



(12) **DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION**

(13) **A1**

(86) Date de dépôt PCT/PCT Filing Date: 2016/12/07
(87) Date publication PCT/PCT Publication Date: 2017/06/15
(85) Entrée phase nationale/National Entry: 2018/05/29
(86) N° demande PCT/PCT Application No.: US 2016/065420
(87) N° publication PCT/PCT Publication No.: 2017/100343
(30) Priorités/Priorities: 2015/12/07 (US62/264,262);
2016/02/23 (US62/298,963)

(51) Cl.Int./Int.Cl. *A61K 48/00* (2006.01),
C12N 15/113 (2010.01), *C12N 15/115* (2010.01),
C12N 15/32 (2006.01), *C12N 15/62* (2006.01),
C12N 15/63 (2006.01), *C12N 15/79* (2006.01),
C12N 9/22 (2006.01)
(71) Demandeur/Applicant:
ARC BIO, LLC, US
(72) Inventeur/Inventor:
GOURGUECHON, STEPHANE B., US
(74) Agent: DEETH WILLIAMS WALL LLP

(54) Titre : PROCÉDES ET COMPOSITIONS POUR LA FABRICATION ET L'UTILISATION D'ACIDES NUCLEIQUES DE GUIDAGE
(54) Title: METHODS AND COMPOSITIONS FOR THE MAKING AND USING OF GUIDE NUCLEIC ACIDS

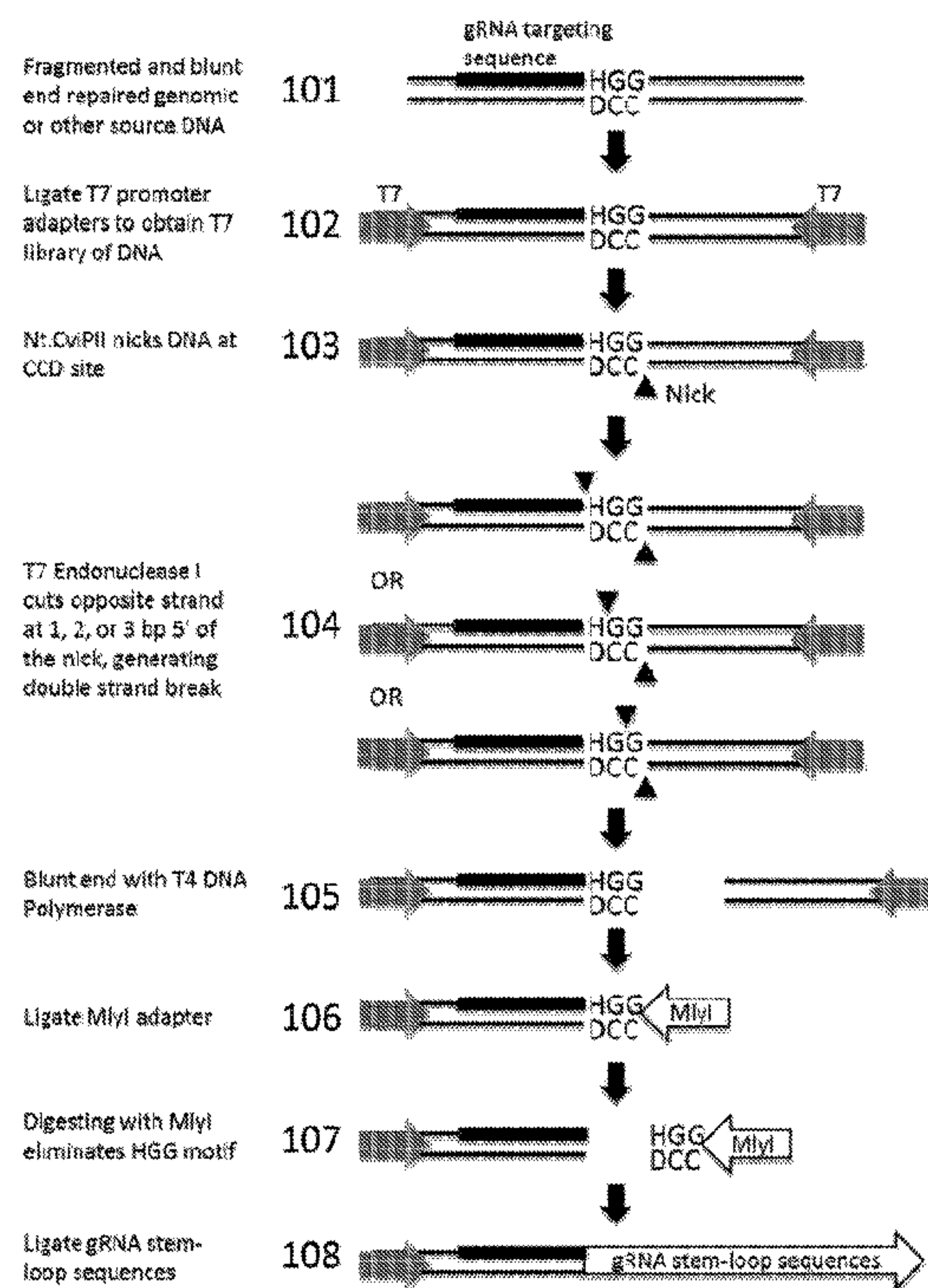


FIG. 1

(57) **Abrégé/Abstract:**

Provided herein are methods and compositions to make guide nucleic acids (gNAs), nucleic acids encoding gNAs, collections of gNAs, and nucleic acids encoding for a collection of gNAs from any source nucleic acid. Also provided herein are methods and compositions to use the resulting gNAs, nucleic acids encoding gNAs, collections of gNAs, and nucleic acids encoding for a collection of gNAs in a variety of applications.

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

(19) World Intellectual Property

Organization

International Bureau

(43) International Publication Date

15 June 2017 (15.06.2017)



(10) International Publication Number

WO 2017/100343 A1

(51) International Patent Classification:

A61K 48/00 (2006.01) C12N 15/63 (2006.01)
 C12N 9/22 (2006.01) C12N 15/79 (2006.01)
 C12N 15/32 (2006.01) C12N 15/113 (2010.01)
 C12N 15/62 (2006.01) C12N 15/115 (2010.01)

(21) International Application Number:

PCT/US2016/065420

(22) International Filing Date:

7 December 2016 (07.12.2016)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/264,262 7 December 2015 (07.12.2015) US
 62/298,963 23 February 2016 (23.02.2016) US

(71) Applicant: ARC BIO, LLC [US/US]; 16 Cherry Street,
 Cambridge, MA 02139 (US).

(72) Inventor: GOURGUECHON, Stephane, B.; c/o ARC
 BIO, LLC, 16 Cherry Street, Cambridge, MA 02139 (US).

(74) Agents: MADHURI, Roy et al.; Cooley LLP, 1299
 Pennsylvania Avenue NW, Suite 700, Washington, DC
 20004 (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, JP, KE, KG, 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, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
 GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))

[Continued on next page]

(54) Title: METHODS AND COMPOSITIONS FOR THE MAKING AND USING OF GUIDE NUCLEIC ACIDS

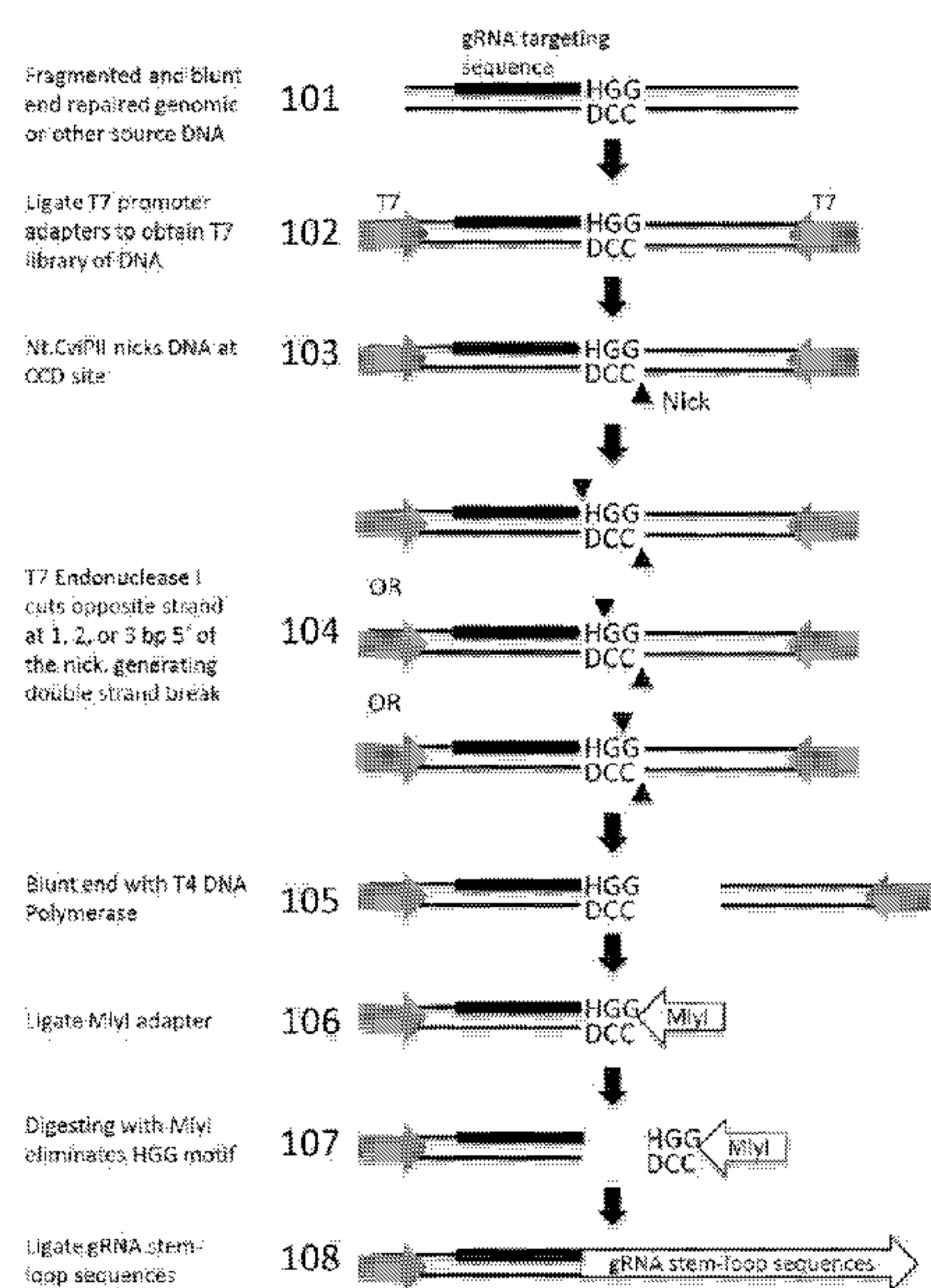


FIG. 1

(57) Abstract: Provided herein are methods and compositions to make guide nucleic acids (gNAs), nucleic acids encoding gNAs, collections of gNAs, and nucleic acids encoding for a collection of gNAs from any source nucleic acid. Also provided herein are methods and compositions to use the resulting gNAs, nucleic acids encoding gNAs, collections of gNAs, and nucleic acids encoding for a collection of gNAs in a variety of applications.

WO 2017/100343 A1

WO 2017/100343 A1



-
- *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))*

METHODS AND COMPOSITIONS FOR THE MAKING AND USING OF GUIDE NUCLEIC ACIDS

CROSS-REFERENCE

[0001] This application claims the benefit of U.S. Provisional Application No. 62/264,262, filed December 7, 2015, and of U.S. Provisional Application No. 62/298,963, filed February 23, 2016, each of which is hereby incorporated by reference in its entirety.

INCORPORATION BY REFERENCE OF SEQUENCE LISTING

[0002] The present application is being filed with a Sequence Listing in electronic format. The Sequence Listing is provided as a file entitled 750592000340SeqList.txt, created December 5, 2016, which is 816,451 bytes in size. The information in electronic format of the Sequence Listing is incorporated by reference in its entirety.

BACKGROUND

[0003] Human clinical DNA samples and sample libraries such as cDNA libraries derived from RNA contain highly abundant sequences that have little informative value and increase the cost of sequencing. While methods have been developed to deplete these unwanted sequences (e.g., via hybridization capture), these methods are often time-consuming and can be inefficient.

[0004] Although a guide nucleic acid (gNA) mediated nuclease systems (such as guide RNA (gRNA)-mediated Cas systems) can efficiently deplete any target DNA, targeted depletion of very high numbers of unique DNA molecules is not feasible. For example, a sequencing library derived from human blood may contain >99% human genomic DNA. Using a gRNA-mediated Cas9 system-based method to deplete this genomic DNA to detect an infectious agent circulating in the human blood would require extremely high numbers of gRNAs (about 10-100 million gRNAs), in order to ensure that a gRNA will be present every 30-50 base pairs (bp), and that no target DNA will be missed. Very large numbers of gRNAs can be predicted computationally and then synthesized chemically, but at a prohibitively expensive cost.

[0005] Therefore, there is a need in the art to provide a cost-effective method of converting any DNA into a gNA (e.g., gRNA) library to enable, for example, genome-wide depletion of unwanted DNA sequences from those of interest, without prior knowledge about their sequences. Provided herein are methods and compositions that address this need.

SUMMARY

[0006] Provided herein are compositions and methods to generate gNAs and collections of gNAs from any source nucleic acid. For example, gRNAs and collections of gRNAs can be generated from source

DNA, such as genomic DNA. Such gNAs and collections of the same are useful for a variety of applications, including depletion, partitioning, capture, or enrichment of target sequences of interest, genome-wide labeling, genome-wide editing, genome-wide functional screens, and genome-wide regulation.

[0007] In one aspect, the invention described herein provides a collection of nucleic acids, a plurality of the nucleic acids in the collection comprising: a first segment comprising a regulatory region; a second segment encoding a targeting sequence; and a third segment encoding a nucleic acid-guided nuclease system protein-binding sequence, wherein at least 10% of the nucleic acids in the collection vary in size. In another aspect, the invention described herein provides a collection of nucleic acids, a plurality of the nucleic acids in the collection comprising: a first segment comprising a regulatory region; a second segment encoding a targeting sequence, wherein the size of the second segment is greater than 21 bp; and a third segment encoding a nucleic acid-guided nuclease system protein-binding sequence. In some embodiments, the nucleic acid-guided nuclease system protein is a CRISPR/Cas system protein. In some embodiments, the size of the second segment varies from 15-250 bp across the collection of nucleic acids. In some embodiments, at least 10% of the second segments in the collection are greater than 21 bp. In some embodiments, the size of the second segment is not 20 bp. In some embodiments, the size of the second segment is not 21 bp. In some embodiments, the collection of nucleic acids is a collection of DNA. In some embodiments, the second segment is single stranded DNA. In some embodiments, the third segment is single stranded DNA. In some embodiments, the second segment is double stranded DNA. In some embodiments, the third segment is double stranded DNA. In some embodiments, the regulatory region is a region capable of binding a transcription factor. In some embodiments, the regulatory region comprises a promoter. In some embodiments, the promoter is selected from the group consisting of T7, SP6, and T3. In some embodiments, the targeting sequence is directed at a mammalian genome, eukaryotic genome, prokaryotic genome, or a viral genome. In some embodiments, the targeting sequence is directed at repetitive or abundant DNA. In some embodiments, the targeting sequence is directed at mitochondrial DNA, ribosomal DNA, Alu DNA, centromeric DNA, SINE DNA, LINE DNA, or STR DNA. In some embodiments, the sequence of the second segments is selected from Table 3 and/or Table 4. In some embodiments, the collection comprises at least 10^2 unique nucleic acid molecules. In some embodiments, the targeting sequence is at least 80% complementary to the strand opposite to a sequence of nucleotides 5' to a PAM sequence. In some embodiments, the collection comprises targeting sequences directed to sequences of interest spaced about every 10,000 bp or less across the genome of an organism. In some embodiments, the PAM sequence is AGG, CGG, or TGG. In some embodiments, the PAM sequence is specific for a CRISPR/Cas system protein selected from the group consisting of Cas9, Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, and Cm5. In some embodiments, the third segment comprises DNA encoding a gRNA stem-loop sequence. In some

embodiments, the third segment encodes for a RNA comprising the sequence
 GUUUUAGAGCUAGAAAUAGCAAGUUA AAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAG
 UGGCACCGAGUCGGUGCUUUUUUUU (SEQ ID NO: 1) or encodes for a RNA comprising the
 sequence

GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUA AAAUAAGGCUAGUCCGUUAUCA
 CUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUUUC (SEQ ID NO: 2). In some embodiments,
 the sequence of the third segment encodes for a crRNA and a tracrRNA. In some embodiments, the
 nucleic acid-guided nuclease system protein is from a bacterial species. In some embodiments, the
 nucleic acid-guided nuclease system protein is from an archaea species. In some embodiments, the
 CRISPR/Cas system protein is a Type I, Type II, or Type III protein. In some embodiments, the
 CRISPR/Cas system protein is selected from the group consisting of Cas9, Cpf1, Cas3, Cas8a-c, Cas10,
 Cse1, Csy1, Csn2, Cas4, Csm2, Cms5, dCas9 and cas9 nickase. In some embodiments, the third segment
 comprises DNA encoding a Cas9-binding sequence. In some embodiments, a plurality of third segments
 of the collection encode for a first nucleic acid-guided nuclease system protein binding sequence, and a
 plurality of the third segments of the collection encode for a second nucleic acid-guided nuclease system
 protein binding sequence. In some embodiments, the third segments of the collection encode for a
 plurality of different binding sequences of a plurality of different binding sequences of a plurality of
 different nucleic acid-guided nuclease system proteins.

[0008] In another aspect, the invention described herein provides for a collection of guide RNAs
 (gRNAs), comprising: a first RNA segment a targeting sequence; and a second RNA segment comprising
 a nucleic acid-guided nuclease system protein-binding sequence, wherein at least 10% of the gRNAs in
 the collection vary in size. In some embodiments, the nucleic acid-guided nuclease system protein is a
 CRISPR/Cas system protein. In some embodiments, the size of the first segment varies from 15-250 bp
 across the collection of gRNAs. In some embodiments, the at least 10% of the first segments in the
 collection are greater than 21 bp. In some embodiments, the size of the first segment is not 20 bp. In
 some embodiments, the size of the first segment is not 21 bp. In some embodiments, the targeting
 sequence is directed at a mammalian genome, eukaryotic genome, prokaryotic genome, or viral genome.
 In some embodiments, the targeting sequence is directed at repetitive or abundant DNA. In some
 embodiments, the targeting sequence is directed at mitochondrial DNA, ribosomal DNA, Alu DNA,
 centromeric DNA, SINE DNA, LINE DNA, or STR DNA. In some embodiments, the sequence of the
 first segments is RNA encoded by sequences selected from Table 3 and/or Table 4. In some
 embodiments, the collection comprises at least 10^2 unique gRNAs. In some embodiments, the gRNAs
 comprise cytosine, guanine, and adenine. In some embodiments, a subset of the gRNAs further comprises
 thymine. In some embodiments, a subset of the gRNAs further comprises uracil. In some embodiments,
 the first segment is at least 80% complementary to a target genomic sequence of interest. In some

embodiments, the targeting sequence is at least 80% complementary to the strand opposite to a sequence of nucleotides 5' to a PAM sequence. In some embodiments the PAM sequence is AGG, CGG, or TGG. In some embodiments, the PAM sequence is specific for a CRISPR/Cas system protein selected from the group consisting of Cas9, Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, and Cm5. In some embodiments, the second segment comprises a gRNA stem-loop sequence. In some embodiments, the third segment comprises DNA encoding a gRNA stem-loop sequence. In some embodiments, the third segment comprises the sequence

GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAG
UGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 1) or comprises the sequence

GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAA
CUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUUC (SEQ ID NO: 2). In some embodiments,

the second segment comprises a crRNA and a tracrRNA. In some embodiments, the nucleic acid-guided nuclease system protein is from a bacterial species. In some embodiments, the nucleic acid-guided nuclease system protein is from an archaea species. In some embodiments, the CRISPR/Cas system protein is a Type I, Type II, or Type III protein. In some embodiments, the CRISPR/Cas system protein is selected from the group consisting of Cas9, Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, Cm5, dCas9 and cas9 nickase. In some embodiments, the second segment comprises a Cas9-binding sequence. In some embodiments, at least 10% of the gRNAs in the collection vary in their 5' terminal-end sequence. In some embodiments, the collection comprises targeting sequences directed to sequences of interest spaced every 10,000 bp or less across the genome of an organism. In some embodiments, a plurality of second segments of the collection comprise a first nucleic acid-guided nuclease system protein binding sequence, and a plurality of the second segments of the collection comprise a second nucleic acid-guided nuclease system protein binding sequence. In some embodiments, the second segments of the collection comprise a plurality of different binding sequences of a plurality of different nucleic acid-guided nuclease system proteins. In some embodiments, a plurality of the gRNAs of the collection are attached to a substrate. In some embodiments, a plurality of the gRNAs of the collection comprise a label. In some particular embodiments, a plurality of the gRNAs of the collection comprise different labels.

[0009] In another aspect, the invention described herein provides nucleic acid comprising: a first segment comprising a regulatory region; a second segment encoding a targeting sequence, wherein the targeting sequence is greater than 30bp; and a third segment encoding a nucleic acid encoding a nucleic acid-guided nuclease system protein-binding sequence. In some embodiments, the nucleic acid-guided nuclease is a CRISPR/Cas system protein. In some embodiments, the nucleic acid is DNA. In some embodiments, the second segment is single stranded DNA. In some embodiments, the third segment is single stranded DNA. In some embodiments, the second segment is double stranded DNA. In some

embodiments, the third segment is double stranded DNA. In some embodiments, the regulatory region is a region capable of binding a transcription factor. In some embodiments, the regulatory region comprises a promoter. In some embodiments, the promoter is selected from the group consisting of T7, SP6, and T3. In some embodiments, the targeting sequence is directed at a mammalian genome, eukaryotic genome, prokaryotic genome, or a viral genome. In some embodiments, the targeting sequence is directed at abundant or repetitive DNA. In some embodiments, the targeting sequence is directed at mitochondrial DNA, ribosomal DNA, Alu DNA, centromeric DNA, SINE DNA, LINE DNA, or STR DNA. In some embodiments, the sequence of the second segments is selected from Table 3 and/or Table 4. In some embodiments, the targeting sequence is at least 80% complementary to the strand opposite to a sequence of nucleotides 5' to a PAM sequence. In some embodiments, the target genomic sequence of interest is 5' upstream of a PAM sequence. In some embodiments, the PAM sequence is specific for a CRISPR/Cas system protein selected from the group consisting of Cas9, Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, and Cm5. In some embodiments, the third segment comprises DNA encoding a gRNA stem-loop sequence. In some embodiments, the third segment comprises DNA encoding a gRNA stem-loop sequence. In some embodiments, the third segment encodes for a RNA comprising the sequence

GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAG
UGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 1) or encodes for a RNA comprising the
sequence

GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAA
CUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUUC (SEQ ID NO: 2). In some embodiments,
the nucleic acid-guided nuclease system protein is from a bacterial species. In some embodiments, the
nucleic acid-guided nuclease system protein is from an archaea species. In some embodiments, the
CRISPR/Cas system protein is a Type I, Type II, or Type III protein. In some embodiments, the
CRISPR/Cas system protein is selected from the group consisting of Cas9, Cpf1, Cas3, Cas8a-c, Cas10,
Cse1, Csy1, Csn2, Cas4, Csm2, Cm5, dCas9 and cas9 nickase. In some embodiments, the third segment
comprises DNA encoding a Cas9-binding sequence.

[0010] In another aspect, the invention described herein provides a guide RNA comprising a first
segment comprising a targeting sequence, wherein the size of the first segment is greater than 30 bp; and
a second segment comprising a nucleic acid-guided nuclease system protein-binding sequence. In some
embodiments, the nucleic acid-guided nuclease is a CRISPR/Cas system protein. In some embodiments,
the gRNA comprises an adenine, a guanine, and a cytosine. In some embodiments, the gRNA further
comprises a thymine. In some embodiments, the gRNA further comprises a uracil. In some embodiments,
the size of the first RNA segment is between 30 and 250 bp. In some embodiments, the targeting
sequence is directed at a mammalian genome, eukaryotic genome, prokaryotic genome, or viral genome.

In some embodiments, the targeting sequence is directed at repetitive or abundant DNA. In some embodiments, the targeting sequence is directed at mitochondrial DNA, ribosomal DNA, Alu DNA, centromeric DNA, SINE DNA, LINE DNA, or STR DNA. In some embodiments, the first segment is at least 80% complementary to the target genomic sequence of interest. In some embodiments, the targeting sequence is at least 80% complementary to the strand opposite to a sequence of nucleotides 5' to a PAM sequence. In some embodiments, the second segment comprises a gRNA stem-loop sequence. In some embodiments, the sequence of the second segment comprises
 GUUUUAGAGCUAGAAAUAGCAAGUUA AAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAG
 UGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 1) or comprises the sequence
 GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUA AAAUAAGGCUAGUCCGUUAUCAA
 CUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUUC (SEQ ID NO: 2). In some embodiments, the sequence of the third segment comprises a crRNA and a tracrRNA. In some embodiments, the nucleic acid-guided nuclease system protein is from a bacterial species. In some embodiments, the nucleic acid-guided nuclease system protein is from an archaea species. In some embodiments, the CRISPR/Cas system protein is a Type I, Type II, or Type III protein. In some embodiments, the CRISPR/Cas system protein is selected from the group consisting of Cas9, Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, Csm5, dCas9 and cas9 nickase. In some embodiments, the second segment is a Cas9-binding sequence.

[0011] In another aspect, the invention provides a complex comprising a nucleic acid-guided nuclease system protein and a comprising a first segment comprising a targeting sequence, wherein the size of the first segment is greater than 30 bp; and a second segment comprising a nucleic acid-guided nuclease system protein-binding sequence.

[0012] In another aspect, the invention described herein provides a method for depleting and partitioning of targeted sequences in a sample, enriching a sample for non-host nucleic acids, or serially depleting targeted nucleic acids in a sample comprising: providing nucleic acids extracted from a sample; and contacting the sample with a plurality of complexes comprising (i) any one of the collection of gRNAs provided herein; and (ii) nucleic acid-guided nuclease system proteins. In some embodiments, the nucleic acid-guided nuclease system proteins are CRISPR/Cas system proteins. In some embodiments, the CRISPR/Cas system proteins are Cas9 proteins.

[0013] In another aspect, the invention provides a method of making a collection of nucleic acids, each comprising a DNA encoding a targeting sequence ligated to a DNA encoding a nucleic acid-guided nuclease system protein-binding sequence, comprising: (a) providing double-stranded DNA molecules, each comprising a sequence of interest 5' to a PAM sequence, and its reverse complementary sequence on the opposite strand; (b) performing an enzymatic digestion reaction on the double stranded DNA molecules, wherein cleavages are generated at the PAM sequence and/or its reverse complementary

sequence on the opposite strand, but never completely remove the PAM sequence and/or its reverse complementary sequence on the opposite strand from the double stranded DNA; (c) ligating adapters comprising a recognition sequence to the resulting DNA molecules of step b; (d) contacting the DNA molecules of step c with a restriction enzyme that recognizes the recognition sequence of step c, whereby generating DNA fragments comprising blunt-ended double strand breaks immediately 5' to the PAM sequence, whereby removing the PAM sequence and the adapter containing the enzyme recognition site; and (e) ligating the resulting double stranded DNA fragments of step d with a DNA encoding a nucleic acid-guided nuclease system protein-binding sequence, whereby generating a plurality of DNA fragments, each comprising a DNA encoding a targeting sequence ligated to a DNA encoding a nucleic acid-guided nuclease system protein-binding sequence. In some embodiments, the nucleic acid-guided nuclease is a CRISPR/Cas nucleic acid-guided nuclease system protein. In some embodiments, the starting DNA molecules of the collection further comprise a regulatory sequence upstream of the sequence of interest 5' to the PAM sequence. In some embodiments, the regulatory sequence comprises a promoter. In some embodiments, the promoter comprises a T7, Sp6, or T3 sequence. In some embodiments, the double stranded DNA molecules are genomic DNA, intact DNA, or sheared DNA. In some embodiments, the genomic DNA is human, mouse, avian, fish, plant, insect, bacterial, or viral. In some embodiments, the DNA segments encoding a targeting sequence are at least 22 bp. In some embodiments, the DNA segments encoding a targeting sequence are 15-250 bp in size range. In some embodiments, the PAM sequence is AGG, CGG, or TGG. In some embodiments, the PAM sequence is specific for a CRISPR/Cas system protein selected from the group consisting of Cas9, Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, and Cm5. In some embodiments, step (b) further comprises (1) contacting the DNA molecules with an enzyme capable of creating a nick in a single strand at a CCD site, whereby generating a plurality of nicked double stranded DNA molecules, each comprising a sequence of interest followed by an HGG sequence, wherein the DNA molecules are nicked at the CCD sites; and (2) contacting the nicked double stranded DNA molecules with an endonuclease, whereby generating a plurality of double stranded DNA fragments, each comprising a sequence of interest followed by an HGG sequence wherein residual nucleotides from HGG and/or CCD sequences is (are) left behind. In some embodiments, step (d) further comprises PCR amplification of the adaptor-ligated DNA fragments from step (c) before cutting with the restriction enzyme recognizing the recognition sequence of step (c), wherein after PCR, the recognition sequence is positioned 3' of the PAM sequence, and a regulatory sequence is positioned at the 5' distal end of the PAM sequence. In some embodiments, the enzymatic reaction of step (b) comprises the use of a Nt.CviPII enzyme, and a T7 Endonuclease I enzyme. In some embodiments, step (c) further comprises a blunt-end reaction with a T4 DNA Polymerase, if the adapter to be ligated does not comprise an overhang. In some embodiments, the adapter of step (c) is either (1) double stranded, comprising a restriction enzyme recognition sequence in

one strand, and a regulatory sequence in the other strand, if the adapter is Y-shaped and comprises an overhang; or (2) has a palindromic enzyme recognition sequence in both strands, if the adapter is not Y-shaped. In some embodiments, the restriction enzyme of step (d) is MlyI. In some embodiments, the restriction enzyme of step (d) is BaeI. In some embodiments, step (d) further comprises contacting the DNA molecules with an XhoI enzyme. In some embodiments, in step (e) the DNA encoding a nucleic acid-guided nuclease system-protein binding sequence encodes for a RNA comprising the sequence GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAG UGGCACCGAGUCGGUGCUUUUUUUU (SEQ ID NO: 1) or encodes for a RNA comprising the sequence

GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUUC (SEQ ID NO: 2). In some embodiments, the targeted sequences of interest are spaced every 10,000 bp or less across the genome of an organism.

[0014] In another aspect, the invention provides a method of making a collection of nucleic acids, each comprising a DNA encoding a targeting sequence ligated to a DNA encoding a nucleic acid-guided nuclease system protein-binding sequence, comprising: (a) providing a plurality of double stranded DNA molecules, each comprising a sequence of interest, an NGG site, and its complement CCN site; (b) contacting the molecules with an enzyme capable of creating a nick in a single strand at a CCN site, whereby generating a plurality of nicked double stranded DNA molecules, each comprising a sequence of interest 5' to the NGG site, wherein the DNA molecules are nicked at the CCD sites; (c) contacting the nicked double stranded DNA molecules with an endonuclease, whereby generating a plurality of double stranded DNA fragments, each comprising a sequence of interest, wherein the fragments comprise an terminal overhang; (d) contacting the double stranded DNA fragments with an enzyme without 5' to 3' exonuclease activity to blunt end the double stranded DNA fragments, whereby generating a plurality of blunt ended double stranded fragments, each comprising a sequence of interest; (e) contacting the blunt ended double stranded fragments of step d with an enzyme that cleaves the terminal NGG site; and (f) ligating the resulting double stranded DNA fragments of step e with a DNA encoding a nucleic acid-guided nuclease system-protein binding sequence, whereby generating a plurality of DNA fragments, each comprising a targeting sequence ligated to a DNA encoding a nucleic acid-guided nuclease system protein-binding sequence. In some embodiments, the nucleic acid-guided nuclease is a CRISPR/Cas system protein. In some embodiments, the plurality of double stranded DNA molecules have a regulatory sequence 5' upstream of the NGG sites. In some embodiments, the regulatory sequence comprises a T7, SP6, or T3 sequence. In some embodiments, the NGG site comprises AGG, CGG, or TGG, and the CCN site comprises CCT, CCG, or CCA. In some embodiments, the plurality of double stranded DNA molecules, each comprising a sequence of interest comprise sheared fragments of genomic DNA. In some embodiments, the genomic DNA is mammalian, prokaryotic, eukaryotic, avian, bacterial or viral.

In some embodiments, the plurality of double stranded DNA molecules in step (a) are at least 500 bp. In some embodiments, the enzyme in step b is a Nt.CviPII enzyme. In some embodiments, the enzyme in step c is a T7 Endonuclease I. In some embodiments, the enzyme in step d is a T4 DNA Polymerase. In some embodiments, in step f the DNA encoding a nucleic acid-guided nuclease system-protein binding sequence encodes for a RNA comprising the sequence

GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAG
UGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 1) or encodes for a RNA comprising the
sequence

GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAA
CUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUUC (SEQ ID NO: 2). In some embodiments,
the step e additionally comprises ligating adaptors carrying a MlyI recognition site and digesting with
MlyI enzyme. In some embodiments, the sequence of interest is spaced every 10,000 bp or less across the
genome.

[0015] In another aspect, the invention provides a method of making a collection of nucleic acids, each comprising a DNA encoding a targeting sequence and a DNA encoding a nucleic acid-guided nuclease system protein-binding sequence, comprising: (a) providing genomic DNA comprising a plurality of sequences of interest, comprising NGG and CCN sites; (b) contacting the genomic DNA with an enzyme capable of creating nicks in the genomic DNA, whereby generating nicked genomic DNA, nicked at CCN sites; (c) contacting the nicked genomic DNA with an endonuclease, whereby generating double stranded DNA fragments, with an overhang; (d) ligating the DNA with overhangs from step c to a Y-shaped adapter, thereby introducing a restriction enzyme recognition sequence only at 3' of the NGG site and a regulatory sequence 5' of the sequence of interest; (e) contacting the product from step d with an enzyme that cleaves away the NGG site together with the adaptor carrying the enzyme recognition sequence; and (f) ligating the resulting double stranded DNA fragments of step e with a DNA encoding a nucleic acid-guided nuclease system protein-binding sequence, whereby generating a plurality of DNA fragments, each comprising a sequence of interest ligated to a DNA encoding a nucleic acid-guided nuclease system protein-binding sequence. In some embodiments, the nucleic acid-guided nuclease is a CRISPR/Cas system protein. In some embodiments, the NGG site comprises AGG, CGG, or TGG, and CCN site comprises CCT, CCG, or CCA. In some embodiments, the regulatory sequence comprises a promoter sequence. In some embodiments, the promoter sequence comprises a T7, SP6, or T3 sequence. In some embodiments, the DNA fragments are sheared fragments of genomic DNA.

[0016] In some embodiments, the genomic DNA is mammalian, prokaryotic, eukaryotic, or viral. In some embodiments, the fragments are at least 200 bp. In some embodiments, the enzyme in step b is a Nt.CviPII enzyme. In some embodiments, the enzyme in step c is a T7 Endonuclease I. In some embodiments, step d further comprises PCR amplification of the adaptor-ligated DNA. In some

embodiments, in step f, the DNA encoding nucleic acid-guided nuclease system protein-binding sequence encodes for a RNA comprising the sequence
 GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAG
 UGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 1) or encodes for a RNA comprising the
 sequence
 GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAA
 CUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUUC (SEQ ID NO: 2). In some embodiments,
 the enzyme removing NGG site in step e is MlyI. In some embodiments, the target of interest of the
 collection is spaced every 10,000 bp or less across the genome.

[0017] In another aspect, the invention provides kits and/or reagents useful for performing a method of making a collection of nucleic acids, each comprising a DNA encoding a targeting sequence ligated to a DNA encoding a nucleic acid-guided nuclease system protein-binding sequence, as described in the embodiments herein.

[0018] In another aspect, the invention described herein provides kit comprising a collection of nucleic acids, a plurality of the nucleic acids in the collection comprising: a first segment comprising a regulatory region; a second segment encoding a targeting sequence; and a third segment encoding a CRISPR/Cas system protein-binding sequence, wherein at least 10% of the nucleic acids in the collection vary in size.

[0019] In another aspect, the invention described herein provides a kit comprising a collection of nucleic acids, a plurality of the nucleic acids in the collection comprising: a first segment comprising a regulatory region; a second segment encoding a targeting sequence, wherein the size of the second segment is greater than 21 bp; and a third segment encoding a CRISPR/Cas system protein-binding sequence.

[0020] In another aspect, the invention described herein provides a kit comprising a collection of guide RNAs comprising a first RNA segment a targeting sequence; and a second RNA segment comprising a CRISPR/Cas system protein-binding sequence, wherein at least 10% of the gRNAs in the collection vary in size.

[0021] In another aspect, the invention described herein provides a method of making a collection of guide nucleic acids, comprising: a. obtaining abundant cells in a source sample; b. collecting nucleic acids from said abundant cells; and c. preparing a collection of guide nucleic acids (gNAs) from said nucleic acids. In some embodiments, said abundant cells comprise cells from one or more most abundant bacterial species in said source sample. In some embodiments, said abundant cells comprise cells from more than one species. In some embodiments, said abundant cells comprise human cells. In some embodiments, said abundant cells comprise animal cells. In some embodiments, said abundant cells comprise plant cells. In some embodiments, said abundant cells comprise bacterial cells. In some

embodiments, the method further comprises contacting nucleic acid-guided nucleases with said library of gNAs to form nucleic acid-guided nuclease-gNA complexes. In some embodiments, the method further comprises using said nucleic acid-guided nuclease-gNA complexes to cleave target nucleic acids at target sites, wherein said gNAs are complementary to said target sites. In some embodiments, said target nucleic acids are from said source sample. In some embodiments, a species of said target nucleic acids is the same as a species of said source sample. In some embodiments, said species of said target nucleic acids and said species of said source sample is human. In some embodiments, said species of said target nucleic acids and said species of said source sample is animal. In some embodiments, said species of said target nucleic acids and said species of said source sample is plant.

[0022] In another aspect, the invention described herein provides a method of making a collection of nucleic acids, each comprising a targeting sequence, comprising: a. obtaining source DNA; b. nicking said source DNA with a nicking enzyme at nicking enzyme recognition sites, thereby producing double-stranded breaks at proximal nicks; and c. repairing overhangs of said double-stranded breaks, thereby producing a double-stranded fragment comprising (i) a targeting sequence and (ii) said nicking enzyme recognition site. In another aspect, the invention described herein provides a method of making a collection of nucleic acids, each comprising a targeting sequence, comprising: a. obtaining source DNA; b. nicking said source DNA with a nicking enzyme at nicking enzyme recognition sites, thereby producing a nick; and c. synthesizing a new strand from said nick, thereby producing a single-stranded fragment of said source DNA comprising a targeting sequence. In some embodiments, the method further comprises producing a double-stranded fragment comprising said targeting sequence from said single-stranded fragment. In some embodiments, said producing said double-stranded fragment comprises random priming and extension. In some embodiments, said random priming is conducted with a primer comprising a random n-mer region and a promoter region. In some embodiments, said random n-mer region is a random hexamer region. In some embodiments, said random n-mer region is a random octamer region. In some embodiments, said promoter region is a T7 promoter region. In some embodiments, the method further comprises ligating a nuclease recognition site nucleic acid comprising a nuclease recognition site to said double-stranded fragment. In some embodiments, said nuclease recognition site corresponds to a nuclease that cuts at a distance from said nuclease recognition site equal to the length of said nicking enzyme recognition sites. In some embodiments, said nuclease recognition site is a MlyI recognition site. In some embodiments, said nuclease recognition site is a BaeI recognition site. In some embodiments, the method further comprises digesting said double-stranded fragment with said nuclease, thereby removing said nicking enzyme recognition site from said double-stranded fragment. In some embodiments, the method further comprises ligating said double-stranded fragment to a nucleic acid-guided nuclease system protein recognition site nucleic acid comprising a nucleic acid-guided nuclease system protein recognition site. In some embodiments, said nucleic acid-guided nuclease

system protein recognition site comprises a guide RNA stem-loop sequence. In some embodiments, said nuclease recognition site corresponds to a nuclease that cuts at a distance from said nuclease recognition site equal to a length of said targeting sequence. In some embodiments, said length of said targeting sequence is 20 base pairs. In some embodiments, said nuclease recognition site is a MmeI recognition site. In some embodiments, the method further comprises digesting said double-stranded fragment with said nuclease. In some embodiments, said nuclease recognition site corresponds to a nuclease that cuts at a distance from said nuclease recognition site equal to a length of said targeting sequence plus a length of said nicking enzyme recognition sites. In some embodiments, said length of said targeting sequence plus a length of said nicking enzyme recognition sites is 23 base pairs. In some embodiments, said nuclease recognition site is a EcoPI5I recognition site. In some embodiments, the method further comprises digesting said double-stranded fragment with said nuclease. In some embodiments, the method further comprises ligating said double-stranded fragment to a nucleic acid-guided nuclease system protein recognition site nucleic acid comprising a nucleic acid-guided nuclease system protein recognition site. In some embodiments, said nucleic acid-guided nuclease system protein recognition site comprises a guide RNA stem-loop sequence.

[0023] In another aspect, the invention described herein provides a kit comprising all essential reagents and instructions for carrying out the methods of aspects of the invention described herein.

BRIEF DESCRIPTION OF THE DRAWINGS

[0024] **FIG. 1** illustrates an exemplary scheme for producing a collection of gRNAs (a gRNA library) from genomic DNA.

[0025] **FIG. 2** illustrates another exemplary scheme for producing a collection of gRNAs (a gRNA library) from genomic DNA.

[0026] **FIG. 3** illustrates an exemplary scheme for nicking of DNA and subsequent treatment with polymerase to generate blunt ends.

[0027] **FIG. 4** illustrates an exemplary scheme for sequential production of a library of gNAs using three adapters.

[0028] **FIG. 5** illustrates an exemplary scheme for sequential production of a library of gNAs using one adapter and one oligo.

[0029] **FIG. 6** illustrates an exemplary scheme for generation of a large pool of DNA fragments with blunt ends using Nicking Enzyme Mediated DNA Amplification (NEMDA).

[0030] **FIG. 7** illustrates an exemplary scheme for generation of a large pool of gNAs using Nicking Enzyme Mediated DNA Amplification (NEMDA).

DETAILED DESCRIPTION OF THE INVENTION

[0031] There is a need in the art for a scalable, low-cost approach to generate large numbers of diverse guide nucleic acids (gNAs) (e.g., gRNAs, gDNAs) for a variety of downstream applications.

[0032] Unless defined otherwise herein, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are described.

[0033] Numeric ranges are inclusive of the numbers defining the range.

[0034] For purposes of interpreting this specification, the following definitions will apply and whenever appropriate, terms used in the singular will also include the plural and vice versa. In the event that any definition set forth below conflicts with any document incorporated herein by reference, the definition set forth shall control.

[0035] As used herein, the singular form “a”, “an”, and “the” includes plural references unless indicated otherwise.

[0036] It is understood that aspects and embodiments of the invention described herein include “comprising,” “consisting,” and “consisting essentially of” aspects and embodiments.

[0037] The term “about” as used herein refers to the usual error range for the respective value readily known to the skilled person in this technical field. Reference to “about” a value or parameter herein includes (and describes) embodiments that are directed to that value or parameter per se.

[0038] The term “nucleic acid,” as used herein, refers to a molecule comprising one or more nucleic acid subunits. A nucleic acid can include one or more subunits selected from adenosine (A), cytosine (C), guanine (G), thymine (T) and uracil (U), and modified versions of the same. A nucleic acid comprises deoxyribonucleic acid (DNA), ribonucleic acid (RNA), combinations, or derivatives thereof. A nucleic acid may be single-stranded and/or double-stranded.

[0039] The nucleic acids comprise “nucleotides”, which, as used herein, is intended to include those moieties that contain purine and pyrimidine bases, and modified versions of the same. Such modifications include methylated purines or pyrimidines, acylated purines or pyrimidines, alkylated riboses or other heterocycles. In addition, the term “nucleotide” or “polynucleotide” includes those moieties that contain hapten or fluorescent labels and may contain not only conventional ribose and deoxyribose sugars, but other sugars as well. Modified nucleosides, nucleotides or polynucleotides also include modifications on the sugar moiety, e.g., wherein one or more of the hydroxyl groups are replaced with halogen atoms or aliphatic groups, or are functionalized as ethers, amines, or the like.

[0040] The term “nucleic acids” and “polynucleotides” are used interchangeably herein. Polynucleotide is used to describe a nucleic acid polymer of any length, e.g., greater than about 2 bases, greater than about 10 bases, greater than about 100 bases, greater than about 500 bases, greater than 1000

bases, up to about 10,000 or more bases composed of nucleotides, e.g., deoxyribonucleotides or ribonucleotides, and may be produced enzymatically or synthetically (e.g., PNA as described in U.S. Patent No. 5,948,902 and the references cited therein) which can hybridize with naturally occurring nucleic acids in a sequence specific manner analogous to that of two naturally occurring nucleic acids, e.g., can participate in Watson-Crick base pairing interactions. Naturally-occurring nucleotides include guanine, cytosine, adenine and thymine (G, C, A and T, respectively). DNA and RNA have a deoxyribose and ribose sugar backbones, respectively, whereas PNA's backbone is composed of repeating N-(2-aminoethyl)-glycine units linked by peptide bonds. In PNA various purine and pyrimidine bases are linked to the backbone by methylene carbonyl bonds. A locked nucleic acid (LNA), often referred to as inaccessible RNA, is a modified RNA nucleotide. The ribose moiety of an LNA nucleotide is modified with an extra bridge connecting the 2' oxygen and 4' carbon. The bridge "locks" the ribose in the 3'-endo (North) conformation, which is often found in the A-form duplexes. LNA nucleotides can be mixed with DNA or RNA residues in the oligonucleotide whenever desired. The term "unstructured nucleic acid," or "UNA," is a nucleic acid containing non-natural nucleotides that bind to each other with reduced stability. For example, an unstructured nucleic acid may contain a G' residue and a C' residue, where these residues correspond to non-naturally occurring forms, i.e., analogs, of G and C that base pair with each other with reduced stability, but retain an ability to base pair with naturally occurring C and G residues, respectively. Unstructured nucleic acid is described in US20050233340, which is incorporated by reference herein for disclosure of UNA.

[0041] The term "oligonucleotide" as used herein denotes a single-stranded multimer of nucleotides.

[0042] Unless otherwise indicated, nucleic acids are written left to right in 5' to 3' orientation; amino acid sequences are written left to right in amino to carboxy orientation, respectively.

[0043] The term "cleaving," as used herein, refers to a reaction that breaks the phosphodiester bonds between two adjacent nucleotides in both strands of a double-stranded DNA molecule, thereby resulting in a double-stranded break in the DNA molecule.

[0044] The term "nicking" as used herein, refers to a reaction that breaks the phosphodiester bond between two adjacent nucleotides in only one strand of a double-stranded DNA molecule, thereby resulting in a break in one strand of the DNA molecule.

[0045] The term "cleavage site," as used herein, refers to the site at which a double-stranded DNA molecule has been cleaved.

[0046] The "nucleic acid-guided nuclease-gNA complex" refers to a complex comprising a nucleic acid-guided nuclease protein and a guide nucleic acid (gNA, for example a gRNA or a gDNA). For example the "Cas9-gRNA complex" refers to a complex comprising a Cas9 protein and a guide RNA (gRNA). The nucleic acid-guided nuclease may be any type of nucleic acid-guided nuclease, including

but not limited to wild type nucleic acid-guided nuclease, a catalytically dead nucleic acid-guided nuclease, or a nucleic acid-guided nuclease-nickase.

[0047] The term "nucleic acid-guided nuclease-associated guide NA" refers to a guide nucleic acid (guide NA). The nucleic acid-guided nuclease-associated guide NA may exist as an isolated nucleic acid, or as part of a nucleic acid-guided nuclease-gNA complex, for example a Cas9-gRNA complex.

[0048] The terms "capture" and "enrichment" are used interchangeably herein, and refer to the process of selectively isolating a nucleic acid region containing: sequences of interest, targeted sites of interest, sequences not of interest, or targeted sites not of interest.

[0049] The term "hybridization" refers to the process by which a strand of nucleic acid joins with a complementary strand through base pairing as known in the art. A nucleic acid is considered to be "selectively hybridizable" to a reference nucleic acid sequence if the two sequences specifically hybridize to one another under moderate to high stringency hybridization and wash conditions. Moderate and high stringency hybridization conditions are known (see, e.g., Ausubel, et al., Short Protocols in Molecular Biology, 3rd ed., Wiley & Sons 1995 and Sambrook et al., Molecular Cloning: A Laboratory Manual, Third Edition, 2001 Cold Spring Harbor, N.Y.). One example of high stringency conditions includes hybridization at about 42 °C in 50% formamide, 5X SSC, 5X Denhardt's solution, 0.5% SDS and 100 µg/ml denatured carrier DNA followed by washing two times in 2X SSC and 0.5% SDS at room temperature and two additional times in 0.1 X SSC and 0.5% SDS at 42 °C.

[0050] The term "duplex," or "duplexed," as used herein, describes two complementary polynucleotides that are base-paired, i.e., hybridized together.

[0051] The term "amplifying" as used herein refers to generating one or more copies of a target nucleic acid, using the target nucleic acid as a template.

[0052] The term "genomic region," as used herein, refers to a region of a genome, e.g., an animal or plant genome such as the genome of a human, monkey, rat, fish or insect or plant. In certain cases, an oligonucleotide used in the method described herein may be designed using a reference genomic region, i.e., a genomic region of known nucleotide sequence, e.g., a chromosomal region whose sequence is deposited at NCBI's Genbank database or other databases, for example.

[0053] The term "genomic sequence," as used herein, refers to a sequence that occurs in a genome. Because RNAs are transcribed from a genome, this term encompasses sequence that exist in the nuclear genome of an organism, as well as sequences that are present in a cDNA copy of an RNA (e.g., an mRNA) transcribed from such a genome.

[0054] The term "genomic fragment," as used herein, refers to a region of a genome, e.g., an animal or plant genome such as the genome of a human, monkey, rat, fish or insect or plant. A genomic fragment may be an entire chromosome, or a fragment of a chromosome. A genomic fragment may be adapter

ligated (in which case it has an adapter ligated to one or both ends of the fragment, or to at least the 5' end of a molecule), or may not be adapter ligated.

[0055] In certain cases, an oligonucleotide used in the method described herein may be designed using a reference genomic region, i.e., a genomic region of known nucleotide sequence, *e.g.*, a chromosomal region whose sequence is deposited at NCBI's Genbank database or other databases, for example. Such an oligonucleotide may be employed in an assay that uses a sample containing a test genome, where the test genome contains a binding site for the oligonucleotide.

[0056] The term "ligating," as used herein, refers to the enzymatically catalyzed joining of the terminal nucleotide at the 5' end of a first DNA molecule to the terminal nucleotide at the 3' end of a second DNA molecule.

[0057] If two nucleic acids are "complementary," each base of one of the nucleic acids base pairs with corresponding nucleotides in the other nucleic acid. The term "complementary" and "perfectly complementary" are used synonymously herein.

[0058] The term "separating," as used herein, refers to physical separation of two elements (*e.g.*, by size or affinity, etc.) as well as degradation of one element, leaving the other intact. For example, size exclusion can be employed to separate nucleic acids, including cleaved targeted sequences.

[0059] In a cell, DNA usually exists in a double-stranded form, and as such, has two complementary strands of nucleic acid referred to herein as the "top" and "bottom" strands. In certain cases, complementary strands of a chromosomal region may be referred to as "plus" and "minus" strands, the "first" and "second" strands, the "coding" and "noncoding" strands, the "Watson" and "Crick" strands or the "sense" and "antisense" strands. The assignment of a strand as being a top or bottom strand is arbitrary and does not imply any particular orientation, function or structure. Until they become covalently linked, the first and second strands are distinct molecules. For ease of description, the "top" and "bottom" strands of a double-stranded nucleic acid in which the top and bottom strands have been covalently linked will still be described as the "top" and "bottom" strands. In other words, for the purposes of this disclosure, the top and bottom strands of a double-stranded DNA do not need to be separated molecules. The nucleotide sequences of the first strand of several exemplary mammalian chromosomal regions (*e.g.*, BACs, assemblies, chromosomes, etc.) is known, and may be found in NCBI's Genbank database, for example.

[0060] The term "top strand," as used herein, refers to either strand of a nucleic acid but not both strands of a nucleic acid. When an oligonucleotide or a primer binds or anneals "only to a top strand," it binds to only one strand but not the other. The term "bottom strand," as used herein, refers to the strand that is complementary to the "top strand." When an oligonucleotide binds or anneals "only to one strand," it binds to only one strand, *e.g.*, the first or second strand, but not the other strand. If an oligonucleotide binds or anneals to both strands of a double-stranded DNA, the oligonucleotide may have

two regions, a first region that hybridizes with the top strand of the double-stranded DNA, and a second region that hybridizes with the bottom strand of the double-stranded DNA.

[0061] The term “double-stranded DNA molecule” refers to both double-stranded DNA molecules in which the top and bottom strands are not covalently linked, as well as double-stranded DNA molecules in which the top and bottom strands are covalently linked. The top and bottom strands of a double-stranded DNA are base paired with one other by Watson-Crick interactions.

[0062] The term “denaturing,” as used herein, refers to the separation of at least a portion of the base pairs of a nucleic acid duplex by placing the duplex in suitable denaturing conditions. Denaturing conditions are well known in the art. In one embodiment, in order to denature a nucleic acid duplex, the duplex may be exposed to a temperature that is above the T_m of the duplex, thereby releasing one strand of the duplex from the other. In certain embodiments, a nucleic acid may be denatured by exposing it to a temperature of at least 90°C for a suitable amount of time (e.g., at least 30 seconds, up to 30 mins). In certain embodiments, fully denaturing conditions may be used to completely separate the base pairs of the duplex. In other embodiments, partially denaturing conditions (e.g., with a lower temperature than fully denaturing conditions) may be used to separate the base pairs of certain parts of the duplex (e.g., regions enriched for A-T base pairs may separate while regions enriched for G-C base pairs may remain paired). Nucleic acid may also be denatured chemically (e.g., using urea or NaOH).

[0063] The term “genotyping,” as used herein, refers to any type of analysis of a nucleic acid sequence, and includes sequencing, polymorphism (SNP) analysis, and analysis to identify rearrangements.

[0064] The term “sequencing,” as used herein, refers to a method by which the identity of consecutive nucleotides of a polynucleotide are obtained.

[0065] The term “next-generation sequencing” refers to the so-called parallelized sequencing-by-synthesis or sequencing-by-ligation platforms, for example, those currently employed by Illumina, Life Technologies, and Roche, etc. Next-generation sequencing methods may also include nanopore sequencing methods or electronic-detection based methods such as Ion Torrent technology commercialized by Life Technologies.

[0066] The term “complementary DNA” or cDNA refers to a double-stranded DNA sample that was produced from an RNA sample by reverse transcription of RNA (using primers such as random hexamers or oligo-dT primers) followed by second-strand synthesis by digestion of the RNA with RNaseH and synthesis by DNA polymerase.

[0067] The term “RNA promoter adapter” is an adapter that contains a promoter for a bacteriophage RNA polymerase, e.g., the RNA polymerase from bacteriophage T3, T7, SP6 or the like.

[0068] Other definitions of terms may appear throughout the specification.

[0069] For any of the structural and functional characteristics described herein, methods of determining these characteristics are known in the art.

Guide Nucleic Acids (gNAs)

[0070] Provided herein are guide nucleic acids (gNAs) derivable from any nucleic acid source. The gNAs can be guide RNAs (gRNAs) or guide DNAs (gDNAs). The nucleic acid source can be DNA or RNA. Provided herein are methods to generate gNAs from any source nucleic acid, including DNA from a single organism, or mixtures of DNA from multiple organisms, or mixtures of DNA from multiple species, or DNA from clinical samples, or DNA from forensic samples, or DNA from environmental samples, or DNA from metagenomic DNA samples (for example a sample that contains more than one species of organism). Examples of any source DNA include, but are not limited to any genome, any genome fragment, cDNA, synthetic DNA, or a DNA collection (e.g. a SNP collection, DNA libraries). The gNAs provided herein can be used for genome-wide applications.

[0071] In some embodiments, the gNAs are derived from genomic sequences (e.g., genomic DNA). In some embodiments, the gNAs are derived from mammalian genomic sequences. In some embodiments, the gNAs are derived from eukaryotic genomic sequences. In some embodiments, the gNAs are derived from prokaryotic genomic sequences. In some embodiments, the gNAs are derived from viral genomic sequences. In some embodiments, the gNAs are derived from bacterial genomic sequences. In some embodiments, the gNAs are derived from plant genomic sequences. In some embodiments, the gNAs are derived from microbial genomic sequences. In some embodiments, the gNAs are derived from genomic sequences from a parasite, for example a eukaryotic parasite.

[0072] In some embodiments, the gNAs are derived from repetitive DNA. In some embodiments, the gNAs are derived from abundant DNA. In some embodiments, the gNAs are derived from mitochondrial DNA. In some embodiments, the gNAs are derived from ribosomal DNA. In some embodiments, the gNAs are derived from centromeric DNA. In some embodiments, the gNAs are derived from DNA comprising Alu elements (Alu DNA). In some embodiments, the gNAs are derived from DNA comprising long interspersed nuclear elements (LINE DNA). In some embodiments, the gNAs are derived from DNA comprising short interspersed nuclear elements (SINE DNA). In some embodiments the abundant DNA comprises ribosomal DNA. In some embodiments, the abundant DNA comprises host DNA (e.g., host genomic DNA or all host DNA). In an example, the gNAs can be derived from host DNA (e.g., human, animal, plant) for the depletion of host DNA to allow for easier analysis of other DNA that is present (e.g., bacterial, viral, or other metagenomic DNA). In another example, the gNAs can be derived from the one or more most abundant types (e.g., species) in a mixed sample, such as the one or more most abundant bacteria species in a metagenomic sample. The one or more most abundant types (e.g., species) can comprise the two, three, four, five, six, seven, eight, nine, ten, or more than ten

most abundant types (e.g., species). The most abundant types can be the most abundant kingdoms, phyla or divisions, classes, orders, families, genres, species, or other classifications. The most abundant types can be the most abundant cell types, such as epithelial cells, bone cells, muscle cells, blood cells, adipose cells, or other cell types. The most abundant types can be non-cancerous cells. The most abundant types can be cancerous cells. The most abundant types can be animal, human, plant, fungal, bacterial, or viral. gNAs can be derived from both a host and the one or more most abundant non-host types (e.g., species) in a sample, such as from both human DNA and the DNA of the one or more most abundant bacterial species. In some embodiments, the abundant DNA comprises DNA from the more abundant or most abundant cells in a sample. For example, for a specific sample, the highly abundant cells can be extracted and their DNA can be used to produce gNAs; these gNAs can be used to produce depletion library and applied to original sample to enable or enhance sequencing or detection of low abundance targets.

[0073] In some embodiments, the gNAs are derived from DNA comprising short terminal repeats (STRs).

[0074] In some embodiments, the gNAs are derived from a genomic fragment, comprising a region of the genome, or the whole genome itself. In one embodiment, the genome is a DNA genome. In another embodiment, the genome is a RNA genome.

[0075] In some embodiments, the gNAs are derived from a eukaryotic or prokaryotic organism; from a mammalian organism or a non-mammalian organism; from an animal or a plant; from a bacteria or virus; from an animal parasite; from a pathogen.

[0076] In some embodiments, the gNAs are derived from any mammalian organism. In one embodiment the mammal is a human. In another embodiment the mammal is a livestock animal, for example a horse, a sheep, a cow, a pig, or a donkey. In another embodiment, a mammalian organism is a domestic pet, for example a cat, a dog, a gerbil, a mouse, a rat. In another embodiment the mammal is a type of a monkey.

[0077] In some embodiments, the gNAs are derived from any bird or avian organism. An avian organism includes but is not limited to chicken, turkey, duck and goose.

[0078] In some embodiments, the gNAs are derived from a plant. In one embodiment, the plant is rice, maize, wheat, rose, grape, coffee, fruit, tomato, potato, or cotton.

[0079] In some embodiments, the gNAs are derived from a species of bacteria. In one embodiment, the bacteria are tuberculosis-causing bacteria.

[0080] In some embodiments, the gNAs are derived from a virus.

[0081] In some embodiments, the gNAs are derived from a species of fungi.

[0082] In some embodiments, the gNAs are derived from a species of algae.

[0083] In some embodiments, the gNAs are derived from any mammalian parasite.

[0084] In some embodiments, the gNAs are derived from any mammalian parasite. In one embodiment, the parasite is a worm. In another embodiment, the parasite is a malaria-causing parasite. In another embodiment, the parasite is a Leishmaniasis-causing parasite. In another embodiment, the parasite is an amoeba.

[0085] In some embodiments, the gNAs are derived from a nucleic acid target. Contemplated targets include, but are not limited to, pathogens; single nucleotide polymorphisms (SNPs), insertions, deletions, tandem repeats, or translocations; human SNPs or STRs; potential toxins; or animals, fungi, and plants. In some embodiments, the gRNAs are derived from pathogens, and are pathogen-specific gNAs.

[0086] In some embodiments, a guide NA of the invention comprises a first NA segment comprising a targeting sequence, wherein the targeting sequence is 15-250 bp; and a second NA segment comprising a nucleic acid guided nuclease system (e.g., CRISPR/Cas system) protein-binding sequence. In some embodiments, the targeting sequence is greater than 21 bp, greater than 22 bp, greater than 23 bp, greater than 24 bp, greater than 25 bp, greater than 26 bp, greater than 27 bp, greater than 28 bp, greater than 29 bp, greater than 30 bp, greater than 40 bp, greater than 50 bp, greater than 60 bp, greater than 70 bp, greater than 80 bp, greater than 90 bp, greater than 100 bp, greater than 110 bp, greater than 120 bp, greater than 130 bp, greater than 140 bp, or even greater than 150 bp. In an exemplary embodiment, the targeting sequence is greater than 30bp. In some embodiments, the targeting sequences of the present invention range in size from 30-50 bp. In some embodiments, targeting sequences of the present invention range in size from 30-75 bp. In some embodiments, targeting sequences of the present invention range in size from 30-100 bp. For example, a targeting sequence can be at least 15 bp, 20 bp, 25 bp, 30 bp, 35 bp, 40 bp, 45 bp, 50 bp, 55 bp, 60 bp, 65 bp, 70 bp, 75 bp, 80 bp, 85 bp, 90 bp, 95 bp, 100 bp, 110 bp, 120 bp, 130 bp, 140 bp, 150 bp, 160 bp, 170 bp, 180 bp, 190 bp, 200 bp, 210 bp, 220 bp, 230 bp, 240 bp, or 250 bp. In specific embodiments, the targeting sequence is at least 22 bp. In specific embodiments, the targeting sequence is at least 30 bp.

[0087] In some embodiments, target-specific gNAs can comprise a nucleic acid sequence that is complementary to a region on the opposite strand of the targeted nucleic acid sequence 5' to a PAM sequence, which can be recognized by a nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) protein. In some embodiments the targeted nucleic acid sequence is immediately 5' to a PAM sequence. In specific embodiments, the nucleic acid sequence of the gNA that is complementary to a region in a target nucleic acid is 15-250 bp. In specific embodiments, the nucleic acid sequence of the gNA that is complementary to a region in a target nucleic acid is 22, 23, 24, 25, 30, 35, 40, 45, 50, 60, 70, 75, 80, 90, or 100 bp.

[0088] In some particular embodiments, the targeting sequence is not 20 bp. In some particular embodiments, the targeting sequence is not 21 bp.

[0089] In some embodiments, the gNAs comprise any purines or pyrimidines (and/or modified versions of the same). In some embodiments, the gNAs comprise adenine, uracil, guanine, and cytosine (and/or modified versions of the same). In some embodiments, the gNAs comprise adenine, thymine, guanine, and cytosine (and/or modified versions of the same). In some embodiments, the gNAs comprise adenine, thymine, guanine, cytosine and uracil (and/or modified versions of the same).

[0090] In some embodiments, the gNAs comprise a label, are attached to a label, or are capable of being labeled. In some embodiments, the gNA comprises a moiety that is further capable of being attached to a label. A label includes, but is not limited to, enzyme, an enzyme substrate, an antibody, an antigen binding fragment, a peptide, a chromophore, a lumiphore, a fluorophore, a chromogen, a hapten, an antigen, a radioactive isotope, a magnetic particle, a metal nanoparticle, a redox active marker group (capable of undergoing a redox reaction), an aptamer, one member of a binding pair, a member of a FRET pair (either a donor or acceptor fluorophore), and combinations thereof.

[0091] In some embodiments, the gNAs are attached to a substrate. The substrate can be made of glass, plastic, silicon, silica-based materials, functionalized polystyrene, functionalized polyethyleneglycol, functionalized organic polymers, nitrocellulose or nylon membranes, paper, cotton, and materials suitable for synthesis. Substrates need not be flat. In some embodiments, the substrate is a 2-dimensional array. In some embodiments, the 2-dimensional array is flat. In some embodiments, the 2-dimensional array is not flat, for example, the array is a wave-like array. Substrates include any type of shape including spherical shapes (e.g., beads). Materials attached to substrates may be attached to any portion of the substrates (e.g., may be attached to an interior portion of a porous substrates material). In some embodiments, the substrate is a 3-dimensional array, for example, a microsphere. In some embodiments, the microsphere is magnetic. In some embodiments, the microsphere is glass. In some embodiments, the microsphere is made of polystyrene. In some embodiments, the microsphere is silica-based. In some embodiments, the substrate is an array with interior surface, for example, is a straw, tube, capillary, cylindrical, or microfluidic chamber array. In some embodiments, the substrate comprises multiple straws, capillaries, tubes, cylinders, or chambers.

Nucleic Acids Encoding gNAs

[0092] Also provided herein are nucleic acids encoding for gNAs (e.g., gRNAs or gDNAs). In some embodiments, by encoding it is meant that a gNA results from the transcription of a nucleic acid encoding for a gNA (e.g., gRNA). In some embodiments, by encoding, it is meant that the nucleic acid is a template for the transcription of a gNA (e.g., gRNA). In some embodiments, by encoding, it is meant that a gNA results from the reverse transcription of a nucleic acid encoding for a gNA. In some embodiments, by encoding, it is meant that the nucleic acid is a template for the reverse transcription of a gNA. In some embodiments, by encoding, it is meant that a gNA results from the amplification of a

nucleic acid encoding for a gNA. In some embodiments, by encoding, it is meant that the nucleic acid is a template for the amplification of a gNA.

[0093] In some embodiments the nucleic acid encoding for a gNA comprises a first segment comprising a regulatory region; a second segment comprising targeting sequence, wherein the second segment can range from 15 bp – 250 bp; and a third segment comprising a nucleic acid encoding a nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) protein-binding sequence.

[0094] In some embodiments, the nucleic acids encoding for gNAs comprise DNA. In some embodiments, the first segment is double stranded DNA. In some embodiments, the first segment is single stranded DNA. In some embodiments, the second segment is single stranded DNA. In some embodiments, the third segment is single stranded DNA. In some embodiments, the second segment is double stranded DNA. In some embodiments, the third segment is double stranded DNA.

[0095] In some embodiments, the nucleic acids encoding for gNAs comprise RNA.

[0096] In some embodiments the nucleic acids encoding for gNAs comprise DNA and RNA.

[0097] In some embodiments, the regulatory region is a region capable of binding a transcription factor. In some embodiments, the regulatory region comprises a promoter. In some embodiments, the promoter is selected from the group consisting of T7, SP6, and T3.

Collections of gNAs

[0098] Provided herein are collections (interchangeably referred to as libraries) of gNAs.

[0099] As used herein, a collection of gNAs denotes a mixture of gNAs containing at least 10^2 unique gNAs. In some embodiments a collection of gNAs contains at least 10^2 , at least 10^3 , at least 10^4 , at least 10^5 , at least 10^6 , at least 10^7 , at least 10^8 , at least 10^9 , at least 10^{10} unique gNAs. In some embodiments a collection of gNAs contains a total of at least 10^2 , at least 10^3 , at least 10^4 , at least 10^5 , at least 10^6 , at least 10^7 , at least 10^8 , at least 10^9 , at least 10^{10} gNAs.

[0100] In some embodiments, a collection of gNAs comprises a first NA segment comprising a targeting sequence; and a second NA segment comprising a nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) protein-binding sequence, wherein at least 10% of the gNAs in the collection vary in size. In some embodiments, the first and second segments are in 5'- to 3'-order'.

[0101] In some embodiments, the size of the first segment varies from 15-250 bp, or 30-100 bp, or 22-30 bp, or 15-50bp, or 15-75 bp, or 15-100 bp, or 15-125 bp, or 15-150 bp, or 15-175 bp, or 15-200 bp, or 15-225 bp, or 15-250 bp, or 22-50 bp, or 22-75 bp, or 22-100 bp, or 22-125 bp, or 22-150 bp, or 22-175 bp, or 22-200 bp, or 22-225 bp, or 22-250 bp across the collection of gNAs.

[0102] In some embodiments, at least 10%, or at least 15%, or at last 20%, or at least 25%, or at least 30%, or at least 35%, or at least 40%, or at least 45%, or at least 50%, or at least 55%, or at least 60%, or

at least 65%, or at least 70%, or at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 95%, or 100% of the first segments in the collection are greater than 21 bp.

[0103] In some embodiments, at least 10%, or at least 15%, or at least 20%, or at least 25%, or at least 30%, or at least 35%, or at least 40%, or at least 45%, or at least 50%, or at least 55%, or at least 60%, or at least 65%, or at least 70%, or at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 95%, or 100% of the first segments in the collection are greater than 25 bp.

[0104] In some embodiments, at least 10%, or at least 15%, or at least 20%, or at least 25%, or at least 30%, or at least 35%, or at least 40%, or at least 45%, or at least 50%, or at least 55%, or at least 60%, or at least 65%, or at least 70%, or at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 95%, or 100% of the first segments in the collection are greater than 30 bp.

[0105] In some embodiments, at least 10%, or at least 15%, or at least 20%, or at least 25%, or at least 30%, or at least 35%, or at least 40%, or at least 45%, or at least 50%, or at least 55%, or at least 60%, or at least 65%, or at least 70%, or at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 95%, or 100% of the first segments in the collection are 15-50 bp.

[0106] In some embodiments, at least 10%, or at least 15%, or at least 20%, or at least 25%, or at least 30%, or at least 35%, or at least 40%, or at least 45%, or at least 50%, or at least 55%, or at least 60%, or at least 65%, or at least 70%, or at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 95%, or 100% of the first segments in the collection are 30-100 bp.

[0107] In some particular embodiments, the size of the first segment is not 20 bp.

[0108] In some particular embodiments, the size of the first segment is not 21 bp.

[0109] In some embodiments, the gNAs and/or the targeting sequence of the gNAs in the collection of gRNAs comprise unique 5' ends. In some embodiments, the collection of gNAs exhibit variability in sequence of the 5' end of the targeting sequence, across the members of the collection. In some embodiments, the collection of gNAs exhibit variability at least 5%, or at least 10%, or at least 15%, or at least 20%, or at least 25%, or at least 30%, or at least 35%, or at least 40%, or at least 45%, or at least 50%, or at least 55%, or at least 60%, or at least 65%, or at least 70%, or at least 75% variability in the sequence of the 5' end of the targeting sequence, across the members of the collection.

[0110] In some embodiments, the 3' end of the gNA targeting sequence can be any purine or pyrimidine (and/or modified versions of the same). In some embodiments, the 3' end of the gNA targeting sequence is an adenine. In some embodiments, the 3' end of the gNA targeting sequence is a guanine. In some embodiments, the 3' end of the gNA targeting sequence is a cytosine. In some embodiments, the 3' end of the gNA targeting sequence is a uracil. In some embodiments, the 3' end of the gNA targeting sequence is a thymine. In some embodiments, the 3' end of the gNA targeting sequence is not cytosine.

[0111] In some embodiments, the collection of gNAs comprises targeting sequences which can base-pair with the targeted DNA, wherein the target of interest is spaced at least every 1 bp, at least every 2 bp, at least every 3 bp, at least every 4 bp, at least every 5 bp, at least every 6 bp, at least every 7 bp, at least every 8 bp, at least every 9 bp, at least every 10 bp, at least every 11 bp, at least every 12 bp, at least every 13 bp, at least every 14 bp, at least every 15 bp, at least every 16 bp, at least every 17 bp, at least every 18 bp, at least every 19 bp, 20 bp, at least every 25 bp, at least every 30 bp, at least every 40 bp, at least every 50 bp, at least every 100 bp, at least every 200 bp, at least every 300 bp, at least every 400 bp, at least every 500 bp, at least every 600 bp, at least every 700 bp, at least every 800 bp, at least every 900 bp, at least every 1000 bp, at least every 2500 bp, at least every 5000 bp, at least every 10,000 bp, at least every 15,000 bp, at least every 20,000 bp, at least every 25,000 bp, at least every 50,000 bp, at least every 100,000 bp, at least every 250,000 bp, at least every 500,000 bp, at least every 750,000bp, or even at least every 1,000,000 bp across a genome of interest.

[0112] In some embodiments, the collection of gNAs comprises a first NA segment comprising a targeting sequence; and a second NA segment comprising a nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) protein-binding sequence, wherein the gNAs in the collection can have a variety of second NA segments with various specificities for protein members of the nucleic acid-guided nuclease system (e.g., CRISPR/Cas system). For example a collection of gNAs as provided herein, can comprise members whose second segment comprises a nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) protein-binding sequence specific for a first nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) protein; and also comprises members whose second segment comprises a nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) protein-binding sequence specific for a second nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) protein, wherein the first and second nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) proteins are not the same. In some embodiments a collection of gNAs as provided herein comprises members that exhibit specificity to at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, or even at least 20 nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) proteins. In one specific embodiment, a collection of gNAs as provided herein comprises members that exhibit specificity for a Cas9 protein and another protein selected from the group consisting of Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, and Cm5.

[0113] In some embodiments, a plurality of the gNA members of the collection are attached to a label, comprise a label or are capable of being labeled. In some embodiments, the gNA comprises is a moiety that is further capable of being attached to a label. A label includes, but is not limited to, enzyme, an enzyme substrate, an antibody, an antigen binding fragment, a peptide, a chromophore, a lumiphore, a fluorophore, a chromogen, a hapten, an antigen, a radioactive isotope, a magnetic particle, a metal

nanoparticle, a redox active marker group (capable of undergoing a redox reaction), an aptamer, one member of a binding pair, a member of a FRET pair (either a donor or acceptor fluorophore), and combinations thereof.

[0114] In some embodiments, a plurality of the gNA members of the collection are attached to a substrate. The substrate can be made of glass, plastic, silicon, silica-based materials, functionalized polystyrene, functionalized polyethyleneglycol, functionalized organic polymers, nitrocellulose or nylon membranes, paper, cotton, and materials suitable for synthesis. Substrates need not be flat. In some embodiments, the substrate is a 2-dimensional array. In some embodiments, the 2-dimensional array is flat. In some embodiments, the 2-dimensional array is not flat, for example, the array is a wave-like array. Substrates include any type of shape including spherical shapes (e.g., beads). Materials attached to substrates may be attached to any portion of the substrates (e.g., may be attached to an interior portion of a porous substrates material). In some embodiments, the substrate is a 3-dimensional array, for example, a microsphere. In some embodiments, the microsphere is magnetic. In some embodiments, the microsphere is glass. In some embodiments, the microsphere is made of polystyrene. In some embodiments, the microsphere is silica-based. In some embodiments, the substrate is an array with interior surface, for example, is a straw, tube, capillary, cylindrical, or microfluidic chamber array. In some embodiments, the substrate comprises multiple straws, capillaries, tubes, cylinders, or chambers.

Collections of Nucleic Acids Encoding gNAs

[0115] Provided herein are collections (interchangeably referred to as libraries) of nucleic acids encoding for gNAs (e.g., gRNAs or gDNAs). In some embodiments, by encoding it is meant that a gNA results from the transcription of a nucleic acid encoding for a gNA. In some embodiments, by encoding, it is meant that the nucleic acid is a template for the transcription of a gNA.

[0116] As used herein, a collection of nucleic acids encoding for gNAs denotes a mixture of nucleic acids containing at least 10^2 unique nucleic acids. In some embodiments a collection of nucleic acids encoding for gNAs contains at least 10^2 , at least 10^3 , at least 10^4 , at least 10^5 , at least 10^6 , at least 10^7 , at least 10^8 , at least 10^9 , at least 10^{10} unique nucleic acids encoding for gNAs. In some embodiments a collection of nucleic acids encoding for gNAs contains a total of at least 10^2 , at least 10^3 , at least 10^4 , at least 10^5 , at least 10^6 , at least 10^7 , at least 10^8 , at least 10^9 , at least 10^{10} nucleic acids encoding for gNAs.

[0117] In some embodiments, a collection of nucleic acids encoding for gNAs comprises a first segment comprising a regulatory region; a second segment comprising a targeting sequence; and a third segment comprising a nucleic acid encoding a nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) protein-binding sequence, wherein at least 10% of the nucleic acids in the collection vary in size.

[0118] In some embodiments, the first, second, and third segments are in 5'- to 3'-order'.

[0119] In some embodiments, the nucleic acids encoding for gNAs comprise DNA. In some embodiments, the first segment is single stranded DNA. In some embodiments, the first segment is double stranded DNA. In some embodiments, the second segment is single stranded DNA. In some embodiments, the third segment is single stranded DNA. In some embodiments, the second segment is double stranded DNA. In some embodiments, the third segment is double stranded DNA.

[0120] In some embodiments, the nucleic acids encoding for gNAs comprise RNA.

[0121] In some embodiments the nucleic acids encoding for gNAs comprise DNA and RNA.

[0122] In some embodiments, the regulatory region is a region capable of binding a transcription factor. In some embodiments, the regulatory region comprises a promoter. In some embodiments, the promoter is selected from the group consisting of T7, SP6, and T3.

[0123] In some embodiments, the size of the second segments (targeting sequence) in the collection varies from 15-250 bp, or 30-100 bp, or 22-30 bp, or 15-50 bp, or 15-75 bp, or 15-100 bp, or 15-125 bp, or 15-150 bp, or 15-175 bp, or 15-200 bp, or 15-225 bp, or 15-250 bp, or 22-50 bp, or 22-75 bp, or 22-100 bp, or 22-125 bp, or 22-150 bp, or 22-175 bp, or 22-200 bp, or 22-225 bp, or 22-250 bp across the collection of gNAs.

[0124] In some embodiments, at least 10%, or at least 15%, or at last 20%, or at least 25%, or at least 30%, or at least 35%, or at least 40%, or at least 45%, or at least 50%, or at least 55%, or at least 60%, or at least 65%, or at least 70%, or at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 95%, or 100% of the second segments in the collection are greater than 21 bp.

[0125] In some embodiments, at least 10%, or at least 15%, or at last 20%, or at least 25%, or at least 30%, or at least 35%, or at least 40%, or at least 45%, or at least 50%, or at least 55%, or at least 60%, or at least 65%, or at least 70%, or at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 95%, or 100% of the second segments in the collection are greater than 25 bp.

[0126] In some embodiments, at least 10%, or at least 15%, or at last 20%, or at least 25%, or at least 30%, or at least 35%, or at least 40%, or at least 45%, or at least 50%, or at least 55%, or at least 60%, or at least 65%, or at least 70%, or at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 95%, or 100% of the second segments in the collection are greater than 30 bp.

[0127] In some embodiments, at least 10%, or at least 15%, or at last 20%, or at least 25%, or at least 30%, or at least 35%, or at least 40%, or at least 45%, or at least 50%, or at least 55%, or at least 60%, or at least 65%, or at least 70%, or at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 95%, or 100% of the second segments in the collection are 15-50bp.

[0128] In some embodiments, at least 10%, or at least 15%, or at last 20%, or at least 25%, or at least 30%, or at least 35%, or at least 40%, or at least 45%, or at least 50%, or at least 55%, or at least 60%, or at least 65%, or at least 70%, or at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 95%, or 100% of the second segments in the collection are 30-100bp.

[0129] In some particular embodiments, the size of the second segment is not 20 bp.

[0130] In some particular embodiments, the size of the second segment is not 21 bp.

[0131] In some embodiments, the gNAs and/or the targeting sequence of the gNAs in the collection of gNAs comprise unique 5' ends. In some embodiments, the collection of gNAs exhibit variability in sequence of the 5' end of the targeting sequence, across the members of the collection. In some embodiments, the collection of gNAs exhibit variability at least 5%, or at least 10%, or at least 15%, or at least 20%, or at least 25%, or at least 30%, or at least 35%, or at least 40%, or at least 45%, or at least 50%, or at least 55%, or at least 60%, or at least 65%, or at least 70%, or at least 75% variability in the sequence of the 5' end of the targeting sequence, across the members of the collection.

[0132] In some embodiments, the collection of nucleic acids comprises targeting sequences, wherein the target of interest is spaced at least every 1 bp, at least every 2 bp, at least every 3 bp, at least every 4 bp, at least every 5 bp, at least every 6 bp, at least every 7 bp, at least every 8 bp, at least every 9 bp, at least every 10 bp, at least every 11 bp, at least every 12 bp, at least every 13 bp, at least every 14 bp, at least every 15 bp, at least every 16 bp, at least every 17 bp, at least every 18 bp, at least every 19 bp, 20 bp, at least every 25 bp, at least every 30 bp, at least every 40 bp, at least every 50 bp, at least every 100 bp, at least every 200 bp, at least every 300 bp, at least every 400 bp, at least every 500 bp, at least every 600 bp, at least every 700 bp, at least every 800 bp, at least every 900 bp, at least every 1000 bp, at least every 2500 bp, at least every 5000 bp, at least every 10,000 bp, at least every 15,000 bp, at least every 20,000 bp, at least every 25,000 bp, at least every 50,000 bp, at least every 100,000 bp, at least every 250,000 bp, at least every 500,000 bp, at least every 750,000 bp, or even at least every 1,000,000 bp across a genome of interest.

[0133] In some embodiments, the collection of nucleic acids encoding for gNAs comprise a third segment encoding for a nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) protein-binding sequence, wherein the segments in the collection vary in their specificity for protein members of the nucleic acid-guided nuclease system (e.g., CRISPR/Cas system). For example, a collection of nucleic acids encoding for gNAs as provided herein, can comprise members whose third segment encode for a nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) protein-binding sequence specific for a first nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) protein; and also comprises members whose third segment encodes for a nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) protein-binding sequence specific for a second nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) protein, wherein the first and second nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) proteins are not the same. In some embodiments, a collection of nucleic acids encoding for gNAs as provided herein comprises members that exhibit specificity to at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, or even at least 20 nucleic

acid-guided nuclease system (e.g., CRISPR/Cas system) proteins. In one specific embodiment, a collection of nucleic acids encoding for gNAs as provided herein comprises members that exhibit specificity for a Cas9 protein and another protein selected from the group consisting of Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, and Cm5.

Sequences of Interest

[0134] Provided herein are gNAs and collections of gNAs, derived from any source DNA (for example from genomic DNA, cDNA, artificial DNA, DNA libraries), that can be used to target sequences of interest in a sample for a variety of applications including, but not limited to, enrichment, depletion, capture, partitioning, labeling, regulation, and editing. The gNAs comprise a targeting sequence, directed at sequences of interest.

[0135] In some embodiments, the sequences of interest are genomic sequences (genomic DNA). In some embodiments, the sequences of interest are mammalian genomic sequences. In some embodiments, the sequences of interest are eukaryotic genomic sequences. In some embodiments, the sequences of interest are prokaryotic genomic sequences. In some embodiments, the sequences of interest are viral genomic sequences. In some embodiments, the sequences of interest are bacterial genomic sequences. In some embodiments, the sequences of interest are plant genomic sequences. In some embodiments, the sequences of interest are microbial genomic sequences. In some embodiments, the sequences of interest are genomic sequences from a parasite, for example a eukaryotic parasite. In some embodiments, the sequences of interest are host genomic sequences (e.g., the host organism of a microbiome, a parasite, or a pathogen). In some embodiments, the sequences of interest are abundant genomic sequences, such as sequences from the genome or genomes of the most abundant species in a sample.

[0136] In some embodiments, the sequences of interest comprise repetitive DNA. In some embodiments, the sequences of interest comprise abundant DNA. In some embodiments, the sequences of interest comprise mitochondrial DNA. In some embodiments, the sequences of interest comprise ribosomal DNA. In some embodiments, the sequences of interest comprise centromeric DNA. In some embodiments, the sequences of interest comprise DNA comprising Alu elements (Alu DNA). In some embodiments, the sequences of interest comprise long interspersed nuclear elements (LINE DNA). In some embodiments, the sequences of interest comprise short interspersed nuclear elements (SINE DNA). In some embodiments, the abundant DNA comprises ribosomal DNA.

[0137] In some embodiments, the sequences of interest comprise single nucleotide polymorphisms (SNPs), short tandem repeats (STRs), cancer genes, inserts, deletions, structural variations, exons, genetic mutations, or regulatory regions.

[0138] In some embodiments, the sequences of interest can be a genomic fragment, comprising a region of the genome, or the whole genome itself. In one embodiment, the genome is a DNA genome. In another embodiment, the genome is a RNA genome.

[0139] In some embodiments, the sequences of interest are from a eukaryotic or prokaryotic organism; from a mammalian organism or a non-mammalian organism; from an animal or a plant; from a bacteria or virus; from an animal parasite; from a pathogen.

[0140] In some embodiments, the sequences of interest are from any mammalian organism. In one embodiment the mammal is a human. In another embodiment the mammal is a livestock animal, for example a horse, a sheep, a cow, a pig, or a donkey. In another embodiment, a mammalian organism is a domestic pet, for example a cat, a dog, a gerbil, a mouse, a rat. In another embodiment the mammal is a type of a monkey.

[0141] In some embodiments, the sequences of interest are from any bird or avian organism. An avian organism includes but is not limited to chicken, turkey, duck and goose.

[0142] In some embodiments, the sequences of interest are from a plant. In one embodiment, the plant is rice, maize, wheat, rose, grape, coffee, fruit, tomato, potato, or cotton.

[0143] In some embodiments, the sequences of interest are from a species of bacteria. In one embodiment, the bacteria are tuberculosis-causing bacteria.

[0144] In some embodiments, the sequences of interest are from a virus.

[0145] In some embodiments, the sequences of interest are from a species of fungi.

[0146] In some embodiments, the sequences of interest are from a species of algae.

[0147] In some embodiments, the sequences of interest are from any mammalian parasite.

[0148] In some embodiments, the sequences of interest are obtained from any mammalian parasite. In one embodiment, the parasite is a worm. In another embodiment, the parasite is a malaria-causing parasite. In another embodiment, the parasite is a Leishmaniasis-causing parasite. In another embodiment, the parasite is an amoeba.

[0149] In some embodiments, the sequences of interest are from a pathogen.

Targeting Sequences

[0150] As used herein, a targeting sequence is one that directs the gNA to the sequences of interest in a sample. For example, a targeting sequence targets a particular sequence of interest, for example the targeting sequence targets a genomic sequence of interest.

[0151] Provided herein are gNAs and collections of gNAs that comprise a segment that comprises a targeting sequence. Also provided herein, are nucleic acids encoding for gNAs, and collections of nucleic acids encoding for gNAs that comprise a segment encoding for a targeting sequence.

[0152] In some embodiments, the targeting sequence comprises DNA.

[0153] In some embodiments, the targeting sequence comprises RNA.

[0154] In some embodiments, the targeting sequence comprises RNA, and shares at least 70% sequence identity, at least 75% sequence identity, at least 80% sequence identity, at least 85% sequence identity, at least 90% sequence identity, at least 95% sequence identity, or shares 100% sequence identity to a sequence 5' to a PAM sequence on a sequence of interest, except that the RNA comprises uracils instead of thymines. In some embodiments, the PAM sequence is AGG, CGG, or TGG.

[0155] In some embodiments, the targeting sequence comprises DNA, and shares at least 70% sequence identity, at least 75% sequence identity, at least 80% sequence identity, at least 85% sequence identity, at least 90% sequence identity, at least 95% sequence identity, or shares 100% sequence identity to a sequence 5' to a PAM sequence on a sequence of interest.

[0156] In some embodiments, the targeting sequence comprises RNA and is complementary to the strand opposite to a sequence of nucleotides 5' to a PAM sequence. In some embodiments, the targeting sequence is at least 70% complementary, at least 75% complementary, at least 80% complementary, at least 85% complementary, at least 90% complementary, at least 95% complementary, or is 100% complementary to the strand opposite to a sequence of nucleotides 5' to a PAM sequence. In some embodiments, the PAM sequence is AGG, CGG, or TGG.

[0157] In some embodiments, the targeting sequence comprises DNA and is complementary to the strand opposite to a sequence of nucleotides 5' to a PAM sequence. In some embodiments, the targeting sequence is at least 70% complementary, at least 75% complementary, at least 80% complementary, at least 85% complementary, at least 90% complementary, at least 95% complementary, or is 100% complementary to the strand opposite to a sequence of nucleotides 5' to a PAM sequence. In some embodiments, the PAM sequence is AGG, CGG, or TGG.

[0158] In some embodiments, a DNA encoding for a targeting sequence of a gRNA shares at least 70% sequence identity, at least 75% sequence identity, at least 80% sequence identity, at least 85% sequence identity, at least 90% sequence identity, at least 95% sequence identity, or shares 100% sequence identity to the strand opposite to a sequence of nucleotides 5' to a PAM sequence. In some embodiments, the PAM sequence is AGG, CGG, or TGG.

[0159] In some embodiments, a DNA encoding for a targeting sequence of a gRNA is complementary to the strand opposite to a sequence of nucleotides 5' to a PAM sequence and is at least 70% complementary, at least 75% complementary, at least 80% complementary, at least 85% complementary, at least 90% complementary, at least 95% complementary, or is 100% complementary to a sequence 5' to a PAM sequence on a sequence of interest. In some embodiments, the PAM sequence is AGG, CGG, or TGG.

Nucleic Acid-Guided Nuclease System Proteins

[0160] Provided herein are gNAs and collections of gNAs comprising a segment that comprises a nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) protein-binding sequence. Also provided herein, are nucleic acids encoding for gNAs, and collections of nucleic acids encoding for gNAs that comprise a segment encoding a nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) protein-binding sequence. A nucleic acid-guided nuclease system can be an RNA-guided nuclease system. A nucleic acid-guided nuclease system can be a DNA-guided nuclease system.

[0161] Methods of the present disclosure can utilize nucleic acid-guided nucleases. As used herein, a “nucleic acid-guided nuclease” is any nuclease that cleaves DNA, RNA or DNA/RNA hybrids, and which uses one or more nucleic acid guide nucleic acids (gNAs) to confer specificity. Nucleic acid-guided nucleases include CRISPR/Cas system proteins as well as non-CRISPR/Cas system proteins.

[0162] The nucleic acid-guided nucleases provided herein can be DNA guided DNA nucleases; DNA guided RNA nucleases; RNA guided DNA nucleases; or RNA guided RNA nucleases. The nucleases can be endonucleases. The nucleases can be exonucleases. In one embodiment, the nucleic acid-guided nuclease is a nucleic acid-guided-DNA endonuclease. In one embodiment, the nucleic acid-guided nuclease is a nucleic acid-guided-RNA endonuclease.

[0163] A nucleic acid-guided nuclease system protein-binding sequence is a nucleic acid sequence that binds any protein member of a nucleic acid-guided nuclease system. For example, a CRISPR/Cas system protein-binding sequence is a nucleic acid sequence that binds any protein member of a CRISPR/Cas system.

[0164] In some embodiments, the nucleic acid-guided nuclease is selected from the group consisting of CAS Class I Type I, CAS Class I Type III, CAS Class I Type IV, CAS Class II Type II, and CAS Class II Type V. In some embodiments, CRISPR/Cas system proteins include proteins from CRISPR Type I systems, CRISPR Type II systems, and CRISPR Type III systems. In some embodiments, the nucleic acid-guided nuclease is selected from the group consisting of Cas9, Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, Cm5, Csf1, C2c2, and NgAgo.

[0165] In some embodiments, nucleic acid-guided nuclease system proteins (e.g., CRISPR/Cas system proteins) can be from any bacterial or archaeal species.

[0166] In some embodiments, the nucleic acid-guided nuclease system proteins (e.g., CRISPR/Cas system proteins) are from, or are derived from nucleic acid-guided nuclease system proteins (e.g., CRISPR/Cas system proteins) from *Streptococcus pyogenes*, *Staphylococcus aureus*, *Neisseria meningitidis*, *Streptococcus thermophiles*, *Treponema denticola*, *Francisella tularensis*, *Pasteurella multocida*, *Campylobacter jejuni*, *Campylobacter lari*, *Mycoplasma gallisepticum*, *Nitratifactor saluginis*, *Parvibaculum lavamentivorans*, *Roseburia intestinalis*, *Neisseria cinerea*, *Gluconacetobacter diazotrophicus*, *Azospirillum*, *Sphaerochaeta globus*, *Flavobacterium columnare*, *Fluviicola taffensis*, *Bacteroides coprophilus*, *Mycoplasma mobile*, *Lactobacillus farciminis*, *Streptococcus pasteurianus*,

Lactobacillus johnsonii, *Staphylococcus pseudintermedius*, *Filifactor alocis*, *Legionella pneumophila*, *Suterella wadsworthensis*, or *Corynebacter diphtheria*.

[0167] In some embodiments, examples of nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) proteins can be naturally occurring or engineered versions.

[0168] In some embodiments, naturally occurring nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) proteins include Cas9, Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, and Csm5. Engineered versions of such proteins can also be employed.

[0169] In some embodiments, engineered examples of nucleic acid-guided nuclease system (e.g., CRISPR/Cas system) proteins include catalytically dead nucleic acid-guided nuclease system proteins. The term "catalytically dead" generally refers to a nucleic acid-guided nuclease system protein that has inactivated nucleases (e.g., HNH and RuvC nucleases). Such a protein can bind to a target site in any nucleic acid (where the target site is determined by the guide NA), but the protein is unable to cleave or nick the target nucleic acid (e.g., double-stranded DNA). In some embodiments, the nucleic acid-guided nuclease system catalytically dead protein is a catalytically dead CRISPR/Cas system protein, such as catalytically dead Cas9 (dCas9). Accordingly, the dCas9 allows separation of the mixture into unbound nucleic acids and dCas9-bound fragments. In one embodiment, a dCas9/gRNA complex binds to targets determined by the gRNA sequence. The dCas9 bound can prevent cutting by Cas9 while other manipulations proceed. In another embodiment, the dCas9 can be fused to another enzyme, such as a transposase, to target that enzyme's activity to a specific site. Naturally occurring catalytically dead nucleic acid-guided nuclease system proteins can also be employed.

[0170] In some embodiments, engineered examples of nucleic acid-guided nuclease (e.g., CRISPR/Cas) system proteins also include nucleic acid-guided nickases (e.g., Cas nickases). A nucleic acid-guided nickase refers to a modified version of a nucleic acid-guided nuclease system protein, containing a single inactive catalytic domain. In one embodiment, the nucleic acid-guided nickase is a Cas nickase, such as Cas9 nickase. A Cas9 nickase may contain a single inactive catalytic domain, for example, either the RuvC- or the HNH-domain. With only one active nuclease domain, the Cas9 nickase cuts only one strand of the target DNA, creating a single-strand break or "nick". Depending on which mutant is used, the guide NA-hybridized strand or the non-hybridized strand may be cleaved. Nucleic acid-guided nickases bound to 2 gNAs that target opposite strands will create a double-strand break in a target double-stranded DNA. This "dual nickase" strategy can increase the specificity of cutting because it requires that both nucleic acid-guided nuclease/gNA (e.g., Cas9/gRNA) complexes be specifically bound at a site before a double-strand break is formed. Naturally occurring nickase nucleic acid-guided nuclease system proteins can also be employed.

[0171] In some embodiments, engineered examples of nucleic acid-guided nuclease system proteins also include nucleic acid-guided nuclease system fusion proteins. For example, a nucleic acid-guided

nuclease (e.g., CRISPR/Cas) system protein may be fused to another protein, for example an activator, a repressor, a nuclease, a fluorescent molecule, a radioactive tag, or a transposase.

[0172] In some embodiments, the nucleic acid-guided nuclease system protein-binding sequence comprises a gNA (e.g., gRNA) stem-loop sequence.

[0173] In some embodiments, a double-stranded DNA sequence encoding the gNA (e.g., gRNA) stem-loop sequence comprises the following DNA sequence on one strand (5'>3',
GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT) (SEQ ID NO: 3), and its reverse-complementary DNA on the other strand (5'>3',
AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC) (SEQ ID NO: 4).

[0174] In some embodiments, a single-stranded DNA sequence encoding the gNA (e.g., gRNA) stem-loop sequence comprises the following DNA sequence: (5'>3',
AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC) (SEQ ID NO: 4), wherein the single-stranded DNA serves as a transcription template.

[0175] In some embodiments, the gNA (e.g., gRNA) stem-loop sequence comprises the following RNA sequence: (5'>3',
GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU) (SEQ ID NO: 1)

[0176] In some embodiments, a double-stranded DNA sequence encoding the gNA (e.g., gRNA) stem-loop sequence comprises the following DNA sequence on one strand (5'>3',
GTTTTAGAGCTATGCTGGAAACAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTC) (SEQ ID NO: 5), and its reverse-complementary DNA on the other strand (5'>3',
GAAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATGCTGTTTCCAGCATAGCTCTAAAAC) (SEQ ID NO: 6).

[0177] In some embodiments, a single-stranded DNA sequence encoding the gNA (e.g., gRNA) stem-loop sequence comprises the following DNA sequence: (5'>3',
GAAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATGCTGTTTCCAGCATAGCTCTAAAAC) (SEQ ID NO: 6), wherein the single-stranded DNA serves as a transcription template.

[0178] In some embodiments, the gNA (e.g., gRNA) stem-loop sequence comprises the following RNA sequence: (5'>3',

GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUA AAAUAAGGCUAGUCCGUUAUCAA
CUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUUC) (SEQ ID NO: 2).

[0179] In some embodiments, provided herein is a nucleic acid encoding for a gNA (e.g., gRNA) comprising a first segment comprising a regulatory region; a second segment encoding a targeting sequence; and a third segment comprising a nucleic acid encoding a nucleic acid-guided nuclease (e.g., CRISPR/Cas) system protein-binding sequence. In some embodiments, the third segment comprises a single transcribed component, which upon transcription yields a NA (e.g., RNA) stem-loop sequence. In some embodiments, the third segment comprising a single transcribed component that encodes for the gNA (e.g., gRNA) stem-loop sequence is double-stranded, comprises the following DNA sequence on one strand (5'>3',

GTTT TAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTG
GCACCGAGTCGGTGCTTTTTTTT) (SEQ ID NO: 3), and its reverse-complementary DNA on the other strand (5'>3',

AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACT
TGCTATTTCTAGCTCTAAAAC) (SEQ ID NO: 4). In some embodiments, the third segment comprising a single transcribed component that encodes for the gNA (e.g., gRNA) stem-loop sequence is single-stranded, and comprises the following DNA sequence: (5'>3',

AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACT
TGCTATTTCTAGCTCTAAAAC) (SEQ ID NO: 4), wherein the single-stranded DNA serves as a transcription template. In some embodiments, upon transcription from the single transcribed component, the resulting gNA (e.g., gRNA) stem-loop sequence comprises the following RNA sequence: (5'>3',
GUUUUAGAGCUAGAAAUAGCAAGUUA AAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAG
UGGCACCGAGUCGGUGCUUUUUUU) (SEQ ID NO: 1). In some embodiments, the third segment comprising a single transcribed component that encodes for the gNA (e.g., gRNA) stem-loop sequence is double-stranded, comprises the following DNA sequence on one strand (5'>3',

GTTT TAGAGCTATGCTGGAACAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTT
GAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT) (SEQ ID NO: 5), and its reverse-complementary DNA on the other strand (5'>3',

GAAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACT
TTGCTATGCTGTTTCCAGCATAGCTCTAAAAC) (SEQ ID NO: 6). In some embodiments, the third segment comprising a single transcribed component that encodes for the gNA (e.g., gRNA) stem-loop sequence is single-stranded, and comprises the following DNA sequence: (5'>3',

GAAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACT
TTGCTATGCTGTTTCCAGCATAGCTCTAAAAC) (SEQ ID NO: 6), wherein the single-stranded DNA serves as a transcription template. In some embodiments, upon transcription from the single

transcribed component, the yielded gRNA stem-loop sequence comprises the following RNA sequence: (5'>3',

GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUA AAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUUC) (SEQ ID NO: 2). In some embodiments, the third segment comprises two sub-segments, which encode for a crRNA and a tracrRNA upon transcription. In some embodiment, the crRNA does not comprise the N20 plus the extra sequence which can hybridize with tracrRNA. In some embodiments, the crRNA comprises the extra sequence which can hybridize with tracrRNA. In some embodiments, the two sub-segments are independently transcribed. In some embodiments, the two sub-segments are transcribed as a single unit. In some embodiments, the DNA encoding the crRNA comprises N_{target}GTTT TAGAGCTATGCTGTTTTG (SEQ ID NO: 7), where N_{target} represents the targeting sequence. In some embodiments, the DNA encoding the tracrRNA comprises the sequence

GGAACCATTC AAAACAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGCTTTTTTTT (SEQ ID NO: 8).

[0180] In some embodiments, provided herein is a nucleic acid encoding for a gNA (e.g., gRNA) comprising a first segment comprising a regulatory region; a second segment encoding a targeting sequence; and a third segment comprising a nucleic acid encoding a nucleic acid-guided nuclease (e.g., CRISPR/Cas) system protein-binding sequence. In some embodiments, the third segment comprises a DNA sequence, which upon transcription yields a gRNA stem-loop sequence capable of binding a nucleic acid-guided nuclease (e.g., CRISPR/Cas) system protein. In one embodiment, the DNA sequence can be double-stranded. In some embodiments, the third segment double stranded DNA comprises the following DNA sequence on one strand (5'>3',

GTTT TAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTG GCACCGAGTCGGTGCTTTTTTTT) (SEQ ID NO: 3), and its reverse-complementary DNA on the other strand (5'>3',

AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACT TGCTATTTCTAGCTCTAAAAC) (SEQ ID NO: 4). In some embodiments, the third segment double stranded DNA comprises the following DNA sequence on one strand (5'>3',

GTTT TAGAGCTATGCTGGAAACAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTT GAAAAAGTGGCACCGAGTCGGTGCTTTTTTTC) (SEQ ID NO: 5), and its reverse-complementary DNA on the other strand (5'>3',

GAAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACT TTGCTATGCTGTTTCCAGCATAGCTCTAAAAC) (SEQ ID NO: 6). In one embodiment, the DNA sequence can be single-stranded. In some embodiments, the third segment single stranded DNA comprises the following DNA sequence (5'>3',

AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC) (SEQ ID NO: 4), wherein the single-stranded DNA serves as a transcription template. In some embodiments, the third segment single stranded DNA comprises the following DNA sequence (5'>3',

GAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATGCTGTTTCCAGCATAGCTCTAAAAC) (SEQ ID NO: 6), wherein the single-stranded DNA serves as a transcription template. In some embodiments, the third segment comprises a DNA sequence which, upon transcription, yields a first RNA sequence that is capable of forming a hybrid with a second RNA sequence, and which hybrid is capable of CRISPR/Cas system protein binding. In some embodiments, the third segment is double-stranded DNA comprising the DNA sequence on one strand: (5'>3', GTTTTAGAGCTATGCTGTTTTG) (SEQ ID NO: 9) and its reverse complementary DNA sequence on the other strand: (5'>3', CAAAACAGCATAGCTCTAAAAC) (SEQ ID NO: 10). In some embodiments, the third segment is single-stranded DNA comprising the DNA sequence of (5'>3', CAAAACAGCATAGCTCTAAAAC) (SEQ ID NO: 10). In some embodiments, the second segment and the third segment together encode for a crRNA sequence. In some embodiments, the second RNA sequence that is capable of forming a hybrid with the first RNA sequence encoded by the third segment of the nucleic acid encoding a gRNA is a tracrRNA. In some embodiments, the tracrRNA comprises the sequence (5'>3',

GGAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCUUUUUUU) (SEQ ID NO: 11). In some embodiments, the tracrRNA is encoded by a double-stranded DNA comprising sequence of (5'>3',

GGAACCATTCAAAACAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGCTTTTTTTT) (SEQ ID NO: 8), and optionally fused with a regulatory sequence at its 5' end. In some embodiments, the regulatory sequence can be bound by a transcription factor. In some embodiments, the regulatory sequence is a promoter. In some embodiments, the regulatory sequence is a T7 promoter, comprising the sequence of (5'>3', GCCTCGAGCTAATACGACTCACTATAGAG) (SEQ ID NO: 12).

[0181] In some embodiments, provided herein is a nucleic acid encoding for a gRNA comprising a first segment comprising a regulatory region; a second segment encoding a targeting sequence; and a third segment comprising a nucleic acid encoding a nucleic acid-guided nuclease (e.g., CRISPR/Cas) system protein-binding sequence. In some embodiments, the third segment encodes for a RNA sequence that, upon post-transcriptional cleavage, yields a first RNA segment and a second RNA segment. In some embodiments, the first RNA segment comprises a crRNA and the second RNA segment comprises a tracrRNA, which can form a hybrid and together, provide for nucleic acid-guided nuclease (e.g., CRISPR/Cas) system protein binding. In some embodiments, the third segment further comprises a

spacer in between the transcriptional unit for the first RNA segment and the second RNA segment, which spacer comprises an enzyme cleavage site.

[0182] In some embodiments, provided herein is a gNA (e.g., gRNA) comprising a first NA segment comprising a targeting sequence and a second NA segment comprising a nucleic acid-guided nuclease (e.g., CRISPR/Cas) system protein-binding sequence. In some embodiments, the size of the first segment is greater than 30 bp. In some embodiments, the second segment comprises a single segment, which comprises the gRNA stem-loop sequence. In some embodiments, the gRNA stem-loop sequence comprises the following RNA sequence: (5'>3',

GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAG
UGGCACCGAGUCGGUGCUUUUUUU) (SEQ ID NO: 1). In some embodiments, the gRNA stem-

loop sequence comprises the following RNA sequence: (5'>3',

GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAA
CUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUUC) (SEQ ID NO: 2). In some embodiments,

the second segment comprises two sub-segments: a first RNA sub-segment (crRNA) that forms a hybrid with a second RNA sub-segment (tracrRNA), which together act to direct nucleic acid-guided nuclease (e.g., CRISPR/Cas) system protein binding. In some embodiments, the sequence of the second sub-segment comprises GUUUUAGAGCUAUGCUGUUUUG. In some embodiments, the first RNA segment and the second RNA segment together forms a crRNA sequence. In some embodiments, the other RNA that will form a hybrid with the second RNA segment is a tracrRNA. In some embodiments the tracrRNA comprises the sequence of 5'>3',

GGAACCAUUCAAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA
AAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 11).

CRISPR/Cas System Nucleic Acid-Guided Nucleases

[0183] In some embodiments, CRISPR/Cas system proteins are used in the embodiments provided herein. In some embodiments, CRISPR/Cas system proteins include proteins from CRISPR Type I systems, CRISPR Type II systems, and CRISPR Type III systems.

[0184] In some embodiments, CRISPR/Cas system proteins can be from any bacterial or archaeal species.

[0185] In some embodiments, the CRISPR/Cas system protein is isolated, recombinantly produced, or synthetic.

[0186] In some embodiments, the CRISPR/Cas system proteins are from, or are derived from CRISPR/Cas system proteins from *Streptococcus pyogenes*, *Staphylococcus aureus*, *Neisseria meningitidis*, *Streptococcus thermophiles*, *Treponema denticola*, *Francisella tularensis*, *Pasteurella multocida*, *Campylobacter jejuni*, *Campylobacter lari*, *Mycoplasma gallisepticum*, *Nitratifactor*

salsuginis, Parvibaculum lavamentivorans, Roseburia intestinalis, Neisseria cinerea, Gluconacetobacter diazotrophicus, Azospirillum, Sphaerochaeta globus, Flavobacterium columnare, Fluviicola taffensis, Bacteroides coprophilus, Mycoplasma mobile, Lactobacillus farciminis, Streptococcus pasteurianus, Lactobacillus johnsonii, Staphylococcus pseudintermedius, Filifactor alocis, Legionella pneumophila, Suterella wadsworthensis, or Corynebacter diphtheria.

[0187] In some embodiments, examples of CRISPR/Cas system proteins can be naturally occurring or engineered versions.

[0188] In some embodiments, naturally occurring CRISPR/Cas system proteins can belong to CAS Class I Type I, III, or IV, or CAS Class II Type II or V, and can include Cas9, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, Cmr5, Csf1, C2c2, and Cpf1.

[0189] In an exemplary embodiment, the CRISPR/Cas system protein comprises Cas9.

[0190] A "CRISPR/Cas system protein-gNA complex" refers to a complex comprising a CRISPR/Cas system protein and a guide NA (e.g. a gRNA or a gDNA). Where the gNA is a gRNA, the gRNA may be composed of two molecules, i.e., one RNA ("crRNA") which hybridizes to a target and provides sequence specificity, and one RNA, the "tracrRNA", which is capable of hybridizing to the crRNA. Alternatively, the guide RNA may be a single molecule (i.e., a gRNA) that contains crRNA and tracrRNA sequences.

[0191] A CRISPR/Cas system protein may be at least 60% identical (e.g., at least 70%, at least 80%, or 90% identical, at least 95% identical or at least 98% identical or at least 99% identical) to a wild type CRISPR/Cas system protein. The CRISPR/Cas system protein may have all the functions of a wild type CRISPR/Cas system protein, or only one or some of the functions, including binding activity, nuclease activity, and nuclease activity.

[0192] The term "CRISPR/Cas system protein-associated guide NA" refers to a guide NA. The CRISPR/Cas system protein -associated guide NA may exist as isolated NA, or as part of a CRISPR/Cas system protein-gNA complex.

Cas9

[0193] In some embodiments, the CRISPR/Cas System protein nucleic acid-guided nuclease is or comprises Cas9. The Cas9 of the present invention can be isolated, recombinantly produced, or synthetic.

[0194] Examples of Cas9 proteins that can be used in the embodiments herein can be found in F.A. Ran, L. Cong, W.X. Yan, D. A. Scott, J.S. Gootenberg, A.J. Kriz, B. Zetsche, O. Shalem, X. Wu, K.S. Makarova, E.V. Koonin, P.A. Sharp, and F. Zhang; "In vivo genome editing using *Staphylococcus aureus* Cas9," Nature **520**, 186–191 (09 April 2015) doi:10.1038/nature14299, which is incorporated herein by reference.

[0195] In some embodiments, the Cas9 is a Type II CRISPR system derived from *Streptococcus pyogenes*, *Staphylococcus aureus*, *Neisseria meningitidis*, *Streptococcus thermophiles*, *Treponema*

denticola, *Francisella tularensis*, *Pasteurella multocida*, *Campylobacter jejuni*, *Campylobacter lari*, *Mycoplasma gallisepticum*, *Nitratifactor salsuginis*, *Parvibaculum lavamentivorans*, *Roseburia intestinalis*, *Neisseria cinerea*, *Gluconacetobacter diazotrophicus*, *Azospirillum*, *Sphaerochaeta globus*, *Flavobacterium columnare*, *Fluviicola taffensis*, *Bacteroides coprophilus*, *Mycoplasma mobile*, *Lactobacillus farciminis*, *Streptococcus pasteurianus*, *Lactobacillus johnsonii*, *Staphylococcus pseudintermedius*, *Filifactor alocis*, *Legionella pneumophila*, *Suterella wadsworthensis*, or *Corynebacter diphtheria*.

[0196] In some embodiments, the Cas9 is a Type II CRISPR system derived from *S. pyogenes* and the PAM sequence is NGG located on the immediate 3' end of the target specific guide sequence. The PAM sequences of Type II CRISPR systems from exemplary bacterial species can also include: *Streptococcus pyogenes* (NGG), *Staph aureus* (NNGRRT), *Neisseria meningitidis* (NNNNGA TT), *Streptococcus thermophilus* (NNAGAA) and *Treponema denticola* (NAAAAC) which are all usable without deviating from the present invention.

[0197] In one exemplary embodiment, Cas9 sequence can be obtained, for example, from the pX330 plasmid (available from Addgene), re-amplified by PCR then cloned into pET30 (from EMD biosciences) to express in bacteria and purify the recombinant 6His tagged protein.

[0198] A “Cas9-gNA complex” refers to a complex comprising a Cas9 protein and a guide NA. A Cas9 protein may be at least 60% identical (e.g., at least 70%, at least 80%, or 90% identical, at least 95% identical or at least 98% identical or at least 99% identical) to a wild type Cas9 protein, e.g., to the *Streptococcus pyogenes* Cas9 protein. The Cas9 protein may have all the functions of a wild type Cas9 protein, or only one or some of the functions, including binding activity, nuclease activity, and nuclease activity.

[0199] The term “Cas9-associated guide NA” refers to a guide NA as described above. The Cas9-associated guide NA may exist isolated, or as part of a Cas9-gNA complex.

Non-CRISPR/Cas System Nucleic Acid-Guided Nucleases

[0200] In some embodiments, non-CRISPR/Cas system proteins are used in the embodiments provided herein.

[0201] In some embodiments, the non-CRISPR/Cas system proteins can be from any bacterial or archaeal species.

[0202] In some embodiments, the non-CRISPR /Cas system protein is isolated, recombinantly produced, or synthetic.

[0203] In some embodiments, the non-CRISPR /Cas system proteins are from, or are derived from *Aquifex aeolicus*, *Thermus thermophilus*, *Streptococcus pyogenes*, *Staphylococcus aureus*, *Neisseria meningitidis*, *Streptococcus thermophiles*, *Treponema denticola*, *Francisella tularensis*, *Pasteurella multocida*, *Campylobacter jejuni*, *Campylobacter lari*, *Mycoplasma gallisepticum*, *Nitratifactor*

salsuginis, Parvibaculum lavamentivorans, Roseburia intestinalis, Neisseria cinerea, Gluconacetobacter diazotrophicus, Azospirillum, Sphaerochaeta globus, Flavobacterium columnare, Fluviicola taffensis, Bacteroides coprophilus, Mycoplasma mobile, Lactobacillus farciminis, Streptococcus pasteurianus, Lactobacillus johnsonii, Staphylococcus pseudintermedius, Filifactor alocis, Legionella pneumophila, Suterella wadsworthensis, Natronobacterium gregoryi, or Corynebacter diphtheria.

[0204] In some embodiments, the non-CRISPR /Cas system proteins can be naturally occurring or engineered versions.

[0205] In some embodiments, a naturally occurring non-CRISPR /Cas system protein is NgAgo (Argonaute from *Natronobacterium gregoryi*).

[0206] A “non-CRISPR /Cas system protein-gNA complex” refers to a complex comprising a non-CRISPR /Cas system protein and a guide NA (e.g. a gRNA or a gDNA). Where the gNA is a gRNA, the gRNA may be composed of two molecules, i.e., one RNA (“crRNA”) which hybridizes to a target and provides sequence specificity, and one RNA, the “tracrRNA”, which is capable of hybridizing to the crRNA. Alternatively, the guide RNA may be a single molecule (i.e., a gRNA) that contains crRNA and tracrRNA sequences.

[0207] A non-CRISPR /Cas system protein may be at least 60% identical (e.g., at least 70%, at least 80%, or 90% identical, at least 95% identical or at least 98% identical or at least 99% identical) to a wild type non-CRISPR /Cas system protein. The non-CRISPR /Cas system protein may have all the functions of a wild type non-CRISPR /Cas system protein, or only one or some of the functions, including binding activity, nuclease activity, and nuclease activity.

[0208] The term “non-CRISPR /Cas system protein-associated guide NA” refers to a guide NA. The non-CRISPR /Cas system protein -associated guide NA may exist as isolated NA, or as part of a non-CRISPR /Cas system protein-gNA complex.

Catalytically Dead Nucleic Acid-Guided Nucleases

[0209] In some embodiments, engineered examples of nucleic acid-guided nucleases include catalytically dead nucleic acid-guided nucleases (CRISPR/Cas system nucleic acid-guided nucleases or non-CRISPR/Cas system nucleic acid-guided nucleases). The term “catalytically dead” generally refers to a nucleic acid-guided nuclease that has inactivated nucleases, for example inactivated HNH and RuvC nucleases. Such a protein can bind to a target site in any nucleic acid (where the target site is determined by the guide NA), but the protein is unable to cleave or nick the nucleic acid.

[0210] Accordingly, the catalytically dead nucleic acid-guided nuclease allows separation of the mixture into unbound nucleic acids and catalytically dead nucleic acid-guided nuclease-bound fragments. In one exemplary embodiment, a dCas9/gRNA complex binds to the targets determined by the gRNA sequence. The dCas9 bound can prevent cutting by Cas9 while other manipulations proceed.

[0211] In another embodiment, the catalytically dead nucleic acid-guided nuclease can be fused to another enzyme, such as a transposase, to target that enzyme's activity to a specific site.

[0212] In some embodiments, the catalytically dead nucleic acid-guided nuclease is dCas9, dCpf1, dCas3, dCas8a-c, dCas10, dCse1, dCsy1, dCsn2, dCas4, dCsm2, dCm5, dCsf1, dC2C2, or dNgAgo.

[0213] In one exemplary embodiment the catalytically dead nucleic acid-guided nuclease protein is a dCas9.

Nucleic Acid-Guided Nuclease Nickases

[0214] In some embodiments, engineered examples of nucleic acid-guided nucleases include nucleic acid-guided nuclease nickases (referred to interchangeably as nickase nucleic acid-guided nucleases).

[0215] In some embodiments, engineered examples of nucleic acid-guided nucleases include CRISPR/Cas system nickases or non-CRISPR/Cas system nickases, containing a single inactive catalytic domain.

[0216] In some embodiments, the nucleic acid-guided nuclease nickase is a Cas9 nickase, Cpf1 nickase, Cas3 nickase, Cas8a-c nickase, Cas10 nickase, Cse1 nickase, Csy1 nickase, Csn2 nickase, Cas4 nickase, Csm2 nickase, Cm5 nickase, Csf1 nickase, C2C2 nickase, or a NgAgo nickase.

[0217] In one embodiment, the nucleic acid-guided nuclease nickase is a Cas9 nickase.

[0218] In some embodiments, a nucleic acid-guided nuclease nickase can be used to bind to target sequence. With only one active nuclease domain, the nucleic acid-guided nuclease nickase cuts only one strand of a target DNA, creating a single-strand break or "nick". Depending on which mutant is used, the guide NA-hybridized strand or the non-hybridized strand may be cleaved. nucleic acid-guided nuclease nickases bound to 2 gNAs that target opposite strands can create a double-strand break in the nucleic acid. This "dual nickase" strategy increases the specificity of cutting because it requires that both nucleic acid-guided nuclease /gNA complexes be specifically bound at a site before a double-strand break is formed.

[0219] In exemplary embodiments, a Cas9 nickase can be used to bind to target sequence. The term "Cas9 nickase" refers to a modified version of the Cas9 protein, containing a single inactive catalytic domain, i.e., either the RuvC- or the HNH-domain. With only one active nuclease domain, the Cas9 nickase cuts only one strand of the target DNA, creating a single-strand break or "nick". Depending on which mutant is used, the guide RNA-hybridized strand or the non-hybridized strand may be cleaved. Cas9 nickases bound to 2 gRNAs that target opposite strands will create a double-strand break in the DNA. This "dual nickase" strategy can increase the specificity of cutting because it requires that both Cas9/gRNA complexes be specifically bound at a site before a double-strand break is formed.

[0220] Capture of DNA can be carried out using a nucleic acid-guided nuclease nickase. In one exemplary embodiment, a nucleic acid-guided nuclease nickase cuts a single strand of double stranded nucleic acid, wherein the double stranded region comprises methylated nucleotides.

Dissociable and Thermostable Nucleic Acid-Guided Nucleases

[0221] In some embodiments, thermostable nucleic acid-guided nucleases are used in the methods provided herein (thermostable CRISPR/Cas system nucleic acid-guided nucleases or thermostable non-CRISPR/Cas system nucleic acid-guided nucleases). In such embodiments, the reaction temperature is elevated, inducing dissociation of the protein; the reaction temperature is lowered, allowing for the generation of additional cleaved target sequences. In some embodiments, thermostable nucleic acid-guided nucleases maintain at least 50% activity, at least 55% activity, at least 60% activity, at least 65% activity, at least 70% activity, at least 75% activity, at least 80% activity, at least 85% activity, at least 90% activity, at least 95% activity, at least 96% activity, at least 97% activity, at least 98% activity, at least 99% activity, or 100% activity, when maintained for at least 75°C for at least 1 minute. In some embodiments, thermostable nucleic acid-guided nucleases maintain at least 50% activity, when maintained for at least 1 minute at least at 75°C, at least at 80°C, at least at 85°C, at least at 90°C, at least at 91°C, at least at 92°C, at least at 93°C, at least at 94°C, at least at 95°C, 96°C, at least at 97°C, at least at 98°C, at least at 99°C, or at least at 100°C. In some embodiments, thermostable nucleic acid-guided nucleases maintain at least 50% activity, when maintained at least at 75°C for at least 1 minute, 2 minutes, 3 minutes, 4 minutes, or 5 minutes. In some embodiments, a thermostable nucleic acid-guided nuclease maintains at least 50% activity when the temperature is elevated, lowered to 25°C-50°C. In some embodiments, the temperature is lowered to 25°C, to 30°C, to 35°C, to 40°C, to 45°C, or to 50°C. In one exemplary embodiment, a thermostable enzyme retains at least 90% activity after 1 min at 95°C.

[0222] In some embodiments, the thermostable nucleic acid-guided nuclease is thermostable Cas9, thermostable Cpf1, thermostable Cas3, thermostable Cas8a-c, thermostable Cas10, thermostable Cse1, thermostable Csy1, thermostable Csn2, thermostable Cas4, thermostable Csm2, thermostable Cm5, thermostable Csf1, thermostable C2C2, or thermostable NgAgo.

[0223] In some embodiments, the thermostable CRISPR/Cas system protein is thermostable Cas9.

[0224] Thermostable nucleic acid-guided nucleases can be isolated, for example, identified by sequence homology in the genome of thermophilic bacteria *Streptococcus thermophilus* and *Pyrococcus furiosus*. Nucleic acid-guided nuclease genes can then be cloned into an expression vector. In one exemplary embodiment, a thermostable Cas9 protein is isolated.

[0225] In another embodiment, a thermostable nucleic acid-guided nuclease can be obtained by *in vitro* evolution of a non-thermostable nucleic acid-guided nuclease. The sequence of a nucleic acid-guided nuclease can be mutagenized to improve its thermostability.

Methods of Making Collections of gNAs

[0226] Provided herein are methods that enable the generation of a large number of diverse gRNAs, collections of gNAs, from any source nucleic acid (e.g., DNA). Methods provided herein can employ enzymatic methods including but not limited to digestion, ligation, extension, overhang filling, transcription, reverse transcription, amplification.

[0227] Generally, the method can comprise providing a nucleic acid (e.g., DNA); employing a first enzyme (or combinations of first enzymes) that cuts at a part of the PAM sequence in the nucleic acid, in a way that a residual nucleotide sequence from the PAM sequence is left; ligating an adapter that positions a restriction enzyme typeIIS site (an enzyme that cuts outside yet near its recognition motif) at a distance to eliminate the PAM sequence; employing a second typeIIS enzyme (or combination of second enzymes) to eliminate the PAM sequence together with the adapter; and fusing a sequence that can be recognized by protein members of the nucleic acid-guided nuclease (e.g., CRISPR/Cas) system, for example, a gRNA stem-loop sequence. In some embodiments, the first enzymatic reactions cuts part of the PAM sequence in a way that residual nucleotide sequence from the PAM sequence is left, and that the nucleotide sequence immediately 5' to the PAM sequence can be any purine or pyrimidine, not just those with a cytosine 5' to the PAM sequence, for example, not just those that are C/NGG or C/TAG, etc.

[0228] Table 1 shows exemplary strategies/protocols to convert any source nucleic acid (e.g., DNA) into a collection of gNAs (e.g., gRNAs) using different restriction enzymes.

Table 1. Exemplary strategies for preparing a collection of guide nucleic acids.

CRISPR/Cas System Species	PAM Sequence	First Enzyme/Components	Strategy	3' Adapter sequence with typeIIS enzyme site (provided with only one strand sequence 5'>3')
Streptococcus pyogenes (SP); SpCas9	NGG	CviPII	Nicks immediately 5' of CCD sequence, nicks the other strand with T7 endonuclease I, blunt with T4 DNA polymerase; ligate to adapter; cut with MlyI to remove PAM and adapter; ligate gRNA stem-loop sequence at 3' end	ggGACTCggatccctatagtc (SEQ ID NO: 4421)
Staphylococcus aureus (SA); SaCas9	NNGRRT or NNGRR(N)	AlwI	Cut, blunt with T4 DNA polymerase; ligate to adapter SA; cut with EcoP15I to remove PAM and adapter; blunt end; ligate gRNA stem-loop sequence at 3' end	ttttagcggccgcctgctgCTCtaacaaagacgatgacgacaagcgt (SEQ ID NO: 4422)
Neisseria meningitidis (NM)	NNNG ATT	TfiI	Cut, blunt with T4 DNA polymerase; ligate to adapter NM; cut with EcoRI to eliminate unwanted DNA and EcoP15I to remove PAM and adapter; blunt end; ligate gRNA stem-loop sequence at 3' end	TCgcggccgctttattctgctgCTCtaacaaagacgatgacgacaagcgt (SEQ ID NO: 4428)

CRISPR/Cas System Species	PAM Sequence	First Enzyme/Components	Strategy	3' Adapter sequence with type IIS enzyme site (provided with only one strand sequence 5'>3')
Streptococcus thermophilus (ST)	NNAGAAW	BsmI	Cut, blunt with T4 DNA polymerase; ligate to adapter ST; cut with EcoP15I to remove PAM and adapter; blunt end; ligate gRNA stem-loop sequence at 3' end	ttgcggccgcttttattctgctgCTC tacaagacgatgacgacaagcgt (SEQ ID NO: 4429)
Treponema denticola (TD)	NAAAAAC	Cly7489 II	Cut, blunt with T4 DNA polymerase; ligate to adapter TD; cut with EcoP15I to remove PAM and adapter	tttagcggccgctgctgCTCtaca aagacgatgacgacaagcgt (SEQ ID NO: 4430)

[0229] Table 2 shows additional exemplary strategies/protocols to convert any source nucleic acid (e.g., DNA) into a collection of gNAs (e.g., gRNAs) using different restriction enzymes.

Table 2. Additional exemplary strategies for preparing a collection of guide nucleic acids.

CRISPR/Cas System Species	PAM Sequence	First Enzyme/Component	Exemplary Strategy	Adapter oligo sequence (with Inosine overhangs, all in 5'>3' direction)
Streptococcus pyogenes (SP); SpCas9	NGG	CviPII	Nicks immediately 5' of CCD sequence, nicks the other strand with T7 endonuclease I; ligate to adapter; cut with MlyI to remove PAM and 3' adapter; ligate gRNA stem-loop sequence at 3' end	Adapter oligo 1: ggggGACTCggatccctatagtata caaagacgatgacgacaagcg (SEQ ID NO: 4404) Adapter oligo 2: gcctcgagc*t*a*atacgactcactata gggatccaagtccc (* denotes a phosphorothioate backbone linkage) (SEQ ID NO: 4405)
Staphylococcus aureus (SA); SaCas9	NNGRRT or NNGRR(N)	AlwI	Cut; ligate to adapter SA; cut with EcoP15I to remove PAM and 3' adapter; blunt end; ligate gRNA stem-loop sequence at 3' end	Adapter oligo 1: IttttagcggccgctgctgCTCtaca agacgatgacgacaagcgt (SEQ ID NO: 4422) Adapter oligo 2: gagatcagcttctgcattgatgcGAGc agcaggcggccgctaata (SEQ ID NO: 4423)
Neisseria meningitidis (NM)	NNNNGATT	TfiI	Cut; ligate to adapter NM; cut with EcoP15I to remove PAM and 3' adapter; blunt end; ligate gRNA stem-loop sequence at 3' end	Adapter oligo 1: attTCgcgccgcttttattctgctgCT Ctacaagacgatgacgacaagcgt (SEQ ID NO: 4424) Adapter oligo 2: gagatcagcttctgcattgatgcGAGc agcagaataaaagcggccgcGA (SEQ ID NO: 4425)

CRISPR/ Cas System Species	PAM Sequence	First Enzyme/ Component	Exemplary Strategy	Adapter oligo sequence (with Inosine overhangs, all in 5'>3' direction)
Streptococcus thermophilus (ST)	NNAGAAW	BsmI	Cut; ligate to adapter ST; cut with EcoP15I to remove PAM and 3' adapter; blunt end; ligate gRNA stem-loop sequence at 3' end	Adapter oligo 1: gcggccgctttattctgctgCTCtaca agacgatgacgacaagcgt (SEQ ID NO: 4426) Adapter oligo 2: gagatcagcttctgcattgatgcGAGc agcagaataaaagcggccgcIG (SEQ ID NO: 4427)

[0230] Exemplary applications of the compositions and methods described herein are provided in **FIG. 1, FIG. 2, FIG. 3, FIG. 4, FIG. 5, FIG. 6, and FIG. 7**. The figures depict non-limiting exemplary embodiments of the present invention that includes a method of constructing a gNA library (e.g., gRNA library) from input nucleic acids (e.g., DNA), such as genomic DNA (e.g., human genomic DNA).

[0231] In **FIG. 1**, the starting material can be fragmented genomic DNA (e.g., human) or other source DNA. These fragments are blunt-ended before constructing the library **101**. T7 promoter adapters are ligated to the blunt-ended DNA fragments **102**, which is then PCR amplified. Nt.CviPII is then used to generate a nick on one strand of the PCR product immediately 5' to the CCD sequence **103**. T7 Endonuclease I cleaves on the opposite strand 1, 2, or 3 bp 5' of the nick **104**. The resulting DNA fragments are blunt-ended with T4 DNA Polymerase, leaving HGG sequence at the end of the DNA fragment **105**. The resulting DNA is cleaned and recovered on beads. An adapter carrying MlyI recognition site is ligated to the blunt-ended DNA fragment immediately 3' of HGG sequence **106**. MlyI generates a blunt-end cleavage immediately 5' to the HGG sequence, removing HGG together with the adapter sequence **107**. The resulting DNA fragments are cleaned and recovered again on beads. A gRNA stem-loop sequence is then ligated to the blunt-end cleaved by MlyI, forming a gRNA library covering the human genome **108**. This library of DNA is then PCR amplified and cleaned on beads, ready for *in vitro* transcription.

[0232] In **FIG. 2**, the starting material can intact genomic DNA (e.g., human) or other source DNA **201**. Nt.CviPII and T7 Endonuclease I are used to generate nicks on each strand of the human genomic DNA, resulting in smaller DNA fragments **202**. DNA fragments of 200-600 bp are size selected on beads, then ligated with Y-shaped adapters carrying a GG overhang on the 5'. One strand of the Y-shaped adapter contains a MlyI recognition site, wherein the other strand contains a mutated MlyI site and a T7 promoter sequence **203**. Because of these features, after PCR amplification, the T7 promoter sequence is at the distal end of the HGG sequence, and the MlyI sequence is at the rear end of HGG **204**. Digestion with MlyI generates a cleavage immediately 5' of HGG sequence **205**. MlyI generates a blunt-end cleavage immediately 5' to the HGG sequence, removing HGG together with the adapter sequence

206. A gRNA stem-loop sequence is then ligated to the blunt-end cleaved by MlyI, forming a gRNA library covering the human genome. This library of DNA is then PCR amplified and cleaned on beads, ready for *in vitro* transcription.

[0233] In FIG. 3, the source DNA (e.g., genomic DNA) can be nicked 301, for example with a nicking enzyme. In some cases, the nicking enzyme can have a recognition site that is three or fewer bases in length. In some cases, CviPII is used, which can recognize and nick at a sequence of CCD (where D represents a base other than C). Nicks can be proximal, surrounding a region containing the sequence (represented by the thicker line) which will be used to yield the guide RNA N20 sequence. When nicks are proximal, a double stranded break can occur and lead to 5' or 3' overhangs 302. These overhangs can be repaired, for example with a polymerase (e.g., T4 polymerase). In some cases, such as with 5' strands, repair can comprise synthesizing a complementary strand. In some case, such as with 3' strands, repair can comprise removing overhangs. Repair can result in a blunt end including the N20 guide sequence and a sequence complementary to the nick recognition sequence (e.g., HGG, where H represents a base other than G).

[0234] In FIG. 4, continuing for example from the end of FIG. 3, different combinations of adapters can be ligated to the DNA to allow for the desired cleaving. Adapters with a recognition site for a nuclease enzyme that cuts 3 base pairs from the site (e.g., MlyI) can be ligated 401, and digestion at that site can be used to remove a left over sequence, such as an HGG sequence 402. Adapters with a recognition site for a nuclease that cuts 20 base pairs from the site (e.g., MmeI) 403. These adapters can also include a second recognition site for a nuclease that cuts the proper number of nucleotides from the site to later remove the first recognition site (e.g., BsaXI). The first enzyme can be used to cut 20 nucleotides down, thereby keeping the N20 sequence 404. Then, a promoter adapter (e.g., T7) can be ligated next to the N20 sequence 405. Then, the nuclease corresponding to the second recognition site (e.g., BsaXI) can be used to remove the adapter for the site that cuts 20 nucleotides away (e.g., MmeI) 406. Finally, the guide RNA stem-loop sequence adapter can be ligated to the N20 sequence 407 to prepare for guide RNA production.

[0235] Alternatively, the protocol shown in FIG. 5 can follow the end of a protocol such as that shown in FIG. 3. Adapters with a recognition site for a nuclease enzyme that cleaves 25 nucleotides from the site (e.g., EcoP15I) can be ligated to the DNA 501. These adapters can also include a second recognition site for a nuclease that cuts the proper number of nucleotides from the site to later remove the first recognition site (e.g., BaeI) and any other left-over sequence, such as HGG. The enzyme corresponding to the first recognition site (e.g., EcoP15I) can then be used to cleave after the N20 sequence 502. Then, a promoter adapter (e.g., T7) can be ligated next to the N20 sequence 503. The enzyme corresponding to the second recognition site (e.g., BaeI) can then be used to remove the

recognition sites and any residual sequence (e.g., HGG) **504**. Finally, the guide RNA stem-loop sequence adapter can be ligated (e.g., by single strand ligation) to the N20 sequence **505**.

[0236] As an alternative to protocols such as that shown in **FIG. 3**, the protocol shown in **FIG. 6** can be used in preparation for protocols such as those shown in **FIG. 4** or **FIG. 5**. A nick can be introduced by a nicking enzyme (e.g., CviPII) **601**. In some cases, the nick recognition site is three or fewer bases in length. In some cases, CviPII is used, which can recognize and nick at a sequence of CCD. A polymerase (e.g., Bst large fragment DNA polymerase) can then be used to synthesize a new DNA strand starting from the nick while displacing the old strand **602**. Because of the DNA synthesis, the nick can be sealed and made available to be nicked again **603**. Subsequent cycles of nicking and synthesis can be used to yield large amounts of target sequences **604**. These single stranded copies of target sequences can be made double stranded, for example by random priming and extension. These double stranded nucleic acids comprising N20 sequences can then be further processed by methods disclosed herein, such as those shown in **FIG. 4** or **FIG. 5**.

[0237] As another alternative to protocols such as that shown in **FIG. 3** or **FIG. 6**, the protocol shown in **FIG. 7** can be used in preparation for protocols such as those shown in **FIG. 4** or **FIG. 5**. A nick can be introduced by a nicking enzyme (e.g., CviPII) **701**. In some cases, the nicking enzyme recognition site is three or fewer bases in length. In some cases, CviPII is used, which can recognize and nick at a sequence of CCD. A polymerase (e.g., Bst large fragment DNA polymerase) can then be used to synthesize a new DNA strand starting from the nick while displacing the old strand (e.g., nicking endonuclease-mediated strand-displacement DNA amplification (NEMDA)). The reaction parameters can be adjusted to control the size of the single stranded DNA produced. For example, the nickase:polymerase ratio (e.g., CviPII:Bst large fragment polymerase ratio) can be adjusted. Reaction temperature can also be adjusted. Next, an oligonucleotide can be added **704** which has (in the 5'→3' direction) a promoter (e.g., T7 promoter) **702** followed by a random n-mer (e.g., random 6-mer, random 8-mer) **703**. The random n-mer region can bind to a region of the single stranded DNA generated previously. For example, binding can be conducted by denaturing at high temperature followed by rapid cool down, which can allow the random n-mer region to bind to the single stranded DNA generated by NEMDA. In some cases, the DNA is denatured at 98 °C for 7 minutes then cooled down rapidly to 10 °C. Extension and/or amplification can be used to produce double-stranded DNA. Blunt ends can be produced, for example enzymatically (e.g., by treatment with DNA polymerase I at 20 °C). This can result in one end ending at the promoter (e.g., T7 promoter) and the other end ending at any nicking enzyme recognition sites (e.g., any CCD sites). These fragments can then be purified, for example by size selection (e.g., by gel purification, capillary electrophoresis, or other fragment separation techniques). In some cases, the target fragments are about 50 base pairs in length (adapter sequence (e.g., T7 adapter) + target N20 sequence + nicking enzyme recognition site or complement (e.g., HGG)). Fragments can then

be ligated to an adapter comprising a nuclease recognition site for a nuclease that cuts an appropriate distance away to remove the nicking enzyme recognition site 705. For example, for a three-nucleotide long nicking enzyme recognition site (e.g., CCD for CviPII), BaeI can be used. The appropriate nuclease (e.g., BaeI) can then be used to remove the nuclease recognition site and the nicking enzyme recognition site 706. The remaining nucleic acid sequence (e.g., the N20 site) can then be ligated to the final stem-loop sequence for the guide RNA 707. Amplification (e.g., PCR) can be conducted. Guide RNAs can be produced.

[0238] In some embodiments, a collection of gNAs (e.g., gRNAs) targeting human mitochondrial DNA (mtDNA) is created, that can be used for directing nucleic acid-guided nuclease (e.g., Cas9) proteins, comprising the nucleic acid-guided nuclease (e.g., Cas9) target sequence. In some embodiments, the targeting sequence of this collection of gNAs (e.g., gRNAs) are encoded by DNA sequences comprising at least the 20 nt sequence provided in the second column from the right of Table 3 (if the NGG sequence is on positive strand) and Table 4 (if the NGG sequence is on negative strand). In some embodiments, a collection of gRNA nucleic acids, as provided herein, with specificity for human mitochondrial DNA, comprise a plurality of members, wherein the members comprise a plurality of targeting sequences provided in the second column from the right column of Table 3 and/or the second column from the right of Table 4.

Table 3. gRNA target sequence for human mtDNA carrying NGG sequence on the (+) strand.

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing gRNA target sequence followed by NGG	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
13	35	ATCACCCCTATTAACCACTCACGG	13	ATCACCCCTATTAACCACTCA	436
14	36	TCACCCTATTAACCAC TCACGGG	14	TCACCCTATTAACCACTCAC	437
32	54	ACGGGAGCTCTCCATG CATTG	15	ACGGGAGCTCTCCATGCATT	438
45	67	ATGCATTTGGTATTTTC GTCTGG	16	ATGCATTTGGTATTTTCGTC	439
46	68	TGCATTTGGTATTTTCG TCTGGG	17	TGCATTTGGTATTTTCGTCT	440
47	69	GCATTTGGTATTTTCGT CTGGGG	18	GCATTTGGTATTTTCGTCTG	441
48	70	CATTTGGTATTTTCGTC TGGGGG	19	CATTTGGTATTTTCGTCTGG	442
49	71	ATTTGGTATTTTCGTCT GGGGGG	20	ATTTGGTATTTTCGTC TGGG	443
79	101	GCGATAGCATTGCGAG ACGCTGG	21	GCGATAGCATTGCGA GACGC	444
85	107	GCATTGCGAGACGCTG GAGCCGG	22	GCATTGCGAGACGCTGGAGC	445
163	185	GCACCTACGTTCAATA TTACAGG	23	GCACCTACGTTCAATATTAC	446
207	229	GTTAATTAATTAATGC TTGTAGG	24	GTTAATTAATTAATGCTTGT	447
301	323	AACCCCCCTCCCCG CTTCTGG	25	AACCCCCCTCCCCGCTTC	448
388	410	AGATTTCAAATTTTAT CTTTTGG	26	AGATTTCAAATTTTATCTTT	449
391	413	TTTCAAATTTTATCTTT TGGCGG	27	TTTCAAATTTTATCTTTTGG	450
604	626	ATACACTGAAAATGTT TAGACGG	28	ATACACTGAAAATGTTTAGA	451
605	627	TACACTGAAAATGTTT AGACGGG	29	TACACTGAAAATGTTTAGAC	452
631	653	ACATCACCCCATAAAC AAATAGG	30	ACATCACCCCATAAA CAAAT	453
636	658	ACCCCATAAACAAATA GGTTTGG	31	ACCCCATAAACAAATAGGTT	454
727	749	TCTAAATCACCACGAT CAAAAGG	32	TCTAAATCACCACGATCAAA	455
788	810	TTAGCCTAGCCACACC CCCACGG	33	TTAGCCTAGCCACACCCCCA	456
789	811	TAGCCTAGCCACACCC CCACGGG	34	TAGCCTAGCCACACCCCCAC	457
851	873	AACTAAGCTATACTAA CCCCAGG	35	AACTAAGCTATACTAACCCC	458

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing gRNA target sequence followed by NGG	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
852	874	ACTAAGCTATACTAAC CCCAGGG	36	ACTAAGCTATACTAA CCCCA	459
856	878	AGCTATACTAACCCCA GGGTTGG	37	AGCTATACTAACCCC AGGGT	460
880	902	CAATTTTCGTGCCAGCC ACCGCGG	38	CAATTTTCGTGCCAGC CACCG	461
912	934	TAACCCCAAGTCAATAG AAGCCGG	39	TAACCCCAAGTCAATA GAAGC	462
1009	1031	CACAAAATAGACTACG AAAGTGG	40	CACAAAATAGACTAC GAAAG	463
1051	1073	ACAATAGCTAAGACCC AAACTGG	41	ACAATAGCTAAGACC CAAAC	464
1052	1074	CAATAGCTAAGACCCA AACTGGG	42	CAATAGCTAAGACCC AAACT	465
1148	1170	AGCCACAGCTTAAAAC TCAAAGG	43	AGCCACAGCTTAAAA CTCAA	466
1154	1176	AGCTTAAAAC TCAAAG GACCTGG	44	AGCTTAAAAC TCAAA GGACC	467
1157	1179	TTAAAAC TCAAAGGAC CTGGCGG	45	TTAAAAC TCAAAGGA CCTGG	468
1178	1200	GGTGCTTCATATCCCT CTAGAGG	46	GGTGCTTCATATCCCT CTAG	469
1267	1289	TCTTCAGCAAACCCTG ATGAAGG	47	TCTTCAGCAAACCCT GATGA	470
1306	1328	AGTACCCACGTAAAGA CGTTAGG	48	AGTACCCACGTAAAG ACGTT	471
1312	1334	CACGTAAAGACGTTAG GTCAAGG	49	CACGTAAAGACGTTA GGTCA	472
1326	1348	AGGTCAAGGTGTAGCC CATGAGG	50	AGGTCAAGGTGTAGC CCATG	473
1329	1351	TCAAGGTGTAGCCCAT GAGGTGG	51	TCAAGGTGTAGCCCA TGAGG	474
1339	1361	GCCCATGAGGTGGCAA GAAATGG	52	GCCCATGAGGTGGCA AGAAA	475
1340	1362	CCCATGAGGTGGCAAG AAATGGG	53	CCCATGAGGTGGCAA GAAAT	476
1389	1411	GATAGCCCTTATGAAA CTTAAGG	54	GATAGCCCTTATGAA ACTTA	477
1390	1412	ATAGCCCTTATGAAAC TTAAGGG	55	ATAGCCCTTATGAAA CTTAA	478
1397	1419	TTATGAAACTTAAGGG TCGAAGG	56	TTATGAAACTTAAGG GTCGA	479
1400	1422	TGAAACTTAAGGGTCG AAGGTGG	57	TGAAACTTAAGGGTC GAAGG	480
1441	1463	AGTAGAGTGCTTAGTT GAACAGG	58	AGTAGAGTGCTTAGT TGAAC	481
1442	1464	GTAGAGTGCTTAGTTG	59	GTAGAGTGCTTAGTT	482

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing gRNA target sequence followed by NGG	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		AACAGGG		GAACA	
1494	1516	CCTCCTCAAGTATACT TCAAAGG	60	CCTCCTCAAGTATAC TTCAA	483
1530	1552	ACCCCTACGCATTTAT ATAGAGG	61	ACCCCTACGCATTTA TATAG	484
1548	1570	AGAGGAGACAAGTCGT AACATGG	62	AGAGGAGACAAGTCG TAACA	485
1560	1582	TCGTAACATGGTAAGT GTACTGG	63	TCGTAACATGGTAAG TGTAC	486
1573	1595	AGTGTACTGGAAAGTG CACTTGG	64	AGTGTACTGGAAAGT GCACT	487
1620	1642	AAAGCACCCAACTTAC ACTTAGG	65	AAAGCACCCAACTTA CACTT	488
1726	1748	CATTTACCCAAATAAA GTATAGG	66	CATTTACCCAAATAA AGTAT	489
1746	1768	AGGCGATAGAAATTGA AACCTGG	67	AGGCGATAGAAATTG AAACC	490
1770	1792	GCAATAGATATAGTAC CGCAAGG	68	GCAATAGATATAGTA CCGCA	491
1771	1793	CAATAGATATAGTACC GCAAGGG	69	CAATAGATATAGTAC CGCAA	492
1809	1831	TAACCAAGCATAATAT AGCAAGG	70	TAACCAAGCATAATA TAGCA	493
1862	1884	TAACTAGAAATAACTT TGCAAGG	71	TAACTAGAAATAACT TTGCA	494
1947	1969	CCGTCTATGTAGCAAA ATAGTGG	72	CCGTCTATGTAGCAA AATAG	495
1948	1970	CGTCTATGTAGCAAAA TAGTGGG	73	CGTCTATGTAGCAAA ATAGT	496
1960	1982	AAAATAGTGGGAAGAT TTATAGG	74	AAAATAGTGGGAAGA TTTAT	497
1966	1988	GTGGGAAGATTTATAG GTAGAGG	75	GTGGGAAGATTTATA GGTAG	498
1987	2009	GGCGACAAACCTACCG AGCCTGG	76	GGCGACAAACCTACC GAGCC	499
1997	2019	CTACCGAGCCTGGTGA TAGCTGG	77	CTACCGAGCCTGGTG ATAGC	500
2086	2108	ATTTAACTGTTAGTCC AAAGAGG	78	ATTTAACTGTTAGTCC AAAG	501
2099	2121	TCCAAAGAGGAACAGC TCTTTGG	79	TCCAAAGAGGAACAG CTCTT	502
2107	2129	GGAACAGCTCTTTGGA CACTAGG	80	GGAACAGCTCTTTGG ACACT	503
2152	2174	AAAAATTTAACACCCA TAGTAGG	81	AAAAATTTAACACCC ATAGT	504
2247	2269	CTGAACTCCTCACACC CAATTGG	82	CTGAACTCCTCACAC CCAAT	505

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing gRNA target sequence followed by NGG	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
2414	2436	CCTCACTGTCAACCCA ACACAGG	83	CCTCACTGTCAACCC AACAC	506
2427	2449	CCAACACAGGCATGCT CATAAGG	84	CCAACACAGGCATGC TCATA	507
2432	2454	ACAGGCATGCTCATAA GGAAAGG	85	ACAGGCATGCTCATA AGGAA	508
2449	2471	GAAAGGTTAAAAAAA GTAAAAGG	86	GAAAGGTTAAAAAAA GTAAA	509
2456	2478	TAAAAAAAGTAAAAG GAACTCGG	87	TAAAAAAAGTAAAAG GAACT	510
2515	2537	TCTAGCATCACCAGTA TTAGAGG	88	TCTAGCATCACCAGT ATTAG	511
2546	2568	GCCCAGTGACACATGT TTAACGG	89	GCCCAGTGACACATG TTTAA	512
2552	2574	TGACACATGTTTAAACG GCCGCGG	90	TGACACATGTTTAAAC GGCCG	513
2571	2593	GCGGTACCCTAACCGT GCAAAGG	91	GCGGTACCCTAACCG TGCAA	514
2599	2621	TAATCACTTGTTTCCTTA AATAGG	92	TAATCACTTGTTTCCTT AAAT	515
2600	2622	AATCACTTGTTTCCTTA AATAGGG	93	AATCACTTGTTTCCTTA AATA	516
2614	2636	TAAATAGGGACCTGTA TGAATGG	94	TAAATAGGGACCTGT ATGAA	517
2624	2646	CCTGTATGAATGGCTC CACGAGG	95	CCTGTATGAATGGCT CCACG	518
2625	2647	CTGTATGAATGGCTCC ACGAGGG	96	CTGTATGAATGGCTC CACGA	519
2676	2698	AAATTGACCTGCCCCGT GAAGAGG	97	AAATTGACCTGCCCCG TGAAG	520
2679	2701	TTGACCTGCCCCGTGAA GAGGCGG	98	TTGACCTGCCCCGTGA AGAGG	521
2680	2702	TGACCTGCCCCGTGAAG AGGCGGG	99	TGACCTGCCCCGTGAA GAGGC	522
2711	2733	AGCAAGACGAGAAGA CCCTATGG	100	AGCAAGACGAGAAG ACCCTA	523
2755	2777	ACAGTACCTAACAAAC CCACAGG	101	ACAGTACCTAACAAA CCCAC	524
2789	2811	CAAACCTGCATTAAAA ATTTCGG	102	CAAACCTGCATTAAA AATTT	525
2793	2815	CCTGCATTAAAAATTT CGGTTGG	103	CCTGCATTAAAAATT TCGGT	526
2794	2816	CTGCATTAAAAATTT GGTTGGG	104	CTGCATTAAAAATTT CGGTT	527
2795	2817	TGCATTAAAAATTTTCG GTTGGGG	105	TGCATTAAAAATTTTC GGTTG	528
2804	2826	AATTTTCGGTTGGGGCG	106	AATTTTCGGTTGGGGC	529

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing gRNA target sequence followed by NGG	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		ACCTCGG		GACCT	
2895	2917	TGATCCAATAACTTGA CCAACGG	107	TGATCCAATAACTTG ACCAA	530
2911	2933	CCAACGGAACAAGTTA CCCTAGG	108	CCAACGGAACAAGTT ACCCT	531
2912	2934	CAACGGAACAAGTTAC CCTAGGG	109	CAACGGAACAAGTTA CCCTA	532
2954	2976	CTAGAGTCCATATCAA CAATAGG	110	CTAGAGTCCATATCA ACAAT	533
2955	2977	TAGAGTCCATATCAAC AATAGGG	111	TAGAGTCCATATCAA CAATA	534
2974	2996	AGGGTTTACGACCTCG ATGTTGG	112	AGGGTTTACGACCTC GATGT	535
2980	3002	TACGACCTCGATGTTG GATCAGG	113	TACGACCTCGATGTT GGATC	536
2992	3014	GTTGGATCAGGACATC CCGATGG	114	GTTGGATCAGGACAT CCCGA	537
3010	3032	GATGGTGCAGCCGCTA TTAAAGG	115	GATGGTGCAGCCGCT ATTAA	538
3058	3080	TACGTGATCTGAGTTC AGACCGG	116	TACGTGATCTGAGTT CAGAC	539
3069	3091	AGTTCAGACCGGAGTA ATCCAGG	117	AGTTCAGACCGGAGT AATCC	540
3073	3095	CAGACCGGAGTAATCC AGGTCGG	118	CAGACCGGAGTAATC CAGGT	541
3110	3132	CAAATTCCTCCCTGTA CGAAAGG	119	CAAATTCCTCCCTGT ACGAA	542
3125	3147	ACGAAAGGACAAGAG AAATAAGG	120	ACGAAAGGACAAGA GAAATA	543
3203	3225	ACCCACACCCACCCAA GAACAGG	121	ACCCACACCCACCCA AGAAC	544
3204	3226	CCCACACCCACCCAA AACAGGG	122	CCCACACCCACCCAA GAACA	545
3217	3239	AAGAACAGGGTTTGT AAGATGG	123	AAGAACAGGGTTTGT TAAGA	546
3227	3249	TTTGTTAAGATGGCAG AGCCCGG	124	TTTGTTAAGATGGCA GAGCC	547
3262	3284	ACTTAAACTTTACAG TCAGAGG	125	ACTTAAACTTTACA GTCAG	548
3294	3316	TCTTCTTAACAACATA CCCATGG	126	TCTTCTTAACAACAT ACCCA	549
3336	3358	TGTACCCATTCTAATC GCAATGG	127	TGTACCCATTCTAATC GCAA	550
3370	3392	CTTACCGAACGAAAA TTCTAGG	128	CTTACCGAACGAAAA ATTCT	551
3391	3413	GGCTATATACAAC TACGCAAAGG	129	GGCTATATACAAC TACGCAA	552

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing gRNA target sequence followed by NGG	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
3406	3428	CGCAAAGGCCCAACG TTGTAGG	130	CGCAAAGGCCCAAC GTTGT	553
3415	3437	CCCAACGTTGTAGGCC CCTACGG	131	CCCAACGTTGTAGGC CCCTA	554
3416	3438	CCAACGTTGTAGGCCC CTACGGG	132	CCAACGTTGTAGGCC CCTAC	555
3570	3592	CCTCCCCATACCCAAC CCCCTGG	133	CCTCCCCATACCCAA CCCCC	556
3586	3608	CCCCTGGTCAACCTCA ACCTAGG	134	CCCCTGGTCAACCTC AACCT	557
3643	3665	GTTTACTCAATCCTCTG ATCAGG	135	GTTTACTCAATCCTCT GATC	558
3644	3666	TTTACTCAATCCTCTGA TCAGGG	136	TTTACTCAATCCTCTG ATCA	559
3676	3698	AACTCAAACCTACGCCC TGATCGG	137	AACTCAAACCTACGCC CTGAT	560
3757	3779	CTATCAACATTACTAA TAAGTGG	138	CTATCAACATTACTA ATAAG	561
3828	3850	ACTCCTGCCATCATGA CCCTTGG	139	ACTCCTGCCATCATG ACCCT	562
3892	3914	ACCCCCTTCGACCTTG CCGAAGG	140	ACCCCCTTCGACCTT GCCGA	563
3893	3915	CCCCCTTCGACCTTGC CGAAGGG	141	CCCCCTTCGACCTTGC CGAA	564
3894	3916	CCCCTTCGACCTTGCC GAAGGGG	142	CCCCTTCGACCTTGCC GAAG	565
3913	3935	GGGGAGTCCGAAC TAG TCTCAGG	143	GGGGAGTCCGAAC TA GTCTC	566
3937	3959	TTCAACATCGAATACG CCGCAGG	144	TTCAACATCGAATAC GCCGC	567
4015	4037	CTCACC ACTACAATCT TCCTAGG	145	CTCACC ACTACAATC TTCCT	568
4287	4309	ACTTTGATAGAGTAAA TAATAGG	146	ACTTTGATAGAGTAA ATAAT	569
4311	4333	GCTTAAACCCCCTTAT TTCTAGG	147	GCTTAAACCCCCTTA TTTCT	570
4386	4408	TCACACCCCATCCTAA AGTAAGG	148	TCACACCCCATCCTA AAGTA	571
4406	4428	AGGTCAGCTAAATAAG CTATCGG	149	AGGTCAGCTAAATAA GCTAT	572
4407	4429	GGTCAGCTAAATAAGC TATCGGG	150	GGTCAGCTAAATAAG CTATC	573
4428	4450	GGCCCATACCCCGAAA ATGTTGG	151	GGCCCATACCCCGAA AATGT	574
4460	4482	TCCCGTACTAATTAAT CCCCTGG	152	TCCCGTACTAATTA TCCCC	575
4494	4516	ATCTACTCTACCATCTT	153	ATCTACTCTACCATCT	576

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing gRNA target sequence followed by NGG	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		TGCAGG		TTGC	
4542	4564	CACTGATTTTTTACCTGAGTAGG	154	CACTGATTTTTTACCTGAGT	577
4692	4714	CTCTTCAACAATATAC TCTCCGG	155	CTCTTCAACAATATA CTCTC	578
4767	4789	ATAGCTATAGCAATAA AACTAGG	156	ATAGCTATAGCAATA AAAC	579
4799	4821	CTTTCACCTTCTGAGTCC CAGAGG	157	CTTTCACCTTCTGAGTC CCAG	580
4809	4831	TGAGTCCCAGAGGTTA CCCAAGG	158	TGAGTCCCAGAGGTT ACCCA	581
4827	4849	CAAGGCACCCCTCTGA CATCCGG	159	CAAGGCACCCCTCTG ACATC	582
4941	4963	TCAATCTTATCCATCAT AGCAGG	160	TCAATCTTATCCATCA TAGC	583
4950	4972	TCCATCATAGCAGGCA GTTGAGG	161	TCCATCATAGCAGGC AGTTG	584
4953	4975	ATCATAGCAGGCAGTT GAGGTGG	162	ATCATAGCAGGCAGT TGAGG	585
5010	5032	TACTCCTCAATTACCC ACATAGG	163	TACTCCTCAATTACCC ACAT	586
5202	5224	CCATCCACCCTCCTCTC CCTAGG	164	CCATCCACCCTCCTCT CCCT	587
5205	5227	TCCACCCTCCTCTCCCT AGGAGG	165	TCCACCCTCCTCTCCC TAGG	588
5223	5245	GGAGGCCTGCCCCCGC TAACCGG	166	GGAGGCCTGCCCCCG CTAAC	589
5239	5261	TAACCGGCTTTTTTGCC CAAATGG	167	TAACCGGCTTTTTTGCC CAAA	590
5240	5262	AACCGGCTTTTTTGCCC AAATGGG	168	AACCGGCTTTTTTGCC CAAAT	591
5500	5522	TAATAATCTTATAGAA ATTTAGG	169	TAATAATCTTATAGA AATTT	592
5569	5591	CTTAATTTCTGTAACA GCTAAGG	170	CTTAATTTCTGTAACA GCTA	593
5646	5668	CTAAGCCCTTACTAGA CCAATGG	171	CTAAGCCCTTACTAG ACCAA	594
5647	5669	TAAGCCCTTACTAGAC CAATGGG	172	TAAGCCCTTACTAGA CCAAT	595
5697	5719	AGCTAAGCACCTAAT CAACTGG	173	AGCTAAGCACCTAA TCAAC	596
5723	5745	CAATCTACTTCTCCCG CCGCCGG	174	CAATCTACTTCTCCCG CCGC	597
5724	5746	AATCTACTTCTCCCGC CGCCGGG	175	AATCTACTTCTCCCGC CGCC	598
5732	5754	TCTCCCGCCGCGGGGA AAAAAGG	176	TCTCCCGCCGCGGG AAAAA	599

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing gRNA target sequence followed by NGG	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
5735	5757	CCCGCCGCCGGGAAAA AAGGCGG	177	CCCGCCGCCGGGAAA AAAGG	600
5736	5758	CCGCCGCCGGGAAAA AGGCGGG	178	CCGCCGCCGGGAAAA AAGGC	601
5747	5769	AAAAAAGGCGGGAGA AGCCCCGG	179	AAAAAAGGCGGGAG AAGCCC	602
5751	5773	AAGGCGGGAGAAGCC CCGGCAGG	180	AAGGCGGGAGAAGC CCCGGC	603
5800	5822	ATTCAATATGAAAATC ACCTCGG	181	ATTCAATATGAAAAT CACCT	604
5806	5828	TATGAAAATCACCTCG GAGCTGG	182	TATGAAAATCACCTC GGAGC	605
5816	5838	ACCTCGGAGCTGGTAA AAAGAGG	183	ACCTCGGAGCTGGTA AAAAG	606
5928	5950	TCTACAAACCACAAAG ACATTGG	184	TCTACAAACCACAAA GACAT	607
5949	5971	GGAACACTATACCTAT TATTCGG	185	GGAACACTATACCTA TTATT	608
5961	5983	CTATTATTCGGCGCAT GAGCTGG	186	CTATTATTCGGCGCA TGAGC	609
5970	5992	GGCGCATGAGCTGGAG TCCTAGG	187	GGCGCATGAGCTGGA GTCCT	610
6005	6027	CCTCCTTATTCGAGCC GAGCTGG	188	CCTCCTTATTCGAGCC GAGC	611
6006	6028	CTCCTTATTCGAGCCG AGCTGGG	189	CTCCTTATTCGAGCC GAGCT	612
6027	6049	GGCCAGCCAGGCAACC TTCTAGG	190	GGCCAGCCAGGCAAC CTTCT	613
6108	6130	ATAGTAATACCCATCA TAATCGG	191	ATAGTAATACCCATC ATAAT	614
6111	6133	GTAATACCCATCATAA TCGGAGG	192	GTAATACCCATCATA ATCGG	615
6117	6139	CCCATCATAATCGGAG GCTTTGG	193	CCCATCATAATCGGA GGCTT	616
6144	6166	TGACTAGTTCCCCTAA TAATCGG	194	TGACTAGTTCCCCTA ATAAT	617
6158	6180	AATAATCGGTGCCCCC GATATGG	195	AATAATCGGTGCCCC CGATA	618
6236	6258	CCTGCTCGCATCTGCT ATAGTGG	196	CCTGCTCGCATCTGCT ATAG	619
6239	6261	GCTCGCATCTGCTATA GTGGAGG	197	GCTCGCATCTGCTAT AGTGG	620
6243	6265	GCATCTGCTATAGTGG AGGCCGG	198	GCATCTGCTATAGTG GAGGC	621
6249	6271	GCTATAGTGGAGGCCG GAGCAGG	199	GCTATAGTGGAGGCC GGAGC	622
6255	6277	GTGGAGGCCGGAGCA	200	GTGGAGGCCGGAGCA	623

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing gRNA target sequence followed by NGG	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		GGAACAGG		GGAAC	
6282	6304	ACAGTCTACCCTCCCTTAGCAGG	201	ACAGTCTACCCTCCC TTAGC	624
6283	6305	CAGTCTACCCTCCCTTAGCAGGG	202	CAGTCTACCCTCCCTTAGCA	625
6300	6322	GCAGGGAAGTACTCCCAACCCTGG	203	GCAGGGAAGTACTCC CACCC	626
6342	6364	ATCTTCTCCTTACACCTAGCAGG	204	ATCTTCTCCTTACACC TAGC	627
6360	6382	GCAGGTGTCTCCTCTATCTTAGG	205	GCAGGTGTCTCCTCT ATCTT	628
6361	6383	CAGGTGTCTCCTCTATCTTAGGG	206	CAGGTGTCTCCTCTAT CTTA	629
6362	6384	AGGTGTCTCCTCTATCTTAGGGG	207	AGGTGTCTCCTCTATC TTAG	630
6495	6517	TCTCTCCCAGTCCTAGCTGCTGG	208	TCTCTCCCAGTCCTAG CTGC	631
6552	6574	ACCACCTTCTTCGACCCCGCCGG	209	ACCACCTTCTTCGAC CCCGC	632
6555	6577	ACCTTCTTCGACCCCGCCGGAGG	210	ACCTTCTTCGACCCC GCCGG	633
6558	6580	TTCTTCGACCCCGCCGGAGGAGG	211	TTCTTCGACCCCGCC GGAGG	634
6597	6619	CAACACCTATTCTGATTTTTCGG	212	CAACACCTATTCTGA TTTTT	635
6630	6652	GTTTATATTCTTATCCTACCAGG	213	GTTTATATTCTTATCC TACC	636
6636	6658	ATTCTTATCCTACCAGGCTTCGG	214	ATTCTTATCCTACCAG GCTT	637
6669	6691	CATATTGTAAGTACTACTCCGG	215	CATATTGTAAGTACT ACTC	638
6687	6709	TCCGGAAAAAAGAAACCATTTGG	216	TCCGGAAAAAAGAA CCATT	639
6696	6718	AAAGAACCATTGATACATAGG	217	AAAGAACCATTGGA TACAT	640
6701	6723	ACCATTGATACATAGGATGG	218	ACCATTGATACAT AGGTA	641
6723	6745	GTCTGAGCTATGATATCAATTGG	219	GTCTGAGCTATGATA TCAAT	642
6732	6754	ATGATATCAATTGGCTTCCTAGG	220	ATGATATCAATTGGC TTCCT	643
6733	6755	TGATATCAATTGGCTTCCTAGGG	221	TGATATCAATTGGCT TCCTA	644
6768	6790	GCACACCATATATTTACAGTAGG	222	GCACACCATATATTT ACAGT	645
6831	6853	ATAATCATCGCTATCCCAACCGG	223	ATAATCATCGCTATC CCCAC	646

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing gRNA target sequence followed by NGG	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
6867	6889	AGCTGACTCGCCACAC TCCACGG	224	AGCTGACTCGCCACA CTCCA	647
6909	6931	GCTGCAGTGCTCTGAG CCCTAGG	225	GCTGCAGTGCTCTGA GCCCT	648
6933	6955	TTCATCTTTCTTTTCAC CGTAGG	226	TTCATCTTTCTTTTCA CCGT	649
6936	6958	ATCTTTCTTTTCACCGT AGGTGG	227	ATCTTTCTTTTCACCG TAGG	650
6945	6967	TTCACCGTAGGTGGCC TGACTGG	228	TTCACCGTAGGTGGC CTGAC	651
7032	7054	TTCCACTATGTCCTATC AATAGG	229	TTCCACTATGTCCTAT CAAT	652
7053	7075	GGAGCTGTATTTGCCA TCATAGG	230	GGAGCTGTATTTGCC ATCAT	653
7056	7078	GCTGTATTTGCCATCA TAGGAGG	231	GCTGTATTTGCCATC ATAGG	654
7086	7108	CACTGATTTCCCCTATT CTCAGG	232	CACTGATTTCCCCTAT TCTC	655
7140	7162	CATTTCACTATCATATT CATCGG	233	CATTTCACTATCATAT TCAT	656
7176	7198	TTCTTCCCACAACACTT TCTCGG	234	TTCTTCCCACAACACT TTCT	657
7185	7207	CAACACTTTCTCGGCC TATCCGG	235	CAACACTTTCTCGGC CTATC	658
7205	7227	CGGAATGCCCCGACGT TACTCGG	236	CGGAATGCCCCGACG TTACT	659
7251	7273	TGAAACATCCTATCAT CTGTAGG	237	TGAAACATCCTATCA TCTGT	660
7358	7380	AGAAGAACCCTCCATA AACCTGG	238	AGAAGAACCCTCCAT AAACC	661
7371	7393	ATAAACCTGGAGTGAC TATATGG	239	ATAAACCTGGAGTGA CTATA	662
7432	7454	ACATAAAATCTAGACA AAAAAGG	240	ACATAAAATCTAGAC AAAAA	663
7436	7458	AAAATCTAGACAAAAA AGGAAGG	241	AAAATCTAGACAAAA AAGGA	664
7457	7479	GGAATCGAACCCCCCA AAGCTGG	242	GGAATCGAACCCCCC AAAGC	665
7476	7498	CTGGTTTCAAGCCAAC CCCATGG	243	CTGGTTTCAAGCCAA CCCCCA	666
7499	7521	CCTCCATGACTTTTTCA AAAAGG	244	CCTCCATGACTTTTTC AAAA	667
7544	7566	CTTTGTCAAAGTTAAA TTATAGG	245	CTTTGTCAAAGTTAA ATTAT	668
7567	7589	CTAAATCCTATATATC TTAATGG	246	CTAAATCCTATATAT CTTAA	669
7586	7608	ATGGCACATGCAGCGC	247	ATGGCACATGCAGCG	670

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing gRNA target sequence followed by NGG	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		AAGTAGG		CAAGT	
7741	7763	TACTAACATCTCAGAC GCTCAGG	248	TACTAACATCTCAGA CGCTC	671
7831	7853	CATCCTTTACATAACA GACGAGG	249	CATCCTTTACATAAC AGACG	672
7865	7887	TCCCTTACCATCAAAT CAATTGG	250	TCCCTTACCATCAAA TCAAT	673
7875	7897	TCAAATCAATTGGCCA CCAATGG	251	TCAAATCAATTGGCC ACCAA	674
7904	7926	ACCTACGAGTACACCG ACTACGG	252	ACCTACGAGTACACC GACTA	675
7907	7929	TACGAGTACACCGACT ACGGCGG	253	TACGAGTACACCGAC TACGG	676
7955	7977	CCCCATTATTCCTAG AACCAGG	254	CCCCATTATTCCTAG AACC	677
8069	8091	TCATGAGCTGTCCCA CATTAGG	255	TCATGAGCTGTCCCC ACATT	678
8093	8115	TTAAAAACAGATGCAA TTCCCGG	256	TTAAAAACAGATGCA ATTCC	679
8131	8153	CACTTTCACCGCTACA CGACCGG	257	CACTTTCACCGCTAC ACGAC	680
8132	8154	ACTTTCACCGCTACAC GACCGGG	258	ACTTTCACCGCTACA CGACC	681
8133	8155	CTTTCACCGCTACACG ACCGGGG	259	CTTTCACCGCTACAC GACCG	682
8134	8156	TTTCACCGCTACACGA CCGGGGG	260	TTTCACCGCTACACG ACCGG	683
8144	8166	ACACGACCGGGGGTAT ACTACGG	261	ACACGACCGGGGGTA TACTA	684
8165	8187	GGTCAATGCTCTGAAA TCTGTGG	262	GGTCAATGCTCTGAA ATCTG	685
8228	8250	CCCCTAAAAATCTTTG AAATAGG	263	CCCCTAAAAATCTTT GAAAT	686
8229	8251	CCCTAAAAATCTTTGA AATAGGG	264	CCCTAAAAATCTTTG AAATA	687
8370	8392	CCCAACTAAATACTAC CGTATGG	265	CCCAACTAAATACTA CCGTA	688
8551	8573	TTCATTGCCCCCACA TCCTAGG	266	TTCATTGCCCCCACA ATCCT	689
8698	8720	ATAACCATACACAACA CTAAAGG	267	ATAACCATACACAAC ACTAA	690
8761	8783	ATTGCCACAATAACC TCCTCGG	268	ATTGCCACAATAAC CTCCT	691
8817	8839	ACTATCTATAAACCTA GCCATGG	269	ACTATCTATAAACCT AGCCA	692
8835	8857	CATGGCCATCCCCTTA TGAGCGG	270	CATGGCCATCCCCTT ATGAG	693

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing gRNA target sequence followed by NGG	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
8836	8858	ATGGCCATCCCCTTATGAGCGGG	271	ATGGCCATCCCCTTAGAGC	694
8851	8873	TGAGCGGGCACAGTGTATAGG	272	TGAGCGGGCACAGTGATTAT	695
8899	8921	CTAGCCCACTTCTTACCACAAGG	273	CTAGCCCACTTCTTACCACA	696
8973	8995	ACTCATTCAACCAATAGCCCTGG	274	ACTCATTCAACCAATAGCCC	697
9004	9026	CTAACCGCTAACATTA CTGCAGG	275	CTAACCGCTAACATTACTGC	698
9028	9050	CACCTACTCATGCACCTAATTGG	276	CACCTACTCATGCACCTAAT	699
9243	9265	CCCAGCCCATGACCCCTAACAGG	277	CCCAGCCCATGACCCCTAAC	700
9244	9266	CCAGCCCATGACCCCTAACAGGG	278	CCAGCCCATGACCCCTAACACA	701
9245	9267	CAGCCCATGACCCCTAACAGGGG	279	CAGCCCATGACCCCTAACAG	702
9273	9295	TCAGCCCTCCTAATGACCTCCGG	280	TCAGCCCTCCTAATGACCTC	703
9321	9343	TCCATAACGCTCCTCATACTAGG	281	TCCATAACGCTCCTCATACT	704
9358	9380	CACTAACCATATACCAATGATGG	282	CACTAACCATATACC AATGA	705
9390	9412	ACACGAGAAAGCACATACCAAGG	283	ACACGAGAAAGCACATACCA	706
9417	9439	CACACACCACCTGTCCAAAAAGG	284	CACACACCACCTGTCCAAAA	707
9429	9451	GTCCAAAAAGGCCTTCGATACGG	285	GTCCAAAAAGGCCTTCGATA	708
9430	9452	TCCAAAAAGGCCTTCGATACGGG	286	TCCAAAAAGGCCTTCGATAC	709
9471	9493	TCAGAAGTTTTTTTCTTCGCAGG	287	TCAGAAGTTTTTTTCTTCGC	710
9522	9544	CTAGCCCCTACCCCCCAATTAGG	288	CTAGCCCCTACCCCCCAATT	711
9525	9547	GCCCCTACCCCCCAATTAGGAGG	289	GCCCCTACCCCCCAATTAGG	712
9526	9548	CCCCTACCCCCCAATTAGGAGGG	290	CCCCTACCCCCCAATTAGGA	713
9532	9554	CCCCCAATTAGGAGGGCACTGG	291	CCCCCAATTAGGAGGGCAC	714
9543	9565	GGAGGGCACTGGCCCCCAACAGG	292	GGAGGGCACTGGCCCCCAAC	715
9606	9628	ACATCCGTATTACTCGCATCAGG	293	ACATCCGTATTACTCGCATC	716
9692	9714	ACTGCTTATTACAATTT	294	ACTGCTTATTACAATT	717

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing gRNA target sequence followed by NGG	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		TACTGG		TTAC	
9693	9715	CTGCTTATTACAATTTT ACTGGG	295	CTGCTTATTACAATTT TACT	718
9756	9778	TCTCCCTTCACCATTTC CGACGG	296	TCTCCCTTCACCATTTC CCGA	719
9765	9787	ACCATTTCCGACGGCA TCTACGG	297	ACCATTTCCGACGGC ATCTA	720
9789	9811	TCAACATTTTTTTGTAGC CACAGG	298	TCAACATTTTTTTGTAG CCAC	721
9798	9820	TTTGTAGCCACAGGCT TCCACGG	299	TTTGTAGCCACAGGC TTCCA	722
9816	9838	CACGGACTTCACGTCA TTATTGG	300	CACGGACTTCACGTC ATTAT	723
9885	9907	TTTACATCCAAACATC ACTTTGG	301	TTTACATCCAAACAT CACTT	724
9910	9932	TCGAAGCCGCCGCCTG ATACTGG	302	TCGAAGCCGCCGCCT GATAC	725
9926	9948	ATACTGGCATTTTTGT GATGTGG	303	ATACTGGCATTTTGT AGATG	726
9963	9985	TATGTCTCCATCTATTG ATGAGG	304	TATGTCTCCATCTATT GATG	727
9964	9986	ATGTCTCCATCTATTG ATGAGGG	305	ATGTCTCCATCTATTG ATGA	728
10122	10144	TTTTGACTACCACAAC TCAACGG	306	TTTTGACTACCACAA CTCAA	729
10155	10177	AAATCCACCCCTTACG AGTGCGG	307	AAATCCACCCCTTAC GAGTG	730
10343	10365	CATCATCCTAGCCCTA AGTCTGG	308	CATCATCCTAGCCCT AAGTC	731
10365	10387	GCCTATGAGTGACTAC AAAAAGG	309	GCCTATGAGTGACTA CAAAA	732
10385	10407	AGGATTAGACTGAACC GAATTGG	310	AGGATTAGACTGAAC CGAAT	733
10500	10522	GCATTTACCATCTCAC TTCTAGG	311	GCATTTACCATCTCA CTTCT	734
10551	10573	TCCTCCCTACTATGCCT AGAAGG	312	TCCTCCCTACTATGCC TAGA	735
10664	10686	CTTTGCCGCCTGCGAA GCAGCGG	313	CTTTGCCGCCTGCGA AGCAG	736
10667	10689	TGCCGCCTGCGAAGCA GCGGTGG	314	TGCCGCCTGCGAAGC AGCGG	737
10668	10690	GCCGCCTGCGAAGCAG CGGTGGG	315	GCCGCCTGCGAAGCA GCGGT	738
10704	10726	GTCTCAATCTCCAACA CATATGG	316	GTCTCAATCTCCAAC ACATA	739
10972	10994	ACTCCTACCCCTCACA ATCATGG	317	ACTCCTACCCCTCAC AATCA	740

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing gRNA target sequence followed by NGG	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
11128	11150	AACCACACTTATCCCC ACCTTGG	318	AACCACACTTATCCC CACCT	741
11147	11169	TTGGCTATCATCACC GATGAGG	319	TTGGCTATCATCACC CGATG	742
11174	11196	CAGCCAGAACGCCTGA ACGCAGG	320	CAGCCAGAACGCCTG AACGC	743
11204	11226	TTCCTATTCTACACCCT AGTAGG	321	TTCCTATTCTACACCC TAGT	744
11252	11274	ATTTACACTCACAACA CCCTAGG	322	ATTTACACTCACAAC ACCCT	745
11369	11391	ATAGTAAAGATACCTC TTTACGG	323	ATAGTAAAGATACCT CTTTA	746
11417	11439	CATGTCGAAGCCCCCA TCGCTGG	324	CATGTCGAAGCCCCC ATCGC	747
11418	11440	ATGTCGAAGCCCCCAT CGCTGGG	325	ATGTCGAAGCCCCCA TCGCT	748
11453	11475	GCCGCAGTACTCTTAA AACTAGG	326	GCCGCAGTACTCTTA AAACT	749
11456	11478	GCAGTACTCTTAAAAC TAGGCGG	327	GCAGTACTCTTAAAA CTAGG	750
11462	11484	CTCTTAAAACTAGGCG GCTATGG	328	CTCTTAAAACTAGGC GGCTA	751
11540	11562	TTCCTTGTACTATCCCT ATGAGG	329	TTCCTTGTACTATCCC TATG	752
11669	11691	CAAACCCCCTGAAGCT TCACCGG	330	CAAACCCCCTGAAGC TTCAC	753
11696	11718	GTCATTCTCATAATCG CCCACGG	331	GTCATTCTCATAATC GCCCA	754
11697	11719	TCATTCTCATAATCGC CCACGGG	332	TCATTCTCATAATCGC CCAC	755
11777	11799	CGCATCATAATCCTCT CTCAAGG	333	CGCATCATAATCCTC TCTCA	756
11866	11888	ACCCCCCACTATTAAC CTACTGG	334	ACCCCCCACTATTAA CCTAC	757
11867	11889	CCCCCACTATTAACC TACTGGG	335	CCCCCACTATTAAC CTACT	758
11927	11949	AATATCACTCTCCTAC TTACAGG	336	AATATCACTCTCCTA CTTAC	759
11985	12007	ACATATTTACCACAAC ACAATGG	337	ACATATTTACCACAA CACAA	760
11986	12008	CATATTTACCACAACA CAATGGG	338	CATATTTACCACAAC ACAAT	761
11987	12009	ATATTTACCACAACAC AATGGGG	339	ATATTTACCACAACA CAATG	762
12104	12126	CTCAACCCCGACATCA TTACCGG	340	CTCAACCCCGACATC ATTAC	763
12105	12127	TCAACCCCGACATCAT	341	TCAACCCCGACATCA	764

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing gRNA target sequence followed by NGG	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		TACCGGG		TTACC	
12164	12186	GATTGTGAATCTGACA ACAGAGG	342	GATTGTGAATCTGAC AACAG	765
12235	12257	TGCCCCCATGTCTAAC AACATGG	343	TGCCCCCATGTCTAA CAACA	766
12254	12276	ATGGCTTTCTCAACTTT TAAAGG	344	ATGGCTTTCTCAACTT TTAA	767
12272	12294	AAAGGATAACAGCTAT CCATTGG	345	AAAGGATAACAGCTA TCCAT	768
12279	12301	AACAGCTATCCATTGG TCTTAGG	346	AACAGCTATCCATTG GTCTT	769
12294	12316	GTCTTAGGCCCCAAAA ATTTTGG	347	GTCTTAGGCCCCAAA AATTT	770
12608	12630	CTGTAGCATTGTTTCGT ACATGG	348	CTGTAGCATTGTTTCG TACA	771
12742	12764	AACCTATTCCAAGTGT TCATCGG	349	AACCTATTCCAAGTG TTCAT	772
12750	12772	CCAAGTTCATCGGC TGAGAGG	350	CCAAGTTCATCGG CTGAG	773
12751	12773	CAACTGTTTCATCGGCT GAGAGGG	351	CAACTGTTTCATCGGC TGAGA	774
12757	12779	TTCATCGGCTGAGAGG GCGTAGG	352	TTCATCGGCTGAGAG GGCGT	775
12847	12869	GCAATCCTATACAACC GTATCGG	353	GCAATCCTATACAAC CGTAT	776
12856	12878	TACAACCGTATCGGCG ATATCGG	354	TACAACCGTATCGGC GATAT	777
12958	12980	CCAAGCCTCACCAC TACTAGG	355	CCAAGCCTCACCACA CTACT	778
12979	13001	GGCCTCCTCCTAGCAG CAGCAGG	356	GGCCTCCTCCTAGCA GCAGC	779
12997	13019	GCAGGCAAATCAGCCC AATTAGG	357	GCAGGCAAATCAGCC CAATT	780
13030	13052	TGACTCCCCTCAGCCA TAGAAGG	358	TGACTCCCCTCAGCC ATAGA	781
13081	13103	TCAAGCACTATAGTTG TAGCAGG	359	TCAAGCACTATAGTT GTAGC	782
13156	13178	CAAACCTAACAACAT GCTTAGG	360	CAAACCTAACAACAT TGCTT	783
13246	13268	TTCTCCAATTCAAGTC AACTAGG	361	TTCTCCAATTCAAGTC AACT	784
13267	13289	GGACTCATAATAGTTA CAATCGG	362	GGACTCATAATAGTT ACAAT	785
13345	13367	GCCATACTATTTATGT GCTCCGG	363	GCCATACTATTTATGT GCTC	786
13346	13368	CCATACTATTTATGTG CTCCGGG	364	CCATACTATTTATGTG CTCC	787

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing gRNA target sequence followed by NGG	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
13393	13415	GAACAAGATATTCGAA AAATAGG	365	GAACAAGATATTCGA AAAAT	788
13396	13418	CAAGATATTCGAAAAA TAGGAGG	366	CAAGATATTCGAAAA ATAGG	789
13441	13463	ACTTCAACCTCCCTCA CCATTGG	367	ACTTCAACCTCCCTC ACCAT	790
13459	13481	ATTGGCAGCCTAGCAT TAGCAGG	368	ATTGGCAGCCTAGCA TTAGC	791
13477	13499	GCAGGAATACCTTTCC TCACAGG	369	GCAGGAATACCTTTC CTCAC	792
13612	13634	ATAATTCTTCTCACCT AACAGG	370	ATAATTCTTCTCACCC TAAC	793
13686	13708	ACTAAACCCCATTA CGCCTGG	371	ACTAAACCCCATTA ACGCC	794
13693	13715	CCCATTAACGCCTGG CAGCCGG	372	CCCATTAACGCCTG GCAGC	795
13708	13730	GCAGCCGGAAGCCTAT TCGCAGG	373	GCAGCCGGAAGCCTA TTCGC	796
13804	13826	GCCCTCGCTGTCACTT CCTAGG	374	GCCCTCGCTGTCACT TCCT	797
13894	13916	TTTTATTTCTCCAACAT ACTCGG	375	TTTTATTTCTCCAACA TACT	798
13936	13958	CACCGCACAAATCCCCT ATCTAGG	376	CACCGCACAAATCCCC TATCT	799
14059	14081	ATCATCACCTCAACCC AAAAAGG	377	ATCATCACCTCAACC CAAAA	800
14237	14259	TACAAAGCCCCCGCAC CAATAGG	378	TACAAAGCCCCCGCA CCAAT	801
14417	14439	ACCCCTGACCCCCATG CCTCAGG	379	ACCCCTGACCCCCAT GCCTC	802
14579	14601	AATACTAAACCCCCAT AAATAGG	380	AATACTAAACCCCCA TAAAT	803
14585	14607	AAACCCCCATAAATAG GAGAAGG	381	AAACCCCCATAAATA GGAGA	804
14664	14686	CATACATCATTATTCTC GCACGG	382	CATACATCATTATTCT CGCA	805
14825	14847	ATCTCCGCATGATGAA ACTTCGG	383	ATCTCCGCATGATGA AACTT	806
14837	14859	TGAAACTTCGGCTCAC TCCTTGG	384	TGAAACTTCGGCTCA CTCCT	807
14867	14889	CTGATCCTCCAAATCA CCACAGG	385	CTGATCCTCCAAATC ACCAC	808
14951	14973	ATCACTCGAGACGTAA ATTATGG	386	ATCACTCGAGACGTA AATTA	809
14981	15003	ATCCGCTACCTTCACG CCAATGG	387	ATCCGCTACCTTCAC GCCAA	810
15020	15042	ATCTGCCTCTTCCTACA	388	ATCTGCCTCTTCCTAC	811

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing gRNA target sequence followed by NGG	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CATCGG		ACAT	
15021	15043	TCTGCCTCTTCCTACACATCGGG	389	TCTGCCTCTTCCTACACATC	812
15026	15048	CTCTTCCTACACATCGGGCGAGG	390	CTCTTCCTACACATCGGGCG	813
15038	15060	ATCGGGCGAGGCCTATTACGG	391	ATCGGGCGAGGCCTATTATTA	814
15071	15093	TACTCAGAAACCTGAAACATCGG	392	TACTCAGAAACCTGAACAT	815
15113	15135	ACTATAGCAACAGCCTTCATAGG	393	ACTATAGCAACAGCCTTCAT	816
15131	15153	ATAGGCTATGTCCTCCCGTGAGG	394	ATAGGCTATGTCCTCCCGTG	817
15149	15171	TGAGGCCAAATATCATCTGAGG	395	TGAGGCCAAATATCATTTCTG	818
15150	15172	GAGGCCAAATATCATTCTGAGGG	396	GAGGCCAAATATCATTCTGA	819
15151	15173	AGGCCAAATATCATTC TGAGGGG	397	AGGCCAAATATCATTCTGAG	820
15194	15216	CTATCCGCCATCCCATACATTGG	398	CTATCCGCCATCCCATACAT	821
15195	15217	TATCCGCCATCCCATA CATTGGG	399	TATCCGCCATCCCATACATT	822
15221	15243	GACCTAGTTCAATGAATCTGAGG	400	GACCTAGTTCAATGAATCTG	823
15224	15246	CTAGTTCAATGAATCTGAGGAGG	401	CTAGTTCAATGAATCTTGAGG	824
15334	15356	CCTCCTATTCTTGCACGAAACGG	402	CCTCCTATTCTTGCACGAAA	825
15335	15357	CTCCTATTCTTGCACGAAACGGG	403	CTCCTATTCTTGCACGAAAC	826
15353	15375	ACGGGATCAAACAACCCCTAGG	404	ACGGGATCAAACAACCCCCT	827
15416	15438	TACACAATCAAAGACGCCCTCGG	405	TACACAATCAAAGACGCCCT	828
15476	15498	CTATTCTCACCAGACCTCCTAGG	406	CTATTCTCACCAGACCTCCT	829
15590	15612	CGATCCGTCCTTAACA AACTAGG	407	CGATCCGTCCTTAACA AACT	830
15593	15615	TCCGTCCCTAACAAC TAGGAGG	408	TCCGTCCCTAACAACCTAGG	831
15740	15762	CTCCTCATTCTAACCTGAATCGG	409	CTCCTCATTCTAACCTGAAT	832
15743	15765	CTCATTCTAACCTGAATCGGAGG	410	CTCATTCTAACCTGAATCGG	833
15776	15798	AGCTACCCTTTTACCATCATTGG	411	AGCTACCCTTTTACCATCAT	834

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing gRNA target sequence followed by NGG	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
15861	15883	TTGAAAACAAAATACTCAAATGG	412	TTGAAAACAAAATAC TCAAA	835
15862	15884	TGAAAACAAAATACTCAAATGGG	413	TGAAAACAAAATACTCAAAT	836
15906	15928	AATACACCAGTCTTGTAAACCGG	414	AATACACCAGTCTTGTAAAC	837
15928	15950	GAGATGAAAACCTTTT TCCAAGG	415	GAGATGAAAACCTTT TCCA	838
16012	16034	AACTATTCTCTGTTCTTTCATGG	416	AACTATTCTCTGTTCTTCA	839
16013	16035	ACTATTCTCTGTTCTTTCATGGG	417	ACTATTCTCTGTTCTTTCAT	840
16014	16036	CTATTCTCTGTTCTTTCATGGGG	418	CTATTCTCTGTTCTTTCATG	841
16026	16048	CTTTCATGGGGAAGCAGATTG	419	CTTTCATGGGGAAGCAGATT	842
16027	16049	TTTCATGGGGAAGCAGATTGGG	420	TTTCATGGGGAAGCAGATT	843
16108	16130	CAGCCACCATGAATATGTACGG	421	CAGCCACCATGAATA TTGTA	844
16252	16274	AAAGCCACCCCTCACC CACTAGG	422	AAAGCCACCCCTCAC CCACT	845
16348	16370	CAAATCCCTTCTCGTCCCACATGG	423	CAAATCCCTTCTCGTCCCA	846
16367	16389	ATGGATGACCCCCCTCAGATAGG	424	ATGGATGACCCCCCTCAGAT	847
16368	16390	TGGATGACCCCCCTCAGATAGGG	425	TGGATGACCCCCCTCAGATA	848
16369	16391	GGATGACCCCCCTCAGATAGGGG	426	GGATGACCCCCCTCAGATAG	849
16434	16456	GAGTGCTACTCTCCTCGCTCCGG	427	GAGTGCTACTCTCCTCGCTC	850
16435	16457	AGTGCTACTCTCCTCGCTCCGGG	428	AGTGCTACTCTCCTCGCTCC	851
16449	16471	CGCTCCGGGCCCATAACACTTGG	429	CGCTCCGGGCCCATAACACT	852
16450	16472	GCTCCGGGCCCATAACACTTGGG	430	GCTCCGGGCCCATAACACTT	853
16451	16473	CTCCGGGCCCATAACA CTTGGGG	431	CTCCGGGCCCATAACA CTTG	854
16452	16474	TCCGGGCCCATAACAC CTTGGGG	432	TCCGGGCCCATAACA CTTG	855
16482	16504	AGTGAAGTGTATCCGACATCTGG	433	AGTGAAGTGTATCCGACATC	856
16495	16517	CGACATCTGGTTCCTA CTTCAGG	434	CGACATCTGGTTCCTA ACTTC	857
16496	16518	GACATCTGGTTCCTAC	435	GACATCTGGTTCCTA	858

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing gRNA target sequence followed by NGG	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		TTCAGGG		CTTCA	

Table 4. gRNA target sequence for human mtDNA carrying NGG sequence on the (-) strand.

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
17	39	CCCTATTAACCACTCA CGGGAGC	859	GCTCCCGTGAGTGGT TAATA	2628
18	40	CCTATTAACCACTCAC GGGAGCT	860	AGCTCCCGTGAGTGG TTAAT	2629
26	48	CCACTCACGGGAGCTC TCCATGC	861	GCATGGAGAGCTCCC GTGAG	2630
43	65	CCATGCATTTGGTATTT TCGTCT	862	AGACGAAAATACCAA ATGCA	2631
104	126	CCGGAGCACCTATGT CGCAGTA	863	TACTGCGACATAGGG TGCTC	2632
112	134	CCCTATGTCGCAGTAT CTGTCTT	864	AAGACAGATACTGCG ACATA	2633
113	135	CCTATGTCGCAGTATC TGTCTTT	865	AAAGACAGATACTGC GACAT	2634
140	162	CCTGCCTCATCCTATTA TTTATC	866	GATAAATAATAGGAT GAGGC	2635
144	166	CCTCATCCTATTATTTA TCGCAC	867	GTGCGATAAATAATA GGATG	2636
150	172	CCTATTATTTATCGCAC CTACGT	868	ACGTAGGTGCGATAA ATAAT	2637
166	188	CCTACGTTCAATATTA CAGGCGA	869	TCGCCTGTAATATTG AACGT	2638
261	283	CCACTTTCCACACAGA CATCATA	870	TATGATGTCTGTGTG GAAAG	2639
268	290	CCACACAGACATCATA ACAAAAA	871	TTTTTGTTATGATGTC TGTG	2640
298	320	CCAAACCCCCCTCCC CCGCTTC	872	GAAGCGGGGGAGGG GGGGTT	2641
304	326	CCCCCTCCCCGCTTC TGGCCA	873	TGGCCAGAAGCGGGG GAGGG	2642
305	327	CCCCCTCCCCGCTTCT GGCCAC	874	GTGGCCAGAAGCGGG GGAGG	2643
306	328	CCCCTCCCCCGCTTCTG GCCACA	875	TGTGGCCAGAAGCGG GGGAG	2644
307	329	CCCTCCCCCGCTTCTG GCCACAG	876	CTGTGGCCAGAAGCG GGGGA	2645
308	330	CCTCCCCCGCTTCTGG	877	GCTGTGGCCAGAAGC	2646

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CCACAGC		GGGGG	
311	333	CCCCGCTTCTGGCCA CAGCACT	878	AGTGCTGTGGCCAGA AGCGG	2647
312	334	CCCCGCTTCTGGCCAC AGCACTT	879	AAGTGCTGTGGCCAG AAGCG	2648
313	335	CCCGCTTCTGGCCACA GCACTTA	880	TAAGTGCTGTGGCCA GAAGC	2649
314	336	CCGCTTCTGGCCACAG CACTTAA	881	TTAAGTGCTGTGGCC AGAAG	2650
324	346	CCACAGCACTTAAACA CATCTCT	882	AGAGATGTGTTTAAG TGCTG	2651
348	370	CCAAACCCCAAAAACA AAGAACC	883	GGTTCTTTGTTTTTGG GGTT	2652
353	375	CCCCAAAAACAAAGA ACCCTAAC	884	GTTAGGGTTCTTTGTT TTTG	2653
354	376	CCCCAAAAACAAAGAA CCCTAACAC	885	TGTTAGGGTTCTTTGT TTTT	2654
355	377	CCAAAAACAAAGAAC CCTAACAC	886	GTGTTAGGGTTCTTTG TTTT	2655
369	391	CCCTAACACCAGCCTA ACCAGAT	887	ATCTGGTTAGGCTGG TGTTA	2656
370	392	CCTAACACCAGCCTAA CCAGATT	888	AATCTGGTTAGGCTG GTGTT	2657
377	399	CCAGCCTAACAGATT TCAAATT	889	AATTTGAAATCTGGT TAGGC	2658
381	403	CCTAACCCAGATTTCAA ATTTTAT	890	ATAAAATTTGAAATC TGGTT	2659
386	408	CCAGATTTCAAATTTT ATCTTTT	891	AAAAGATAAAATTTG AAATC	2660
433	455	CCCCCAACTAACACA TTATTTT	892	AAAATAATGTGTTAG TTGGG	2661
434	456	CCCCCAACTAACACAT TATTTTC	893	GAAAATAATGTGTTA GTTGG	2662
435	457	CCCCAACTAACACATT ATTTTCC	894	GGAAAATAATGTGTT AGTTG	2663
436	458	CCCAACTAACACATTA TTTTCCC	895	GGGAAAATAATGTGT TAGTT	2664
437	459	CCAATAACACATTAT TTTCCCC	896	GGGGAAAATAATGTG TTAGT	2665
456	478	CCCCTCCCACTCCCAT ACTACTA	897	TAGTAGTATGGGAGT GGGAG	2666
457	479	CCCTCCCACTCCCATA CTACTAA	898	TTAGTAGTATGGGAG TGGGA	2667
458	480	CCTCCCACTCCCATAC TACTAAT	899	ATTAGTAGTATGGGA GTGGG	2668

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
461	483	CCCACTCCCATACTACTAATCTC	900	GAGATTAGTAGTATGGAGT	2669
462	484	CCACTCCCATACTACTAATCTCA	901	TGAGATTAGTAGTATGGGAG	2670
467	489	CCCATACTACTAATCTCATCAAT	902	ATTGATGAGATTAGTAGTAT	2671
468	490	CCATACTACTAATCTCATCAATA	903	TATTGATGAGATTAGTAGTA	2672
494	516	CCCCCGCCCATCCTACCCAGCAC	904	GTGCTGGGTAGGATGGCGG	2673
495	517	CCCCGCCCATCCTACCAGCACA	905	TGTGCTGGGTAGGATGGGCG	2674
496	518	CCCGCCCATCCTACCCAGCACAC	906	GTGTGCTGGGTAGGATGGGC	2675
497	519	CCGCCCATCCTACCCAGCACACA	907	TGTGTGCTGGGTAGGATGGG	2676
500	522	CCCATCCTACCCAGCACACACAC	908	GTGTGTGTGCTGGGTAGGAT	2677
501	523	CCATCCTACCCAGCACACACACA	909	TGTGTGTGTGCTGGGTAGGA	2678
505	527	CCTACCCAGCACACACACACACG	910	GCGGTGTGTGTGTGTGCTGGGT	2679
509	531	CCCAGCACACACACACACACGCTGCT	911	AGCAGCGGTGTGTGTGTGCT	2680
510	532	CCAGCACACACACACACCGCTGCTA	912	TAGCAGCGGTGTGTGTGTGCT	2681
524	546	CCGCTGCTAACCCCATACCCCGA	913	TCGGGGTATGGGGTTAGCAG	2682
534	556	CCCATAACCCGAACCAACCAAAA	914	TTTGGTTGGTTCGGGTATG	2683
535	557	CCCATAACCCGAACCAACCAAAAC	915	GTTTGGTTGGTTCGGGTAT	2684
536	558	CCATAACCCGAACCAACCAAAACC	916	GGTTTGGTTGGTTCGGGTATA	2685
541	563	CCCCGAACCAACCAAAACCCAAA	917	TTTGGGGTTTGGTTGGTTCG	2686
542	564	CCCGAACCAACCAAAACCCAAAG	918	CTTTGGGGTTTGGTTGGTTC	2687
543	565	CCGAACCAACCAAAACCCAAAGA	919	TCTTTGGGGTTTGGTTGGTT	2688
548	570	CCAACCAAAACCCCAAAAGACACC	920	GGGTGTCTTTGGGGTTTGGT	2689
552	574	CCAAACCCCAAAGACACCCCA	921	TGGGGGGTGTCTTTGGGGT	2690
557	579	CCCCAAAGACACCCCA	922	AACTGTGGGGGGGTGT	2691

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CACAGTT		CTTTG	
558	580	CCCAAAGACACCCCCC ACAGTTT	923	AAACTGTGGGGGGTG TCTTT	2692
559	581	CCAAAGACACCCCCCA CAGTTTA	924	TAAACTGTGGGGGGT GTCTT	2693
568	590	CCCCCACAGTTTATG TAGCTTA	925	TAAGCTACATAAACT GTGGG	2694
569	591	CCCCCACAGTTTATGT AGCTTAC	926	GTAAGCTACATAAAC TGTGG	2695
570	592	CCCCACAGTTTATGTA GCTTACC	927	GGTAAGCTACATAAA CTGTG	2696
571	593	CCCACAGTTTATGTAG CTTACCT	928	AGGTAAGCTACATAA ACTGT	2697
572	594	CCACAGTTTATGTAGC TTACCTC	929	GAGGTAAGCTACATA AACTG	2698
591	613	CCTCCTCAAAGCAATA CACTGAA	930	TTCAGTGTATTGCTTT GAGG	2699
594	616	CCTCAAAGCAATACAC TGAAAAT	931	ATTTTCAGTGTATTGC TTTG	2700
637	659	CCCATAAACAAATAG GTTTGGT	932	ACCAAACCTATTTGT TTATG	2701
638	660	CCATAAACAAATAGG TTTGGTC	933	GACCAAACCTATTTG TTTAT	2702
639	661	CCATAAACAAATAGGT TTGGTCC	934	GGACCAAACCTATTT GTTTA	2703
660	682	CCTAGCCTTTCTATTAG CTCTTA	935	TAAGAGCTAATAGAA AGGCT	2704
665	687	CCTTTCTATTAGCTCTT AGTAAG	936	CTTACTAAGAGCTAA TAGAA	2705
705	727	CCCGTTCCAGTGAGT TCACCCT	937	AGGGTGAACCTCACTG GAACG	2706
706	728	CCCGTTCCAGTGAGTT CACCTC	938	GAGGGTGAACCTCACT GGAAC	2707
707	729	CCGTTCCAGTGAGTTC ACCCTCT	939	AGAGGGTGAACCTCAC TGGAA	2708
712	734	CCAGTGAGTTCACCCT CTAAATC	940	GATTTAGAGGGTGAA CTCAC	2709
724	746	CCCTCTAAATCACCAC GATCAAA	941	TTTGATCGTGGTGATT TAGA	2710
725	747	CCTCTAAATCACCACG ATCAAAA	942	TTTTGATCGTGGTGAT TTAG	2711
736	758	CCACGATCAAAAGGAA CAAGCAT	943	ATGCTTGTTCTTTTG ATCG	2712
792	814	CCTAGCCACACCCCCA CGGGAAA	944	TTTCCCGTGGGGGGTG TGGCT	2713

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
797	819	CCACACCCCCACGGGA AACAGCA	945	TGCTGTTTCCCGTGG GGGTG	2714
802	824	CCCCACGGGAAACAG CAGTGAT	946	ATCACTGCTGTTTCCC GTGG	2715
803	825	CCCCACGGGAAACAGC AGTGATT	947	AATCACTGCTGTTTCC CGTG	2716
804	826	CCCACGGGAAACAGCA GTGATTA	948	TAATCACTGCTGTTTC CCGT	2717
805	827	CCACGGGAAACAGCA GTGATTAA	949	TTAATCACTGCTGTTT CCCG	2718
828	850	CCTTTAGCAATAAACG AAAGTTT	950	AAACTTTCGTTTATTG CTAA	2719
867	889	CCCCAGGGTTGGTCAA TTTCGTG	951	CACGAAATTGACCAA CCCTG	2720
868	890	CCCAGGGTTGGTCAAT TTCGTGC	952	GCACGAAATTGACCA ACCCT	2721
869	891	CCAGGGTTGGTCAATT TCGTGCC	953	GGCACGAAATTGACC AACCC	2722
890	912	CCAGCCACCGCGGTCA CACGATT	954	AATCGTGTGACCGCG GTGGC	2723
894	916	CCACCGCGGTCACACG ATTAACC	955	GGTTAATCGTGTGAC CGCGG	2724
897	919	CCGCGGTCACACGATT AACCCAA	956	TTGGGTTAATCGTGT GACCG	2725
915	937	CCCAAGTCAATAGAAG CCGGCGT	957	ACGCCGGCTTCTATT GACTT	2726
916	938	CCAAGTCAATAGAAGC CGGCGTA	958	TACGCCGGCTTCTATT GACT	2727
931	953	CCGGCGTAAAGAGTGT TTTAGAT	959	ATCTAAAACACTCTT TACGC	2728
956	978	CCCCTCCCCAATAAAA GCTAAAA	960	TTTAGCTTTATTGGG GAGG	2729
957	979	CCCCTCCCCAATAAAG CTAAAC	961	GTTTAGCTTTATTGG GGAG	2730
958	980	CCCTCCCCAATAAAGC TAAACT	962	AGTTTAGCTTTATTG GGGA	2731
959	981	CCTCCCCAATAAAGCT AAAACCTC	963	GAGTTTAGCTTTATT GGGG	2732
962	984	CCCCAATAAAGCTAAA ACTCACC	964	GGTGAGTTTAGCTTT ATTG	2733
963	985	CCAATAAAGCTAAAA CTCACCT	965	AGGTGAGTTTAGCT TTATT	2734
964	986	CCAATAAAGCTAAAAC TCACCTG	966	CAGGTGAGTTTAGC TTTAT	2735
983	1005	CCTGAGTTGTAAAAAA	967	ACTGGAGTTTTTTAC	2736

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CTCCAGT		AACTC	
1001	1023	CCAGTTGACACAAAATAGACTAC	968	GTAGTCTATTTTGTGTCAAC	2737
1064	1086	CCCAAAGTGGGATTAGATACCCC	969	GGGGTATCTAATCCCAGTTT	2738
1065	1087	CCAAAGTGGGATTAGATACCCCA	970	TGGGGTATCTAATCCAGTT	2739
1083	1105	CCCCACTATGCTTAGCCCTAAAC	971	GTTTAGGGCTAAGCATAGTG	2740
1084	1106	CCCACTATGCTTAGCCCTAAACC	972	GGTTTAGGGCTAAGCATAGT	2741
1085	1107	CCACTATGCTTAGCCCTAAACCT	973	AGGTTTAGGGCTAAGCATAG	2742
1098	1120	CCCTAAACCTCAACAGTTAAATC	974	GATTAACTGTTGAGGTTTA	2743
1099	1121	CCTAAACCTCAACAGTTAAATCA	975	TGATTAACTGTTGAGGTTT	2744
1105	1127	CCTCAACAGTTAAATCAACAAAA	976	TTTTGTTGATTAACTGTTG	2745
1135	1157	CCAGAACAACACTACGAGCCACAGCT	977	AGCTGTGGCTCGTAGTGTTT	2746
1150	1172	CCACAGCTTAAAGGACAAAGGAC	978	GTCCTTTGAGTTTAAAGCTG	2747
1172	1194	CCTGGCGGTGCTTCATATCCCTC	979	GAGGGATATGAAGCAACCGCC	2748
1190	1212	CCCTCTAGAGGAGCCTGTTCTGT	980	ACAGAACAGGCTCCTCTAGA	2749
1191	1213	CCTCTAGAGGAGCCTGTTCTGTA	981	TACAGAACAGGCTCCTCTAG	2750
1203	1225	CCTGTTCTGTAATCGATAAACCC	982	GGGTTTATCGATTACAGAAC	2751
1223	1245	CCCCGATCAACCTCACACCTCT	983	AGAGGTGGTGAGGTTGATCG	2752
1224	1246	CCCGATCAACCTCACACCTCTT	984	AAGAGGTGGTGAGGTTGATC	2753
1225	1247	CCGATCAACCTCACACCTCTTG	985	CAAGAGGTGGTGAGGTTGAT	2754
1233	1255	CCTCACCACCTCTTGCTCAGCCT	986	AGGCTGAGCAAGAGGTGGTG	2755
1238	1260	CCACCTCTTGCTCAGCCTATATA	987	TATATAGGCTGAGCAAGAGG	2756
1241	1263	CCTCTTGCTCAGCCTATATACCG	988	CGGTATATAGGCTGAGCAAG	2757
1253	1275	CCTATATACCGCCATCTTCAGCA	989	TGCTGAAGATGGCGGTATAT	2758

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
1261	1283	CCGCCATCTTCAGCAA ACCCTGA	990	TCAGGGTTTGCTGAA GATGG	2759
1264	1286	CCATCTTCAGCAAACC CTGATGA	991	TCATCAGGGTTTGCT GAAGA	2760
1278	1300	CCCTGATGAAGGCTAC AAAGTAA	992	TTACTTTGTAGCCTTC ATCA	2761
1279	1301	CCTGATGAAGGCTACA AAGTAAG	993	CTTACTTTGTAGCCTT CATC	2762
1310	1332	CCCACGTAAAGACGTT AGGTCAA	994	TTGACCTAACGTCTTT ACGT	2763
1311	1333	CCACGTAAAGACGTTA GGTCAAG	995	CTTGACCTAACGTCTT TACG	2764
1340	1362	CCCATGAGGTGGCAAG AAATGGG	996	CCCATTCTTGCCACC TCAT	2765
1341	1363	CCATGAGGTGGCAAGA AATGGGC	997	GCCCATTTCTTGCCAC CTCA	2766
1375	1397	CCCCAGAAAACACTACGA TAGCCCT	998	AGGGCTATCGTAGTT TTCTG	2767
1376	1398	CCCAGAAAACACTACGAT AGCCCTT	999	AAGGGCTATCGTAGT TTTCT	2768
1377	1399	CCAGAAAACACTACGATA GCCCTTA	1000	TAAGGGCTATCGTAG TTTTC	2769
1394	1416	CCCTTATGAAACTTAA GGGTCGA	1001	TCGACCCTTAAGTTTC ATAA	2770
1395	1417	CCTTATGAAACTTAAAG GGTCGAA	1002	TTCGACCCTTAAGTTT CATA	2771
1465	1487	CCCTGAAGCGCGTACA CACCGCC	1003	GGCGGTGTGTACGCG CTTCA	2772
1466	1488	CCTGAAGCGCGTACAC ACCGCCC	1004	GGGCGGTGTGTACGC GCTTC	2773
1483	1505	CCGCCCGTCACCCTCC TCAAGTA	1005	TACTTGAGGAGGGTG ACGGG	2774
1486	1508	CCCGTCACCCTCCTCA AGTATAC	1006	GTATACTTGAGGAGG GTGAC	2775
1487	1509	CCGTCACCCTCCTCAA GTATACT	1007	AGTATACTTGAGGAG GGTGA	2776
1493	1515	CCCTCCTCAAGTATAC TTCAAAG	1008	CTTTGAAGTATACTT GAGGA	2777
1494	1516	CCTCCTCAAGTATACT TCAAAGG	1009	CCTTTGAAGTATACTT GAGG	2778
1497	1519	CCTCAAGTATACTTCA AAGGACA	1010	TGTCCTTTGAAGTAT ACTTG	2779
1531	1553	CCCCTACGCATTTATA TAGAGGA	1011	TCCTCTATATAAATG CGTAG	2780
1532	1554	CCCTACGCATTTATAT	1012	CTCCTCTATATAAAT	2781

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		AGAGGAG		GCGTA	
1533	1555	CCTACGCATTTATATA GAGGAGA	1013	TCTCCTCTATATAAAT GCGT	2782
1601	1623	CCAGAGTGTAGCTTAA CACAAAG	1014	CTTTGTGTTAAGCTAC ACTC	2783
1626	1648	CCCAACTTACACTTAG GAGATTT	1015	AAATCTCCTAAGTGT AAGTT	2784
1627	1649	CCAACTTACACTTAGG AGATTTC	1016	GAAATCTCCTAAGTG TAAGT	2785
1662	1684	CCGCTCTGAGCTAAAC CTAGCCC	1017	GGGCTAGGTTTAGCT CAGAG	2786
1677	1699	CCTAGCCCCAAACCCA CTCCACC	1018	GGTGGAGTGGGTTTG GGGCT	2787
1682	1704	CCCCAAACCCACTCCA CCTTACT	1019	AGTAAGGTGGAGTGG GTTTG	2788
1683	1705	CCCAAACCCACTCCAC CTTACTA	1020	TAGTAAGGTGGAGTG GGTTT	2789
1684	1706	CCAAACCCACTCCACC TTACTAC	1021	GTAGTAAGGTGGAGT GGGTT	2790
1689	1711	CCCACTCCACCTTACT ACCAGAC	1022	GTCTGGTAGTAAGGT GGAGT	2791
1690	1712	CCACTCCACCTTACTA CCAGACA	1023	TGTCTGGTAGTAAGG TGGAG	2792
1695	1717	CCACCTTACTACCAGA CAACCTT	1024	AAGGTTGTCTGGTAG TAAGG	2793
1698	1720	CCTTACTACCAGACAA CCTTAGC	1025	GCTAAGGTTGTCTGG TAGTA	2794
1706	1728	CCAGACAACCTTAGCC AAACCAT	1026	ATGGTTTGGCTAAGG TTGTC	2795
1714	1736	CCTTAGCCAAACCATT TACCCAA	1027	TTGGGTAAATGGTTT GGCTA	2796
1720	1742	CCAAACCATTTACCCA AATAAAG	1028	CTTTATTTGGGTAAAT GGTT	2797
1725	1747	CCATTTACCCAAATAA AGTATAG	1029	CTATACTTTATTTGGG TAAA	2798
1732	1754	CCCAAATAAAGTATAG GCGATAG	1030	CTATCGCCTATACTTT ATTT	2799
1733	1755	CCAAATAAAGTATAGG CGATAGA	1031	TCTATCGCCTATACTT TATT	2800
1764	1786	CCTGGCGCAATAGATA TAGTACC	1032	GGTACTATATCTATT GCGCC	2801
1785	1807	CCGCAAGGGAAAGAT GAAAAATT	1033	AATTTTTCATCTTTCC CTTG	2802
1812	1834	CCAAGCATAATATAGC AAGGACT	1034	AGTCCTTGCTATATTA TGCT	2803

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
1837	1859	CCCCTATACCTTCTGC ATAATGA	1035	TCATTATGCAGAAGG TATAG	2804
1838	1860	CCCTATACCTTCTGCAT AATGAA	1036	TTCATTATGCAGAAG GTATA	2805
1839	1861	CCTATACCTTCTGCAT AATGAAT	1037	ATTCATTATGCAGAA GGTAT	2806
1845	1867	CCTTCTGCATAATGAA TTAACTA	1038	TAGTTAATTCATTATG CAGA	2807
1889	1911	CCAAAGCTAAGACCCC CGAAACC	1039	GGTTTCGGGGGTCTT AGCTT	2808
1901	1923	CCCCGAAACCAGACG AGCTACC	1040	GGTAGCTCGTCTGGT TTCGG	2809
1902	1924	CCCCGAAACCAGACGA GCTACCT	1041	AGGTAGCTCGTCTGG TTTCG	2810
1903	1925	CCCGAAACCAGACGAG CTACCTA	1042	TAGGTAGCTCGTCTG GTTTC	2811
1904	1926	CCGAAACCAGACGAGC TACCTAA	1043	TTAGGTAGCTCGTCT GGTTT	2812
1910	1932	CCAGACGAGCTACCTA AGAACAG	1044	CTGTTCTTAGGTAGCT CGTC	2813
1922	1944	CCTAAGAACAGCTAAA AGAGCAC	1045	GTGCTCTTTTAGCTGT TCTT	2814
1946	1968	CCCGTCTATGTAGCAA AATAGTG	1046	CACTATTTTGCTACAT AGAC	2815
1947	1969	CCGTCTATGTAGCAAA ATAGTGG	1047	CCACTATTTTGCTACA TAGA	2816
1996	2018	CCTACCGAGCCTGGTG ATAGCTG	1048	CAGCTATCACCAGGC TCGGT	2817
2000	2022	CCGAGCCTGGTGATAG CTGGTTG	1049	CAACCAGCTATCACC AGGCT	2818
2005	2027	CCTGGTGATAGCTGGT TGTCCAA	1050	TTGGACAACCAGCTA TCACC	2819
2024	2046	CCAAGATAGAATCTTA GTTCAAC	1051	GTTGAACATAAGATTC TATCT	2820
2057	2079	CCCACAGAACCCTCTA AATCCCC	1052	GGGGATTAGAGGGT TCTGT	2821
2058	2080	CCACAGAACCCTCTAA ATCCCCCT	1053	AGGGGATTAGAGGG TTCTG	2822
2066	2088	CCCTCTAAATCCCCTT GTAAATT	1054	AATTTACAAGGGGAT TTAGA	2823
2067	2089	CCTCTAAATCCCCTTGT AAATTT	1055	AAATTTACAAGGGGA TTAG	2824
2076	2098	CCCCTTGTAATTTTAA CTGTTAG	1056	CTAACAGTTAAATTT ACAAG	2825
2077	2099	CCCTTGTAATTTAAC	1057	ACTAACAGTTAAATT	2826

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		TGTTAGT		TACAA	
2078	2100	CCTTGTA AATT TAACT GTTAGTC	1058	GACTAACAGTTAAAT TTACA	2827
2100	2122	CCAAAGAGGAACAGCT CTTTGGA	1059	TCCAAAGAGCTGTTC CTCTT	2828
2136	2158	CCTTG TAGAGAGAGTA AAAAATT	1060	AATTTTTTACTCTCTC TACA	2829
2164	2186	CCCATAGTAGGCCTAA AAGCAGC	1061	GCTGCTTTTAGGCCT ACTAT	2830
2165	2187	CCATAGTAGGCCTAAA AGCAGCC	1062	GGCTGCTTTTAGGCC TACTA	2831
2175	2197	CCTAAAAGCAGCCACC AATTAAG	1063	CTTAATTGGTGGCTG CTTTT	2832
2186	2208	CCACCAATTAAGAAAG CGTTCAA	1064	TTGAACGCTTTCTTAA TTGG	2833
2189	2211	CCAATTAAGAAAGCGT TCAAGCT	1065	AGCTTGAACGCTTTC TTAAT	2834
2217	2239	CCCACTACCTAAAAAA TCCCAA	1066	TTTGGGATTTTTTAGG TAGT	2835
2218	2240	CCCACTACCTAAAAAAT CCCAAAC	1067	GTTTGGGATTTTTTAG GTAG	2836
2224	2246	CCTAAAAAATCCCAA CATATAA	1068	TTATATGTTTGGGATT TTTT	2837
2234	2256	CCCAAACATATAACTG AACTCCT	1069	AGGAGTTCAGTTATA TGTTT	2838
2235	2257	CCAAACATATAACTGA ACTCCTC	1070	GAGGAGTTCAGTTAT ATGTT	2839
2254	2276	CCTCACACCCAATTGG ACCAATC	1071	GATTGGTCCAATTGG GTGTG	2840
2261	2283	CCCAATTGGACCAATC TATCACC	1072	GGTGATAGATTGGTC CAATT	2841
2262	2284	CCAATTGGACCAATCT ATCACCC	1073	GGGTGATAGATTGGT CCAAT	2842
2271	2293	CCAATCTATCACCTA TAGAAGA	1074	TCTTCTATAGGGTGA TAGAT	2843
2282	2304	CCCTATAGAAGAACTA ATGTTAG	1075	CTAACATTAGTTCTTC TATA	2844
2283	2305	CCTATAGAAGAACTAA TGTTAGT	1076	ACTAACATTAGTTCTT CTAT	2845
2328	2350	CCTCCGCATAAGCCTG CGTCAGA	1077	TCTGACGCAGGCTTA TGCGG	2846
2331	2353	CCGCATAAGCCTGCGT CAGATTA	1078	TAATCTGACGCAGGC TTATG	2847
2340	2362	CCTGCGTCAGATTAAA AACTGA	1079	TCAGTGTTTAAATCTG ACGC	2848

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
2378	2400	CCCAATATCTACAATC AACCAAC	1080	GTTGGTTGATTGTAG ATATT	2849
2379	2401	CCAATATCTACAATCA ACCAACA	1081	TGTTGGTTGATTGTA GATAT	2850
2396	2418	CCAACAAGTCATTATT ACCCTCA	1082	TGAGGGTAATAATGA CTTGT	2851
2413	2435	CCCTCACTGTCAACCC AACACAG	1083	CTGTGTTGGGTTGAC AGTGA	2852
2414	2436	CCTCACTGTCAACCCA ACACAGG	1084	CCTGTGTTGGGTTGA CAGTG	2853
2426	2448	CCCAACACAGGCATGC TCATAAG	1085	CTTATGAGCATGCCT GTGTT	2854
2427	2449	CCAACACAGGCATGCT CATAAGG	1086	CCTTATGAGCATGCC TGTGT	2855
2488	2510	CCCCGCCTGTTTACCA AAAACAT	1087	ATGTTTTTGGTAAAC AGGCG	2856
2489	2511	CCCGCCTGTTTACCAA AAACATC	1088	GATGTTTTTGGTAAA CAGGC	2857
2490	2512	CCGCCTGTTTACCAAA AACATCA	1089	TGATGTTTTTGGTAA ACAGG	2858
2493	2515	CCTGTTTACCAAAAAC ATCACCT	1090	AGGTGATGTTTTTGG TAAAC	2859
2501	2523	CCAAAAACATCACCTC TAGCATC	1091	GATGCTAGAGGTGAT GTTTT	2860
2513	2535	CCTCTAGCATCACCAG TATTAGA	1092	TCTAATACTGGTGAT GCTAG	2861
2525	2547	CCAGTATTAGAGGCAC CGCCTGC	1093	GCAGGCGGTGCCTCT AATAC	2862
2540	2562	CCGCCTGCCCAGTGAC ACATGTT	1094	AACATGTGTCACTGG GCAGG	2863
2543	2565	CCTGCCCAGTGACACA TGTTTAA	1095	TTAAACATGTGTCAC TGGGC	2864
2547	2569	CCCAGTGACACATGTT TAACGGC	1096	GCCGTAAACATGTG TCACT	2865
2548	2570	CCAGTGACACATGTTT AACGGCC	1097	GGCCGTAAACATGT GTCAC	2866
2569	2591	CCGCGGTACCCTAACC GTGCAA	1098	TTTGCACGGTTAGGG TACCG	2867
2577	2599	CCCTAACCGTGCAAAG GTAGCAT	1099	ATGCTACCTTTGCAC GGTTA	2868
2578	2600	CCTAACCGTGCAAAGG TAGCATA	1100	TATGCTACCTTTGCAC GGTT	2869
2583	2605	CCGTGCAAAGGTAGCA TAATCAC	1101	GTGATTATGCTACCTT TGCA	2870
2611	2633	CCTTAAATAGGGACCT	1102	TTCATACAGGTCCCT	2871

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		GTATGAA		ATTTA	
2624	2646	CCTGTATGAATGGCTC CACGAGG	1103	CCTCGTGGAGCCATT CATAAC	2872
2639	2661	CCACGAGGGTTCAGCT GTCTCTT	1104	AAGAGACAGCTGAAC CCTCG	2873
2670	2692	CCAGTGAAATTGACCT GCCCGTG	1105	CACGGGCAGGTCAAT TTCAC	2874
2683	2705	CCTGCCCCGTGAAGAGG CGGGCAT	1106	ATGCCCCGCCTCTTCA CGGGC	2875
2687	2709	CCCGTGAAGAGGCGGG CATAACA	1107	TGTTATGCCCCGCCTCT TCAC	2876
2688	2710	CCGTGAAGAGGCGGGC ATAACAC	1108	GTGTTATGCCCCGCCT CTTCA	2877
2726	2748	CCCTATGGAGCTTTAA TTTATTA	1109	TAATAAATTAAAGCT CCATA	2878
2727	2749	CCTATGGAGCTTTAAT TTATTAA	1110	TTAATAAATTAAAGC TCCAT	2879
2761	2783	CCTAACAAACCCACAG GTCCTAA	1111	TTAGGACCTGTGGGT TTGTT	2880
2770	2792	CCCACAGGTCCTAAAC TACCAA	1112	TTTGGTAGTTTAGGA CCTGT	2881
2771	2793	CCACAGGTCCTAAACT ACCAAAC	1113	GTTTGGTAGTTTAGG ACCTG	2882
2779	2801	CCTAAACTACCAAACC TGCATTA	1114	TAATGCAGGTTTGGT AGTTT	2883
2788	2810	CCAAACCTGCATTAAA AATTTTCG	1115	CGAAATTTTAAATGC AGGTT	2884
2793	2815	CCTGCATTAAAAATTT CGGTTGG	1116	CCAACCGAAATTTTT AATGC	2885
2821	2843	CCTCGGAGCAGAACCC AACCTCC	1117	GGAGGTTGGGTTCTG CTCCG	2886
2834	2856	CCCAACCTCCGAGCAG TACATGC	1118	GCATGTACTGCTCGG AGGTT	2887
2835	2857	CCAACCTCCGAGCAGT ACATGCT	1119	AGCATGTACTGCTCG GAGGT	2888
2839	2861	CCTCCGAGCAGTACAT GCTAAGA	1120	TCTTAGCATGTACTG CTCGG	2889
2842	2864	CCGAGCAGTACATGCT AAGACTT	1121	AAGTCTTAGCATGTA CTGCT	2890
2867	2889	CCAGTCAAAGCGAACT ACTATAC	1122	GTATAGTAGTTCGCT TTGAC	2891
2899	2921	CCAATAACTTGACCAA CGGAACA	1123	TGTTCCGTTGGTCAA GTTAT	2892
2911	2933	CCAACGGAACAAGTTA CCCTAGG	1124	CCTAGGGTAACTTGT TCCGT	2893

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
2927	2949	CCCTAGGGATAACAGC GCAATCC	1125	GGATTGCGCTGTTAT CCCTA	2894
2928	2950	CCTAGGGATAACAGCG CAATCCT	1126	AGGATTGCGCTGTTA TCCCT	2895
2948	2970	CCTATTCTAGAGTCCA TATCAAC	1127	GTTGATATGGACTCT AGAAT	2896
2961	2983	CCATATCAACAATAGG GTTTACG	1128	CGTAAACCCTATTGT TGATA	2897
2985	3007	CCTCGATGTTGGATCA GGACATC	1129	GATGTCCTGATCCAA CATCG	2898
3007	3029	CCCGATGGTGCAGCCG CTATTAA	1130	TTAATAGCGGCTGCA CCATC	2899
3008	3030	CCGATGGTGCAGCCGC TATTTAA	1131	TTTAATAGCGGCTGC ACCAT	2900
3020	3042	CCGCTATTAAAGGTTC GTTTGTT	1132	AACAAACGAACCTTT AATAG	2901
3056	3078	CCTACGTGATCTGAGT TCAGACC	1133	GGTCTGAACTCAGAT CACGT	2902
3077	3099	CCGGAGTAATCCAGGT CGGTTTC	1134	GAAACCGACCTGGAT TACTC	2903
3087	3109	CCAGGTCGGTTTCTAT CTACNTT	1135	AANGTAGATAGAAAC CGACC	2904
3116	3138	CCTCCCTGTACGAAAG GACAAGA	1136	TCTTGTCTTTCGTAC AGGG	2905
3119	3141	CCCTGTACGAAAGGAC AAGAGAA	1137	TTCTCTTGTCTTTCG TACA	2906
3120	3142	CCTGTACGAAAGGACA AGAGAAA	1138	TTTCTCTTGTCTTTC GTAC	2907
3148	3170	CCTACTTCACAAAGCG CCTTCCC	1139	GGGAAGGCGCTTTGT GAAGT	2908
3164	3186	CCTTCCCCCGTAAATG ATATCAT	1140	ATGATATCATTTACG GGGGA	2909
3168	3190	CCCCCGTAAATGATAT CATCTCA	1141	TGAGATGATATCATT TACGG	2910
3169	3191	CCCCGTAAATGATATC ATCTCAA	1142	TTGAGATGATATCAT TTACG	2911
3170	3192	CCCGTAAATGATATCA TCTCAAC	1143	GTTGAGATGATATCA TTTAC	2912
3171	3193	CCGTAAATGATATCAT CTCAACT	1144	AGTTGAGATGATATC ATTTA	2913
3204	3226	CCCACACCCACCCAAG AACAGGG	1145	CCCTGTTCTTGGGTG GGTGT	2914
3205	3227	CCACACCCACCCAAGA ACAGGGT	1146	ACCCTGTTCTTGGGT GGGTG	2915
3210	3232	CCCACCCAAGAACAGG	1147	AACAAACCCTGTTCT	2916

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		GTTTGTT		TGGGT	
3211	3233	CCACCCAAGAACAGGGTTTGTTA	1148	TAACAAACCCTGTTC TTGGG	2917
3214	3236	CCCAAGAACAGGGTTTGTTAAGA	1149	TCTTAACAAACCCTG TTCTT	2918
3215	3237	CCAAGAACAGGGTTTGTTAAGAT	1150	ATCTTAACAAACCCT GTTCT	2919
3245	3267	CCCGGTAATCGCATAA AACTTAA	1151	TTAAGTTTTATGCGAT TACC	2920
3246	3268	CCGGTAATCGCATAAA ACTTAAA	1152	TTTAAGTTTTATGCGA TTAC	2921
3292	3314	CCTCTTCTTAACAACA TACCCAT	1153	ATGGGTATGTTGTTA AGAAG	2922
3310	3332	CCCATGGCCAACCTCC TACTCCT	1154	AGGAGTAGGAGGTTG GCCAT	2923
3311	3333	CCATGGCCAACCTCCT ACTCCTC	1155	GAGGAGTAGGAGGTT GGCCA	2924
3317	3339	CCAACCTCCTACTCCT CATTGTA	1156	TACAATGAGGAGTAG GAGGT	2925
3321	3343	CCTCCTACTCCTCATTG TACCCA	1157	TGGGTACAATGAGGA GTAGG	2926
3324	3346	CCTACTCCTCATTGTAC CCATTC	1158	GAATGGGTACAATGA GGAGT	2927
3330	3352	CCTCATTGTACCCATTC TAATCG	1159	CGATTAGAATGGGTA CAATG	2928
3340	3362	CCCATTCTAATCGCAA TGGCATT	1160	AATGCCATTGCGATT AGAAT	2929
3341	3363	CCATTCTAATCGCAAT GGCATTC	1161	GAATGCCATTGCGAT TAGAA	2930
3363	3385	CCTAATGCTTACCGAA CGAAAAA	1162	TTTTTCGTTCGGTAAG CATT	2931
3374	3396	CCGAACGAAAAATTCT AGGCTAT	1163	ATAGCCTAGAATTTT TCGTT	2932
3414	3436	CCCCAACGTTGTAGGC CCCTACG	1164	CGTAGGGGCCTACAA CGTTG	2933
3415	3437	CCCAACGTTGTAGGCC CCTACGG	1165	CCGTAGGGGCCTACA ACGTT	2934
3416	3438	CCAACGTTGTAGGCCC CTACGGG	1166	CCCGTAGGGGCCTAC AACGT	2935
3429	3451	CCCCTACGGGCTACTA CAACCCT	1167	AGGGTTGTAGTAGCC CGTAG	2936
3430	3452	CCCTACGGGCTACTAC AACCCTT	1168	AAGGGTTGTAGTAGC CCGTA	2937
3431	3453	CCTACGGGCTACTACA ACCCTTC	1169	GAAGGGTTGTAGTAG CCCGT	2938

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
3448	3470	CCCTTCGCTGACGCCA TAAAACT	1170	AGTTTTATGGCGTCA GCGAA	2939
3449	3471	CCTTCGCTGACGCCAT AAAACTC	1171	GAGTTTTATGGCGTC AGCGA	2940
3461	3483	CCATAAAACTCTTCAC CAAAGAG	1172	CTCTTTGGTGAAGAG TTTTA	2941
3476	3498	CCAAAGAGCCCCTAAA ACCCGCC	1173	GGCGGGTTTTAGGGG CTCTT	2942
3484	3506	CCCCTAAAACCCGCCA CATCTAC	1174	GTAGATGTGGCGGGT TTTAG	2943
3485	3507	CCCTAAAACCCGCCAC ATCTACC	1175	GGTAGATGTGGCGGG TTTTA	2944
3486	3508	CCTAAAACCCGCCACA TCTACCA	1176	TGGTAGATGTGGCGG GTTTT	2945
3493	3515	CCCGCCACATCTACCA TCACCCT	1177	AGGGTGATGGTAGAT GTGGC	2946
3494	3516	CCGCCACATCTACCAT CACCCTC	1178	GAGGGTGATGGTAGA TGTGG	2947
3497	3519	CCACATCTACCATCAC CCTCTAC	1179	GTAGAGGGTGATGGT AGATG	2948
3506	3528	CCATCACCTCTACAT CACCGCC	1180	GGCGGTGATGTAGAG GGTGA	2949
3512	3534	CCCTCTACATCACCGC CCCGACC	1181	GGTCGGGGCGGTGAT GTAGA	2950
3513	3535	CCTCTACATCACCGCC CCGACCT	1182	AGGTCGGGGCGGTGA TGTAG	2951
3524	3546	CCGCCCCGACCTTAGC TCTCACC	1183	GGTGAGAGCTAAGGT CGGGG	2952
3527	3549	CCCCGACCTTAGCTCT CACCATC	1184	GATGGTGAGAGCTAA GGTCG	2953
3528	3550	CCCGACCTTAGCTCTC ACCATCG	1185	CGATGGTGAGAGCTA AGGTC	2954
3529	3551	CCGACCTTAGCTCTCA CCATCGC	1186	GCGATGGTGAGAGCT AAGGT	2955
3533	3555	CCTTAGCTCTCACCAT CGCTCTT	1187	AAGAGCGATGGTGAG AGCTA	2956
3545	3567	CCATCGCTCTTCTACTA TGAACC	1188	GGTTCATAGTAGAAG AGCGA	2957
3566	3588	CCCCCTCCCCATACC CAACCCC	1189	GGGGTTGGGTATGGG GAGGG	2958
3567	3589	CCCCCTCCCCATACCC AACCCCC	1190	GGGGGTGGGTATGG GGAGG	2959
3568	3590	CCCCTCCCCATACCCA ACCCCCT	1191	AGGGGGTTGGGTATG GGGAG	2960
3569	3591	CCCTCCCCATACCCAA	1192	CAGGGGGTTGGGTAT	2961

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CCCCCTG		GGGGA	
3570	3592	CCTCCCCATACCCAAC CCCCTGG	1193	CCAGGGGGTGTGGTA TGGGG	2962
3573	3595	CCCCATACCCAACCCC CTGGTCA	1194	TGACCAGGGGGTGTG GTATG	2963
3574	3596	CCCATACCCAACCCCC TGGTCAA	1195	TTGACCAGGGGGTGTG GGTAT	2964
3575	3597	CCATACCCAACCCCCT GGTCAAC	1196	GTTGACCAGGGGGTGTG GGTA	2965
3580	3602	CCCAACCCCCTGGTCA ACCTCAA	1197	TTGAGGTTGACCAGG GGGT	2966
3581	3603	CCAACCCCCTGGTCAA CCTCAAC	1198	GTTGAGGTTGACCAG GGGT	2967
3585	3607	CCCCCTGGTCAACCTC AACCTAG	1199	CTAGGTTGAGGTTGA CCAGG	2968
3586	3608	CCCCTGGTCAACCTCA ACCTAGG	1200	CCTAGGTTGAGGTTG ACCAG	2969
3587	3609	CCCTGGTCAACCTCAA CCTAGGC	1201	GCCTAGGTTGAGGTT GACCA	2970
3588	3610	CCTGGTCAACCTCAAC CTAGGCC	1202	GGCCTAGGTTGAGGT TGACC	2971
3597	3619	CCTCAACCTAGGCCTC CTATTTA	1203	TAAATAGGAGGCCTA GGTTG	2972
3603	3625	CCTAGGCCTCCTATTT ATTCTAG	1204	CTAGAATAAATAGGA GGCCT	2973
3609	3631	CCTCCTATTTATTCTAG CCACCT	1205	AGGTGGCTAGAATAA ATAGG	2974
3612	3634	CCTATTTATTCTAGCCA CCTCTA	1206	TAGAGGTGGCTAGAA TAAAT	2975
3626	3648	CCACCTCTAGCCTAGC CGTTTAC	1207	GTAAACGGCTAGGCT AGAGG	2976
3629	3651	CCTCTAGCCTAGCCGT TACTCA	1208	TGAGTAAACGGCTAG GCTAG	2977
3636	3658	CCTAGCCGTTTACTCA ATCCTCT	1209	AGAGGATTGAGTAAA CGGCT	2978
3641	3663	CCGTTTACTCAATCCTC TGATCA	1210	TGATCAGAGGATTGA GTAAA	2979
3654	3676	CCTCTGATCAGGGTGA GCATCAA	1211	TTGATGCTCACCTG ATCAG	2980
3689	3711	CCCTGATCGGCGCACT GCGAGCA	1212	TGCTCGCAGTGCGCC GATCA	2981
3690	3712	CCTGATCGGCGCACTG CGAGCAG	1213	CTGCTCGCAGTGCGC CGATC	2982
3716	3738	CCCAAACAATCTCATA TGAAGTC	1214	GACTTCATATGAGAT TGTTT	2983

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
3717	3739	CCAAACAATCTCATATGAAGTCA	1215	TGACTTCATATGAGATTGTT	2984
3740	3762	CCCTAGCCATCATTCTACTATCA	1216	TGATAGTAGAATGATGGCTA	2985
3741	3763	CCTAGCCATCATTCTACTATCAA	1217	TTGATAGTAGAATGATGGCT	2986
3746	3768	CCATCATTCTACTATCAACATTA	1218	TAATGTTGATAGTAGAATGA	2987
3782	3804	CCTTTAACCTCTCCACCCTTATC	1219	GATAAGGGTGGAGAGGTAA	2988
3789	3811	CCTCTCCACCCTTATCACAACAC	1220	GTGTTGTGATAAGGGTGGAG	2989
3794	3816	CCACCCTTATCACAACACAAGAA	1221	TTCTTGTGTTGTGATAAGGG	2990
3797	3819	CCCTTATCACAACACAAGAACAC	1222	GTGTTCTTGTGTTGTGATAA	2991
3798	3820	CCTTATCACAACACAAGAACACC	1223	GGTGTTCCTTGTGTTGTGATA	2992
3819	3841	CCTCTGATTACTCCTGCATCAT	1224	ATGATGGCAGGAGTATCAG	2993
3831	3853	CCTGCCATCATGACCCTTGGCCA	1225	TGGCCAAGGGTCATGATGGC	2994
3835	3857	CCATCATGACCCTTGGCCATAAT	1226	ATTATGGCCAAGGGTCATGA	2995
3844	3866	CCCTTGGCCATAATATGATTTAT	1227	ATAAATCATATTATGGCCAA	2996
3845	3867	CCTTGGCCATAATATGATTTATC	1228	GATAAATCATATTATGGCCA	2997
3851	3873	CCATAATATGATTTATCTCCACA	1229	TGTGGAGATAAATCATATTA	2998
3869	3891	CCACACTAGCAGAGACCAACCGA	1230	TCGGTTGGTCTCTGCTAGTG	2999
3884	3906	CCAACCGAACCCCCTTCGACCTT	1231	AAGGTCGAAGGGGGTTCGGT	3000
3888	3910	CCGAACCCCCTTCGACCTTGCCG	1232	CGGCAAGGTCGAAGGGGTT	3001
3893	3915	CCCCCTTCGACCTTGCCGAAGGG	1233	CCCTTCGGCAAGGTCGAAGG	3002
3894	3916	CCCCTTCGACCTTGCCGAAGGGG	1234	CCCCTTCGGCAAGGTCGAAG	3003
3895	3917	CCCTTCGACCTTGCCGAAGGGGA	1235	TCCCCTTCGGCAAGGTCGAA	3004
3896	3918	CCTTCGACCTTGCCGAAGGGGAG	1236	CTCCCCTTCGGCAAGGTCGA	3005
3903	3925	CCTTGCCGAAGGGGAG	1237	GTTCGGACTCCCCTTC	3006

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		TCCGAAC		GGCA	
3908	3930	CCGAAGGGGAGTCCGA ACTAGTC	1238	GACTAGTTCGGACTC CCCTT	3007
3920	3942	CCGAAGTAGTCTCAGG CTTCAAC	1239	GTTGAAGCCTGAGAC TAGTT	3008
3953	3975	CCGCAGGCCCTTCGC CCTATTC	1240	GAATAGGGCGAAGG GGCCTG	3009
3960	3982	CCCCTTCGCCCTATTCT TCATAG	1241	CTATGAAGAATAGGG CGAAG	3010
3961	3983	CCCTTCGCCCTATTCTT CATAGC	1242	GCTATGAAGAATAGG GCGAA	3011
3962	3984	CCTTCGCCCTATTCTTC ATAGCC	1243	GGCTATGAAGAATAG GGCGA	3012
3968	3990	CCCTATTCTTCATAGCC GAATAC	1244	GTATTCGGCTATGAA GAATA	3013
3969	3991	CCTATTCTTCATAGCC GAATACA	1245	TGTATTCGGCTATGA AGAAT	3014
3983	4005	CCGAATACACAAACAT TATTATA	1246	TATAATAATGTTTGT GTATT	3015
4013	4035	CCCTCACCCTACAAT CTTCCTA	1247	TAGGAAGATTGTAGT GGTGA	3016
4014	4036	CCTCACCCTACAATC TTCCTAG	1248	CTAGGAAGATTGTAG TGGTG	3017
4019	4041	CCACTACAATCTTCCT AGGAACA	1249	TGTTCTAGGAAGAT TGTAG	3018
4032	4054	CCTAGGAACAACATAT GACGCAC	1250	GTGCGTCATATGTTG TTCCT	3019
4058	4080	CCCCTGAACTCTACAC AACATAT	1251	ATATGTTGTGTAGAG TTCAG	3020
4059	4081	CCCTGAACTCTACACA ACATATT	1252	AATATGTTGTGTAGA GTTCA	3021
4060	4082	CCTGAACTCTACACAA CATATTT	1253	AAATATGTTGTGTAG AGTTC	3022
4088	4110	CCAAGACCCTACTTCT AACCTCC	1254	GGAGGTTAGAAGTAG GGTCT	3023
4094	4116	CCCTACTTCTAACCTCC CTGTTT	1255	GAACAGGGAGGTTAG AAGTA	3024
4095	4117	CCTACTTCTAACCTCCC TGTTCT	1256	AGAACAGGGAGGTTA GAAGT	3025
4106	4128	CCTCCCTGTTCTTATGA ATTCGA	1257	TCGAATTCATAAGAA CAGGG	3026
4109	4131	CCCTGTTCTTATGAATT CGAACA	1258	TGTTCTGAATTCATA GAACA	3027
4110	4132	CCTGTTCTTATGAATTC GAACAG	1259	CTGTTCTGAATTCATA AGAAC	3028

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
4137	4159	CCCCGATTCCGCTAC GACCAAC	1260	GTTGGTCGTAGCGGA ATCGG	3029
4138	4160	CCCCGATTCCGCTACG ACCAACT	1261	AGTTGGTCGTAGCGG AATCG	3030
4139	4161	CCCGATTCCGCTACGA CCAACTC	1262	GAGTTGGTCGTAGCG GAATC	3031
4140	4162	CCGATTCCGCTACGAC CAACTCA	1263	TGAGTTGGTCGTAGC GGAAT	3032
4146	4168	CCGCTACGACCAACTC ATACACC	1264	GGTGTATGAGTTGGT CGTAG	3033
4155	4177	CCAACCTACACCTC CTATGAA	1265	TTCATAGGAGGTGTA TGAGT	3034
4167	4189	CCTCCTATGAAAAAAC TTCCTAC	1266	GTAGGAAGTTTTTTC ATAGG	3035
4170	4192	CCTATGAAAAAACTTC CTACCAC	1267	GTGGTAGGAAGTTTT TTCAT	3036
4185	4207	CCTACCACTCACCCTA GCATTAC	1268	GTAATGCTAGGGTGA GTGGT	3037
4189	4211	CCACTCACCCTAGCAT TACTTAT	1269	ATAAGTAATGCTAGG GTGAG	3038
4196	4218	CCCTAGCATTACTTAT ATGATAT	1270	ATATCATATAAGTAA TGCTA	3039
4197	4219	CCTAGCATTACTTATA TGATATG	1271	CATATCATATAAGTA ATGCT	3040
4223	4245	CCATACCCATTACAAT CTCCAGC	1272	GCTGGAGATTGTAAT GGGTA	3041
4228	4250	CCCATTACAATCTCCA GCATTCC	1273	GGAATGCTGGAGATT GTAAT	3042
4229	4251	CCATTACAATCTCCAG CATTCCC	1274	GGGAATGCTGGAGAT TGTA	3043
4241	4263	CCAGCATTCCCCCTCA AACCTAA	1275	TTAGGTTTGAGGGGG AATGC	3044
4249	4271	CCCCCTCAAACCTAAG AAATATG	1276	CATATTTCTTAGGTTT GAGG	3045
4250	4272	CCCCTCAAACCTAAGA AATATGT	1277	ACATATTTCTTAGGTT TGAG	3046
4251	4273	CCCTCAAACCTAAGAA ATATGTC	1278	GACATATTTCTTAGG TTTGA	3047
4252	4274	CCTCAAACCTAAGAAA TATGTCT	1279	AGACATATTTCTTAG GTTTG	3048
4259	4281	CCTAAGAAATATGTCT GATAAAA	1280	TTTTATCAGACATATT TCTT	3049
4318	4340	CCCCCTTATTTCTAGG ACTATGA	1281	TCATAGTCCTAGAAA TAAGG	3050
4319	4341	CCCCTTATTTCTAGGA	1282	CTCATAGTCCTAGAA	3051

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CTATGAG		ATAAG	
4320	4342	CCCTTATTTCTAGGACT ATGAGA	1283	TCTCATAGTCCTAGA AATAA	3052
4321	4343	CCTTATTTCTAGGACT ATGAGAA	1284	TTCTCATAGTCCTAG AAATA	3053
4349	4371	CCCATCCCTGAGAATC CAAAATT	1285	AATTTTGGATTCTCA GGGAT	3054
4350	4372	CCATCCCTGAGAATCC AAAATTC	1286	GAATTTTGGATTCTC AGGGA	3055
4354	4376	CCCTGAGAATCCAAAA TTCTCCG	1287	CGGAGAATTTTGGAT TCTCA	3056
4355	4377	CCTGAGAATCCAAAAT TCTCCGT	1288	ACGGAGAATTTTGG TTCTC	3057
4364	4386	CCAAAATTCTCCGTGC CACCTAT	1289	ATAGGTGGCACGGAG AATTT	3058
4374	4396	CCGTGCCACCTATCAC ACCCCAT	1290	ATGGGGTGTGATAGG TGGCA	3059
4379	4401	CCACCTATCACACCCC ATCCTAA	1291	TTAGGATGGGGTGTG ATAGG	3060
4382	4404	CCTATCACACCCCATC CTAAAGT	1292	ACTTTAGGATGGGGT GTGAT	3061
4391	4413	CCCCATCCTAAAGTAA GGTCAGC	1293	GCTGACCTTACTTTA GGATG	3062
4392	4414	CCCATCCTAAAGTAAG GTCAGCT	1294	AGCTGACCTTACTTT AGGAT	3063
4393	4415	CCATCCTAAAGTAAGG TCAGCTA	1295	TAGCTGACCTTACTTT AGGA	3064
4397	4419	CCTAAAGTAAGGTCAG CTAAATA	1296	TATTTAGCTGACCTTA CTTT	3065
4430	4452	CCCATACCCCGAAAAT GTTGGTT	1297	AACCAACATTTTCGG GGTAT	3066
4431	4453	CCATACCCCGAAAATG TTGGTTA	1298	TAACCAACATTTTCG GGGTA	3067
4436	4458	CCCCGAAAATGTTGGT TATACCC	1299	GGGTATAACCAACAT TTTCG	3068
4437	4459	CCCGAAAATGTTGGTT ATACCCT	1300	AGGGTATAACCAACA TTTTC	3069
4438	4460	CCGAAAATGTTGGTTA TACCCTT	1301	AAGGGTATAACCAAC ATTTT	3070
4456	4478	CCCTTCCCGTACTAATT AATCCC	1302	GGGATTAATTAGTAC GGGAA	3071
4457	4479	CCTTCCCGTACTAATT AATCCCC	1303	GGGGATTAATTAGTA CGGGA	3072
4461	4483	CCCGTACTAATTAATC CCCTGGC	1304	GCCAGGGGATTAATT AGTAC	3073

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
4462	4484	CCGTACTAATTAATCC CCTGGCC	1305	GGCCAGGGGATTAAT TAGTA	3074
4476	4498	CCCCTGGCCCAACCCG TCATCTA	1306	TAGATGACGGGTTGG GCCAG	3075
4477	4499	CCCTGGCCCAACCCGT CATCTAC	1307	GTAGATGACGGGTTG GGCCA	3076
4478	4500	CCTGGCCCAACCCGTC ATCTACT	1308	AGTAGATGACGGGTT GGGCC	3077
4483	4505	CCCAACCCGTCATCTA CTCTACC	1309	GGTAGAGTAGATGAC GGGTT	3078
4484	4506	CCAACCCGTCATCTAC TCTACCA	1310	TGGTAGAGTAGATGA CGGGT	3079
4488	4510	CCCGTCATCTACTCTA CCATCTT	1311	AAGATGGTAGAGTAG ATGAC	3080
4489	4511	CCGTCATCTACTCTAC CATCTTT	1312	AAAGATGGTAGAGTA GATGA	3081
4504	4526	CCATCTTTGCAGGCAC ACTCATC	1313	GATGAGTGTGCCTGC AAAGA	3082
4555	4577	CCTGAGTAGGCCTAGA AATAAAC	1314	GTTTATTTCTAGGCCT ACTC	3083
4565	4587	CCTAGAAATAAACATG CTAGCTT	1315	AAGCTAGCATGTTTA TTTCT	3084
4593	4615	CCAGTTCTAACCAAAA AAATAAA	1316	TTTATTTTTTTGGTTA GAAC	3085
4603	4625	CCAAAAAATAAACCC TCGTTCC	1317	GGAACGAGGGTTTAT TTTTT	3086
4616	4638	CCCTCGTTCCACAGAA GCTGCCA	1318	TGGCAGCTTCTGTGG AACGA	3087
4617	4639	CCTCGTTCCACAGAAG CTGCCAT	1319	ATGGCAGCTTCTGTG GAACG	3088
4624	4646	CCACAGAAGCTGCCAT CAAGTAT	1320	ATACTTGATGGCAGC TTCTG	3089
4636	4658	CCATCAAGTATTTCT CACGCAA	1321	TTGCGTGAGGAAATA CTTGA	3090
4649	4671	CCTCACGCAAGCAACC GCATCCA	1322	TGGATGCGGTTGCTT GCGTG	3091
4663	4685	CCGCATCCATAATCCT TCTAATA	1323	TATTAGAAGGATTAT GGATG	3092
4669	4691	CCATAATCCTTCTAAT AGCTATC	1324	GATAGCTATTAGAAG GATTA	3093
4676	4698	CCTTCTAATAGCTATC CTCTTCA	1325	TGAAGAGGATAGCTA TTAGA	3094
4691	4713	CCTCTTCAACAATATA CTCTCCG	1326	CGGAGAGTATATTGT TGAAG	3095
4711	4733	CCGGACAATGAACCAT	1327	ATTGGTTATGGTTCAT	3096

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		AACCAAT		TGTC	
4723	4745	CCATAACCAATACTACCAATCAA	1328	TTGATTGGTAGTATTGGTTA	3097
4729	4751	CCAATACTACCAATCAATACTCA	1329	TGAGTATTGATTGGTAGTAT	3098
4738	4760	CCAATCAATACTCATCATTAATA	1330	TATTAATGATGAGTATTGAT	3099
4795	4817	CCCCCTTTCACCTTCTGAGTCCCA	1331	TGGGACTCAGAAGTGAAAGG	3100
4796	4818	CCCCTTTCACCTTCTGAGTCCAG	1332	CTGGGACTCAGAAGTGAAAG	3101
4797	4819	CCCTTTCACCTTCTGAGTCCCAGA	1333	TCTGGGACTCAGAAGTGAAA	3102
4798	4820	CCTTTCACCTTCTGAGTCCAGAG	1334	CTCTGGGACTCAGAAAGTGAA	3103
4814	4836	CCCAGAGGTTACCCAAAGGCACCC	1335	GGGTGCCTTGGGTAACTCT	3104
4815	4837	CCAGAGGTTACCCAAAGGCACCCC	1336	GGGGTGCCTTGGGTAACTCT	3105
4825	4847	CCAAGGCACCCCTCTGACATCC	1337	GGATGTCAGAGGGGTGCCTT	3106
4826	4848	CCAAGGCACCCCTCTGACATCCG	1338	CGGATGTCAGAGGGGTGCCT	3107
4834	4856	CCCCTCTGACATCCGGCCTGCTT	1339	AAGCAGGCCGGATGTCAGAG	3108
4835	4857	CCCTCTGACATCCGGCCTGCTTC	1340	GAAGCAGGCCGGATGTCAGA	3109
4836	4858	CCTCTGACATCCGGCCTGCTTCT	1341	AGAAGCAGGCCGGATGTCAG	3110
4846	4868	CCGGCCTGCTTCTTCTCACATGA	1342	TCATGTGAGAAGAAGCAGGC	3111
4850	4872	CCTGCTTCTTCTCACATGACAAA	1343	TTTGTGTCATGTGAGAAAGC	3112
4879	4901	CCCCATCTCAATCATATACCAA	1344	TTGGTATATGATTGATGG	3113
4880	4902	CCCCATCTCAATCATATACCAA	1345	TTTGGTATATGATTGATAG	3114
4881	4903	CCCATCTCAATCATATACCAAAT	1346	ATTTGGTATATGATTGAGAT	3115
4882	4904	CCATCTCAATCATATACCAAATC	1347	GATTTGGTATATGATTGAG	3116
4898	4920	CCAAATCTCTCCCTCACTAAACG	1348	CGTTTAGTGAGGGAGAGATT	3117
4908	4930	CCCTCACTAAACGTAAAGCCTTCT	1349	AGAAGGCTTACGTTTAGTGA	3118

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
4909	4931	CCTCACTAAACGTAAG CCTTCTC	1350	GAGAAGGCTTACGTT TAGTG	3119
4925	4947	CCTTCTCCTCACTCTCT CAATCT	1351	AGATTGAGAGAGTGA GGAGA	3120
4931	4953	CCTCACTCTCTCAATCT TATCCA	1352	TGGATAAGATTGAGA GAGTG	3121
4951	4973	CCATCATAGCAGGCAG TTGAGGT	1353	ACCTCAACTGCCTGC TATGA	3122
4982	5004	CCAAACCCAGCTACGC AAAATCT	1354	AGATTTTGCCTAGCT GGGT	3123
4987	5009	CCCAGCTACGCAAAAT CTTAGCA	1355	TGCTAAGATTTTGCG TAGCT	3124
4988	5010	CCAGCTACGCAAAATC TTAGCAT	1356	ATGCTAAGATTTTGC GTAGC	3125
5014	5036	CCTCAATTACCCACAT AGGATGA	1357	TCATCCTATGTGGGT AATTG	3126
5023	5045	CCCACATAGGATGAAT AATAGCA	1358	TGCTATTATTCATCCT ATGT	3127
5024	5046	CCACATAGGATGAATA ATAGCAG	1359	CTGCTATTATTCATCC TATG	3128
5052	5074	CCGTACAACCCTAACA TAACCAT	1360	ATGGTTATGTTAGGG TTGTA	3129
5060	5082	CCCTAACATAACCATT CTTAATT	1361	AATTAAGAATGGTTA TGTTA	3130
5061	5083	CCTAACATAACCATTCT TTAATTT	1362	AAATTAAGAATGGTT ATGTT	3131
5071	5093	CCATTCTTAATTAACT ATTTAT	1363	ATAAATAGTTAAATT AAGAA	3132
5099	5121	CCTAACTACTACCGCA TTCCTAC	1364	GTAGGAATGCGGTAG TAGTT	3133
5110	5132	CCGCATTCTACTACT CAACTTA	1365	TAAGTTGAGTAGTAG GAATG	3134
5117	5139	CCTACTACTCAACTTA AACTCCA	1366	TGGAGTTTAAGTTGA GTAGT	3135
5137	5159	CCAGCACCACGACCCT ACTACTA	1367	TAGTAGTAGGGTCGT GGTGC	3136
5143	5165	CCACGACCCTACTACT ATCTCGC	1368	GCGAGATAGTAGTAG GGTCG	3137
5149	5171	CCCTACTACTATCTCG CACCTGA	1369	TCAGGTGCGAGATAG TAGTA	3138
5150	5172	CCTACTACTATCTCGC ACCTGAA	1370	TTCAGGTGCGAGATA GTAGT	3139
5167	5189	CCTGAAACAAGCTAAC ATGACTA	1371	TAGTCATGTTAGCTT GTTTC	3140
5193	5215	CCCTTAATTCCATCCA	1372	AGGAGGGTGGATGGA	3141

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CCCTCCT		ATTAA	
5194	5216	CCTTAATTCCATCCAC CCTCCTC	1373	GAGGAGGGTGGATGG AATTA	3142
5202	5224	CCATCCACCCTCCTCTC CCTAGG	1374	CCTAGGGAGAGGAGG GTGGA	3143
5206	5228	CCACCCTCCTCTCCCTA GGAGGC	1375	GCCTCCTAGGGAGAG GAGGG	3144
5209	5231	CCCTCCTCTCCCTAGG AGGCCTG	1376	CAGGCCTCCTAGGGA GAGGA	3145
5210	5232	CCTCCTCTCCCTAGGA GGCCTGC	1377	GCAGGCCTCCTAGGG AGAGG	3146
5213	5235	CCTCTCCCTAGGAGGC CTGCCCC	1378	GGGGCAGGCCTCCTA GGGAG	3147
5218	5240	CCCTAGGAGGCCTGCC CCCGCTA	1379	TAGCGGGGGCAGGCC TCCTA	3148
5219	5241	CCTAGGAGGCCTGCCC CCGCTAA	1380	TTAGCGGGGGCAGGC CTCCT	3149
5228	5250	CCTGCCCCCGCTAACC GGCTTTT	1381	AAAAGCCGGTTAGCG GGGGC	3150
5232	5254	CCCCGCTAACCGGCT TTTTGCC	1382	GGCAAAAAGCCGGTT AGCGG	3151
5233	5255	CCCCGCTAACCGGCTT TTTGCCC	1383	GGGCAAAAAGCCGGT TAGCG	3152
5234	5256	CCCGCTAACCGGCTTT TTGCCCA	1384	TGGGCAAAAAGCCGG TTAGC	3153
5235	5257	CCGCTAACCGGCTTTT TGCCCAA	1385	TTGGGCAAAAAGCCG GTTAG	3154
5242	5264	CCGGCTTTTGTGCCAA ATGGGCC	1386	GGCCCATTTGGGCAA AAAGC	3155
5253	5275	CCCAAATGGGCCATTA TCGAAGA	1387	TCTTCGATAATGGCC CATTT	3156
5254	5276	CCAAATGGGCCATTAT CGAAGAA	1388	TTCTTCGATAATGGC CCATT	3157
5263	5285	CCATTATCGAAGAATT CACAAAA	1389	TTTTGTGAATTCTTCG ATAA	3158
5294	5316	CCTCATCATCCCCACC ATCATAG	1390	CTATGATGGTGGGGA TGATG	3159
5303	5325	CCCCACCATCATAGCC ACCATCA	1391	TGATGGTGGCTATGA TGGTG	3160
5304	5326	CCCACCATCATAGCCA CCATCAC	1392	GTGATGGTGGCTATG ATGGT	3161
5305	5327	CCACCATCATAGCCAC CATCACC	1393	GGTGATGGTGGCTAT GATGG	3162
5308	5330	CCATCATAGCCACCAT CACCTC	1394	GAGGGTGATGGTGGC TATGA	3163

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
5317	5339	CCACCATCACCTCCTTAACCTC	1395	GAGGTTAAGGAGGGTGATGG	3164
5320	5342	CCATCACCTCCTTAACTCTAC	1396	GTAGAGGTTAAGGAGGGTGA	3165
5326	5348	CCCTCCTTAACCTCTACTTCTAC	1397	GTAGAAGTAGAGGTTAAGGA	3166
5327	5349	CCTCCTTAACCTCTACTTCTACC	1398	GGTAGAAGTAGAGGTTAAGG	3167
5330	5352	CCTTAACCTCTACTTCTACCTAC	1399	GTAGGTAGAAGTAGAGGTTA	3168
5336	5358	CCTCTACTTCTACCTACGCCTAA	1400	TTAGGCGTAGGTAGAGTAG	3169
5348	5370	CCTACGCCTAATCTACTCCACCT	1401	AGGTGGAGTAGATTAGGCGT	3170
5354	5376	CCTAATCTACTCCACCTCAATCA	1402	TGATTGAGGTGGAGTAGATT	3171
5365	5387	CCACCTCAATCACACTACTCCCC	1403	GGGGAGTAGTGTGATTGAGG	3172
5368	5390	CCTCAATCACACTACTCCCCATA	1404	TATGGGGAGTAGTGTGATTG	3173
5384	5406	CCCCATATCTAACAACGTAAAAA	1405	TTTTTACGTTGTTAGATATG	3174
5385	5407	CCCATATCTAACAACGTAAAAAT	1406	ATTTTTACGTTGTTAGATAT	3175
5386	5408	CCATATCTAACAACGTAAAAATA	1407	TATTTTTACGTTGTTAGATATA	3176
5433	5455	CCCACCCCATTCCTCCCTCACACT	1408	AGTGTGGGGAGGAATGGGGT	3177
5434	5456	CCACCCCATTCCTCCCTCACACTC	1409	GAGTGTGGGGAGGAATGGGG	3178
5437	5459	CCCATTCTCTCCCCACACTCATC	1410	GATGAGTGTGGGGAGGAATG	3179
5438	5460	CCCATTCTCTCCCCACACTCATCG	1411	CGATGAGTGTGGGGAGGAAT	3180
5439	5461	CCATTCTCTCCCCACACTCATCGC	1412	GCGATGAGTGTGGGGAGGAAT	3181
5444	5466	CCTCCCCACACTCATCGCCCTTA	1413	TAAGGGCGATGAGTGTGGGG	3182
5447	5469	CCCACACTCATCGCCCTTACCA	1414	TGGTAAGGGCGATGTGTGTG	3183
5448	5470	CCCACACTCATCGCCCTTACCAC	1415	GTGGTAAGGGCGATGTGTGT	3184
5449	5471	CCACACTCATCGCCCTTACCACG	1416	CGTGGTAAGGGCGATGTGTG	3185
5461	5483	CCCTTACCACGCTACT	1417	AGGTAGGAGTAGCGT	3186

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CCTACCT		GGTAA	
5462	5484	CCTTACCACGCTACTCCTACCTA	1418	TAGGTAGGAGTAGCGTGGTA	3187
5467	5489	CCACGCTACTCCTACCTATCTCC	1419	GGAGATAGGTAGGAGTAGCG	3188
5477	5499	CCTACCTATCTCCCCTTTATAC	1420	GTATAAAAGGGGAGATAGGT	3189
5481	5503	CCTATCTCCCCTTTTATACTAAT	1421	ATTAGTATAAAAGGGGAGAT	3190
5488	5510	CCCCTTTTATACTAATAATCTTA	1422	TAAGATTATTAGTATAAAAG	3191
5489	5511	CCCTTTTATACTAATAATCTTAT	1423	ATAAGATTATTAGTATAAAA	3192
5490	5512	CCTTTTATACTAATAATCTTATA	1424	TATAAGATTATTAGTATAAAA	3193
5534	5556	CCAAGAGCCTTCAAAGCCCTCAG	1425	CTGAGGGCTTTGAAGGCTCT	3194
5541	5563	CCTTCAAAGCCCTCAGTAAGTTG	1426	CAACTTACTGAGGGCTTTGA	3195
5550	5572	CCCTCAGTAAGTTGCAATACTTA	1427	TAAGTATTGCAACTTACTGA	3196
5551	5573	CCTCAGTAAGTTGCAATACTTAA	1428	TTAAGTATTGCAACTTACTG	3197
5601	5623	CCCCACTCTGCATCAACTGAACG	1429	CGTTCAGTTGATGCAGAGTG	3198
5602	5624	CCCACTCTGCATCAACTGAACGC	1430	GCGTTCAGTTGATGCAGAGT	3199
5603	5625	CCACTCTGCATCAACTGAACGCA	1431	TGCGTTCAGTTGATGCAGAG	3200
5632	5654	CCACTTTAATTAAGCTAAGCCCT	1432	AGGGCTTAGCTTAATTAAAG	3201
5651	5673	CCCTTACTAGACCAATGGGACTT	1433	AAGTCCCATTGGTCTAGTAA	3202
5652	5674	CCTTACTAGACCAATGGGACTTA	1434	TAAGTCCCATTGGTCTAGTA	3203
5662	5684	CCAATGGGACTTAAACCCACAAA	1435	TTTGTGGGTTTAAGTCCCAT	3204
5677	5699	CCCACAAACACTTAGTTAACAGC	1436	GCTGTTAACCTAAGTGTTTGT	3205
5678	5700	CCACAAACACTTAGTTAACAGCT	1437	AGCTGTTAACCTAAGTGTTTG	3206
5706	5728	CCCTAATCAACTGGCTTCAATCT	1438	AGATTGAAGCCAGTTGATTA	3207
5707	5729	CCTAATCAACTGGCTTCAATCTA	1439	TAGATTGAAGCCAGTTGATT	3208

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
5735	5757	CCCGCCGCCGGGAAAAAAGGCGG	1440	CCGCCTTTTTTCCCGGCGGC	3209
5736	5758	CCGCCGCCGGGAAAAAAGGCGGG	1441	CCCGCCTTTTTTCCCGGCGG	3210
5739	5761	CCGCCGGGAAAAAAGGCGGGAGA	1442	TCTCCCGCCTTTTTTC	3211
5742	5764	CCGGGAAAAAAGGCGGGAGAAGC	1443	GCTTCTCCCGCCTTTT	3212
5764	5786	CCCCGGCAGGTTTGAACTGCTT	1444	AAGCAGCTTCAAACCTGCCG	3213
5765	5787	CCCGGCAGGTTTGAACTGCTTC	1445	GAAGCAGCTTCAAACCTGCC	3214
5766	5788	CCGGCAGGTTTGAAAGCTGCTTCT	1446	AGAAGCAGCTTCAAACTGC	3215
5817	5839	CCTCGGAGCTGGTAAAAAGAGGC	1447	GCCTCTTTTTACCAGCTCCG	3216
5839	5861	CCTAACCCCTGTCTTTAGATTTA	1448	TAAATCTAAAGACAGGGGT	3217
5844	5866	CCCCTGTCTTTAGATTACAGTC	1449	GACTGTAAATCTAAAGACAG	3218
5845	5867	CCCTGTCTTTAGATTACAGTCC	1450	GGACTGTAAATCTAAAGACA	3219
5846	5868	CCTGTCTTTAGATTACAGTCCA	1451	TGGACTGTAAATCTAAAGAC	3220
5866	5888	CCAATGCTTCACTCAGCCATTTT	1452	AAAATGGCTGAGTGAAGCAT	3221
5882	5904	CCATTTTACCTCACCCCCTACTGA	1453	TCAGTGGGGGTGAGGTAAAA	3222
5890	5912	CCTCACCCCCACTGATGTTGCGC	1454	GGCGAACATCAGTGGGGGTG	3223
5895	5917	CCCCACTGATGTTGCGCCGACCG	1455	CGGTCGGCGAACATCAGTGG	3224
5896	5918	CCCCACTGATGTTGCGCCGACCGT	1456	ACGGTCGGCGAACATCAGTG	3225
5897	5919	CCCACTGATGTTGCGCCGACCGTT	1457	AACGGTCGGCGAACATCAGT	3226
5898	5920	CCCACTGATGTTGCGCCGACCGTTG	1458	CAACGGTCGGCGAACATCAG	3227
5911	5933	CCGACCGTTGACTATTCTCTACA	1459	TGTAGAGAATAGTCAACGGT	3228
5915	5937	CCGTTGACTATTCTCTACAAACC	1460	GGTTTGTAGAGAATAGTCAA	3229
5936	5958	CCACAAAGACATTGGAACACTAT	1461	ATAGTGTTCCAATGTTCTTG	3230
5960	5982	CCTATTATTCGGCGCA	1462	CAGCTCATGCGCCGA	3231

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		TGAGCTG		ATAAT	
5987	6009	CCTAGGCACAGCTCTA AGCCTCC	1463	GGAGGCTTAGAGCTG TGCCT	3232
6005	6027	CCTCCTTATTCGAGCC GAGCTGG	1464	CCAGCTCGGCTCGAA TAAGG	3233
6008	6030	CCTTATTCGAGCCGAG CTGGGCC	1465	GGCCCAGCTCGGCTC GAATA	3234
6019	6041	CCGAGCTGGGCCAGCC AGGCAAC	1466	GTTGCCTGGCTGGCC CAGCT	3235
6029	6051	CCAGCCAGGCAACCTT CTAGGTA	1467	TACCTAGAAGGTTGC CTGGC	3236
6033	6055	CCAGGCAACCTTCTAG GTAACGA	1468	TCGTTACCTAGAAGG TTGCC	3237
6041	6063	CCTTCTAGGTAACGAC CACATCT	1469	AGATGTGGTCGTTAC CTAGA	3238
6056	6078	CCACATCTACAACGTT ATCGTCA	1470	TGACGATAACGTTGT AGATG	3239
6082	6104	CCCATGCATTTGTAAT AATCTTC	1471	GAAGATTATTACAAA TGCAT	3240
6083	6105	CCATGCATTTGTAATA ATCTTCT	1472	AGAAGATTATTACAA ATGCA	3241
6117	6139	CCCATCATAATCGGAG GCTTTGG	1473	CCAAAGCCTCCGATT ATGAT	3242
6118	6140	CCATCATAATCGGAGG CTTTGGC	1474	GCCAAAGCCTCCGAT TATGA	3243
6153	6175	CCCCTAATAATCGGTG CCCCCGA	1475	TCGGGGGCACCGATT ATTAG	3244
6154	6176	CCCTAATAATCGGTGC CCCCGAT	1476	ATCGGGGGCACCGAT TATTA	3245
6155	6177	CCTAATAATCGGTGCC CCCCGATA	1477	TATCGGGGGCACC GA TTATT	3246
6169	6191	CCCCCGATATGGCGTT TCCCCGC	1478	GCGGGGAAACGCCAT ATCGG	3247
6170	6192	CCCCGATATGGCGTTT CCCCGCA	1479	TGCGGGGAAACGCCA TATCG	3248
6171	6193	CCCGATATGGCGTTTC CCCGCAT	1480	ATGCGGGGAAACGCC ATATC	3249
6172	6194	CCGATATGGCGTTTCC CCGCATA	1481	TATGCGGGGAAACGC CATAT	3250
6186	6208	CCCCGCATAAACAACA TAAGCTT	1482	AAGCTTATGTTGTTTA TGCG	3251
6187	6209	CCCGCATAAACAACAT AAGCTTC	1483	GAAGCTTATGTTGTTT ATGC	3252
6188	6210	CCGCATAAACAACATA AGCTTCT	1484	AGAAGCTTATGTTGT TTATG	3253

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
6219	6241	CCTCCCTCTCTCCTACTCCTGCT	1485	AGCAGGAGTAGGAGAGAGGG	3254
6222	6244	CCCTCTCTCCTACTCCTGCTCGC	1486	GCGAGCAGGAGTAGGAGAGA	3255
6223	6245	CCTCTCTCCTACTCCTGCTCGCA	1487	TGCGAGCAGGAGTAGGAGAG	3256
6230	6252	CCTACTCCTGCTCGCATCTGCTA	1488	TAGCAGATGCGAGCAGGAGT	3257
6236	6258	CCTGCTCGCATCTGCTATAGTGG	1489	CCACTATAGCAGATGCGAGC	3258
6262	6284	CCGGAGCAGGAACAGGTTGAACA	1490	TGTTCAACCTGTTCTCCTGCTC	3259
6290	6312	CCCTCCCTTAGCAGGGAACACTACT	1491	AGTAGTTCCTGCTAAGGGA	3260
6291	6313	CCTCCCTTAGCAGGGAACACTACTC	1492	GAGTAGTTCCTGCTAAGGG	3261
6294	6316	CCCTTAGCAGGGAACACTCCCA	1493	TGGGAGTAGTTCCTGCTAA	3262
6295	6317	CCTTAGCAGGGAACACTCCAC	1494	GTGGGAGTAGTTCCTGCTA	3263
6313	6335	CCCACCCTGGAGCCTCGTAGAC	1495	GTCTACGGAGGCTCAGGGT	3264
6314	6336	CCACCCTGGAGCCTCGTAGACC	1496	GGTCTACGGAGGCTCAGGG	3265
6317	6339	CCCTGGAGCCTCCGTAGACCTAA	1497	TTAGGTCTACGGAGGCTCCA	3266
6318	6340	CCTGGAGCCTCCGTAGACCTAAC	1498	GTTAGGTCTACGGAGGCTCC	3267
6325	6347	CCTCCGTAGACCTAACCATCTTC	1499	GAAGATGGTTAGGCTACGG	3268
6328	6350	CCGTAGACCTAACCATCTTCTCC	1500	GGAGAAGATGGTTAGGCTA	3269
6335	6357	CCTAACCATCTTCTCCTTACACC	1501	GGTGTAAGGAGAAGATGGTT	3270
6340	6362	CCATCTTCTCCTTACACCTAGCA	1502	TGCTAGGTGTAAGGAGAAGA	3271
6349	6371	CCTTACACCTAGCAGGTGTCTCC	1503	GGAGACACCTGCTAGGTGTA	3272
6356	6378	CCTAGCAGGTGTCTCTCTATCT	1504	AGATAGAGGAGACACCTGCT	3273
6370	6392	CCTCTATCTTAGGGGCATCAAT	1505	ATTGATGGCCCCTAAGATAG	3274
6385	6407	CCATCAATTCATCACACAATT	1506	AATTGTTGTGATGAATTGA	3275
6420	6442	CCCCCTGCCATAACCC	1507	TGGTATTGGGTTATG	3276

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		AATACCA		GCAGG	
6421	6443	CCCCTGCCATAACCCA ATACCAA	1508	TTGGTATTGGGTTAT GGCAG	3277
6422	6444	CCCTGCCATAACCCAA TACCAA	1509	TTTGGTATTGGGTTAT GGCA	3278
6423	6445	CCTGCCATAACCCAAT ACCAAC	1510	GTTTGGTATTGGGTT ATGGC	3279
6427	6449	CCATAACCCAATACCA AACGCCC	1511	GGGCGTTTGGTATTG GGTTA	3280
6433	6455	CCCAATACCAAACGCC CCTCTTC	1512	GAAGAGGGGCGTTTG GTATT	3281
6434	6456	CCAATACCAAACGCCC CTCTTCG	1513	CGAAGAGGGGCGTTT GGTAT	3282
6440	6462	CCAAACGCCCCTCTTC GTCTGAT	1514	ATCAGACGAAGAGGG GCGTT	3283
6447	6469	CCCCTCTTCGTCTGATC CGTCCT	1515	AGGACGGATCAGACG AAGAG	3284
6448	6470	CCCTCTTCGTCTGATCC GTCCTA	1516	TAGGACGGATCAGAC GAAGA	3285
6449	6471	CCTCTTCGTCTGATCCG TCCTAA	1517	TTAGGACGGATCAGA CGAAG	3286
6463	6485	CCGTCCTAATCACAGC AGTCCTA	1518	TAGGACTGCTGTGAT TAGGA	3287
6467	6489	CCTAATCACAGCAGTC CTACTTC	1519	GAAGTAGGACTGCTG TGATT	3288
6482	6504	CCTACTTCTCCTATCTC TCCCAG	1520	CTGGGAGAGATAGGA GAAGT	3289
6491	6513	CCTATCTCTCCCAGTCC TAGCTG	1521	CAGCTAGGACTGGGA GAGAT	3290
6500	6522	CCCAGTCCTAGCTGCT GGCATCA	1522	TGATGCCAGCAGCTA GGACT	3291
6501	6523	CCAGTCCTAGCTGCTG GCATCAC	1523	GTGATGCCAGCAGCT AGGAC	3292
6506	6528	CCTAGCTGCTGGCATC ACTATAC	1524	GTATAGTGATGCCAG CAGCT	3293
6539	6561	CCGCAACCTCAACACC ACCTTCT	1525	AGAAGGTGGTGTGGA GGTTG	3294
6545	6567	CCTCAACACCACCTTC TTCGACC	1526	GGTCGAAGAAGGTGG TGTTG	3295
6553	6575	CCACCTTCTTCGACCC CGCCGGA	1527	TCCGGCGGGGTCGAA GAAGG	3296
6556	6578	CCTTCTTCGACCCCGC CGGAGGA	1528	TCCTCCGGCGGGGTC GAAGA	3297
6566	6588	CCCCGCCGGAGGAGGA GACCCCA	1529	TGGGGTCTCCTCCTCC GGCG	3298

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
6567	6589	CCCGCCGGAGGAGGA GACCCCAT	1530	ATGGGGTCTCCTCCT CCGGC	3299
6568	6590	CCGCCGGAGGAGGAG ACCCCAT	1531	AATGGGGTCTCCTCC TCCGG	3300
6571	6593	CCGGAGGAGGAGACC CCATTCTA	1532	TAGAATGGGGTCTCC TCCTC	3301
6584	6606	CCCCATTCTATACCAA CACCTAT	1533	ATAGGTGTTGGTATA GAATG	3302
6585	6607	CCCATTCTATACCAAC ACCTATT	1534	AATAGGTGTTGGTAT AGAAT	3303
6586	6608	CCATTCTATACCAACA CCTATTC	1535	GAATAGGTGTTGGTA TAGAA	3304
6596	6618	CCAACACCTATTCTGA TTTTTCG	1536	CGAAAAATCAGAATA GGTGT	3305
6602	6624	CCTATTCTGATTTTTCG GTCACC	1537	GGTGACCGAAAAATC AGAAT	3306
6623	6645	CCCTGAAGTTTATATT CTTATCC	1538	GGATAAGAATATAAA CTTCA	3307
6624	6646	CCTGAAGTTTATATTCT TATCCT	1539	AGGATAAGAATATAA ACTTC	3308
6644	6666	CCTACCAGGCTTCGGA ATAATCT	1540	AGATTATTCCGAAGC CTGGT	3309
6648	6670	CCAGGCTTCGGAATAA TCTCCCA	1541	TGGGAGATTATTCCG AAGCC	3310
6667	6689	CCCATATTGTAACCTA CTACTCC	1542	GGAGTAGTAAGTTAC AATAT	3311
6668	6690	CCATATTGTAACCTAC TACTCCG	1543	CGGAGTAGTAAGTTA CAATA	3312
6688	6710	CCGGAAAAAAGAAC CATTTGGA	1544	TCCAAATGGTTCTTTT TTTC	3313
6702	6724	CCATTTGGATACATAG GTATGGT	1545	ACCATACCTATGTAT CCAAA	3314
6749	6771	CCTAGGGTTTATCGTG TGAGCAC	1546	GTGCTCACACGATAA ACCCT	3315
6773	6795	CCATATATTTACAGTA GGAATAG	1547	CTATTCCTACTGTAA ATATA	3316
6820	6842	CCTCCGCTACCATAAT CATCGCT	1548	AGCGATGATTATGGT AGCGG	3317
6823	6845	CCGCTACCATAATCAT CGCTATC	1549	GATAGCGATGATTAT GGTAG	3318
6829	6851	CCATAATCATCGCTAT CCCCACC	1550	GGTGGGGATAGCGAT GATTA	3319
6845	6867	CCCCACCGGCGTCAAA GTATTTA	1551	TAAATACTTTGACGC CGGTG	3320
6846	6868	CCCACCGGCGTCAAAG	1552	CTAAATACTTTGACG	3321

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		TATTTAG		CCGGT	
6847	6869	CCACCGGCGTCAAAGT ATTTAGC	1553	GCTAAATACTTTGAC GCCGG	3322
6850	6872	CCGGCGTCAAAGTATT TAGCTGA	1554	TCAGCTAAATACTTT GACGC	3323
6877	6899	CCACACTCCACGGAAG CAATATG	1555	CATATTGCTTCCGTG GAGTG	3324
6884	6906	CCACGGAAGCAATATG AAATGAT	1556	ATCATTTTCATATTGCT TCCG	3325
6925	6947	CCCTAGGATTCATCTTT CTTTTC	1557	GAAAAGAAAGATGA ATCCTA	3326
6926	6948	CCTAGGATTCATCTTTC TTTTCA	1558	TGAAAAGAAAGATGA ATCCT	3327
6949	6971	CCGTAGGTGGCCTGAC TGGCATT	1559	AATGCCAGTCAGGCC ACCTA	3328
6959	6981	CCTGACTGGCATTGTA TTAGCAA	1560	TTGCTAATACAATGC CAGTC	3329
7027	7049	CCCACTTCCACTATGT CCTATCA	1561	TGATAGGACATAGTG GAAGT	3330
7028	7050	CCACTTCCACTATGTC CTATCAA	1562	TTGATAGGACATAGT GGAAG	3331
7034	7056	CCACTATGTCCTATCA ATAGGAG	1563	CTCCTATTGATAGGA CATAG	3332
7043	7065	CCTATCAATAGGAGCT GTATTTG	1564	CAAATACAGCTCCTA TTGAT	3333
7066	7088	CCATCATAGGAGGCTT CATTCAC	1565	GTGAATGAAGCCTCC TATGA	3334
7095	7117	CCCCTATTCTCAGGCT ACACCCT	1566	AGGGTGTAGCCTGAG AATAG	3335
7096	7118	CCCTATTCTCAGGCTA CACCCTA	1567	TAGGGTGTAGCCTGA GAATA	3336
7097	7119	CCTATTCTCAGGCTAC ACCCTAG	1568	CTAGGGTGTAGCCTG AGAAT	3337
7114	7136	CCCTAGACCAAACCTA CGCCAAA	1569	TTTGGCGTAGGTTTG GTCTA	3338
7115	7137	CCTAGACCAAACCTAC GCCAAAA	1570	TTTTGGCGTAGGTTTG GTCT	3339
7121	7143	CCAAACCTACGCCAAA ATCCATT	1571	AATGGATTTTGGCGT AGGTT	3340
7126	7148	CCTACGCCAAAATCCA TTTCACT	1572	AGTGAAATGGATTTT GGCGT	3341
7132	7154	CCAAAATCCATTTTAC TATCATA	1573	TATGATAGTGAAATG GATTT	3342
7139	7161	CCATTTCACTATCATAT TCATCG	1574	CGATGAATATGATAG TGAAA	3343

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
7181	7203	CCCACAACACTTTCTC GGCCTAT	1575	ATAGGCCGAGAAAGT GTTGT	3344
7182	7204	CCACAACACTTTCTCG GCCTATC	1576	GATAGGCCGAGAAAG TGTTG	3345
7199	7221	CCTATCCGGAATGCCC CGACGTT	1577	AACGTCGGGGCATTC CGGAT	3346
7204	7226	CCGGAATGCCCCGACG TTACTCG	1578	CGAGTAACGTCGGGG CATTC	3347
7212	7234	CCCCGACGTTACTCGG ACTACCC	1579	GGGTAGTCCGAGTAA CGTCG	3348
7213	7235	CCCGACGTTACTCGGA CTACCCC	1580	GGGGTAGTCCGAGTA ACGTC	3349
7214	7236	CCGACGTTACTCGGAC TACCCCG	1581	CGGGGTAGTCCGAGT AACGT	3350
7232	7254	CCCCGATGCATACACC ACATGAA	1582	TTCATGTGGTGTATG CATCG	3351
7233	7255	CCCGATGCATACACCA CATGAAA	1583	TTTCATGTGGTGTATG CATC	3352
7234	7256	CCGATGCATACACCAC ATGAAAC	1584	GTTTCATGTGGTGTAT GCAT	3353
7246	7268	CCACATGAAACATCCT ATCATCT	1585	AGATGATAGGATGTT TCATG	3354
7259	7281	CCTATCATCTGTAGGC TCATTCA	1586	TGAATGAGCCTACAG ATGAT	3355
7327	7349	CCTTCGCTTCGAAGCG AAAAGTC	1587	GACTTTTCGCTTCGA AGCGA	3356
7349	7371	CCTAATAGTAGAAGAA CCCTCCA	1588	TGGAGGGTTCTTCTA CTATT	3357
7365	7387	CCCTCCATAAACCTGG AGTGACT	1589	AGTCACTCCAGGTTT ATGGA	3358
7366	7388	CCTCCATAAACCTGGA GTGACTA	1590	TAGTCACTCCAGGTT TATGG	3359
7369	7391	CCATAAACCTGGAGTG ACTATAT	1591	ATATAGTCACTCCAG GTTTA	3360
7376	7398	CCTGGAGTGACTATAT GGATGCC	1592	GGCATCCATATAGTC ACTCC	3361
7397	7419	CCCCCACCCTACCAC ACATTCTG	1593	CGAATGTGTGGTAGG GTGGG	3362
7398	7420	CCCCCACCCTACCACA CATTCGA	1594	TCGAATGTGTGGTAG GGTGG	3363
7399	7421	CCCCACCCTACCACAC ATTCGAA	1595	TTCGAATGTGTGGTA GGGTG	3364
7400	7422	CCCACCCTACCACACA TTCGAAG	1596	CTTCGAATGTGTGGT AGGGT	3365
7401	7423	CCACCCTACCACACAT	1597	TCTTCGAATGTGTGG	3366

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		TCGAAGA		TAGGG	
7404	7426	CCCTACCACACATTTCG AAGAACC	1598	GGTTCTTCGAATGTG TGGTA	3367
7405	7427	CCTACCACACATTTCGA AGAACCC	1599	GGGTTCTTCGAATGT GTGGT	3368
7409	7431	CCACACATTTCGAAGAA CCCGTAT	1600	ATACGGGTTCTTCGA ATGTG	3369
7425	7447	CCCGTATACATAAAAT CTAGACA	1601	TGTCTAGATTTTATGT ATAC	3370
7426	7448	CCGTATACATAAAATC TAGACAA	1602	TTGTCTAGATTTTATG TATA	3371
7466	7488	CCCCCCTAAAGCTGGTT TCAAGCC	1603	GGCTTGAAACCAGCT TTGGG	3372
7467	7489	CCCCCCTAAAGCTGGTTT CAAGCCA	1604	TGGCTTGAAACCAGC TTTGG	3373
7468	7490	CCCCCCTAAAGCTGGTTTC AAGCCAA	1605	TTGGCTTGAAACCAG CTTTG	3374
7469	7491	CCCCCCTAAAGCTGGTTTCA AGCCAAC	1606	GTTGGCTTGAAACCA GCTTT	3375
7470	7492	CCCCCCTAAAGCTGGTTTCAA GCCAACC	1607	GGTTGGCTTGAAACC AGCTT	3376
7487	7509	CCAACCCCATGGCCTC CATGACT	1608	AGTCATGGAGGCCAT GGGGT	3377
7491	7513	CCCCATGGCCTCCATG ACTTTTT	1609	AAAAAGTCATGGAGG CCATG	3378
7492	7514	CCCCATGGCCTCCATGA CTTTTTTC	1610	GAAAAAGTCATGGAG GCCAT	3379
7493	7515	CCATGGCCTCCATGAC TTTTTCA	1611	TGAAAAAGTCATGGA GGCCA	3380
7499	7521	CCTCCATGACTTTTTTCA AAAAGG	1612	CCTTTTTGAAAAAGT CATGG	3381
7502	7524	CCATGACTTTTTTCAAA AAGGTAT	1613	ATACCTTTTTGAAAA AGTCA	3382
7533	7555	CCATTTTCATAACTTTGT CAAAGT	1614	ACTTTGACAAAGTTA TGAAA	3383
7573	7595	CCTATATATCTTAATG GCACATG	1615	CATGTGCCATTAAGA TATAT	3384
7626	7648	CCCCTATCATAGAAGA GCTTATC	1616	GATAAGCTCTTCTAT GATAG	3385
7627	7649	CCCTATCATAGAAGAG CTTATCA	1617	TGATAAGCTCTTCTAT GATA	3386
7628	7650	CCTATCATAGAAGAGC TTATCAC	1618	GTGATAAGCTCTTCT ATGAT	3387
7650	7672	CCTTTCATGATCACGC CCTCATA	1619	TATGAGGGCGTGATC ATGAA	3388

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
7665	7687	CCCTCATAATCATTTTCCTTATC	1620	GATAAGGAAAATGAT TATGA	3389
7666	7688	CCTCATAATCATTTTCC TTATCT	1621	AGATAAGGAAAATGA TTATG	3390
7681	7703	CCTTATCTGCTTCCTAG TCCTGT	1622	ACAGGACTAGGAAGC AGATA	3391
7693	7715	CCTAGTCCTGTATGCC CTTTTCC	1623	GGAAAAGGGCATACA GGACT	3392
7699	7721	CCTGTATGCCCTTTTCC TAACAC	1624	GTGTTAGGAAAAGGG CATAAC	3393
7707	7729	CCCTTTTCCTAACACTC ACAACA	1625	TGTTGTGAGTGTTAG GAAAA	3394
7708	7730	CCTTTTCCTAACACTCA CAACAA	1626	TTGTTGTGAGTGTTA GGAAA	3395
7714	7736	CCTAACACTCACAACA AAACATA	1627	TTAGTTTTGTTGTGAG TGTT	3396
7773	7795	CCGTCTGAACTATCCT GCCCGCC	1628	GGCGGGCAGGATAGT TCAGA	3397
7786	7808	CCTGCCCCGCCATCATC CTAGTCC	1629	GGACTAGGATGATGG CGGGC	3398
7790	7812	CCCGCCATCATCCTAG TCCTCAT	1630	ATGAGGACTAGGATG ATGGC	3399
7791	7813	CCGCCATCATCCTAGT CCTCATC	1631	GATGAGGACTAGGAT GATGG	3400
7794	7816	CCATCATCCTAGTCCT CATCGCC	1632	GGCGATGAGGACTAG GATGA	3401
7801	7823	CCTAGTCCTCATCGCC CTCCCAT	1633	ATGGGAGGGCGATGA GGACT	3402
7807	7829	CCTCATCGCCCTCCCA TCCCTAC	1634	GTAGGGATGGGAGGG CGATG	3403
7815	7837	CCCTCCCATCCCTACG CATCCTT	1635	AAGGATGCGTAGGGA TGGGA	3404
7816	7838	CCTCCCATCCCTACGC ATCCTTT	1636	AAAGGATGCGTAGGG ATGGG	3405
7819	7841	CCCATCCCTACGCATC CTTTACA	1637	TGTAAAGGATGCGTA GGGAT	3406
7820	7842	CCATCCCTACGCATCC TTTACAT	1638	ATGTAAAGGATGCGT AGGGA	3407
7824	7846	CCCTACGCATCCTTTA CATAACA	1639	TGTTATGTAAAGGAT GCGTA	3408
7825	7847	CCTACGCATCCTTTAC ATAACAG	1640	CTGTTATGTAAAGGA TCGCT	3409
7834	7856	CCTTTACATAACAGAC GAGGTCA	1641	TGACCTCGTCTGTTAT GTAA	3410
7862	7884	CCCTCCCTTACCATCA	1642	ATTGATTTGATGGTA	3411

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		AATCAAT		AGGGA	
7863	7885	CCTCCCTTACCATCAA ATCAATT	1643	AATTGATTTGATGGT AAGGG	3412
7866	7888	CCCTTACCATCAAATC AATTGGC	1644	GCCAATTGATTTGAT GGTAA	3413
7867	7889	CCTTACCATCAAATCA ATTGGCC	1645	GGCCAATTGATTTGA TGGTA	3414
7872	7894	CCATCAAATCAATTGG CCACCAA	1646	TTGGTGGCCAATTGA TTTGA	3415
7888	7910	CCACCAATGGTACTGA ACCTACG	1647	CGTAGGTTTCAGTACC ATTGG	3416
7891	7913	CCAATGGTACTGAACC TACGAGT	1648	ACTCGTAGGTTTCAGT ACCAT	3417
7905	7927	CCTACGAGTACACCGA CTACGGC	1649	GCCGTAGTCGGTGTA CTCGT	3418
7917	7939	CCGACTACGGCGGACT AATCTTC	1650	GAAGATTAGTCCGCC GTAGT	3419
7944	7966	CCTACATACTTCCCCC ATTATTC	1651	GAATAATGGGGGAAG TATGT	3420
7955	7977	CCCCATTATTCCTAG AACCAGG	1652	CCTGGTTCTAGGAAT AATGG	3421
7956	7978	CCCCATTATTCCTAGA ACCAGGC	1653	GCCTGGTTCTAGGAA TAATG	3422
7957	7979	CCCATTATTCCTAGAA CCAGGCG	1654	CGCCTGGTTCTAGGA ATAAT	3423
7958	7980	CCATTATTCCTAGAAC CAGGCGA	1655	TCGCCTGGTTCTAGG AATAA	3424
7966	7988	CCTAGAACCAGGCGAC CTGCGAC	1656	GTCGCAGGTCGCCTG GTTCT	3425
7973	7995	CCAGGCGACCTGCGAC TCCTTGA	1657	TCAAGGAGTCGCAGG TCGCC	3426
7981	8003	CCTGCGACTCCTTGAC GTTGACA	1658	TGTCAACGTCAAGGA GTCGC	3427
7990	8012	CCTTGACGTTGACAAT CGAGTAG	1659	CTACTCGATTGTCAA CGTCA	3428
8017	8039	CCCGATTGAAGCCCCC ATTCGTA	1660	TACGAATGGGGGCTT CAATC	3429
8018	8040	CCGATTGAAGCCCCCA TTCGTAT	1661	ATACGAATGGGGGCT TCAAT	3430
8028	8050	CCCCATTTCGTATAAT AATTACA	1662	TGTAATTATTATACG AATGG	3431
8029	8051	CCCCATTTCGTATAATA ATTACAT	1663	ATGTAATTATTATAC GAATG	3432
8030	8052	CCCATTCGTATAATAA TTACATC	1664	GATGTAATTATTATA CGAAT	3433

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
8031	8053	CCATTCGTATAATAAT TACATCA	1665	TGATGTAATTATTAT ACGAA	3434
8080	8102	CCCCACATTAGGCTTA AAAACAG	1666	CTGTTTTTAAGCCTAA TGTG	3435
8081	8103	CCCACATTAGGCTTAA AAACAGA	1667	TCTGTTTTTAAGCCTA ATGT	3436
8082	8104	CCACATTAGGCTTAAA AACAGAT	1668	ATCTGTTTTTAAGCCT AATG	3437
8111	8133	CCCGGACGTCTAAACC AAACCAC	1669	GTGGTTTGGTTTAGA CGTCC	3438
8112	8134	CCGGACGTCTAAACCA AACCCT	1670	AGTGGTTTGGTTTAG ACGTC	3439
8125	8147	CCAAACCACTTTCACC GCTACAC	1671	GTGTAGCGGTGAAAG TGGTT	3440
8130	8152	CCACTTTCACCGCTAC ACGACCG	1672	CGGTCGTGTAGCGGT GAAAG	3441
8139	8161	CCGCTACACGACCGGG GGTATAC	1673	GTATACCCCGGTCTG TGTAG	3442
8150	8172	CCGGGGGTATACTACG GTCAATG	1674	CATTGACCGTAGTAT ACCCC	3443
8194	8216	CCACAGTTTCATGCCC ATCGTCC	1675	GGACGATGGGCATGA AACTG	3444
8207	8229	CCCATCGTCCTAGAAT TAATTCC	1676	GGAATTAATTCTAGG ACGAT	3445
8208	8230	CCATCGTCCTAGAATT AATTCCC	1677	GGGAATTAATTCTAG GACGA	3446
8215	8237	CCTAGAATTAATTCCC CTAAAAA	1678	TTTTTAGGGGAATTA ATTCT	3447
8228	8250	CCCCTAAAAATCTTTG AAATAGG	1679	CCTATTTCAAAGATTT TTAG	3448
8229	8251	CCCTAAAAATCTTTGA AATAGGG	1680	CCCTATTTCAAAGAT TTTTA	3449
8230	8252	CCTAAAAATCTTTGAA ATAGGGC	1681	GCCCTATTTCAAAGA TTTTT	3450
8252	8274	CCCGTATTTACCCTAT AGCACCC	1682	GGGTGCTATAGGGTA AATAC	3451
8253	8275	CCGTATTTACCCTATA GCACCCC	1683	GGGGTGCTATAGGGT AAATA	3452
8262	8284	CCCTATAGCACCCCCT CTACCCC	1684	GGGGTAGAGGGGGTG CTATA	3453
8263	8285	CCTATAGCACCCCCTC TACCCC	1685	GGGGGTAGAGGGGGT GCTAT	3454
8272	8294	CCCCTCTACCCCCTCT AGAGCC	1686	GGCTCTAGAGGGGGT AGAGG	3455
8273	8295	CCCCTCTACCCCCTCTA	1687	GGGCTCTAGAGGGGG	3456

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		GAGCCC		TAGAG	
8274	8296	CCCTCTACCCCCTCTA GAGCCA	1688	TGGGCTCTAGAGGGG GTAGA	3457
8275	8297	CCTCTACCCCCTCTAG AGCCCAC	1689	GTGGGCTCTAGAGGG GGTAG	3458
8281	8303	CCCCCTCTAGAGCCCA CTGTAA	1690	TTTACAGTGGGCTCT AGAGG	3459
8282	8304	CCCCTCTAGAGCCCAC TGTAAG	1691	CTTTACAGTGGGCTC TAGAG	3460
8283	8305	CCCTCTAGAGCCCCT GTAAAGC	1692	GCTTTACAGTGGGCT CTAGA	3461
8284	8306	CCTCTAGAGCCCCTG TAAAGCT	1693	AGCTTTACAGTGGGC TCTAG	3462
8293	8315	CCCACTGTAAAGCTAA CTTAGCA	1694	TGCTAAGTTAGCTTT ACAGT	3463
8294	8316	CCACTGTAAAGCTAAC TTAGCAT	1695	ATGCTAAGTTAGCTT TACAG	3464
8320	8342	CCTTTTAAGTTAAAGA TTAAGAG	1696	CTCTTAATCTTTAACT TAAA	3465
8345	8367	CCAACACCTCTTTACA GTGAAAT	1697	ATTTCACTGTAAAGA GGTGT	3466
8351	8373	CCTCTTTACAGTGAAA TGCCCCA	1698	TGGGGCATTTCCTG TAAAG	3467
8369	8391	CCCCAACTAAATACTA CCGTATG	1699	CATACGGTAGTATTT AGTTG	3468
8370	8392	CCCAACTAAATACTAC CGTATGG	1700	CCATACGGTAGTATT TAGTT	3469
8371	8393	CCAATACTAAATACTACC GTATGGC	1701	GCCATACGGTAGTAT TTAGT	3470
8385	8407	CCGTATGGCCCACCAT AATTACC	1702	GGTAATTATGGTGGG CCATA	3471
8393	8415	CCCACCATAATTACCC CATACT	1703	AGTATGGGGGTAAATT ATGGT	3472
8394	8416	CCACCATAATTACCCC CATACTC	1704	GAGTATGGGGGTAAAT TATGG	3473
8397	8419	CCATAATTACCCCCAT ACTCCTT	1705	AAGGAGTATGGGGGT AATTA	3474
8406	8428	CCCCCATACTCCTTAC ACTATTC	1706	GAATAGTGTAAGGAG TATGG	3475
8407	8429	CCCCATACTCCTTACA CTATTCC	1707	GGAATAGTGTAAGGA GTATG	3476
8408	8430	CCCATACTCCTTACAC TATTCCT	1708	AGGAATAGTGTAAGG AGTAT	3477
8409	8431	CCATACTCCTTACACT ATTCCTC	1709	GAGGAATAGTGTAAG GAGTA	3478

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
8416	8438	CCTTACACTATTCCTCA TCACCC	1710	GGGTGATGAGGAATA GTGTA	3479
8428	8450	CCTCATCACCCAACTA AAAATAT	1711	ATATTTTTAGTTGGGT GATG	3480
8436	8458	CCCAACTAAAAATATT AACACA	1712	TGTGTTTAATATTTTT AGTT	3481
8437	8459	CCAACTAAAAATATTA AACACAA	1713	TTGTGTTTAATATTTTT TAGT	3482
8464	8486	CCACCTACCTCCCTCA CCAAAGC	1714	GCTTTGGTGAGGGAG GTAGG	3483
8467	8489	CCTACCTCCCTCACCA AAGCCCA	1715	TGGGCTTTGGTGAGG GAGGT	3484
8471	8493	CCTCCCTCACCAAAGC CCATAAA	1716	TTTATGGGCTTTGGTG AGGG	3485
8474	8496	CCCTCACCAAAGCCCA TAAAAAT	1717	ATTTTTATGGGCTTTG GTGA	3486
8475	8497	CCTCACCAAAGCCCAT AAAAATA	1718	TATTTTTATGGGCTTT GGTG	3487
8480	8502	CCAAAGCCCATAAAAA TAAAAAA	1719	TTTTTTATTTTTATGG GCTT	3488
8486	8508	CCCATAAAAATAAAAA ATTATAA	1720	TTATAATTTTTTATTT TTAT	3489
8487	8509	CCATAAAAAATAAAAAA TTATAAC	1721	GTTATAATTTTTTATT TTTA	3490
8513	8535	CCCTGAGAACCAAAT GAACGAA	1722	TTCGTTCATTTTGGTT CTCA	3491
8514	8536	CCTGAGAACCAAATG AACGAAA	1723	TTTCGTTCATTTTGGT TCTC	3492
8522	8544	CCAAAATGAACGAAA ATCTGTTC	1724	GAACAGATTTTCGTT CATT	3493
8558	8580	CCCCACAATCCTAGG CCTACCC	1725	GGGTAGGCCTAGGAT TGTGG	3494
8559	8581	CCCCACAATCCTAGGC CTACCCG	1726	CGGGTAGGCCTAGGA TTGTG	3495
8560	8582	CCCACAATCCTAGGCC TACCCGC	1727	GCGGGTAGGCCTAGG ATTGT	3496
8561	8583	CCACAATCCTAGGCCT ACCCGCC	1728	GGCGGGTAGGCCTAG GATTG	3497
8568	8590	CCTAGGCCTACCCGCC GCAGTAC	1729	GTACTGCGGCGGGTA GGCCT	3498
8574	8596	CCTACCCGCCGCAGTA CTGATCA	1730	TGATCAGTACTGCGG CGGGT	3499
8578	8600	CCCGCCGCAGTACTGA TCATTCT	1731	AGAATGATCAGTACT GCGGC	3500
8579	8601	CCGCCGCAGTACTGAT	1732	TAGAATGATCAGTAC	3501

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CATTCTA		TGCGG	
8582	8604	CCGCAGTACTGATCAT TCTATTT	1733	AAATAGAATGATCAG TACTG	3502
8605	8627	CCCCCTCTATTGATCCC CACCTC	1734	GAGGTGGGGATCAAT AGAGG	3503
8606	8628	CCCCTCTATTGATCCCC ACCTCC	1735	GGAGGTGGGGATCAA TAGAG	3504
8607	8629	CCCTCTATTGATCCCC ACCTCCA	1736	TGGAGGTGGGGATCA ATAGA	3505
8608	8630	CCTCTATTGATCCCCA CCTCCA	1737	TTGGAGGTGGGGATC AATAG	3506
8619	8641	CCCCACCTCCAAATAT CTCATCA	1738	TGATGAGATATTTGG AGGTG	3507
8620	8642	CCCACCTCCAAATATC TCATCAA	1739	TTGATGAGATATTTG GAGGT	3508
8621	8643	CCACCTCCAAATATCT CATCAAC	1740	GTTGATGAGATATTT GGAGG	3509
8624	8646	CCTCCAAATATCTCAT CAACAAC	1741	GTTGTTGATGAGATA TTTGG	3510
8627	8649	CCAAATATCTCATCAA CAACCGA	1742	TCGGTTGTTGATGAG ATATT	3511
8646	8668	CCGACTAATCACCACC CAACAAT	1743	ATTGTTGGGTGGTGA TTAGT	3512
8657	8679	CCACCCAACAATGACT AATCAAA	1744	TTTGATTAGTCATTGT TGGG	3513
8660	8682	CCCAACAATGACTAAT CAAACTA	1745	TAGTTTGATTAGTCAT TGTT	3514
8661	8683	CCAACAATGACTAATC AAACATA	1746	TTAGTTTGATTAGTCA TTGT	3515
8684	8706	CCTCAAAACAAATGAT AACCATA	1747	TATGGTTATCATTTGT TTTG	3516
8702	8724	CCATACACAACACTAA AGGACGA	1748	TCGTCCTTTAGTGTTG TGTA	3517
8726	8748	CCTGATCTCTTATACTA GTATCC	1749	GGATACTAGTATAAG AGATC	3518
8747	8769	CCTTAATCATTTTTTATT GCCACA	1750	TGTGGCAATAAAAAT GATTA	3519
8765	8787	CCACAACCTCCTCGGACTC	1751	GAGTCCGAGGAGGTT AGTTG	3520
8775	8797	CCTCCTCGGACTCCTG CCTCACT	1752	AGTGAGGCAGGAGTC CGAGG	3521
8778	8800	CCTCGGACTCCTGCCT CACTCAT	1753	ATGAGTGAGGCAGGA GTCCG	3522
8787	8809	CCTGCCTCACTCATTTA CACCAA	1754	TTGGTGTAATGAGT GAGGC	3523

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
8791	8813	CCTCACTCATTACAC CAACCAC	1755	GTGGTTGGTGTAAT GAGTG	3524
8806	8828	CCAACCACCCAACTAT CTATAAA	1756	TTTATAGATAGTTGG GTGGT	3525
8810	8832	CCACCCAACTATCTAT AAACCTA	1757	TAGGTTTATAGATAG TTGGG	3526
8813	8835	CCCAACTATCTATAAA CCTAGCC	1758	GGCTAGGTTTATAGA TAGTT	3527
8814	8836	CCAACTATCTATAAAC CTAGCCA	1759	TGGCTAGGTTTATAG ATAGT	3528
8829	8851	CCTAGCCATGGCCATC CCCTTAT	1760	ATAAGGGGATGGCCA TGGCT	3529
8834	8856	CCATGGCCATCCCCTT ATGAGCG	1761	CGCTCATAAGGGGAT GGCCA	3530
8840	8862	CCATCCCCTTATGAGC GGGCACA	1762	TGTGCCCGCTCATAA GGGGA	3531
8844	8866	CCCCTTATGAGCGGGC ACAGTGA	1763	TCACTGTGCCCGCTC ATAAG	3532
8845	8867	CCCTTATGAGCGGGCA CAGTGAT	1764	ATCACTGTGCCCGCT CATAA	3533
8846	8868	CCTTATGAGCGGGCAC AGTGATT	1765	AATCACTGTGCCCGC TCATA	3534
8897	8919	CCCTAGCCCCTTCTT ACCACAA	1766	TTGTGGTAAGAAGTG GGCTA	3535
8898	8920	CCTAGCCCCTTCTTA CCACAAG	1767	CTTGTGGTAAGAAGT GGGCT	3536
8903	8925	CCCACTTCTTACCACA AGGCACA	1768	TGTGCCTTGTGGTAA GAAGT	3537
8904	8926	CCACTTCTTACCACAA GGCACAC	1769	GTGTGCCTTGTGGTA AGAAG	3538
8914	8936	CCACAAGGCACACCTA CACCCCT	1770	AGGGGTGTAGGTGTG CCTTG	3539
8926	8948	CCTACACCCCTTATCC CCATACT	1771	AGTATGGGGATAAGG GGTGT	3540
8932	8954	CCCCTTATCCCCATACT AGTTAT	1772	ATAACTAGTATGGGG ATAAG	3541
8933	8955	CCCTTATCCCCATACT AGTTATT	1773	AATAACTAGTATGGG GATAA	3542
8934	8956	CCTTATCCCCATACTA GTTATTA	1774	TAATAACTAGTATGG GGATA	3543
8940	8962	CCCATACTAGTTATT ATCGAAA	1775	TTTCGATAATAACTA GTATG	3544
8941	8963	CCCATACTAGTTATTA TCGAAAC	1776	GTTTCGATAATAACT AGTAT	3545
8942	8964	CCATACTAGTTATTAT	1777	GGTTTCGATAATAAC	3546

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CGAAACC		TAGTA	
8963	8985	CCATCAGCCTACTCAT TCAACCA	1778	TGGTTGAATGAGTAG GCTGA	3547
8970	8992	CCTACTCATTCAACCA ATAGCCC	1779	GGGCTATTGGTTGAA TGAGT	3548
8983	9005	CCAATAGCCCTGGCCG TACGCCT	1780	AGGCGTACGGCCAGG GCTAT	3549
8990	9012	CCCTGGCCGTACGCCT AACCGCT	1781	AGCGGTTAGGCGTAC GGCCA	3550
8991	9013	CCTGGCCGTACGCCTA ACCGCTA	1782	TAGCGGTTAGGCGTA CGGCC	3551
8996	9018	CCGTACGCCTAACCGC TAACATT	1783	AATGTTAGCGGTTAG GCGTA	3552
9003	9025	CCTAACCGCTAACATT ACTGCAG	1784	CTGCAGTAATGTTAG CGGTT	3553
9008	9030	CCGCTAACATTACTGC AGGCCAC	1785	GTGGCCTGCAGTAAT GTTAG	3554
9027	9049	CCACCTACTCATGCAC CTAATTG	1786	CAATTAGGTGCATGA GTAGG	3555
9030	9052	CCTACTCATGCACCTA ATTGGAA	1787	TTCCAATTAGGTGCA TGAGT	3556
9042	9064	CCTAATTGGAAGCGCC ACCCTAG	1788	CTAGGGTGGCGCTTC CAATT	3557
9056	9078	CCACCCTAGCAATATC AACCATT	1789	AATGGTTGATATTGC TAGGG	3558
9059	9081	CCCTAGCAATATCAAC CATTAAC	1790	GTTAATGGTTGATAT TGCTA	3559
9060	9082	CCTAGCAATATCAACC ATTAACC	1791	GGTTAATGGTTGATA TTGCT	3560
9074	9096	CCATTAACCTTCCCTCT ACACTT	1792	AAGTGTAGAGGGAAG GTTAA	3561
9081	9103	CCTTCCCTCTACACTTA TCATCT	1793	AGATGATAAGTGTAG AGGGA	3562
9085	9107	CCCTCTACACTTATCAT CTTCAC	1794	GTGAAGATGATAAGT GTAGA	3563
9086	9108	CCTCTACACTTATCATC TTCACA	1795	TGTGAAGATGATAAG TGTAAG	3564
9129	9151	CCTAGAAATCGCTGTC GCCTTAA	1796	TTAAGGCGACAGCGA TTTCT	3565
9146	9168	CCTTAATCCAAGCCTA CGTTTTTC	1797	GAAAACGTAGGCTTG GATTA	3566
9153	9175	CCAAGCCTACGTTTTTC ACACTTC	1798	GAAGTGTGAAAACGT AGGCT	3567
9158	9180	CCTACGTTTTTCACACTT CTAGTA	1799	TACTAGAAGTGTGAA AACGT	3568

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
9183	9205	CCTCTACCTGCACGAC AACACAT	1800	ATGTGTTGTCGTGCA GGTAG	3569
9189	9211	CCTGCACGACAACACA TAATGAC	1801	GTCATTATGTGTTGTC GTGC	3570
9211	9233	CCCACCAATCACATGC CTATCAT	1802	ATGATAGGCATGTGA TTGGT	3571
9212	9234	CCACCAATCACATGCC TATCATA	1803	TATGATAGGCATGTG ATTGG	3572
9215	9237	CCAATCACATGCCTAT CATATAG	1804	CTATATGATAGGCAT GTGAT	3573
9226	9248	CCTATCATATAGTAAA ACCCAGC	1805	GCTGGGTTTTACTAT ATGAT	3574
9243	9265	CCCAGCCCATGACCCC TAACAGG	1806	CCTGTTAGGGGTCAT GGGCT	3575
9244	9266	CCAGCCCATGACCCCT AACAGGG	1807	CCCTGTTAGGGGTCA TGGGC	3576
9248	9270	CCCATGACCCCTAACA GGGGCC	1808	GGGCCCTGTTAGGG GTCAT	3577
9249	9271	CCATGACCCCTAACAG GGGCCCT	1809	AGGGCCCCTGTTAGG GGTCA	3578
9255	9277	CCCCTAACAGGGGCCC TCTCAGC	1810	GCTGAGAGGGGCCCT GTTAG	3579
9256	9278	CCCTAACAGGGGCCCT CTCAGCC	1811	GGCTGAGAGGGCCCC TGTTA	3580
9257	9279	CCTAACAGGGGCCCTC TCAGCCC	1812	GGGCTGAGAGGGCCC CTGTT	3581
9268	9290	CCCTCTCAGCCCTCCT AATGACC	1813	GGTCATTAGGAGGGC TGAGA	3582
9269	9291	CCTCTCAGCCCTCCTA ATGACCT	1814	AGGTCATTAGGAGGG CTGAG	3583
9277	9299	CCCTCCTAATGACCTC CGGCCTA	1815	TAGGCCGGAGGTCAT TAGGA	3584
9278	9300	CCTCCTAATGACCTCC GGCCTAG	1816	CTAGGCCGGAGGTCA TTAGG	3585
9281	9303	CCTAATGACCTCCGGC CTAGCCA	1817	TGGCTAGGCCGGAGG TCATT	3586
9289	9311	CCTCCGGCCTAGCCAT GTGATTT	1818	AAATCACATGGCTAG GCCGG	3587
9292	9314	CCGGCCTAGCCATGTG ATTTCAC	1819	GTGAAATCACATGGC TAGGC	3588
9296	9318	CCTAGCCATGTGATTT CACTTCC	1820	GGAAGTGAAATCACA TGGCT	3589
9301	9323	CCATGTGATTTCACTTC CACTCC	1821	GGAGTGGAAGTGAAA TCACA	3590
9317	9339	CCACTCCATAACGCTC	1822	GTATGAGGAGCGTTA	3591

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CTCATAC		TGGAG	
9322	9344	CCATAACGCTCCTCAT ACTAGGC	1823	GCCTAGTATGAGGAG CGTTA	3592
9332	9354	CCTCATACTAGGCCTA CTAACCA	1824	TGGTTAGTAGGCCTA GTATG	3593
9344	9366	CCTACTAACCAACACA CTAACCA	1825	TGGTTAGTGTGTTGG TTAGT	3594
9352	9374	CCAACACACTAACCAT ATACCAA	1826	TTGGTATATGGTTAG TGTGT	3595
9364	9386	CCATATACCAATGATG GCGCGAT	1827	ATCGCGCCATCATTG GTATA	3596
9371	9393	CCAATGATGGCGCGAT GTAACAC	1828	GTGTTACATCGCGCC ATCAT	3597
9407	9429	CCAAGGCCACCACACA CCACCTG	1829	CAGGTGGTGTGTGGT GGCCT	3598
9413	9435	CCACCACACACCACCT GTCCAAA	1830	TTTGGACAGGTGGTG TGTGG	3599
9416	9438	CCACACACCACCTGTC CAAAAAG	1831	CTTTTTGGACAGGTG GTGTG	3600
9423	9445	CCACCTGTCCAAAAG GCCTTCG	1832	CGAAGGCCTTTTTGG ACAGG	3601
9426	9448	CCTGTCCAAAAGGCC TTCGATA	1833	TATCGAAGGCCTTTTT GGAC	3602
9431	9453	CCAAAAGGCCTTCGA TACGGGA	1834	TCCCGTATCGAAGGC CTTTT	3603
9440	9462	CCTTCGATACGGGATA ATCCTAT	1835	ATAGGATTATCCCGT ATCGA	3604
9458	9480	CCTATTTATTACCTCAG AAGTTT	1836	AAACTTCTGAGGTAA TAAAT	3605
9469	9491	CCTCAGAAGTTTTTTTC TTCGCA	1837	TGCGAAGAAAAAAC TTCTG	3606
9505	9527	CCTTTTACCACTCCAG CCTAGCC	1838	GGCTAGGCTGGAGTG GTAAA	3607
9512	9534	CCACTCCAGCCTAGCC CCTACCC	1839	GGGTAGGGGCTAGGC TGGAG	3608
9517	9539	CCAGCCTAGCCCCTAC CCCCCAA	1840	TTGGGGGGGTAGGGGC TAGGC	3609
9521	9543	CCTAGCCCCTACCCCC CAATTAG	1841	CTAATTGGGGGGGTAG GGGCT	3610
9526	9548	CCCCTACCCCCCAATT AGGAGGG	1842	CCCTCCTAATTGGGG GGTAG	3611
9527	9549	CCCTACCCCCCAATTA GGAGGGC	1843	GCCCTCCTAATTGGG GGGTA	3612
9528	9550	CCTACCCCCCAATTAG GAGGGCA	1844	TGCCCTCCTAATTGG GGGGT	3613

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
9532	9554	CCCCCAATTAGGAGG GCACTGG	1845	CCAGTGCCCTCCTAA TTGGG	3614
9533	9555	CCCCCAATTAGGAGGG CACTGGC	1846	GCCAGTGCCCTCCTA ATTGG	3615
9534	9556	CCCCAATTAGGAGGGC ACTGGCC	1847	GGCCAGTGCCCTCCT AATTG	3616
9535	9557	CCCAATTAGGAGGGCA CTGGCCC	1848	GGGCCAGTGCCCTCC TAATT	3617
9536	9558	CCAATTAGGAGGGCAC TGGCCCC	1849	GGGGCCAGTGCCCTC CTAAT	3618
9555	9577	CCCCAACAGGCATCA CCCCGCT	1850	AGCGGGGTGATGCCT GTTGG	3619
9556	9578	CCCCAACAGGCATCAC CCCGCTA	1851	TAGCGGGGTGATGCC TGTTG	3620
9557	9579	CCCAACAGGCATCACC CCGCTAA	1852	TTAGCGGGGTGATGC CTGTT	3621
9558	9580	CCAACAGGCATCACCC CGCTAAA	1853	TTAGCGGGGTGATG CCTGT	3622
9571	9593	CCCGCTAAATCCCCT AGAAGTC	1854	GACTTCTAGGGGATT TAGCG	3623
9572	9594	CCCGCTAAATCCCCTA GAAGTCC	1855	GGACTTCTAGGGGAT TTAGC	3624
9573	9595	CCGCTAAATCCCCTAG AAGTCCC	1856	GGGACTTCTAGGGGA TTTAG	3625
9582	9604	CCCCTAGAAGTCCCAC TCCTAAA	1857	TTTAGGAGTGGGACT TCTAG	3626
9583	9605	CCCTAGAAGTCCCCTC CTAAAC	1858	GTTTAGGAGTGGGAC TTCTA	3627
9584	9606	CCTAGAAGTCCCCTC CTAAACA	1859	TGTTTAGGAGTGGGA CTTCT	3628
9593	9615	CCCACTCCTAAACACA TCCGTAT	1860	ATACGGATGTGTTTA GGAGT	3629
9594	9616	CCCACTCCTAAACACAT CCGTATT	1861	AATACGGATGTGTTT AGGAG	3630
9599	9621	CCTAAACACATCCGTA TTACTCG	1862	CGAGTAATACGGATG TGTTT	3631
9610	9632	CCGTATTACTCGCATC AGGAGTA	1863	TACTCCTGATGCGAG TAATA	3632
9640	9662	CCTGAGCTCACCATAG TCTAATA	1864	TATTAGACTATGGTG AGCTC	3633
9650	9672	CCATAGTCTAATAGAA AACAACC	1865	GGTTGTTTTCTATTAG ACTA	3634
9671	9693	CCGAAACCAAATAATT CAAGCAC	1866	GTGCTTGAATTATTG GTTT	3635
9677	9699	CCAAATAATTCAAGCA	1867	TAAGCAGTGCTTGAA	3636

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CTGCTTA		TTATT	
9727	9749	CCCTCCTACAAGCCTCAGAGTAC	1868	GTACTCTGAGGCTTG TAGGA	3637
9728	9750	CCTCCTACAAGCCTCAGAGTACT	1869	AGTACTCTGAGGCTT GTAGG	3638
9731	9753	CCTACAAGCCTCAGAGTACTTCG	1870	CGAAGTACTCTGAGG CTTGT	3639
9739	9761	CCTCAGAGTACTTCGAGTCTCCC	1871	GGGAGACTCGAAGTA CTCTG	3640
9759	9781	CCCTTCACCATTTCGACGGCAT	1872	ATGCCGTCGGAAATG GTGAA	3641
9760	9782	CCTTCACCATTTCGACGGCATC	1873	GATGCCGTCGGAAAT GGTGA	3642
9766	9788	CCATTTCCGACGGCATCTACGGC	1874	GCCGTAGATGCCGTC GGAAA	3643
9772	9794	CCGACGGCATCTACGGCTCAACA	1875	TGTTGAGCCGTAGAT GCCGT	3644
9805	9827	CCACAGGCTTCCACGGACTTCAC	1876	GTGAAGTCCGTGGAA GCCTG	3645
9815	9837	CCACGGACTTCACGTCATTATTG	1877	CAATAATGACGTGAA GTCCG	3646
9848	9870	CCTCACTATCTGCTTCA TCCGCC	1878	GGCGGATGAAGCAGA TAGTG	3647
9866	9888	CCGCCAACTAATATTTCACTTTA	1879	TAAAGTGAAATATTA GTTGG	3648
9869	9891	CCAATAATATTTTACATTTACAT	1880	ATGTAAAGTGAAATA TTAGT	3649
9892	9914	CCAAACATCACTTTGGCTTCGAA	1881	TTCGAAGCCAAAGTG ATGTT	3650
9916	9938	CCGCCGCCTGATACTG GCATTTT	1882	AAAATGCCAGTATCA GGCGG	3651
9919	9941	CCGCCTGATACTGGCA TTTTGTA	1883	TACAAAATGCCAGTA TCAGG	3652
9922	9944	CCTGATACTGGCATT TGTAGAT	1884	ATCTACAAAATGCCA GTATC	3653
9970	9992	CCATCTATTGATGAGG GTCTTAC	1885	GTAAGACCCTCATCA ATAGA	3654
10012	10034	CCGTAACTTCCAATT AACTAGT	1886	ACTAGTTAATTGGAA GTTAA	3655
10022	10044	CCAATTAAGTAGTTTT GACAACA	1887	TGTTGTCAAACTAG TTAAT	3656
10069	10091	CCTTAATTTTAATAATC AACACC	1888	GGTGTTGATTATTAA AATTA	3657
10090	10112	CCCTCCTAGCCTTACT ACTAATA	1889	TATTAGTAGTAAGGC TAGGA	3658

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
10091	10113	CCTCCTAGCCTTACTA CTAATAA	1890	TTATTAGTAGTAAGG CTAGG	3659
10094	10116	CCTAGCCTTACTACTA ATAATTA	1891	TAATTATTAGTAGTA AGGCT	3660
10099	10121	CCTTACTACTAATAAT TATTACA	1892	TGTAATAATTATTAG TAGTA	3661
10131	10153	CCACAACCTCAACGGCT ACATAGA	1893	TCTATGTAGCCGTTG AGTTG	3662
10159	10181	CCACCCCTTACGAGTG CGGCTTC	1894	GAAGCCGCACTCGTA AGGGG	3663
10162	10184	CCCCTTACGAGTGCGG CTTCGAC	1895	GTCGAAGCCGCACTC GTAAG	3664
10163	10185	CCCTTACGAGTGCGGC TTCGACC	1896	GGTCGAAGCCGCACT CGTAA	3665
10164	10186	CCTTACGAGTGCGGCT TCGACCC	1897	GGGTCGAAGCCGCAC TCGTA	3666
10184	10206	CCCTATATCCCCCGCC CGCGTCC	1898	GGACGCGGGCGGGG GATATA	3667
10185	10207	CCTATATCCCCCGCCC GCGTCCC	1899	GGGACGCGGGCGGG GGATAT	3668
10192	10214	CCCCGCCCCGCGTCCC TTCTCTCC	1900	GGAGAAAGGGACGC GGGCGG	3669
10193	10215	CCCCGCCCCGCGTCCCT TTCTCCA	1901	TGGAGAAAGGGACGC GGGCG	3670
10194	10216	CCCGCCCCGCGTCCCTT TCTCCAT	1902	ATGGAGAAAGGGAC GCGGGC	3671
10195	10217	CCGCCCCGCGTCCCTTT CTCCATA	1903	TATGGAGAAAGGGAC GCGGG	3672
10198	10220	CCCGCGTCCCTTTCTCC ATAAAA	1904	TTTTATGGAGAAAGG GACGC	3673
10199	10221	CCGCGTCCCTTTCTCCA TAAAT	1905	ATTTTATGGAGAAAG GGACG	3674
10205	10227	CCCTTTCTCCATAAAA TTCTTCT	1906	AGAAGAATTTTATGG AGAAA	3675
10206	10228	CCTTTCTCCATAAAATT CTCTT	1907	AAGAAGAATTTTATG GAGAA	3676
10213	10235	CCATAAAATTCTTCTT AGTAGCT	1908	AGCTACTAAGAAGAA TTTTA	3677
10240	10262	CCTTCTTATTATTTGAT CTAGAA	1909	TTCTAGATCAAATAA TAAGA	3678
10267	10289	CCCTCCTTTTACCCCTA CCATGA	1910	TCATGGTAGGGGTAA AAGGA	3679
10268	10290	CCTCCTTTTACCCCTAC CATGAG	1911	CTCATGGTAGGGGTA AAAGG	3680
10271	10293	CCTTTTACCCCTACCAT	1912	GGGCTCATGGTAGGG	3681

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		GAGCCC		GTAAA	
10278	10300	CCCCTACCATGAGCCC TACAAAC	1913	GTTTGTAGGGCTCAT GGTAG	3682
10279	10301	CCCTACCATGAGCCCT ACAAACA	1914	TGTTTGTAGGGCTCA TGGTA	3683
10280	10302	CCTACCATGAGCCCTA CAAACAA	1915	TTGTTTGTAGGGCTC ATGGT	3684
10284	10306	CCATGAGCCCTACAAA CAACTAA	1916	TTAGTTGTTTGTAGG GCTCA	3685
10291	10313	CCCTACAAACAATAA CCTGCCA	1917	TGGCAGGTTAGTTGT TTGTA	3686
10292	10314	CCTACAAACAATAAC CTGCCAC	1918	GTGGCAGGTTAGTTG TTTGT	3687
10307	10329	CCTGCCACTAATAGTT ATGTCAT	1919	ATGACATAACTATTA GTGGC	3688
10311	10333	CCACTAATAGTTATGT CATCCCT	1920	AGGGATGACATAACT ATTAG	3689
10330	10352	CCCTCTTATTAATCATC ATCCTA	1921	TAGGATGATGATTAA TAAGA	3690
10331	10353	CCTCTTATTAATCATCA TCCTAG	1922	CTAGGATGATGATTAA ATAAG	3691
10349	10371	CCTAGCCCTAAGTCTG GCCTATG	1923	CATAGGCCAGACTTA GGGCT	3692
10354	10376	CCCTAAGTCTGGCCTA TGAGTGA	1924	TCACTCATAGGCCAG ACTTA	3693
10355	10377	CCTAAGTCTGGCCTAT GAGTGAC	1925	GTCACTCATAGGCCA GACTT	3694
10366	10388	CCTATGAGTGACTACA AAAAGGA	1926	TCCTTTTTGTAGTCAC TCAT	3695
10399	10421	CCGAATTGGTATATAG TTTAAAC	1927	GTTTAAACTATATAC CAATT	3696
10466	10488	CCAAATGCCCTCATT TACATAA	1928	TTATGTAAATGAGGG GCATT	3697
10473	10495	CCCCTCATTTACATAA ATATTAT	1929	ATAATATTTATGTAA ATGAG	3698
10474	10496	CCCTCATTTACATAAA TATTATA	1930	TATAATATTTATGTA AATGA	3699
10475	10497	CCTCATTTACATAAAT ATTATAC	1931	GTATAATATTTATGT AAATG	3700
10507	10529	CCATCTCACTTCTAGG AATACTA	1932	TAGTATTCCTAGAAG TGAGA	3701
10544	10566	CCTCATATCCTCCCTAC TATGCC	1933	GGCATAGTAGGGAGG ATATG	3702
10552	10574	CCTCCCTACTATGCCT AGAAGGA	1934	TCCTTCTAGGCATAG TAGGG	3703

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
10555	10577	CCCTACTATGCCTAGA AGGAATA	1935	TATTCCTTCTAGGCAT AGTA	3704
10556	10578	CCTACTATGCCTAGAA GGAATAA	1936	TTATTCCTTCTAGGCA TAGT	3705
10565	10587	CCTAGAAGGAATAATA CTATCGC	1937	GCGATAGTATTATTC CTTCT	3706
10612	10634	CCCTCAACACCCACTC CCTCTTA	1938	TAAGAGGGAGTGGGT GTTGA	3707
10613	10635	CCTCAACACCCACTCC CTCTTAG	1939	CTAAGAGGGAGTGGG TGTTG	3708
10621	10643	CCCACTCCCTCTTAGC CAATATT	1940	AATATTGGCTAAGAG GGAGT	3709
10622	10644	CCACTCCCTCTTAGCC AATATTG	1941	CAATATTGGCTAAGA GGGAG	3710
10627	10649	CCCTCTTAGCCAATAT TGTGCCT	1942	AGGCACAATATTGGC TAAGA	3711
10628	10650	CCTCTTAGCCAATATT GTGCCTA	1943	TAGGCACAATATTGG CTAAG	3712
10636	10658	CCAATATTGTGCCTAT TGCCATA	1944	TATGGCAATAGGCAC AATAT	3713
10647	10669	CCTATTGCCATACTAG TCTTTGC	1945	GCAAAGACTAGTATG GCAAT	3714
10654	10676	CCATACTAGTCTTTGC CGCCTGC	1946	GCAGGCGGCAAAGAC TAGTA	3715
10669	10691	CCGCCTGCGAAGCAGC GGTGGGC	1947	GCCCACCGCTGCTTC GCAGG	3716
10672	10694	CCTGCGAAGCAGCGGT GGGCCTA	1948	TAGGCCCACCGCTGC TTCGC	3717
10691	10713	CCTAGCCCTACTAGTC TCAATCT	1949	AGATTGAGACTAGTA GGGCT	3718
10696	10718	CCCTACTAGTCTCAAT CTCCAAC	1950	GTTGGAGATTGAGAC TAGTA	3719
10697	10719	CCTACTAGTCTCAATC TCCAACA	1951	TGTTGGAGATTGAGA CTAGT	3720
10714	10736	CCAACACATATGGCCT AGACTAC	1952	GTAGTCTAGGCCATA TGTGT	3721
10727	10749	CCTAGACTACGTACAT AACCTAA	1953	TTAGGTTATGTACGT AGTCT	3722
10745	10767	CCTAAACCTACTCCAA TGCTAAA	1954	TTTAGCATTGGAGTA GGTTT	3723
10751	10773	CCTACTCCAATGCTAA AACTAAT	1955	ATTAGTTTTAGCATTG GAGT	3724
10757	10779	CCAATGCTAAAACTAA TCGTCCC	1956	GGGACGATTAGTTTT AGCAT	3725
10777	10799	CCCAACAATTATATTA	1957	GTGGTAGTAATATAA	3726

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CTACCAC		TTGTT	
10778	10800	CCAACAATTATATTAC TACCACT	1958	AGTGGTAGTAATATA ATTGT	3727
10796	10818	CCACTGACATGACTTT CCAAAAA	1959	TTTTTGGAAAGTCAT GTCAG	3728
10812	10834	CCAAAAAACACATAAT TTGAATC	1960	GATTCAAATTATGTG TTTT	3729
10842	10864	CCACCCACAGCCTAAT TATTAGC	1961	GCTAATAATTAGGCT GTGGG	3730
10845	10867	CCCACAGCCTAATTAT TAGCATC	1962	GATGCTAATAATTAG GCTGT	3731
10846	10868	CCACAGCCTAATTATT AGCATCA	1963	TGATGCTAATAATTA GGCTG	3732
10852	10874	CCTAATTATTAGCATC ATCCCTC	1964	GAGGGATGATGCTAA TAATT	3733
10870	10892	CCCTCTACTATTTTTTA ACCAAA	1965	TTTGGTTAAAAAATA GTAGA	3734
10871	10893	CCTCTACTATTTTTTAA CCAAAT	1966	ATTTGGTTAAAAAAT AGTAG	3735
10888	10910	CCAAATCAACAACAAC CTATTTA	1967	TAAATAGGTTGTTGT TGATT	3736
10903	10925	CCTATTTAGCTGTTCCC CAACCT	1968	AGGTTGGGGAACAGC TAAAT	3737
10917	10939	CCCCAACCTTTTCCTCC GACCCC	1969	GGGGTCGGAGGAAA AGGTTG	3738
10918	10940	CCCAACCTTTTCCTCCG ACCCCC	1970	GGGGGTCGGAGGAA AAGGTT	3739
10919	10941	CCAACCTTTTCCTCCG ACCCCCT	1971	AGGGGGTCGGAGGA AAAGGT	3740
10923	10945	CCTTTTCCTCCGACCCC CTAACA	1972	TGTTAGGGGGTCGGA GGAAA	3741
10929	10951	CCTCCGACCCCCTAAC AACCCCC	1973	GGGGGTTGTTAGGGG GTCGG	3742
10932	10954	CCGACCCCCTAACAAC CCCCCTC	1974	GAGGGGGGTTGTTAG GGGGT	3743
10936	10958	CCCCCTAACAACCCCC CTCCTAA	1975	TTAGGAGGGGGGTTG TTAGG	3744
10937	10959	CCCCTAACAACCCCCC TCCTAAT	1976	ATTAGGAGGGGGGTT GTTAG	3745
10938	10960	CCCTAACAACCCCCCT CTAATA	1977	TATTAGGAGGGGGGT TGTTA	3746
10939	10961	CCTAACAACCCCCCTC CTAATAC	1978	GTATTAGGAGGGGGG TTGTT	3747
10947	10969	CCCCCTCCTAATACT AACTACC	1979	GGTAGTTAGTATTAG GAGGG	3748

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
10948	10970	CCCCCTCCTAATACTA ACTACCT	1980	AGGTAGTTAGTATTA GGAGG	3749
10949	10971	CCCCTCCTAATACTAA CTACCTG	1981	CAGGTAGTTAGTATT AGGAG	3750
10950	10972	CCCTCCTAATACTAAC TACCTGA	1982	TCAGGTAGTTAGTAT TAGGA	3751
10951	10973	CCTCCTAATACTAACT ACCTGAC	1983	GTCAGGTAGTTAGTA TTAGG	3752
10954	10976	CCTAATACTAACTACC TGACTCC	1984	GGAGTCAGGTAGTTA GTATT	3753
10968	10990	CCTGACTCCTACCCCT CACAATC	1985	GATTGTGAGGGGTAG GAGTC	3754
10975	10997	CCTACCCCTCACAATC ATGGCAA	1986	TTGCCATGATTGTGA GGGGT	3755
10979	11001	CCCCTCACAATCATGG CAAGCCA	1987	TGGCTTGCCATGATT GTGAG	3756
10980	11002	CCCTCACAATCATGGC AAGCCAA	1988	TTGGCTTGCCATGATT GTGA	3757
10981	11003	CCTCACAATCATGGCA AGCCAAC	1989	GTTGGCTTGCCATGA TTGTG	3758
10999	11021	CCAACGCCACTTATCC AGTGAAC	1990	GTTCACTGGATAAGT GGCGT	3759
11005	11027	CCACTTATCCAGTGAA CCACTAT	1991	ATAGTGGTTCACTGG ATAAG	3760
11013	11035	CCAGTGAACCACTATC ACGAAAA	1992	TTTTTCGTGATAGTGGT TCAC	3761
11021	11043	CCACTATCACGAAAAA AACTCTA	1993	TAGAGTTTTTTTCGTG ATAG	3762
11044	11066	CCTCTCTATACTAATCT CCCTAC	1994	GTAGGGAGATTAGTA TAGAG	3763
11061	11083	CCCTACAAATCTCCTT AATTATA	1995	TATAATTAAGGAGAT TTGTA	3764
11062	11084	CCTACAAATCTCCTTA ATTATAA	1996	TTATAATTAAGGAGA TTTGT	3765
11073	11095	CCTTAATTATAACATT CACAGCC	1997	GGCTGTGAATGTTAT AATTA	3766
11094	11116	CCACAGAACTAATCAT ATTTTAT	1998	ATAAAATATGATTAG TTCTG	3767
11130	11152	CCACACTTATCCCCAC CTTGGCT	1999	AGCCAAGGTGGGGAT AAGTG	3768
11140	11162	CCCCACCTTGGCTATC ATCACCC	2000	GGGTGATGATAGCCA AGGTG	3769
11141	11163	CCCACCTTGGCTATCA TCACCCG	2001	CGGGTGATGATAGCC AAGGT	3770
11142	11164	CCACCTTGGCTATCAT	2002	TCGGGTGATGATAGC	3771

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CACCCGA		CAAGG	
11145	11167	CCTTGGCTATCATCAC CCGATGA	2003	TCATCGGGTGATGAT AGCCA	3772
11160	11182	CCCGATGAGGCAACCA GCCAGAA	2004	TTCTGGCTGGTTGCCT CATC	3773
11161	11183	CCGATGAGGCAACCAG CCAGAAC	2005	GTTCTGGCTGGTTGC CTCAT	3774
11173	11195	CCAGCCAGAACGCCTG AACGCAG	2006	CTGCGTTCAGGCGTT CTGGC	3775
11177	11199	CCAGAACGCCTGAACG CAGGCAC	2007	GTGCCTGCGTTCAGG CGTTC	3776
11185	11207	CCTGAACGCAGGCACA TACTTCC	2008	GGAAGTATGTGCCTG CGTTC	3777
11206	11228	CCTATTCTACACCCTA GTAGGCT	2009	AGCCTACTAGGGTGT AGAAT	3778
11217	11239	CCCTAGTAGGCTCCCT TCCCCTA	2010	TAGGGGAAGGGAGCC TACTA	3779
11218	11240	CCTAGTAGGCTCCCTT CCCCTAC	2011	GTAGGGGAAGGGAG CCTACT	3780
11229	11251	CCCTTCCCCTACTCATC GACTA	2012	TAGTGCGATGAGTAG GGGAA	3781
11230	11252	CCTTCCCCTACTCATCG CACTAA	2013	TTAGTGCGATGAGTA GGGGA	3782
11234	11256	CCCCTACTCATCGCAC TAATTTA	2014	TAAATTAGTGCGATG AGTAG	3783
11235	11257	CCCTACTCATCGCACT AATTTAC	2015	GTAAATTAGTGCGAT GAGTA	3784
11236	11258	CCTACTCATCGCACTA ATTTACA	2016	TGTAAATTAGTGCGA TGAGT	3785
11268	11290	CCCTAGGCTCACTAAA CATTCTA	2017	TAGAATGTTTAGTGA GCCTA	3786
11269	11291	CCTAGGCTCACTAAAC ATTCTAC	2018	GTAGAATGTTTAGTG AGCCT	3787
11307	11329	CCCAAGAACTATCAAA CTCCTGA	2019	TCAGGAGTTTGATAG TTCTT	3788
11308	11330	CCAAGAACTATCAAAC TCCTGAG	2020	CTCAGGAGTTTGATA GTTCT	3789
11325	11347	CCTGAGCCAACAACCTT AATATGA	2021	TCATATTAAGTTGTTG GCTC	3790
11331	11353	CCAACAACCTTAATATG ACTAGCT	2022	AGCTAGTCATATTAA GTTGT	3791
11381	11403	CCTCTTTACGGACTCC ACTTATG	2023	CATAAGTGGAGTCCG TAAAG	3792
11395	11417	CCACTTATGACTCCCT AAAGCCC	2024	GGGCTTTAGGGAGTC ATAAG	3793

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
11407	11429	CCCTAAAGCCCATGTCGAAGCCC	2025	GGGCTTCGACATGGGCTTTA	3794
11408	11430	CCTAAAGCCCATGTCGAAGCCCC	2026	GGGGCTTCGACATGGGCTTT	3795
11415	11437	CCCATGTCGAAGCCCCCATCGCT	2027	AGCGATGGGGGCTTCGACAT	3796
11416	11438	CCATGTCGAAGCCCCCATCGCTG	2028	CAGCGATGGGGGCTTCGACA	3797
11427	11449	CCCCATCGCTGGGTC AATAGTA	2029	TACTATTGACCCAGCGATGG	3798
11428	11450	CCCCATCGCTGGGTCAATAGTAC	2030	GTACTATTGACCCAGCGATG	3799
11429	11451	CCCATCGCTGGGTCAA TAGTACT	2031	AGTACTATTGACCCA GCGAT	3800
11430	11452	CCATCGCTGGGTCAATAGTACTT	2032	AAGTACTATTGACCCAGCGA	3801
11454	11476	CCGCAGTACTCTTAAACTAGGC	2033	GCCTAGTTTAAAGAGTACTG	3802
11494	11516	CCTCACACTCATTCTCAACCCCC	2034	GGGGGTTGAGAATGAGTGTG	3803
11512	11534	CCCCTGACAAAACACATAGCCT	2035	AGGCTATGTGTTTTGT CAGG	3804
11513	11535	CCCCTGACAAAACACATAGCCTA	2036	TAGGCTATGTGTTTTGTCAG	3805
11514	11536	CCCTGACAAAACACATAGCCTAC	2037	GTAGGCTATGTGTTTTGTCA	3806
11515	11537	CCTGACAAAACACATAGCCTACC	2038	GGTAGGCTATGTGTTTGTGTC	3807
11532	11554	CCTACCCCTTCCTTGTA CTATCC	2039	GGATAGTACAAGGAAGGGGT	3808
11536	11558	CCCCTTCCTTGTA CTATCCCTAT	2040	ATAGGGATAGTACAAGGAAG	3809
11537	11559	CCCTTCCTTGTA CTATCCCTATG	2041	CATAGGGATAGTACAAGGAAG	3810
11538	11560	CCTTCCTTGTA CTATCCCTATGA	2042	TCATAGGGATAGTACAAGGAAG	3811
11542	11564	CCTTGTA CTATCCCTATGAGGCA	2043	TGCCTCATAGGGATAGTACA	3812
11553	11575	CCCTATGAGGCATAATTATAACA	2044	TGTTATAATTATGCCTCATA	3813
11554	11576	CCTATGAGGCATAATTATAACAA	2045	TTGTTATAATTATGCCTCAT	3814
11580	11602	CCATCTGCCTACGACAACAGAC	2046	GTCTGTTTGTCTAGGCAGA	3815
11587	11609	CCTACGACAAACAGAC	2047	ATTTTAGGTCTGTTTG	3816

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CTAAAAT		TCGT	
11602	11624	CCTAAAATCGCTCATTGCATACT	2048	AGTATGCAATGAGCGATTTT	3817
11635	11657	CCACATAGCCCTCGTAGTAACAG	2049	CTGTTACTACGAGGGCTATG	3818
11643	11665	CCCTCGTAGTAACAGCATTCTC	2050	GAGAATGGCTGTTACTACGA	3819
11644	11666	CCTCGTAGTAACAGCCATTCTCA	2051	TGAGAATGGCTGTTA CTACG	3820
11658	11680	CCATTCTCATCCAAACCCCCTGA	2052	TCAGGGGGTTTGGATGAGAA	3821
11668	11690	CCAAACCCCCTGAAGCTTCACCG	2053	CGGTGAAGCTTCAGGGGGTT	3822
11673	11695	CCCCCTGAAGCTTCACCGGCGCA	2054	TGCGCCGGTGAAGCTTCAGG	3823
11674	11696	CCCCTGAAGCTTCACCGGCGCAG	2055	CTGCGCCGGTGAAGCTTCAG	3824
11675	11697	CCCTGAAGCTTCACCGGCGCAGT	2056	ACTGCGCCGGTGAAGCTTCA	3825
11676	11698	CCTGAAGCTTCACCGGCGCAGTC	2057	GACTGCGCCGGTGAA GCTTC	3826
11688	11710	CCGGCGCAGTCATTCTCATAATC	2058	GATTATGAGAATGAC TGCGC	3827
11712	11734	CCCACGGGCTTACATCTCATTA	2059	TAATGAGGATGTAAG CCCGT	3828
11713	11735	CCACGGGCTTACATCTCATTAC	2060	GTAATGAGGATGTAA GCCCG	3829
11727	11749	CCTCATTACTATTCTGCCTAGCA	2061	TGCTAGGCAGAATAG TAATG	3830
11743	11765	CCTAGCAAAC TCAAAC TACGAAC	2062	GTTCGTAGTTTGAGTT TGCT	3831
11788	11810	CCTCTCTCAAGGACTTCAAAC TC	2063	GAGTTTGAAGTCCTT GAGAG	3832
11815	11837	CCCACTAATAGCTTTTT GATGAC	2064	GTCATCAAAAAGCTA TTAGT	3833
11816	11838	CCACTAATAGCTTTTT GATGACT	2065	AGTCATCAAAAAGCT ATTAG	3834
11848	11870	CCTCGCTAACCTCGCC TTACCCC	2066	GGGGTAAGGCGAGGT TAGCG	3835
11857	11879	CCTCGCCTTACCCCC ACTATTA	2067	TAATAGTGGGGGGTA AGGCG	3836
11862	11884	CCTTACCCCCCACTATT AACCTA	2068	TAGGTTAATAGTGGG GGGTA	3837
11867	11889	CCCCCACTATTAACC TACTGGG	2069	CCCAGTAGGTTAATA GTGGG	3838

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
11868	11890	CCCCACTATTAACCT ACTGGGA	2070	TCCCAGTAGGTTAAT AGTGG	3839
11869	11891	CCCCACTATTAACCTA CTGGGAG	2071	CTCCCAGTAGGTTAA TAGTG	3840
11870	11892	CCCACTATTAACCTAC TGGGAGA	2072	TCTCCCAGTAGGTTA ATAGT	3841
11871	11893	CCACTATTAACCTACT GGGAGAA	2073	TTCTCCCAGTAGGTT AATAG	3842
11881	11903	CCTACTGGGAGAACTC TCTGTGC	2074	GCACAGAGAGTTCTC CCAGT	3843
11910	11932	CCACGTTCTCCTGATC AAATATC	2075	GATATTTGATCAGGA GAACG	3844
11919	11941	CCTGATCAAATATCAC TCTCCTA	2076	TAGGAGAGTGATATT TGATC	3845
11938	11960	CCTACTTACAGGACTC AACATAC	2077	GTATGTTGAGTCCTG TAAGT	3846
11970	11992	CCCTATACTCCCTCTAC ATATTT	2078	AAATATGTAGAGGGA GTATA	3847
11971	11993	CCTATACTCCCTCTAC ATATTTA	2079	TAAATATGTAGAGGG AGTAT	3848
11979	12001	CCCTCTACATATTTACC ACAACA	2080	TGTTGTGGTAAATAT GTAGA	3849
11980	12002	CCTCTACATATTTACC ACAACAC	2081	GTGTTGTGGTAAATA TGTAG	3850
11994	12016	CCACAACACAATGGGG CTCACTC	2082	GAGTGAGCCCCATTG TGTTG	3851
12018	12040	CCCACCACATTAACAA CATAAAA	2083	TTTTATGTTGTTAATG TGGT	3852
12019	12041	CCACCACATTAACAAC ATAAAAC	2084	GTTTTATGTTGTTAAT GTGG	3853
12022	12044	CCACATTAACAACATA AAACCCT	2085	AGGGTTTTATGTTGTT AATG	3854
12041	12063	CCCTCATTCACACGAG AAAACAC	2086	GTGTTTTCTCGTGTGA ATGA	3855
12042	12064	CCTCATTCACACGAGA AAACACC	2087	GGTGTTCCTCGTGTG AATG	3856
12063	12085	CCCTCATGTTTCATACA CCTATCC	2088	GGATAGGTGTATGAA CATGA	3857
12064	12086	CCTCATGTTTCATACAC CTATCCC	2089	GGGATAGGTGTATGA ACATG	3858
12079	12101	CCTATCCCCCATTCTCC TCCTAT	2090	ATAGGAGGAGAATGG GGGAT	3859
12084	12106	CCCCATTCTCCTCCTA TCCCTC	2091	GAGGGATAGGAGGA GAATGG	3860
12085	12107	CCCCATTCTCCTCCTAT	2092	TGAGGGATAGGAGGA	3861

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CCCTCA		GAATG	
12086	12108	CCCATTCTCCTCCTATC CCTCAA	2093	TTGAGGGATAGGAGG AGAAT	3862
12087	12109	CCATTCTCCTCCTATCC CTCAAC	2094	GTTGAGGGATAGGAG GAGAA	3863
12094	12116	CCTCCTATCCCTCAAC CCCGACA	2095	TGTCGGGGTTGAGGG ATAGG	3864
12097	12119	CCTATCCCTCAACCCC GACATCA	2096	TGATGTCGGGGTTGA GGGAT	3865
12102	12124	CCCTCAACCCCGACAT CATTACC	2097	GGTAATGATGTCGGG GTTGA	3866
12103	12125	CCTCAACCCCGACATC ATTACCG	2098	CGGTAATGATGTCGG GGTTG	3867
12109	12131	CCCCGACATCATTACC GGGTTTT	2099	AAAACCCGGTAATGA TGTCG	3868
12110	12132	CCCGACATCATTACCG GGTTTTTC	2100	GAAAACCCGGTAATG ATGTC	3869
12111	12133	CCGACATCATTACCGG GTTTTCC	2101	GGAAAACCCGGTAAT GATGT	3870
12123	12145	CCGGGTTTTCTCTTGT AAATAT	2102	ATATTTACAAGAGGA AAACC	3871
12132	12154	CCTCTTGTAATATAG TTTAACC	2103	GGTTAAACTATATTT ACAAG	3872
12153	12175	CCAAAACATCAGATTG TGAATCT	2104	AGATTCACAATCTGA TGTTT	3873
12194	12216	CCCCTTATTTACCGAG AAAGCTC	2105	GAGCTTTCTCGGTAA ATAAG	3874
12195	12217	CCCTTATTTACCGAGA AAGCTCA	2106	TGAGCTTTCTCGGTA AATAA	3875
12196	12218	CCTTATTTACCGAGAA AGCTCAC	2107	GTGAGCTTTCTCGGT AAATA	3876
12205	12227	CCGAGAAAGCTCACAA GAACTGC	2108	GCAGTTCTTGTGAGC TTTCT	3877
12237	12259	CCCCCATGTCTAACAA CATGGCT	2109	AGCCATGTTGTTAGA CATGG	3878
12238	12260	CCCCATGTCTAACAAC ATGGCTT	2110	AAGCCATGTTGTTAG ACATG	3879
12239	12261	CCCATGTCTAACAACA TGGCTTT	2111	AAAGCCATGTTGTTA GACAT	3880
12240	12262	CCATGTCTAACAACAT GGCTTTC	2112	GAAAGCCATGTTGTT AGACA	3881
12288	12310	CCATTGGTCTTAGGCC CCAAAAA	2113	TTTTTGGGGCCTAAG ACCAA	3882
12302	12324	CCCCAAAAATTTTGGT GCAACTC	2114	GAGTTGCACCAAAAT TTTTG	3883

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
12303	12325	CCCAAAAATTTTGGTGCAACTCC	2115	GGAGTTGCACCAAAA TTTT	3884
12304	12326	CCAAAAATTTTGGTGC AACTCCA	2116	TGGAGTTGCACCAAA ATTTT	3885
12324	12346	CCAAATAAAAAGTAATA ACCATGC	2117	GCATGGTTATTACTTT TATT	3886
12341	12363	CCATGCACACTACTAT AACCACC	2118	GGTGGTTATAGTAGT GTGCA	3887
12359	12381	CCACCCTAACCCTGAC TTCCCTA	2119	TAGGGAAGTCAGGGT TAGGG	3888
12362	12384	CCCTAACCCTGACTTC CCTAATT	2120	AATTAGGGAAGTCAG GGTTA	3889
12363	12385	CCTAACCCTGACTTCC CTAATTC	2121	GAATTAGGGAAGTCA GGGTT	3890
12368	12390	CCCTGACTTCCCTAATT CCCCCC	2122	GGGGGGAATTAGGGA AGTCA	3891
12369	12391	CCTGACTTCCCTAATTC CCCCCA	2123	TGGGGGGAATTAGGG AAGTC	3892
12377	12399	CCCTAATTCCCCCAT CTTACC	2124	GGTAAGGATGGGGGG AATTA	3893
12378	12400	CCTAATTCCCCCATC CTTACCA	2125	TGGTAAGGATGGGGG GAATT	3894
12385	12407	CCCCCATCCTTACCA CCTCGT	2126	ACGAGGGTGGTAAGG ATGGG	3895
12386	12408	CCCCCATCCTTACCAC CCTCGTT	2127	AACGAGGGTGGTAAG GATGG	3896
12387	12409	CCCCATCCTTACCACC CTCGTTA	2128	TAACGAGGGTGGTAA GGATG	3897
12388	12410	CCCATCCTTACCACCC TCGTTAA	2129	TTAACGAGGGTGGTA AGGAT	3898
12389	12411	CCATCCTTACCACCCT CGTTAAC	2130	GTTAACGAGGGTGGT AAGGA	3899
12393	12415	CCTTACCACCCTCGTT AACCCTA	2131	TAGGGTTAACGAGGG TGGTA	3900
12398	12420	CCACCCTCGTTAACCC TAACAAA	2132	TTTGTTAGGGTTAAC GAGGG	3901
12401	12423	CCCTCGTTAACCCTAA CAAAAAA	2133	TTTTTTGTTAGGGTTA ACGA	3902
12402	12424	CCTCGTTAACCCTAAC AAAAAAA	2134	TTTTTTTGTAGGGTT AACG	3903
12411	12433	CCCTAACAAAAAAAC TCATACC	2135	GGTATGAGTTTTTTTT GTTA	3904
12412	12434	CCTAACAAAAAAACT CATACCC	2136	GGGTATGAGTTTTTTT TGTT	3905
12432	12454	CCCCCATTATGTAAAA	2137	CAATGGATTTTACAT	3906

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		TCCATTG		AATGG	
12433	12455	CCCCATTATGTAAAATCCATTGT	2138	ACAATGGATTTTACA TAATG	3907
12434	12456	CCCATTATGTAAAATCCATTGTC	2139	GACAATGGATTTTAC ATAAT	3908
12435	12457	CCATTATGTAAAATCCATTGTCG	2140	CGACAATGGATTTTA CATAA	3909
12449	12471	CCATTGTCGCATCCACCTTTATT	2141	AATAAAGGTGGATGC GACAA	3910
12461	12483	CCACCTTTATTATCAGTCTCTTC	2142	GAAGAGACTGATAAT AAAGG	3911
12464	12486	CCTTTATTATCAGTCTCTTCCCC	2143	GGGGAAGAGACTGAT AATAA	3912
12483	12505	CCCCACAACAATATTCATGTGCC	2144	GGCACATGAATATTG TTGTG	3913
12484	12506	CCCACAACAATATTCA TGTGCCT	2145	AGGCACATGAATATT GTTGT	3914
12485	12507	CCACAACAATATTCATGTGCCTA	2146	TAGGCACATGAATAT TGTTG	3915
12504	12526	CCTAGACCAAGAAGTTATTATCT	2147	AGATAATAACTTCTT GGTCT	3916
12510	12532	CCAAGAAGTTATTATCTCGAACT	2148	AGTTCGAGATAATAA CTTCT	3917
12542	12564	CCACAACCCAAACAACCCAGCTC	2149	GAGCTGGGTTGTTTG GGTTG	3918
12548	12570	CCCAAACAACCCAGCTCTCCCTA	2150	TAGGGAGAGCTGGGT TGTTT	3919
12549	12571	CCAAACAACCCAGCTCTCCCTAA	2151	TTAGGGAGAGCTGGG TTGTT	3920
12557	12579	CCCAGCTCTCCCTAAGCTTCAA	2152	TTTGAAGCTTAGGGA GAGCT	3921
12558	12580	CCAGCTCTCCCTAAGCTTCAAAC	2153	GTTTGAAGCTTAGGG AGAGC	3922
12566	12588	CCCTAAGCTTCAAACCTAGACTAC	2154	GTAGTCTAGTTTGAA GCTTA	3923
12567	12589	CCTAAGCTTCAAACCTAGACTACT	2155	AGTAGTCTAGTTTGA AGCTT	3924
12593	12615	CCATAATATTCATCCCTGTAGCA	2156	TGCTACAGGGATGAA TATTA	3925
12606	12628	CCCTGTAGCATTGTTCGTTACAT	2157	ATGTAACGAACAATG CTACA	3926
12607	12629	CCTGTAGCATTGTTCGTTACATG	2158	CATGTAACGAACAAT GCTAC	3927
12632	12654	CCATCATAGAATTCTCCTGTGA	2159	TCACAGTGAGAATTC TATGA	3928

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
12669	12691	CCCAAACATTAATCAGTTCTTCA	2160	TGAAGAACTGATTAA TGTTT	3929
12670	12692	CCAAACATTAATCAGTTCTTCAA	2161	TTGAAGAACTGATTAA ATGTT	3930
12708	12730	CCTAATTACCATACTAATCTTAG	2162	CTAAGATTAGTATGGTAATT	3931
12716	12738	CCATACTAATCTTAGTTACCGCT	2163	AGCGGTAACATAAGATTAGTA	3932
12734	12756	CCGCTAACAACCTATTCCAACCTG	2164	CAGTTGGAATAGGTTGTTAG	3933
12744	12766	CCTATTCCAACCTGTTCATCGGCT	2165	AGCCGATGAACAGTTGGAAT	3934
12750	12772	CCAACTGTTTCATCGGCTGAGAGG	2166	CCTCTCAGCCGATGAACAGT	3935
12788	12810	CCTTCTTGCTCATCAGTTGATGA	2167	TCATCAACTGATGAGCAAGA	3936
12815	12837	CCCGAGCAGATGCCAACACAGCA	2168	TGCTGTGTTGGCATCTGCTC	3937
12816	12838	CCGAGCAGATGCCAACACAGCAG	2169	CTGCTGTGTTGGCATCTGCT	3938
12827	12849	CCAACACAGCAGCCATTCAAGCA	2170	TGCTTGAATGGCTGCTGTGT	3939
12839	12861	CCATTCAAGCAATCCTATACAAC	2171	GTTGTATAGGATTGCTTGAA	3940
12852	12874	CCTATACAACCGTATC GGCGATA	2172	TATCGCCGATACGGTTGTAT	3941
12861	12883	CCGTATCGGCGATATC GGTTTCA	2173	TGAAACCGATATCGCCGATA	3942
12885	12907	CCTCGCCTTAGCATGATTTATCC	2174	GGATAAATCATGCTAAGGCG	3943
12890	12912	CCTTAGCATGATTTATCCTACAC	2175	GTGTAGGATAAATCATGCTA	3944
12906	12928	CCTACACTCCAACCTCATGAGACC	2176	GGTCTCATGAGTTGGAGTGT	3945
12914	12936	CCAACTCATGAGACCCACAACAA	2177	TTGTTGTGGGTCTCATGAGT	3946
12927	12949	CCCACAACAAATAGCCCTTCTAA	2178	TTAGAAGGGCTATTTGTTGT	3947
12928	12950	CCACAACAAATAGCCCTTCTAAA	2179	TTTAGAAGGGCTATTGTTG	3948
12941	12963	CCCTTCTAAACGCTAATCCAAGC	2180	GCTTGGATTAGCGTTTAGAA	3949
12942	12964	CCTTCTAAACGCTAATCCAAGCC	2181	GGCTTGGATTAGCGTTTAGA	3950
12958	12980	CCAAGCCTCACCCAC	2182	CCTAGTAGTGGGGTG	3951

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		TACTAGG		AGGCT	
12963	12985	CCTCACCCCACTACTA GGCCTCC	2183	GGAGGCCTAGTAGTG GGGTG	3952
12968	12990	CCCCACTACTAGGCCT CCTCCTA	2184	TAGGAGGAGGCCTAG TAGTG	3953
12969	12991	CCCACTACTAGGCCTC CTCCTAG	2185	CTAGGAGGAGGCCTA GTAGT	3954
12970	12992	CCACTACTAGGCCTCC TCCTAGC	2186	GCTAGGAGGAGGCCT AGTAG	3955
12981	13003	CCTCCTCCTAGCAGCA GCAGGCA	2187	TGCCTGCTGCTGCTA GGAGG	3956
12984	13006	CCTCCTAGCAGCAGCA GGCAAAT	2188	ATTTGCCTGCTGCTGC TAGG	3957
12987	13009	CCTAGCAGCAGCAGGC AAATCAG	2189	CTGATTTGCCTGCTGC TGCT	3958
13010	13032	CCCAATTAGGTCTCCA CCCCTGA	2190	TCAGGGGTGGAGACC TAATT	3959
13011	13033	CCAATTAGGTCTCCAC CCCTGAC	2191	GTCAGGGGTGGAGAC CTAAT	3960
13023	13045	CCACCCCTGACTCCCC TCAGCCA	2192	TGGCTGAGGGGAGTC AGGGG	3961
13026	13048	CCCCTGACTCCCCTCA GCCATAG	2193	CTATGGCTGAGGGGA GTCAG	3962
13027	13049	CCCTGACTCCCCTCAG CCATAGA	2194	TCTATGGCTGAGGGG AGTCA	3963
13028	13050	CCTGACTCCCCTCAGC CATAGAA	2195	TTCTATGGCTGAGGG GAGTC	3964
13035	13057	CCCCTCAGCCATAGAA GGCCCCA	2196	TGGGGCCTTCTATGG CTGAG	3965
13036	13058	CCCTCAGCCATAGAAG GCCCCAC	2197	GTGGGGCCTTCTATG GCTGA	3966
13037	13059	CCTCAGCCATAGAAGG CCCCACC	2198	GGTGGGGCCTTCTAT GGCTG	3967
13043	13065	CCATAGAAGGCCCCAC CCCAGTC	2199	GACTGGGGTGGGGCC TTCTA	3968
13053	13075	CCCCACCCAGTCTCA GCCCTAC	2200	GTAGGGCTGAGACTG GGGTG	3969
13054	13076	CCCACCCAGTCTCAG CCCTACT	2201	AGTAGGGCTGAGACT GGGGT	3970
13055	13077	CCACCCAGTCTCAGC CCTACTC	2202	GAGTAGGGCTGAGAC TGGGG	3971
13058	13080	CCCAGTCTCAGCCCT ACTCCAC	2203	GTGGAGTAGGGCTGA GACTG	3972
13059	13081	CCCAGTCTCAGCCCTA CTCCACT	2204	AGTGGAGTAGGGCTG AGACT	3973

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
13060	13082	CCAGTCTCAGCCCTAC TCCACTC	2205	GAGTGGAGTAGGGCT GAGAC	3974
13070	13092	CCCTACTCCACTCAAG CACTATA	2206	TATAGTGCTTGAGTG GAGTA	3975
13071	13093	CCTACTCCACTCAAGC ACTATAG	2207	CTATAGTGCTTGAGT GGAGT	3976
13077	13099	CCACTCAAGCACTATA GTTGTAG	2208	CTACAACTATAGTGC TTGAG	3977
13119	13141	CCGCTTCCACCCCCTA GCAGAAA	2209	TTTCTGCTAGGGGGT GGAAG	3978
13125	13147	CCACCCCCTAGCAGAA AATAGCC	2210	GGCTATTTTCTGCTAG GGGG	3979
13128	13150	CCCCCTAGCAGAAAAT AGCCAC	2211	GTGGGCTATTTTCTGC TAGG	3980
13129	13151	CCCCTAGCAGAAAATA GCCCACT	2212	AGTGGGCTATTTTCT GCTAG	3981
13130	13152	CCCTAGCAGAAAATAG CCCCTA	2213	TAGTGGGCTATTTTCT GCTA	3982
13131	13153	CCTAGCAGAAAATAGC CCACTAA	2214	TTAGTGGGCTATTTTC TGCT	3983
13146	13168	CCACTAATCCAAACT CTAACAC	2215	GTGTTAGAGTTTGGA TTAGT	3984
13147	13169	CCACTAATCCAAACTC TAACACT	2216	AGTGTTAGAGTTTGGA ATTAG	3985
13155	13177	CCAAACTCTAACACTA TGCTTAG	2217	CTAAGCATAGTGTTA GAGTT	3986
13187	13209	CCACTCTGTTCGCAGC AGTCTGC	2218	GCAGACTGCTGCGAA CAGAG	3987
13211	13233	CCCTTACACAAAATGA CATCAAA	2219	TTTGATGTCATTTTGT GTAA	3988
13212	13234	CCTTACACAAAATGAC ATCAAAA	2220	TTTTGATGTCATTTTG TGTA	3989
13244	13266	CCTTCTCCACTTCAAGT CAACTA	2221	TAGTTGACTTGAAGT GGAGA	3990
13250	13272	CCACTTCAAGTCAACT AGGACTC	2222	GAGTCCTAGTTGACT TGAAG	3991
13296	13318	CCAACCACACCTAGCA TTCCTGC	2223	GCAGGAATGCTAGGT GTGGT	3992
13300	13322	CCACACCTAGCATTCC TGCACAT	2224	ATGTGCAGGAATGCT AGGTG	3993
13305	13327	CCTAGCATTCCTGCAC ATCTGTA	2225	TACAGATGTGCAGGA ATGCT	3994
13314	13336	CCTGCACATCTGTACC CACGCCT	2226	AGGCGTGGGTACAGA TGTGC	3995
13328	13350	CCCACGCCTTCTTCAA	2227	TATGGCTTTGAAGAA	3996

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		AGCCATA		GGCGT	
13329	13351	CCACGCCTTCTTCAAA GCCATAC	2228	GTATGGCTTTGAAGA AGGCG	3997
13334	13356	CCTTCTTCAAAGCCAT ACTATTT	2229	AAATAGTATGGCTTT GAAGA	3998
13346	13368	CCATACTATTTATGTG CTCCGGG	2230	CCCGGAGCACATAAA TAGTA	3999
13364	13386	CCGGGTCCATCATCCA CAACCTT	2231	AAGGTTGTGGATGAT GGACC	4000
13370	13392	CCATCATCCACAACCT TAACAAT	2232	ATTGTTAAGGTTGTG GATGA	4001
13377	13399	CCACAACCTTAACAAT GAACAAG	2233	CTTGTTCAATTGTTAAG GTTG	4002
13383	13405	CCTTAACAATGAACAA GATATTC	2234	GAATATCTTGTTCAAT GTTA	4003
13430	13452	CCATACCTCTCACTTC AACCTCC	2235	GGAGGTTGAAGTGAG AGGTA	4004
13435	13457	CCTCTCACTTCAACCTC CCTCAC	2236	GTGAGGGAGGTTGAA GTGAG	4005
13448	13470	CCTCCCTCACCATTGG CAGCCTA	2237	TAGGCTGCCAATGGT GAGGG	4006
13451	13473	CCCTCACCATTGGCAG CCTAGCA	2238	TGCTAGGCTGCCAAT GGTGA	4007
13452	13474	CCTCACCATTGGCAGC CTAGCAT	2239	ATGCTAGGCTGCCAA TGGTG	4008
13457	13479	CCATTGGCAGCCTAGC ATTAGCA	2240	TGCTAATGCTAGGCT GCCAA	4009
13467	13489	CCTAGCATTAGCAGGA ATACCTT	2241	AAGGTATTCCTGCTA ATGCT	4010
13486	13508	CCTTTCCTCACAGGTTT CTACTC	2242	GAGTAGAAACCTGTG AGGAA	4011
13491	13513	CCTCACAGGTTTCTAC TCCAAAG	2243	CTTTGGAGTAGAAAC CTGTG	4012
13508	13530	CCAAAGACCACATCAT CGAAACC	2244	GGTTTCGATGATGTG GTCTT	4013
13515	13537	CCACATCATCGAAACC GCAAACA	2245	TGTTTGCGGTTTCGAT GATG	4014
13529	13551	CCGCAAACATATCATA CACAAAC	2246	GTTTGTGTATGATAT GTTTG	4015
13553	13575	CCTGAGCCCTATCTAT TACTCTC	2247	GAGAGTAATAGATAG GGCTC	4016
13559	13581	CCCTATCTATTACTCTC ATCGCT	2248	AGCGATGAGAGTAAT AGATA	4017
13560	13582	CCTATCTATTACTCTCA TCGCTA	2249	TAGCGATGAGAGTAA TAGAT	4018

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
13583	13605	CCTCCCTGACAAGCGCCTATAGC	2250	GCTATAGGCGCTTGT CAGGG	4019
13586	13608	CCCTGACAAGCGCCTA TAGCACT	2251	AGTGCTATAGGCGCT TGTCA	4020
13587	13609	CCTGACAAGCGCCTAT AGCACTC	2252	GAGTGCTATAGGCGC TTGTC	4021
13598	13620	CCTATAGCACTCGAAT AATTCTT	2253	AAGAATTATTCGAGT GCTAT	4022
13625	13647	CCCTAACAGGTCAACC TCGCTTC	2254	GAAGCGAGGTTGACC TGTTA	4023
13626	13648	CCTAACAGGTCAACCT CGCTTCC	2255	GGAAGCGAGGTTGAC CTGTT	4024
13639	13661	CCTCGCTTCCCCACCCT TACTAA	2256	TTAGTAAGGGTGGGG AAGCG	4025
13647	13669	CCCCACCCTTACTAAC ATTAACG	2257	CGTTAATGTTAGTAA GGGTG	4026
13648	13670	CCCACCCTTACTAACA TTAACGA	2258	TCGTTAATGTTAGTA AGGGT	4027
13649	13671	CCACCCTTACTAACAT TAACGAA	2259	TTCGTTAATGTTAGTA AGGG	4028
13652	13674	CCCTTACTAACATTAA CGAAAAT	2260	ATTTTCGTTAATGTTA GTAA	4029
13653	13675	CCTTACTAACATTAAAC GAAAATA	2261	TATTTTCGTTAATGTT AGTA	4030
13677	13699	CCCCACCCTACTAAAC CCCATTA	2262	TAATGGGGTTTAGTA GGGTG	4031
13678	13700	CCCACCCTACTAAACC CCATTAA	2263	TTAATGGGGTTTAGT AGGGT	4032
13679	13701	CCACCCTACTAAACCC CATTAAA	2264	TTTAATGGGGTTTAG TAGGG	4033
13682	13704	CCCTACTAAACCCCAT TAAACGC	2265	GCGTTTAATGGGGTT TAGTA	4034
13683	13705	CCTACTAAACCCCAT TAAACGCC	2266	GGCGTTTAATGGGGT TTAGT	4035
13692	13714	CCCCATTAAACGCCTG GCAGCCG	2267	CGGCTGCCAGGCGTT TAATG	4036
13693	13715	CCCATTAACGCCTGG CAGCCGG	2268	CCGGCTGCCAGGCGT TTAAT	4037
13694	13716	CCATTAAACGCCTGGC AGCCGGA	2269	TCCGGCTGCCAGGCG TTAA	4038
13704	13726	CCTGGCAGCCGGAAGC CTATTTCG	2270	CGAATAGGCTTCCGG CTGCC	4039
13712	13734	CCGGAAGCCTATTTCG AGGATTT	2271	AAATCCTGCGAATAG GCTTC	4040
13719	13741	CCTATTCGCAGGATTT	2272	TAATGAGAAATCCTG	4041

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CTCATTA		CGAAT	
13754	13776	CCCCCGCATCCCCCTT CCAAACA	2273	TGTTTGG AAGGGGGA TGCGG	4042
13755	13777	CCCCGCATCCCCCTTC CAAACAA	2274	TTGTTTGG AAGGGGG ATGCG	4043
13756	13778	CCCGCATCCCCCTTCC AAACAAC	2275	GTTGTTTGG AAGGGG GATGC	4044
13757	13779	CCGCATCCCCCTTCCA AACAACA	2276	TGTTGTTTGG AAGGG GGATG	4045
13763	13785	CCCCCTTCCAAACAAC AATCCCC	2277	GGGGATTGTTGTTTG GAAGG	4046
13764	13786	CCCCTTCCAAACAACA ATCCCCC	2278	GGGGGATTGTTGTTT GGAAG	4047
13765	13787	CCCTTCCAAACAACAA TCCCCCT	2279	AGGGGGATTGTTGTT TGGA	4048
13766	13788	CCTTCCAAACAACAAT CCCCCTC	2280	GAGGGGGATTGTTGT TTGGA	4049
13770	13792	CCAAACAACAATCCCC CTCTACC	2281	GGTAGAGGGGGATTG TTGTT	4050
13782	13804	CCCCCTCTACCTAAAA CTCACAG	2282	CTGTGAGTTTTAGGT AGAGG	4051
13783	13805	CCCCTCTACCTAAAAC TCACAGC	2283	GCTGTGAGTTTTAGG TAGAG	4052
13784	13806	CCCTCTACCTAAAAC CACAGCC	2284	GGCTGTGAGTTTTAG GTAGA	4053
13785	13807	CCTCTACCTAAAAC ACAGCCC	2285	GGGCTGTGAGTTTTA GGTAG	4054
13791	13813	CCTAAAACCTCACAGCC CTCGCTG	2286	CAGCGAGGGCTGTGA GTTTT	4055
13805	13827	CCCTCGCTGTCACTTTC CTAGGA	2287	TCCTAGGAAAGTGAC AGCGA	4056
13806	13828	CCTCGCTGTCACTTTCC TAGGAC	2288	GTCCTAGGAAAGTGA CAGCG	4057
13821	13843	CCTAGGACTTCTAACA GCCCTAG	2289	CTAGGGCTGTTAGAA GTCCT	4058
13838	13860	CCCTAGACCTCAACTA CCTAACC	2290	GGTTAGGTAGTTGAG GTCTA	4059
13839	13861	CCTAGACCTCAACTAC CTAACCA	2291	TGGTTAGGTAGTTGA GGTCT	4060
13845	13867	CCTCAACTACCTAACC AACAAAC	2292	GTTTGTGTTAGGT AGTTG	4061
13854	13876	CCTAACCAACAACTT AAAATAA	2293	TTATTTTAAGTTTGTT GGTT	4062
13859	13881	CCAACAACTTAAAAT AAAATCC	2294	GGATTTTATTTTAAGT TTGT	4063

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
13880	13902	CCCCACTATGCACATT TTATTTC	2295	GAAATAAAATGTGCA TAGTG	4064
13881	13903	CCCCTATGCACATTT TATTCT	2296	AGAAATAAAATGTGC ATAGT	4065
13882	13904	CCACTATGCACATTTT ATTCTC	2297	GAGAAATAAAATGTG CATAG	4066
13904	13926	CCAACATACTCGGATT CTACCCT	2298	AGGGTAGAATCCGAG TATGT	4067
13923	13945	CCCTAGCATCACACAC CGCACAA	2299	TTGTGCGGTGTGTGA TGCTA	4068
13924	13946	CCTAGCATCACACACC GCACAAT	2300	ATTGTGCGGTGTGTG ATGCT	4069
13938	13960	CCGCACAATCCCCTAT CTAGGCC	2301	GGCCTAGATAGGGGA TTGTG	4070
13947	13969	CCCCTATCTAGGCCTT CTTACGA	2302	TCGTAAGAAGGCCTA GATAG	4071
13948	13970	CCCTATCTAGGCCTTCT TACGAG	2303	CTCGTAAGAAGGCCT AGATA	4072
13949	13971	CCTATCTAGGCCTTCTT ACGAGC	2304	GCTCGTAAGAAGGCC TAGAT	4073
13959	13981	CCTTCTTACGAGCCAA AACCTGC	2305	GCAGGTTTTGGCTCG TAAGA	4074
13971	13993	CCAAAACCTGCCCCTA CTCCTCC	2306	GGAGGAGTAGGGGC AGGTTT	4075
13977	13999	CCTGCCCCTACTCCTCC TAGACC	2307	GGTCTAGGAGGAGTA GGGGC	4076
13981	14003	CCCCTACTCCTCCTAG ACCTAAC	2308	GTTAGGTCTAGGAGG AGTAG	4077
13982	14004	CCCTACTCCTCCTAGA CCTAACC	2309	GGTTAGGTCTAGGAG GAGTA	4078
13983	14005	CCTACTCCTCCTAGAC CTAACCT	2310	AGGTTAGGTCTAGGA GGAGT	4079
13989	14011	CCTCCTAGACCTAACC TGACTAG	2311	CTAGTCAGGTTAGGT CTAGG	4080
13992	14014	CCTAGACCTAACCTGA CTAGAAA	2312	TTTCTAGTCAGGTTA GGTCT	4081
13998	14020	CCTAACCTGACTAGAA AAGCTAT	2313	ATAGCTTTTCTAGTCA GGTT	4082
14003	14025	CCTGACTAGAAAAGCT ATTACCT	2314	AGGTAATAGCTTTTC TAGTC	4083
14023	14045	CCTAAAACAATTTTAC AGCACCA	2315	TGGTGCTGTGAAATT GTTTT	4084
14043	14065	CCAAATCTCCACCTCC ATCATCA	2316	TGATGATGGAGGTGG AGATT	4085
14051	14073	CCACCTCCATCATCAC	2317	GGTTGAGGTGATGAT	4086

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CTCAACC		GGAGG	
14054	14076	CCTCCATCATCACCTC AACCCAA	2318	TTGGGTTGAGGTGAT GATGG	4087
14057	14079	CCATCATCACCTCAAC CCAAAAA	2319	TTTTTGGGTTGAGGT GATGA	4088
14066	14088	CCTCAACCCAAAAAGG CATAATT	2320	AATTATGCCTTTTTGG GTTG	4089
14072	14094	CCCAAAAAGGCATAAT TAAACTT	2321	AAGTTTAATTATGCC TTTTT	4090
14073	14095	CCAAAAAGGCATAATT AAACTTT	2322	AAAGTTTAATTATGC CTTTT	4091
14100	14122	CCTCTCTTTCTTCTTCC CACTCA	2323	TGAGTGGGAAGAAGA AAGAG	4092
14115	14137	CCCACTCATCCTAACC CTACTCC	2324	GGAGTAGGGTTAGGA TGAGT	4093
14116	14138	CCCACTCATCCTAACCC TACTCCT	2325	AGGAGTAGGGTTAGG ATGAG	4094
14124	14146	CCTAACCCTACTCCTA ATCACAT	2326	ATGTGATTAGGAGTA GGGTT	4095
14129	14151	CCCTACTCCTAATCAC ATAACCT	2327	AGGTTATGTGATTAG GAGTA	4096
14130	14152	CCTACTCCTAATCACA TAACCTA	2328	TAGGTTATGTGATTA GGAGT	4097
14136	14158	CCTAATCACATAACCT ATTCCCC	2329	GGGGAATAGGTTATG TGATT	4098
14149	14171	CCTATTCCCCCGAGCA ATCTCAA	2330	TTGAGATTGCTCGGG GGAAT	4099
14155	14177	CCCCCGAGCAATCTCA ATTACAA	2331	TTGTAATTGAGATTG CTCGG	4100
14156	14178	CCCCGAGCAATCTCAA TTACAAT	2332	ATTGTAATTGAGATT GCTCG	4101
14157	14179	CCCGAGCAATCTCAAT TACAATA	2333	TATTGTAATTGAGAT TGCTC	4102
14158	14180	CCGAGCAATCTCAATT ACAATAT	2334	ATATTGTAATTGAGA TTGCT	4103
14186	14208	CCAACAAACAATGTTC AACCAGT	2335	ACTGGTTGAACATTG TTTGT	4104
14204	14226	CCAGTAACTACTACTA ATCAACG	2336	CGTTGATTAGTAGTA GTTAC	4105
14227	14249	CCCATAATCATACAAA GCCCCCG	2337	CGGGGGCTTTGTATG ATTAT	4106
14228	14250	CCATAATCATACAAAG CCCCCGC	2338	GCGGGGGCTTTGTAT GATTA	4107
14244	14266	CCCCCGCACCAATAGG ATCCTCC	2339	GGAGGATCCTATTGG TGCGG	4108

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
14245	14267	CCCCGCACCAATAGGA TCCTCCC	2340	GGGAGGATCCTATTG GTGCG	4109
14246	14268	CCCGCACCAATAGGAT CCTCCCG	2341	CGGGAGGATCCTATT GGTGC	4110
14247	14269	CCGCACCAATAGGATC CTCCCGA	2342	TCGGGAGGATCCTAT TGGTG	4111
14252	14274	CCAATAGGATCCTCCC GAATCAA	2343	TTGATTCGGGAGGAT CCTAT	4112
14262	14284	CCTCCCGAATCAACCC TGACCCC	2344	GGGGTCAGGGTTGAT TCGGG	4113
14265	14287	CCCGAATCAACCCTGA CCCCTCT	2345	AGAGGGGTCAGGGTT GATTC	4114
14266	14288	CCGAATCAACCCTGAC CCCTCTC	2346	GAGAGGGGTCAGGGT TGATT	4115
14275	14297	CCCTGACCCCTCTCCTT CATAAA	2347	TTTATGAAGGAGAGG GGTCA	4116
14276	14298	CCTGACCCCTCTCCTTC ATAAAT	2348	ATTTATGAAGGAGAG GGGTC	4117
14281	14303	CCCCTCTCCTTCATAA ATTATTC	2349	GAATAATTTATGAAG GAGAG	4118
14282	14304	CCCTCTCCTTCATAAAT TATTCA	2350	TGAATAATTTATGAA GGAGA	4119
14283	14305	CCTCTCCTTCATAAATT ATTCAG	2351	CTGAATAATTTATGA AGGAG	4120
14288	14310	CCTTCATAAATTATTC AGCTTCC	2352	GGAAGCTGAATAATT TATGA	4121
14309	14331	CCTACACTATTAAAGT TTACCAC	2353	GTGGTAAACTTTAAT AGTGT	4122
14328	14350	CCACAACCACCACCCC ATCATAAC	2354	GTATGATGGGGTGGT GGTTC	4123
14334	14356	CCACCACCCCATCATA CTCTTTC	2355	GAAAGAGTATGATGG GGTGG	4124
14337	14359	CCACCCCATCATACTC TTTCACC	2356	GGTGAAAGAGTATGA TGGGG	4125
14340	14362	CCCCATCATACTCTTTC ACCCAC	2357	GTGGGTGAAAGAGTA TGATG	4126
14341	14363	CCCATCATACTCTTTC A CCCACA	2358	TGTGGGTGAAAGAGT ATGAT	4127
14342	14364	CCATCATACTCTTTCAC CCACAG	2359	CTGTGGGTGAAAGAG TATGA	4128
14358	14380	CCCACAGCACCAATCC TACCTCC	2360	GGAGGTAGGATTGGT GCTGT	4129
14359	14381	CCACAGCACCAATCCT ACCTCCA	2361	TGGAGGTAGGATTGG TGCTG	4130
14367	14389	CCAATCCTACCTCCAT	2362	GTTAGCGATGGAGGT	4131

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CGCTAAC		AGGAT	
14372	14394	CCTACCTCCATCGCTA ACCCAC	2363	GTGGGGTTAGCGATG GAGGT	4132
14376	14398	CCTCCATCGCTAACCC CACTAAA	2364	TTTAGTGGGGTTAGC GATGG	4133
14379	14401	CCATCGCTAACCCAC TAAAACA	2365	TGTTTTAGTGGGGTT AGCGA	4134
14389	14411	CCCCACTAAAACACTC ACCAAGA	2366	TCTTGGTGAGTGTTTT AGTG	4135
14390	14412	CCCACTAAAACACTCA CCAAGAC	2367	GTCTTGGTGAGTGTTT TAGT	4136
14391	14413	CCACTAAAACACTCAC CAAGACC	2368	GGTCTTGGTGAGTGT TTAG	4137
14406	14428	CCAAGACCTCAACCCC TGACCCC	2369	GGGGTCAGGGGTTGA GGTCT	4138
14412	14434	CCTCAACCCCTGACCC CCATGCC	2370	GGCATGGGGGTCAGG GGTTG	4139
14418	14440	CCCCTGACCCCCATGC CTCAGGA	2371	TCCTGAGGCATGGGG GTCAG	4140
14419	14441	CCCTGACCCCCATGCC TCAGGAT	2372	ATCCTGAGGCATGGG GGTCA	4141
14420	14442	CCTGACCCCCATGCCT CAGGATA	2373	TATCCTGAGGCATGG GGGTC	4142
14425	14447	CCCCCATGCCTCAGGA TACTCCT	2374	AGGAGTATCCTGAGG CATGG	4143
14426	14448	CCCCATGCCTCAGGAT ACTCCTC	2375	GAGGAGTATCCTGAG GCATG	4144
14427	14449	CCCATGCCTCAGGATA CTCCTCA	2376	TGAGGAGTATCCTGA GGCAT	4145
14428	14450	CCATGCCTCAGGATAC TCCTCAA	2377	TTGAGGAGTATCCTG AGGCA	4146
14433	14455	CCTCAGGATACTCCTC AATAGCC	2378	GGCTATTGAGGAGTA TCCTG	4147
14445	14467	CCTCAATAGCCATCGC TGTAGTA	2379	TACTACAGCGATGGC TATTG	4148
14454	14476	CCATCGCTGTAGTATA TCCAAAG	2380	CTTTGGATATACTAC AGCGA	4149
14471	14493	CCAAAGACAACCATCA TTCCCCC	2381	GGGGGAATGATGGTT GTCTT	4150
14481	14503	CCATCATTTCCCCTAA ATAAATT	2382	AATTTATTTAGGGGG AATGA	4151
14489	14511	CCCCCTAAATAAATTA AAAAAAC	2383	GTTTTTTTAATTTATT TAGG	4152
14490	14512	CCCCTAAATAAATTAA AAAAACT	2384	AGTTTTTTTAATTTAT TTAG	4153

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
14491	14513	CCCTAAATAAATTTAAA AAAATA	2385	TAGTTTTTTTAATTTA TTTA	4154
14492	14514	CCTAAATAAATTTAAAA AAACTAT	2386	ATAGTTTTTTTAATTT ATTT	4155
14519	14541	CCCATATAACCTCCCC CAAATT	2387	AATTTTGGGGGAGGT TATAT	4156
14520	14542	CCATATAACCTCCCC AAAATTC	2388	GAATTTTGGGGGAGG TTATA	4157
14528	14550	CCTCCCCCAAATTC GAATAAT	2389	ATTATTCTGAATTTG GGGG	4158
14531	14553	CCCCCAAATTCAGAA TAATAAC	2390	GTTATTATTCTGAATT TTGG	4159
14532	14554	CCCCCAAATTCAGAAT AATAACA	2391	TGTTATTATTCTGAAT TTTG	4160
14533	14555	CCCAAATTCAGAATA ATAACAC	2392	GTGTTATTATTCTGAA TTTT	4161
14534	14556	CCAAAATTCAGAATA TAACACA	2393	TGTGTTATTATTCTGA ATTT	4162
14557	14579	CCCGACCACACCGCTA ACAATCA	2394	TGATTGTTAGCGGTG TGGTC	4163
14558	14580	CCGACCACACCGCTAA CAATCAA	2395	TTGATTGTTAGCGGT GTGGT	4164
14562	14584	CCACACCGCTAACAAT CAATACT	2396	AGTATTGATTGTTAG CGGTG	4165
14567	14589	CCGCTAACAATCAATA CTAAACC	2397	GGTTTAGTATTGATT GTTAG	4166
14588	14610	CCCCATAAATAGGAG AAGGCTT	2398	AAGCCTTCTCCTATTT ATGG	4167
14589	14611	CCCCATAAATAGGAGA AGGCTTA	2399	TAAGCCTTCTCCTATT TATG	4168
14590	14612	CCCATAAATAGGAGAA GGCTTAG	2400	CTAAGCCTTCTCCTAT TTAT	4169
14591	14613	CCATAAATAGGAGAAG GCTTAGA	2401	TCTAAGCCTTCTCCTA TTTA	4170
14620	14642	CCCCACAAACCCCAT ACTAAAC	2402	GTTTAGTAATGGGGT TTGTG	4171
14621	14643	CCCACAAACCCCATTA CTAAACC	2403	GGTTTAGTAATGGGG TTTGT	4172
14622	14644	CCACAAACCCCATTA TAAACCC	2404	GGGTTTAGTAATGGG GTTTG	4173
14629	14651	CCCATTACTAAACCC ACACTCA	2405	TGAGTGTGGGTTTAG TAATG	4174
14630	14652	CCATTACTAAACCCA CACTCAA	2406	TTGAGTGTGGGTTTA GTAAT	4175
14631	14653	CCATTACTAAACCCAC	2407	GTTGAGTGTGGGTTT	4176

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		ACTCAAC		AGTAA	
14642	14664	CCCACACTCAACAGAA ACAAAGC	2408	GCTTTGTTTCTGTTGA GTGT	4177
14643	14665	CCACACTCAACAGAAA CAAAGCA	2409	TGCTTTGTTTCTGTTG AGTG	4178
14694	14716	CCACGACCAATGATAT GAAAAAC	2410	GTTTTTCATATCATTG GTCG	4179
14700	14722	CCAATGATATGAAAAA CCATCGT	2411	ACGATGGTTTTTCAT ATCAT	4180
14716	14738	CCATCGTTGTATTTCA ACTACAA	2412	TTGTAGTTGAAATAC AACGA	4181
14744	14766	CCAATGACCCCAATAC GCAAAAC	2413	GTTTTGCGTATTGGG GTCAT	4182
14751	14773	CCCAATACGCAAAAC TAACCCC	2414	GGGGTTAGTTTTGCG TATTG	4183
14752	14774	CCCAATACGCAAAACT AACCCCC	2415	GGGGGTTAGTTTTGC GTATT	4184
14753	14775	CCAATACGCAAAACTA ACCCCCT	2416	AGGGGGTTAGTTTTG CGTAT	4185
14770	14792	CCCCCTAATAAAATTA ATTAACC	2417	GGTTAATTAATTTTAT TAGG	4186
14771	14793	CCCCTAATAAAATTAA TTAACCA	2418	TGGTTAATTAATTTTA TTAG	4187
14772	14794	CCCTAATAAAATTAAT TAACCAC	2419	GTGGTTAATTAATTTT ATTA	4188
14773	14795	CCTAATAAAATTAATT AACCACT	2420	AGTGGTTAATTAATT TTATT	4189
14791	14813	CCACTCATTTCATCGAC CTCCCCA	2421	TGGGGAGGTCGATGA ATGAG	4190
14806	14828	CCTCCCCACCCCATCC AACATCT	2422	AGATGTTGGATGGGG TGGGG	4191
14809	14831	CCCCACCCCATCCAAC ATCTCCG	2423	CGGAGATGTTGGATG GGGTG	4192
14810	14832	CCCACCCCATCCAACA TCTCCGC	2424	GCGGAGATGTTGGAT GGGGT	4193
14811	14833	CCACCCCATCCAACAT CTCCGCA	2425	TGCGGAGATGTTGGA TGGGG	4194
14814	14836	CCCATCCAACATCTC CGCATGA	2426	TCATGCGGAGATGTT GGATG	4195
14815	14837	CCCATCCAACATCTCC GCATGAT	2427	ATCATGCGGAGATGT TGGAT	4196
14816	14838	CCATCCAACATCTCCG CATGATG	2428	CATCATGCGGAGATG TTGGA	4197
14820	14842	CCAACATCTCCGCATG ATGAAAC	2429	GTTTCATCATGCGGA GATGT	4198

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
14829	14851	CCGCATGATGAACTT CGGCTCA	2430	TGAGCCGAAGTTTCA TCATG	4199
14854	14876	CCTTGGCGCCTGCCTG ATCCTCC	2431	GGAGGATCAGGCAGG CGCCA	4200
14862	14884	CCTGCCTGATCCTCCA AATCACC	2432	GGTGATTTGGAGGAT CAGGC	4201
14866	14888	CCTGATCCTCCAAATC ACCACAG	2433	CTGTGGTGATTTGGA GGATC	4202
14872	14894	CCTCCAAATCACCACA GGACTAT	2434	ATAGTCCTGTGGTGA TTTGG	4203
14875	14897	CCAAATCACCACAGGA CTATTCC	2435	GGAATAGTCCTGTGG TGATT	4204
14883	14905	CCACAGGACTATTCCT AGCCATG	2436	CATGGCTAGGAATAG TCCTG	4205
14896	14918	CCTAGCCATGCACTAC TCACCAG	2437	CTGGTGAGTAGTGCA TGGCT	4206
14901	14923	CCATGCACTACTCACC AGACGCC	2438	GGCGTCTGGTGAGTA GTGCA	4207
14915	14937	CCAGACGCCTCAACCG CCTTTTC	2439	GAAAAGGCGGTTGAG GCGTC	4208
14922	14944	CCTCAACCGCCTTTTC ATCAATC	2440	GATTGATGAAAAGGC GGTTG	4209
14928	14950	CCGCCTTTTCATCAATC GCCCAC	2441	GTGGGCGATTGATGA AAAGG	4210
14931	14953	CCTTTTCATCAATCGCC CACATC	2442	GATGTGGGCGATTGA TGAAA	4211
14946	14968	CCCACATCACTCGAGA CGTAAAT	2443	ATTTACGTCTCGAGT GATGT	4212
14947	14969	CCACATCACTCGAGAC GTAAATT	2444	AATTTACGTCTCGAG TGATG	4213
14983	15005	CCGCTACCTTCACGCC AATGGCG	2445	CGCCATTGGCGTGAA GGTAG	4214
14989	15011	CCTTCACGCCAATGGC GCCTCAA	2446	TTGAGGCGCCATTGG CGTGA	4215
14997	15019	CCAATGGCGCCTCAAT ATTCTTT	2447	AAAGAATATTGAGGC GCCAT	4216
15006	15028	CCTCAATATTCTTTATC TGCCTC	2448	GAGGCAGATAAAGA ATATTG	4217
15025	15047	CCTCTTCCTACACATC GGGCGAG	2449	CTCGCCCGATGTGTA GGAAG	4218
15031	15053	CCTACACATCGGGCGA GGCCTAT	2450	ATAGGCCTCGCCCGA TGTGT	4219
15049	15071	CCTATATTACGGATCA TTTCTCT	2451	AGAGAAATGATCCGT AATAT	4220
15081	15103	CCTGAAACATCGGCAT	2452	GAGGATAATGCCGAT	4221

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		TATCCTC		GTTTC	
15100	15122	CCTCCTGCTTGCAACT ATAGCAA	2453	TTGCTATAGTTGCAA GCAGG	4222
15103	15125	CCTGCTTGCAACTATA GCAACAG	2454	CTGTTGCTATAGTTGC AAGC	4223
15126	15148	CCTTCATAGGCTATGT CCTCCCG	2455	CGGGAGGACATAGCC TATGA	4224
15142	15164	CCTCCCGTGAGGCCAA ATATCAT	2456	ATGATATTTGGCCTC ACGGG	4225
15145	15167	CCCGTGAGGCCAAATA TCATTCT	2457	AGAATGATATTTGGC CTCAC	4226
15146	15168	CCGTGAGGCCAAATAT CATTCTG	2458	CAGAATGATATTTGG CCTCA	4227
15154	15176	CCAAATATCATTCTGA GGGGCCA	2459	TGGCCCCTCAGAATG ATATT	4228
15174	15196	CCACAGTAATTACAAA CTTACTA	2460	TAGTAAGTTTGTAAT TACTG	4229
15198	15220	CCGCCATCCCATACAT TGGGACA	2461	TGTCCCAATGTATGG GATGG	4230
15201	15223	CCATCCCATACATTGG GACAGAC	2462	GTCTGTCCCAATGTA TGGGA	4231
15205	15227	CCCATACATTGGGACA GACCTAG	2463	CTAGGTCTGTCCCAA TGTAT	4232
15206	15228	CCATACATTGGGACAG ACCTAGT	2464	ACTAGGTCTGTCCCA ATGTA	4233
15223	15245	CCTAGTTCAATGAATC TGAGGAG	2465	CTCCTCAGATTCATTG AACT	4234
15263	15285	CCCACCCTCACACGAT TCTTTAC	2466	GTAAAGAATCGTGTG AGGGT	4235
15264	15286	CCACCCTCACACGATT CTTTACC	2467	GGTAAAGAATCGTGT GAGGG	4236
15267	15289	CCCTCACACGATTCTTT ACCTTT	2468	AAAGGTAAAGAATCG TGTGA	4237
15268	15290	CCTCACACGATTCTTT ACCTTTC	2469	GAAAGGTAAAGAATC GTGTG	4238
15285	15307	CCTTTCACTTCATCTTG CCCTTC	2470	GAAGGGCAAGATGA AGTGAA	4239
15302	15324	CCCTTCATTATTGCAG CCCTAGC	2471	GCTAGGGCTGCAATA ATGAA	4240
15303	15325	CCTTCATTATTGCAGC CCTAGCA	2472	TGCTAGGGCTGCAAT AATGA	4241
15318	15340	CCCTAGCAAACTCCA CCTCCTA	2473	TAGGAGGTGGAGTGT TGCTA	4242
15319	15341	CCTAGCAAACTCCAC CTCCTAT	2474	ATAGGAGGTGGAGTG TTGCT	4243

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
15331	15353	CCACCTCCTATTCTTGCACGAAA	2475	TTTCGTGCAAGAATAGGAGG	4244
15334	15356	CCTCCTATTCTTGCACGAAACGG	2476	CCGTTTCGTGCAAGATAGG	4245
15337	15359	CCTATTCTTGCACGAAACGGGAT	2477	ATCCCGTTTCGTGCAAGAAT	4246
15367	15389	CCCCCTAGGAATCACCTCCCATTC	2478	AATGGGAGGTGATTCTAGG	4247
15368	15390	CCCCTAGGAATCACCTCCCATTC	2479	GAATGGGAGGTGATTCTAG	4248
15369	15391	CCCTAGGAATCACCTCCCATTC	2480	GGAATGGGAGGTGATTCTTA	4249
15370	15392	CCTAGGAATCACCTCCCATTCG	2481	CGGAATGGGAGGTGATTCTCT	4250
15381	15403	CCTCCCATTCCGATAAATCACCC	2482	GGTGATTTTATCGGAATGGG	4251
15384	15406	CCCATTCCGATAAAATCACCTTC	2483	GAAGGTGATTTTATCGGAAT	4252
15385	15407	CCATTCCGATAAAATCACCTTC	2484	GGAAGGTGATTTTATCGGAA	4253
15390	15412	CCGATAAAATCACCTTCACCCCT	2485	AGGGTGGAAGGTGATTAT	4254
15402	15424	CCTTCCACCCTTACTACACAATC	2486	GATTGTGTAGTAAGGTGGA	4255
15406	15428	CCACCCTTACTACACAATCAAAG	2487	CTTTGATTGTGTAGTAAGGG	4256
15409	15431	CCCTTACTACACAATCAAAGACG	2488	CGTCTTTGATTGTGTAGTAA	4257
15410	15432	CCTTACTACACAATCAAAGACGC	2489	GCGTCTTTGATTGTGTAGTA	4258
15432	15454	CCCTCGGCTTACTTCTCTTCCTT	2490	AAGGAAGAGAAGTAGCCGA	4259
15433	15455	CCTCGGCTTACTTCTCTCTCCTTC	2491	GAAGGAAGAGAAGTAGCCG	4260
15451	15473	CCTTCTCTCCTTAATGACATTAA	2492	TTAATGTCATTAAGGAGAGA	4261
15459	15481	CCTTAATGACATTAACTATTC	2493	GAATAGTGTTAATGTCATTA	4262
15485	15507	CCAGACCTCCTAGGCGACCCAGA	2494	TCTGGGTCGCCTAGGAGGTC	4263
15490	15512	CCTCCTAGGCGACCCAGACAATT	2495	AATTGTCTGGGTCGCCTAGG	4264
15493	15515	CCTAGGCGACCCAGACAATTATA	2496	TATAATTGTCTGGGTCGCCT	4265
15502	15524	CCCAGACAATTATACC	2497	TGGCTAGGGTATAAT	4266

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		CTAGCCA		TGTCT	
15503	15525	CCAGACAATTATACCC TAGCCAA	2498	TTGGCTAGGGTATAA TTGTC	4267
15516	15538	CCCTAGCCAACCCCTT AACACCC	2499	GGTGTTTAAGGGGTT GGCTA	4268
15517	15539	CCTAGCCAACCCCTTA AACACCC	2500	GGGTGTTTAAGGGGT TGGCT	4269
15522	15544	CCAACCCCTTAAACAC CCCTCCC	2501	GGGAGGGGTGTTTAA GGGGT	4270
15526	15548	CCCCTTAAACACCCCT CCCACACA	2502	TGTGGGGAGGGGTGT TTAAG	4271
15527	15549	CCCTTAAACACCCCTC CCCACAT	2503	ATGTGGGGAGGGGTG TTAA	4272
15528	15550	CCTTAAACACCCCTCC CCACATC	2504	GATGTGGGGAGGGGT GTTAA	4273
15537	15559	CCCCTCCCCACATCAA GCCCGAA	2505	TTCGGGCTTGATGTG GGGAG	4274
15538	15560	CCCTCCCCACATCAAG CCCGAAT	2506	ATTCGGGCTTGATGT GGGGA	4275
15539	15561	CCTCCCCACATCAAGC CCGAATG	2507	CATTCGGGCTTGATG TGGGG	4276
15542	15564	CCCCACATCAAGCCCG AATGATA	2508	TATCATTCGGGCTTG ATGTG	4277
15543	15565	CCCACATCAAGCCCGA ATGATAT	2509	ATATCATTCGGGCTT GATGT	4278
15544	15566	CCACATCAAGCCCGAA TGATATT	2510	AATATCATTCGGGCT TGATG	4279
15554	15576	CCCGAATGATATTTCC TATTCGC	2511	GCGAATAGGAAATAT CATT	4280
15555	15577	CCGAATGATATTTCTT ATTCGCC	2512	GGCGAATAGGAAATA TCATT	4281
15568	15590	CCTATTCGCCTACACA ATTCTCC	2513	GGAGAATTGTGTAGG CGAAT	4282
15576	15598	CCTACACAATTCTCCG ATCCGTC	2514	GACGGATCGGAGAAT TGTGT	4283
15589	15611	CCGATCCGTCCCTAAC AAACCTAG	2515	CTAGTTTGTTAGGGA CGGAT	4284
15594	15616	CCGTCCCTAACAACT AGGAGGC	2516	GCCTCCTAGTTTGTTA GGA	4285
15598	15620	CCCTAACAACTAGGA GGCGTCC	2517	GGACGCCTCCTAGTT TGTTA	4286
15599	15621	CCTAACAACTAGGAG GCGTCCT	2518	AGGACGCCTCCTAGT TTGTT	4287
15619	15641	CCTTGCCCTATTACTAT CCATCC	2519	GGATGGATAGTAATA GGGCA	4288

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
15624	15646	CCCTATTACTATCCATCCTCATC	2520	GATGAGGATGGATAGTAATA	4289
15625	15647	CCTATTACTATCCATCCTCATCC	2521	GGATGAGGATGGATAGTAAT	4290
15636	15658	CCATCCTCATCCTAGCAATAATC	2522	GATTATTGCTAGGATGAGGA	4291
15640	15662	CCTCATCCTAGCAATAATCCCCA	2523	TGGGGATTATTGCTAGGATG	4292
15646	15668	CCTAGCAATAATCCCCATCCTCC	2524	GGAGGATGGGGATTATTGCT	4293
15658	15680	CCCCATCCTCCATATATCCAAAC	2525	GTTTGGATATATGGAAGGATG	4294
15659	15681	CCCATCCTCCATATATCCAAACA	2526	TGTTTGGATATATGGAAGGAT	4295
15660	15682	CCATCCTCCATATATCAAAACAA	2527	TTGTTTGGATATATGGAAGGA	4296
15664	15686	CCTCCATATATCCAAACAACAAA	2528	TTTGTTGTTTGGATATATGG	4297
15667	15689	CCATATATCCAAACAACAAGCA	2529	TGCTTTGTTGTTTGGATATA	4298
15675	15697	CCAAACAACAAGCATAATATTT	2530	AAATATTATGCTTTGTGTGT	4299
15700	15722	CCACTAAGCCAATCACTTTATT	2531	AATAAAGTGATTGGCTTAGT	4300
15701	15723	CCACTAAGCCAATCACCTTTATTG	2532	CAATAAAGTGATTGGCTTAG	4301
15709	15731	CCAATCACTTTATTGATCCTAG	2533	CTAGGAGTCAATAAAGTGAT	4302
15727	15749	CCTAGCCGCAGACCTCCTCATTC	2534	GAATGAGGAGGTCTGCGGCT	4303
15732	15754	CCGCAGACCTCCTCATCTAACC	2535	GGTTAGAATGAGGAGGTCTG	4304
15739	15761	CCTCCTCATTCTAACCTGAATCG	2536	CGATTCAGGTTAGAAATGAGG	4305
15742	15764	CCTCATTCTAACCTGATCGGAG	2537	CTCCGATTCAGGTTAATGATG	4306
15753	15775	CCTGAATCGGAGGACAACAGTA	2538	TACTGGTTGTCCTCCGATTC	4307
15770	15792	CCAGTAAGCTACCCTTTTACCAT	2539	ATGGTAAAAGGGTAGCTTAC	4308
15781	15803	CCCTTTTACCATCATTGGACAAG	2540	CTTGTCCAATGATGGTAAAA	4309
15782	15804	CCTTTTACCATCATTGGACAAGT	2541	ACTTGTCCAATGATGGTAAA	4310
15789	15811	CCATCATTGGACAAGT	2542	GGATGCTACTTGTCC	4311

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		AGCATCC		AATGA	
15810	15832	CCGTACTATACTTCAC AACAATC	2543	GATTGTTGTGAAGTA TAGTA	4312
15832	15854	CCTAATCCTAATACCA ACTATCT	2544	AGATAGTTGGTATTA GGATT	4313
15838	15860	CCTAATACCAACTATC TCCCTAA	2545	TTAGGGAGATAGTTG GTATT	4314
15845	15867	CCAACATCTCCCTAA TTGAAAA	2546	TTTTCAATTAGGGAG ATAGT	4315
15855	15877	CCCTAATTGAAAACAA AATACTC	2547	GAGTATTTTGTTTTCA ATTA	4316
15856	15878	CCTAATTGAAAACAAA ATACTCA	2548	TGAGTATTTTGTTTTC AATT	4317
15885	15907	CCTGTCCTTGTAAGTAT AAACTAA	2549	TTAGTTTATACTACA AGGAC	4318
15890	15912	CCTTGTAGTATAAACT AATACAC	2550	GTGTATTAGTTTATAC TACA	4319
15912	15934	CCAGTCTTGTAACCG GAGATGA	2551	TCATCTCCGGTTTACA AGAC	4320
15925	15947	CCGGAGATGAAAACCT TTTTCCA	2552	TGGAAAAAGGTTTTTC ATCTC	4321
15938	15960	CCTTTTTCCAAGGACA AATCAGA	2553	TCTGATTTGTCCTTGG AAAA	4322
15945	15967	CCAAGGACAAATCAGA GAAAAAG	2554	CTTTTTCTCTGATTTG TCCT	4323
15977	15999	CCACCATTAGCACCCA AAGCTAA	2555	TTAGCTTTGGGTGCT AATGG	4324
15980	16002	CCATTAGCACCCAAAG CTAAGAT	2556	ATCTTAGCTTTGGGT GCTAA	4325
15989	16011	CCCAAAGCTAAGATTC TAATTTA	2557	TAAATTAGAATCTTA GCTTT	4326
15990	16012	CCAAAGCTAAGATTCT AATTTAA	2558	TTAAATTAGAATCTT AGCTT	4327
16052	16074	CCACCCAAGTATTGAC TCACCCA	2559	TGGGTGAGTCAATAC TTGGG	4328
16055	16077	CCCAAGTATTGACTCA CCCATCA	2560	TGATGGGTGAGTCAA TACTT	4329
16056	16078	CCAAGTATTGACTCAC CCATCAA	2561	TTGATGGGTGAGTCA ATACT	4330
16071	16093	CCCATCAACAACCGCT ATGTATT	2562	AATACATAGCGGTTG TTGAT	4331
16072	16094	CCATCAACAACCGCTA TGTATTT	2563	AAATACATAGCGGTT GTTGA	4332
16082	16104	CCGCTATGTATTCGT ACATTAC	2564	GTAATGTACGAAATA CATAG	4333

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
16107	16129	CCAGCCACCATGAATA TTGTACG	2565	CGTACAATATTCATG GTGGC	4334
16111	16133	CCACCATGAATATTGT ACGGTAC	2566	GTACCGTACAATATT CATGG	4335
16114	16136	CCATGAATATTGTACG GTACCAT	2567	ATGGTACCGTACAAT ATTCA	4336
16133	16155	CCATAAATACTTGACC ACCTGTA	2568	TACAGGTGGTCAAGT ATTTA	4337
16147	16169	CCACCTGTAGTACATA AAAACCC	2569	GGGTTTTTATGTACTA CAGG	4338
16150	16172	CCTGTAGTACATAAAA ACCCAAT	2570	ATTGGGTTTTTATGTA CTAC	4339
16167	16189	CCCAATCCACATCAAAA ACCCCCT	2571	AGGGGGTTTTTGATGT GGATT	4340
16168	16190	CCAATCCACATCAAAA CCCCTC	2572	GAGGGGGTTTTTGATG TGGAT	4341
16173	16195	CCACATCAAAAACCCCC TCCCCAT	2573	ATGGGGAGGGGGTTT TGATG	4342
16184	16206	CCCCCTCCCCATGCTT ACAAGCA	2574	TGCTTGTAAGCATGG GGAGG	4343
16185	16207	CCCCTCCCCATGCTTA CAAGCAA	2575	TTGCTTGTAAGCATG GGGAG	4344
16186	16208	CCCTCCCCATGCTTAC AAGCAAG	2576	CTTGCTTGTAAGCAT GGGGA	4345
16187	16209	CCTCCCCATGCTTACA AGCAAGT	2577	ACTTGCTTGTAAGCA TGGGG	4346
16190	16212	CCCCATGCTTACAAGC AAGTACA	2578	TGTACTTGCTTGTAAGCATG	4347
16191	16213	CCCATGCTTACAAGCA AGTACAG	2579	CTGTACTTGCTTGTAAGCAT	4348
16192	16214	CCATGCTTACAAGCAA GTACAGC	2580	GCTGTACTTGCTTGTAAGCA	4349
16221	16243	CCCTCAACTATCACAC ATCAACT	2581	AGTTGATGTGTGATA GTTGA	4350
16222	16244	CCTCAACTATCACACA TCAACTG	2582	CAGTTGATGTGTGAT AGTTG	4351
16250	16272	CCAAAGCCACCCCTCA CCCACTA	2583	TAGTGGGTGAGGGGT GGCTT	4352
16256	16278	CCACCCCTCACCCACT AGGATAC	2584	GTATCCTAGTGGGTG AGGGG	4353
16259	16281	CCCCTCACCCACTAGG ATACCAA	2585	TTGGTATCCTAGTGG GTGAG	4354
16260	16282	CCCTCACCCACTAGGA TACCAAC	2586	GTTGGTATCCTAGTG GGTGA	4355
16261	16283	CCTCACCCACTAGGAT	2587	TGTTGGTATCCTAGT	4356

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
		ACCAACA		GGGTG	
16266	16288	CCCACTAGGATACCAA CAAACCT	2588	AGGTTTGTGTTGGTATC CTAGT	4357
16267	16289	CCACTAGGATACCAAC AAACCTA	2589	TAGGTTTGTGTTGGTATC CTAG	4358
16278	16300	CCAACAAACCTACCCA CCCTTAA	2590	TTAAGGGTGGGTAGG TTTGT	4359
16286	16308	CCTACCCACCCTTAAC AGTACAT	2591	ATGTACTGTTAAGGG TGGGT	4360
16290	16312	CCCACCCTTAACAGTA CATAGTA	2592	TACTATGTACTGTTA AGGGT	4361
16291	16313	CCACCCTTAACAGTAC ATAGTAC	2593	GTACTATGTACTGTT AAGGG	4362
16294	16316	CCCTTAACAGTACATA GTACATA	2594	TATGTACTATGTACT GTAA	4363
16295	16317	CCTTAACAGTACATAG TACATAA	2595	TTATGTACTATGTACT GTAA	4364
16320	16342	CCATTTACCGTACATA GCACATT	2596	AATGTGCTATGTACG GTAAA	4365
16327	16349	CCGTACATAGCACATT ACAGTCA	2597	TGACTGTAATGTGCT ATGTA	4366
16353	16375	CCCTTCTCGTCCCCATG GATGAC	2598	GTCATCCATGGGGAC GAGAA	4367
16354	16376	CCTTCTCGTCCCCATG GATGACC	2599	GGTCATCCATGGGGA CGAGA	4368
16363	16385	CCCCATGGATGACCCC CCTCAGA	2600	TCTGAGGGGGGTCAT CCATG	4369
16364	16386	CCCATGGATGACCCCC CTCAGAT	2601	ATCTGAGGGGGGTCA TCCAT	4370
16365	16387	CCATGGATGACCCCC TCAGATA	2602	TATCTGAGGGGGGTC ATCCA	4371
16375	16397	CCCCCTCAGATAGGG GTCCCTT	2603	AAGGGACCCCTATCT GAGGG	4372
16376	16398	CCCCCTCAGATAGGGG TCCCTTG	2604	CAAGGGACCCCTATC TGAGG	4373
16377	16399	CCCCTCAGATAGGGGT CCCTTGA	2605	TCAAGGGACCCCTAT CTGAG	4374
16378	16400	CCCTCAGATAGGGGTC CCTTGAC	2606	GTCAAGGGACCCCTA TCTGA	4375
16379	16401	CCTCAGATAGGGGTCC CTTGACC	2607	GGTCAAGGGACCCCT ATCTG	4376
16393	16415	CCCTTGACCACCATCC TCCGTGA	2608	TCACGGAGGATGGTG GTCAA	4377
16394	16416	CCTTGACCACCATCCT CCGTGAA	2609	TTCACGGAGGATGGT GGTCA	4378

Chr start position (+ strand)	Chr end position (+ strand)	nt sequence on the (+) strand containing CCN sequence followed by the reverse complementary sequence of gRNA target sequence	SEQ ID NO	20 nt gRNA target sequence (will encode the gRNA targeting sequence)	SEQ ID NO
16400	16422	CCACCATCCTCCGTGA AATCAAT	2610	ATTGATTTACGGAG GATGG	4379
16403	16425	CCATCCTCCGTGAAAT CAATATC	2611	GATATTGATTTACG GAGGA	4380
16407	16429	CCTCCGTGAAATCAAT ATCCCGC	2612	GCGGGATATTGATTT CACGG	4381
16410	16432	CCGTGAAATCAATATC CCGCACA	2613	TGTGCGGGATATTGA TTTCA	4382
16425	16447	CCCGCACAAGAGTGCT ACTCTCC	2614	GGAGAGTAGCACTCT TGTGC	4383
16426	16448	CCGCACAAGAGTGCTA CTCTCCT	2615	AGGAGAGTAGCACTC TTGTG	4384
16446	16468	CCTCGCTCCGGGCCCA TAACACT	2616	AGTGTTATGGGCCCCG GAGCG	4385
16453	16475	CCGGGCCCATAACACT TGGGGGT	2617	ACCCCCAAGTGTTAT GGGCC	4386
16458	16480	CCCATAACACTTGGGG GTAGCTA	2618	TAGCTACCCCCAAGT GTTAT	4387
16459	16481	CCATAACACTTGGGGG TAGCTAA	2619	TTAGCTACCCCCAAG TGTTA	4388
16494	16516	CCGACATCTGGTTCCT ACTTCAG	2620	CTGAAGTAGGAACCA GATGT	4389
16507	16529	CCTACTTCAGGGTCAT AAAGCCT	2621	AGGCTTTATGACCCT GAAGT	4390
16527	16549	CCTAAATAGCCCACAC GTTCCCC	2622	GGGGAACGTGTGGGC TATTT	4391
16536	16558	CCCACACGTTCCCCTT AAATAAG	2623	CTTATTTAAGGGGAA CGTGT	4392
16537	16559	CCACACGTTCCCCTTA AATAAGA	2624	TCTTATTTAAGGGGA ACGTG	4393
16546	16568	CCCCTTAAATAAGACA TCACGAT	2625	ATCGTGATGTCTTATT TAAG	4394
16547	16569	CCCTTAAATAAGACAT CACGATG	2626	CATCGTGATGTCTTAT TTAA	4395
16548	16570	CCTTAAATAAGACATC ACGATGG	2627	CCATCGTGATGTCTT ATTTA	4396

Applications

[0239] The gNAs (e.g., gRNAs) and collections of gNAs (e.g., gRNAs) provided herein are useful for a variety of applications, including depletion, partitioning, capture, or enrichment of target sequences of interest; genome-wide labeling; genome-wide editing; genome-wide function screens; and genome-wide regulation.

[0240] In one embodiment, the gNAs are selective for host nucleic acids in a biological sample from a host, but are not selective for non-host nucleic acids in the sample from a host. In one embodiment, the gNAs are selective for non-host nucleic acids from a biological sample from a host but are not selective for the host nucleic acids in the sample. In one embodiment, the gNAs are selective for both host nucleic acids and a subset of the non-host nucleic acids in a biological sample from a host. For example, where a complex biological sample comprises host nucleic acids and nucleic acids from more than one non-host organisms, the gRNAs may be selective for more than one of the non-host species. In such embodiments, the gNAs are used to serially deplete or partition the sequences that are not of interest. For example, saliva from a human contains human DNA, as well as the DNA of more than one bacterial species, but may also contain the genomic material of an unknown pathogenic organism. In such an embodiment, gNAs directed at the human DNA and the known bacteria can be used to serially deplete the human DNA, and the DNA of the known bacterial, thus resulting in a sample comprising the genomic material of the unknown pathogenic organism.

[0241] In an exemplary embodiment, the gNAs are selective for human host DNA obtained from a biological sample from the host, but do not hybridize with DNA from an unknown pathogen(s) also obtained from the sample.

[0242] In some embodiments, the gNAs are useful for depleting and partitioning of targeted sequences in a sample, enriching a sample for non-host nucleic acids, or serially depleting targeted nucleic acids in a sample comprising: providing nucleic acids extracted from a sample; and contacting the sample with a plurality of complexes comprising (i) any one of the collection of gNAs described herein and (ii) nucleic acid-guided nuclease (e.g., CRISPR/Cas) system proteins.

[0243] In some embodiments, the gNAs are useful for method of depletion and partitioning of targeted sequences in a sample comprising: providing nucleic acids extracted from a sample, wherein the extracted nucleic acids comprise sequences of interest and targeted sequences for one of depletion and partitioning; contacting the sample with a plurality of complexes comprising (i) a collection of gNAs provided herein; and (ii) nucleic acid-guided nuclease (e.g., CRISPR/Cas) system proteins, under conditions in which the nucleic acid-guided nuclease system proteins cleave the nucleic acids in the sample.

[0244] In some embodiments, the gNAs are useful for enriching a sample for non-host nucleic acids comprising: providing a sample comprising host nucleic acids and non-host nucleic acids; contacting the sample with a plurality of complexes comprising (i) a collection of gNAs provided herein comprising targeting sequences directed at the host nucleic acids; and (ii) nucleic acid-guided nuclease (e.g., CRISPR/Cas) system proteins, under conditions in which the nucleic acid-guided nuclease system proteins cleave the host nucleic acids in the sample, thereby depleting the sample of host nucleic acids, and allowing for the enrichment of non-host nucleic acids..

[0245] In some embodiments, the gNAs are useful for one method for serially depleting targeted nucleic acids in a sample comprising: providing a biological sample from a host comprising host nucleic acids and non-host nucleic acids, wherein the non-host nucleic acids comprise nucleic acids from at least one known non-host organism and nucleic acids from an unknown non-host organism; providing a plurality of complexes comprising (i) a collection of gNAs provided herein, directed at the host nucleic acids; and (ii) nucleic acid-guided nuclease (e.g., CRISPR/Cas) system proteins; mixing the nucleic acids from the biological sample with the gNA-nucleic acid-guided nuclease system protein complexes (e.g., gRNA-CRISPR/Cas system protein complexes) configured to hybridize to targeted sequences in the host nucleic acids, wherein at least a portion of the complexes hybridizes to the targeted sequences in the host nucleic acids, and wherein at least a portion of the host nucleic acids are cleaved; mixing the remaining nucleic acids from the biological sample with the gNA-nucleic acid-guided nuclease system protein complexes configured to hybridize to targeted sequences in the at least one known non-host nucleic acids, wherein at least a portion of the complexes hybridizes to the targeted sequences in the at least one non-host nucleic acids, and wherein at least a portion of the non-host nucleic acids are cleaved; and isolating the remaining nucleic acids from the unknown non-host organism and preparing for further analysis.

[0246] In some embodiments, the gNAs generated herein are used to perform genome-wide or targeted functional screens in a population of cells. In such an embodiment, libraries of *in vitro*-transcribed gNAs (e.g., gRNAs) or vectors encoding the gNAs can be introduced into a population of cells via transfection or other laboratory techniques known in the art, along with a nucleic acid-guided nuclease (e.g., CRISPR/Cas) system protein, in a way that gNA-directed nucleic acid-guided nuclease system protein editing can be achieved to sequences across the entire genome or to a specific region of the genome. In one embodiment, the nucleic acid-guided nuclease system protein can be introduced as a DNA. In one embodiment, the nucleic acid-guided nuclease system protein can be introduced as mRNA. In one embodiment, the nucleic acid-guided nuclease system protein can be introduced as protein. In one exemplary embodiment, the nucleic acid-guided nuclease system protein is Cas9.

[0247] In some embodiments, the gNAs generated herein are used for the selective capture and/or enrichment of nucleic acid sequences of interest. For example, in some embodiments, the gNAs generated herein are used for capturing target nucleic acid sequences comprising: providing a sample comprising a plurality of nucleic acids; and contacting the sample with a plurality of complexes comprising (i) a collection of gNAs provided herein; and (ii) nucleic acid-guided nuclease (e.g., CRISPR/Cas) system proteins. Once the sequences of interest are captured, they can be further ligated to create, for example, a sequencing library.

[0248] In some embodiments, the gNAs generated herein are used for introducing labeled nucleotides at targeted sites of interest comprising: (a) providing a sample comprising a plurality of nucleic acid

fragments; (b) contacting the sample with a plurality of complexes comprising (i) a collection of gNAs provided herein; and (ii) nucleic acid-guided nuclease (e.g., CRISPR/Cas) system protein-nickases (e.g. Cas9-nickases), wherein the gNAs are complementary to targeted sites of interest in the nucleic acid fragments, thereby generating a plurality of nicked nucleic acid fragments at the targeted sites of interest; and (c) contacting the plurality of nicked nucleic acid fragments with an enzyme capable of initiating nucleic acid synthesis at a nicked site, and labeled nucleotides, thereby generating a plurality of nucleic acid fragments comprising labeled nucleotides in the targeted sites of interest.

[0249] In some embodiments, the gNAs generated herein are used for capturing target nucleic acid sequences of interest comprising: (a) providing a sample comprising a plurality of adapter-ligated nucleic acids, wherein the nucleic acids are ligated to a first adapter at one end and are ligated to a second adapter at the other end; and (b) contacting the sample with a collection of gNAs which comprise a plurality of dead nucleic acid-guided nuclease-gNA complexes (e.g., dCas9-gRNA complexes), wherein the dead nucleic acid-guided nuclease (e.g., dCas9) is fused to a transposase, wherein the gNAs are complementary to targeted sites of interest contained in a subset of the nucleic acids, and wherein the dead nucleic acid-guided nuclease-gNA transposase complexes (e.g., dCas9-gRNA transposase complexes) are loaded with a plurality of third adapters, to generate a plurality of nucleic acids fragments comprising either a first or second adapter at one end and a third adapter at the other end. In one embodiment the method further comprises amplifying the product of step (b) using first or second adapter and third adapter-specific PCR.

[0250] In some embodiments, the gNAs generated herein are used to perform genome-wide or targeted activation or repression in a population of cells. In such an embodiment, libraries of *in vitro*-transcribed gNAs (e.g., gRNAs) or vectors encoding the gNAs can be introduced into a population of cells via transfection or other laboratory techniques known in the art, along with a catalytically dead nucleic acid-guided nuclease (e.g., CRISPR/Cas) system protein fused to an activator or repressor domain (catalytically dead nucleic acid-guided nuclease system protein-fusion protein), in a way that gNA-directed catalytically dead nucleic acid-guided nuclease system protein-mediated activation or repression can be achieved at sequences across the entire genome or to a specific region of the genome. In one embodiment, the catalytically dead nucleic acid-guided nuclease system protein -fusion protein can be introduced as DNA. In one embodiment, the catalytically dead nucleic acid-guided nuclease system protein -fusion protein can be introduced as mRNA. In one embodiment, the catalytically dead nucleic acid-guided nuclease system protein-fusion protein can be introduced as protein. In some embodiments, the collection of gNAs or nucleic acids encoding for gNAs exhibit specificity for more than one nucleic acid-guided nuclease system protein. In one exemplary embodiment, the catalytically dead nucleic acid-guided nuclease system protein is dCas9.

[0251] In some embodiments, the collection comprises gRNAs or nucleic acids encoding for gRNAs with specificity for Cas9 and one or more CRISPR/Cas system proteins selected from the group consisting of Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, and Cm5. In some embodiments, the collection comprises gRNAs or nucleic acids encoding for gRNAs with specificity for various catalytically dead CRISPR/Cas system proteins fused to different fluorophores, for example for use in the labeling and/or visualization of different genomes or portions of genomes, for use in the labeling and/or visualization of different chromosomal regions, or for use in the labeling and/or visualization of the integration of viral genes/genomes into a genome.

[0252] In some embodiments, the collection of gNAs (or nucleic acids encoding for gNAs) have specificity for different nucleic acid-guided nuclease (e.g., CRISPR/Cas) system proteins, and target different sequences of interest, for example from different species. For example, a first subset of gNAs from a collection of gNAs (or transcribed from a population of nucleic acids encoding such gNAs) targeting a genome from a first species can be first mixed with a first nucleic acid-guided nuclease system protein member (or an engineered version); and a second subset of gNAs from a collection of gNAs (or transcribed from a population of nucleic acids encoding such gNAs) targeting a genome from a second species can be mixed with a second different nucleic acid-guided nuclease system protein member (or an engineered version). In one embodiment, the nucleic acid-guided nuclease system proteins can be a catalytically dead version (for example dCas9) fused with different fluorophores, so that different targeted sequence of interest, e.g. different species genome, or different chromosomes of one species, can be labeled by different fluorescent labels. For example, different chromosomal regions can be labeled by different gRNA-targeted dCas9-fluorophores, for visualization of genetic translocations. For example, different viral genomes can be labeled by different gRNA-targeted dCas9-fluorophores, for visualization of integration of different viral genomes into the host genome. In another embodiment, the nucleic acid-guided nuclease system protein can be dCas9 fused with either activation or repression domain, so that different targeted sequence of interest, e.g. different chromosomes of a genome, can be differentially regulated. In another embodiment, the nucleic acid-guided nuclease system protein can be dCas9 fused different protein domain which can be recognized by different antibodies, so that different targeted sequence of interest, e.g. different DNA sequences within a sample mixture, can be differentially isolated.

Exemplary Compositions of the Invention

[0253] In one embodiment, provided herein is a composition comprising a nucleic acid fragment, a nickase nucleic acid-guided nuclease-gNA complex, and labeled nucleotides. In one exemplary embodiment, provided herein is a composition comprising a nucleic acid fragment, a nickase Cas9-gRNA complex, and labeled nucleotides. In such embodiments, the nucleic acid may comprise DNA.

The nucleotides can be labeled, for example with biotin. The nucleotides can be part of an antibody-conjugate pair.

[0254] In one embodiment, provided herein is a composition comprising a nucleic acid fragment and a catalytically dead nucleic acid-guided nuclease-gNA complex, wherein the catalytically dead nucleic acid-guided nuclease is fused to a transposase. In one exemplary embodiment, provided herein is a composition comprising a DNA fragment and a dCas9-gRNA complex, wherein the dCas9 is fused to a transposase.

[0255] In one embodiment, provided herein is a composition comprising a nucleic acid fragment comprising methylated nucleotides, a nickase nucleic acid-guided nuclease-gNA complex, and unmethylated nucleotides. In an exemplary embodiment, provided herein is a composition comprising a DNA fragment comprising methylated nucleotides, a nickase Cas9-gRNA complex, and unmethylated nucleotides.

[0256] In one embodiment, provided herein is a gDNA complexed with a nucleic acid-guided -DNA endonuclease. In an exemplary embodiment, the nucleic acid-guided -DNA endonuclease is NgAgo.

[0257] In one embodiment, provided herein is a gDNA complexed with a nucleic acid-guided -RNA endonuclease.

[0258] In one embodiment, provided herein is a gRNA complexed with a nucleic acid-guided -DNA endonuclease.

[0259] In one embodiment, provided herein is a gRNA complexed with a nucleic acid-guided -RNA endonuclease. In one embodiment, the nucleic acid-guided -RNA endonuclease comprises C2c2.

Kits and Articles of Manufacture

[0260] The present application provides kits comprising any one or more of the compositions described herein, not limited to adapters, gNAs (e.g., gRNAs), gNA collections (e.g., gRNA collections), nucleic acid molecules encoding the gNA collections, and the like.

[0261] In one exemplary embodiment, the kit comprises a collection of DNA molecules capable of transcribing into a library of gRNAs wherein the gRNAs are targeted to human genomic or other sources of DNA sequences.

[0262] In one embodiment, the kit comprises a collection of gNAs wherein the gNAs are targeted to human genomic or other sources of DNA sequences.

[0263] In some embodiments, provided herein are kits comprising any of the collection of nucleic acids encoding gNAs, as described herein. In some embodiments, provided herein are kits comprising any of the collection of gNAs, as described herein.

[0264] The present application also provides all essential reagents and instructions for carrying out the methods of making the gNAs and the collection of nucleic acids encoding gNAs, as described herein. In

some embodiments, provided herein are kits that comprise all essential reagents and instructions for carrying out the methods of making individual gNAs and collections of gNAs as described herein.

[0265] Also provided herein is computer software monitoring the information before and after contacting a sample with a gNA collection produced herein. In one exemplary embodiment, the software can compute and report the abundance of non-target sequence in the sample before and after providing gNA collection to ensure no off-target targeting occurs, and wherein the software can check the efficacy of targeted-depletion/enrichment/capture/partitioning/labeling/regulation/editing by comparing the abundance of the target sequence before and after providing gNA collection to the sample.

[0266] The following examples are included for illustrative purposes and are not intend to limit the scope of the invention.

EXAMPLES

Example 1: Construction of a gRNA library from a T7 promoter human DNA library

T7 promoter library construction

[0267] Human genomic DNA (400 ng) was fragmented using an S2 Covaris sonicator (Covaris) for 8 cycles, to yield fragments of 200-300 bp in length. Fragmented DNA was repaired using the NEBNext End Repair Module (NEB) and incubated at 25 °C for 30 min, then heat inactivated at 75 °C for 20 min. To make T7 promoter adapters, oligos T7-1 (5'GCCTCGAGC*T*A*ATACGACTCACTATAGAG3', * denotes a phosphorothioate backbone linkage)(SEQ ID NO: 4397) and T7-2 (sequence 5'Phos-CTCTATAGTGAGTCGTATTA3') (SEQ ID NO: 4398) were admixed at 15 µM, heated to 98 °C for 3 min then cooled slowly (0.1 °C/min) to 30 °C. T7 promoter blunt adapters (15 pmol total) were then added to the blunt-ended human genomic DNA fragments, and incubated with Blunt/TA Ligase Master Mix (NEB) at 25 °C for 30 min ((2) in FIG. 1). Ligations were amplified with 2 µM oligo T7-1, using Hi-Fidelity 2X Master Mix (NEB) for 10 cycles of PCR (98 °C for 20 s, 63 °C for 20 s, 72 °C for 35 s). Amplification was verified by running a small aliquot on agarose gel electrophoresis. PCR amplified products were recovered using 0.6X AxyPrep beads (Axygen) according to the manufacturer's instructions, and resuspended in 15 µL of 10 mM Tris-HCl pH 8.

Digestion of DNA

[0268] PCR amplified T7 promoter DNA (2 µg total per digestion) was digested with 0.1 µL of Nt.CviPII (NEB) in 10 µL of NEB buffer 2 (50 mM NaCl, 10 mM Tris-HCl pH 7.9, 10 mM MgCl₂, 100 µg/mL BSA) for 10 min at 37 °C ((3) in FIG. 1), then heat inactivated at 75 °C for 20 min. An additional 10 µL of NEB buffer 2 with 1 µL of T7 Endonuclease I (NEB) was added to the reaction, and incubated at 37 °C for 20 min ((4) in FIG. 1). Enzymatic digestion of DNA was verified by agarose gel electrophoresis. Digested DNA was recovered by adding 0.6X AxyPrep beads (Axygen), according to the manufacturer's instructions, and resuspended in 15 µL of 10 mM Tris-HCl pH 8.

Ligation of adapters and removal of HGG

[0269] DNA was then blunted using T4 DNA Polymerase (NEB) for 20 min at 25 °C, followed by heat inactivation at 75 °C for 20 min ((5) in FIG. 1).

[0270] To make MlyI adapters, oligos MlyI-1 (sequence 5'>3', 5'Phos-GGGACTCGGATCCCTATAGTGATACAAAGACGATGACGACAAGCG) (SEQ ID NO: 4399) and MlyI-2 (sequence 5'>3', TCACTATAGGGATCCGAGTCCC) (SEQ ID NO: 4400) were admixed at 15 μM, heated to 98 °C for 3 min then cooled slowly (0.1 °C/min) to 30 °C. MlyI adapters (15 pmol total) were then added to T4 DNA Polymerase-blunted DNA, and incubated with Blunt/TA Ligase Master Mix (NEB) at 25 °C for 30 min ((6) in FIG. 1). Ligations were heat inactivated at 75 °C for 20 min, then digested with MlyI and XhoI (NEB) for 1 hr at 37 °C, so that HGG motifs are eliminated ((7) in FIG. 1). Digests were then cleaned using 0.8X AxyPrep beads (Axygen), and DNA was resuspended in 10 μL of 10 mM Tris-Cl pH 8.

[0271] To make StlgR adapters, oligos stlgR (sequence 5'>3', 5'Phos-GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTTGGATCCGATGC) (SEQ ID NO: 4401) and stlgRev (sequence 5'>3', GGATCCAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATT TAACTTGCTATTTCTAGCTCTAAAC) (SEQ ID NO: 4402) were admixed at 15 μM, heated to 98 °C for 3 min then cooled slowly (0.1 °C/min) to 60 °C. StlgR adapters (5 pmol total) were added to HGG-removed DNA fragments, and incubated with Blunt/TA Ligase Master Mix (NEB) at 25 °C for 30 min ((8) in FIG. 1). Ligations were then incubated with Hi-Fidelity 2X Master Mix (NEB), using 2 μM of both oligos T7-1 and gRU (sequence 5'>3', AAAAAAAGCACCGACTCGGTG) (SEQ ID NO: 4403), and amplified using 20 cycles of PCR (98 °C for 20 s, 60 °C for 20 s, 72 °C for 35 s). Amplification was verified by running a small aliquot on agarose gel electrophoresis. PCR amplified products were recovered using 0.6X AxyPrep beads (Axygen) according to the manufacturer's instructions, and resuspended in 15 μL of 10mM Tris-HCl pH 8.

In vitro transcription

[0272] The T7/gRU amplified library of PCR products was then used as template for *in vitro* transcription, using the HiScribe T7 *In Vitro* Transcription Kit (NEB). 500-1000 ng of template was incubated overnight at 37 °C according to the manufacturer's instructions. To transcribe the guide libraries into gRNAs, the following *in vitro* transcription reaction mixture was assembled: 10 μL of purified library (~500 ng), 6.5 μL of H₂O, 2.25 μL of ATP, 2.25 μL of CTP, 2.25 μL of GTP, 2.25 μL of UTP, 2.25 μL of 10X reaction buffer (NEB) and 2.25 μL of T7 RNA Polymerase mix. The reaction was incubated at 37 °C for 24 hr, then purified using the RNA cleanup kit (Life Technologies), eluted with 100 μL of RNase-free water, quantified and stored at -20 °C until use.

Example 2: Construction of gRNA library from intact human genomic DNA***Digestion of DNA***

[0273] Human genomic DNA ((1) in FIG. 2; 20 µg total per digestion) was digested with 0.1 µL of Nt.CviPII (NEB) in 40 µL of NEB buffer 2 (50 mM NaCl, 10 mM Tris-HCl pH 7.9, 10 mM MgCl₂, 100 µg/mL BSA) for 10 min at 37 °C, then heat inactivated at 75 °C for 20 min. An additional 40 µL of NEB buffer 2 and 1 µL of T7 Endonuclease I (NEB) was added to the reaction, with 20 min incubation at 37 °C (e.g., (2) in FIG. 2). Fragmentation of genomic DNA was verified with a small aliquot by agarose gel electrophoresis. DNA fragments between 200 and 600 bp were recovered by adding 0.3X AxyPrep beads (Axygen), incubating at 25 °C for 5 min, capturing beads on a magnetic stand and transferring the supernatant to a new tube. DNA fragments below 600 bp do not bind to beads at this bead/DNA ratio and remain in the supernatant. 0.7X AxyPrep beads (Axygen) were then added to the supernatant (this will bind all DNA molecules longer than 200 bp), allowed to bind for 5 min. Beads were captured on a magnetic stand and washed twice with 80% ethanol, air dried. DNA was then resuspended in 15 µL of 10 mM Tris-HCl pH 8. DNA concentration was determined using a Qbit assay (Life Technologies).

Ligation of adapters

[0274] To make T7/MlyI adapters, oligos MlyI-1 (sequence 5'>3', 5'Phos-GGGGGACTCGGATCCCTATAGTGATACAAAGACGATGACGACAAGCG) (SEQ ID NO: 4404) and T7-7 (sequence 5'>3', GCCTCGAGC*T*A*ATACGACTCACTATAGGGATCCAAGTCCC, * denotes a phosphorothioate backbone linkage) (SEQ ID NO: 4405) were admixed at 15 µM, heated to 98 °C for 3 min then cooled slowly (0.1 °C/min) to 30 °C. The purified, Nt.CviPII/T7 Endonuclease I digested DNA (100 ng) was then ligated to 15 pmol of T7/MlyI adapters using Blunt/TA Ligase Master Mix (NEB) at 25 °C for 30 min ((3) in FIG. 2). Ligations were then amplified by 10 cycles of PCR (98 °C for 20 s, 60 °C for 20 s, 72 °C for 35 s) using Hi-Fidelity 2X Master Mix (NEB), and 2 µM of both oligos T7-17 (GCCTCGAGC*T*A*ATACGACTCACTATAGGG * denotes a phosphorothioate backbone linkage) (SEQ ID NO: 4406) and Flag (sequence 5'>3', CGCTTGTCGTCATCGTCTTTGTA) (SEQ ID NO: 4407). PCR amplification increases the yield of DNA and, given the nature of the Y-shaped adapters we used, always resulted in T7 promoter being added distal to the HGG site and MlyI site being added next to the HGG motif ((4) in FIG. 2).

[0275] PCR products were then digested with MlyI and XhoI (NEB) for 1 hr at 37 °C, and heat inactivated at 75 °C for 20 min ((5) in FIG. 2). Following that, 5 pmol of adapter StlgR (in Example 1) was ligated using Blunt/TA Ligase Master Mix (NEB) at 25 °C for 30 min ((6) in FIG. 2). Ligations were then amplified by PCR using Hi-Fidelity 2X Master Mix (NEB), 2 µM of both oligos T7-7 and gRU (in Example 1) and 20 cycles of PCR (98 °C for 20 s, 60 °C for 20 s, 72 °C for 35 s). Amplification was verified by running a small aliquot on agarose gel electrophoresis. PCR amplified products were

recovered using 0.6X AxyPrep beads (Axygen) according to the manufacturer’s instructions, and resuspended in 15 µL of 10 mM Tris-HCl pH 8.

[0276] Samples were then used as templates for *in vitro* transcription reaction as described in Example 1.

Example 3: Direct Cutting with CviPII

[0277] 30 µg of human genomic DNA was digested with 2 units of NtCviPII (New England Biolabs) for 1 hour at 37 °C, followed by heat inactivation at 75 °C for 20 minutes. The size of the fragments was verified to be 200-1,000 base pairs using a fragment analyzer instrument (Advanced Analytical). The 5’ or 3’ protruding ends (as shown, for example, in **FIG. 3**) were converted to blunt ends by adding 100 units of T4 DNA polymerase (New England Biolabs), 100 µM dNTPs and incubating at 12 °C for 30 minutes. DNA was then recovered using a PCR cleanup kit (Zymo) and eluted in 20 µL elution buffer. The DNA was then ligated to MlyI adapter (see, for example, Example 4) or BaeI/EcoP15I adapters (see, for example, Example 4) or BaeI/EcoP15I adapters (see, for example, Example 5)

Example 4: Use of MlyI Adapter

[0278] Adapter MlyI was made by combining 2 µmoles of MlyI Ad1 and MlyAd2 in 40 µL water. Adapter BsaXI/MmeI was made by combining 2 µmoles oligo BsMm-Ad1 and 2 µmoles oligo BsMm-Ad2 in 40 µL water. T7 adapter was made by combining 1.5 µmoles of T7-Ad1 and T7-Ad2 oligos in 100 µL water. Stem-loop adapter was made by combining 1.5 µmoles of gR-top and gR-bot oligos in 100 µL water. In all cases, after mixing adapters were heated to 98 °C for 3 min then cooled to room temperature at a cooling rate of 1 °C/min in a thermal cycler.

[0279] Table 5. Oligonucleotides used with MlyI Adapter.

SEQ ID NO	Oligo name	Sequence (5'>3')	Modification
4408	MlyI-Ad1	gagatcagcttctgcattgatgccagcagcccgagtcag	none
4409	MlyI-Ad2	ctgactcgggctgctgtacaaagacgatgacgacaagcgta	5'phosphate
4410	BsMm-Ad1	gagatcagcttctgcattgatgcGGAGCCGCAGTACACTATCCAAC	none
4411	BsMm-Ad2	GTTGGATAGTGTACTGCGGCTCtacaaagacgatgacgacaagcg	5'phosphate
4412	T7-Ad1	gcctcgagctaatacgactcactatagagNN	none
4398	T7-Ad2	Ctctatagtgagtcgtatta	5'phosphate
4413	gR-top	ttagagctagaaatagcaagttaaaataaggctagtccgttatcaactgaaaaagtggcaccgagtcggtgctttttt	5'phosphate
4414	gR-bot	aaaaaagcaccgactcgggtgccactttttcaagttgataacggactagccttattttaacttgctatttctagct	none

		ctaaaac	
--	--	---------	--

[0280] The DNA containing the CCD blunt ends (from earlier section) was then ligated to 50 pmoles of adapter MlyI, using the blunt/TA ligation master mix (New England Biolabs) at room temperature for 30 minutes. The DNA was then recovered by incubating with 0.6X Kapa SPRI beads (Kapa Biosystems) for 5 minutes, capturing the beads with a magnetic rack, washing twice with 80% ethanol, air drying the beads for 5 minutes and finally resuspending the DNA in 50 µL buffer 4 (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 100 µg/mL BSA, pH 7.9). These steps eliminate small (<100 nucleotides) DNA and MlyI adapter dimers.

[0281] Purified DNA was then digested by adding 20 units of MlyI (New England Biolabs) and incubating at 37 °C for 1 hour to eliminate both the adapter derived sequences and the CCD (and complementary HGG) motifs. DNA was recovered from the digest by incubating with 0.6X Kapa SPRI beads (Kapa Biosystems) for 5 minutes, capturing the beads with a magnetic rack, washing twice with 80% ethanol, air drying the beads for 5 minutes and finally resuspending the DNA in 30 µL buffer 4.

[0282] The purified DNA was then ligated to 50 pmoles of adapter BsaXI/MmeI, using the blunt/TA ligation master mix (New England Biolabs) at room temperature for 30 minutes. The DNA was then recovered by incubating with 0.6X Kapa SPRI beads (Kapa Biosystems) for 5 minutes, capturing the beads with a magnetic rack, washing twice with 80% ethanol, air drying the beads for 5 minutes and finally resuspending the DNA in 50 µL buffer 4 (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 100 µg/mL BSA, pH 7.9). DNA was then digested by addition of 20 units MmeI (New England Biolabs) and 40 pmol/µL SAM (S-adenosyl methionine) at 37 °C for 1 hour, followed by heat inactivation at 75 °C for 20 minutes. DNA was then ligated to 30 pmoles T7 adapter using the blunt/TA ligation master mix (New England Biolabs) at room temperature for 30 minutes. DNA was then recovered using a PCR cleanup kit (Zymo) and eluted in 20 µL buffer 4, then digested with 20 units of BsaXI for 1 hour at 37 °C. The guide RNA stem-loop sequences were added by adding 15 pmoles stem-loop adapter and using the blunt/TA ligation master mix (New England Biolabs) at room temperature for 30 min. DNA was then recovered using a PCR cleanup kit (Zymo), eluted in 20 µL elution buffer and PCR amplified using HiFidelity 2X master mix (New England Biolabs). Primers T7-Ad1 and gRU (sequence 5'>3' AAAAAAGCACCGACTCGGTG) (SEQ ID NO: 4419) were used to amplify with the following settings (98 °C 3 min; 98 °C for 20 sec, 60 °C for 30 secs, 72 °C for 20 sec, 30 cycles). The PCR amplicon was cleaned up using the PCR cleanup kit and verified by DNA sequencing, then used as template for an in vitro transcription reaction to generate guide RNAs.

Example 5: Use of BaeI/EcoP15I Adapter

[0283] Adapter BaeI/EcoP15I was made by combining 2 µmoles of BE Ad1 and BE Ad2 in 40 µL water. T7-E adapter was made by combining 1.5 µmoles of T7-Ad3 and T7-Ad4 oligos in 100 µL water.

In all cases, after mixing adapters were heated to 98 °C for 3 min then cooled to room temperature at a cooling rate of 1 °C/min in a thermal cycler.

[0284] Table 6. Oligonucleotides used with BaeI/EcoP15I Adapter.

SEQ ID NO:	Oligo name	Sequence (5'>3')	Modification
4415	BE Ad1	ActgctgacACAgtatcTTTTTTTTTgtttaaactTTTTTTTTTgatacACAgtcagcagA	5'phosphate
4416	Be Ad2	TctgctgacTTGTgtatcAAAAAAAAAAgtttaaactAAAAAAAAAAgatacTTGTgtcagcagT	5'phosphate
12	T7-Ad3	gcctcgagctaatacgactcactatagag	none
4417	T7-Ad4	NNctctatagtgcgtatta	5'phosphate
4418	stlgR	ttagagctagaaatagcaagttaaataaggctagtcggttatcaactgaaaaagtggcaccgagtcggtgcttttt	5'adenylation

[0285] The DNA containing the CCD blunt ends (from earlier section) was then ligated to 50 pmoles of adapter BaeI/EcoP15I, using the blunt/TA ligation master mix (New England Biolabs) at room temperature for 30 minutes. The DNA was then recovered by incubating with 0.6X Kapa SPRI beads (Kapa Biosystems) for 5 minutes, capturing the beads with a magnetic rack, washing twice with 80% ethanol, air drying the beads for 5 minutes and finally resuspending the DNA in 50 µL buffer 4 (50 mM potassium acetate 20 mM Tris-acetate, 10 mM magnesium acetate, 100 µg/mL BSA, pH 7.9). Recovered DNA was then digested with 20 units PmeI for 30 min at 37 °C; DNA was then recovered by incubating with 1.2X Kapa SPRI beads (Kapa Biosystems) for 5 minutes, capturing the beads with a magnetic rack, washing twice with 80% ethanol, air drying the beads for 5 minutes and finally resuspending the DNA in 50 µL buffer 4. These steps eliminate small (<100 nucleotides) DNA and BaeI/EcoP15I adapter multimers.

[0286] DNA was then digested by addition of 20 units EcoP15I (New England Biolabs) and 1 mM ATP at 37 °C for 1 hour, followed by heat inactivation at 75 °C for 20 minutes. DNA was then ligated to 30 pmoles T7-E adapter using the blunt/TA ligation master mix (New England Biolabs) at room temperature for 30 minutes. DNA was then recovered using a PCR cleanup kit (Zymo) and eluted in 20 µL buffer 4.

[0287] Purified DNA was then digested by adding 20 units of BaeI (New England Biolabs), 40 pmol/µL SAM (S-adenosyl methionine) and incubating at 37 °C for 1 hour to eliminate both the adapter derived sequences and the CCD (and complementary HGG) motifs. DNA was then recovered using a PCR cleanup kit (Zymo) and eluted in 20 µL elution buffer.

[0288] Recovered DNA was then ligated to the stlgR oligo using Thermostable 5' AppDNA/RNA Ligase

[0289] (New England Biolabs) by adding 20 units ligase, 20 pmol stlgR oligo, in 20 µL ss ligation buffer (10 mM Bis-Tris-Propane-HCl, 10 mM MgCl₂, 1 mM DTT, 2.5 mM MnCl₂, pH 7 @ 25 °C) and incubating at 65 °C for 1 hour followed by heat inactivation at 90 °C for 5 min. DNA product was then PCR amplified using HiFidelity 2X master mix (New England Biolabs). Primers T7-Ad3 and gRU (sequence 5'>3' AAAAAAGCACCGACTCGGTG) (SEQ ID NO: 4419) were used to amplify with the following settings (98 °C 3 min; 98 °C for 20 sec, 60 °C for 30 secs, 72 °C for 20 sec, 30 cycles). The PCR amplicon was cleaned up using the PCR cleanup kit and verified by DNA sequencing, then used as template for an in vitro transcription reaction to generate the guide RNAs.

Example 6: NEMDA Method

[0290] NEMDA (Nicking Endonuclease Mediated DNA Amplification) was performed using 50 ng of human genomic DNA. The DNA was incubated in 100 µL thermo polymerase buffer (20 mM Tris-HCl, 10 mM (NH₄)₂SO₄, 10 mM KCl, 6 mM MgSO₄, 0.1% Triton® X-100, pH 8.8) supplemented with 0.3 mM dNTPs, 40 units of Bst large fragment DNA polymerase, and 0.1 units of NtCviPII (New England Biolabs) at 55 °C for 45 min, followed by 65 °C for 30 min and finally 80 °C for 20 min in a thermal cycler.

[0291] The DNA was then diluted with 300 µL of buffer 4 supplemented with 200 pmoles of T7-RND8 oligo (sequence 5'>3' gcctcgagctaatacgactcactatagagnnnnnnnn) (SEQ ID NO: 4420) and boiled at 98 °C for 10 min followed by rapid cooling to 10 °C for 5 min. The reaction was then supplemented with 40 units of *E. coli* DNA polymerase I and 0.1 mM dNTPs (New England Biolabs) and incubated at room temperature for 20 min followed by heat inactivation at 75 °C for 20 min. DNA was then recovered using a PCR cleanup kit (Zymo) and eluted in 30 µL elution buffer.

[0292] DNA was then ligated to 50 pmoles of adapter BaeI/EcoP15I, using the blunt/TA ligation master mix (New England Biolabs) at room temperature for 30 minutes. The DNA was then recovered by incubating with 0.6X Kapa SPRI beads (Kapa Biosystems) for 5 minutes, capturing the beads with a magnetic rack, washing twice with 80% ethanol, air drying the beads for 5 minutes and finally resuspending the DNA in 50 µL buffer 4 (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 100 µg/mL BSA, pH 7.9). Recovered DNA was then digested with 20 units PmeI for 30 min at 37 °C; DNA was then recovered by incubating with 1.2X Kapa SPRI beads (Kapa Biosystems) for 5 minutes, capturing the beads with a magnetic rack, washing twice with 80% ethanol, air drying the beads for 5 minutes and finally resuspending the DNA in 50 µL buffer 4. These steps eliminate small (<100 nucleotides) DNA and BaeI/EcoP15I adapter multimers.

[0293] Purified DNA was then digested by adding 20 units of BaeI (New England Biolabs), 40 pmol/ μ L SAM (S-adenosyl methionine) and incubating at 37 °C for 1 hour to eliminate both the adapter derived sequences and the CCD (and complementary HGG) motifs. DNA was then recovered using a PCR cleanup kit (Zymo) and eluted in 20 μ L elution buffer.

[0294] Recovered DNA was then ligated to the stlgR oligo using Thermostable 5' AppDNA/RNA Ligase (New England Biolabs) by adding 20 units ligase, 20 pmol stlgR oligo, in 20 μ L ss ligation buffer (10 mM Bis-Tris-Propane-HCl, 10 mM MgCl₂, 1 mM DTT, 2.5 mM MnCl₂, pH 7 @ 25 °C) and incubating at 65 °C for 1 hour followed by heat inactivation at 90 °C for 5 min. DNA product was then PCR amplified using HiFidelity 2X master mix (New England Biolabs). Primers T7-Ad3 (sequence 5'>3' gcctcgagctaatacgactcactatagag) (SEQ ID NO: 12) and gRU (sequence 5'>3' AAAAAGCACCGACTCGGTG) (SEQ ID NO: 4419) were used to amplify with the following settings (98 °C for 3 min; 98 °C for 20 sec, 60°C for 30 secs, 72 °C for 20 sec, 30 cycles). The PCR amplicon was cleaned up using the PCR cleanup kit and verified by DNA sequencing, then used as template for an in vitro transcription reaction to generate the guide RNAs.

CLAIMS

1. A collection of nucleic acids, a plurality of the nucleic acids in the collection comprising:
 - a. a first segment comprising a regulatory region;
 - b. a second segment encoding a targeting sequence; and
 - c. a third segment encoding a nucleic acid-guided nuclease system protein-binding sequence, wherein at least 10% of the nucleic acids in the collection vary in size.
2. The collection of claim 1, wherein the nucleic acid-guided nuclease system protein is a CRISPR/Cas system protein.
3. The collection of claim 1, wherein the size of the second segment varies from 15-250 bp across the collection of nucleic acids.
4. The collection of claim 1, wherein at least 10% of the second segments in the collection are greater than 21 bp.
5. The collection of claim 1, wherein the size of the second segment is not 20 bp.
6. The collection of claim 1, wherein the size of the second segment is not 21 bp.
7. The collection of claim 1, wherein the collection of nucleic acids is a collection of DNA.
8. The collection of claim 7, wherein the second segment is single stranded DNA.
9. The collection of claim 7, wherein the third segment is single stranded DNA.
10. The collection of claim 7, wherein the second segment is double stranded DNA.
11. The collection of claim 7, wherein the third segment is double stranded DNA.
12. The collection of claim 1, wherein the regulatory region is a region capable of binding a transcription factor.
13. The collection of claim 1, wherein the regulatory region comprises a promoter.
14. The collection of claim 13, wherein the promoter is selected from the group consisting of T7, SP6, and T3.
15. The collection of claim 1, wherein the targeting sequence is directed at a mammalian genome, eukaryotic genome, prokaryotic genome, or a viral genome.
16. The collection of claim 1, wherein the targeting sequence is directed at repetitive or abundant DNA.
17. The collection of claim 1, wherein the targeting sequence is directed at mitochondrial DNA, ribosomal DNA, Alu DNA, centromeric DNA, SINE DNA, LINE DNA, or STR DNA.
18. The collection of claim 1, wherein the sequence of the second segments is selected from Table 3 and/or Table 4.
19. The collection of claim 1, wherein the collection comprises at least 10^2 unique nucleic acid molecules.

20. The collection of claim 1, wherein the targeting sequence is at least 80% complementary to the strand opposite to a sequence of nucleotides 5' to a PAM sequence.
21. The collection of claim 1, wherein the collection comprises targeting sequences directed to sequences of interest spaced about every 10,000 bp or less across the genome of an organism.
22. The collection of claim 20, wherein the PAM sequence is AGG, CGG, or TGG.
23. The collection of claim 20, wherein the PAM sequence is specific for a CRISPR/Cas system protein selected from the group consisting of Cas9, Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, and Cm5.
24. The collection of claim 1, wherein the third segment comprises DNA encoding a gRNA stem-loop sequence.
25. The collection of claim 1, wherein the sequence of the third segment encodes for a RNA comprising the sequence
 GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA
 AAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 1) or encodes for a RNA comprising the sequence
 GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUC
 AACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUUC (SEQ ID NO: 2).
26. The collection of claim 1, wherein the sequence of the third segment encodes for a crRNA and a tracrRNA.
27. The collection of claim 1, wherein the nucleic acid-guided nuclease system protein is from a bacterial species.
28. The collection of claim 1, wherein the nucleic acid-guided nuclease system protein is from an archaea species.
29. The collection of claim 2, wherein the CRISPR/Cas system protein is a Type I, Type II, or Type III protein.
30. The collection of claim 2, wherein the CRISPR/Cas system protein is selected from the group consisting of Cas9, Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, Cm5, dCas9 and cas9 nickase.
31. The collection of claim 1, wherein the third segment comprises DNA encoding a Cas9-binding sequence.
32. The collection of claim 1, wherein a plurality of third segments of the collection encode for a first nucleic acid-guided nuclease system protein binding sequence, and a plurality of the third segments of the collection encode for a second nucleic acid-guided nuclease system protein binding sequence.

33. The collection of claim 1, wherein the third segments of the collection encode for a plurality of different binding sequences of a plurality of different nucleic acid-guided nuclease system proteins.
34. A collection of nucleic acids, a plurality of the nucleic acids in the collection comprising:
 - a. a first segment comprising a regulatory region;
 - b. a second segment encoding a targeting sequence, wherein the size of the second segment is greater than 21 bp; and
 - c. a third segment encoding a nucleic acid-guided nuclease system protein-binding sequence.
35. The collection of claim 34, wherein the nucleic acid-guided nuclease system protein is a CRISPR/Cas system protein.
36. The collection of claim 34, wherein at least 10% of the nucleic acids in the collection vary in size.
37. The collection of claim 34, wherein the size of the second segment varies from 22-250 bp across the collection of nucleic acids.
38. The collection of claim 34, wherein at least 10% of the second segments in the collection are greater than 30 bp.
39. The collection of claim 34, wherein the collection of nucleic acids is a collection of DNA.
40. The collection of claim 39, wherein the second segment is single stranded DNA.
41. The collection of claim 39, wherein the third segment is single stranded DNA.
42. The collection of claim 39, wherein the second segment is double stranded DNA.
43. The collection of claim 39, wherein the third segment is double stranded DNA.
44. The collection of claim 34, wherein the regulatory region is a region capable of binding a transcription factor.
45. The collection of claim 34, wherein the regulatory region comprises a promoter.
46. The collection of claim 45, wherein the promoter is selected from the group consisting of T7, SP6, and T3.
47. The collection of claim 34, wherein the targeting sequence is directed at a mammalian genome, eukaryotic genome, prokaryotic genome, or a viral genome.
48. The collection of claim 34, wherein the targeting sequence is directed at repetitive or abundant DNA.
49. The collection of claim 34, wherein the targeting sequence is directed at mitochondrial DNA, ribosomal DNA, Alu DNA, centromeric DNA, SINE DNA, LINE DNA, or STR DNA.
50. The collection of claim 34, wherein the sequence of the second segments is selected from Table 3 and/or Table 4.

51. The collection of claim 34, wherein the collection comprises at least 10^2 unique nucleic acid molecules.
52. The collection of claim 33, wherein the collection comprises targeting sequences directed to sequences of interest spaced about every 10,000 bp or less across the genome of an organism.
53. The collection of claim 34, wherein the targeting sequence is at least 80% complementary to the strand opposite to a sequence of nucleotides 5' to a PAM sequence.
54. The collection of claim 53, wherein the PAM sequence is AGG, CGG, or TGG.
55. The collection of claim 53, wherein the PAM sequence is specific for a CRISPR/Cas system protein selected from the group consisting of Cas9, Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, and Cm5.
56. The collection of claim 34, wherein the third segment comprises DNA encoding a gRNA stem-loop sequence.
57. The collection of claim 34, wherein the third segment encodes for a RNA comprising the sequence
GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAG
UGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 1) or encodes for a RNA comprising the sequence
GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAA
CUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUUC (SEQ ID NO: 2).
58. The collection of claim 34, wherein the third segment encodes for a crRNA and a tracrRNA.
59. The collection of claim 34, wherein the nucleic acid-guided nuclease system protein is from a bacterial species.
60. The collection of claim 34, wherein the nucleic acid-guided nuclease system protein is from an archaea species.
61. The collection of claim 35, wherein the CRISPR/Cas system protein is a Type I, Type II, or Type III protein.
62. The collection of claim 35, wherein the CRISPR/Cas system protein is selected from the group consisting of Cas9, Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, Cm5, dCas9 and cas9 nickase.
63. The collection of claim 34, wherein the third segment comprises DNA encoding a Cas9-binding sequence.
64. The collection of claim 34, wherein a plurality of third segments of the collection encode for a first nucleic acid-guided nuclease system protein binding sequence, and a plurality of the third segments of the collection encode for a second nucleic acid-guided nuclease system protein binding sequence.

65. The collection of claim 34, wherein the third segments of the collection encode for a plurality of different binding sequences of a plurality of different nucleic acid-guided nuclease system proteins.
66. A collection of guide RNAs (gRNAs), comprising:
- a. a first RNA segment comprising a targeting sequence; and
 - b. a second RNA segment comprising a nucleic acid-guided nuclease system protein-binding sequence,
- wherein at least 10% of the gRNAs in the collection vary in size.
67. The collection of claim 66, wherein the nucleic acid-guided nuclease system protein is a CRISPR/Cas system protein.
68. The collection of claim 66, wherein the size of the first segment varies from 15-250 bp across the collection of gRNAs.
69. The collection of claim 66, wherein at least 10% of the first segments in the collection are greater than 21 bp.
70. The collection of claim 66, wherein the size of the first segment is not 20 bp.
71. The collection of claim 66, wherein the size of the first segment is not 21 bp.
72. The collection of claim 66, wherein the targeting sequence is directed at a mammalian genome, eukaryotic genome, prokaryotic genome, or viral genome.
73. The collection of claim 66, wherein the targeting sequence is directed at repetitive or abundant DNA.
74. The collection of claim 66, wherein the targeting sequence is directed at mitochondrial DNA, ribosomal DNA, Alu DNA, centromeric DNA, SINE DNA, LINE DNA, or STR DNA.
75. The collection of claim 66, wherein the sequence of the first segments is RNA encoded by sequences selected from Table 3 and/or Table 4.
76. The collection of claim 66, wherein the collection comprises at least 10^2 unique gRNAs.
77. The collection of claim 66, wherein the gRNAs comprise cytosine, guanine, and adenine.
78. The collection of claim 77, wherein a subset of the gRNAs further comprise thymine.
79. The collection of claim 77, wherein a subset of the gRNAs further comprise uracil.
80. The collection of claim 66, wherein the first segment is at least 80% complementary to a target genomic sequence of interest.
81. The collection of claim 66, wherein the targeting sequence is at least 80% complementary to the strand opposite to a sequence of nucleotides 5' to a PAM sequence.
82. The collection of claim 66, wherein the PAM sequence is specific for a CRISPR/Cas system protein selected from the group consisting of Cas9, Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, and Cm5.

83. The collection of claim 66, wherein the second segment comprises a gRNA stem-loop sequence.
84. The collection of claim 66, wherein the sequence of the second segment comprises
 GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA
 AAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 1) or comprises
 GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUC
 AACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUUC (SEQ ID NO: 2).
85. The collection of claim 66, wherein the second segment comprises a crRNA and a tracrRNA.
86. The collection of claim 66, wherein the nucleic acid-guided nuclease system protein is from a bacterial species.
87. The collection of claim 66, wherein the nucleic acid-guided nuclease system protein is from an archaea species.
88. The collection of claim 67, wherein the CRISPR/Cas system protein is a Type I, Type II, or Type III protein.
89. The collection of claim 67, wherein the CRISPR/Cas system protein is selected from the group consisting of Cas9, Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, Cm5, dCas9 and cas9 nickase, and their engineered versions
90. The collection of claim 66, wherein the second segment comprises a Cas9-binding sequence.
91. The collection of claim 66, wherein at least 10% of the gRNAs in the collection vary in their 5' terminal-end sequence.
92. The collection of claim 66, wherein the collection comprises targeting sequences directed to sequences of interest spaced every 10,000 bp or less across the genome of an organism.
93. The collection of claim 66, wherein a plurality of second segments of the collection comprise a first nucleic acid-guided nuclease system protein binding sequence, and a plurality of the second segments of the collection comprise a second nucleic acid-guided nuclease system protein binding sequence.
94. The collection of claim 66, wherein the second segments of the collection comprise a plurality of different binding sequences of a plurality of different nucleic acid-guided nuclease system proteins.
95. The collection of claim 66, wherein a plurality of the gRNAs of the collection are attached to a substrate.
96. The collection of claim 66, wherein a plurality of the gRNAs of the collection comprise a label.
97. The collection of claim 96, wherein the plurality of the gRNAs of the collection comprise different labels.
98. A nucleic acid comprising:
 a. a first segment comprising a regulatory region;

- b. a second segment encoding a targeting sequence, wherein the targeting sequence is greater than 30 bp; and
 - c. a third segment encoding a nucleic acid encoding a nucleic acid-guided nuclease system protein-binding sequence.
99. The nucleic acid of claim 98, wherein the nucleic-acid guided nuclease system protein is a CRISPR/Cas system protein.
100. The nucleic acid of claim 98, wherein the nucleic acid is DNA.
101. The nucleic acid of claim 98, wherein the second segment is single stranded DNA.
102. The nucleic acid of claim 98, wherein the third segment is single stranded DNA.
103. The nucleic acid of claim 98, wherein the second segment is double stranded DNA.
104. The nucleic acid of claim 98, wherein the third segment is double stranded DNA.
105. The nucleic acid of claim 98, wherein the regulatory region is a region capable of binding a transcription factor.
106. The nucleic acid of claim 98, wherein the regulatory region comprises a promoter.
107. The collection of claim 106, wherein the promoter is selected from the group consisting of T7, SP6, and T3.
108. The nucleic acid of claim 98, wherein the targeting sequence is directed at a mammalian genome, eukaryotic genome, prokaryotic genome, or a viral genome.
109. The nucleic acid of claim 98, wherein the targeting sequence is directed at abundant or repetitive DNA.
110. The nucleic acid of claim 98, wherein the targeting sequence is directed at mitochondrial DNA, ribosomal DNA, Alu DNA, centromeric DNA, SINE DNA, LINE DNA, or STR DNA.
111. The nucleic acid of claim 98, targeting sequence is at least 80% complementary to the strand opposite to a sequence of nucleotides 5' to a PAM sequence.
112. The nucleic acid of claim 111, wherein the target genomic sequence of interest is 5' upstream of a PAM sequence.
113. The nucleic acid of claim 112, wherein the PAM sequence is specific for a CRISPR/Cas system protein selected from the group consisting of Cas9, Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, and Cm5.
114. The nucleic acid of claim 98, wherein the third segment comprises DNA encoding a gRNA stem-loop sequence.
115. The nucleic acid of claim 98, wherein the sequence of the third segment encodes for a RNA comprising the sequence
GUUUUAGAGCUAGAAAUAGCAAGUUA AAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAG

UGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 1) or encodes for a RNA comprising the sequence

GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUA AAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUUUC (SEQ ID NO: 2).

116. The nucleic acid of claim 98, wherein the third segment encodes for a crRNA and a tracrRNA.
117. The nucleic acid of claim 98, wherein the nucleic acid-guided nuclease system protein is from a bacterial species.
118. The nucleic acid of claim 98, wherein the nucleic acid-guided nuclease system protein is from an archaea species.
119. The nucleic acid of claim 99, wherein the CRISPR/Cas system protein is a Type I, Type II, or Type III protein.
120. The nucleic acid of claim 99, wherein the CRISPR/Cas system protein is selected from the group consisting of Cas9, Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, Cm5, dCas9 and cas9 nickase.
121. The nucleic acid of claim 98, wherein the third segment comprises DNA encoding a Cas9-binding sequence.
122. A guide RNA comprising:
 - a. a first segment comprising a targeting sequence, wherein the size of the first segment is greater than 30 bp; and
 - b. a second segment comprising a nucleic acid-guided nuclease system protein-binding sequence.
123. The guide RNA of claim 122, wherein the nucleic acid-guided nuclease system protein is a CRISPR/Cas system protein.
124. The guide RNA of claim 122, wherein the gRNA comprises an adenine, a guanine, and a cytosine.
125. The guide RNA of claim 123, wherein the gRNA further comprises a thymine.
126. The guide RNA of claim 123, wherein the gRNA further comprises a uracil.
127. The guide RNA of claim 122, wherein the size of the first RNA segment is between 30 and 250 bp.
128. The guide RNA of claim 122, wherein the targeting sequence is directed at a mammalian genome, eukaryotic genome, prokaryotic genome, or viral genome.
129. The guide RNA of claim 122, wherein the targeting sequence is directed at repetitive or abundant DNA.

130. The guide RNA of claim 122, wherein the targeting sequence is directed at mitochondrial DNA, ribosomal DNA, Alu DNA, centromeric DNA, SINE DNA, LINE DNA, or STR DNA.
131. The guide RNA of claim 122, wherein the first segment is at least 80% complementary to the target genomic sequence of interest.
132. The guide RNA of claim 122, wherein the targeting sequence is at least 80% complementary to the strand opposite to a sequence of nucleotides 5' to a PAM sequence.
133. The guide RNA of claim 122, wherein the second segment comprises a gRNA stem-loop sequence.
134. The guide RNA of claim 122, wherein the sequence of the second segment comprises
 GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA
 AAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 1) or comprises
 GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUC
 AACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUUC (SEQ ID NO: 2).
135. The guide RNA of claim 122, wherein the sequence of the second segment comprises a crRNA and a tracrRNA.
136. The guide RNA of claim 122, wherein the nucleic acid-guided nuclease system protein is from a bacterial species.
137. The guide RNA of claim 122, wherein the nucleic acid-guided nuclease system protein is from an archaea species.
138. The guide RNA of claim 123, wherein the CRISPR/Cas system protein is a Type I, Type II, or Type III protein.
139. The guide RNA of claim 123, wherein the CRISPR/Cas system protein is selected from the group consisting of Cas9, Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, Cm5, dCas9 and cas9 nickase, and their engineered versions.
140. The guide RNA of claim 122, wherein the second segment is a Cas9-binding sequence.
141. A complex comprising a Cas9 protein and a gRNA of claim 122.
142. A method for depleting and partitioning of targeted sequences in a sample, enriching a sample for non-host nucleic acids, or serially depleting targeted nucleic acids in a sample comprising:
- a. providing nucleic acids extracted from a sample; and
 - b. contacting the sample with a plurality of complexes comprising (i) the collection of gRNAs of claim 66; and (ii) nucleic-acid guided nuclease system proteins.
143. The method of claim 142, wherein the nucleic-acid guided nuclease system proteins comprise CRISPR/Cas system proteins.

144. The method of claim 143, wherein the CRISPR/Cas system proteins comprise Cas9.
145. A method of making a collection of nucleic acids, each comprising a DNA encoding a targeting sequence ligated to a DNA encoding a nucleic acid-guided nuclease system protein-binding sequence, comprising:
- a. providing double-stranded DNA molecules, each comprising a sequence of interest 5' to a PAM sequence, and its reverse complementary sequence on the opposite strand;
 - b. performing an enzymatic digestion reaction on the double stranded DNA molecules, wherein cleavages are generated at the PAM sequence and/or its reverse complementary sequence on the opposite strand, but never completely remove the PAM sequence and/or its reverse complementary sequence on the opposite strand from the double stranded DNA;
 - c. ligating adapters comprising a recognition sequence to the resulting DNA molecules of step (b);
 - d. contacting the DNA molecules of step (c) with a restriction enzyme that recognizes the recognition sequence of step (c), whereby generating DNA fragments comprising blunt-ended double strand breaks immediately 5' to the PAM sequence, whereby removing the PAM sequence and the adapter containing the enzyme recognition site; and
 - e. ligating the resulting double stranded DNA fragments of step (d) with a DNA encoding a nucleic acid-guided nuclease system protein-binding sequence, whereby generating a plurality of DNA fragments, each comprising a DNA encoding a targeting sequence ligated to a DNA encoding a nucleic acid-guided nuclease system protein-binding sequence.
146. The method of claim 145, wherein the nucleic acid-guided nuclease system protein is a CRISPR/Cas system protein.
147. The method of claim 145, wherein the starting DNA molecules of the collection further comprise a regulatory sequence upstream of the sequence of interest 5' to the PAM sequence.
148. The method of claim 146, wherein the regulatory sequence comprises a promoter.
149. The method of claim 148, wherein the promoter comprises a T7, Sp6, or T3 sequence.
150. The method of claim 145, wherein the double stranded DNA molecules are genomic DNA, intact DNA, or sheared DNA.
151. The method of claim 150, wherein the genomic DNA is human, mouse, avian, fish, plant, insect, bacterial, or viral.
152. The method of claim 145, wherein the DNA segments encoding a targeting sequence are at least 22 bp.
153. The method of claim 145, wherein the DNA segments encoding a targeting sequence are 15-250 bp in size range.

154. The method of claim 145, wherein the PAM sequence is AGG, CGG, or TGG.
155. The method of claim 145, wherein the PAM sequence is specific for a CRISPR/Cas system protein selected from the group consisting of Cas9, Cpf1, Cas3, Cas8a-c, Cas10, Cse1, Csy1, Csn2, Cas4, Csm2, and Cm5.
156. The method of claim 145, wherein step (b) further comprises (1) contacting the DNA molecules with an enzyme capable of creating a nick in a single strand at a CCD site, whereby generating a plurality of nicked double stranded DNA molecules, each comprising a sequence of interest followed by an HGG sequence, wherein the DNA molecules are nicked at the CCD sites; and (2) contacting the nicked double stranded DNA molecules with an endonuclease, whereby generating a plurality of double stranded DNA fragments, each comprising a sequence of interest followed by an HGG sequence wherein residual nucleotides from HGG and/or CCD sequences is (are) left behind .
157. The method of claim 145, wherein step (d) further comprises PCR amplification of the adaptor-ligated DNA fragments from step (c) before cutting with the restriction enzyme recognizing the recognition sequence of step (c), wherein after PCR, the recognition sequence is positioned 3' of the PAM sequence, and a regulatory sequence is positioned at the 5' distal end of the PAM sequence.
158. The method of claim 145, wherein the enzymatic reaction of step (b) comprises the use of a Nt.CviPII enzyme, and a T7 Endonuclease I enzyme.
159. The method of claim 145, wherein step (c) further comprises a blunt-end reaction with a T4 DNA Polymerase, if the adapter to be ligated does not comprise an overhang.
160. The method of claim 145, wherein the adapter of step (c) is either (1) double stranded, comprising a restriction enzyme recognition sequence in one strand, and a regulatory sequence in the other strand, if the adapter is Y-shaped and comprises an overhang; or (2) has a palindromic enzyme recognition sequence in both strands, if the adapter is not Y-shaped.
161. The method of claim 145, wherein the restriction enzyme of step (d) is MlyI.
162. The method of claim 145, wherein step (d) further comprises contacting the DNA molecules with an XhoI enzyme
163. The method of claim 145, wherein in step (e) the DNA encoding a nucleic acid-guided nuclease system-protein binding sequence encodes for a RNA comprising the sequence
 GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA
 AAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 1) or encodes for a RNA comprising the sequence
 GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUC
 AACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUUUC (SEQ ID NO: 2).
164. The method of claim 145, wherein the targeted sequences of interest are spaced every 10,000 bp or less across the genome of an organism.

165. A method of making a collection of nucleic acids, each comprising a DNA encoding a targeting sequence ligated to a DNA encoding a nucleic acid-guided nuclease system protein-binding sequence, comprising:

- a. providing a plurality of double stranded DNA molecules, each comprising a sequence of interest, an NGG site, and its complement CCN site;
- b. contacting the molecules with an enzyme capable of creating a nick in a single strand at a CCN site, whereby generating a plurality of nicked double stranded DNA molecules, each comprising a sequence of interest 5' to the NGG site, wherein the DNA molecules are nicked at the CCD sites;
- c. contacting the nicked double stranded DNA molecules with an endonuclease, whereby generating a plurality of double stranded DNA fragments, each comprising a sequence of interest, wherein the fragments comprise an terminal overhang;
- d. contacting the double stranded DNA fragments with an enzyme without 5' to 3' exonuclease activity to blunt end the double stranded DNA fragments, whereby generating a plurality of blunt ended double stranded fragments, each comprising a sequence of interest;
- e. contacting the blunt ended double stranded fragments of step d with an enzyme that cleaves the terminal NGG site; and
- f. ligating the resulting double stranded DNA fragments of step e with a DNA encoding a nucleic acid-guided nuclease system-protein binding sequence, whereby generating a plurality of DNA fragments, each comprising a targeting sequence ligated to a DNA encoding a nucleic acid-guided nuclease system protein-binding sequence.

166. The method of claim 165, wherein the nucleic acid-guided nuclease system protein is a CRISPR/Cas system protein.

167. The method of claim 165, wherein the plurality of double stranded DNA molecules have a regulatory sequence 5' upstream of the NGG sites.

168. The method of claim 166, wherein the regulatory sequence comprises a T7, SP6, or T3 sequence.

169. The method of claim 153, wherein the NGG site comprises AGG, CGG, or TGG, and the CCN site comprises CCT, CCG, or CCA.

170. The method of claim 165, wherein the plurality of double stranded DNA molecules, each comprising a sequence of interest comprise sheared fragments of genomic DNA.

171. The method of claim 170, wherein the genomic DNA is mammalian, prokaryotic, eukaryotic, avian, bacterial or viral.

172. The method of claim 165, wherein the plurality of double stranded DNA molecules in step (a) are at least 500 bp.

173. The method of claim 165, wherein the enzyme in step b is a Nt.CviPII enzyme
174. The method of claim 165, wherein the enzyme in step c is a T7 Endonuclease I.
175. The method of claim 165, wherein the enzyme in step d is a T4 DNA Polymerase.
176. The method of claim 165, wherein in step f the DNA encoding a nucleic acid-guided nuclease system-protein binding sequence encodes for a RNA comprising the sequence
GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAG
UGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 1) or encodes for a RNA comprising the
sequence
GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAA
CUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUUC (SEQ ID NO: 2).
177. The method of claim 165, wherein the step e additionally comprises ligating adaptors carrying a MlyI recognition site and digesting with MlyI enzyme.
178. The method of claim 165, wherein the sequence of interest is spaced every 10,000 bp or less across the genome.
179. A method of making a collection of nucleic acids, each comprising a DNA encoding a targeting sequence and a DNA encoding a nucleic acid-guided nuclease system protein-binding sequence, comprising:
- providing genomic DNA comprising a plurality of sequences of interest, comprising NGG and CCN sites;
 - contacting the genomic DNA with an enzyme capable of creating nicks in the genomic DNA, whereby generating nicked genomic DNA, nicked at CCN sites;
 - contacting the nicked genomic DNA with an endonuclease, whereby generating double stranded DNA fragments, with an overhang;
 - ligating the DNA with overhangs from (step c) to a Y-shaped adapter, thereby introducing a restriction enzyme recognition sequence only at 3' of the NGG site and a regulatory sequence 5' of the sequence of interest;
 - contacting the product from step d with an enzyme that cleaves away the NGG site together with the adaptor carrying the enzyme recognition sequence; and
 - ligating the resulting double stranded DNA fragments of step e with a DNA encoding a nucleic acid-guided nuclease system protein-binding sequence, whereby generating a plurality of DNA fragments, each comprising a sequence of interest ligated to a DNA encoding a nucleic acid-guided nuclease system protein-binding sequence.
180. The method of claim 179, wherein the nucleic acid-guided nuclease system protein is a CRISPR/Cas system protein.

181. The method of claim 179, wherein the NGG site comprises AGG, CGG, or TGG, and CCN site comprises CCT, CCG, or CCA.
182. The method of claim 179, wherein the regulatory sequence comprises a promoter sequence.
183. The method of claim 182, wherein the promoter sequence comprises a T7, SP6, or T3 sequence.
184. The method of claim 179, wherein the DNA fragments are sheared fragments of genomic DNA.
185. The method of claim 179, wherein the genomic DNA is mammalian, prokaryotic, eukaryotic, or viral.
186. The method of claim 179, wherein the fragments are at least 200 bp.
187. The method of claim 179, wherein the enzyme in step b is a Nt.CviPII enzyme.
188. The method of claim 179, wherein the enzyme in step c is a T7 Endonuclease I.
189. The method of claim 179, wherein step d further comprises PCR amplification of the adaptor-ligated DNA.
190. The method of claim 179, wherein in step f, the DNA encoding nucleic acid-guided nuclease system protein-binding sequence encodes for a RNA comprising the sequence
GUUUUAGAGCUAGAAAUAGCAAGUUA AAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAG
UGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO: 1) or encodes for a RNA comprising the
sequence
GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUA AAAUAAGGCUAGUCCGUUAUCAA
CUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUUC (SEQ ID NO: 2).
191. The method of claim 179, wherein the enzyme removing NGG site in step e is MlyI.
192. The method of claim 179, wherein the target of interest of the collection is spaced every 10,000 bp or less across the genome.
193. A kit comprising the collection of nucleic acids of claim 1.
194. A kit comprising the collection of nucleic acids of claim 34.
195. A kit comprising the collection of guide RNAs of claim 66.
196. A method of making a collection of guide nucleic acids, comprising:
 - a. obtaining abundant cells in a source sample;
 - b. collecting nucleic acids from said abundant cells; and
 - c. preparing a collection of guide nucleic acids (gNAs) from said nucleic acids.
197. The method of claim 196, wherein said abundant cells comprise cells from one or more most abundant bacterial species in said source sample.
198. The method of claim 196, wherein said abundant cells comprise cells from more than one species.
199. The method of claim 196, wherein said abundant cells comprise human cells.

200. The method of claim 196, wherein said abundant cells comprise animal cells.
201. The method of claim 196, wherein said abundant cells comprise plant cells.
202. The method of claim 196, wherein said abundant cells comprise bacterial cells.
203. The method of claim 196, further comprising contacting nucleic acid-guided nucleases with said library of gNAs to form nucleic acid-guided nuclease-gNA complexes.
204. The method of claim 203, further comprising using said nucleic acid-guided nuclease-gNA complexes to cleave target nucleic acids at target sites, wherein said gNAs are complementary to said target sites.
205. The method of claim 204, wherein said target nucleic acids are from said source sample.
206. The method of claim 204, wherein a species of said target nucleic acids is the same as a species of said source sample.
207. The method of claim 206, wherein said species of said target nucleic acids and said species of said source sample is human.
208. The method of claim 206, wherein said species of said target nucleic acids and said species of said source sample is animal.
209. The method of claim 206, wherein said species of said target nucleic acids and said species of said source sample is plant.
210. A method of making a collection of nucleic acids, each comprising a targeting sequence, comprising:
- a. obtaining source DNA;
 - b. nicking said source DNA with a nicking enzyme at nicking enzyme recognition sites, thereby producing double-stranded breaks at proximal nicks; and
 - c. repairing overhangs of said double-stranded breaks, thereby producing a double-stranded fragment comprising (i) a targeting sequence and (ii) said nicking enzyme recognition site.
211. A method of making a collection of nucleic acids, each comprising a targeting sequence, comprising:
- a. obtaining source DNA;
 - b. nicking said source DNA with a nicking enzyme at nicking enzyme recognition sites, thereby producing a nick; and
 - c. synthesizing a new strand from said nick, thereby producing a single-stranded fragment of said source DNA comprising a targeting sequence.
212. The method of claim 211, further comprising producing a double-stranded fragment comprising said targeting sequence from said single-stranded fragment.
213. The method of claim 212, wherein said producing said double-stranded fragment comprises random priming and extension.

214. The method of claim 213, wherein said random priming is conducted with a primer comprising a random n-mer region and a promoter region.
215. The method of claim 214, wherein said random n-mer region is a random hexamer region.
216. The method of claim 214, wherein said random n-mer region is a random octamer region.
217. The method of claim 214, wherein said promoter region is a T7 promoter region.
218. The method of claim 210 or 212, further comprising ligating a nuclease recognition site nucleic acid comprising a nuclease recognition site to said double-stranded fragment.
219. The method of claim 218, wherein said nuclease recognition site corresponds to a nuclease that cuts at a distance from said nuclease recognition site equal to the length of said nicking enzyme recognition sites.
220. The method of claim 219, wherein said nuclease recognition site is a MlyI recognition site.
221. The method of claim 219, wherein said nuclease recognition site is a BaeI recognition site.
222. The method of claim 219, further comprising digesting said double-stranded fragment with said nuclease, thereby removing said nicking enzyme recognition site from said double-stranded fragment.
223. The method of claim 222, further comprising ligating said double-stranded fragment to a nucleic acid-guided nuclease system protein recognition site nucleic acid comprising a nucleic acid-guided nuclease system protein recognition site.
224. The method of claim 223, wherein said nucleic acid-guided nuclease system protein recognition site comprises a guide RNA stem-loop sequence.
225. The method of claim 218, wherein said nuclease recognition site corresponds to a nuclease that cuts at a distance from said nuclease recognition site equal to a length of said targeting sequence.
226. The method of claim 225, wherein said length of said targeting sequence is 20 base pairs.
227. The method of claim 225, wherein said nuclease recognition site is a MmeI recognition site.
228. The method of claim 225, further comprising digesting said double-stranded fragment with said nuclease.
229. The method of claim 218, wherein said nuclease recognition site corresponds to a nuclease that cuts at a distance from said nuclease recognition site equal to a length of said targeting sequence plus a length of said nicking enzyme recognition sites.
230. The method of claim 229, wherein said length of said targeting sequence plus a length of said nicking enzyme recognition sites is 23 base pairs.
231. The method of claim 229, wherein said nuclease recognition site is a EcoP15I recognition site.
232. The method of claim 229, further comprising digesting said double-stranded fragment with said nuclease.

233. The method of claim 212, further comprising ligating said double-stranded fragment to a nucleic acid-guided nuclease system protein recognition site nucleic acid comprising a nucleic acid-guided nuclease system protein recognition site.

234. The method of claim 233, wherein said nucleic acid-guided nuclease system protein recognition site comprises a guide RNA stem-loop sequence.

235. A kit comprising all essential reagents and instructions for carrying out the methods of claim 145, 165, 179, 196, 210 or 211.

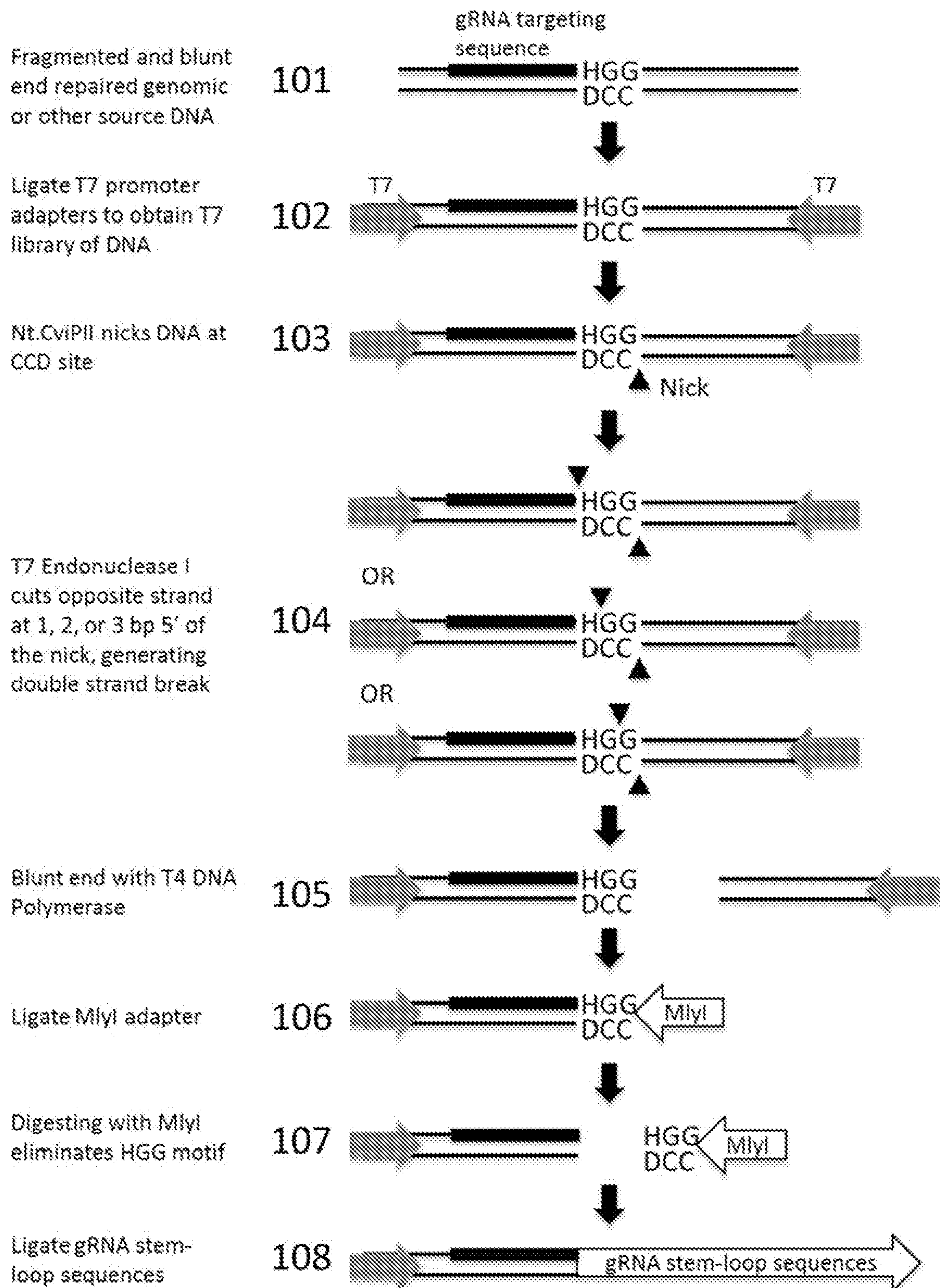
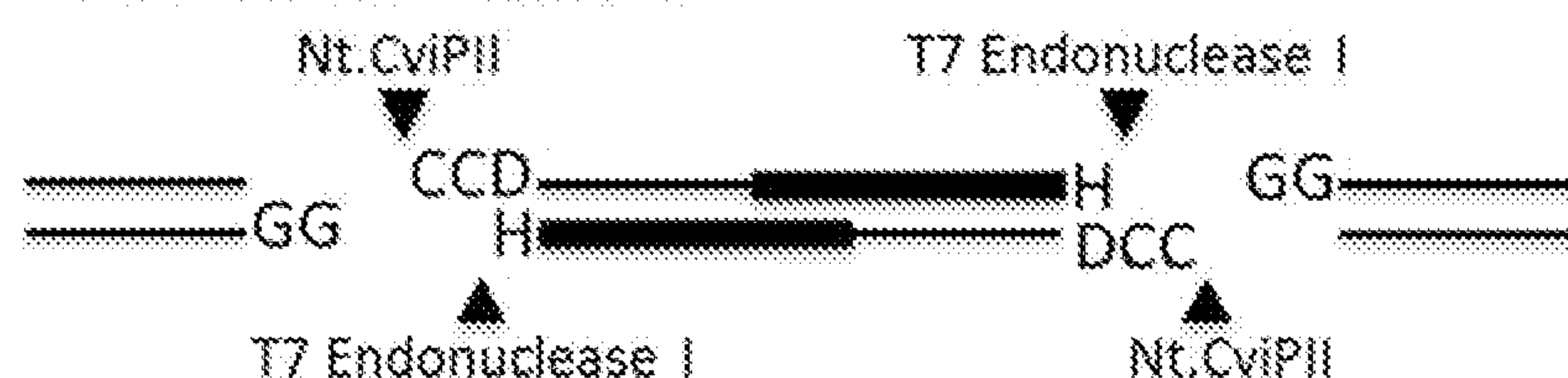


FIG. 1

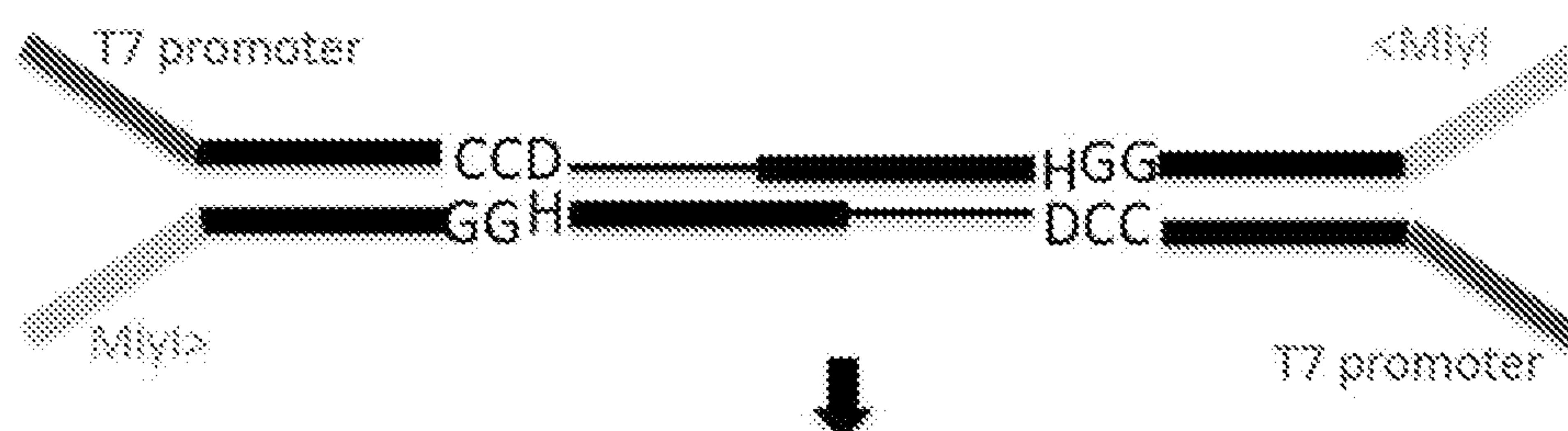
201 Genomic or other source DNA



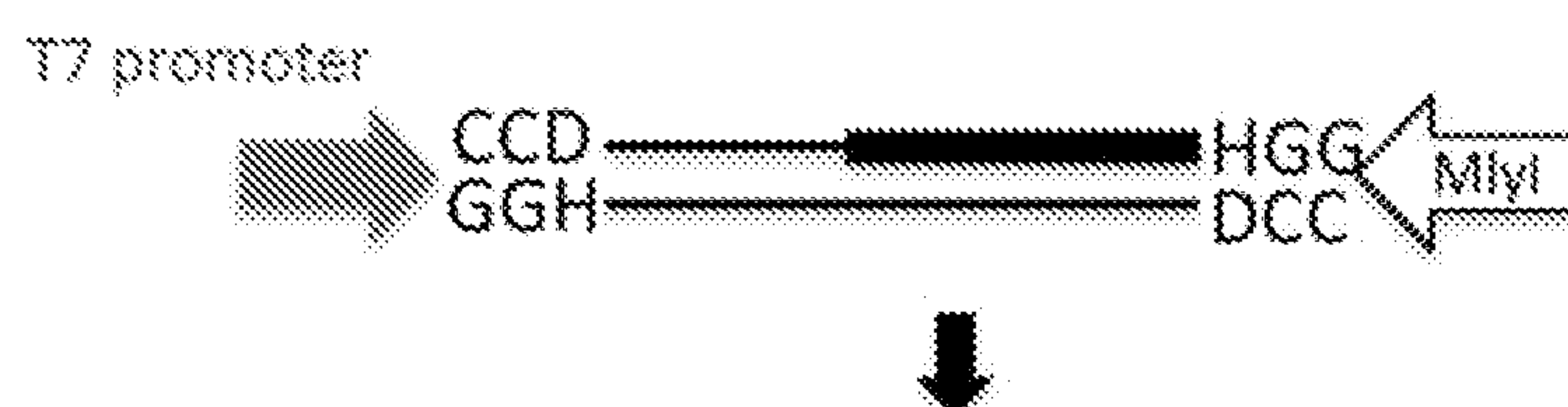
202 Nt.CviPII and T7 Endonuclease I treatment



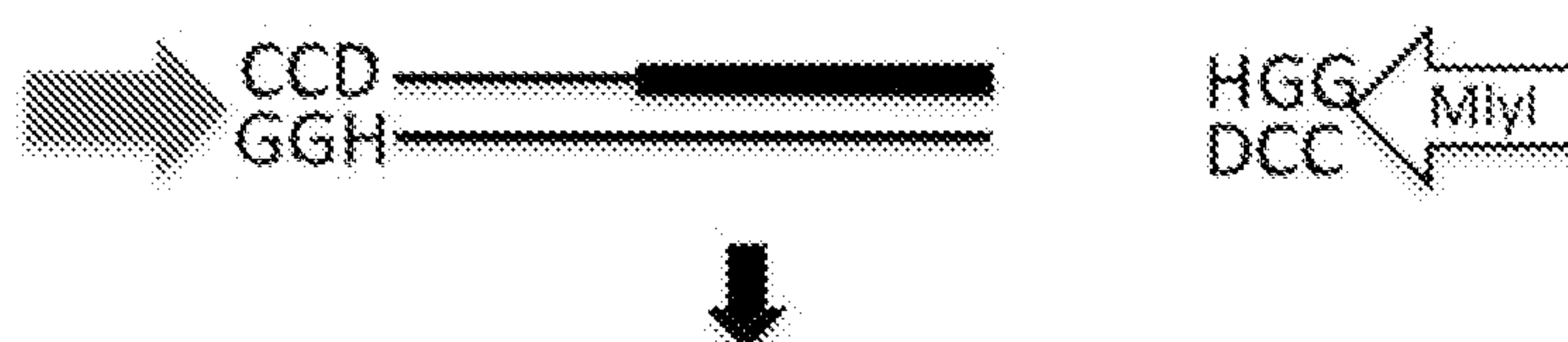
203 Ligate T7/MlyI adapters



204 PCR amplify with T7 and MlyI primers



205 Digest with MlyI, separates HGG from gRNA targeting sequence



206 Ligate gRNA stem-loop sequences



FIG. 2

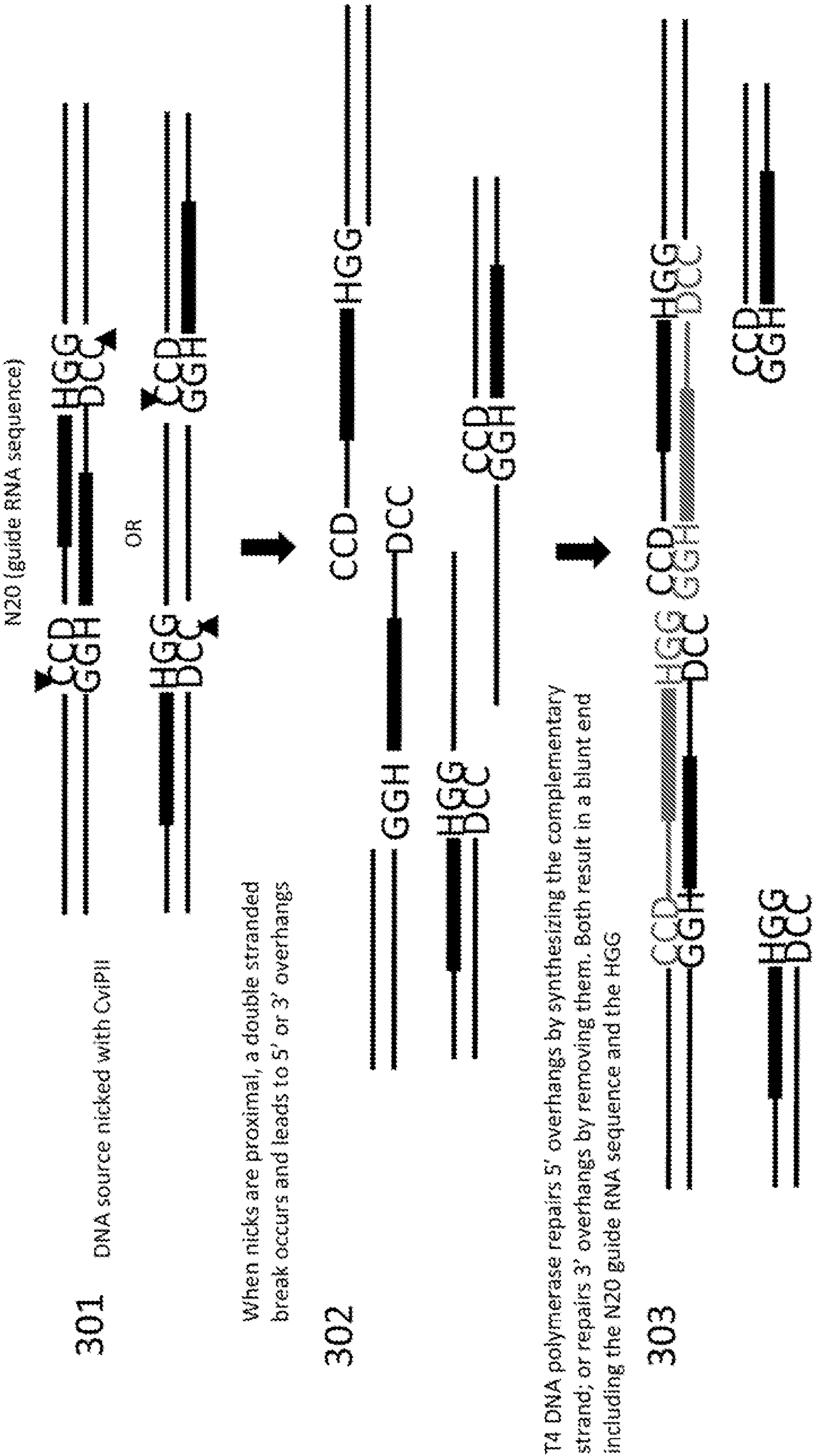


FIG. 3

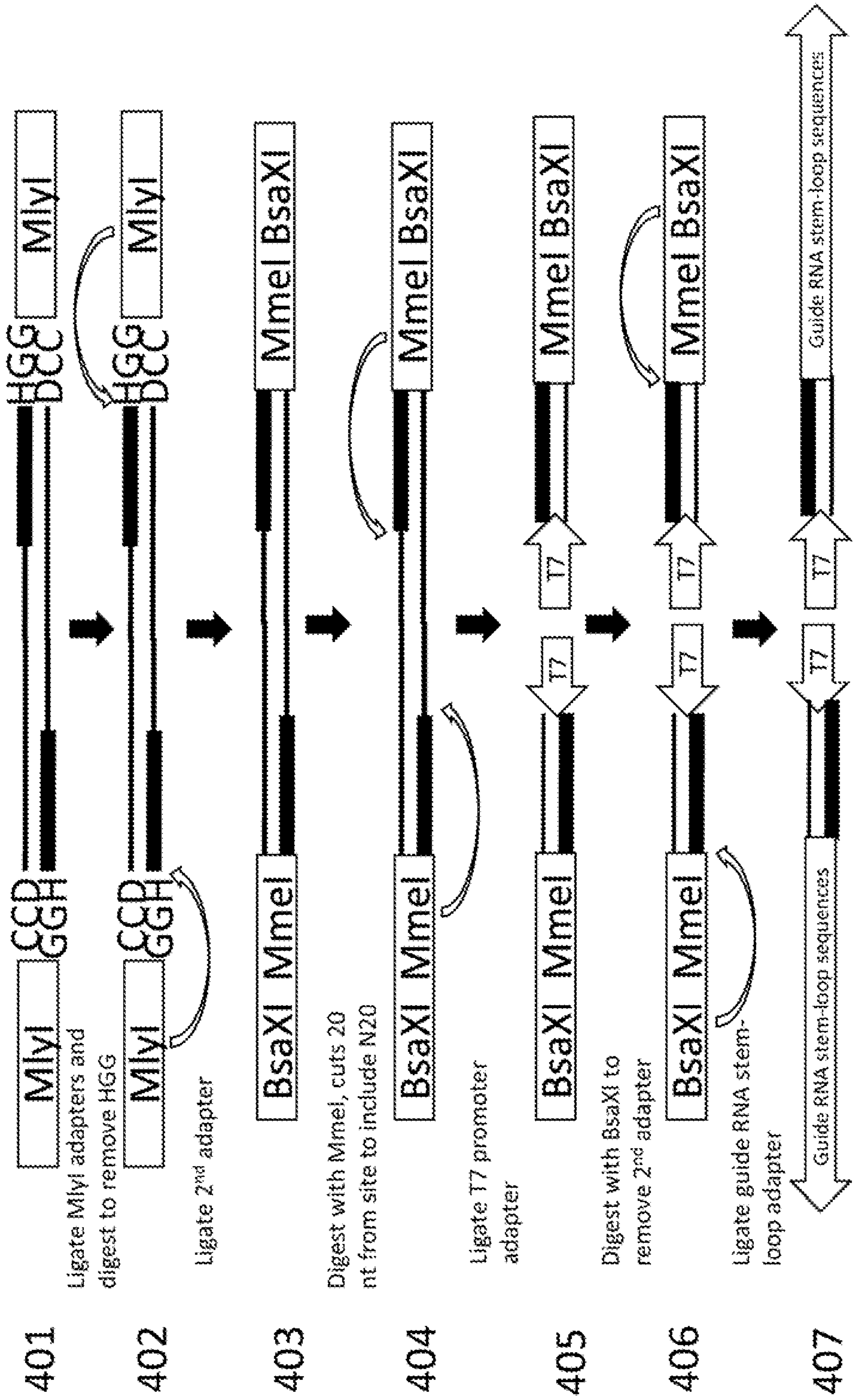


FIG. 4

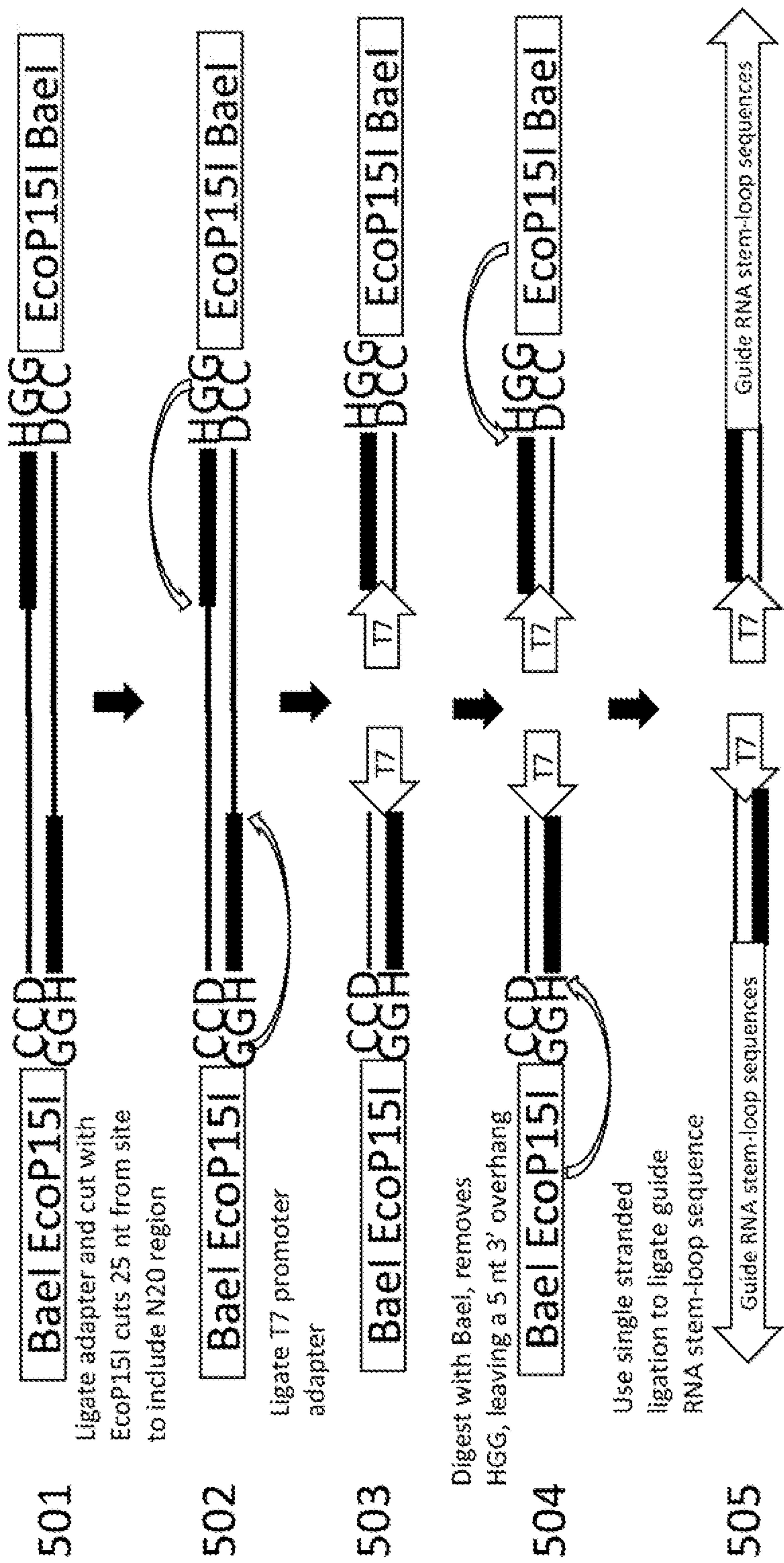


FIG. 5

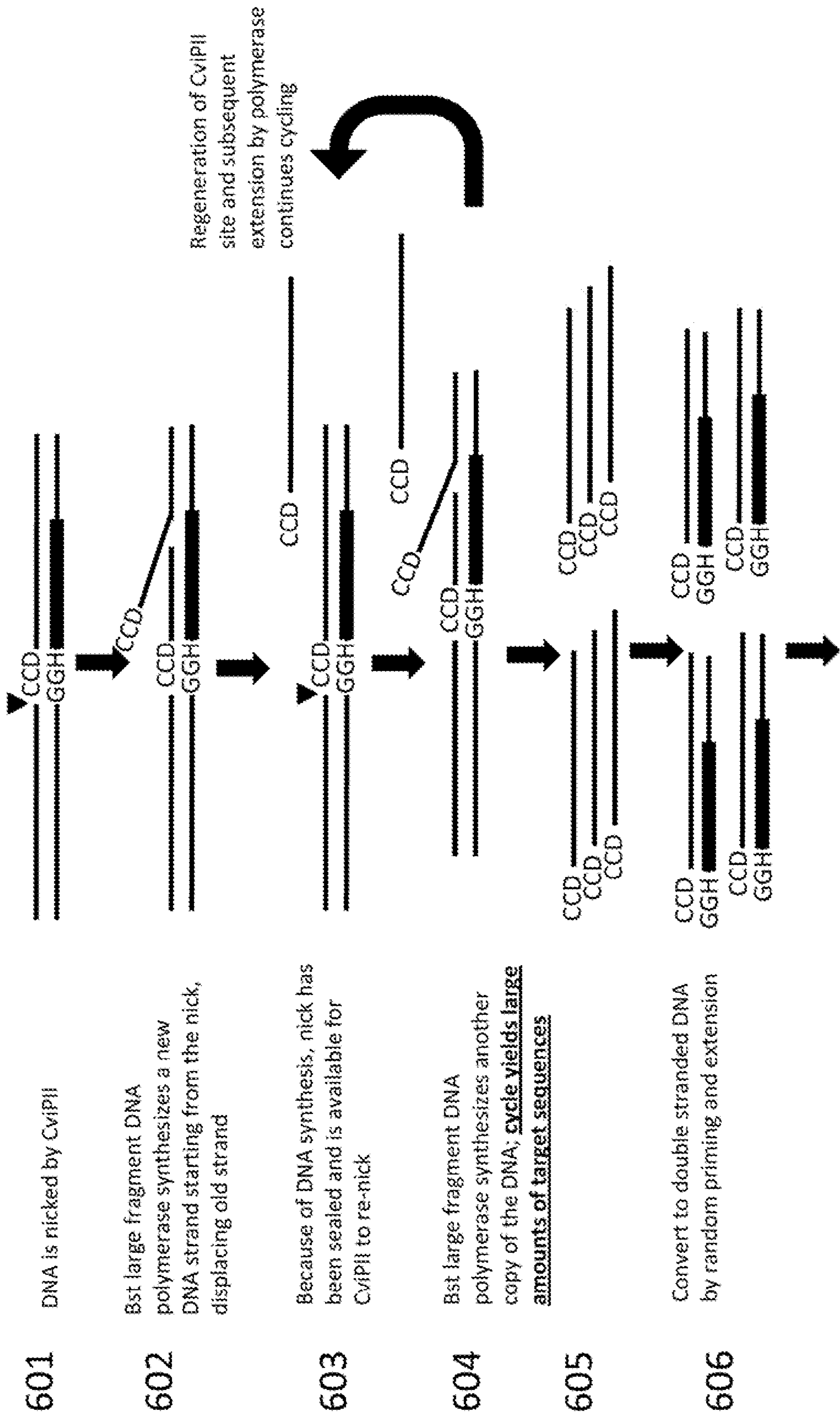
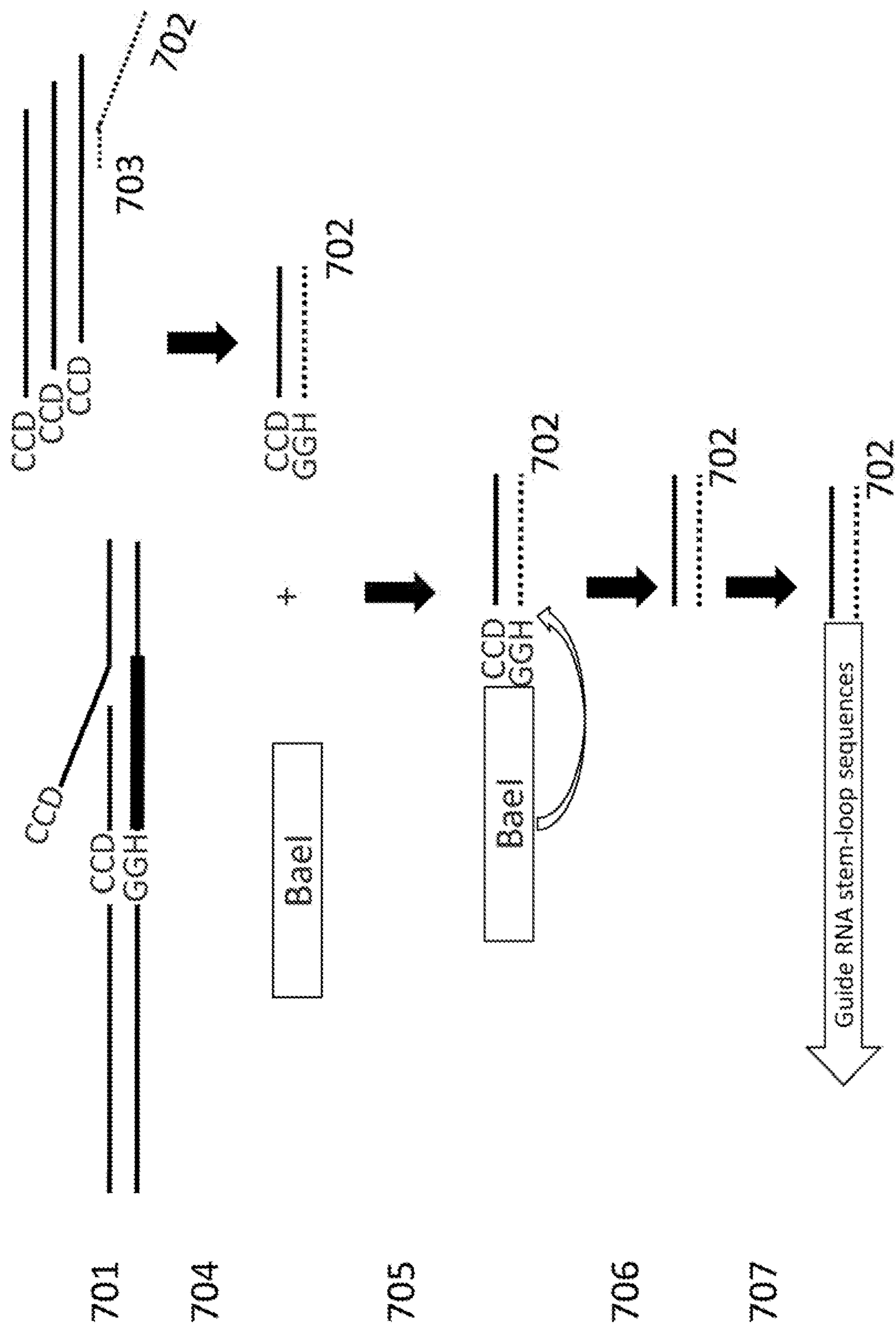


FIG. 6



7. G. E.

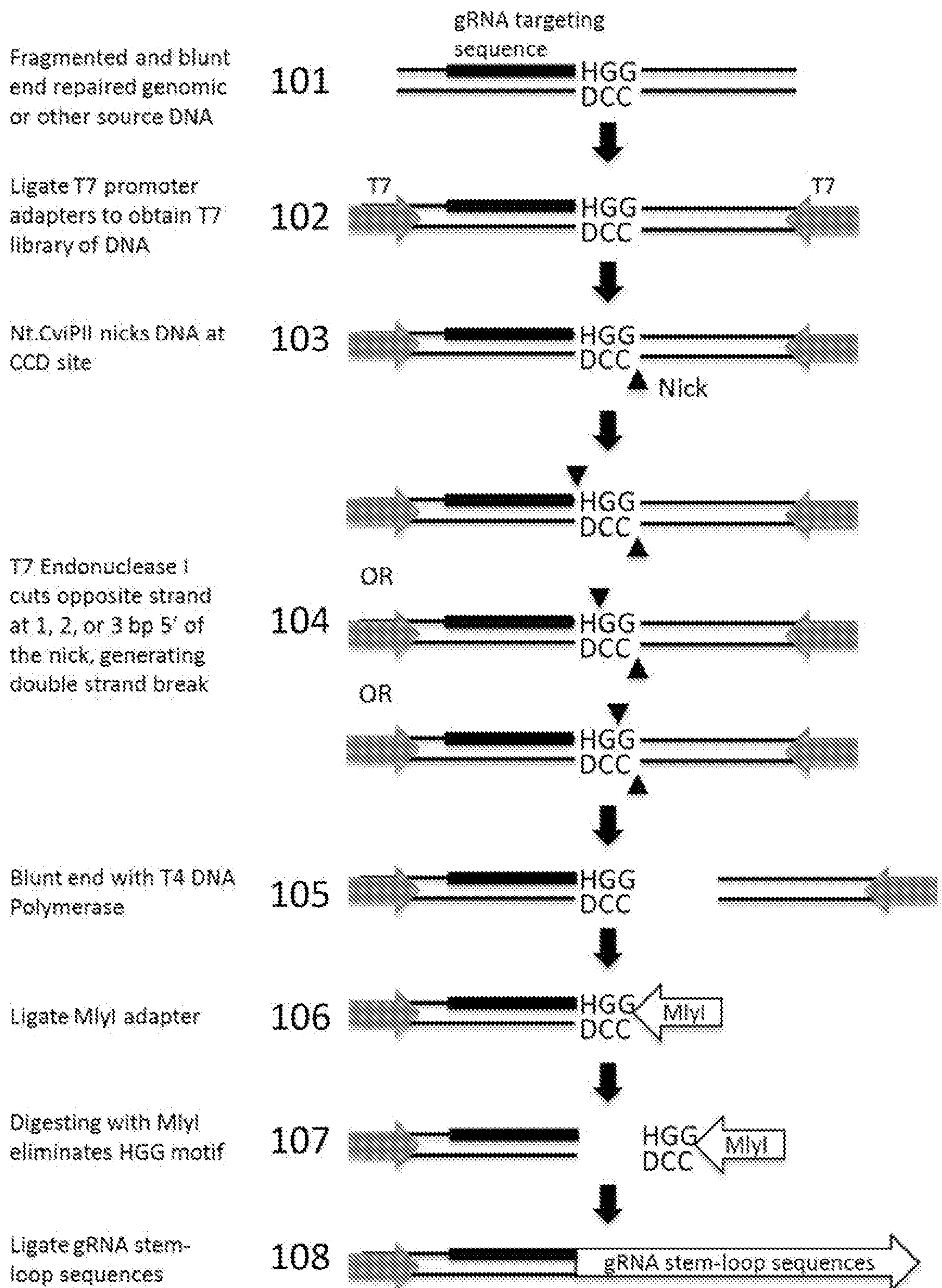


FIG. 1