition of embodiment 95, wherein the second RNA binding protein comprises or consists of an endonuclease D1 (ENDOD1).

Embodiment 101. The composition of embodiment 100, wherein the ENDOD1 protein comprises or consists of SEQ ID NO: 34.

Embodiment 102. The composition of embodiment 95, wherein the second RNA binding protein comprises or consists of a Human flap endonuclease-1 (hFEN1).

Embodiment 103. The composition of embodiment 102, wherein the hFEN1 protein comprises or consists of SEQ ID NO: 35.

Embodiment 104. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a human Schlafen 14 (hSLFN14) polypeptide.

Embodiment 105. The composition of embodiment 104, wherein the hSLFN14 polypeptide comprises or consists of SEQ ID NO: 36.

Embodiment 106. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a human beta-lactamase-like protein 2 (hLACTB2) polypeptide.

Embodiment 107. The composition of embodiment 106, wherein the hLACTB2 polypeptide comprises or consists of SEQ ID NO: 37.

Embodiment 108. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of an apurinic/apyrimidinic (AP) endodeoxyribonuclease (APEX2) polypeptide.

Embodiment 109. The composition of embodiment 108, wherein the APEX2 polypeptide comprises or consists of SEQ ID NO: 38.

Embodiment 110. The composition of embodiment 108, wherein the APEX2 polypeptide comprises or consists of: SEQ ID NO: 39.

Embodiment 111. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of an angiogenin (ANG) polypeptide.

Embodiment 112. The composition of embodiment 111, wherein the ANG polypeptide comprises or consists of SEQ ID NO: 40.

Embodiment 113. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a heat responsive protein 12 (HRSP12) polypeptide.

Embodiment 114. The composition of embodiment 113, wherein the HRSP12 polypeptide comprises or consists of SEQ ID NO: 41.

Embodiment 115. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a Zinc Finger CCCH-Type Containing 12A (ZC3H12A) polypeptide.

Embodiment 116. The composition of embodiment 115, wherein the ZC3H12A polypeptide comprises or consists of SEQ ID NO: 42.

Embodiment 117. The composition of embodiment 115, wherein the ZC3H12A polypeptide comprises or consists of SEQ ID NO: 43.

Embodiment 118. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a Reactive Intermediate Imine Deaminase A (RIDA) polypeptide.

Embodiment 119. The composition of embodiment 118, wherein the RIDA polypeptide comprises or consists of SEQ ID NO: 44.

Embodiment 120. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a Phospholipase D Family Member 6 (PDL6) polypeptide.

Embodiment 121. The composition of embodiment 120, wherein the PDL6 polypeptide comprises or consists of: (SEQ ID NO: 126.

Embodiment 122. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a Endonuclease III-like **protein** 1 (NTHL) polypeptide.

Embodiment 123. The composition of embodiment 122, wherein the NTHL polypeptide comprises or consists of SEQ ID NO: 123.

Embodiment 124. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a Mitochondrial ribonuclease P catalytic subunit (KIAA0391) polypeptide.

Embodiment 125. The composition of embodiment 124, wherein the KIAA0391 polypeptide comprises or consists of SEQ ID NO: 127.

Embodiment 126. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of an apurinic or apyrimidinic site lyase (APEX1) polypeptide.

Embodiment 127. The composition of embodiment 126, wherein the APEX1 polypeptide comprises or consists of SEQ ID NO: 125.

Embodiment 128. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of an argonaute 2 (AGO2) polypeptide.

Embodiment 129. The composition of embodiment 128, wherein the AGO2 polypeptide comprises or consists of SEQ ID NO: 128.

Embodiment 130. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a mitochondrial nuclease EXOG (EXOG) polypeptide.

Embodiment 131. The composition of embodiment 130, wherein the EXOG polypeptide comprises or consists of SEQ ID NO: 129.

Embodiment 132. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a Zinc Finger CCCH-Type Containing 12D (ZC3H12D) polypeptide.

Embodiment 133. The composition of embodiment 132, wherein the ZC3H12D polypeptide comprises or consists of SEQ ID NO: 130.

Embodiment 134. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of an endoplasmic reticulum to nucleus signaling 2 (ERN2) polypeptide.

Embodiment 135. The composition of embodiment 134, wherein the ERN2 polypeptide comprises or consists of SEQ ID NO: 131.

Embodiment 136. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a pelota mRNA surveillance and ribosome rescue factor (PELO) polypeptide.

Embodiment 137. The composition of embodiment 136, wherein the PELO polypeptide comprises or consists of SEQ ID NO: 132.

Embodiment 138. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a YBEY metallopeptidase (YBEY) polypeptide.

Embodiment 139. The composition of embodiment 138, wherein the YBEY polypeptide comprises or consists of SEQ ID NO: 133.

Embodiment 140. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a cleavage and polyadenylation specific factor 4 like (CPSF4L) polypeptide.

Embodiment 141. The composition of embodiment 140, wherein the CPSF4L comprises or consists of SEQ ID NO: 134.

Embodiment 142. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of an hCG_2002731polypeptide.

Embodiment 143. The composition of embodiment 142, wherein the hCG_2002731 polypeptide comprises or consists of SEQ ID NO: 135.

Embodiment 144. The composition of embodiment 142, wherein the hCG_2002731 polypeptide comprises or consists of SEQ ID NO: 136.

Embodiment 145. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of an Excision Repair Cross-Complementation Group 1 (ERCC1) polypeptide.

Embodiment 146. The composition of embodiment 145, wherein the ERCC1 polypeptide comprises or consists of SEQ ID NO: 137.

Embodiment 147. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a ras-related C3 botulinum toxin substrate 1 isoform (RAC1) polypeptide.

Embodiment 148. The composition of embodiment 147, wherein the RAC1 polypeptide comprises or consists of SEQ ID NO: 138.

Embodiment 149. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a Ribonuclease A A1 (RAA1) polypeptide.

Embodiment 150. The composition of embodiment 149, wherein the RAA1 polypeptide comprises or consists of SEQ ID NO: 139.

Embodiment 151. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a Ras Related Protein (RAB1) polypeptide.

Embodiment 152. The composition of embodiment 151, wherein the RAB1 polypeptide comprises or consists of SEQ ID NO: 140.

Embodiment 153. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a DNA Replication Helicase/Nuclease 2 (DNA2) polypeptide.

Embodiment 154. The composition of embodiment 153, wherein the DNA2 polypeptide comprises or consists of SEQ ID NO: 141.

Embodiment 155. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a FLJ35220 polypeptide.

Embodiment 156. The composition of embodiment 155, wherein the FLJ35220 polypeptide comprises or consists of SEQ ID NO: 142.

Embodiment 157. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a FLJ13173 polypeptide.

Embodiment 158. The composition of embodiment 157, wherein the FLJ13173 polypeptide comprises or consists of: (SEQ ID NO: 143.

Embodiment 159. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a DNA repair endonuclease XPF (ERCC4) polypeptide.

Embodiment 160. The composition of embodiment 159, wherein the ERCC4 polypeptide comprises or consists of SEQ ID NO: 64.

Embodiment 161. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(K41R)) polypeptide.

Embodiment 162. The composition of embodiment 161, wherein the Rnase1(K41R) polypeptide comprises or consists of SEQ ID NO: 116.

Embodiment 163. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(K41R, D121E)) polypeptide.

Embodiment 164. The composition of embodiment 163, wherein the Rnase1 (Rnase1(K41R, D121E)) polypeptide comprises or consists of SEQ ID NO: 117.

Embodiment 165. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(K41R, D121E, H119N)) polypeptide.

Embodiment 166. The composition of embodiment 165, wherein the Rnase1 (Rnase1(K41R, D121E, H119N)) polypeptide comprises or consists of SEQ ID NO: 118.

Embodiment 167. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(H119N)) polypeptide.

Embodiment 168. The composition of embodiment 167, wherein the Rnase1 (Rnase1(H119N)) polypeptide comprises or consists of SEQ ID NO: 119.

Embodiment 169. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(R39D, N67D, N88A, G89D, R91D, H119N)) polypeptide.

Embodiment 170. The composition of embodiment 169, wherein the Rnase1 (Rnase1(R39D, N67D, N88A, G89D, R91D, H119N)) polypeptide comprises or consists of SEQ ID NO: 120.

Embodiment 171. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(R39D, N67D, N88A, G89D, R91D, H119N)) polypeptide.

Embodiment 172. The composition of embodiment 171, wherein the Rnase1 (Rnase1(R39D, N67D, N88A, G89D, R91D, H119N, K41R, D121E)) polypeptide comprises or consists of SEQ ID NO: 121.

Embodiment 173. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of a mutated Rnase1 (Rnase1(R39D, N67D, N88A, G89D, R91D, H119N)) polypeptide.

Embodiment 174. The composition of embodiment 173, wherein the Rnase1 (Rnase1(R39D, N67D, N88A, G89D, R91D)) polypeptide comprises or consists of SEQ ID NO: 122.

Embodiment 175. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of Teneurin Transmembrane Protein 1 (TENM1) polypeptide.

Embodiment 176. The composition of embodiment 175, wherein the TENM1 polypeptide comprises or consists of SEQ ID NO: 144.

Embodiment 177. The composition of embodiment 69, wherein the second RNA binding protein comprises or consists of Teneurin Transmembrane Protein 2 (TENM2) polypeptide.

Embodiment 178. The composition of embodiment 177, wherein the TENM2 polypeptide comprises or consists of SEQ ID NO: 145.

Embodiment 179. A composition comprising a sequence encoding a target RNA-binding fusion protein comprising (a) a sequence encoding a first RNA-binding polypeptide or portion thereof; and (b) a sequence encoding a second RNA-binding polypeptide, wherein the first RNA-biding polypeptide binds a target RNA not guided by a gRNA sequence, and wherein the second RNA-binding polypeptide comprises RNA-nuclease activity.

Embodiment 180. The composition of embodiment 179, wherein the first RNA-binding polypeptide or portion thereof is a PUF, PUMBY, or PPR polypeptide or portion thereof.

Embodiment 181. A method for modifying the level of expression of an RNA molecule or a protein encoded by the RNA molecule, the method comprising contacting the composition of embodiments 1 or 179 and the RNA molecule under conditions suitable for binding of the fusion protein or a portion thereof to the RNA molecule.

EXAMPLES

Example 1: Methods

[0344] HEK-293 cells were cultured in DMEM with 10% FBS and 1% penicillin/streptomycin (GIBCO) and passaged at 90%–100% confluency. Cells were seeded at 1x10^5 cells per well of a 24-well plate for RNA isolation or .5x10^5 cells per well of a 96-well plate for luciferase assays. RNA isolations were carried out with RNAeasy columns (Qiagen) according to the manufacturer's protocol. RNA quality and concentrations were estimated using the Nanodrop spectrophotometer. cDNA preparation was done using Superscript III (Thermo) with random primers according to the manufacturer's protocol. qPCR was carried out with primers in a sequence adjacent to the CTG repeat in the reporter plasmid using the following primers:

Forward Primer	TetCTG_DMPK_E15_F	TCGGAGCGGTTGTGAACT	SEQ ID NO: 83

Reverse Primer	TetCTG_DMPK_E15_R	GTTCGCCGTTGTTCTGTC	SEQ ID NO: 84

[0345] Relative abundance of the CTG repeat reporter was determined by normalization to GAPDH. Next, levels of the CTG-targeting sgRNA were normalized to a non-targeting sgRNA to generate a final value reported in the associated data package.

CTG-targeting spacer	AGCAGCAGCAGCAGCAGCAG	SEQ ID NO: 85
Non-targeting control spacer (λ2)	GTGATAAGTGGAATGCCATG	SEQ ID NO: 86
sgRNA scaffold (N's indicate spacer)	GNNNNNNNNNNNNNNNNNNNGUUUAAGAGCUAUGCUG GAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUA UCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU	SEQ ID NO: 87

[0346] Luciferase assays were conducted with the Promega Dual Luciferase kit according to manufacturer's directions. Reported values are a ratio of firefly and renilla luciferase luminescence readings.

Example 2: RNA-guided cleavage of repetitive RNA molecules and mRNA molecules

[0347] Experimental Design: Various fusions of human proteins with annotated RNA endonuclease activity and Cas9 (*Streptococcus pyogenes* or *Campylobacter jejuni*) were constructed. Plasmids encoding the above fusions were co-transfected with either a repeat-containing plasmid or a luciferase assay plasmid (comprising an mRNA sequence encoding a luciferase protein). A level of CTG repeat-containing RNA was measured with qPCR in the condition in which an RNA endonuclease/Cas9 fusion was co-transfected with a repetitive RNA. A level of luciferase protein was measured using a luminescence assay in the condition in which an RNA endonuclease/Cas9 fusion was co-transfected with a luciferase assay plasmid. All measurements were normalized to a non-targeting sgRNA control construct (Figures 3A-5 and Figure 9).

Example 3: RNA-guided cleavage of Viral RNA Molecules

[0348] A549 cells were cultured in DMEM with 10% FBS and 1% penicillin/streptomycin (GIBCO) and passaged at 90%–100% confluency. Cells were seeded at 1x10^5 cells per well of a 24-well plate for RNA isolation or .5x10^5 cells per well. Cells were transfected with plasmids encoding *Campylobacter jejuni* Cas9 (CjeCas9) fused to the gene NTHL1 (residues 31-312, E43) or CPSF4L (full length, E67) with plasmids encoding one of four sites in Zika NS5 RNA. CjeCas9 was driven by an EFS promoter while the guide RNAs were driven by U6 promoter. The sequences of the sgRNAs are presented in Table 1. The sequences of the constructs used in this study are presented below.

[0349] RNA isolations were carried out with RNAeasy columns (Qiagen) according to the manufacturer's protocol. RNA quality and concentrations were estimated using the Nanodrop spectrophotometer. cDNA preparation was done using Superscript III (Thermo) with random primers according to the manufacturer's protocol. qPCR was carried out with the following primers as listed in Table 2.

[0350] Figure 7 shows expression levels of Zika NS5 assessed in the presence of both E43 and E67 endonucleases with sgRNAs containing the various NS5-targeting spacer sequences as indicated in Table 2. Zika NS5 expression is displayed as fold change relative to the endonuclease loaded with an sgRNA containing a control (Lambda) spacer sequence.

[0351] Immunofluorescence microscopy was used to visualize Zika NS5 expression in the presence of E43 or E67 endonucleases fused to CjeCas9. Figure 8A shows a fluorescence microscopy image of cells transfected with CjeCas9-endonuclease fusions loaded with an sgRNA containing a Zika NS5-targeting spacer sequence. Expression of Zika NS5 is markedly decreased in the presence of CjeCas9-endonuclease fusions loaded with the appropriate Zika NS5-targeting sgRNA as compared to CjeCas9-endonuclease fusions loaded with a non-Zika NS5 targeting sgRNA (Figures 8A and 8B). Figure 6 is a list of exemplary endonucleases for use in the compositions of the disclosure.

[0352] **Table 1**: qPCR primers

GAPDH_F	CAGCCTCAAGATCATCAGCAA (SEQ ID NO: 192)
GAPDH_R	TGTGGTCATGAGTCCTTCCA (SEQ ID NO: 193)
NS5_F	GAGGAGAGTGCCAGAGTTGT (SEQ ID NO: 194)

NS5_R	TCTCTCTCCCCATCCAGTGA (SEQ ID NO: 195)

[0353] Table 2: sgRNA sequences

NS5-targeting spacer 1	gcaatgatcttcatgttgggagc (SEQ ID NO: 196)
NS5-targeting spacer 2	gaaccttgttgatgaactcttc (SEQ ID NO: 197)
NS5-targeting spacer 3	gttggtgattagagcttcattc (SEQ ID NO: 198)
NS5-targeting spacer 4	gagtgatcctcgttcaagaatcc (SEQ ID NO: 199)
Non-targeting control spacer (λ2)	GTGATAAGTGGAATGCCATG (SEQ ID NO: 200)
sgRNA scaffold (N's indicate spacer)	GNNNNNNNNNNNNNNNNNNGUUUAAGAGCUAUG CUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGU CCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGU GCUUUUUUU (SEQ ID NO: 201)

[0354] A **E43-CjeCas9 and sgRNA plasmid** may comprise or consist of the sequence **(**U6**:**

N's=sgRNA spacer, **E43**, **CjeCas9**):

gtttattacagggacagcagagatccagtttggttaattaaggtaccgagggcctatttcccatgattccttcatatttgcatatacgatacaagg
ctgttagagagataattagaattaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttg
cagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttggctttatatatcttGTGGAAAGG
ACGAAACACCNNNNNNNNNNNNNNNNNNGTTTTAGTCCCTGAAGGGACTAAAAT
AAAGAGTTTGCGGGACTCTGCGGGGTTACAATCCCCTAAAACCGCTTTTTTTCCTGC
AGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTT
TTGTTCCCTTTAGTGAGGGTTAATTGCGCGAATTCGCTAGCTAGGTCTTGAAAGGAG
TGGGAATTGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCC
GAGAAGTTGGGGGGAGGGGTCGGCAATTGATCCGGTGCCTAGAGAAGGTGGCGCG
GGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGG
GAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTG
CCGCCAGAACACAGGACCGGTTCTAGAGCGCTATTTAGAACCatg**TGTTCTCCCCAA**
GAATCTGGCATGACCGCTCTTTCAGCGAGGATGTTGACGCGAAGCAGATCCCT
GGGACCTGGGGCCGGGCCACGAGGGTGTCGGGAAGAACCAGGACCGTTGCGA
CGGAGGGAAGCAGCAGCGGAAGCTCGGAAATCCCATTCTCCGGTTAAACGACC
CCGCAAGGCACAACGGCTCAGGGTTGCTTACGAGGGGAGCGATTCCGAAAAGG
GTGAAGGAGCAGAGCCCTTGAAGGTTCCAGTATGGGAACCCCAGGATTGGCAG
CAGCAGCTTGTAAACATCCGAGCAATGAGGAACAAAAAAGATGCACCTGTTGA
TCACCTCGGAACCGAACATTGTTATGATTCTAGTGCGCCGCCAAAAGTCCGCC
GGTATCAGGTTCTGTTGAGTTTGATGCTGAGTAGTCAGACTAAGGACCAGGTT
ACGGCCGGAGCAATGCAACGGCTTCGGGCACGGGGACTCACGGTCGATAGCAT
TTTGCAGACCGATGACGCAACATTGGGTAAACTCATATATCCAGTTGGCTTCTG
GCGGAGCAAAGTGAAGTACATCAAGCAGACCTCAGCCATTCTCCAACAACATT
ACGGAGGTGATATACCCGCAAGCGTAGCTGAACTGGTAGCACTGCCGGGCGTC
GGTCCCAAAATGGCACATCTGGCTATGGCGGTTGCTTGGGGAACGGTGTCTGG
TATCGCAGTTGATACGCATGTCCACCGCATCGCCAATCGGCTGAGGTGGACTA

AAAAAGCCACTAAGTCTCCTGAAGAAACACGGGCTGCTCTGGAAGAGTGGCTT
CCACGAGAGCTGTGGCATGAAATCAATGGATTGCTGGTTGGTTTCGGGCAGCA
GACATGCTTGCCCGTGCACCCCCGGTGTCATGCTTGCTTGAACCAGGCTTTGT
GCCCAGCTGCCCAGGGCCTGAGTGGAAGTGAGACACCGGGAACATCTGAGTCTGC
GACCCCGGAGAGCacaaacGCGCGAATCCTGGCCTTCGcgATTGGCATTAGCAGCAT
CGGCTGGGCATTCTCTGAAAACGACGAACTGAAGGATTGCGGCGTGCGAATTT
TCACTAAGGTCGAAAATCCCAAAACTGGTGAATCACTCGCTCTCCCTAGACGAC
TGGCACGCTCCGCACGAAAGAGGCTTGCCCGCCGCAAGGCACGCTTGAACCAT
CTTAAACACCTTATTGCAAATGAGTTTAAACTGAATTATGAGGACTACCAATCC
TTTGACGAGTCTCTTGCTAAAGCCTACAAAGGGAGCCTTATATCCCCGTATGAG
CTCCGGTTCAGAGCACTCAACGAACTGCTGTCCAAACAGGATTTTGCTCGCGT
GATTCTCCACATAGCGAAGAGGCGAGGATACGATGACATTAAAAACAGTGATG
ATAAGGAAAAAGGGGCCATACTCAAAGCGATTAAGCAAAATGAAGAGAAGCTC
GCTAACTATCAATCAGTAGGGGAGTATCTCTATAAAGAGTACTTCCAGAAGTTC
AAAGAAAATAGCAAGGAATTTACTAATGTCCGGAATAAAAAGGAGTCTTACGA
AAGATGTATTGCGCAATCTTTCCTCAAGGACGAGCTCAAATTGATTTTCAAGAA
ACAAAGGGAATTTGGGTTCAGCTTCTCAAAAAAATTTGAGGAAGAGGTTCTGA
GCGTTGCCTTTTACAAACGCGCCCTTAAGGACTTCTCACATCTCGTAGGGAATT
GTAGTTTCTTCACCGATGAAAAACGGGCGCCAAAAAATAGCCCTTTGGCTTTTA
TGTTTGTCGCTCTGACTCGCATCATTAATCTGCTCAACAACCTTAAAAACACGG
AAGGGATTCTGTACACAAAGGATGATCTGAACGCTCTGCTTAACGAAGTTTTGA
AGAACGGGACTTTGACCTACAAACAAACCAAAAAGCTTCTTGGTCTCAGTGATG
ACTACGAATTCAAGGGAGAAAAAGGGACATATTTCATCGAATTCAAGAAGTATA
AGGAGTTCATCAAAGCCTTGGGCGAGCACAACTTGTCTCAAGATGATCTCAAC
GAAATTGCTAAGGATATCACTCTGATTAAAGACGAGATCAAGCTCAAAAAGGC
GTTGGCGAAGTATGACCTTAACCAAAACCAAATAGATAGCCTCAGCAAGTTGG
AATTTAAAGATCACTTGAATATAAGTTTCAAGGCCCTTAAGTTGGTCACCCCCT
TGATGCTTGAAGGAAAGAAATATGATGAGGCATGTAATGAGCTGAATCTCAAG
GTTGCTATTAACGAAGACAAAAAAGATTTCCTCCCAGCTTTCAATGAGACTTAC
TATAAGGACGAGGTTACCAATCCTGTGGTGCTCCGAGCCATCAAAGAGTATCG
AAAGGTCCTGAATGCTTTGCTCAAAAAATACGGTAAGGTACACAAAATAAATAT
TGAGCTCGCAAGGGAGGTCGGTAAGAACCACTCCCAGCGCGCCAAAATAGAAA
AGGAACAGAATGAAAATTACAAAGCGAAAAAGGACGCCGAGCTCGAGTGCGAA
AAGCTGGGCCTGAAAATAAACAGCAAGAACATTCTCAAACTCCGCCTCTTCAAA
GAACAAAAAGAATTTTGTGCTTATAGTGGTGAGAAAATAAAAATCTCCGATCTT
CAAGACGAGAAGATGCTCGAAATAGACgcgATATATCCATATAGCAGGTCTTTTG
ACGATTCTTACATGAATAAAGTGCTTGTTTTCACTAAGCAGAATCAGGAAAAGT
TGAATCAGACCCCCTTTGAGGCCTTTGGCAACGACTCAGCAAAGTGGCAGAAG
ATCGAGGTCTTGGCTAAGAATCTTCCTACTAAGAAACAGAAAAGGATATTGGAT
AAGAACTATAAAGACAAAGAACAAAAGAACTTTAAAGACCGCAACCTCAATGA
CACCAGATACATAGCAAGATTGGTTCTGAACTACACAAAAGATTATTTGGACTT
CTTGCCGCTGTCTGATGATGAGAACACGAAACTCAACGACACGCAAAAGGGGT
CTAAAGTCCACGTCGAAGCTAAATCTGGGATGCTCACCTCAGCATTGAGGCAT
ACGTGGGGATTCTCAGCAAAGGACCGAAACAATCACCTGCACCATGCCATTGA
CGCAGTTATCATAGCGTATGCCAATAATTCAATAGTAAAAGCGTTTAGCGACTT

**CAAGAAGGAACAAGAGTCCAACAGCGCCGAGCTCTACGCAAAAAAGATTAGTG
AACTCGACTACAAAAACAAAAGAAAATTCTTTGAGCCGTTCAGCGGATTTCGAC
AGAAGGTATTGGATAAAATAGATGAAATTTTCGTGAGCAAACCCGAAAGGAAA
AAGCCCTCAGGCGCCTTGCACGAAGAGACTTTCAGGAAGGAAGAGGAATTCTA
CCAAAGCTACGGCGGAAAAGAGGGAGTTTTGAAGGCTCTCGAACTTGGAAAGA
TTAGGAAGGTGAACGGCAAGATAGTGAAAAACGGCGATATGTTCCGGGTTGAT
ATCTTCAAACATAAAAAAACGAATAAATTTTATGCTGTGCCTATATACACTATG
GACTTCGCACTTAAGGTCCTGCCGAATAAGGCGGTAGCCCGATCTAAAAAAGG
CGAAATTAAGGACTGGATTTTGATGGATGAAAATTACGAGTTCTGCTTTTCTCT
CTACAAGGATTCCCTTATATTGATACAGACGAAAGATATGCAGGAACCGGAATT
CGTGTATTACAACGCTTTTACTTCCTCTACGGTATCTTTGATTGTCTCCAAACAT
GACAACAAATTCGAAACACTCAGTAAAAACCAAAAGATTCTCTTTAAAAATGCG
AACGAGAAAGAAGTAATTGCAAAATCAATTGGCATCCAAAATTTGAAAGTTTTT
GAAAAATATATAGTATCTGCCCTCGGAGAGGTTACTAAAGCGGAATTTAGACA
GCGAGAGGACTTCAAAAAATCAGGTCCA**CCCAAGAAAAAACGCAAGGTGGAAGA
TCCGAAGAAAAAGCGAAAAGTGGATGTGtaaCGTTTTCCGGGACGCCGGCTGGATGA
TCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGC
AGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCAT
TTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTC
TGTATACCG(SEQ ID NO: 202).

[0355] A **E67-CjeCas9 and sgRNA plasmid** may comprise or consist of the sequence **(**U6**:**

N's=sgRNA spacer, **E67**, **CjeCas9**):

gtttattacagggacagcagagatccagtttggttaattaaggtaccgagggcctatttcccatgattccttcatatttgcatatacgatacaagg
ctgttagagagataattagaattaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttg
cagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttggctttatatatcttGTGGAAAGG
ACGAAACACCNNNNNNNNNNNNNNNNNNGTTTTAGTCCCTGAAGGGACTAAAAT
AAAGAGTTTGCGGGACTCTGCGGGGTTACAATCCCCTAAAACCGCTTTTTTTCCTGC
AGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTT
TTGTTCCCTTTAGTGAGGGTTAATTGCGCGAATTCGCTAGCTAGGTCTTGAAAGGAG
TGGGAATTGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCC
GAGAAGTTGGGGGGAGGGGTCGGCAATTGATCCGGTGCCTAGAGAAGGTGGCGCG
GGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGG
GAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTG
CCGCCAGAACACAGGACCGGTTCTAGAGCGCTATTTAGAACCatg**CAGGAGGTAATA
GCGGGGCTTGAGCGATTTACCTTTGCCTTCGAAAAAGACGTAGAGATGCAGAA
GGGAACCGGCCTGCTCCCATTTCAAGGTATGGACAAATCAGCATCTGCCGTGT
GCAATTTTTTCACCAAGGGTCTGTGTGAAAAGGGGAAGCTCTGTCCATTTCGCC
ATGATCGCGGAGAGAAGATGGTGGTGTGTAAGCACTGGCTGAGAGGGCTTTGC
AAAAAAGGCGACCACTGCAAATTTCTTCACCAATATGACCTGACTCGAATGCCT
GAGTGTTATTTTTACAGTAAGTTCGGTGACTGTAGCAACAAAGAATGCAGCTTC
TTGCATGTCAAACCAGCATTCAAGTCACAGGATTGCCCGTGGTACGATCAGGG
TTTTTGCAAGGACGGTCCCCTCTGCAAATATCGACACGTACCCAGAATTATGTG
CCTTAATTACCTGGTCGGCTTCTGTCCTGAAGGGCCAAAATGTCAGTTTGCTCA**

AAAAATTCGCGAGTTCAAATTGCTCCCTGGGTCTAAAATTTGGGAACCCCAGGA
TTGGCAGCAGCAGCTTGTAAACATCCGAGCAATGAGGAACAAAAAAGATGCAC
CTGTTGATCACCTCGGAACCGAACATTGTTATGATTCTAGTGCGCCGCCAAAAG
TCCGCCGGTATCAGGTTCTGTTGAGTTTGATGCTGAGTAGTCAGACTAAGGAC
CAGGTTACGGCCGGAGCAATGCAACGGCTTCGGGCACGGGGACTCACGGTCG
ATAGCATTTTGCAGACCGATGACGCAACATTGGGTAAACTCATATATCCAGTTG
GCTTCTGGCGGAGCAAAGTGAAGTACATCAAGCAGACCTCAGCCATTCTCCAA
CAACATTACGGAGGTGATATACCCGCAAGCGTAGCTGAACTGGTAGCACTGCC
GGGCGTCGGTCCCAAAATGGCACATCTGGCTATGGCGGTTGCTTGGGGAACGG
TGTCTGGTATCGCAGTTGATACGCATGTCCACCGCATCGCCAATCGGCTGAGG
TGGACTAAAAAAGCCACTAAGTCTCCTGAAGAAACACGGGCTGCTCTGGAAGA
GTGGCTTCCACGAGAGCTGTGGCATGAAATCAATGGATTGCTGGTTGGTTTCG
GGCAGCAGACATGCTTGCCCGTGCACCCCCGGTGTCATGCTTGCTTGAACCAG
GCTTTGTGCCCAGCTGCCCAGGGCCTGAGTGGAAGTGAGACACCGGGAACATCT
GAGTCTGCGACCCCGGAGAGCacaaac**GCGCGAATCCTGGCCTTCGcgATTGGCATT**
AGCAGCATCGGCTGGGCATTCTCTGAAAACGACGAACTGAAGGATTGCGGCGT
GCGAATTTTCACTAAGGTCGAAAATCCCAAAACTGGTGAATCACTCGCTCTCCC
TAGACGACTGGCACGCTCCGCACGAAAGAGGCTTGCCCGCCGCAAGGCACGCT
TGAACCATCTTAAACACCTTATTGCAAATGAGTTTAAACTGAATTATGAGGACT
ACCAATCCTTTGACGAGTCTCTTGCTAAAGCCTACAAAGGGAGCCTTATATCCC
CGTATGAGCTCCGGTTCAGAGCACTCAACGAACTGCTGTCCAAACAGGATTTT
GCTCGCGTGATTCTCCACATAGCGAAGAGGCGAGGATACGATGACATTAAAAA
CAGTGATGATAAGGAAAAAGGGGCCATACTCAAAGCGATTAAGCAAAATGAAG
AGAAGCTCGCTAACTATCAATCAGTAGGGGAGTATCTCTATAAAGAGTACTTCC
AGAAGTTCAAAGAAAATAGCAAGGAATTTACTAATGTCCGGAATAAAAAGGAG
TCTTACGAAAGATGTATTGCGCAATCTTTCCTCAAGGACGAGCTCAAATTGATT
TTCAAGAAACAAAGGGAATTTGGGTTCAGCTTCTCAAAAAAATTTGAGGAAGA
GGTTCTGAGCGTTGCCTTTTACAAACGCGCCCTTAAGGACTTCTCACATCTCGT
AGGGAATTGTAGTTTCTTCACCGATGAAAAACGGGCGCCAAAAAATAGCCCTTT
GGCTTTTATGTTTGTCGCTCTGACTCGCATCATTAATCTGCTCAACAACCTTAA
AAACACGGAAGGGATTCTGTACACAAAGGATGATCTGAACGCTCTGCTTAACG
AAGTTTTGAAGAACGGGACTTTGACCTACAAACAAACCAAAAAGCTTCTTGGTC
TCAGTGATGACTACGAATTCAAGGGAGAAAAAGGGACATATTTCATCGAATTCA
AGAAGTATAAGGAGTTCATCAAAGCCTTGGGCGAGCACAACTTGTCTCAAGAT
GATCTCAACGAAATTGCTAAGGATATCACTCTGATTAAAGACGAGATCAAGCTC
AAAAAGGCGTTGGCGAAGTATGACCTTAACCAAAACCAAATAGATAGCCTCAG
CAAGTTGGAATTTAAAGATCACTTGAATATAAGTTTCAAGGCCCTTAAGTTGGT
CACCCCCTTGATGCTTGAAGGAAAGAAATATGATGAGGCATGTAATGAGCTGA
ATCTCAAGGTTGCTATTAACGAAGACAAAAAAGATTTCCTCCCAGCTTTCAATG
AGACTTACTATAAGGACGAGGTTACCAATCCTGTGGTGCTCCGAGCCATCAAA
GAGTATCGAAAGGTCCTGAATGCTTTGCTCAAAAAATACGGTAAGGTACACAA
AATAAATATTGAGCTCGCAAGGGAGGTCGGTAAGAACCACTCCCAGCGCGCCA
AAATAGAAAAGGAACAGAATGAAAATTACAAAGCGAAAAAGGACGCCGAGCTC
GAGTGCGAAAAGCTGGGCCTGAAAATAAACAGCAAGAACATTCTCAAACTCCG
CCTCTTCAAAGAACAAAAAGAATTTTGTGCTTATAGTGGTGAGAAAATAAAAAT

**CTCCGATCTTCAAGACGAGAAGATGCTCGAAATAGACgcgATATATCCATATAGC
AGGTCTTTTGACGATTCTTACATGAATAAAGTGCTTGTTTTCACTAAGCAGAAT
CAGGAAAAGTTGAATCAGACCCCCTTTGAGGCCTTTGGCAACGACTCAGCAAA
GTGGCAGAAGATCGAGGTCTTGGCTAAGAATCTTCCTACTAAGAAACAGAAAA
GGATATTGGATAAGAACTATAAAGACAAAGAACAAAAGAACTTTAAAGACCGC
AACCTCAATGACACCAGATACATAGCAAGATTGGTTCTGAACTACACAAAAGAT
TATTTGGACTTCTTGCCGCTGTCTGATGATGAGAACACGAAACTCAACGACACG
CAAAAGGGGTCTAAAGTCCACGTCGAAGCTAAATCTGGGATGCTCACCTCAGC
ATTGAGGCATACGTGGGGATTCTCAGCAAAGGACCGAAACAATCACCTGCACC
ATGCCATTGACGCAGTTATCATAGCGTATGCCAATAATTCAATAGTAAAAGCGT
TTAGCGACTTCAAGAAGGAACAAGAGTCCAACAGCGCCGAGCTCTACGCAAAA
AAGATTAGTGAACTCGACTACAAAAACAAAAGAAAATTCTTTGAGCCGTTCAGC
GGATTTCGACAGAAGGTATTGGATAAAATAGATGAAATTTTCGTGAGCAAACCC
GAAAGGAAAAAGCCCTCAGGCGCCTTGCACGAAGAGACTTTCAGGAAGGAAGA
GGAATTCTACCAAAGCTACGGCGGAAAAGAGGGAGTTTTGAAGGCTCTCGAAC
TTGGAAAGATTAGGAAGGTGAACGGCAAGATAGTGAAAAACGGCGATATGTTC
CGGGTTGATATCTTCAAACATAAAAAAACGAATAAATTTTATGCTGTGCCTATA
TACACTATGGACTTCGCACTTAAGGTCCTGCCGAATAAGGCGGTAGCCCGATC
TAAAAAAGGCGAAATTAAGGACTGGATTTTGATGGATGAAAATTACGAGTTCTG
CTTTTCTCTCTACAAGGATTCCCTTATATTGATACAGACGAAAGATATGCAGGA
ACCGGAATTCGTGTATTACAACGCTTTTACTTCCTCTACGGTATCTTTGATTGT
CTCCAAACATGACAACAAATTCGAAACACTCAGTAAAAACCAAAAGATTCTCTT
TAAAAATGCGAACGAGAAAGAAGTAATTGCAAAATCAATTGGCATCCAAAATTT
GAAAGTTTTTGAAAAATATATAGTATCTGCCCTCGGAGAGGTTACTAAAGCGGA
ATTTAGACAGCGAGAGGACTTCAAAAAATCAGGTCCA**CCCAAGAAAAAACGCAA
GGTGGAAGATCCGAAGAAAAAGCGAAAAGTGGATGTGtaaCGTTTTCCGGGACGCCG
GCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACT
TGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAA
ATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATC
TTATCATGTCTGTATACCG** (SEQ ID NO: 203).

INCORPORATION BY REFERENCE

[0356] Every document cited herein, including any cross referenced or related patent or application is hereby incorporated herein by reference in its entirety unless expressly excluded or otherwise limited. The citation of any document is not an admission that it is prior art with respect to any invention disclosed or embodimented herein or that it alone, or in any combination with any other reference or references, teaches, suggests or discloses any such invention. Further, to the extent that any meaning or definition of a term in this document conflicts with any meaning or definition of the same term in a document incorporated by reference, the meaning or definition assigned to that term in this document shall govern.

OTHER EMBODIMENTS

[0357] While particular embodiments of the disclosure have been illustrated and described, various other changes and modifications can be made without departing from the spirit and scope of the disclosure. The scope of the appended claims includes all such changes and modifications that are within the scope of this disclosure.

CLAIMS

What is claimed is:

1. A composition comprising a nucleic acid sequence encoding an RNA-guided target RNA-binding fusion protein comprising (a) a first RNA-binding polypeptide or portion thereof; and (b) a second RNA-binding polypeptide, wherein the first RNA-binding polypeptide binds a target RNA when guided by a gRNA sequence, and wherein the second RNA-binding polypeptide comprises RNA-nuclease activity.
2. The composition of claim 1, wherein the first RNA-binding polypeptide or portion thereof is a CRISPR/Cas polypeptide or portion thereof.
3. The composition of claim 2, wherein the CRISPR/Cas polypeptide or portion thereof is selected from the group consisting of Cas9, Cpf1, Cas13a, Cas13b, Cas13c and CasRX/Cas13d, wherein the CRISPR/Cas polypeptide has native, reduced or null activity.
4. The composition of claim 1, wherein the second RNA-binding polypeptide binds RNA in a manner in which it associates with RNA.
5. The composition of claim 4, wherein the second RNA-binding polypeptide associates with RNA in a manner in which it cleaves RNA.
6. The composition of claim 1, wherein the nucleic acid sequence comprises a promoter.
7. The composition of claim 6, wherein the promoter is a constitutive promoter or a tissue-specific promoter.
8. The composition of claim 1, wherein the nucleic acid sequence further comprises a gRNA sequence, wherein the gRNA sequence comprises a spacer sequence that specifically binds a target sequence within an RNA molecule and a scaffold sequence that specifically binds to the first RNA-binding polypeptide.
9. The composition of claim 8, wherein the spacer sequence comprises a sequence comprising at least 1, 2, 3, 4, 5, 6, or 7 repeats of a sequence selected from the group consisting of: CUG (SEQ ID NO: 18), CCUG (SEQ ID NO: 19), CAG (SEQ ID NO: 80), GGGGCC (SEQ ID NO: 81), and a combination thereof.

10. The composition of claim 8, wherein the nucleic acid sequence comprises a promoter which drives expression of the gRNA sequence.
11. The composition of claim 9, wherein the promoter is a polymerase III promoter.
12. The composition of claim 10, wherein the polymerase III promoter is a U6 promoter.
13. The composition of claims 1 or 9, wherein the promoter is a tRNA promoter.
14. The composition of claims 1 or 9, wherein the fusion protein comprises an NLS, NES or tag.
15. A vector comprising the composition of claim 1 or 8.
16. The vector of claim 15, wherein the vector is selected from the group consisting of: adeno-associated virus, retrovirus, lentivirus, adenovirus, nanoparticle, micelle, liposome, lipoplex, polymersome, polyplex, and dendrimer.
17. A cell comprising the vector of claim 15.
18. The composition of claim 1, wherein the second RNA-binding polypeptide is selected from the group consisting of: RNAse1, RNAse4, RNAse6, RNAse7, RNAse8, RNAse2, RNAse6PL, RNAseL, RNAseT2, RNAse11, RNAseT2-like, NOB1, ENDOV, ENDOG, ENDOD1, hFEN1, hSLFN14, hLACTB2, APEX2, ANG, HRSP12, ZC3H12A, RIDA, PDL6, NTHL, KIAA0391, APEX1, AGO2, EXOG, ZC3H12D, ERN2, PELO, YBEY, CPSF4L, hCG_2002731, ERCC1, RAC1, RAA1, RAB1, DNA2, FLJ35220, FLJ13173, ERCC4, Rnase1(K41R), Rnase1(K41R, D121E), Rnase1(K41R, D121E, H119N), Rnase1(H119N), Rnase1(R39D, N67D, N88A, G89D, R91D, H119N), Rnase1(R39D, N67D, N88A, G89D, R91D, H119N, K41R, D121E), Rnase1(R39D, N67D, N88A, G89D, R91D), TENM1, TENM2, RNAseK, TALEN, and ZNF638.
19. A composition comprising:

(a) a guide RNA (gRNA) sequence comprising a spacer sequence that specifically binds a target sequence within an RNA molecule and a scaffold sequence that specifically binds to the first RNA-binding polypeptide;

(b) a nucleic acid sequence encoding a fusion protein, the fusion protein comprising a first RNA-binding polypeptide and a sequence encoding a second RNA-binding polypeptide,

wherein neither the first RNA-binding polypeptide nor the second RNA-binding polypeptide comprises a significant DNA-nuclease activity,

wherein the first RNA-binding polypeptide and the second RNA-binding polypeptide are not identical, and

wherein the second RNA-binding polypeptide comprises an RNA-nuclease activity.

20. A method for modifying the level of expression of a target RNA molecule or a protein encoded by the RNA molecule, the method comprising contacting the composition of claims 19 and the RNA molecule under conditions suitable for binding of the fusion protein or a portion thereof to the RNA molecule.

FIGURE 1

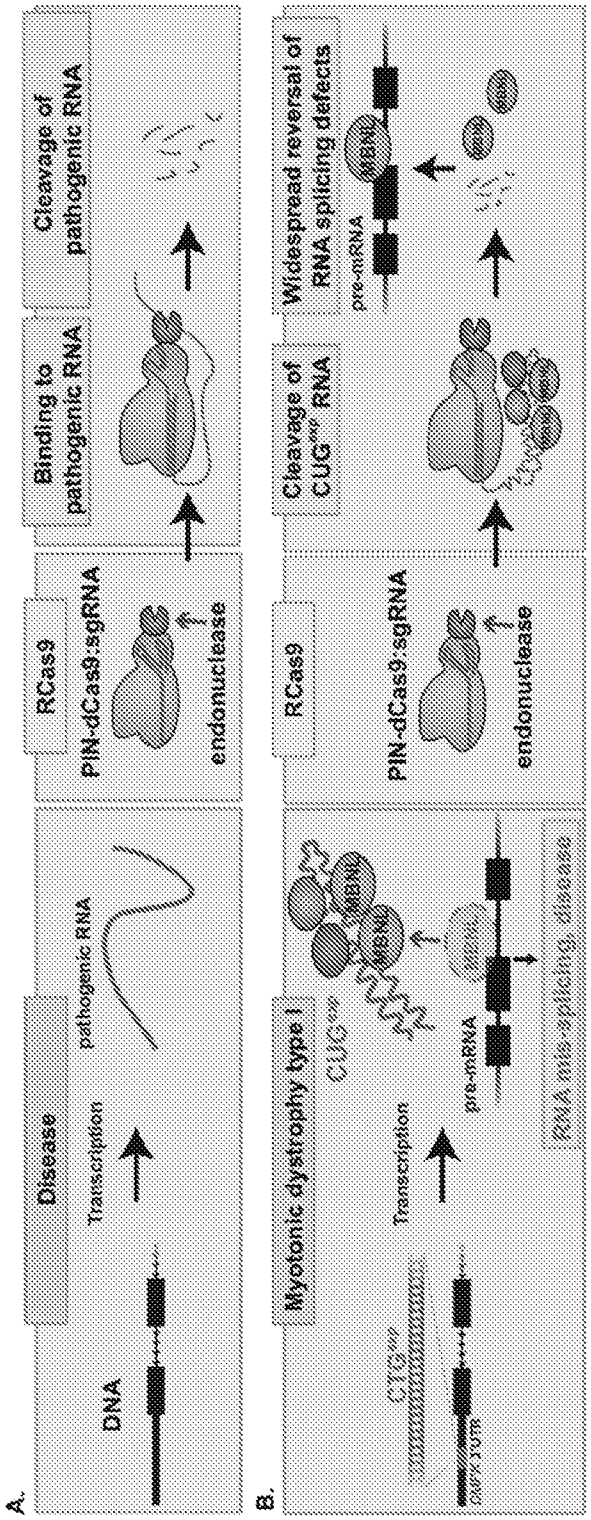

A.
Disease
RCas9
Binding to pathogenic RNA
Cleavage of pathogenic RNA
DNA
Transcription
pathogenic RNA
PIN-dCas9:sgRNA
endonuclease
B.
Myotonic dystrophy type I
RCas9
Cleavage of CUG^exp RNA
Widespread reversal of RNA splicing defects
CUG^exp
Transcription
pre-mRNA
PIN-dCas9:sgRNA
endonuclease
RNA mis-splicing disease

FIGURE 2

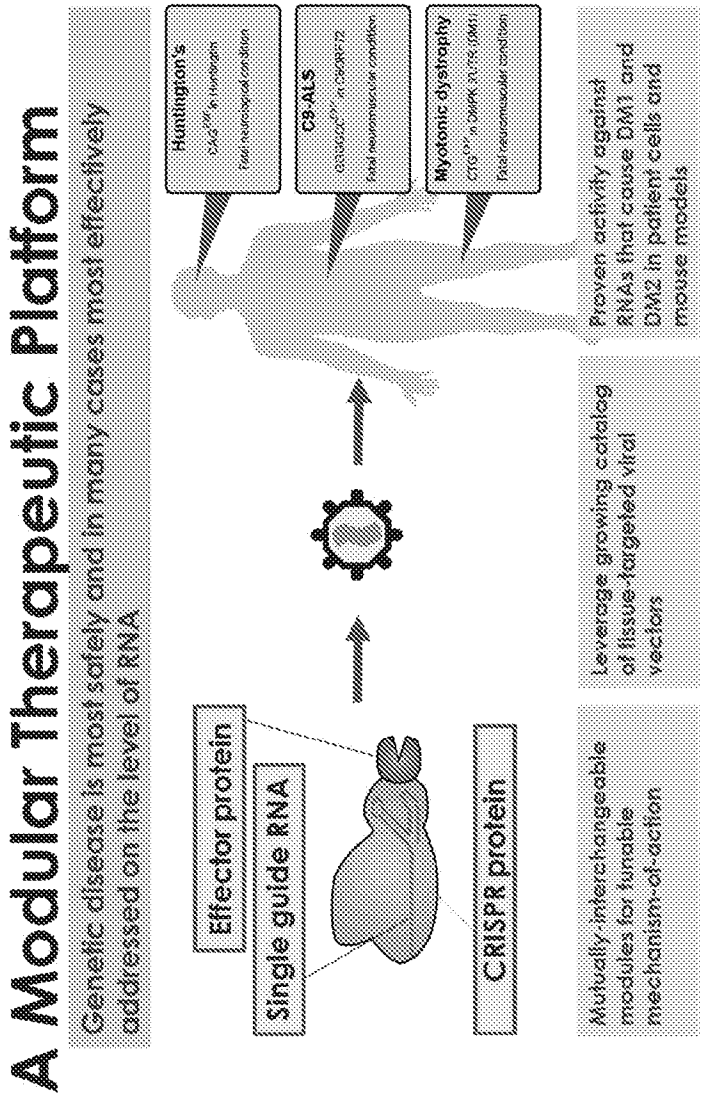
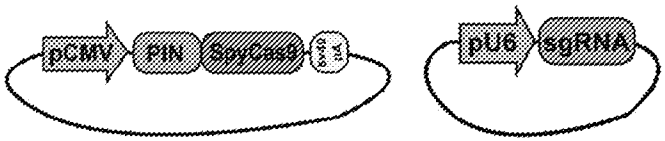

A Modular Therapeutic Platform
Genetic disease is most safely and in many cases most effectively addressed on the level of RNA
Effector protein
Single guide RNA
CRISPR protein
Huntington's
Fatal neurodegeneration
C9-ALS
Fatal neuromuscular condition
Myotonic dystrophy
Fatal neuromuscular condition
Proven activity against RNAs that cause DM1 and DM2 in patient cells and mouse models
Leverage growing catalog of tissue-targeted viral vectors
Mutually-interchangeable modules for tunable mechanism-of-action

FIGURE 3A

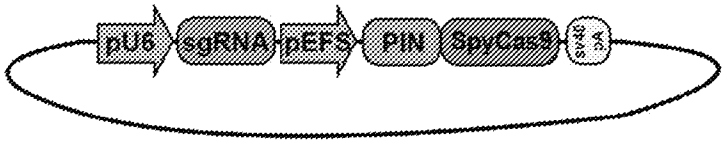
High expression positive control: "pos control"



Two plasmid system with PIN-SpyCas9 driven by a highly-active CMV promoter. This promoter is 5-20x as active as EFS.

FIGURE 3B

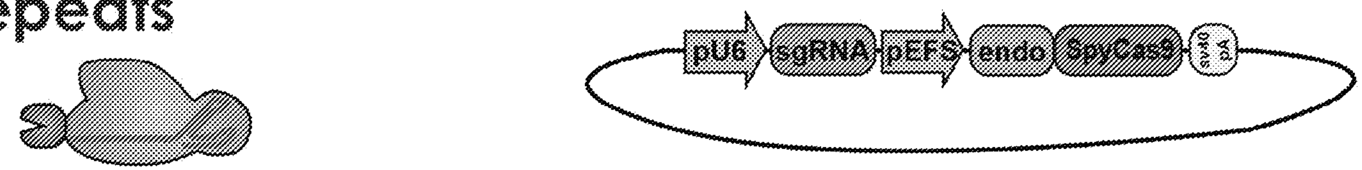
Low expression positive control: "P13"



Single plasmid system with lower expression of the fusion. Same architecture as the fusions involving new endonucleases

FIGURE 4A

CjeCas9 with N-terminal endonuclease: activity against CTG repeats



Endonuclease-C. jejuni Cas9 (nuclease inactive)

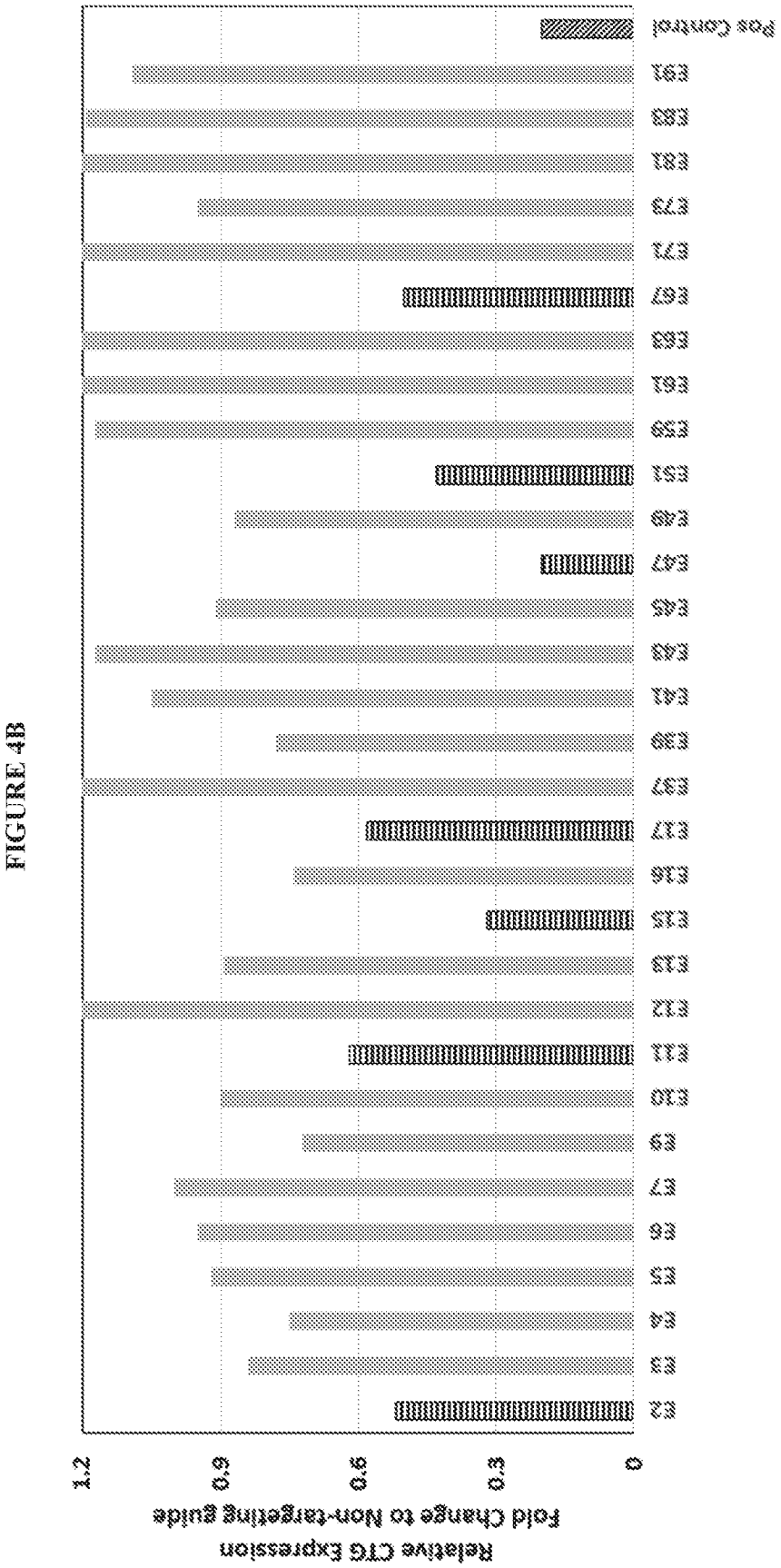
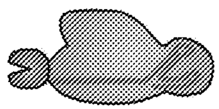

Relative CTG Expression
Fold Change to Non-targeting guide
1.2
0.9
0.6
0.3
0
E2
E3
E4
E5
E6
E7
E9
E10
E11
E12
E13
E15
E16
E17
E37
E39
E41
E43
E45
E47
E49
E51
E59
E61
E63
E67
E71
E73
E81
E83
E91
Pos Control

FIGURE 4B

FIGURE 5A

SpyCas9 with N-terminal endonuclease: activity against luciferase mRNA



Endonuclease-C. jejuni Cas9 (nuclease inactive)

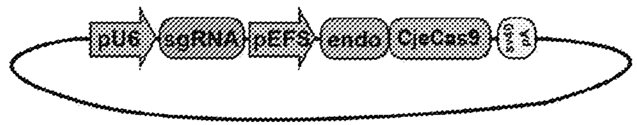


FIGURE 5B

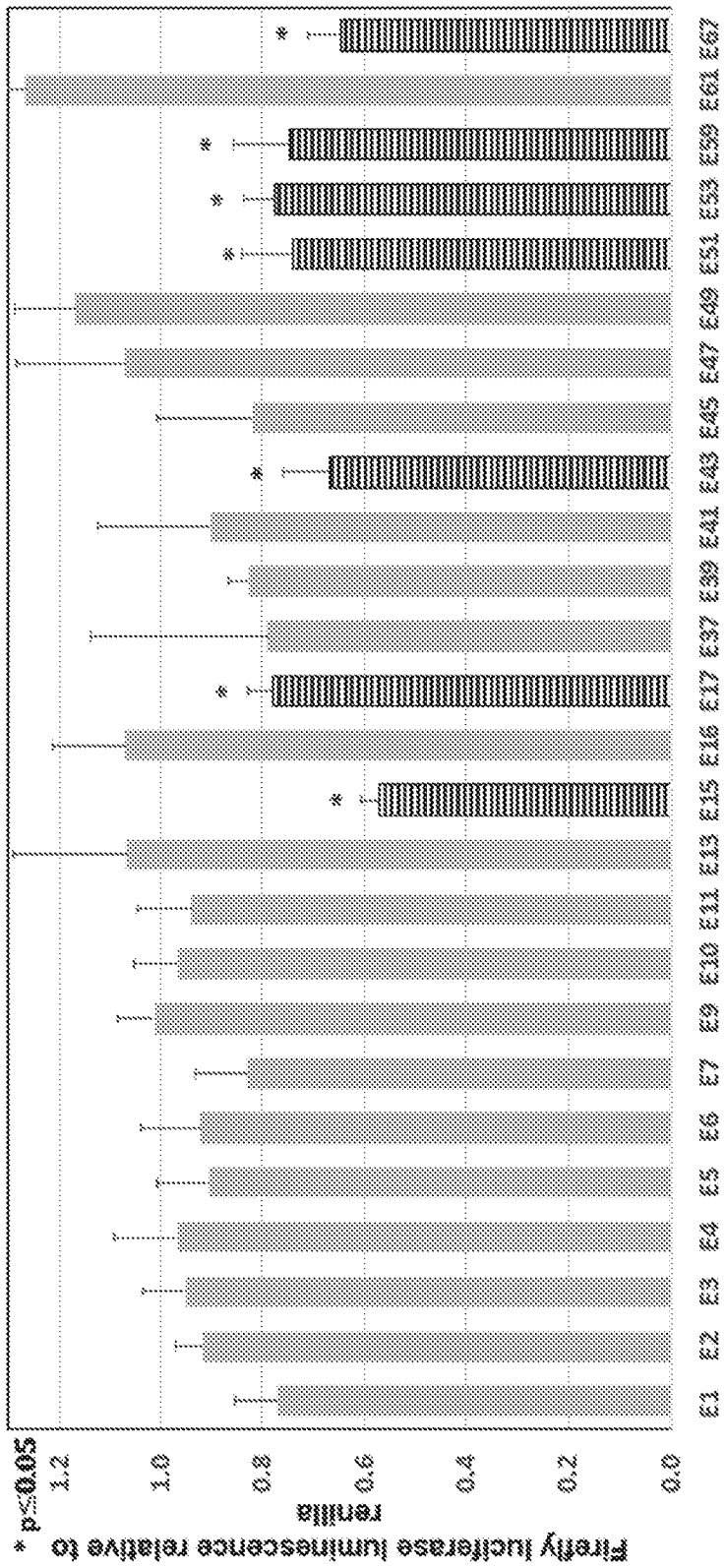

Firefly luciferase luminescence relative to renilla
* p≤0.05
0.0
0.2
0.4
0.6
0.8
1.0
1.2
E1
E2
E3
E4
E5
E6
E7
E9
E10
E11
E13
E15
E16
E17
E37
E39
E41
E43
E45
E47
E49
E51
E53
E59
E61
E67

FIGURE 6

Endonuclease Name Key

E1	RNAse1	E53	EXOG
E2	RNAse4	E55	ZC3H12D
E3	RNAse6	E57	ERN2
E4	NOB1	E59	ENDOG
E5	EndoV	E61	PELO
E6	FEN1	E63	YBEY
E7	SLFN14	E65	ENDOD1
E8	LACTB2	E67	CPSF4L
E9	RNAse7	E69	hCG_2002731
E10	RNAse8	E71	hCG_2002731
E11	RNAse2	E73	ERCC1
E12	ANG	E75	RAC1
E13	HRSP12	E77	APEX2
E14	RNAse6PL	E79	APEX2_1-350
E15	RNAseL	E81	RAA1_25-156
E16	RNAseT2	E83	RAB1
E17	ZC3H12A	E85	RNAseK
E37	RIDA	E87	DNA2_FL
E39	RNAse11	E89	RNASE1(K41R)
E41	PDL6	E91	RNASE1(K41R, D121E)
E43	NTHL1	E93	RNASE1(K41R, D121E, H119N)
E45	KIAA0391	E95	RNASE1(H119N)
E47	APEX1	E97	RNASE1(R39D, N67D, N88A, G89D, R91D, H119N)
E49	Ago2	E99	RNASE1(R39D, N67D, N88A, G89D, R91D, H119N, K41R, D121E)
E51	ZC3H12A	E100	RNAse1 (R39D, N67D, N88A, G89D, R91D, H119N, K41R, D121E)
		E101	RNASE1(R39D, N67D, N88A, G89D, R91D)

FIGURE 7A

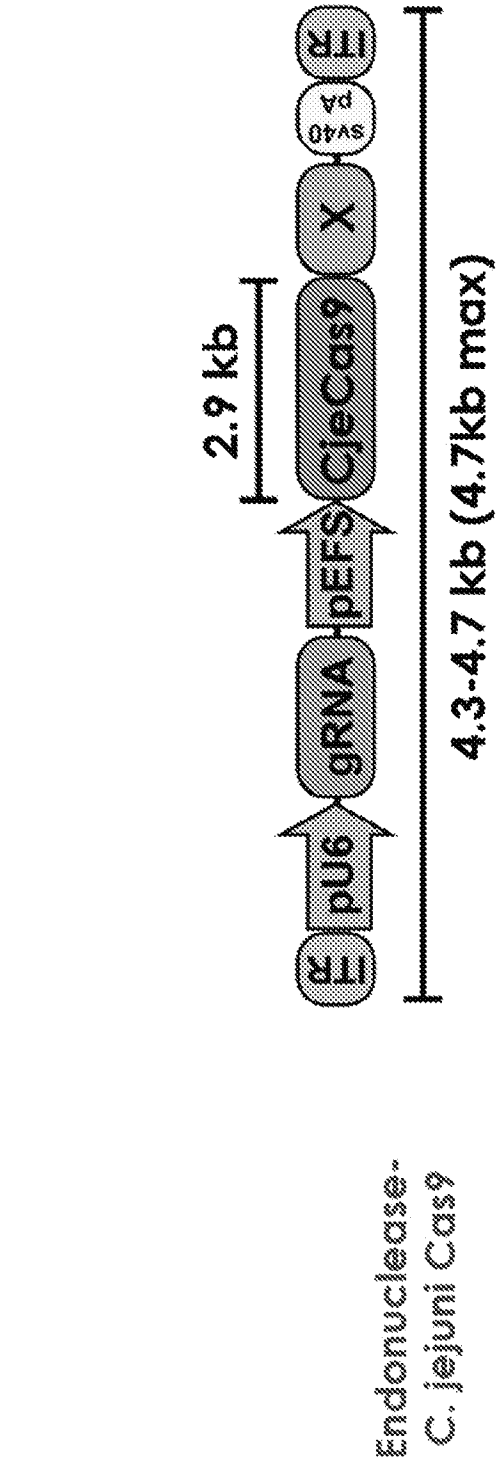

ITR
pU6
gRNA
pEFS
CjeCas9
X
SV40 pA
ITR
2.9 kb
4.3-4.7 kb (4.7kb max)
Endonuclease-
C. jejuni Cas9

FIGURE 7B

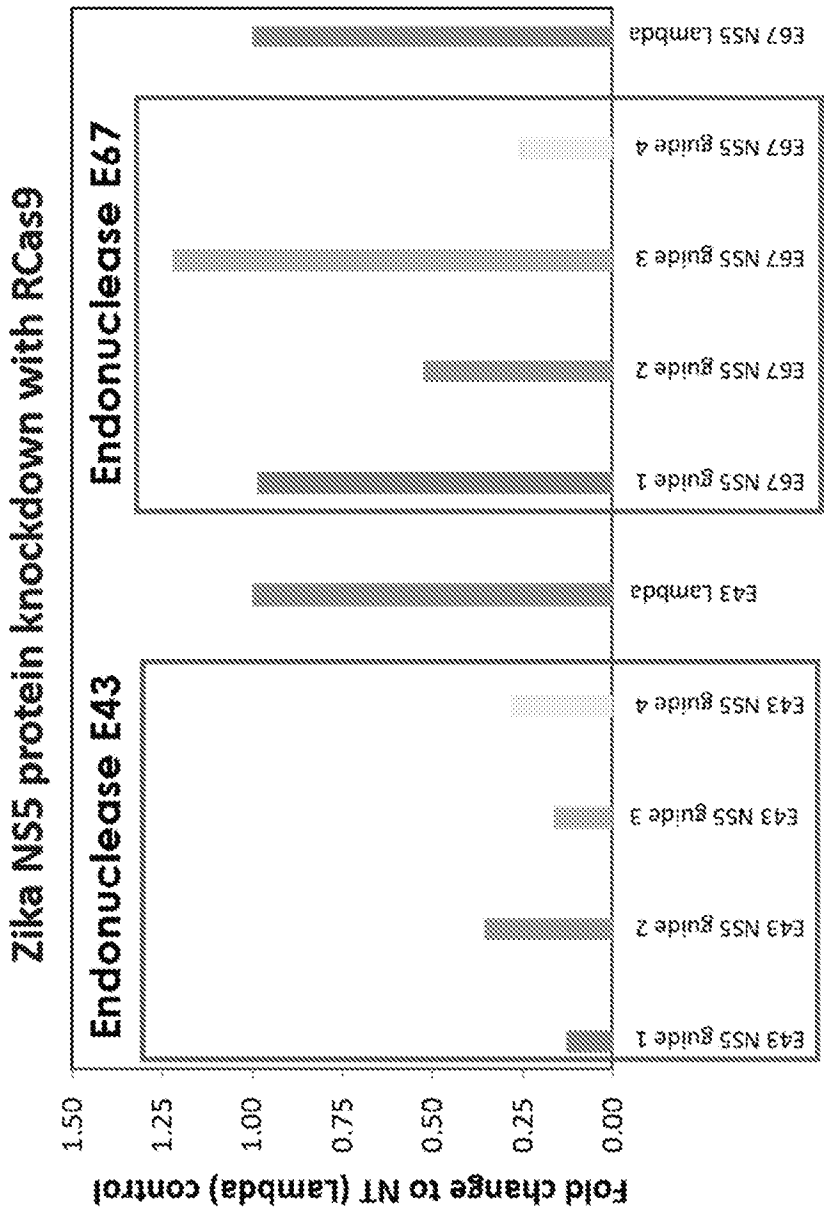
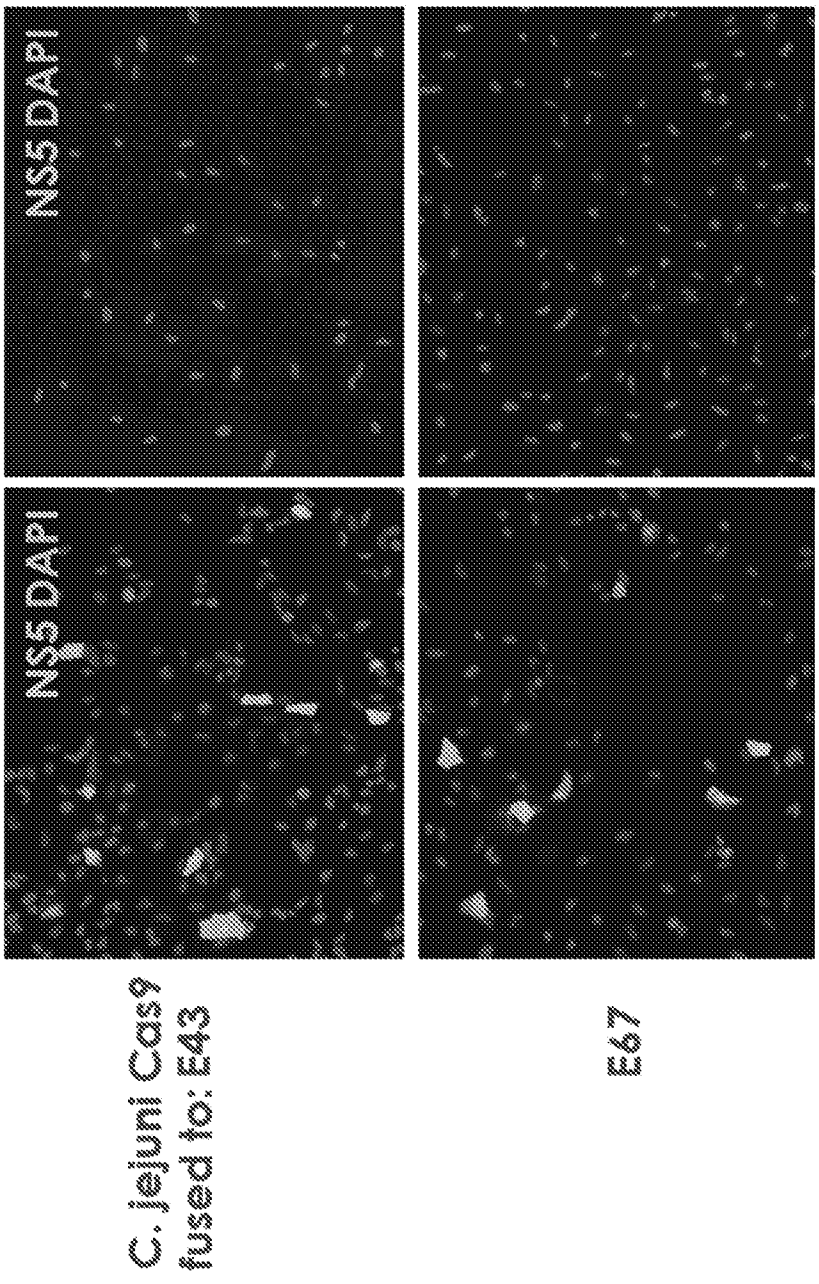

Zika NS5 protein knockdown with RCas9
Fold change to NT (lambda) control
1.50
1.25
1.00
0.75
0.50
0.25
0.00
Endonuclease E43
E43 NS5 guide 1
E43 NS5 guide 2
E43 NS5 guide 3
E43 NS5 guide 4
E43 lambda
Endonuclease E67
E67 NS5 guide 1
E67 NS5 guide 2
E67 NS5 guide 3
E67 NS5 guide 4
E67 NS5 lambda

FIGURE 8A


sgRNA targeting: non-targeting
Zika NS5
NS5 DAPI
NS5 DAPI
C. jejuni Cas9 fused to: E43
E67

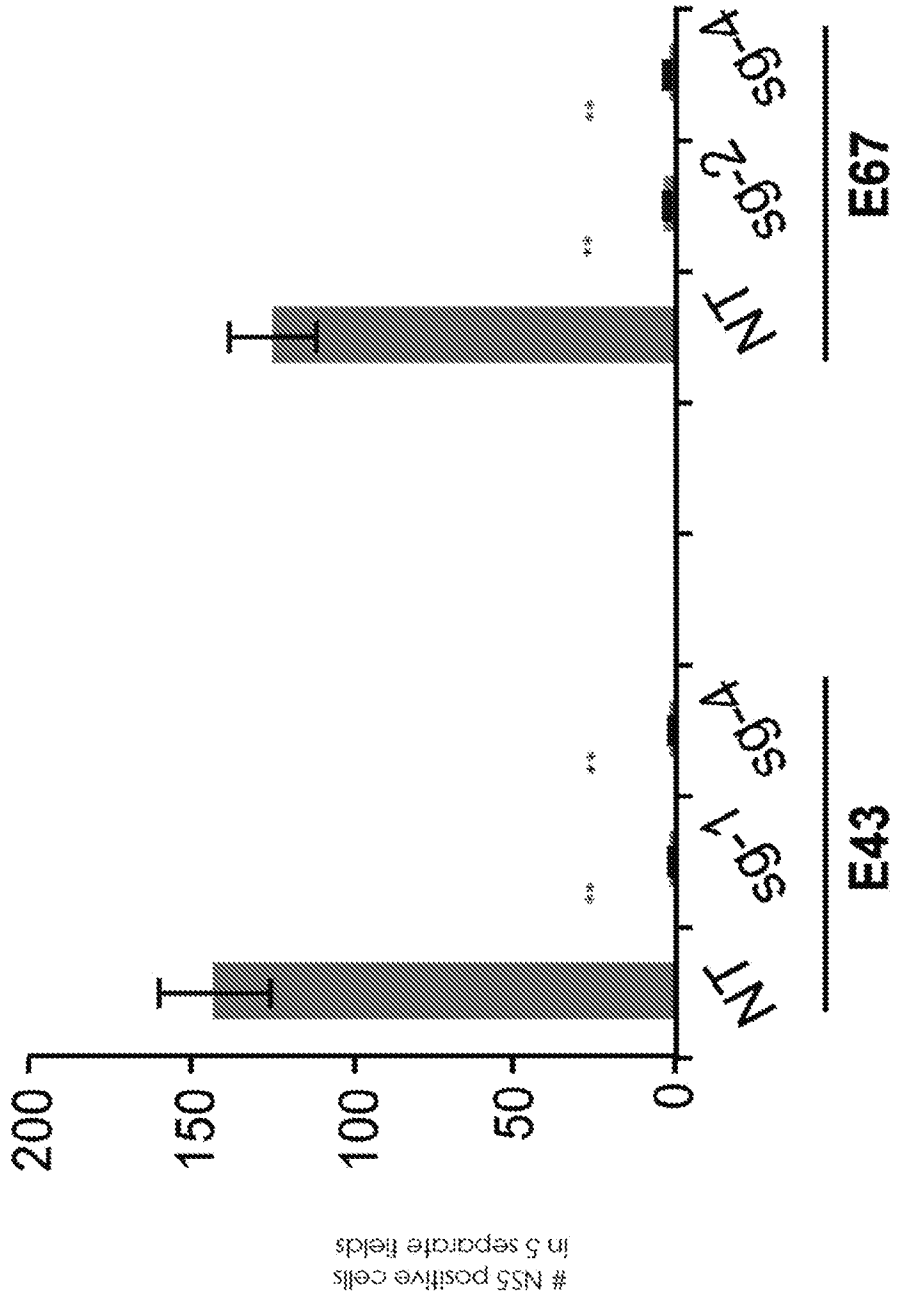
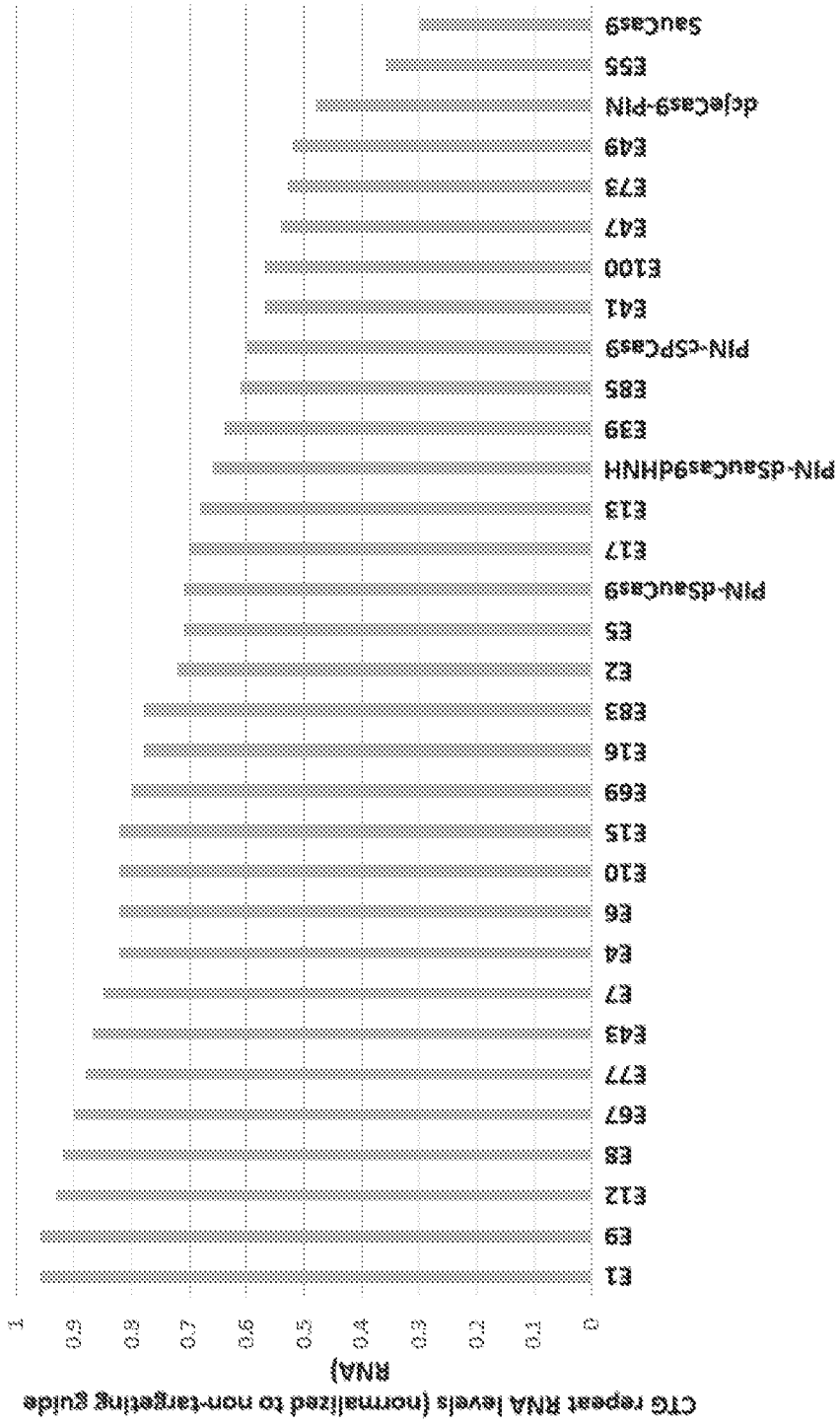

N55 positive cells
in 5 separate fields
200
150
100
50
0
NT
sg-1
sg-4
NT
sg-2
sg-4
E43
E67
**
**
**
**

FIGURE 8B

FIGURE 9


CTG repeat RNA levels (normalized to non-targeting guide RNA)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
E1
E9
E12
E8
E67
E77
E43
E7
E4
E6
E10
E15
E69
E16
E83
E2
E5
PIN-dSauCas9
E17
E13
PIN-dSauCas9dHNH
E39
E85
PIN-cSPCas9
E41
E100
E47
E73
E49
dCjeCas9-PIN
E55
SauCas9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2019/036021

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C12N 9/22; C12N 15/113 (2019.01)

CPC - C07K 2319/85; C12N 9/22; C12N 15/113; C12N 2310/20 (2019.08)

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

USPC - 435/199; 435/69.7 (keyword delimited)

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.
X -- Y	US 2017/0145394 A1 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA) 25 May 2017 (25.05.2017) entire document	1-12, 14-17, 19, 20 ---------- 13, 18
Y	US 2017/0088845 A1 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA et al) 30 March 2017 (30.03.2017) entire document	13
Y	US 2013/0178513 A1 (ISIS PHARMACEUTICALS, INC. et al) 11 July 2013 (11.07.2013) entire document	18
A	US 2017/0314002 A1 (BIO-RAD LABORATORIES, INC.) 02 November 2017 (02.11.2017) entire document	1-20
A	WO 2015/089486 A2 (THE BROAD INSTITUTE INC. et al) 18 June 2015 (18.06.2015) entire document	1-20

☐ 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

"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	Date of mailing of the international search report
04 September 2019	16 OCT 2019

Name and mailing address of the ISA/US	Authorized officer
Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450 Facsimile No. 571-273-8300	Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774