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

(19) **World Intellectual Property Organization**
International Bureau

(10) **International Publication Number**
WO 2023/172966 A1

(43) **International Publication Date**
14 September 2023 (14.09.2023)

(51) **International Patent Classification:**
A61K 38/43 (2006.01) *A61K 38/46* (2006.01)
C12N 5/10 (2006.01) *A61K 38/48* (2006.01)
A61K 35/76 (2015.01) *A61K 48/00* (2006.01)
C12N 9/50 (2006.01)

(21) **International Application Number:** PCT/US2023/063954

(22) **International Filing Date:** 08 March 2023 (08.03.2023)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**
63/317,720 08 March 2022 (08.03.2022) US

(72) **Inventors; and**
(71) **Applicants: LU, Baisong** [US/US]; 169 Brooks Landing Drive, Winston-Salem, NC 27106 (US). **ATALA, Anthony** [US/US]; 345 N. Stratford Road, Winston-Salem, NC 27104 (US).

(74) **Agent: FERNANDEZ, Lizette, M.** et al.; Mailstop IP Docketing - 22, Kilpatrick Townsend & Stockton LLP, 1100 Peachtree St., Suite 2800, Atlanta, GA 30309 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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))*
— *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))*

(54) **Title:** COMPOSITIONS, SYSTEMS AND METHODS FOR EUKARYOTIC GENE EDITING

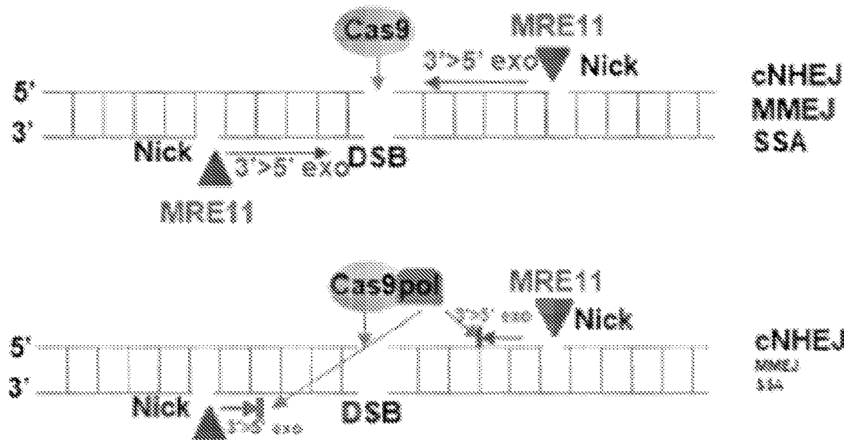


FIG. 1A

(57) **Abstract:** Provided herein are mammalian expression plasmids for delivering CRISPR component coding sequences, i.e., an sgRNA and a fusion protein comprising a CRISPR-associated endonuclease and a DNA polymerase domain, to a cell. Methods for using any of the mammalian expression plasmids described herein are also provided.

WO 2023/172966 A1

COMPOSITIONS, SYSTEMS AND METHODS FOR EUKARYOTIC GENE EDITING

PRIOR RELATED APPLICATION

[0001] This application claims the benefit of and priority to U.S. Provisional Application No. 63/317,720 filed on March 8, 2022, which is hereby incorporated by reference in its entirety.

SEQUENCE LISTING

[0002] The instant application contains a Sequence Listing in XML format. The Sequence Listing, named 095199-1375708.xml was created on March 7, 2023, is 175 Kilobytes in size, and is hereby incorporated by reference in its entirety.

FIELD

[0003] This disclosure describes compositions and methods of using same for eukaryotic gene editing.

BACKGROUND

[0001] CRISPR-based genome editing can unpredictably generate on-target deletions, for example, long DNA deletions (e.g., >500 bp) in the genome of a cell. The possibility of generating these DNA deletions poses safety risks to somatic genome editing, and makes the outcomes of genome editing less predictable. Therefore, compositions and methods for editing a cellular genome while reducing on-target deletions are necessary for safe and efficient genome editing.

SUMMARY

[0002] Provided herein is a mammalian expression plasmid comprising a eukaryotic promoter operably linked to a non-viral nucleic acid sequence, wherein the non-viral nucleic acid sequence comprises: (i) a nucleic acid sequence encoding a fusion protein, wherein the fusion protein comprises: (a) a polypeptide comprising a DNA polymerase domain; and (b) a CRISPR-associated endonuclease coding sequence; and (ii) a guide RNA (gRNA) coding sequence, wherein the gRNA coding sequence comprises at least one aptamer coding sequence.

[0003] In some embodiments, the polypeptide comprising a DNA polymerase domain comprises an *E. coli* DNA polymerase I (DNA Pol I) or a fragment thereof. In some embodiments, the polypeptide comprising a DNA polymerase domain comprises the Klenow fragment of *E. coli* DNA polymerase I (DNA Pol I).

[0004] In some embodiments, the CRISPR-associated endonuclease coding sequence encodes a Cas9 protein or a Cpf1 protein.

[0005] In some embodiments, the at least one aptamer coding sequence encodes an aptamer sequence bound specifically by an ABP selected from the group consisting of MS2 coat protein, PP7 coat protein, lambda N RNA-binding domain, or Com protein. In some embodiments, the aptamer is an MS2 aptamer sequence or a com aptamer sequence. In some embodiments, the sgRNA coding sequence comprises at least one aptamer inserted into the tetraloop or the ST2 loop of the sgRNA coding sequence. In some embodiments, the sgRNA coding comprises at least one com aptamer inserted into the ST2 loop of the gRNA coding sequence.

[0006] Also provided is a lentiviral packaging system comprising: a) a packaging plasmid comprising a eukaryotic promoter operably linked to a Gag nucleotide sequence, wherein the Gag nucleotide sequence comprises a nucleocapsid (NC) coding sequence and a matrix protein (MA) coding sequence, wherein one or both of the NC coding sequence or the MA coding sequence comprises at least one non-viral aptamer-binding protein (ABP) nucleotide sequence, and wherein the packaging plasmid does not encode a functional integrase protein; b) at least one mammalian expression plasmid provided herein; and c) an envelope plasmid comprising an envelope glycoprotein coding sequence.

[0007] In some embodiments, the packaging plasmid further comprises a Rev nucleotide sequence and a Tat nucleotide sequence. In some embodiments, the system further comprises a second packaging plasmid comprising a Rev nucleotide sequence. In some embodiments, the at least one non-viral ABP nucleotide sequence encodes MS2 coat protein, PP7 coat protein, lambda N peptide, or Com protein.

[0008] Further provided is a lentiviral particle comprising: (a) a fusion protein comprising a nucleocapsid (NC) protein or a matrix (MA) protein, wherein the NC protein or MA protein comprises at least one non-viral aptamer binding protein (ABP); and (b) a ribonucleotide protein (RNP) complex comprising: (i) a nucleic acid sequence encoding a fusion protein, wherein the fusion protein comprises: (a) a polypeptide comprising a DNA polymerase domain, and (b) a CRISPR-associated endonuclease coding sequence; and (ii) a guide RNA (gRNA)

coding sequence, wherein the gRNA coding sequence comprises at least one aptamer coding sequence wherein the lentivirus-like particle does not comprise a functional integrase protein.

[0009] In some embodiments, the CRISPR-associated endonuclease coding sequence encodes a Cas9 protein or a Cpf1 protein.

[0010] In some embodiments, the fusion polypeptide comprises an *E. coli* DNA Pol I. In some embodiments, the polypeptide comprises the Klenow fragment of DNA Pol I.

[0011] Also provided is a method of producing a lentiviral particle, the method comprising: a) transfecting a plurality of eukaryotic cells with the packaging plasmid, the at least one mammalian expression plasmid, and the envelope plasmid of any system described herein; and b) culturing the transfected eukaryotic cells for sufficient time for lentiviral particles to be produced.

[0012] In some embodiments, the lentiviral particle comprises a ribonucleotide protein (RNP) complex comprising: (i) nucleic acid sequence encoding a fusion protein, wherein the fusion protein comprises:(a) a polypeptide comprising a DNA polymerase domain and (b) a CRISPR-associated endonuclease coding sequence; and (ii) a guide RNA. In some embodiments, the plurality of eukaryotic cells are mammalian cells.

[0013] Also provided are cells comprising any of the plasmids, lentiviral packaging systems or lentivirus-like particles described herein. Cells modified by any of the methods provided herein are also provided.

[0014] Further provided is a method of modifying a genomic target sequence in a cell, the method comprising transducing a plurality of eukaryotic cells with a plurality of viral particles, wherein the plurality of viral particles comprise a lentivirus-like particle as provided herein comprising (a) a fusion protein comprising a NC protein or a MA protein wherein the NC protein or MA protein comprises at least one non-viral ABP; and (b) a RNP complex comprising: (i) a nucleic acid sequence encoding a fusion protein, wherein the fusion protein comprises: (a) a polypeptide comprising a DNA polymerase domain, and (b) a CRISPR-associated endonuclease coding sequence; and (ii) a gRNA coding sequence, wherein the gRNA coding sequence comprises at least one aptamer coding sequence wherein the lentivirus-like particle does not comprise a functional integrase protein, wherein the RNP complex binds to the genomic target sequence in genomic DNA of the cell, wherein the CRISPR-associated endonuclease cleaves the genomic target sequence to create a double-stranded break, thereby modifying the genomic target sequence

[0015] In some embodiments, NJEH is increased as compared to non-NHEJ end joining in the cell as compared to a cell modified by a CRISPR-associated endonuclease that is not fused to a DNA polymerase domain. In some embodiments, non-NJEH is microhomology-mediated end joining (MMEJ) and/or single-stranded annealing (SSA).

[0016] In some embodiments, the number of on-target deletions greater than 500 base pairs in size is decreased in the cell as compared to a cell. In some embodiments, the ratio of on-target one base pair (1-bp) deletions to on-target deletions greater than 1-bp is increased in the cell. In some embodiments, the ratio of on-target one base pair (1-bp) deletions to deletions greater than 500 base pairs is increased in the cell. In some embodiments, the number of templated insertions (TIS) increases in the cell. In some embodiments, the ratio of TIS to non-TIS increases in the cell.

[0017] In some embodiments, the plurality of eukaryotic cells are mammalian cells. In some embodiments, the plurality of eukaryotic cells are cells present in subject. In some embodiments, the subject is a human subject. In some embodiments, the subject is injected with the plurality of viral particles.

[0018] Also provided is a method for treating a disease in a subject comprising: a) obtaining cells from the subject; b) modifying the cells of the subject using the method provided herein using the lentivirus-like particles as provided herein; and c) administering the modified cells to the subject. In some embodiments, the disease is cancer. In some embodiments, the disease is Duchenne muscular dystrophy. In some embodiments, the cells are T cells.

DESCRIPTION OF THE FIGURES

[0019] The present application includes the following figures. The figures are intended to illustrate certain embodiments and/or features of the compositions and methods, and to supplement any description(s) of the compositions and methods. The figures do not limit the scope of the compositions and methods, unless the written description expressly indicates that such is the case.

[0020] FIG. 1A is a diagram showing a counteracting DNA resection (e.g., by MRE11) by DNA polymerase I (DNA Pol I or pol I) fused to Cas9. The expected result is the suppression of microhomology-mediated end joining (MMEJ) and single stranded annealing (SSA) DNA repair pathways that require DNA resection.

[0021] FIG. 1B shows that DNA polymerase 1 generates one base pair (1-bp) insertions (TIS) via filling in 5' overhangs. The red nucleotides are filled in by DNA polymerase. The target site of *CLCN5* sgRNA ((GAGGACAAGTCGTACAATGGTGG) (SEQ ID NO: 111)

and its complement (CTCCTGTTCAGCATGTTACCACC) (SEQ ID NO: 112)) was used as an example. The cleavage sites on both strands generate one nucleotide (1-nt) 5' overhangs, as indicated by small arrows.

[0022] FIG. 1C is a graph showing that fusion of DNA pol I to Cas9 increased the percentage of 1-bp deletions and decreased >1-bp deletions targeting *CLCN5* gene in HEK293T cells. Two-way ANOVA was followed by Bonferroni posttests. Replicate numbers are listed in parenthesis.

[0023] FIG. 1D is a graph showing that fusion of DNA pol I to Cas9 (Cas9-pol I) increased the ratio of 1-bp TIS versus 1-bp non-TIS (two tailed t-test).

[0024] FIG. 2A is an illustration showing Cas9-pol I and various mutant fusion proteins tested in the studies described in the Examples. Dashed lines indicate the deleted regions.

[0025] FIG. 2B is a graph showing the effects of different DNA pol I mutants on deletion profiles targeting *CLCN5* gene in HEK293T cells. Numbers in parentheses indicate biological replicate numbers. Groups above the red dashed lines showed statistical significance compared to those below the lines

[0026] FIG. 2C is a graph showing the effects of different DNA pol I mutants on insertions targeting *CLCN5* gene in HEK293T cells. Numbers in parentheses indicate biological replicate numbers. Groups above the red dashed lines showed statistical significance compared to those below the lines

[0027] FIG. 2D is a graph showing the effects of different DNA pol I mutants on deletion profiles targeting *CLCN5* gene in IMR90 cells. Replicate numbers were 3 for all groups. * and ** indicate $p<0.05$ and $p<0.01$, respectively, between the indicated group and all other groups (Bonferroni posttests following two-way ANOVA).

[0028] FIG. 2E is a graph showing the effects of different DNA pol I mutants on insertions targeting *CLCN5* gene in IMR90 cells. Replicate numbers were 3 for all groups. * and ** indicate $p<0.05$ and $p<0.01$, respectively, between the indicated group and all other groups (Bonferroni posttests following two-way ANOVA).

[0029] FIG. 3A is a graph showing the effects of *RBBP8* knockdown on Cas9-induced DNA mutation profiles targeting *CLCN5* in HEK293T cells. Replicate numbers were 6 for Cas9 and 3 for the remaining groups. *, ** and *** indicate $p<0.05$, $p<0.01$ and $p<0.001$ between the indicated groups (Bonferroni post-tests following two-way ANOVA).

[0030] FIG. 3B is a graph showing the effects of *RBBP8* knockdown on Cas9-induced insertions targeting *CLCN5* in HEK293T cells. Replicate numbers were 6 for Cas9 and 3 for

the remaining groups. *, ** and *** indicate $p<0.05$, $p<0.01$ and $p<0.001$ between the indicated groups (Bonferroni post-tests following two-way ANOVA).

[0031] FIG. 3C is a graph showing the effects of Effects of *RBBP8* knockdown on deletions with various sizes targeting *CLCN5* in HEK293T cells. The single and double arrows indicate the resection-dependent and independent deletions, respectively.

[0032] FIG. 3D shows the most frequently observed deletions generated by Cas9 targeting *CLCN5* in HEK293T cells. A partial wildtype *CLCN5* sequence (SEQ ID NO: 113) is shown, as well as SEQ ID NO: 113 comprising an 11 base pair deletion (SEQ ID NO: 114), SEQ ID NO: 113 comprising a different 11 base pair deletion (SEQ ID NO: 115), SEQ ID NO: 113 comprising an 8 base pair deletion (SEQ ID NO: 116), SEQ ID NO: 113 comprising a different 8 base pair deletion (SEQ ID NO: 117), and SEQ ID NO: 113 comprising a 16 base pair deletion (SEQ ID NO: 118). Dashed lines indicate deletions. Microhomology is underlined. Microhomology away from cleavage sites is indicated with a caret symbol (^). Asterisks indicate microhomology at the predicted cleavage size, which is indicated by a vertical dashed line.

[0033] FIG. 4A shows examination of suppressed deletions targeting *HBB* gene in hematopoietic cells. * and *** indicate $p<0.05$ and $p<0.001$, respectively, between Cas9 and Cas9-Klenow (Bonferroni posttests following two-way ANOVA). A partial wildtype HBB sequence (SEQ ID NO: 119) as well as SEQ ID NO: 119 comprising a 3 base pair deletion (SEQ ID NO: 120) and SEQ ID NO: 119 comprising a 12 base pair deletion (SEQ ID NO: 121) are shown. The top image shows the peaks of deletions and the bottom image show the most frequently observed deletions. Microhomology was underlined. Asterisks (*) indicate microhomology at the predicted cleavage size, which is indicated by a vertical dashed line. Carets (^) indicate microhomology away from the predicted cleavage site. Each group has three biological replicates.

[0034] FIG. 4B shows examination of suppressed deletions targeting *DMD* exon 53 in HEK293T cells. * and *** indicate $p<0.05$ and $p<0.001$, respectively, between Cas9 and Cas9-Klenow (Bonferroni posttests following two-way ANOVA). A partial wildtype DMD exon 53 sequence (SEQ ID NO: 122) as well as SEQ ID NO: 122 comprising an 11 base pair deletion (SEQ ID NO: 123), SEQ ID NO: 122 comprising a 9 base pair deletion (SEQ ID NO: 124), and SEQ ID NO: 122 with a 6 base pair deletion are shown (SEQ ID NO: 125). The top image shows the peaks of deletions and the bottom image shows the most frequently observed deletions. Microhomology is underlined. Asterisks (*) indicate microhomology at the predicted

cleavage size, which is indicated by a vertical dashed line. Carets (^) indicate microhomology away from the predicted cleavage site. Each group has three biological replicates.

[0035] FIG. 4C shows examination of suppressed deletions targeting *HBB* gene in IMR90 cells. * and *** indicate p<0.05 and p<0.001, respectively, between Cas9 and Cas9-Klenow (Bonferroni posttests following two-way ANOVA). A partial wildtype *HBB* sequence (SEQ ID NO: 126 is shown, as well as SEQ ID NO: 126 comprising a 3 base pair deletion (SEQ ID NO: 127) and SEQ ID NO: 126 comprising a 5 base pair deletion (SEQ ID NO: 128). The top image show the peaks of deletions and the bottom image shows the most frequently observed deletions. Microhomology is underlined. Asterisks (*) indicate microhomology at the predicted cleavage size, which is indicated by a vertical dashed line. Carets (^) indicate microhomology away from the predicted cleavage site. Each group has three biological replicates.

[0036] FIG. 4D shows examination of suppressed deletions targeting *DMD* exon 44 in human myoblasts. * and *** indicate p<0.05 and p<0.001, respectively, between Cas9 and Cas9-Klenow (Bonferroni posttests following two-way ANOVA). A partial wildtype *DMD44* sequence (SEQ ID NO: 129) is shown, as well as SEQ ID NO: 129 comprising a 10 base pair deletion (SEQ ID NO: 130), SEQ ID NO: 129 comprising a 7 base pair deletion (SEQ ID NO: 131), and SEQ ID NO: 129 comprising a different 7 base pair deletion (SEQ ID NO: 132). The top image show the peaks of deletions and the bottom image shows the most frequently observed deletions. Microhomology is underlined. Asterisks (*) indicate microhomology at the predicted cleavage size, which is indicated by a vertical dashed line. Carets (^) indicate microhomology away from the predicted cleavage site. Each group has three biological replicates.

[0037] FIG. 5A shows the large deletions generated by Cas9 and Cas9-Klenow targeting *CLCN5* gene in HEK293T cells. The data were combined from three replicates. The regions labeled 434 asnd 4534 indicate the two 25-bp sequence used for calculating distance for deletion detection. The regions labeled 2471, 2481, and 2487indicate the sgRNA target and the region labeled 2494 indicates the PAM. The asterisks indicate the identical deletions independently observed in two different experiments.

[0038] FIG. 5B shows the large deletions generated by Cas9 and Cas9-Klenow targeting *CLCN5* gene in IMR90 cells. The data were combined from three replicates. The regions labeled 434 and 4534 indicate the two 25-bp sequence used for calculating distance for deletion detection. The regions labeled 2471, 2489, and 2488 indicate the sgRNA target and the region

labeled 2494 indicates the PAM. The asterisks indicate the identical deletions independently observed in two different experiments.

[0039] FIG. 6 is a next generation sequencing (NGS) analysis of the integrated target sequence in GFP-reporter cells treated with *CLCN5* sgRNA and Cas9-pol. SEQ ID NOs: 133-141 are shown.

[0040] FIG. 7 is a graph showing that Cas9 and various exonuclease fusions had similar mutation profiles as Cas9. Numbers in parentheses indicate replicate numbers.

[0041] FIG. 8A is a graph showing the effects of different pol I domains on 2-bp TIS. The data for exo, 3' exo and 5' exo fusion proteins targeting *CLCN5* in HEK293T cells were pooled into one group. Each dot indicates one datum point. ** and *** indicate $p<0.05$ and $p<0.001$, respectively, compared with Cas9 group. Tukey's Multiple Comparison Test was performed following one-way ANOVA. Cas9-pol showed a trend of increase but did not reach statistical significance due to large intra group variation.

[0042] FIG. 8B is a graph showing the effects of different pol I domains on 3-bp TIS. The data for exo, 3' exo and 5' exo fusion proteins targeting *CLCN5* in HEK293T cells were pooled into one group. Each dot indicates one datum point. ** and *** indicate $p<0.05$ and $p<0.001$, respectively, compared with Cas9 group. Tukey's Multiple Comparison Test was performed following one-way ANOVA. Cas9-pol showed a trend of increase but did not reach statistical significance due to large intra group variation.

[0043] FIG. 9 is a graph showing that the overall INDEL rates did not have an effect on the percentages of different mutation types.

[0044] FIG. 10A is a graph showing the Effects of *MRE11* or *RBBP8* Knockdown on *CLCN5* mutation profiles (% of all INDELs) in IMR90 cells. Each group had 3 biological replicates. Two-way ANOVA was followed by Bonferroni posttests. ** and *** indicate $p<0.01$ and $p<0.001$, respectively, between the indicated groups.

[0045] FIG. 10B is a graph showing the Effects of *MRE11* or *RBBP8* Knockdown on *CLCN5* mutation profiles (% of all insertions) in IMR90 cells. Each group had 3 biological replicates. Two-way ANOVA was followed by Bonferroni posttests. ** and *** indicate $p<0.01$ and $p<0.001$, respectively, between the indicated groups.

[0046] FIG. 11A shows deletions that were decreased by Cas9-Klenow when targeting *HBB* in HEK293T cells. The graph (top image) shows the percentages of deletion sizes. The bottom image show the sequences of the most commonly observed deletions. A partial wildtype *HBB* sequence (SEQ ID NO: 142) is shown, as well as SEQ ID NO: 142 comprising

a 3 base pair deletion (SEQ ID NO: 143), SEQ ID NO: 142 comprising a 5 base pair deletion (SEQ ID NO: 144), and SEQ ID NO: 142 comprising a 10 base pair deletion (SEQ ID NO: 145). The regions underlined with green lines indicate microhomology at the predicted cleavage site which is indicated by vertical dashed lines. Two-way ANOVA was followed by Bonferroni posttests. * and *** indicates $p<0.05$ and $P<0.001$, respectively, between the two groups.

[0047] FIG. 11B shows deletions that were decreased by Cas9-Klenow when targeting Intergenic 1 in IMR90 cells. The graph (top image) show the percentages of deletion sizes. The bottom image show the sequences of the most commonly observed deletions. A partial wildtype Intergenic 1 sequence (SEQ ID NO: 146) is shown, as well as SEQ ID NO: 146 comprising a 4 base pair deletion (SEQ ID NO: 147). The sgRNA is shown shaded (in gray). PAM is indicated by overline.

[0048] FIG. 11C shows deletions that were decreased by Cas9-Klenow when targeting *CLCN5* in IMR90 cells. The graph (top image) shows the percentages of deletion sizes. The bottom image show the sequences of the most commonly observed deletions. No sequence could be listed for *CLCN5*/IMR90 since no major deletion peaks were observed. The sgRNA is shown shaded (in gray). PAM is indicated by overline.

Definitions

[0049] As used in this specification and the appended claims, the singular forms “a,” “an,” and “the” include plural reference unless the context clearly dictates otherwise.

[0050] The use of any and all examples or exemplary language (e.g., “such as”) provided herein, is intended merely to better illustrate the invention and does not pose a limitation on the scope of the invention unless otherwise claimed.

[0051] The terms “may,” “may be,” “can,” and “can be,” and related terms are intended to convey that the subject matter involved is optional (that is, the subject matter is present in some examples and is not present in other examples), not a reference to a capability of the subject matter or to a probability, unless the context clearly indicates otherwise.

[0052] The terms “optional” and “optionally” mean that the subsequently described event, circumstance, or material may or may not occur or be present, and that the description includes instances where the event, circumstance, or material occurs or is present as well as instances where it does not occur or is not present.

[0053] The use herein of the terms "including," "comprising," or "having," and variations thereof, is meant to encompass the elements listed thereafter and equivalents thereof as well as

additional elements. Embodiments recited as "including," "comprising," or "having" certain elements are also contemplated as "consisting essentially of and "consisting of those certain elements. As used herein, "and/or" refers to and encompasses any and all possible combinations of one or more of the associated listed items, as well as the lack of combinations where interpreted in the alternative ("or").

[0054] As used herein, the transitional phrase "consisting essentially of" (and grammatical variants) is to be interpreted as encompassing the recited materials or steps "and those that do not materially affect the basic and novel characteristic(s)" of the claimed invention. See In re Herz, 537 F.2d 549, 551-52, 190 U.S.P.Q. 461, 463 (CCPA 1976) (emphasis in the original); see also MPEP §2111.03. Thus, the term "consisting essentially of" as used herein should not be interpreted as equivalent to "comprising."

[0055] The term "nucleic acid" or "nucleotide" refers to deoxyribonucleic acids (DNA) or ribonucleic acids (RNA) (e.g., mRNA) and polymers thereof in either single- or double-stranded form. It is understood that when an RNA is described, its corresponding DNA is also described, wherein uridine is represented as thymidine. Similarly, when a DNA is described, its corresponding RNA is also described wherein thymidine is represented by uridine. Unless specifically limited, the term encompasses nucleic acids containing known analogues of natural nucleotides that have similar binding properties as the reference nucleic acid and are metabolized in a manner similar to naturally occurring nucleotides. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions), alleles, orthologs, SNPs, and complementary sequences as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer et al., Nucleic Acid Res. 19:5081 (1991); Ohtsuka et al., J. Biol. Chem. 260:2605-2608 (1985); and Rossolini et al., Mol. Cell. Probes 8:91-98 (1994)). The polynucleotides of the invention also encompass all forms of sequences including, but not limited to, single-stranded forms, double-stranded forms, hairpins, stem-and-loop structures, and the like.

[0056] The term "gene" can refer to the segment of DNA involved in producing or encoding a polypeptide chain. It may include regions preceding and following the coding region (leader and trailer) as well as intervening sequences (introns) between individual coding segments (exons). Alternatively, the term "gene" can refer to the segment of DNA involved in

producing or encoding a non-translated RNA, such as an rRNA, tRNA, guide RNA, or micro RNA.

[0057] "Treating" refers to any indicia of success in the treatment or amelioration or prevention of the disease, condition, or disorder, including any objective or subjective parameter such as abatement; remission; diminishing of symptoms or making the disease condition more tolerable to the patient; slowing in the rate of degeneration or decline; or making the final point of degeneration less debilitating. The treatment or amelioration of symptoms can be based on objective or subjective parameters; including the results of an examination by a physician. Accordingly, the term "treating" includes the administration of the compounds or agents of the present disclosure to prevent or delay, to alleviate, or to arrest or inhibit development of the symptoms or conditions associated with a disease, condition or disorder as described herein. The term "therapeutic effect" refers to the reduction, elimination, or prevention of the disease, symptoms of the disease, or side effects of the disease in the subject. "Treating" or "treatment" using the methods of the present disclosure includes preventing the onset of symptoms in a subject that can be at increased risk of a disease or disorder associated with a disease, condition or disorder as described herein, but does not yet experience or exhibit symptoms, inhibiting the symptoms of a disease or disorder (slowing or arresting its development), providing relief from the symptoms or side effects of a disease (including palliative treatment), and relieving the symptoms of a disease (causing regression). Treatment can be prophylactic (to prevent or delay the onset of the disease, or to prevent the manifestation of clinical or subclinical symptoms thereof) or therapeutic suppression or alleviation of symptoms after the manifestation of the disease or condition. The term "treatment," as used herein, includes preventative (e.g., prophylactic), curative, or palliative treatment.

[0058] A "promoter" is defined as one or more a nucleic acid control sequences that direct transcription of a nucleic acid. As used herein, a promoter includes necessary nucleic acid sequences near the start site of transcription, such as, in the case of a polymerase II type promoter, a TATA element. A promoter also optionally includes distal enhancer or repressor elements, which can be located as much as several thousand base pairs from the start site of transcription.

[0059] "Polypeptide," "peptide," and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. All three terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally

occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymers. As used herein, the terms encompass full-length proteins, truncated proteins, and fragments thereof, and amino acid chains, wherein the amino acid residues are linked by covalent peptide bonds. As used throughout, the term "fusion polypeptide" or "fusion protein" is a polypeptide comprising two or more proteins or fragments thereof. In some embodiments, a linker comprising about 3 to 10 amino acids can be positioned between any two proteins or fragments thereof to help facilitate proper folding of the proteins upon expression.

[0060] The term "identity" or "substantial identity", as used in the context of a polynucleotide or polypeptide sequence described herein, refers to a sequence that has at least 60% sequence identity to a reference sequence. Alternatively, percent identity can be any integer from 60% to 100%. Exemplary embodiments include at least: 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%, as compared to a reference sequence using the programs described herein; preferably BLAST using standard parameters, as described below. It is understood that sequences having at 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to any nucleotide or polypeptide sequence set forth herein, for example, any one of SEQ ID NOs: 1-110, can be used in the compositions and methods provided herein. It is understood that a nucleic acid sequence can comprise, consist of, or consist essentially of any nucleic acid sequence described herein. Similarly, a polypeptide can comprise, consist of, or consist essentially of, any polypeptide sequence described herein. For sequence comparison, typically one sequence acts as a reference sequence to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

[0061] A "comparison window", as used herein, includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 20 to 600, about 20 to 50, about 20 to 100, about 50 to about 200 or about 100 to about 150, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for

comparison are well-known in the art. Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman Add. APL. Math. 2:482 (1981), by the homology alignment algorithm of Needleman and Wunsch J. Mol. Biol. 48:443 (1970), by the search for similarity method of Pearson and Lipman Proc. Natl. Acad. Sci. (U.S.A.) 85: 2444 (1988), by computerized implementations of these algorithms (e.g., BLAST), or by manual alignment and visual inspection.

[0062] Algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al. (1990) J. Mol. Biol. 215: 403-410 and Altschul et al. (1977) Nucleic Acids Res. 25: 3389-3402, respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (NCBI) web site. The algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al, supra). These initial neighborhood word hits acts as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a word size (W) of 28, an expectation (E) of 10, M=1, N=-2, and a comparison of both strands.

[0063] The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin & Altschul, Proc. Nat'l. Acad. Sci. USA 90:5873-5787 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of

the test nucleic acid to the reference nucleic acid is less than about 0.01, more preferably less than about 10-5, and most preferably less than about 10-20.

[0064] As used throughout, by subject is meant an individual. For example, the subject is a mammal, such as a primate, and, more specifically, a human. Non-human primates are subjects as well. The term subject includes domesticated animals, such as cats, dogs, etc., livestock (for example, cattle, horses, pigs, sheep, goats, etc.) and laboratory animals (for example, ferret, chinchilla, mouse, rabbit, rat, gerbil, guinea pig, etc.). Thus, veterinary uses and medical uses and formulations are contemplated herein. The term does not denote a particular age or sex. Thus, adult and newborn subjects, whether male or female, are intended to be covered. As used herein, patient or subject may be used interchangeably and can refer to a subject afflicted with a disease or disorder.

[0065] An "expression cassette" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements that permit transcription of a particular polynucleotide sequence in a host cell. An expression cassette may be part of a plasmid, viral genome, or nucleic acid fragment. Typically, an expression cassette includes a polynucleotide to be transcribed, operably linked to a promoter, followed by a transcription termination signal sequence. An expression cassette may or may not include specific regulatory sequences, such as 5' or 3' untranslated regions from human globin genes.

[0066] A "reporter gene" encodes proteins that are readily detectable due to their biochemical characteristics, such as enzymatic activity or chemifluorescent features. These reporter proteins can be used as selectable markers. One specific example of such a reporter is green fluorescent protein. Fluorescence generated from this protein can be detected with various commercially-available fluorescent detection systems. Other reporters can be detected by staining. The reporter can also be an enzyme that generates a detectable signal when contacted with an appropriate substrate. The reporter can be an enzyme that catalyzes the formation of a detectable product. Suitable enzymes include, but are not limited to, proteases, nucleases, lipases, phosphatases and hydrolases. The reporter can encode an enzyme whose substrates are substantially impermeable to eukaryotic plasma membranes, thus making it possible to tightly control signal formation. Specific examples of suitable reporter genes that encode enzymes include, but are not limited to, CAT (chloramphenicol acetyl transferase; Alton and Vapnek (1979) Nature 282: 864-869); luciferase (lux); β-galactosidase; LacZ; β.-glucuronidase; and alkaline phosphatase (Toh, et al. (1980) Eur. J. Biochem. 182: 231-238; and Hall et al. (1983) J. Mol. Appl. Gen. 2: 101), each of which are incorporated by reference

herein in its entirety. Other suitable reporters include those that encode for a particular epitope that can be detected with a labeled antibody that specifically recognizes the epitope.

[0067] The "CRISPR/Cas" system refers to a widespread class of bacterial systems for defense against foreign nucleic acid. CRISPR/Cas systems are found in a wide range of eubacterial and archaeal organisms. CRISPR/Cas systems include type I, II, and III sub-types. The CRISPR/Cas system classification as described in by Makarova, et al. (*Nat Rev Microbiol.* 2015 Nov; 13(11):722-36) defines five types and 16 subtypes based on shared characteristics and evolutionary similarity. These are grouped into two large classes based on the structure of the effector complex that cleaves genomic DNA. The Type II CRISPR/Cas system was the first used for genome engineering, with Type V following in 2015. Wild-type type II CRISPR/Cas systems utilize an RNA-mediated nuclease Cas protein or homolog (referred to herein as a "CRISPR-associated endonuclease") in complex with guide RNA to recognize and cleave foreign nucleic acid. Cas9 proteins also use an activating RNA (also referred to as a transactivating or tracr RNA). Guide RNAs having the activity of either a guide RNA or both a guide RNA and an activating RNA, depending on the type of CRISPR-associated endonuclease used therewith, are also known in the art. In some cases, such dual activity guide RNAs are referred to as a single guide RNA (sgRNA). Synthetic guide RNAs that do not contain an activating RNA sequence may also be referred to as sgRNAs. In this disclosure, the terms sgRNA and gRNA are used interchangeably to refer to an RNA molecule that complexes with a CRISPR-associated endonuclease and localizes the CRISPR-associated endonuclease, for example, a CRISPR-associated endonuclease in a ribonucleoprotein complex, to a target DNA sequence.

[0068] As used herein, "activity" in the context of sgRNA activity, or RNP activity, i.e., RNP activity of a complex comprising: (1) a gRNA and (2) a fusion protein comprising a CRISPR-associated endonuclease and a DNA polymerase domain, refers to the ability of a sgRNA to bind to a target genetic element. Typically, activity also refers to the ability of an RNP (i.e., an sgRNA complexed with a fusion protein comprising a CRISPR-associated endonuclease and a DNA polymerase domain) to edit the genome of a cell. As used herein, the phrase "editing" in the context of editing of a genome of a cell refers to inducing a structural change in the sequence of the genome at a target genomic region, for example, cleaving a genomic sequence and inserting a donor sequence into the genome of a cell, at the cleavage site, via homology directed repair (HDR), or cleaving a sequence and allowing repair via non-homologous end joining (NHEJ).

[0069] As used herein, the term "ribonucleoprotein complex" "RNPs", and the like refers to a complex between: (1) a fusion protein comprising a CRISPR-associated endonuclease (e.g., Cas9) and a DNA polymerase domain, and a crRNA (e.g., guide RNA or single guide RNA), (2) a fusion protein comprising a CRISPR-associated endonuclease and a DNA polymerase domain; and a trans-activating crRNA (tracrRNA), (3) a fusion protein comprising a CRISPR-associated endonuclease and a DNA polymerase domain, and a guide RNA, or (4) a combination thereof (e.g., a complex containing the fusion protein comprising a CRISPR-associated endonuclease and a DNA polymerase domain, a tracrRNA, and a crRNA guide).

[0070] As used herein, a "cell" can be any eukaryotic cell, for example, human T cell or a cell capable of differentiating into a T cell, for example, a T cell that expresses a TCR receptor molecule. These include hematopoietic stem cells and cells derived from hematopoietic stem cells. Populations of cells, for example, populations of cells comprising viral particles or genetically modified cells made by any of the genomic editing methods provided herein, are also provided.

[0071] As used herein, the phrase "hematopoietic stem cell" refers to a type of stem cell that can give rise to a blood cell. Hematopoietic stem cells can give rise to cells of the myeloid or lymphoid lineages, or a combination thereof. Hematopoietic stem cells are predominantly found in the bone marrow, although they can be isolated from peripheral blood, or a fraction thereof. Various cell surface markers can be used to identify, sort, or purify hematopoietic stem cells. In some cases, hematopoietic stem cells are identified as c-kit+ and lin-. In some cases, human hematopoietic stem cells are identified as CD34+, CD59+, Thy1/CD90+, CD38lo/-, C-kit/CD117+, lin-. In some cases, human hematopoietic stem cells are identified as CD34-, CD59+, Thy1/CD90+, CD38lo/-, C-kit/CD117+, lin-. In some cases, human hematopoietic stem cells are identified as CD133+, CD59+, Thy1/CD90+, CD38lo/-, C-kit/CD117+, lin-. In some cases, mouse hematopoietic stem cells are identified as CD34lo/-, SCA-1+, Thy1+/lo, CD38+, C-kit +, lin-. In some cases, the hematopoietic stem cells are CD150+CD48-CD244-.

[0072] As used herein, the phrase "hematopoietic cell" refers to a cell derived from a hematopoietic stem cell. The hematopoietic cell may be obtained or provided by isolation from an organism, system, organ, or tissue (e.g., blood, or a fraction thereof). Alternatively, an hematopoietic stem cell can be isolated and the hematopoietic cell obtained or provided by differentiating the stem cell. Hematopoietic cells include cells with limited potential to differentiate into further cell types. Such hematopoietic cells include, but are not limited to, multipotent progenitor cells, lineage-restricted progenitor cells, common myeloid progenitor

cells, granulocyte-macrophage progenitor cells, or megakaryocyte-erythroid progenitor cells. Hematopoietic cells include cells of the lymphoid and myeloid lineages, such as lymphocytes, erythrocytes, granulocytes, monocytes, and thrombocytes. In some embodiments, the hematopoietic cell is an immune cell, such as a T cell, B cell, macrophage, a natural killer (NK) cell or dendritic cell. In some embodiments the cell is an innate immune cell.

[0073] As used herein, the phrase "T cell" refers to a lymphoid cell that expresses a T cell receptor molecule. T cells include human alpha beta (αβ) T cells and human gamma delta (γδ) T cells. T cells include, but are not limited to, naïve T cells, stimulated T cells, primary T cells (e.g., uncultured), cultured T cells, immortalized T cells, helper T cells, cytotoxic T cells, memory T cells, regulatory T cells, natural killer T cells, combinations thereof, or sub-populations thereof. T cells can be CD4+, CD8+, or CD4+ and CD8+. T cells can also be CD4-, CD8-, or CD4- and CD8-. T cells can be helper cells, for example helper cells of type TH1, TH2, TH3, TH9, TH17, or TFH. T cells can be cytotoxic T cells. Regulatory T cells can be FOXP3+ or FOXP3-. T cells can be alpha/beta T cells or gamma/delta T cells. In some cases, the T cell is a CD4+CD25hiCD127lo regulatory T cell. In some cases, the T cell is a regulatory T cell selected from the group consisting of type 1 regulatory (Tr1), TH3, CD8+CD28-, Treg17, and Qa-1 restricted T cells, or a combination or sub-population thereof. In some cases, the T cell is a FOXP3+ T cell. In some cases, the T cell is a CD4+CD25loCD127hi effector T cell. In some cases, the T cell is a CD4+CD25loCD127hiCD45RAhiCD45RO- naïve T cell. A T cell can be a recombinant T cell that has been genetically manipulated.

[0074] As used herein, the phrase "primary" in the context of a primary cell is a cell that has not been transformed or immortalized. Such primary cells can be cultured, sub-cultured, or passaged a limited number of times (e.g., cultured 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 times). In some cases, the primary cells are adapted to in vitro culture conditions. In some cases, the primary cells are isolated from an organism, system, organ, or tissue, optionally sorted, and utilized directly without culturing or sub-culturing. In some cases, the primary cells are stimulated, activated, or differentiated. For example, primary T cells can be activated by contact with (e.g., culturing in the presence of) CD3, CD28 agonists, IL-2, IFN-γ, or a combination thereof.

DETAILED DESCRIPTION

[0075] The following description recites various aspects and embodiments of the present compositions and methods. No particular embodiment is intended to define the scope of the compositions and methods. Rather, the embodiments merely provide non-limiting examples of

various compositions and methods that are at least included within the scope of the disclosed compositions and methods. The description is to be read from the perspective of one of ordinary skill in the art; therefore, information well known to the skilled artisan is not necessarily included.

[0076] The CRISPR/Cas9 (Clustered Regularly Interspaced Short Palindromic Repeats and CRISPR-associated protein 9) system uses a single effector protein to make DNA double-strand breaks (DSBs) guided by single guide RNA (sgRNA), and has been used to make specific genetic changes in human cells. *Streptococcus pyogenes* Cas9 (SpCas9) mainly generates blunt ends via cleaving the both DNA strands 3 nucleotides (nt) upstream of the "NGG" Protospacer Adjacent Motif (PAM). It can also generate staggered ends with 1, 2 or 3 nt 5' overhangs via cleaving the targeting strand 3 nt, and the non-targeting strand 4, 5, or 6 nt upstream of the PAM.

[0077] Human cells have multiple pathways to repair the double stranded breaks (DSBs) created by CRISPR/Cas9 nucleases: homology-directed repair (HDR), canonic non-homologous end joining (cNHEJ or NHEJ), and alternative end joining pathways including microhomology-mediated end joining (MMEJ) and single strand annealing (SSA). In the cNHEJ pathway, the two ends of the DSBs are processed for re-ligation without the involvement of resection or template. Pol X family members (Pol λ, μ, and β and terminal transferase) may function to fill in 5' overhangs. This fill-in reaction can explain the widely observed predictable insertions or templated insertions (TISs) induced by SpCas9.

[0078] HDR, MMEJ and SSA repair pathways all depend on DNA resection to generate 3' overhangs. DNA resection is initiated by the MRE11-RAD50-NBS1 complex and stimulated by CtIP (encoded by *RBBP8*). MRE11's endonuclease activity generates a nick 3' to the DSB, and the 3'-5' exonucleolytic activity generates a 3'-single stranded DNA overhang from the nick. HDR uses long 3' overhangs and the DNA template to faithfully repair the DSBs. MMEJ and SSA use microhomology (2-5 nt) and short homology (10-15 nt), respectively, in the two 3' overhangs to facilitate DNA synthesis, generating DNA deletions with sizes depending on the distances between the homologous regions. Long 3' overhangs from excessive DNA resection can also be filled in by DNA polα and associated complexes. This fill-in reaction explains the observed small tandem duplicates at DSBs with 3' overhangs generated by Cas9 nickases.

[0079] Although CRISPR/Cas9-induced mutation profiles are predictable to a certain degree, and small insertions and deletions (INDELs) are the dominant mutation types,

unpredictable on-target large deletions are widely reported. Currently, methods for suppressing the generation of large deletions are lacking. Although genome editing methods without generating DSBs, such as base editing and prime editing, have been developed, many applications including eradicating integrated provirus DNA from human cells and removing disease-causing DNA repeats still involve DSB generation. Thus, a method that can generate refined mutation profiles will improve safety.

[0080] Provided herein are compositions, systems, methods of manufacture, and methods for genome editing. Using the compositions and methods described herein, the genome of a cell can be efficiently edited using a CRISPR-associated endonuclease, while reducing the number of long deletions generated in the genome of the cell, for example, by increasing non-homologous end joining (NHEJ) in the cell as compared to non-NHEJ (for example, microhomology-mediated end joining (MMEJ) or single strand annealing (SSA)) in the cell. Exemplary components, systems, methods of manufacture, and methods for editing the genome of a cell using an RNP comprising (1) a fusion protein comprising a CRISPR-associated endonuclease and a polypeptide comprising a DNA polymerase domain; and (2) an sgRNA are provided herein.

Mammalian Expression Plasmids

[0081] Provided herein are mammalian expression plasmids that are used to deliver CRISPR component coding sequences, i.e., an sgRNA and a fusion protein comprising a CRISPR-associated endonuclease and a DNA polymerase domain, into mammalian cells being used to generate the lentivirus-like particles of this disclosure. For example, provided herein is a mammalian expression plasmid comprising a eukaryotic promoter operably linked to a non-viral nucleic acid sequence, wherein the non-viral nucleic acid sequence comprises: (i) a nucleic acid sequence encoding a fusion protein, wherein the fusion protein comprises: (a) a polypeptide comprising a DNA polymerase domain; and (b) a CRISPR-associated endonuclease coding sequence; and (ii) a guide RNA (gRNA) coding sequence, wherein the gRNA coding sequence comprises at least one aptamer coding sequence.

[0082] In the mammalian expression plasmids provided herein, the polypeptide comprising a DNA polymerase domain can be fused or linked to a CRISPR-associate endonuclease. Optionally, the CRISPR-associated endonuclease is linked to the polypeptide comprising a DNA polymerase domain via a peptide linker. The linker can be between about 2 and about 25 amino acids in length.

[0083] In some embodiments, the CRISPR-associated endonuclease coding sequence encodes a Cas9 protein. Full-length Cas9 is an endonuclease comprising a recognition domain and two nuclease domains (HNH and RuvC, respectively) that creates double-stranded breaks in DNA sequences. Cas9 is targeted to a genomic site in a cell by interacting with a guide RNA that hybridizes to a 20-nucleotide DNA sequence that immediately precedes an NGG motif recognized by Cas9. This results in a double-strand break in the genomic DNA of the cell. In some examples, a Cas9 nuclease that requires an NGG protospacer adjacent motif (PAM) immediately 3' of the region targeted by the guide RNA can be utilized. As another example, Cas9 proteins with orthogonal PAM motif requirements can be utilized to target sequences that do not have an adjacent NGG PAM sequence. Exemplary Cas9 proteins with orthogonal PAM sequence specificities include, but are not limited to those described in Esvelt *et al.*, Nature Methods 10: 1116–1121 (2013).

[0084] Various Cas9 nucleases can be utilized in the methods described herein. For example, a Cas9 nuclease that requires an NGG protospacer adjacent motif (PAM) immediately 3' of the region targeted by the guide RNA, such as SpCas9, can be utilized. Such Cas9 nucleases can be targeted to any region of a genome that contains an NGG sequence. In another example, a Cas9 nuclease that requires an NNGRRT or NNGRR(N) PAM immediately 3' of the region targeted by the guide RNA, such as SaCas9, can be utilized. As another example, Cas9 proteins with orthogonal PAM motif requirements can be utilized to target sequences that do not have an adjacent NGG PAM sequence. Exemplary Cas9 proteins with orthogonal PAM sequence specificities include, but are not limited to those described in Nature Methods 10, 1116–1121 (2013), and those described in Zetsche *et al.*, Cell, Volume 163, Issue 3, p759–771, 22 October 2015. An exemplary amino acid sequence for a Cas9 protein is set forth herein as SEQ ID NO: 29.

[0085] In some cases, the Cas9 protein is a nickase, such that when bound to target nucleic acid as part of a complex with a guide RNA, a single strand break or nick is introduced into the target nucleic acid. A pair of Cas9 nickases, each bound to a structurally different guide RNA, can be targeted to two proximal sites of a target genomic region and thus introduce a pair of proximal single stranded breaks into the target genomic region. Nickase pairs can provide enhanced specificity because off-target effects are likely to result in single nicks, which are generally repaired without lesion by base-excision repair mechanisms. Exemplary Cas9 nickases include Cas9 nucleases having a D10A or H840A mutation.

[0086] In some embodiments, the CRISPR-associated endonuclease is a Cpf1 polypeptide. Cpf1 protein is a Class II, Type V CRISPR/Cas system protein. Cpf1 is a smaller and simpler endonuclease than Cas9 (such as the spCas9). The Cpf1 protein has a RuvC-like endonuclease domain that is similar to the RuvC domain of Cas9 but does not have a HNH endonuclease domain. The N-terminal domain of Cpf1 also does not have the alpha-helical recognition lobe like the Cas9 protein. When cleaving DNA, Cpf1 introduces a sticky-end-like DNA double-stranded break with a 4 or 5 nucleotide overhang. The Cpf1 protein does not need a tracrRNA; rather, the Cpf1 protein functions with only a crRNA. In the context of this disclosure, where the CRISPR-associated endonuclease is a Cpf1 protein, the sgRNA does not comprise a tracr sequence. The sgRNA used with the Cpf1 protein may comprise only a crRNA sequence (constant region). In some examples, a Cpf1 protein that requires an TTTN or TTN PAM (depending on the species, where "N" is an nucleobase) immediately 5' of the region targeted by the guide RNA can be utilized. Known Cpf1 proteins and derivatives thereof may be used in the context of this disclosure. For example, in some instances, the CRISPR-associated endonuclease is FnCpf1p and the PAM is 5′ TTN, where N is A/C/G or T. In some instances, the CRISPR-associated endonuclease is PaCpf1p and the PAM is 5′ TTTV, where V is A/C or G In certain instances, the CRISPR-associated endonuclease is FnCpf1p and the PAM is 5′ TTN, where N is A/C/G or T, and the PAM is located upstream of the 5′ end of the protospacer. In certain instances, the CRISPR-associated endonuclease is FnCpf1p and the PAM is 5′ CTA and is located upstream of the 5′ end of the protospacer or the target locus. In one example, the CRISPR-associated endonuclease is AsCpf1p and the PAM is 5′ TTTN. An exemplary amino acid sequence for Cpf1 is set forth herein as SEQ ID NO: 30.

[0087] As used throughout, a DNA polymerase domain is a fragment of a full-length DNA polymerase that catalyzes the 5'-3' polymerization of nucleotides into duplex DNA, in the presence of a nucleic acid template. Any DNA polymerase domain or polypeptide comprising a DNA polymerase domain can be used to make the fusion proteins described herein. By fusing the DNA polymerase domain to the CRISPR-associated endonuclease, the DNA polymerase domain is targeted to the site of the double-stranded break made by the CRISPR-associated endonuclease in the genome of the cell. In some embodiments the DNA polymerase domain is catalytically active. In some embodiments, the catalytic activity of the DNA polymerase domain is reduced (e.g., relative to a corresponding wild-type DNA polymerase domain), for example, but at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%,

80%, 90% or 100%. In some embodiments, the DNA polymerase domain comprises one or more mutations, for example, a D705A mutation, that abolishes catalytic activity.

[0088] In some cases, the polypeptide comprising the DNA polymerase domain has polymerase activity and 3'-5' exonuclease activity. In some cases, the polypeptide comprising the DNA polymerase domain has 3'-5' exonuclease activity, but has reduced or lacks polymerase activity (e.g., relative to a corresponding wild-type DNA polymerase domain).

[0089] An exemplary polypeptide comprising a DNA polymerase domain is DNA polymerase I (DNA Pol I). DNA Pol I is an enzyme that has 5'-3' DNA dependent DNA polymerase activity, 3'-5' exonuclease activity, 5'-3' exonuclease activity and 5'-3' RNA-dependent DNA polymerase activity. An exemplary sequence for DNA Pol I is provided herein as SEQ ID NO: 27. In some embodiments, a DNA pol I fragment comprising a DNA polymerase domain is fused to the CRISPR-associated endonuclease. In some cases, the fragment comprises an amino acid sequence having 5'-3'DNA polymerase activity and an amino acid sequence having 3'-5' exonuclease activity (e.g., the Klenow fragment of DNA Pol I). An exemplary sequence for the Klenow fragment of DNA Pol I is provided herein as SEQ ID NO: 28. Other exemplary DNA polymerases include, but are not limited to *Thermus aquaticus* DNA Pol I and *Bacillus stearothermophilus* DNA pol I.

[0090] The mammalian expression plasmids provided herein comprise CRISPR component coding sequences, e.g., the coding sequence for a fusion protein comprising a CRISPR-associated endonuclease and a polypeptide comprising a DNA polymerase domain; and a gRNA. In some instances, the gRNA coding sequence comprises at least one aptamer coding sequence. In some instances, the at least one aptamer coding sequence may be positioned at the 5' end or the 3' end of the gRNA. In some instances, the at least one aptamer coding sequence may be inserted at an internal position within the gRNA such as, for example, at one or more of the loops formed in the folded gRNA. For example, where the gRNA is for the Cas9 protein, the at least one aptamer coding sequence may be positioned at the tetra loop, the stem loop 2 (ST2), or the 3' end of the gRNA. In some instances, a spacer of 1-30 nucleotides may be positioned between the gRNA the at least one aptamer coding sequence, or flanking the at least one aptamer coding sequence.

[0091] In some instances, the mammalian expression vector comprises at least one aptamer coding sequence that encodes an aptamer sequence that is bound specifically by an aptamer-binding protein (ABP). In the context of this disclosure, an aptamer sequence is an RNA

sequence that forms a tertiary loop structure that is specifically bound by an ABP. ABPs are RNA-binding proteins or RNA-binding protein domains. Suitable aptamer coding sequences include polynucleotide sequences that encode known bacteriophage aptamer sequences. Exemplary aptamer coding sequences include those encoding the aptamer sequences provided above in Table 1. In some instances, the aptamers are bound by a dimer of ABP. These aptamer sequences are RNA sequences known to be bound specifically by bacteriophage proteins. In some circumstances, the at least one aptamer coding sequence encodes an aptamer sequence bound specifically by an ABP selected from the group consisting of MS2 coat protein, PP7 coat protein, lambda N RNA-binding domain, or Com protein.

Table 1. Aptamer-Binding Proteins and Corresponding Aptamer Sequences

	Aptamer-Binding Proteins			
	MS2 coat protein	PP7 coat protein	lambda N peptide (amino acids 1-22)	Com protein
Nucleic Acid Sequence	SEQ ID NO: 1	SEQ ID NO :3	SEQ ID NO:5	SEQ ID NO:7
Amino Acid Sequence	SEQ ID NO: 2	SEQ ID NO: 4	SEQ ID NO:6	SEQ ID NO:8
Aptamer (RNA)	SEQ ID NO:9	SEQ ID NO:11	SEQ ID NO:13 (Box-B aptamer)	SEQ ID NO:15
Aptamer (DNA)	SEQ ID NO:10	SEQ ID NO:12	SEQ ID NO:14	SEQ ID NO:16

[0092] In some instances, the mammalian expression vector comprises a sgRNA that comprises one aptamer coding sequence downstream thereof. In other instances, the gRNA may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 aptamer coding sequences. For example, in some instances, the gRNA may comprise two aptamer coding sequences in tandem.

[0093] As used throughout, a sgRNA is a single guide RNA sequence that interacts with a CRISPR-associated endonuclease (a CRISPR site-directed nuclease) and specifically binds to or hybridizes to a target nucleic acid within the genome of a cell (genomic target sequence), such that the sgRNA and the CRISPR-associated endonuclease co-localize to the target nucleic acid in the genome of the cell. Each sgRNA includes a DNA targeting sequence or protospacer sequence of about 10 to 50 nucleotides in length that specifically binds to or hybridizes to a target DNA sequence in the genome. For example, the DNA targeting sequence may be about 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 nucleotides in length. For example, the DNA targeting sequence may be about 15-30 nucleotides, about 15-25 nucleotides, about 10-25 nucleotides, or about 18-23 nucleotides. In one example, the DNA targeting sequence is about 20 nucleotides. In some embodiments, the sgRNA comprises a

crRNA sequence and a transactivating crRNA (tracrRNA) sequence. In some embodiments, the sgRNA does not comprise a tracrRNA sequence.

[0094] Generally, the DNA targeting sequence is designed to complement (e.g., perfectly complement) or substantially complement (e.g., having 1-4 mismatches) to the target DNA sequence. In some cases, the DNA targeting sequence can incorporate wobble or degenerate bases to bind multiple genetic elements. In some cases, the 19 nucleotides at the 3' or 5' end of the binding region are perfectly complementary to the target genetic element or elements. In some cases, the binding region can be altered to increase stability. For example, non-natural nucleotides, can be incorporated to increase RNA resistance to degradation. In some cases, the binding region can be altered or designed to avoid or reduce secondary structure formation in the binding region. In some cases, the binding region can be designed to optimize G-C content. In some cases, G-C content is preferably between about 40% and about 60% (e.g., 40%, 45%, 50%, 55%, 60%). In some cases, the binding region, can be selected to begin with a sequence that facilitates efficient transcription of the sgRNA. For example, the binding region can begin at the 5' end with a G nucleotide. In some cases, the binding region can contain modified nucleotides such as, without limitation, methylated or phosphorylated nucleotides.

[0095] As used herein, the term "complementary" or "complementarity" refers to base pairing between nucleotides or nucleic acids, for example, and not to be limiting, base pairing between a sgRNA and a target sequence. Complementary nucleotides are, generally, A and T (or A and U), and G and C. The guide RNAs described herein can comprise sequences, for example, DNA targeting sequence that are perfectly complementary or substantially complementary (e.g., having 1-4 mismatches) to a genomic sequence.

[0096] The sgRNA includes a sgRNA constant region that interacts with or binds to the CRISPR-associated endonuclease. In the constructs provided herein, the constant region of an sgRNA can be from about 75 to 250 nucleotides in length. In some examples, the constant region is a modified constant region comprising one, two, three, four, five, six, seven, eight, nine, ten or more nucleotide substitutions in the stem, the stem loop, a hairpin, a region in between hairpins, and/or the nexus of a constant region. In some instances, a modified constant region that has at least 80%, 85%, 90%, or 95% activity, as compared to the activity of the natural or wild-type sgRNA constant region from which the modified constant region is derived, may be used in the constructs described herein. In particular, modifications should not be made at nucleotides that interact directly with a CRISPR-associated endonuclease or at nucleotides that are important for the secondary structure of the constant region.

[0097] The mammalian expression plasmids comprise a eukaryotic promoter operably linked to the non-viral nucleic acid sequence. In some instances, a RNA polymerase II promoter is operably linked to the nucleic acid encoding the fusion protein comprising the CRISPR-associated endonuclease and the polypeptide comprising the DNA polymerase domain; and a RNA polymerase III promoter is operably linked to the gRNA coding sequence.

[0098] The RNA polymerase II promoter sequence is selected from a mammalian species. The RNA polymerase III promoter sequences is selected from a mammalian species. For example, these promoter sequences can be selected from a human, cow, sheep, buffalo, pig, or mouse, to name a few. In some examples, the RNA polymerase II promoter sequence is a CMV, FE1α, or SV40 sequence. In some examples, the RNA polymerase III promoter sequence is a U6 or an H1 sequence. In some examples, the RNA polymerase II sequence is a modified RNA polymerase II sequence. For example, the RNA polymerase II sequences having at least 80%, 85%, 90%, 95%, or 99% identity to a wild-type RNA polymerase II promoter sequence from any mammalian species can be used in the constructs provided herein. In some examples, the RNA polymerase III sequence is a modified RNA polymerase III sequence. For example, the RNA polymerase III sequences having at least 80%, 85%, 90%, 95%, or 99% identity to a wild-type RNA polymerase III promoter sequence from any mammalian species can be used in the constructs provided herein. Those of skill in the art readily understand how to determine the identity of two polypeptides or nucleic acids. For example, the identity can be calculated after aligning the two sequences so that the identity is at its highest level. Another way of calculating identity can be performed by published algorithms. For example, optimal alignment of sequences for comparison can be conducted using the algorithm of Needleman and Wunsch, J. Mol. Biol. 48(3): 443-453 (1970). In some instances, the eukaryotic promoter is an inducible or regulatable promoter.

[0099] Coding sequences transcribed from a RNA pol II promoter include a poly(A) signal and a transcription terminator sequence downstream of the coding sequence. Commonly used mammalian terminators (SV40, hGH, BGH, and rbGlob) include the sequence motif AAUAAA (SEQ ID NO: 31) which promotes both polyadenylation and termination. Coding sequences transcribed from a RNA pol III promoter include a simple run of T residues downstream of the coding sequence as a terminator sequence. The role of the terminator, a sequence-based element, is to define the end of a transcriptional unit (such as a gene) and initiate the process of releasing the newly synthesized RNA from the transcription machinery.

Terminators are found downstream of the gene to be transcribed, and typically occur directly after any 3' regulatory elements, such as the polyadenylation or poly(A) signal.

[0100] In some instances, the mammalian expression plasmid may also include at least one polynucleotide sequence encoding a RNA-stabilizing sequence positioned downstream of the CRISPR component coding sequence or the aptamer coding sequence if positioned downstream of the CRISPR component coding sequence. The polynucleotide sequence encoding the RNA-stabilizing sequence is transcribed downstream of the CRISPR/Cas system component coding sequence and stabilizes the longevity of the transcribed RNA sequence. In one example, the polynucleotide sequence encoding the RNA-stabilizing sequence is positioned downstream of the catalytically impaired CRISPR-associated endonuclease coding sequence. In another example, the polynucleotide sequence encoding the RNA-stabilizing sequence is positioned downstream of the gRNA coding sequence. An exemplary RNA-stabilizing sequence is the sequence of the 3' UTR of human beta globin gene as set forth in SEQ ID NO: 20 (DNA) and SEQ ID NO: 21 (RNA). Another example of an RNA-stabilizing sequence is SEQ ID NO: 22 which comprises two copies of SEQ ID NO: 20. Other RNA-stabilizing sequences are described in Hayashi, T. et al., Developmental Dynamics 239(7):2034-2040 (2010) and Newbury, S. et al., Cell 48(2):297-310 (1987). In some instances, a spacer of 1-30 nucleotides may be positioned between the CRISPR component coding sequence and the at least one polynucleotide sequence encoding RNA-stabilizing sequence.

[0101] In some instances, the mammalian expression plasmid may comprise one or more expression cassettes. In some instances the mammalian expression plasmid comprises a first expression cassette that encodes any of the fusion proteins described herein and a second expression cassette that encodes the gRNA comprising at least one aptamer. In some instances, the mammalian expression plasmid may also comprise a reporter gene.

Systems

[0102] Another aspect of this disclosure are lentiviral packaging systems. Such systems include the mammalian expression plasmids described in this disclosure. These systems are useful in providing components for introduction into mammalian cells to generate the lentivirus-like particles described in this disclosure.

[0103] In some instances, the system includes a lentiviral packaging plasmid comprising a eukaryotic promoter operably linked to a viral sequence, for example, a Gag nucleotide sequence, wherein the Gag nucleotide sequence comprises a nucleocapsid (NC) coding sequence and a matrix protein (MA) coding sequence, wherein one or both of the NC coding

sequence or the MA coding sequence comprise at least one non-viral aptamer-binding protein (ABP) nucleotide sequence, and wherein the packaging plasmid does not encode a functional integrase protein.

[0104] For example, provided herein is a lentiviral packaging system comprising: (a) a packaging plasmid comprising a eukaryotic promoter operably linked to a Gag nucleotide sequence, wherein the Gag nucleotide sequence comprises a nucleocapsid (NC) coding sequence and a matrix protein (MA) coding sequence, wherein one or both of the NC coding sequence or the MA coding sequence comprises at least one non-viral aptamer-binding protein (ABP) nucleotide sequence, and wherein the packaging plasmid does not encode a functional integrase protein; (b) at least one mammalian expression plasmid comprising (i) a nucleic acid sequence encoding a fusion protein comprising a CRISPR-associated endonuclease and a DNA polymerase domain; and (ii) a gRNA described herein; and (c) an envelope plasmid comprising an envelope glycoprotein coding sequence.

[0105] The system may include a second generation packaging plasmid or third generation packaging plasmids or modified versions thereof. In some instances, the packaging plasmid includes the Gag nucleotide sequence as described above and further comprises a Rev nucleotide sequence and a Tat nucleotide sequence. In other instances, the system includes a first packaging plasmid including a Gag nucleotide sequence as described above and a second packaging plasmid comprising a Rev nucleotide sequence. In each of the packaging plasmids, the viral protein coding sequences are operably linked to a eukaryotic promoter for example, each individually or one promoter for multiple protein coding sequences. The system may include a second generation packaging plasmid or third generation packaging plasmids or modified versions thereof.

[0106] In some instances, the ABP coding sequence is at the 5' end or 3' end of the viral protein coding sequence, i.e., at the 5' end or the 3' end of the NC or MA coding sequence. In some instances, the ABP coding sequence may be inserted into the viral protein coding sequence such that the encoded ABP is fused to the viral protein. The ABP coding sequence may be inserted in frame at an internal position within the viral protein coding sequence. When positioned in frame at an internal position near the 5' or 3' end of the viral protein coding sequence, the ABP coding sequence is positioned so as not to disrupt processing sequences such as those described in Tritch, R.J. et al., *J. Virol.* 65(2):922-30 (1991) and Scarlata, S. and Carter, C., *Biochimica et Biophysica Acta* – Biomembranes 1614(1):62-72 (2003), which are incorporated herein by reference in their entirety. For example, the Gag nucleotide sequence

encodes, *inter alia*, the NC coding sequence and the MA coding sequence, and the Gag precursor protein is processed by proteolytic cleavage into separate mature viral proteins. The in frame insertion of the ABP coding sequence would not disrupt the nucleotides encoding the processing sequences for proteolytic cleavage. In some instances, nucleotides in the viral protein coding sequence may be replaced with the ABP protein coding sequence. In some instances, a linker sequence encoding 3-6 amino acids may be positioned between the viral protein coding sequence and the ABP coding sequence, or flanking the ABP coding sequence, to help facilitate proper folding of the protein domains upon expression.

[0107] In one example, the modified viral protein is NC and the ABP coding sequence is inserted at the 5' end or the 3' end of the NC coding sequence. In another example, the modified viral protein is NC and the ABP coding sequence is inserted before or after one of the zinc finger (ZF) domains. For example, the ABP coding sequence may be inserted after the last codon of the second ZF (ZF2) domain. In another example, the ABP coding sequence may be inserted before the first codon of the ZF2 domain. In another example, the ABP coding sequence may be inserted before the first codon of the first ZF (ZF1) domain. In another example, the ABP coding sequence may be inserted after the last codon of the first ZF (ZF1) domain. In some instances, the ABP coding sequence is inserted into the NC coding sequence in a manner that does not disrupt the highly positive stretch of amino acids in the NC protein.

[0108] In another example, the modified viral protein is MA and the ABP coding sequence is inserted at the 5' end or the 3' end of the MA coding sequence. In another example, the ABP coding sequence is inserted in frame at an internal position within the MA coding sequence. In some instances, nucleotides in the MA coding sequence may be replaced with the ABP protein coding sequence. For example, the nucleotides encoding amino acids 44-132 of the MA protein may be replaced with the ABP coding sequence. In another example, the ABP coding sequence is inserted prior to the codon encoding amino acid 44 of the MA protein. In another example, the ABP coding sequence is inserted after the codon encoding amino acid 132 of the MA protein.

[0109] In some instances, the system includes a packaging plasmid comprising a eukaryotic promoter operably linked to a NEF coding sequence or a VPR coding sequence, wherein the NEF coding sequence or the VPR coding sequence comprises at least one non-viral ABP nucleotide sequence. The system may include a second generation packaging plasmid or third generation packaging plasmids or modified versions thereof. In some instances, the packaging plasmid includes a Gag nucleotide sequence, a Rev nucleotide

sequence, and a Tat nucleotide sequence. In other instances, the system includes a first packaging plasmid including a Gag nucleotide sequence and a second packaging plasmid comprising a Rev nucleotide sequence.

[0110] In some instances, the modified viral protein is VPR and the ABP coding sequence is inserted at the 5' end or the 3' end of the VPR coding sequence. In one example, the ABP coding sequence is inserted at the 5' end of the VPR coding sequence.

[0111] In other instances, the modified viral protein is NEF and the ABP coding sequence is inserted at the 5' end or the 3' end of the NEF coding sequence. In one example, the ABP coding sequence is inserted at the 3' end of the NEF coding sequence.

[0112] In some instances, the coding sequence of the viral protein may be one of SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, or SEQ ID NO:25. In some instances, the amino acid sequence of the the viral protein may be one of SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, or SEQ ID NO:26. In some instances, the lentiviral packaging plasmid comprises a sequence encoding at least one of SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, or SEQ ID NO:26 operably linked to a eukaryotic promoter. In some instances, if the viral protein is NEF, the polypeptide may comprise three mutations that enhances packaging in the viral capsid such as, for example, the following substitution mutations: G3C, V153L, and E177G.

[0113] In some instances, the plasmids may encode one or more viral proteins that comprise two or more aptamer-binding proteins fused thereto. In certain instances, the Gag nucleotide sequence of the lentiviral packaging plasmid may comprise a NC coding sequence and a MA coding sequence and where one or both of the NC coding sequence or the MA coding sequence comprises a first non-viral ABP nucleotide sequence and a second non-viral ABP nucleotide sequence. The first non-viral ABP nucleotide sequence and the second non-viral ABP nucleotide sequence may both encode the same ABP. Alternatively, the first non-viral ABP nucleotide sequence and the second non-viral ABP nucleotide sequence encode different ABPs. In some instances, the Gag nucleotide sequence of the lentiviral packaging plasmid may comprise a NC coding sequence comprising at least one first non-viral ABP nucleotide sequence and a MA coding sequence comprising at least one second non-viral ABP nucleotide sequence . The at least one first non-viral ABP nucleotide sequence and the at least one second non-viral ABP nucleotide sequence may both encode the same ABP. Alternatively, the at least one first non-viral ABP nucleotide sequence and the at least one second non-viral ABP nucleotide sequence encode different ABPs.

[0114] In certain instances, the packaging plasmid may encode a VPR coding sequence or a NEF coding sequence and where the VPR coding sequence or the NEF coding sequence comprises a first non-viral ABP nucleotide sequence and a second non-viral ABP nucleotide sequence. The first non-viral ABP nucleotide sequence and the second non-viral ABP nucleotide sequence may both encode the same ABP. Alternatively, the first non-viral ABP nucleotide sequence and the second non-viral ABP nucleotide sequence encode different ABPs.

[0115] A non-viral aptamer-binding protein (ABP) nucleotide sequence encodes a polypeptide sequence that binds to an RNA aptamer sequence. Several non-viral ABPs are suitable for use in this disclosure. In particular, suitable ABPs include bacteriophage RNA-binding proteins that bind specifically to RNA sequences that form stem-loop structures referred to as RNA aptamer sequences. Exemplary non-viral aptamer binding protein include MS2 coat protein, PP7 coat protein, lambda N peptide, and Com (Control of mom) protein. The lambda N peptide may be amino acids 1-22 of the lambda N protein, which are the RNA-binding domain of the protein. In some instances, the ABPs bind to their aptamers as dimers. Information about these ABP and the aptamer sequences to which they bind is provided in Table 1. In some embodiments, the at least one non-viral ABP nucleotide sequence encodes a polypeptide having the sequence set forth in any of SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, or SEQ ID NO:16. In some embodiments, the at least one non-viral ABP nucleotide sequence comprises any of SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, or SEQ ID NO:15.

[0116] A feature of the lentiviral packaging plasmids provided herein is that they may not encode a functional integrase protein. When the packaging plasmids do not encode a functional integrase protein and they are used in the systems and methods described herein, there is substantially reduced risk the nucleic acid molecules carried by the lentivirus-like particles produced using these packaging plasmids will integrate into the genome of the transduced eukaryotic cell. In some instances, the lentiviral packaging plasmid comprises an integrase coding sequence with an integrase-inactivating mutation therein. For example, the integrase-inactivating mutation may be an aspartic acid to valine mutation at amino acid position 64 (D64V) of the integrase protein encoded by the integrase coding sequence. In some instances, the lentiviral packaging plasmid comprises a deletion of all or a portion of an integrase coding sequence.

[0117] In some embodiments, the lentiviral packaging plasmids comprise a eukaryotic promoter operably linked to the Gag nucleotide sequence. In some embodiments, the

mammalian expression plasmids comprise a eukaryotic promoter operably linked to the VPR coding sequence or the NEF coding sequence. In some instances, the eukaryotic promoter is a RNA polymerase II promoter. The RNA polymerase II promoter sequence is selected from a mammalian species. For example, the promoter sequence can be selected from a human, cow, sheep, buffalo, pig, or mouse, to name a few. In some examples, the RNA polymerase II promoter sequence is a CMV, FE1α, or SV40 sequence. In some examples, the RNA polymerase II sequence is a modified RNA polymerase II sequence. For example, the RNA polymerase II sequences having at least 80%, 85%, 90%, 95%, or 99% identity to a wild-type RNA polymerase II promoter sequence from any mammalian species can be used in the constructs provided herein. Those of skill in the art readily understand how to determine the identity of two polypeptides or nucleic acids, as described above.

[0118] Coding sequences transcribed from a RNA pol II promoter include a poly(A) signal and a transcription terminator sequence downstream of the coding sequence. Commonly used mammalian terminators (e.g., SV40, hGH, BGH, and rbGlob) include the sequence motif AAUAAA which promotes both polyadenylation and termination. The role of the terminator, a sequence-based element, is to define the end of a transcriptional unit (such as a gene) and initiate the process of releasing the newly synthesized RNA from the transcription machinery. Terminators are found downstream of the gene to be transcribed, and typically occur directly after any 3' regulatory elements, such as the polyadenylation or poly(A) signal.

[0119] In some instances, the lentiviral packaging plasmids may comprise one or more expression cassettes.

[0120] The system also can include an envelope plasmid having an envelope coding sequence that encodes a viral envelope glycoprotein. For example, the Env nucleotide sequence may encode VSV-G. The envelope coding sequence is operably linked to a eukaryotic promoter. Appropriate eukaryotic promoters are described above. In some instances, the eukaryotic promoter is a RNA pol II promoter.

[0121] The system can comprise any of the packaging plasmids, envelope plasmids and mammalian expression plasmids, i.e., a mammalian expresson plasmid comprising (i) a nucleic acid sequence encoding a fusion protein comprising a CRISPR-associated endonuclease and a polpeptide comprising a DNA polymerase domain; and (ii) a gRNA comprising at least one aptamer, described herein. When any of the packaging plasmids, mammalian expression plasmids and envelope plasmids described herein are delivered to eukaryotic cells as a system, the gRNA expressed by the mammalian expression plasmid forms a complex with the CRISPR-

associated endonuclease expressed by the mammalian expression plasmids to form an RNP that is packaged by the viral particles produced by the eukaryotic cells, via the interaction between the aptamer fused or linked to the gRNA and the ABP linked to the viral protein expressed by the packaging plasmid.

[0122] Also provided herein are kits the include the components of the systems described in this disclosure. In some embodiments, the kits include one or more of the plasmids described herein.

Lentivirus-like Particles

[0123] In another aspect, provided are lentivirus-like particles, for example, lentivirus-like particles made by any of the methods described herein. As used herein, a lentivirus-like particle is multiprotein structure that mimics the organization and conformation of authentic native viruses but lacks the viral genome. A plurality of lentivirus-like particles are also provided. The lentivirus-like particles contain a modified lentiviral protein that is a fusion protein in which at least one aptamer-binding protein is fused to one or more viral proteins. In the context of this disclosure, the modified viral protein may be structural or non-structural. Exemplary structural proteins are lentiviral nucleocapsid (NC) protein and matrix (MA) protein. Exemplary non-structural proteins are viral protein R (VPR) and negative regulatory factor (NEF). In some instances, the particles contain a fusion protein comprising a NC protein and a MA protein where one or both thereof are fused with at least one non-viral aptamer binding protein (ABP). The NC protein of the particles may have two functional zinc finger protein domains. In particular, retention of the second NC zinc finger domain may preserve the efficiency of viral assembly and budding. In some instances, the particles contain a fusion protein comprising a VPR protein or a NEF protein where the VPR protein or the NEF protein are fused with at least one non-viral ABP. The particles also contain an RNP comprising: (i) a fusion protein comprising a CRISPR-associated endonuclease and a polypeptide comprising a DNA polymerase domain; and (ii) a gRNA. Any of the mammalian expression plasmids described herein comprising a non-viral nucleic acid sequence, wherein at least one aptamer is attached or inserted into the gRNA sequence, can be used to generated lentivirus-like particles containing RNPs. In some instances, the lentivirus-like particles do not contain a functional integrase protein. These virus-like particles are useful to transduce eukaryotic cells of interest.

[0124] The particles may comprise a viral fusion protein comprising one or more ABPs. In some instances, the particles contain a NC protein, a MA protein, or both, where one or both of the NC protein or MA protein are fused with one or more non-viral ABP. In some instances,

lentivirus-like particles comprise a NC protein fused with at least one non-viral ABP. In some instances, lentivirus-like particles comprise a MA protein fused with at least one non-viral ABP. In some instances, the lentivirus-like particles may comprise a NC protein and a MA protein, where one or both of the NC protein or the MA protein may be fused with two non-viral ABP proteins, a first non-viral ABP and a second non-viral ABP fused to a C' terminal end of the first non-viral ABP (i.e. in tandem). In certain instances, the particles may contain one or both of a NC protein or a MA protein fused with a first non-viral ABP and a second non-viral ABP.

[0125] In some instances, the lentivirus-like particle contains a VPR protein or a NEF protein, where the VPR protein or the NEF protein is fused to one or more non-viral ABP. In some instances, the lentivirus-like particle contains a VPR protein or a NEF protein fused to two non-viral ABP, a first non-viral ABP and a second non-viral ABP fused to a C' terminal end of the first non-viral ABP (i.e. in tandem). In some instances, the lentivirus-like particle contains a VPR protein or a NEF protein fused to a first non-viral ABP and a second non-viral ABP. The first non-viral ABP and the second non-viral ABP may both be the same ABP. Alternatively, the first non-viral ABP and the second non-viral ABP may be different ABPs. In some instances, the lentivirus-like particles may comprise a NC protein with at least one first non-viral ABP fused to MA protein with at least one second non-viral ABP fused to its C' terminal end. The at least one first non-viral ABP and the at least one second non-viral ABP both be the same ABP. Alternatively, the at least one first non-viral ABP protein and the at least one second non-viral ABP may be different ABPs. The first non-viral ABP and the second non-viral ABP may both be the same ABP. Alternatively, the first non-viral ABP and the second non-viral ABP may be different ABPs.

[0126] A non-viral ABP is a polypeptide sequence that binds to an RNA aptamer sequence. Several non-viral ABPs are suitable for use in this disclosure. In particular, suitable ABPs include bacteriophage RNA-binding proteins that bind specifically to known RNA aptamer sequences, which are RNA sequences that form stem-loop structures. Exemplary non-viral aptamer binding protein include MS2 coat protein, PP7 coat protein, lambda N peptide, and Com (Control of mom) protein. The lambda N peptide may be amino acids 1-22 of the lambda N protein, which are the RNA-binding domain of the protein. Information about these ABP and the aptamer sequences to which they bind is provided above in Table 1.

[0127] The lentivirus-like particles may comprise various lentiviral proteins. However, in some instances, the lentivirus-like particles do not comprise all of the types of proteins or

nucleic acids found in native lentiviruses. In some instances, the particles may contain NC, MA, CA, SP1, SP2, P6, POL, ENV, TAT, REV, VIF, VPU, VPR, and/or NEF proteins, or a derivative, combination, or portion of any thereof. In some instances, the particles may contain NC, MA, CA, SP1, SP2, P6, and POL. In some instances, the lentivirus-like particles may comprise only those proteins that form the viral shell (capsid). In some instances, one or more lentiviral proteins may be excluded in full or in part from the lentivirus-like particles. For example, in some instances, the lentivirus-like particles may not contain a POL protein or may comprise a non-functional version of a POL protein such as, for example, a POL protein with an inactivating point mutation or an inactivating truncation. In another example, the lentivirus-like particles may not contain an integrase protein or may comprise a non-functional version of an integrase protein such as, for example, an integrase protein with an inactivating point mutation or an inactivating truncation. For example, the lentivirus-like particle may contain a non-functional integrase protein comprising an aspartic acid to valine mutation at amino acid position 64 (D64V). In another example, the lentivirus-like particles may not contain a reverse transcriptase protein or may comprise a non-functional version of a reverse transcriptase protein such as, for example, a reverse transcriptase protein with an inactivating point mutation or an inactivating truncation.

[0128] As set forth above, gRNA generally comprises a DNA targeting sequence and a constant region that interacts with the CRISPR-associated endonuclease. In some instances, the gRNA may comprise a transactivating crRNA (tracrRNA) sequence. For example, the gRNA may comprise a tracrRNA where it is to be used in conjunction with a Cas9 protein or derivative. In other instances, the gRNA does not comprise a tracrRNA sequence. For example, the gRNA may not comprise a tracrRNA sequence where it is to be used in conjunction with a Cpf1 protein or derivative.

[0129] In some instances, the gRNA comprises at least one aptamer sequence. In some instances, the at least one aptamer sequence may be positioned at the 5' end or the 3' end of the gRNA. In some instances, the at least one aptamer sequence may be inserted at an internal position within the gRNA such as, for example, at one or more of the loops formed in the folded gRNA. For example, where the gRNA is for a Cas9 protein, the at least one aptamer sequence may be positioned at the tetra loop, the stem loop 2 (ST2), or the 3' end of the gRNA. In some instances, a spacer of 1-30 ribonucleotides may be positioned between the gRNA and the at least one aptamer sequence, or flanking the at least one aptamer sequence. In certain instances, at least one aptamer sequence does not interfere with lentivirus-like particle transduction of

eukaryotic cells. For example, at least one non-viral ABP fused to one or more of the NC protein, the MA protein, the VPR protein, or the NEF protein may not interfere with lentivirus-like particle transduction of eukaryotic cells.

Gene Editing Methods

[0130] Described herein are methods of using the plasmids and systems provided in this disclosure in CRISPR/Cas systems for editing DNA targets, for example, a gene, in the genome of a eukaryotic cell. In the methods provided herein, eukaryotic cells comprising a target genomic sequence of interest to be modified are transduced with lentivirus-like particles that contain a viral fusion protein comprising a viral protein fused to at least one aptamer-binding protein (ABP) and an RNP comprising (1) a gRNA and (2) a fusion protein comprising a CRISPR-associated endonuclease and a polypeptide comprising a DNA polymerase domain.

[0131] The methods described herein can be used to edit the genome of the cell, while reducing on-target deletions of any size. For example, deletions of less than 100 bp, 90bp, 80 bp, 70 bp, 60 bp, 50 bp, 40 bp 30 bp, 20 bp, 10 bp, 9 bp, 8 bp, 7 bp, 6 bp, 5 bp, 4 bp, 3 bp, 2 bp or 1 bp can be reduced. In some embodiments, large genomic deletions of over 100 bp, 200 bp, 300 bp, 400 bp, or 500 bp in size can be reduced. As used herein, the term "on-target deletion," refers to a deletion that occurs at or near an sgRNA target site in the genome of the cell, i.e., in the proximity of the target genomic sequence of interest to be modified. In the methods provided herein, large genomic deletions can be reduced by at least 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 90%, 95%, 99% or greater, as compared to the number of large genomic deletions when the cells are edited with a CRISPR-associated endonuclease that is not fused or linked to a polypeptide comprising a DNA polymerase domain.

[0132] In some instances, the provided methods increase non-homologous end joining (NHEJ), as compared to non-NHEJ in a cell edited with a CRISPR-associated endonuclease that is not fused or associated with a polypeptide comprising a DNA polymerase domain. In some embodiments, NJEH is increased as compared to non-NHEJ end joining in the cell. In some embodiments, non-NJEH is microhomology-mediated end joining (MMEJ) and/or single-stranded annealing (SSA). In some embodiments, the increase is at least a 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 200%, 300%, or 400% increase. In some embodiments, the increase is a 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 20-fold, 30-fold, 40-fold, 50-fold increase or greater.

[0133] In some embodiments, the ratio of NHEJ to non-NHEJ increases in the cell. In some methods, the ratio of on-target one base pair (1-bp) deletions to on-target deletions greater than

1-bp is increased in the cell. For example, in some embodiments, the ratio of on-target one base pair (1-bp) deletions to deletions greater than 500 base pairs is increased in the cell.

[0134] In some embodiments, the number of non-templated insertions (non-TIS) decreases in the cell. In some embodiments, the decrease is at least 10%, 20%, 30%, 40%, 50%, 60%, 80%, 90%, 95%, 99%, or 100%,

[0135] In some embodiments, the number of 1-bp templated insertions (TIS) increases in the cell. In some embodiments, the increase is at least a 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 200%, 300%, or 400% increase. In some embodiments, the increase is a 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 20-fold, 30-fold, 40-fold, 50-fold increase or greater.

[0136] In some embodiments, the ratio of 1-bp TIS to 1-bp non-TIS increases in the cell. As used herein, "a templated insertion" is a nonrandom insertion that occurs after Cas9-induced double stranded breakage leaves a 1-nt 5′ overhanging base that is subsequently filled in by a DNA polymerase, and ligated.

[0137] In some embodiments, the exonuclease activity of MRE11 Double Strand Break Repair Nuclease (MRE11) is reduced (e.g., relative to wild-type MRE11). For example, in the methods provided herein, the exonuclease activity of MRE11 can be reduced by at least 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 90%, 95%, 99% or more.

[0138] When lentivirus-like particles lacking integrase activity are used in the method, there is reduced risk of integration into the cell genome of any of the nucleic acids carried by the particles. In some instances, the lentiviral-particles used lack portions of the lentiviral genomic sequences that are essential for viral replication and, as such, reduce the risk of continued particle production. Another advantage of the provided components is that the viral fusion protein may increase packaging of RNPs, into the lentivirus-like particles, which in turn increase genome editing efficiency.

[0139] In some instances, the transduced eukaryotic cells are mammalian cells. In some instances, the eukaryotic cells may be *in vitro* cultured cells. In some instances, the eukaryotic cells may be *ex vivo* cells obtained from a subject. In other instances, the eukaryotic cells are present in a subject. As used throughout, by subject is meant an individual. For example, the subject is a mammal, such as a primate, and, more specifically, a human. Non-human primates are subjects as well. The term subject includes domesticated animals, such as cats, dogs, etc., livestock (for example, cattle, horses, pigs, sheep, goats, etc.) and laboratory animals (for example, ferret, chinchilla, mouse, rabbit, rat, gerbil, guinea pig, etc.). Thus, veterinary uses

and medical uses and formulations are contemplated herein. The term does not denote a particular age or sex. Thus, adult and newborn subjects, whether male or female, are intended to be covered. As used herein, patient or subject may be used interchangeably and can refer to a subject afflicted with a disease or disorder. The lentivirus-like particles provided herein may be administered to the subject, for example, injected into a subject, according to known, routine methods. Exemplary modes of administration include oral, rectal, transmucosal, topical, intranasal, inhalation (e.g., via an aerosol), buccal (e.g., sublingual), vaginal, intrathecal, intraocular, transdermal, intradermal, intrapleural, intracerebral, and intraarticular), topical, and the like, as well as direct tissue or organ injection. Administration can also be to a tumor. The most suitable route in any given case will depend on the nature and severity of the condition being treated and on the nature of the particular lentivirus-like particle that is being used. In some instances, the lentivirus-like particles are injected intravenously (IV), intraperitoneally (IP), intramuscularly, or into a specific organ or tissue. In some embodiments, more than one administration (e.g., two, three, four or more administrations) may be employed to achieve the desired level of gene editing over a period of various intervals, e.g., daily, weekly, monthly, yearly, etc.

[0140] An effective amount of any of the recombinant lentivirus-like particles described herein will vary and can be determined by one of skill in the art through experimentation and/or clinical trials. For example, an effective dose can be from about 10^6 to about 10^{15} lentivirus-like particles, for example, from about 10^6 to about 10^{14}, from about 10^6 to about 10^{13}, from about 10^6 to about 10^{12} lentivirus-like particles, from about 10^6 to about 10^{12}, from about 10^6 to about 10^{11}, or from about 10^6 to about 10^{11} lentivirus-like particles. Other effective dosages can be readily established by one of ordinary skill in the art through routine trials establishing dose response curves. See, for example, Mangeot et al. "Genome editing in primary cells and *in vivo* using viral-derived Nanoblades loaded with Cas9-sgRNA ribonucleoproteins," *Nat Commun* 10, 45 (2019). doi.org/10.1038/s41467-018-07845-z.

[0141] In some instances, the provided methods are for modifying a target locus of interest, the method comprising transducing a plurality of eukaryotic cells with a plurality of viral particles, wherein the plurality of viral particles comprise (i) a fusion protein comprising a viral protein, for example, NC, MA, VRP, or NEF, wherein the viral protein comprises at least one non-viral aptamer binding protein (ABP); and (ii) a ribonucleotide protein (RNP) complex comprising (1) a gRNA and (2) a fusion protein comprising a CRISPR-associated endonuclease and a polypeptide comprising a DNA polymerase domain, wherein the RNP complex is capable

of binding (e.g., preferentially binding) via the gRNA, to the genomic target sequence in genomic DNA of the cell and the CRISPR-associated endonuclease alters the genomic DNA of the cell. As described above, the RNP complexes are packaged into the viral particles via the interaction of an aptamer sequence attached to or inserted into a gRNA sequence that forms a complex with the CRISPR-associated endonuclease.

[0142] Generally, the sgRNA is targeted to specific regions at or near a gene. In some instances, the sgRNA can be targeted to a region where single base changes are necessary, for example, to correct a single base mutation in the human beta-globin gene that causes sickle cell anemia. The sgRNA allows the RNP complexes described herein to a specific site in the genomic sequence of a cell.

[0143] In some instances, the modifications to the system components as described in this disclosure do not impair how the system components function following transduction into eukaryotic cells. Rather, the components may function similarly or better than unmodified components upon transduction into eukaryotic cells. For example, the viral fusion proteins in the lentivirus-like particles may not interfere with the lentivirus-like particle transduction of eukaryotic cells. Similarly, if the RNP complexes packaged in the lentivirus-like particles comprise at least one aptamer sequence, the at least one aptamer sequence may not interfere with the lentivirus-like particle transduction of eukaryotic cells. In some instances, the lentivirus-like proteins containing viral fusion protein may result in greater gene editing upon transduction into eukaryotic cells relative to lentivirus-like particles that do not comprise a viral fusion protein. In one example the viral fusion protein may be a NC-ABP fusion protein, such as a NC-MS2 fusion protein or NC-PP7 fusion protein. In one example, the NC fusion protein is fused to one or two ABPs, such as one or two MS2 proteins, one or two PP7 proteins, or one MS2 protein and one PP7 protein.

[0144] The eukaryotic cells can be *in vitro*, *ex vivo* or *in vivo*. In some embodiments, the cell is a primary cell (isolated from a subject). As used herein, a primary cell is a cell that has not been transformed or immortalized. Such primary cells can be cultured, sub-cultured, or passaged a limited number of times (e.g., cultured 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 times). In some cases, the primary cells are adapted to *in vitro* culture conditions. In some cases, the primary cells are isolated from an organism, system, organ, or tissue, optionally sorted, and utilized directly without culturing or sub-culturing. In some cases, the primary cells are stimulated, activated, or differentiated. In some embodiments, the cells are cultured under conditions effective for expanding the population of modified cells. In some

embodiments, cells modified by any of the methods provided herein are purified. In some cases, cells are removed from a subject, modified using any of the methods described herein and re-administered to the patient.

[0145] In some instances, once the cells have been transduced with the viral particles described above, the cells are cultured for a sufficient amount of time to allow for gene editing to occur, such that a pool of cells expressing a detectable phenotype can be selected from the plurality of transduced cells. The phenotype can be, for example, cell growth, survival, or proliferation. In some examples, the phenotype is cell growth, survival, or proliferation in the presence of an agent, such as a cytotoxic agent, an oncogene, a tumor suppressor, a transcription factor, a kinase (e.g., a receptor tyrosine kinase), a gene (e.g., an exogenous gene) under the control of a promoter (e.g., a heterologous promoter), a checkpoint gene or cell cycle regulator, a growth factor, a hormone, a DNA damaging agent, a drug, or a chemotherapeutic. The phenotype can also be protein expression, RNA expression, protein activity, or cell motility, migration, or invasiveness. In some examples, the selecting the cells on the basis of the phenotype comprises fluorescence activated cell sorting, affinity purification of cells, or selection based on cell motility.

[0146] In some examples, the selecting the cells comprises analysis of the genomic DNA of the cells such as by amplification, sequencing, SNP analysis, etc. Sequencing methods include, but are not limited to, shotgun sequencing, bridge PCR, Sanger sequencing (including microfluidic Sanger sequencing), pyrosequencing, massively parallel signature sequencing, nanopore DNA sequencing, single molecule real-time sequencing (SMRT) (Pacific Biosciences, Menlo Park, CA), ion semiconductor sequencing, ligation sequencing, sequencing by synthesis (Illumina, San Diego, Ca), Polony sequencing, 454 sequencing, solid phase sequencing, DNA nanoball sequencing, heliscope single molecule sequencing, mass spectroscopy sequencing, pyrosequencing, Supported Oligo Ligation Detection (SOLiD) sequencing, DNA microarray sequencing, RNAP sequencing, tunneling currents DNA sequencing, and any other DNA sequencing method identified in the future. One or more of the sequencing methods described herein can be used in high throughput sequencing methods. As used herein, the term "high throughput sequencing" refers to all methods related to sequencing nucleic acids where more than one nucleic acid sequence is sequenced at a given time.

Methods of Treatment

[0147] Any of the methods and compositions described herein can be used to treat a disease (e.g., cancer, a blood disorder (for example, sickle cell anemia or beta thalassemia), an infectious disease, an autoimmune disease, transplantation rejection, graft vs. host disease or other inflammatory disorder) in a subject. In some embodiments, the infectious disease is a viral infection, for example, but not limited to human immunodeficiency virus (HIV), herpes simplex virus type 1 (herpetic stromal keratitis), and human papilloma virus (HPV). In some embodiments, the disease is a neurodegerative or myodegemerative disease, for example, Duchenne muscular dystrophy.

[0148] In some methods, the cancer to be treated is selected from a cancer of B-cell origin, breast cancer, gastric cancer, neuroblastoma, osteosarcoma, lung cancer, colon cancer, chronic myeloid cancer, leukemia (e.g., acute myeloid leukemia, chronic lymphocytic leukemia (CLL) or acute lymphocytic leukemia (ALL)), prostate cancer, colon cancer, renal cell carcinoma, liver cancer, kidney cancer, ovarian cancer, stomach cancer, testicular cancer, rhabdomyosarcoma, and Hodgkin's lymphoma. In some embodiments, the cancer of B-cell origin is selected from the group consisting of B-lineage acute lymphoblastic leukemia, B-cell chronic lymphocytic leukemia, and B-cell non-Hodgkin's lymphoma.

[0149] In some methods, the cells of the subject are modified *in vivo*. In some methods, the method of treating a disease in a subject comprises: a) obtaining cells from the subject; b) modifying the cells using any of the methods provided herein; and c) administering the modified cells to the subject. See, for example, Milone and O'Doherty "Clinical sue of lentiviral vectors," *Leukemia* 32, 1529-1541 (2018). Optionally, the disease is selected from the group consisting of cancer, a blood disorder (for example, sickle cell anemia or beta thalassemia), an infectious disease, an autoimmune disease, transplantation rejection, graft vs. host disease or other inflammatory disorder in a subject. In some methods for treating cancer, the cells obtained from the subject are modified to express a tumor specific antigen. As used throughout, the phrase "tumor-specific antigen" means an antigen that is unique to cancer cells or is expressed more abundantly in cancer cells than in in non-cancerous cells. Optionally, the cells obtained from the subject are T cells. Optionally, the modified cells are expanded prior to administration to the subject.

[0150] The lentivirus-like particles or cells described herein can be formulated as a pharmaceutical composition. Therefore, provided herein is a pharmaceutical composition comprising any of the lentivirus-like particles described herein. Also provided is a

pharmaceutical composition comprising any of the modified cells described herein Optionally, the pharmaceutical composition can further comprise a carrier. The term carrier means a compound, composition, substance, or structure that, when in combination with lentivirus-like particles or cells, aids or facilitates preparation, storage, administration, delivery, effectiveness, selectivity, or any other feature of the lentivirus-like particles or cells for its intended use or purpose. For example, a carrier can be selected to minimize any degradation of the active ingredient and to minimize any adverse side effects in the subject. Such pharmaceutically acceptable carriers include sterile biocompatible pharmaceutical carriers, including, but not limited to, saline, buffered saline, artificial cerebral spinal fluid, dextrose, and water. By pharmaceutically acceptable is meant a material that is not biologically or otherwise undesirable, which can be administered to an individual along with the selected agent without causing unacceptable biological effects or interacting in a deleterious manner with the other components of the pharmaceutical composition in which it is contained.

Embodiments

[0151] 1. A mammalian expression plasmid comprising a eukaryotic promoter operably linked to a non-viral nucleic acid sequence, wherein the non-viral nucleic acid sequence comprises: (i) a nucleic acid sequence encoding a fusion protein, wherein the fusion protein comprises: (a) a polypeptide comprises a DNA polymerase domain; and (b) a CRISPR-associated endonuclease coding sequence; and (ii) a guide RNA (gRNA) coding sequence, wherein the gRNA coding sequence comprises at least one aptamer coding sequence.

[0152] 2. The mammalian expression plasmid of embodiment 1, wherein the CRISPR-associated endonuclease coding sequence encodes a Cas9 protein.

[0153] 3. The mammalian expression plasmid of embodiment 1 or 2, wherein the polypeptide comprising the DNA polymerase domain is an E. coli DNA polymerase I (DNA PolI).

[0154] 4. The mammalian expression plasmid of embodiment 1 or 2, wherein the polypeptide comprising the DNA polymerase domain is a Klenow fragment of DNA PolI.

[0155] 5. The mammalian expression plasmid of any one of embodiments 1-4, wherein the polypeptide comprising the DNA polymerase domain has reduced polymerase activity.

[0156] 6. The mammalian expression plasmid of any one of embodiments 1-5, wherein the at least one aptamer coding sequence encodes an aptamer sequence bound specifically by an ABP selected from the group consisting of MS2 coat protein, PP7 coat protein, lambda N RNA-binding domain, or Com protein.

[0157] 7. The mammalian expression plasmid of any one of embodiments 1-6, wherein the aptamer is an MS2 aptamer sequence or a com aptamer sequence.

[0158] 8. The mammalian expression plasmid of any one of embodiments 1-7, wherein the sgRNA coding sequence comprises at least one aptamer inserted into the tetraloop or the ST2 loop of the sgRNA coding sequence.

[0159] 9. The mammalian expression plasmid of embodiment 8, wherein the the sgRNA coding comprises at least one com aptamer inserted into the ST2 loop of the gRNA coding sequence.

[0160] 10. A lentiviral packaging system comprising: a) a packaging plasmid comprising a eukaryotic promoter operably linked to a Gag nucleotide sequence, wherein the Gag nucleotide sequence comprises a nucleocapsid (NC) coding sequence and a matrix protein (MA) coding sequence, wherein one or both of the NC coding sequence or the MA coding sequence comprises at least one non-viral aptamer-binding protein (ABP) nucleotide sequence, and wherein the packaging plasmid does not encode a functional integrase protein; b) at least one mammalian expression plasmid of any one of embodiments 1-9; and c) an envelope plasmid comprising an envelope glycoprotein coding sequence.

[0161] 11. The lentiviral packaging system of embodiment 10, wherein the packaging plasmid further comprises a Rev nucleotide sequence and a Tat nucleotide sequence.

[0162] 12. The lentiviral packaging system of embodiment 10 or 0, further comprising a second packaging plasmid comprising a Rev nucleotide sequence.

[0163] 13. The lentiviral packaging system of any one of embodiments 10-12, wherein the at least one non-viral ABP nucleotide sequence encodes MS2 coat protein, PP7 coat protein, lambda N peptide, or Com protein.

[0164] 14. A lentiviral particle comprising: A) a fusion protein comprising a nucleocapsid (NC) protein or a matrix (MA) protein wherein the NC protein or MA protein comprises at least one non-viral aptamer binding protein (ABP); and B) a ribonucleotide protein (RNP) complex comprising: (i) a nucleic acid sequence encoding a fusion protein, wherein the fusion protein comprises: (a) a polypeptide comprising a DNA polymerase domain; and (b) a CRISPR-associated endonuclease coding sequence; and (ii) a guide RNA (gRNA) coding sequence, wherein the gRNA coding sequence comprises at least one aptamer coding sequence wherein the lentivirus-like particle does not comprise a functional integrase protein.

[0165] 15. The lentiviral particle of embodiment 14, wherein the wherein the CRISPR-associated endonuclease coding sequence encodes a Cas9 protein.

[0166] 16. The lentiviral particle of embodiment 14 or 15, wherein the the polypeptide comprising a DNA polymerase domain is an *E. coli* DNA polymerase I (DNA PolI).

[0167] 17. The lentiviral particle of embodiment 14 or 15, wherein the polypeptide comprising a DNA polymerase domain is a Klenow fragment of DNA PolI).

[0168] 18. The lentiviral particle of any one of embodiments 14-17, wherein the polypeptide comprising a DNA polymerase domain has reduced polymerase activity.

[0169] 19. A method of producing a lentiviral particle, the method comprising: a) transfecting a plurality of eukaryotic cells with the packaging plasmid, the at least one mammalian expression plasmid, and the envelope plasmid of the system of any one of embodiments 10-13; and b) culturing the transfected eukaryotic cells for sufficient time for lentiviral particles to be produced.

[0170] 20. The method of embodiment 19, wherein the lentiviral particle comprises a ribonucleotide protein (RNP) complex comprising: (i) nucleic acid sequence encoding a fusion protein, wherein the fusion protein comprises: (a) a polypeptide comprising a DNA polymerase domain; and (b) a CRISPR-associated endonuclease coding sequence; and (ii) a guide RNA.

[0171] 21. The method of embodiment 20, wherein the plurality of eukaryotic cells are mammalian cells.

[0172] 22. A lentiviral particle made by the method of any one of embodiments 19-21.

[0173] 23. A method of modifying a genomic target sequence in a cell, the method comprising transducing a plurality of eukaryotic cells with a plurality of viral particles, wherein the plurality of viral particles comprise a lentivirus-like particle according to any one of embodiments 14-18, wherein the RNP complex binds to the genomic target sequence in genomic DNA of the cell, wherein the CRISPR-associated endonuclease cleaves the genomic target sequence to create a double-stranded break, thereby modifying the genomic target sequence.

[0174] 24. The method of embodiment 23, wherein non-homologous end joining (NHEJ) increases in the cell, as compared to a cell modified by a CRISPR-associated endonuclease that is not fused to a DNA polymerase domain.

[0175] 25. The method of embodiment 23, wherein the ratio of NHEJ to non-NHEJ increases in the cell.

[0176] 26. The method of embodiment 25, wherein the non-NHEJ end joining is microhomology-mediated end joining (MMEJ) and/or single-stranded annealing (SSA).

[0177] 27. The method of embodiments 23-26, wherein the number of on-target deletions is decreased in the cell.

[0178] 28. The method of embodiment 27, wherein the number of on-target deletions greater than 500 base pairs in size is decreased in the cell.

[0179] 29. The method of any one of embodiments 23-27, wherein the ratio of on-target one base pair (1-bp) deletions to on-target deletions greater than 1-bp is increased in the cell.

[0180] 30. The method of embodiment29, wherein the ratio of on-target one base pair (1-bp) deletions to deletions greater than 500 base pairs is increased in the cell.

[0181] 31. The method of any one of embodiments 23-30, wherein the number of templated insertions (TIS) increases in the cell.

[0182] 32. The method of embodiment 31, wherein the ratio of TIS to non-TIS increases in the cell.

[0183] 34. The method of any of embodiments 23-32, wherein the plurality of eukaryotic cells are mammalian cells.

[0184] 35. The method of any of embodiments 23-33, wherein the plurality of eukaryotic cells are cells present in subject.

[0185] 36. The method of embodiment 34, wherein the subject is a human subject.

[0186] 37. The method of embodiment 35, wherein the subject is injected with the plurality of viral particles.

[0187] 38. A cell containing the plasmid of any one of embodiments 1-9.

[0188] 39. A cell containing the lentiviral packaging system of any one of embodiments 10-13.

[0189] 40. A cell containing the lentiviral particle of any one of embodiments 14-16.

[0190] 41. A cell modified using the method of any one of embodiments 23-36.

[0191] 42. A method for treating a disease in a subject comprising: a) obtaining cells from the subject; b) modifying the cells of the subject using the method of any one of embodiments 23-36; and c) administering the modified cells to the subject.

[0192] 43. The method of embodiment 41, wherein the disease is cancer.

[0193] 44. The method of embodiment 42, wherein the disease is Duchenne muscular dystrophy.

[0194] 45. The method of any one of embodiments 41-43, wherein the cells are T cells.

[0195] All patents, patent publications, patent applications, journal articles, books, technical references, and the like discussed in the instant disclosure are incorporated herein by reference in their entirety for all purposes.

[0196] It is to be understood that the figures and descriptions of the disclosure have been simplified to illustrate elements that are relevant for a clear understanding of the disclosure. It should be appreciated that the figures are presented for illustrative purposes and not as construction drawings. Omitted details and modifications or alternative embodiments are within the purview of persons of ordinary skill in the art.

[0197] It can be appreciated that, in certain aspects of the disclosure, a single component may be replaced by multiple components, and multiple components may be replaced by a single component, to provide an element or structure or to perform a given function or functions. Except where such substitution would not be operative to practice certain embodiments of the disclosure, such substitution is considered within the scope of the disclosure.

[0198] The examples presented herein are intended to illustrate potential and specific implementations of the disclosure. It can be appreciated that the examples are intended primarily for purposes of illustration of the disclosure for those skilled in the art. There may be variations to these diagrams or the operations described herein without departing from the spirit of the disclosure. For instance, in certain cases, method steps or operations may be performed or executed in differing order, or operations may be added, deleted or modified.

[0199] Where a range of values is provided, it is understood that each intervening value, to the smallest fraction of the unit of the lower limit, unless the context clearly dictates otherwise, between the upper and lower limits of that range is also specifically disclosed. Any narrower range between any stated values or unstated intervening values in a stated range and any other stated or intervening value in that stated range is encompassed. The upper and lower limits of those smaller ranges may independently be included or excluded in the range, and each range where either, neither, or both limits are included in the smaller ranges is also encompassed within the technology, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included.

[0200] Different arrangements of the components depicted in the drawings or described above, as well as components and steps not shown or described are possible. Similarly, some features and sub-combinations are useful and may be employed without reference to other features and sub-combinations. Embodiments of the disclosure have been described for

illustrative and not restrictive purposes, and alternative embodiments will become apparent to readers of this patent. Accordingly, the present disclosure is not limited to the embodiments described above or depicted in the drawings, and various embodiments and modifications can be made without departing from the scope of the claims below.

EXAMPLES

[0201] *Escherichia. coli* DNA polymerase I (pol I) functions in DNA repair and in replication of the lagging-strand chromosomal DNA. Its small N-terminal domain contains the 5'-3' exonuclease activity, and the large C-terminal Klenow fragment, which can be generated by partial digestion of the full length polymerase carries the polymerase and 3'-5' exonuclease activities Since MMEJ and SSA can generate large DNA deletions and both rely on DNA resection, it was hypothesized that counteracting DNA resection with DNA polymerase I could favor canonic NHEJ over MMEJ and SSA, and decrease the generation of large deletions. Therefore, *E. coli* pol I was targeted to DSMs to counteract DNA resection, and decrease the chances of generating large deletions (Figure 1A). Studies were also conducted to determine if 5' overhangs generated by CRISPR/Cas9 were filled in and the proportion of TISs (Figure 1B) was increased, further refining the INDELs.

A. Materials and Methods

[0202] **Constructs.** Constructs used in this study are described in Table 2. The envelope plasmid pMD2.G for lentiviral pseudotyping was purchased from Addgene (Addgene 12259, Watertown, MA). Some of the plasmids used for this study are available from Addgene (plasmids ID: 176234, 76235, 176236, 176237, 176238). Others are available from authors upon request. Sequences for primers used in this study are listed in Table 3. Single guide RNA (sgRNA) sequences are listed in Table 4.

Table 2. Constructs

No.	Name	Purpose	Generation strategy
1	pspCas9-POLI-Tetra-com vector	Vector plasmid for expressing spCas9-POLI fusion protein and sgRNA. There are two copies of HBB 3' UTR in the 3'UTR of spCas9-POLI to enhance expression. The Tetraloop of the sgRNA scaffold was replaced by com-aptamer to enable Cas9 RNP encapsulation into viral capsids via com-COM interaction.	The DNA sequence of E. coli DNA POLI (encoded by POLA) was optimized for human and mouse expression as "GGATCCTCTGAGACACCAGGCACCTCCGAGTCTGCC ACCCCTGAGGGAGGCAGCGGAGGCAGCGGCTCCGTG CAGATCCCACAGAACCCCCTGATCCTGGTGGACGGC AGCTCCTACCTGTATCGGGCCTACCACGCCTTCCCAC CTCTGACAAACTCCGCCGGAGAGCCAACCGGAGCCA TGTATGGCGTGCTGAATATGCTGAGGAGCCTGATCAT GCAGTACAAGCCTACACACGCCGCCGTGGTGTTTGAT GCCAAGGGCAAGACCTTCCGCGACGAGCTGTTTGAG CACTACAAGAGCCACAGGCCACCAATGCCTGACGAT CTGAGGGCACAGATCGAGCCACTGCACGCAATGGTG AAGGCCATGGGCCTGCCTCTGCTGGCCGTGAGCGGA GTGGAGGCCGACGATGTGATCGGCACACTGGCAAGG

No.	Name	Purpose	Generation strategy
			GAGGCAGAGAAGGCAGGCCGCCCAGTGCTGATCTCC ACCGGCGACAAGGATATGGCCCAGCTGGTGACACCA AACATCACCCTGATCAACACCATGACAAATACCATCC TGGGCCCCGAGGAGGTGGTGAATAAGTATGGCGTGC CTCCAGAGCTGATCATCGATTTCCTGGCCCTGATGGG CGACTCTAGCGATAACATCCCTGGAGTGCCAGGAGT GGGAGAAAAGACCGCACAGGCCCTGCTGCAGGGCCT GGGAGGCCTGGACACCCTGTACGCCGAGCCAGAGAA GATCGCCGGCCTGTCCTTTAGGGGCGCCAAGACAATG GCCGCCAAGCTGGAGCAGAATAAGGAGGTGGCCTAC CTGTCTTATCAGCTGGCCACAATCAAGACCGACGTGG AGCTGGAGCTGACCTGCGAGCAGCTGGAGGTGCAGC AGCCTGCAGCAGAGGAGCTGCTGGGCCTGTTCAAGA AGTACGAGTTTAAGAGATGGACAGCCGATGTGGAGG CCGGCAAGTGGCTGCAGGCAAAGGGAGCAAAGCCAG CAGCAAAGCCACAGGAGACAAGCGTGGCAGACGAG GCACCAGAGGTGACAGCCACCGTGATCTCCTACGAT AACTATGTGACAATCCTGGACGAGGAGACACTGAAG GCCTGGATCGCCAAGCTGGAGAAGGCCCCCGTGTTC GCCTTTGATACAGAGACAGACAGCCTGGATAACATCT CCGCCAATCTGGTGGGCCTGTCTTTCGCCATCGAGCC TGGCGTGGCCGCCTATATCCCAGTGGCCCACGACTAC CTGGATGCCCCCGACCAGATCAGCAGGGAGAGAGCC CTGGAGCTGCTGAAGCCTCTGCTGGAGGATGAGAAG GCCCTGAAGGTCGGCCAGAACCTGAAGTATGACAGG GGCATCCTGGCCAATTACGGCATCGAGCTGAGAGGC ATCGCCTTTGACACCATGCTGGAGTCTTATATCCTGA ATAGCGTGGCAGGCCGGCACGACATGGATTCCCTGG CCGAGAGGTGGCTGAAGCACAAGACAATCACCTTCG AGGAGATCGCCGGCAAGGGCAAGAACCAGCTGACCT TCAACCAGATCGCCCTGGAGGAGGCAGGCAGGTACG CAGCAGAGGACGCAGATGTGACCCTGCAGCTGCACC TGAAGATGTGGCCAGATCTGCAGAAGCACAAGGGCC CCCTGAACGTGTTCGAGAATATCGAGATGCCCCTGGT GCCTGTGCTGAGCCGGATCGAGCGCAACGGCGTGAA GATCGACCCTAAGGTGCTGCACAATCACTCCGAGGA GCTGACCCTGAGACTGGCCGAGCTGGAGAAGAAGGC CCACGAGATCGCCGGCGAGGAGTTCAACCTGTCCTCT ACAAAGCAGCTGCAGACCATCCTGTTTGAGAAGCAG GGCATCAAGCCCCTGAAGAAAACCCCTGGAGGAGCA CCATCTACCAGCGAGGAGGTGCTGGAGGAGCTGGCC CTGGATTATCCCCTGCCTAAAGTGATCCTGGAGTACC GGGGCCTGGCCAAGCTGAAGTCTACATATACCGACA AGCTGCCCCTGATGATCAACCCTAAGACAGGAAGGG TGCACACCAGCTACCACCAGGCAGTGACAGCAACCG GCCGCCTGAGCTCCACCGATCCAAACCTGCAGAATAT CCCCGTGAGGAATGAGGAGGGCAGGAGAATCAGACA GGCCTTCATCGCCCCCGAGGATTATGTGATCGTGTCC GCCGACTACTCTCAGATCGAGCTGAGGATCATGGCCC ACCTGTCCAGAGATAAGGGCCTGCTGACAGCCTTCGC CGAGGGCAAGGACATCCACAGGGCAACCGCAGCAGA GGTGTTTGGCCTGCCTCTGGAGACAGTGACCTCCGAG CAGCGGCGCTCTGCCAAGGCCATCAACTTCGGCCTGA TCTATGGCATGTCTGCCTTTGGCCTGGCCAGGCAGCT GAATATCCCTAGAAAGGAGGCCCAGAAGTACATGGA CCTGTATTTCGAGAGGTACCCAGGCGTGCTGGAGTAC ATGGAGAGGACAAGGGCACAGGCAAAGGAGCAGGG CTATGTGGAGACACTGGATGGCAGGAGACTGTACCT GCCAGACATCAAGTCTAGCAACGGAGCAAGGAGGGC

No.	Name	Purpose	Generation strategy
			AGCAGCAGAGAGGGCCGCCATCAATGCCCCCATGCA GGGCACAGCCGCCGATATCATCAAGAGAGCCATGAT CGCAGTGGACGCATGGCTGCAGGCAGAGCAGCCAAG GGTGAGAATGATCATGCAGGTGCACGATGAGCTGGT GTTTGAGGTGCACAAGGACGATGTGGACGCCGTGGC CAAGCAGATCCACCAGCTGATGGAGAACTGTACCAG GCTGGATGTGCCACTGCTGGTGGAAGTGGGCAGCGG AGAGAATTGGGACCAGGCCCACTGAGGATCC" (SEQ ID NO: 36) and was inserted into BamH1 site of pspCas9-Tetra-com vector. The polypeptide sequence is (SEQ ID NO: 37): *MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVP AADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDR HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYL QEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGV DAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGL TPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQ DLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEE FYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGP LARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIER MTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGM RKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF DSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILE DIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHD DSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQT VKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKY DENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHA HDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKS EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNG ETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSK ESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK VEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEV KKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQ ISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTL TNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDKRPAATKKAGQAKK*KKGSSETPGTSESAT PEGGSGGSGS**VQIPQNPLILVDGSSYLYRAYHAFPPLT NSAGEPTGAMYGVLNMLRSLIMQYKPTHAAVVFDA KGKTFRDELFEHYKSHRPPMPDDLRAQIEPLHAMV KAMGLPLLAVSGVEADDVIGTLAREAEKAGRPVLIS TGDKDMAQLVTPNITLINTMTNTILGPEEVVNKYGV PPELIIDFLALMGDSSDNIPGVPGVGEKTAQALLQGL GGLDTLYAEPEKIAGLSFRGAKTMAAKLEQNKEVA YLSYQLATIKTDVELELTCEQLEVQQPAAEELLGLF KKYEFKRWTADVEAGKWLQAKGAKPAAKPQETSV ADEAPEVTATVISYDNYVTILDEETLKAWIAKLEKAP VFAFDTETDSLDNISANLVGLSFAIEPGVAAYIPVAHD YLDAPDQISRERALELLKPLLEDEKALKVGQNLKYD**

No.	Name	Purpose	Generation strategy
			RGILANYGIELRGIAFDTMLESYILNSVAGRHDMDSL AERWLKHKTITFEEIAGKGKNQLTFNQIALEEAGRY AAEDADVTLQLHLKMWPDLQKHKGPLNVFENIEMP LVPVLSRIERNGVKIDPKVLHNHSEELTLRLAELEKK AHEIAGEEFNLSSTKQLQTILFEKQGIKPLKKTPGGA PSTSEEVLEELALDYPLPKVILEYRGLAKLKSTYTDK LPLMINPKTGRVHTSYHQAVTATGRLSSTDPNLQNIP VRNEEGRRIRQAFIAPEDYVIVSADYSQIELRIMAHLS RDKGLLTAFAEGKDIHRATAAEVFGLPLETVTSEQR RSAKAINFGLIYGMSAFGLARQLNIPRKEAQKYMDL YFERYPGVLEYMERTRAQAKEQGYVETLDGRRLYL PDIKSSNGARRAAAERAAINAPMQGTAADIIKRAMIA VDAWLQAEQPRVRMIMQVHDELVFEVHKDDVDAV AKQIHQLMENCTRLDVPLLVEVGSGENWDQAH (Key: *Cas9* (SEQ ID NO: 29)-linker (SEQ ID NO: 32)-**Polymerase (SEQ ID NO: 27)**
2	pspCas9-poli-Tetra-com-clcn5-sp-g1	Vector plasmid for expressing spCas9-polI fusion protein and sgRNA targeting human CLCN5 gene.	The annealed product of hCLCN5-sp-g1-F (ACCGGAGGACAAGTCGTACAATGG) (SEQ ID NO: 38) and hCLCN5-sp-g1-R (AAACCCATTGTACGACTTGTCCTC) (SEQ ID NO: 39) was inserted between the BbsI sites of pspCas9-POLI-Tetra-com vector by T4 DNA ligase.
3	pspCas9-Klenow-Tetra-com-CLCN5-sp-g1	Plasmid for expressing spCas9-Klenow fusion protein and sgRNA targeting human CLCN5 gene.	Use pspCas9-POLI-tetra-com-hCLCN5-sp-g1 as the template and Klenow-F (GTCTGCCACCCCTGAGGGAGGAGCACCAGAGGTGAC AGCCAC) (SEQ ID NO: 40) and Klenow-R (TTGTGCAGCACCTTAGGGTCG) (SEQ ID NO: 41) as the primers to amplify a 735 bp fragment. This DNA was inserted between the two Bsu36I sites of pspCas9-POLI-tetra-com-hCLCN5-sp-g1 by Infusion reaction. The end result is the removal of the DNA coding for the 5' exonuclease activity of E. coli POLI. Nucleic acid sequence (SEQ ID NO: 42): ATGGACTATAAGGACCACGACGGAGACTACAAGGAT CATGATATTGATTACAAAGACGATGACGATAAGATG GCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGA GTCCCAGCAGCCGACAAGAAGTACAGCATCGGCCTG GACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCA CCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGG TGCTGGGCAACACCGACCGGCACAGCATCAAGAAGA ACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAA CAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAA GAAGATACACCAGACGGAAGAACCGGATCTGCTATC TGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGG ACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCT GGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCAT CTTCGGCAACATCGTGGACGAGGTGGCCTACCACGA GAAGTACCCCACCATCTACCACCTGAGAAAGAAACT GGTGGACAGCACCGACAAGGCCGACCTGCGGCTGAT CTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGC CACTTCCTGATCGAGGGCGACCTGAACCCCGACAAC AGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAG ACCTACAACCAGCTGTTCGAGGAAAACCCCATCAAC GCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCC AGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATC GCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTC

No.	Name	Purpose	Generation strategy
			GGAAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCA ACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCA AACTGCAGCTGAGCAAGGACACCTACGACGACGACC TGGACAACCTGCTGGCCCAGATCGGCGACCAGTACG CCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGC CATCCTGCTGAGCGACATCCTGAGAGTGAACACCGA GATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAG AGATACGACGAGCACCACCAGGACCTGACCCTGCTG AAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTAC AAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTAC GCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAG TTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGG ACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAG AGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACG GCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGC ACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATT CCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCT GACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCC AGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAG AGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAA GTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCG AGCGGATGACCAACTTCGATAAGAACCTGCCCAACG AGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGT ACTTCACCGTGTATAACGAGCTGACCAAAGTGAAAT ACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGA GCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGT TCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGA AAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACT CCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACG CCTCCCTGGGCACATACCACGATCTGCTGAAAATTAT CAAGGACAAGGACTTCCTGGACAATGAGGAAAACGA GGACATTCTGGAAGATATCGTGCTGACCCTGACACTG TTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAA ACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGC AGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGC TGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGC AGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGA CGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCAC GACGACAGCCTGACCTTTAAAGAGGACATCCAGAAA GCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAG CACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGA AGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGC TCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACA TCGTGATCGAAATGGCCAGAGAGAACCAGACCACCC AGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAG CGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAG ATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTG CAGAACGAGAAGCTGTACCTGTACTACCTGCAGAAT GGGCGGGATATGTACGTGGACCAGGAACTGGACATC AACCGGCTGTCCGACTACGATGTGGACCATATCGTGC CTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAA GGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAG CGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGAT GAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCT GATTACCCAGAGAAAGTTCGACAATCTGACCAAGGC CGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGG CTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGAT CACAAAGCACGTGGCACAGATCCTGGACTCCCGGAT GAACACTAAGTACGACGAGAATGACAAGCTGATCCG

No.	Name	Purpose	Generation strategy
			GGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGGT GTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTG CGCGAGATCAACAACTACCACCACGCCCACGACGCC TACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAA AAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGC GACTACAAGGTGTACGACGTGCGGAAGATGATCGCC AAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAG TACTTCTTCTACAGCAACATCATGAACTTTTTCAAGA CCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGC GGCCTCTGATCGAGACAAACGGCGAAACCGGGGAGA TCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCG GAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAA AAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGA GTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGAT CGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGG CGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTG GTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAA CTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATC ATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGAC TTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAG GACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCG AGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTG CCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGC CCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCA CTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGA GCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTA CCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCC AAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAA GTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCC ATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTA CCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTA CTTTGACACCACCATCGACCGGAAGAGGTACACCAG CACCAAAGAGGTGCTGGACGCCACCCTGATCCACCA GAGCATCACCGGCCTGTACGAGACACGGATCGACCT GTCTCAGCTGGGAGGCGACAAAAGGCCGGCGGCCAC GAAAAAGGCCGGccaggcaaaaaagaaaaaggGATCCTCTGAG ACACCAGGCACCTCCGAGTCTGCCACCCCTGAGGGA GGAGCACCAGAGGTGACAGCCACCGTGATCTCCTAC GATAACTATGTGACAATCCTGGACGAGGAGACACTG AAGGCCTGGATCGCCAAGCTGGAGAAGGCCCCCGTG TTCGCCTTTGATACAGAGACAGACAGCCTGGATAACA TCTCCGCCAATCTGGTGGGCCTGTCTTTCGCCATCGA GCCTGGCGTGGCCGCCTATATCCCAGTGGCCCACGAC TACCTGGATGCCCCCGACCAGATCAGCAGGGAGAGA GCCCTGGAGCTGCTGAAGCCTCTGCTGGAGGATGAG AAGGCCCTGAAGGTCGGCCAGAACCTGAAGTATGAC AGGGGCATCCTGGCCAATTACGGCATCGAGCTGAGA GGCATCGCCTTTGACACCATGCTGGAGTCTTATATCC TGAATAGCGTGGCAGGCCGGCACGACATGGATTCCC TGGCCGAGAGGTGGCTGAAGCACAAGACAATCACCT TCGAGGAGATCGCCGGCAAGGGCAAGAACCAGCTGA CCTTCAACCAGATCGCCCTGGAGGAGGCAGGCAGGT ACGCAGCAGAGGACGCAGATGTGACCCTGCAGCTGC ACCTGAAGATGTGGCCAGATCTGCAGAAGCACAAGG GCCCCCTGAACGTCTTTGAGAATATCGAGATGCCCCT GGTGCCTGTGCTGAGCCGGATCGAGCGCAACGGCGT GAAGATCGACCCTAAGGTGCTGCACAATCACTCCGA GGAGCTGACCCTGAGACTGGCCGAGCTGGAGAAGAA GGCCCACGAGATCGCCGGCGAGGAGTTCAACCTGTC

No.	Name	Purpose	Generation strategy
			CTCTACAAAGCAGCTGCAGACCATCCTGTTTGAGAAG CAGGGCATCAAGCCCCTGAAGAAAACCCCTGGAGGA GCACCATCTACCAGCGAGGAGGTGCTGGAGGAGCTG GCCCTGGATTATCCCCTGCCTAAAGTGATCCTGGAGT ACCGGGGCCTGGCCAAGCTGAAGTCTACATATACCG ACAAGCTGCCCCTGATGATCAACCCTAAGACAGGCC GGGTGCACACCAGCTACCACCAGGCAGTGACAGCAA CCGGCCGCCTGAGCTCCACCGATCCAAACCTGCAGA ATATCCCCGTGAGGAATGAGGAGGGCAGGAGAATCA GACAGGCCTTCATCGCCCCCGAGGATTATGTGATCGT GTCCGCCGACTACTCTCAGATCGAGCTGAGGATCATG GCCCACCTGTCCAGAGATAAGGGCCTGCTGACAGCCT TCGCCGAGGGCAAGGACATCCACAGGGCAACCGCAG CAGAGGTGTTTGGCCTGCCTCTGGAGACAGTGACCTC CGAGCAGCGGCGCTCTGCCAAGGCCATCAACTTCGG CCTGATCTATGGCATGTCTGCCTTTGGCCTGGCCAGG CAGCTGAATATCCCTAGAAAGGAGGCCCAGAAGTAC ATGGACCTGTATTTCGAGAGATACCCCGGCGTGCTGG AGTACATGGAGAGGACACGCGCACAGGCAAAGGAGC AGGGCTATGTGGAGACACTGGATGGCAGGAGACTGT ACCTGCCAGACATCAAGTCTAGCAACGGAGCAAGGA GGGCAGCAGCAGAGAGGGCCGCCATCAATGCCCCCA TGCAGGGCACAGCCGCCGATATCATCAAGAGAGCCA TGATCGCAGTGGACGCCTGGCTGCAGGCAGAGCAGC CAAGGGTGAGAATGATCATGCAGGTGCACGATGAGC TGGTGTTTGAGGTGCACAAGGACGATGTGGACGCCG TGGCCAAGCAGATCCACCAGCTGATGGAGAACTGTA CCCGCCTGGATGTGCCACTGCTGGTGGAAGTGGGAA GCGGAGAGAATTGGGACCAGGCCCACTGAG. The polypeptide sequence is SEQ ID NO: 43: *MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVP AADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDR HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYL QEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGV DAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGL TPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQ DLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEE FYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGP LARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIER MTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGM RKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF DSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILE DIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHD DSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQT VKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKY DENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHA HDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKS EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNG ETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSK*

No.	Name	Purpose	Generation strategy
			ESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK *VEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEV* *KKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK* *YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQ* *ISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTL* *TNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE* *TRIDLSQLGGDKRPAATKKAGQAKKKK*GSSETPGTSESAT PEGG**APEVTATVISYDNYVTILDEETLKAWIAKLEKA** **PVFAFDTETDSLDNISANLVGLSFAIEPGVAAYIPVAH** **DYLDAPDQISRERALELLKPLLEDEKALKVGQNLKY** **DRGILANYGIELRGIAFDTMLESYILNSVAGRHDMDS** **LAERWLKHKTITFEEIAGKGKNQLTFNQIALEEAGR** **YAAEDADVTLQLHLKMWPDLQKHKGPLNVFENIEM** **PLVPVLSRIERNGVKIDPKVLHNHSEELTLRLAELEK** **KAHEIAGEEFNLSSTKQLQTILFEKQGIKPLKKTPGG** **APSTSEEVLEELALDYPLPKVILEYRGLAKLKSTYTD** **KLPLMINPKTGRVHTSYHQAVTATGRLSSTDPNLQN** **IPVRNEEGRRIRQAFIAPEDYVIVSADYSQIELRIMAH** **LSRDKGLLTAFAEGKDIHRATAAEVFGLPLETVTSE** **QRRSAKAINFGLIYGMSAFGLARQLNIPRKEAQKYM** **DLYFERYPGVLEYMERTRAQAKEQGYVETLDGRRL** **YLPDIKSSNGARRAAAERAAINAPMQGTAADIIKRA** **MIAVDAWLQAEQPRVRMIMQVHDELVFEVHKDDV** **DAVAKQIHQLMENCTRLDVPLLVEVGSGENWDQAH** (Key: *Cas9* (SEQ ID NO: 27)-linker (SEQ ID NO: 32)- **Klenow** (SEQ ID NO: 28)
4	pspCas9-3exo-Tetra-com-CLCN5-sp-g1	Plasmid for expressing a fusion protein of spCas9 and the 3' exonuclease domain of POLI, and sgRNA targeting human CLCN5 gene.	Use pspCas9-exo-tetra-com-hCLCN5-sp-g1 as the template and Klenow-F (GTCTGCCACCCCTGAGGGAGGAGCACCAGAGGTGAC AGCCAC) (SEQ ID NO: 44) and 3exo-R (CtaggaattttctagactaGTGCTTCTGCAGATCTGGCCAC) (SEQ ID NO: 45) to amplify a 650 bp fragment. This fragment was inserted between the Bsu36I and XbaI sites of pspCas9-exo-tetra-com-hCLCN5-sp-g1 Infusion reaction. Nucleic acid sequence (SEQ ID NO: 46): ATGGACTATAAGGACCACGACGGAGACTACAAGGAT CATGATATTGATTACAAAGACGATGACGATAAGATG GCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGA GTCCCAGCAGCCGACAAGAAGTACAGCATCGGCCTG GACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCA CCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGG TGCTGGGCAACACCGACCGGCACAGCATCAAGAAGA ACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAA CAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAA GAAGATACACCAGACGGAAGAACCGGATCTGCTATC TGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGG ACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCT GGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCAT CTTCGGCAACATCGTGGACGAGGTGGCCTACCACGA GAAGTACCCCACCATCTACCACCTGAGAAAGAAACT GGTGGACAGCACCGACAAGGCCGACCTGCGGCTGAT CTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGC CACTTCCTGATCGAGGGCGACCTGAACCCCGACAAC AGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAG ACCTACAACCAGCTGTTCGAGGAAAACCCCATCAAC GCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCC AGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATC GCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTC GGAAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCA

No.	Name	Purpose	Generation strategy
			ACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCA AACTGCAGCTGAGCAAGGACACCTACGACGACGACC TGGACAACCTGCTGGCCCAGATCGGCGACCAGTACG CCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGC CATCCTGCTGAGCGACATCCTGAGAGTGAACACCGA GATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAG AGATACGACGAGCACCACCAGGACCTGACCCTGCTG AAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTAC AAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTAC GCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAG TTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGG ACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAG AGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACG GCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGC ACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATT CCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCT GACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCC AGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAG AGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAA GTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCG AGCGGATGACCAACTTCGATAAGAACCTGCCCAACG AGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGT ACTTCACCGTGTATAACGAGCTGACCAAAGTGAAAT ACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGA GCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGT TCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGA AAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACT CCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACG CCTCCCTGGGCACATACCACGATCTGCTGAAAATTAT CAAGGACAAGGACTTCCTGGACAATGAGGAAAACGA GGACATTCTGGAAGATATCGTGCTGACCCTGACACTG TTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAA ACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGC AGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGC TGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGC AGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGA CGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCAC GACGACAGCCTGACCTTTAAAGAGGACATCCAGAAA GCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAG CACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGA AGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGC TCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACA TCGTGATCGAAATGGCCAGAGAGAACCAGACCACCC AGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAG CGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAG ATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTG CAGAACGAGAAGCTGTACCTGTACTACCTGCAGAAT GGGCGGGATATGTACGTGGACCAGGAACTGGACATC AACCGGCTGTCCGACTACGATGTGGACCATATCGTGC CTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAA GGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAG CGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGAT GAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCT GATTACCCAGAGAAAGTTCGACAATCTGACCAAGGC CGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGG CTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGAT CACAAAGCACGTGGCACAGATCCTGGACTCCCGGAT GAACACTAAGTACGACGAGAATGACAAGCTGATCCG GGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGGT

No.	Name	Purpose	Generation strategy
			GTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTG CGCGAGATCAACAACTACCACCACGCCCACGACGCC TACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAA AAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGC GACTACAAGGTGTACGACGTGCGGAAGATGATCGCC AAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAG TACTTCTTCTACAGCAACATCATGAACTTTTTCAAGA CCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGC GGCCTCTGATCGAGACAAACGGCGAAACCGGGGAGA TCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCG GAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAA AAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGA GTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGAT CGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGG CGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTG GTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAA CTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATC ATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGAC TTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAG GACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCG AGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTG CCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGC CCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCA CTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGA GCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTA CCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCC AAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAA GTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCC ATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTA CCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTA CTTTGACACCACCATCGACCGGAAGAGGTACACCAG CACCAAAGAGGTGCTGGACGCCACCCTGATCCACCA GAGCATCACCGGCCTGTACGAGACACGGATCGACCT GTCTCAGCTGGGAGGCGACAAAAGGCCGGCGGCCAC GAAAAAGGCCGGccaggcaaaaaagaaaaaggGATCCTCTGAG ACACCAGGCACCTCCGAGTCTGCCACCCCTGAGGGA GGAGCACCAGAGGTGACAGCCACCGTGATCTCCTAC GATAACTATGTGACAATCCTGGACGAGGAGACACTG AAGGCCTGGATCGCCAAGCTGGAGAAGGCCCCCGTG TTCGCCTTTGATACAGAGACAGACAGCCTGGATAACA TCTCCGCCAATCTGGTGGGCCTGTCTTTCGCCATCGA GCCTGGCGTGGCCGCCTATATCCCAGTGGCCCACGAC TACCTGGATGCCCCCGACCAGATCAGCAGGGAGAGA GCCCTGGAGCTGCTGAAGCCTCTGCTGGAGGATGAG AAGGCCCTGAAGGTCGGCCAGAACCTGAAGTATGAC AGGGGCATCCTGGCCAATTACGGCATCGAGCTGAGA GGCATCGCCTTTGACACCATGCTGGAGTCTTATATCC TGAATAGCGTGGCAGGCCGGCACGACATGGATTCCC TGGCCGAGAGGTGGCTGAAGCACAAGACAATCACCT TCGAGGAGATCGCCGGCAAGGGCAAGAACCAGCTGA CCTTCAACCAGATCGCCCTGGAGGAGGCAGGCAGGT ACGCAGCAGAGGACGCAGATGTGACCCTGCAGCTGC ACCTGAAGATGTGGCCAGATCTGCAGAAGCACtag. The polypeptide is (SEQ ID NO: 47) *MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVP AADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDR HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYL QEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR*

No.	Name	Purpose	Generation strategy
			*GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGV DAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGL TPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQ DLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEE FYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGP LARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIER MTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGM RKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF DSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILE DIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHD DSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQT VKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKY DENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHA HDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKS EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNG ETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSK ESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK VEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEV KKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQ ISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTL TNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDKRPAATKKAGQAKKKK*GSSETPGTSESAT PEGG**APEVTATVISYDNYVTILDEETLKAWIAKLEKA PVFAFDTETDSLDNISANLVGLSFAIEPGVAAYIPVAH DYLDAPDQISRERALELLKPLLEDEKALKVGQNLKY DRGILANYGIELRGIAFDTMLESYILNSVAGRHDMDS LAERWLKHKTITFEEIAGKGKNQLTFNQIALEEAGR YAAEDADVTLQLHLKMWPDLQKH** (Key: *Cas9* (SEQ ID NO: 27)-linker (SEQ ID NO: 32)-**3' exo** (SEQ ID NO: 33)
5	pspCas9-Exo-tetra-com-hCLCN5-sp-g1	Plasmid for expressing a fusion protein of spCas9 and the 5' and 3' exonuclease domain of POLI, and sgRNA targeting human CLCN5 gene.	Cut pspCas9-POLI-tetra-com-hCLCN5-sp-g1 with AccI+NheI to remove the DNA coding for the DNA polymerase activity. Treat the DNA with the Klenow fragment of DNA POLI to make blunt ends. Then use T4 DNA ligase to ligate the ends. Nucleic acid sequence (SEQ ID NO: 48): ATGGACTATAAGGACCACGACGGAGACTACAAGGAT CATGATATTGATTACAAAGACGATGACGATAAGATG GCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGA GTCCCAGCAGCCGACAAGAAGTACAGCATCGGCCTG GACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCA CCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGG TGCTGGGCAACACCGACCGGCACAGCATCAAGAAGA ACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAA CAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAA GAAGATACACCAGACGGAAGAACCGGATCTGCTATC TGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGG ACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCT GGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCAT CTTCGGCAACATCGTGGACGAGGTGGCCTACCACGA GAAGTACCCCACCATCTACCACCTGAGAAAGAAACT GGTGGACAGCACCGACAAGGCCGACCTGCGGCTGAT

No.	Name	Purpose	Generation strategy
			CTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGC CACTTCCTGATCGAGGGCGACCTGAACCCCGACAAC AGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAG ACCTACAACCAGCTGTTCGAGGAAAACCCCATCAAC GCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCC AGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATC GCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTC GGAAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCA ACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCA AACTGCAGCTGAGCAAGGACACCTACGACGACGACC TGGACAACCTGCTGGCCCAGATCGGCGACCAGTACG CCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGC CATCCTGCTGAGCGACATCCTGAGAGTGAACACCGA GATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAG AGATACGACGAGCACCACCAGGACCTGACCCTGCTG AAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTAC AAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTAC GCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAG TTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGG ACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAG AGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACG GCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGC ACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATT CCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCT GACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCC AGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAG AGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAA GTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCG AGCGGATGACCAACTTCGATAAGAACCTGCCCAACG AGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGT ACTTCACCGTGTATAACGAGCTGACCAAAGTGAAAT ACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGA GCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGT TCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGA AAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACT CCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACG CCTCCCTGGGCACATACCACGATCTGCTGAAAATTAT CAAGGACAAGGACTTCCTGGACAATGAGGAAAACGA GGACATTCTGGAAGATATCGTGCTGACCCTGACACTG TTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAA ACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGC AGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGC TGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGC AGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGA CGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCAC GACGACAGCCTGACCTTTAAAGAGGACATCCAGAAA GCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAG CACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGA AGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGC TCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACA TCGTGATCGAAATGGCCAGAGAGAACCAGACCACCC AGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAG CGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAG ATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTG CAGAACGAGAAGCTGTACCTGTACTACCTGCAGAAT GGGCGGGATATGTACGTGGACCAGGAACTGGACATC AACCGGCTGTCCGACTACGATGTGGACCATATCGTGC CTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAA GGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAG

No.	Name	Purpose	Generation strategy
			CGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGAT GAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCT GATTACCCAGAGAAAGTTCGACAATCTGACCAAGGC CGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGG CTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGAT CACAAAGCACGTGGCACAGATCCTGGACTCCCGGAT GAACACTAAGTACGACGAGAATGACAAGCTGATCCG GGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGGT GTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTG CGCGAGATCAACAACTACCACCACGCCCACGACGCC TACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAA AAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGC GACTACAAGGTGTACGACGTGCGGAAGATGATCGCC AAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAG TACTTCTTCTACAGCAACATCATGAACTTTTTCAAGA CCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGC GGCCTCTGATCGAGACAAACGGCGAAACCGGGGAGA TCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCG GAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAA AAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGA GTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGAT CGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGG CGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTG GTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAA CTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATC ATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGAC TTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAG GACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCG AGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTG CCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGC CCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCA CTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGA GCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTA CCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCC AAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAA GTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCC ATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTA CCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTA CTTTGACACCACCATCGACCGGAAGAGGTACACCAG CACCAAAGAGGTGCTGGACGCCACCCTGATCCACCA GAGCATCACCGGCCTGTACGAGACACGGATCGACCT GTCTCAGCTGGGAGGCGACAAAAGGCCGGCGGCCAC GAAAAAGGCCGGccaggcaaaaaagaaaaaggGATCCTCTGAG ACACCAGGCACCTCCGAGTCTGCCACCCCTGAGGGA GGAAGCGGAGGAAGCGGCTCCGTGCAGATCCCACAG AACCCCCTGATCCTGGTGGACGGCAGCTCCTACCTGT ATCGGGCCTACCACGCCTTCCCACCTCTGACAAACTC CGCCGGAGAGCCAACCGGAGCCATGTATGGCGTGCT GAATATGCTGAGGAGCCTGATCATGCAGTACAAGCC TACACACGCCGCCGTGGTGTTTGATGCCAAGGGCAA GACCTTCCGCGACGAGCTGTTTGAGCACTACAAGAGC CACAGGCCACCAATGCCTGACGATCTGAGGGCACAG ATCGAGCCACTGCACGCAATGGTGAAGGCCATGGGC CTGCCTCTGCTGGCCGTGAGCGGAGTGGAGGCCGAC GATGTGATCGGCACACTGGCAAGGGAGGCAGAGAAG GCAGGCCGCCCAGTGCTGATCTCCACCGGCGACAAG GATATGGCCCAGCTGGTGACACCAAACATCACCCTG ATCAACACCATGACAAATACCATCCTGGGCCCCGAG GAGGTGGTGAATAAGTATGGCGTGCCTCCAGAGCTG

No.	Name	Purpose	Generation strategy
			ATCATCGATTTCCTGGCCCTGATGGGCGACTCTAGCG ATAACATCCCTGGCGTGCCAGGAGTGGGAGAAAAGA CCGCACAGGCCCTGCTGCAGGGCCTGGGAGGCCTGG ACACCCTGTACGCCGAGCCAGAGAAGATCGCCGGCC TGTCCTTTAGGGGCGCCAAGACAATGGCCGCCAAGCT GGAGCAGAATAAGGAGGTGGCCTACCTGTCTTATCA GCTGGCCACAATCAAGACCGACGTGGAGCTGGAGCT GACCTGCGAGCAGCTGGAGGTGCAGCAGCCTGCAGC AGAGGAGCTGCTGGGCCTGTTCAAGAAGTACGAGTT TAAGAGATGGACAGCCGATGTGGAGGCCGGCAAGTG GCTGCAGGCAAAGGGAGCAAAGCCAGCAGCAAAGCC ACAGGAGACAAGCGTGGCAGACGAGGCACCAGAGGT GACAGCCACCGTGATCTCCTACGATAACTATGTGACA ATCCTGGACGAGGAGACACTGAAGGCCTGGATCGCC AAGCTGGAGAAGGCCCCCGTGTTCGCCTTTGATACAG AGACAGACAGCCTGGATAACATCTCCGCCAATCTGGT GGGCCTGTCTTTCGCCATCGAGCCTGGCGTGGCCGCC TATATCCCAGTGGCCCACGACTACCTGGATGCCCCCG ACCAGATCAGCAGGGAGAGAGCCCTGGAGCTGCTGA AGCCTCTGCTGGAGGATGAGAAGGCCCTGAAGGTCG GCCAGAACCTGAAGTATGACAGGGGCATCCTGGCCA ATTACGGCATCGAGCTGAGAGGCATCGCCTTTGACAC CATGCTGGAGTCTTATATCCTGAATAGCGTGGCAGGC CGGCACGACATGGATTCCCTGGCCGAGAGGTGGCTG AAGCACAAGACAATCACCTTCGAGGAGATCGCCGGC AAGGGCAAGAACCAGCTGACCTTCAACCAGATCGCC CTGGAGGAGGCAGGCAGGTACGCAGCAGAGGACGCA GATGTGACCCTGCAGCTGCACCTGAAGATGTGGCCA GATCTGCAGAAGCACAAGGGCCCCCTGAACGTCTTTG AGAATATCGAGATGCCCCTGGTGCCTGTGCTGAGCCG GATCGAGCGCAACGGCGTGAAGATCGACCCTAAGGT GCTGCACAATCACTCCGAGGAGCTGACCCTGAGACT GGCCGAGCTGGAGAAGAAGGCCCACGAGATCGCCGG CGAGGAGTTCAACCTGTCCTCTACAAAGCAGCTGCAG ACCATCCTGTTTGAGAAGCAGGGCATCAAGCCCCTGA AGAAAACCCCTGGAGGAGCACCATCTACCAGCGAGG AGGTGCTGGAGGAGCTGGCCCTGGATTATCCCCTGCC TAAAGTGATCCTGGAGTACCGGGGCCTGGCCAAGCT GAAGTCTctagctcgctttcttgctgtccaatttctattaaaggttcctttgttccctaag tccaactactaa. The polypeptide sequence is (SEQ ID NO: 49): *MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVP AADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDR HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYL QEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGV DAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGL TPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQ DLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEE FYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGP LARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIER MTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGM RKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF DSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILE*

No.	Name	Purpose	Generation strategy
			*DIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHD DSLTFKEDIQKAQVSGQGDSLIIEIIIANLAGSPAIKKGILQT VKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKY DENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHA HDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKS EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNG ETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSK ESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK VEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEV KKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQ ISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTL TNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDKRPAATKKAGQAKKKK*GSSETPGTSESAT PEGGSGGSGS**VQIPQNPLILVDGSSYLYRAYHAFPPLT NSAGEPTGAMYGVLNMLRSLIMQYKPTHAAVVFDA KGKTFRDELFEHYKSHRPPMPDDLRAQIEPLHAMV KAMGLPLLAVSGVEADDVIGTLAREAEKAGRPVLIS TGDKDMAQLVTPNITLINTMTNTILGPEEVVNKYGV PPELIIDFLALMGDSSDNIPGVPGVGEKTAQALLQGL GGLDTLYAEPEKIAGLSFRGAKTMAAKLEQNKEVA YLSYQLATIKTDVELELTCEQLEVQQPAAEELLGLF KKYEFKRWTADVEAGKWLQAKGAKPAAKPQETSV ADEAPEVTATVISYDNYVTILDEETLKAWIAKLEKAP VFAFDTETDSLDNISANLVGLSFAIEPGVAAYIPVAHD YLDAPDQISRERALELLKPLLEDEKALKVGQNLKYD RGILANYGIELRGIAFDTMLESYILNSVAGRHDMDSL AERWLKHKTITFEEIAGKGKNQLTFNQIALEEAGRY AAEDADVTLQLHLKMWPDLQKHKGPLNVFENIEMP LVPVLSRIERNGVKIDPKVLHNHSEELTLRLAELEKK AHEIAGEEFNLSSTKQLQTILFEKQGIKPLKKTPGGA PSTSEEVLEELALDYPLPKVILEYRGLAKLKSLARFL AVQFLLKVPLFPKSNY**(Key: *Cas9* (SEQ ID NO: 27)-linker (SEQ ID NO: 32)-**5'-3' exo** (SEQ ID NO: 34)
6	pspCas9-5Exo-tetra-com-hCLCN5-sp-g1	Plasmid for expressing a fusion protein of spCas9 and the 5' exonuclease domain of POLI, and sgRNA targeting human CLCN5 gene.	Cut pspCas9-POLI-tetra-com-hCLCN5-sp-g1 with ClaI+NheI to remove majority of the POLI coding DNA. Then use 5Exo-F (CAGAGCTGATCATCGATTTCCTGG) (SEQ ID NO: 50) and 5Exo-R (AagaaagcgagctagccCACGGTGGCTGTCACCTCTGG) (SEQ ID NO: 51) to amplify a 400 bp band coding for the 5' exonuclease domain from pspCas9-POLI-tetra-com-hCLCN5-sp-g1. This DNA was inserted into the ClaI+NheI linearized vector by infusion reaction. Nucleic acid sequence (SEQ ID NO: 52): ATGGACTATAAGGACCACGACGGAGACTACAAGGAT CATGATATTGATTACAAAGACGATGACGATAAGATG GCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGA GTCCCAGCAGCCGACAAGAAGTACAGCATCGGCCTG GACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCA CCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGG TGCTGGGCAACACCGACCGGCACAGCATCAAGAAGA ACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAA CAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAA

No.	Name	Purpose	Generation strategy
			GAAGATACACCAGACGGAAGAACCGGATCTGCTATC TGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGG ACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCT GGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCAT CTTCGGCAACATCGTGGACGAGGTGGCCTACCACGA GAAGTACCCCACCATCTACCACCTGAGAAAGAAACT GGTGGACAGCACCGACAAGGCCGACCTGCGGCTGAT CTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGC CACTTCCTGATCGAGGGCGACCTGAACCCCGACAAC AGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAG ACCTACAACCAGCTGTTCGAGGAAAACCCCATCAAC GCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCC AGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATC GCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTC GGAAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCA ACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCA AACTGCAGCTGAGCAAGGACACCTACGACGACGACC TGGACAACCTGCTGGCCCAGATCGGCGACCAGTACG CCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGC CATCCTGCTGAGCGACATCCTGAGAGTGAACACCGA GATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAG AGATACGACGAGCACCACCAGGACCTGACCCTGCTG AAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTAC AAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTAC GCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAG TTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGG ACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAG AGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACG GCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGC ACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATT CCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCT GACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCC AGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAG AGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAA GTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCG AGCGGATGACCAACTTCGATAAGAACCTGCCCAACG AGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGT ACTTCACCGTGTATAACGAGCTGACCAAAGTGAAAT ACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGA GCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGT TCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGA AAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACT CCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACG CCTCCCTGGGCACATACCACGATCTGCTGAAAATTAT CAAGGACAAGGACTTCCTGGACAATGAGGAAAACGA GGACATTCTGGAAGATATCGTGCTGACCCTGACACTG TTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAA ACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGC AGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGC TGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGC AGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGA CGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCAC GACGACAGCCTGACCTTTAAAGAGGACATCCAGAAA GCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAG CACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGA AGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGC TCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACA TCGTGATCGAAATGGCCAGAGAGAACCAGACCACCC AGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAG

No.	Name	Purpose	Generation strategy
			CGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAG ATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTG CAGAACGAGAAGCTGTACCTGTACTACCTGCAGAAT GGGCGGGATATGTACGTGGACCAGGAACTGGACATC AACCGGCTGTCCGACTACGATGTGGACCATATCGTGC CTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAA GGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAG CGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGAT GAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCT GATTACCCAGAGAAAGTTCGACAATCTGACCAAGGC CGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGG CTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGAT CACAAAGCACGTGGCACAGATCCTGGACTCCCGGAT GAACACTAAGTACGACGAGAATGACAAGCTGATCCG GGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGGT GTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTG CGCGAGATCAACAACTACCACCACGCCCACGACGCC TACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAA AAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGC GACTACAAGGTGTACGACGTGCGGAAGATGATCGCC AAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAG TACTTCTTCTACAGCAACATCATGAACTTTTTCAAGA CCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGC GGCCTCTGATCGAGACAAACGGCGAAACCGGGGAGA TCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCG GAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAA AAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGA GTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGAT CGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGG CGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTG GTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAA CTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATC ATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGAC TTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAG GACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCG AGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTG CCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGC CCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCA CTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGA GCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTA CCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCC AAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAA GTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCC ATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTA CCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTA CTTTGACACCACCATCGACCGGAAGAGGTACACCAG CACCAAAGAGGTGCTGGACGCCACCCTGATCCACCA GAGCATCACCGGCCTGTACGAGACACGGATCGACCT GTCTCAGCTGGGAGGCGACAAAAGGCCGGCGGCCAC GAAAAAGGCCGGccaggcaaaaaagaaaaaggGATCCTCTGAG ACACCAGGCACCTCCGAGTCTGCCACCCCTGAGGGA GGAAGCGGAGGAAGCGGCTCCGTGCAGATCCCACAG AACCCCCTGATCCTGGTGGACGGCAGCTCCTACCTGT ATCGGGCCTACCACGCCTTCCCACCTCTGACAAACTC CGCCGGAGAGCCAACCGGAGCCATGTATGGCGTGCT GAATATGCTGAGGAGCCTGATCATGCAGTACAAGCC TACACACGCCGCCGTGGTGTTTGATGCCAAGGGCAA GACCTTCCGCGACGAGCTGTTTGAGCACTACAAGAGC CACAGGCCACCAATGCCTGACGATCTGAGGGCACAG

No.	Name	Purpose	Generation strategy
			ATCGAGCCACTGCACGCAATGGTGAAGGCCATGGGC CTGCCTCTGCTGGCCGTGAGCGGAGTGGAGGCCGAC GATGTGATCGGCACACTGGCAAGGGAGGCAGAGAAG GCAGGCCGCCCAGTGCTGATCTCCACCGGCGACAAG GATATGGCCCAGCTGGTGACACCAAACATCACCCTG ATCAACACCATGACAAATACCATCCTGGGCCCCGAG GAGGTGGTGAATAAGTATGGCGTGCCTCCAGAGCTG ATCATCGATTTCCTGGCCCTGATGGGCGACTCTAGCG ATAACATCCCTGGCGTGCCAGGAGTGGGAGAAAAGA CCGCACAGGCCCTGCTGCAGGGCCTGGGAGGCCTGG ACACCCTGTACGCCGAGCCAGAGAAGATCGCCGGCC TGTCCTTTAGGGGCGCCAAGACAATGGCCGCCAAGCT GGAGCAGAATAAGGAGGTGGCCTACCTGTCTTATCA GCTGGCCACAATCAAGACCGACGTGGAGCTGGAGCT GACCTGCGAGCAGCTGGAGGTGCAGCAGCCTGCAGC AGAGGAGCTGCTGGGCCTGTTCAAGAAGTACGAGTT TAAGAGATGGACAGCCGATGTGGAGGCCGGCAAGTG GCTGCAGGCAAAGGGAGCAAAGCCAGCAGCAAAGCC ACAGGAGACAAGCGTGGCAGACGAGGCACCAGAGGT GACAGCCACCGTGggctag. The polypeptide sequence is (SEQ ID NO: 53): *MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVP* *AADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDR* *HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYL* *QEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV* *DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR* *GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGV* *DAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGL* *TPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD* *LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQ* *DLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEE* *FYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH* *QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGP* *LARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIER* *MTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGM* *RKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF* *DSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILE* *DIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG* *WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHD* *DSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQT* *VKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER* *MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD* *MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK* *NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK* *AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKY* *DENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHA* *HDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKS* *EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNG* *ETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSK* *ESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK* *VEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEV* *KKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK* *YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQ* *ISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTL* *TNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE* *TRIDLSQLGGDKRPAATKKAGQAKKKK*GSSETPGTSESAT PEGGSGGSGS**VQIPQNPLILVDGSSYLYRAYHAFPPLT** **NSAGEPTGAMYGVLNMLRSLIMQYKPTHAAVVFDA**

No.	Name	Purpose	Generation strategy
			KGKTFRDELFEHYKSHRPPMPDDLRAQIEPLHAMV KAMGLPLLAVSGVEADDVIGTLAREAEKAGRPVLIS TGDKDMAQLVTPNITLINTMTNTILGPEEVVNKYGV PPELIIDFLALMGDSSDNIPGVPGVGEKTAQALLQGL GGLDTLYAEPEKIAGLSFRGAKTMAAKLEQNKEVA YLSYQLATIKTDVELELTCEQLEVQQPAAEELLGLF KKYEFKRWTADVEAGKWLQAKGAKPAAKPQETSV ADEAPEVTATVG (Key: (Key: *Cas9* (SEQ ID NO: 27)-linker (SEQ ID NO: 32)--**5' exo** (SEQ ID NO: 35))
7	pspCas9-POLI-ST2-com-53-sp-g2	Vector plasmid for expressing spCas9-POLI fusion protein and sgRNA targeting human DMD exon 53.	The AflIII+Acc65I fragment of pspCas9-3'UTR-ST2-com-52-sp-g2 was used to replace that of pspCas9-POLI-tetra-com-hCLCN5-sp-g1 by restriction enzyme digestion and T4 DNA ligase.
8	pspCas9-Klenow-ST2-com-53-sp-g2	Plasmid for expressing spCas9-Klenow fusion protein and sgRNA targeting human DMD exon 53.	The AflIII+Acc65I fragment of pspCas9-3'UTR-ST2-com-52-sp-g2 was used to replace that of pspCas9-Klenow-tetra-com-hCLCN5-sp-g1 by restriction enzyme digestion and T4 DNA ligase.
9	pspCas9-POLI-ST2-com vector	Vector plasmid for expressing spCas9-POLI fusion protein and sgRNA. There are two copies of HBB 3' UTR in the 3'UTR of spCas9-POLI to enhance expression. The ST2 loop of the sgRNA scaffold was replaced by com-aptamer to enable Cas9 RNP encapsulation into viral capsids via com-COM interaction.	The AflIII and Acc65I fragment of pspCas9-3'UTR-ST2-com vector was used to replace the AflIII and Acc65I fragment of pspCas9-POLI-Tetra-com vector by restriction enzyme digestion and T4 DNA ligase.
10	pspCas9-3'UTR-ST2-com-HBB-W	Plasmid for expressing spCas9 and wild type HBB sgRNA. The ST2 loop of the sgRNA scaffold was replaced by com-aptamer to enable Cas9 RNP encapsulation into viral capsids via com-COM interaction.	The annealed HBB-W oligoes were inserted into the BbsI site of pspCas9-3'UTR-ST2-com-vector by T4 DNA ligase.
11	pspCas9-Klenow-ST2-com-vector	Vector plasmid for expressing spCas9-Klenow fusion protein and sgRNA. There are two copies of HBB 3' UTR in the 3'UTR of spCas9-Klenow to enhance expression. The ST2 loop of the sgRNA scaffold was replaced by com-aptamer to enable Cas9 RNP encapsulation into viral capsids via com-Com interaction.	The AflIII-Acc65I fragment containing the U6-sgRNA cassette was cut from pspCas9-poli-ST2-com vector and was inserted into the AflIII-Acc65I Sites of pspCas9-Klenow-ST2-com-53-sp-g2 by T4 DNA ligase.
12	pspCas9-Klenow-	Plasmid for expressing spCas9-Klenow fusion	The annealed HBB oligonucleotides (HBB-sp-WF and HBB-sp-WR) were inserted into the BbsI site of pspCas9-Klenow-ST2-com-vector by T4 DNA ligase.

No.	Name	Purpose	Generation strategy
	ST2-com-HBB-W	protein and sgRNA targeting human HBB.	
13	pspCas9-3'UTR-ST2-com-DMD44	Plasmid for expressing spCas9 and sgRNA targeting human DMD exon 44.	The annealed DMD44 oligonucleotides (DMD44-g1-F and DMD44-g1-R) were inserted into the BbsI site of pspCas9-3'UTR-ST2-com-vector by T4 DNA ligase.
14	pspCas9-Klenow-ST2-com-DMD44	Plasmid for expressing spCas9-Klenow fusion protein and sgRNA targeting human DMD exon 44.	The annealed DMD44 oligonucleotides were inserted into the BbsI site of pspCas9-Klenow-ST2-com-vector by T4 DNA ligase.
15	pspCas9-3'UTR-ST2-com-g5	Plasmid for expressing spCas9 and sgRNA targeting human Intragenic 1 (GRCh38.p13, chromosome 20, 32752960–32752979).	See CRISPR J. 2021 Dec;4(6):914-928. doi: 10.1089/crispr.2020.0106. Epub 2021 Mar 16.
16	pspCas9-Klenow-ST2-com-g5	Plasmid for expressing spCas9-Klenow fusion protein and sgRNA targeting Intragenic 1.	The annealed oligonucleotides ABE-g5-F and ABE-g5-R were inserted into the BbsI site of pspCas9-Klenow-ST2-com-vector by T4 DNA ligase.
17	pspCas9-3'UTR-ST2-com-P53	Plasmid for expressing spCas9 and sgRNA targeting human P53.	See CRISPR J. 2021 Dec;4(6):914-928. doi: 10.1089/crispr.2020.0106. Epub 2021 Mar 16.
18	pspCas9-Klenow-ST2-com-P53	Plasmid for expressing spCas9-Klenow fusion protein and sgRNA targeting human P53.	The annealed P53 oligonucleotides (P53-g1F1 and P53-g1R1) were inserted into the BbsI site of pspCas9-Klenow-ST2-com-vector by T4 DNA ligase.
19	pspCas9-3'UTR-ST2-com-MRE11	Plasmid for expressing spCas9 and sgRNA targeting human MRE11.	The annealed MRE11 oligonucleotides (Mre11-gF and Mre11-gR) were inserted into the BbsI site of pspCas9-3'UTR-ST2-com-vector by T4 DNA ligase.
20	pspCas9-Klenow-ST2-com-MRE11	Plasmid for expressing spCas9-Klenow fusion protein and sgRNA targeting human MRE11.	The annealed MRE11 oligonucleotides were inserted into the BbsI site of pspCas9-Klenow-ST2-com-vector by T4 DNA ligase.
21	pspCas9-3'UTR-ST2-com-CtIP	Plasmid for expressing spCas9 and sgRNA targeting human RBBP8.	The annealed CtIP oligonucleotides (CtIP-gF and CtIP-gR) were inserted into the BbsI site of pspCas9-3'UTR-ST2-com-vector by T4 DNA ligase.
22	pspCas9-Klenow-ST2-com-CtIP	Plasmid for expressing spCas9-Klenow fusion protein and sgRNA targeting human TBBP8.	The annealed CtIP oligonucleotides were inserted into the BbsI site of pspCas9-Klenow-ST2-com-vector by T4 DNA ligase.
23	Ppoli-CLCN5-g1	Plasmid for expressing *E. coli* polymerase I with nuclear targeting signals and sgRNA for *CLCN5*.	pspCas9-poli-Tetra-com-clcn5-sp-g1 was cut with PpuMI to delete the 2.6 kb encoding the Cas9 functional domains, and the rest DNA was self-ligated.
24	pspCas9-poli(D705A)-Tetra-com-clcn5-sp-g1	Plasmid for expressing sgRNA for *CLCN5* and Cas9-pol fusion protein with D705A mutation to inactivate the polymerase activity.	Cut pspCas9-poli-Tetra-com-clcn5-sp-g1 with AccI and NheI, and recover the vector (remove the 1 bk band); Use Pol-F1 (GCCAAGCTGAAGTCTACATATAC) (SEQ ID NO: 54) and Pol-MR (CTGAGAGTAGGCGGCGGACACG) (SEQ ID NO: 55) as primers and use pspCas9-poli-Tetra-com-clcn5-sp-g1 as the template to amplify a fragment of 225 bp (Fragment 1); Use Pol-MF

No.	Name	Purpose	Generation strategy
			(CGTGTCCGCCGCCTACTCTCAG) (SEQ ID NO: 56) and pol-R1 (aagaaagcgagctagcaatgaa) as primers and use pspCas9-poli-Tetra-com-clcn5-sp-g1 as the template to amplify a fragment of 840 bp (Fragment 2). Insert Fragment 1 and 2 into the recovered vector by Infusion reaction.
25	pspCas9-Klenow(D705A)-Tetra-com-clcn5-sp-g1	Plasmid for expressing sgRNA for *CLCN5* and Cas9-Klenow fusion protein with D705A mutation to inactivate the polymerase activity.	The AccI-FseI fragment of pspCas9-poli(D705A)-Tetra-com-clcn5-sp-g1 was replaced by the AccI-FseI fragment of pspCas9-Klenow-Tetra-com-clcn5-sp-g1 to remove the 5' exonuclease coding region by T4 DNA ligase.

Table 3. Primer Sequence Information

Primer name	Sequence	Purpose
Reporter-F	Tccatttcaggtgtcgtgag (SEQ ID NO: 57)	To amplify the DNA in the GFP reporter cassette.
Reporter-R2	TCCAGCTCGACCAGGATG (SEQ ID NO: 58)	
HBB-F1	AGCCAGGGCTGGGCATAAAAG (SEQ ID NO: 59)	Used to amplify the HBB region for NGS with HBB-R3
HBB-R3	TGGGAAAATAGACCAATAGGCAGAG (SEQ ID NO: 60)	Used to amplify the HBB region for NGS with HBB-R or HBB-R
DMD53-F	TCCTGTTGTTCATCATCCTAGC (SEQ ID NO: 61)	To amplify the DMD 53 exon target region for NGS
DMD53-R	TCCAGCCATTGTGTTGAATC (SEQ ID NO: 62)	
hCLCN5-F	GTTTAAGGGCCCGCCTTTTG (SEQ ID NO: 63)	To amplify the CLCN5 target region for NGS
hCLCN5-R	TGTCTTACCTCTCGGTGCCT (SEQ ID NO: 64)	
hCLCN5-F3	GACCCAGGTTTCCTGAGCTG (SEQ ID NO: 65)	To amplify the CLCN5 target region for SMRT sequencing
hCLCN5-R3	TTCAGAGCTTCCTCCCAAGC (SEQ ID NO: 66)	
OT4-F	CCCTATACCTGGGCTCCGTT (SEQ ID NO: 67)	To amplify the Off-target 4 region for NGS
OT4-R	GAAAGGGCCTCTCTCTTTGTAATG (SEQ ID NO: 68)	
OT5-F	AAGCTCTACAAGGGCAGAGAATG (SEQ ID NO: 69)	To amplify the Off-target 8 region for NGS
OT5-R	TCAAAGCTCCCAGATTCACGTT (SEQ ID NO: 70)	
OT8-F	CTGCTCTTTGCCTGTTGGAG (SEQ ID NO: 71)	To amplify the Off-target 8 region for NGS
OT8-R	GCTAAAGCTGGAAGGCTGTG (SEQ ID NO: 72)	
HBD-F2	AAAAGGCAGGGCAGAGTCGACTG (SEQ ID NO: 73)	To amplify HBD region for NGS
HBD-R2	GGTAGGAAAACAGCCCAAGGGAC (SEQ ID NO: 74)	
DMD44-F	CCATCACCCTTCAGAACCTGA (SEQ ID NO: 75)	To amplify DMD exon 44 region for NGS

DMD44-R	Tcagtggctaacagaagctga (SEQ ID NO: 76)	
CLCN5 5' Sequence	GTAGGATTCTAATCACTGCCTGCTC (SEQ ID NO: 77)	Sequences used to find deletions in analyzing SMRT sequencing data
CLCN5 3' Sequence	CTTTCTGCACCTCCTGATAGCCTTG (SEQ ID NO: 78)	
MRE11-onF	GCCAAGTGTGAATGTGCACA (SEQ ID NO: 79)	To amplify the on-target of for MRE11 NGS
MRE11-onR	CCTCTTAGGCTATGACCAGGG (SEQ ID NO: 80)	
Ctip-onF	TCATTGGGAGGCCGAACATC (SEQ ID NO: 81)	To amplify the on-target of RBBP8 for NGS
CtIP-onR	AAGGGCTGAAGGATGATGCA (SEQ ID NO: 82)	
P53-onF	CTGGCATTCTGGGAGCTTCA (SEQ ID NO: 83)	To amplify the on-target of p53 for NGS
P53-onR	GAGACCTGTGGGAAGCGAAA (SEQ ID NO: 84)	
g5-onF	GTCTGAGGTCACACAGTGGG (SEQ ID NO: 85)	To amplify the on-target of intragenic 1 for NGS
g5-onR	CTGAGAGCAGGGACCACATC (SEQ ID NO: 86)	
ABE-g5-F	ACCGGATGAGATAATGATGAGTCA (SEQ ID NO: 87)	To make construct for expressing intragenic 1 sgRNA
ABE-g5-R	aaacTGACTCATCATTATCTCATC (SEQ ID NO: 88)	
P53-g1F1	ACCGCCATTGTTCAATATCGTCCG (SEQ ID NO: 89)	To make construct for expressing *P53* sgRNA
P53-g1R1	AAACCGGACGATATTGAACAATGG (SEQ ID NO: 90)	
DMD44-g1-F	CACCGtttagcatgttcccaattctc (SEQ ID NO: 91)	To make construct for expressing DMD exon 44 sgRNA
DMD44-g1-R	AAACgagaattgggaacatgctaaaC (SEQ ID NO: 92)	
HBB-sp-WF	ACCGGTAACGGCAGACTTCTCCTC (SEQ ID NO: 93)	To make construct for expressing *HBB* sgRNA
HBB-sp-WR	AAACGAGGAGAAGTCTGCCGTTAC (SEQ ID NO: 94)	
CtIP-gF	accgTTGCCCAAAGATTCCCCAG (SEQ ID NO: 95)	To make construct for expressing *RBBP8* sgRNA
CtIP-gR	aaacCTGGGGAATCTTTGGGCAA (SEQ ID NO: 96)	
Mre11-gF	accgTGACTGAGATCTGAGTGCTC (SEQ ID NO: 97)	To make construct for expressing *MRE11* sgRNA
Mre11-gR	aaacGAGCACTCAGATCTCAGTCA (SEQ ID NO: 98)	
53-sp-g2-F	ACCGactgttgcctccggttctga (SEQ ID NO: 99)	To make construct for expressing DMD exon 53 sgRNA
53-sp-g2-R	AAACtcagaaccggaggcaacagt (SEQ ID NO: 100)	
hCLCN5-sp-g1-F	ACCGGAGGACAAGTCGTACAATGG (SEQ ID NO: 101)	To make construct for expressing *CLCN5* sgRNA

hCLCN5-sp-g1-R	AAACCCATTGTACGACTTGTC CTC (SEQ ID NO: 102)	

Table 4. Single guide RNA target sequences

Name	Protospacer	PAM
CLCN5	GAGGACAAGTCGTACAATGG (SEQ ID NO: 103)	TGG
DMD53	ACTGTTGCCTCCGGTTCTGA (SEQ ID NO: 104)	AGG
DMD44	TTTAGCATGTTCCCAATTCTC (SEQ ID NO: 105)	AGG
HBB	GTAACGGCAGACTTCTCCTC (SEQ ID NO: 106)	AGG
Intragenic site 1	GATGAGATAATGATGAGTCA (SEQ ID NO: 107)	GGG
P53	CCATTGTTCAATATCGTCCG (SEQ ID NO: 108)	GGG
MRE11	TGACTGAGATCTGAGTGCTC (SEQ ID NO: 109)	TGG
RBBP8	TTGCCCAAAGATTCCCCAG (SEQ ID NO: 110)	GGG

[0203] Cell culture. HEK293T (ATCC CRL-3216™) and HEK293T-derived *CLCN5* GFP-reporter cells reported recently (Lu et al. Lentiviral Capsid-Mediated *Streptococcus pyogenes* Cas9 Ribonucleoprotein Delivery for Efficient and Safe Multiplex Genome Editing. *CRISPR J* 4(6):914-928 (2021)) were cultured in DMEM with 10% FBS, 2 mM L-glutamine, 100 U/ml penicillin and 100 µg/ml streptomycin (ThermoFisher Scientific, Waltham, MA) at 37°C in an incubator with 5% CO_2. Human lung fibroblast IMR90 cells (ATCC CCL-186) were cultured in Eagle's Minimum Essential Medium supplemented with 10% FBS, 2 mM L-glutamine, 100 U/ml penicillin and 100 µg/ml streptomycin. Human CD34+ Progenitor Cells from Mobilized Peripheral Blood (Lonza, Catalog #: 4Y-101C, Basel, Switzerland) were cultured in serum-free medium (Stemcell Technology, Catalog # 09605, Vancouver, CA) supplemented with 1x StemSpan™ CD34+ Expansion Supplement (Stemcell Technology, Catalog #02691). Human skeletal muscle myoblasts (Lonza, Catalog #: CC-2580) were cultured in SkGMTM-2 Skeletal Muscle Cell Growth Medium-2 BulletKit™ (Lonza, Catalog #: CC-3245).

[0204] Transfection of HEK293T cells. HEK293T cells were transfected in 24-well plates using FuGENE HD (Promega, Catalog#: E2312, Madison, WI). The day before transfection, 1.25×10^5 cells were seeded in 24-well plates. For DNA transfection, 0.5 µg plasmid DNA was added to 50 µl of OPTI-MEM. In a different tube, 1.5 µl FuGENE HD was

added to 50 µl OPTI-MEM. The two mixtures were mixed and incubated at room temperature for 15 mins before adding to the cells, whose medium was changed to OPTI-MEM, just before DNA transfection. Twenty-four hours after transfection, the medium was changed to normal growth medium and the cells were analyzed 72 hours after transfection.

[0205] Nucleofection of human primary cells. The Nucleofector™ 2b device (Lonza) was used for nucleofection of human primary cells. IMR90 cells, human myoblasts and Human CD34+ hematopoietic cells were nucleofected with the Cell Line Nucleofector™ Kit R (Lonza, Catalog #: VCA-1001, program X-001), Human Dermal Fibroblast Nucleofector™ Kit (Lonza, Catalog #: VPD-1001, program P-022), and the Human CD34+ Cell Nucleofector™ Kit (Lonza, Catalog #: VPA-1003), respectively. Cell number for each nucleofection was 2 x 10^5. 4.5 µg of target plasmid DNA (expressing sgRNA/Cas9 or sgRNA/Cas9-Klenow) and 0.5 µg GFP-expressing plasmid DNA (CmiR0001-MR03, GeneCopoeia, Inc., Rockville, MD) were used for each nucleofection, where the GFP-expressing plasmid DNA was used as an indicator for nucleofection efficiency. The transfected cells were checked for similar GFP positive percentage under a fluorescent microscope before further experiments.

[0206] Knocking down *MRE11* or *RBBP8* in human HEK293T cells and IMR90 cells. CRISPR/Cas9 nucleases were used to knock down *MRE11* or *RBBP8* to observe their roles in *CLCN5* mutation profiles induced by Cas9 or Cas9-Klenow. The sgRNA sequences were validated as described in Shou et al., "Precise and Predictable CRISPR Chromosomal Rearrangements Reveal Principles of Cas9-Mediated Nucleotide Insertion. *Mol Cell*, 71, 498-509 e494 (2018) and listed in Table 4. To increase the chances of *MRE11* or *RBBP8* knockdown, the DNA ratio for *CLCN5* sgRNA- and *MRE11/RBBP8* sgRNA-expressing plasmids was 1:2. The DNA mixture was co-transfected into HEK293T cells by FuGENE HD and into IMR90 cells by nucleofection as described above.

[0207] Production of lentivirus-like particles. Lentivirus-like particles were produced as described previously (Lu et al. (2021)). The packaging plasmid, pspAX2-D64V-NC-COM, has the aptamer-binding protein Com inserted in the nucleocapsid protein, and the sgRNA's ST2 loop was replaced by a com aptamer to enable the packaging of the Cas9 RNP into the lentiviral capsids via the interactions between aptamer com and aptamer-binding protein Com. In brief, 5 million HEK293T cells were seeded in 10-cm tissue culture dishes. Twenty-four hours after cell seeding, the following DNA mixture was added to 500 µl of OPTI-MEM: 7.5 µg pspAX2-D64V-NC-COM, 7.5 µg plasmids DNA expressing *CLCN5* sgRNA and Cas9 (or Cas9-Klenow), and 3 µg pMD2.G. In the meantime, 500 µl of OPTI-MEM was mixed with 54

μl of 1 mg/ml Polyethylenimine (PEI, Polysciences Inc., Warrington, PA). The mixture was incubated at room temperature for 15 mins before adding the mixture to the cells seeded the previous day, with the medium changed to OPTI-MEM just before DNA transfection. 24 hours after transfection, the medium was changed to normal growth medium and the culture medium containing the virus-like particles was collected 48 hours after medium change. The concentrations of the particles were quantitated by p24 based ELISA (Cell Biolabs, QuickTiter™ Lentivirus Titer Kit Catalog Number VPK-107, San Diego, CA).

[0208] Transduction of lentivirus-like particles. To transduce the lentivirus-like particles into HEK293T cells or *CLCN5* GFP reporter cells, the virus-like particles were added to the cells in the presence of 8 μg/ml polybrene. Up to 200 ng p24 of particles were added to 2.5 x 10^4 cells grown in 24-well plates. Twenty-four hours after transduction, the medium was changed to normal growth medium. GFP expression in *CLCN5* GFP reporter cells could be detected 36 hours after transduction. The cells were collected for DNA isolation 72 hours after transduction.

[0209] PCR amplification of target DNA for INDEL profile analyses. Genomic DNA was isolated with the DNeasy Blood & Tissue Kit (Qiagen, Germantown, MD) according to the manufacturer's instructions. The primer sequences used for amplifying target DNAs are set forth in Table 3. Genomic DNA template input for PCR was up to 0.5 μg. For samples with low DNA concentration, 0.2 μg DNA was used. Pre-determined minimal cycle numbers (25-30) were used to reduce amplification bias. The proofreading CloneAmp HiFi PCR Premix (Takara, Mountain View, USA; catalog #639298) was used for PCR.

[0210] Off-target analysis. Four potential off-targets for *HBB* sgRNA1 were analyzed to compare the off-target activities between Cas9 and Cas9-Klenow. These included G1-OT4, G1-OT5, HBD and Off-8. The regions of the predicted off-targets were amplified with their respective specific primers (Table 3) and subjected to next-generation sequencing (NGS) analysis.

[0211] NGS and data analysis. NGS analyses were done by Genewiz Inc. (Morrisville, NC) using their "Amplicon EZ" service. Approximately 50,000 reads were obtained per sample. After removing the 3' linker and 5' barcode sequences, the resulting reads were submitted to the online software Cas-Analyzer (Park et al., "Cas-analyzer: an online tool for assessing genome editing results using NGS data. *Bioinformatics*, 33, 286-288 (2017)), and CRISPResso2 (Clement et al., CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat Biotechnol*, 37, 224-226 (2019)) for mutation analysis. The two

programs gave similar results in most cases. However, CRISPResso2 did not perform well when the INDEL rates were smaller than 5%. Data presented were analyzed with Cas-Analyzer unless otherwise stated.

[0212] **Single-molecule real-time (SMRT) sequencing (PacBio).** Single-molecule real-time (SMRT) sequencing was performed to detect large deletions targeting the human *CLCN5* gene. A region of 4862 bp was amplified by LongAmp® Hot Start Taq 2X Master Mix (New England Biolab Catalog Number: M0533, Ipswich, MA) with primers hCLCN5-F3 and hCLCN5-R3. The DNA was submitted to GCB Sequencing and Genomic Technologies Shared Resource (Duke University, Durham, NC) for SMRT sequencing (Sequel I). One SMRTcell was used for eight barcoded samples. Demultiplexed Circular Consensus (CCS) reads received from the Sequencing Center were used to search for large deletions by comparing with the reference sequence through R programming. Specifically, all sequences were searched for the presence of a 25-bp 5' sequence and a 25-bp 3' sequence (see Table 3 for sequence), each allowing 2 nucleotide difference (92% identity) to accommodate possible sequencing error. These two regions are at the 5' and 3' end of the sequenced DNA. The distances between the two 25-bp regions were then calculated for each read. Those reads with distances different from that of the reference sequence, and without intact sgRNA target sites were regarded as reads with deletions. The counts and the sequences of the reads with deletions were listed.

[0213] **Statistical analysis.** GraphPad Prism software (V5) was used for t-tests, ANOVA and Chi-square tests. $P<0.05$ was regarded as statistically significant. Means and s.e.m (standard error of the mean) are reported.

B. Results

[0214] **Targeting *E. coli* DNA polymerase I to DNA DSBs increased 1-bp deletions and TISs, and decreased large deletions induced by CRISPR/Cas9.** *E. coli* DNA polymerase I (pol I) was targeted to DNA double-stranded breaks (DSBs) by fusing it to the C-terminus of *Streptococcus pyogenes* Cas9 with a linker peptide used in EvolvR in *E. coli* and yeast (FIG. 1B). To enhance fusion protein expression in human cells, the codons of the *polA* gene coding for pol I were optimized. A fusion protein was used to target the 5' untranslated region (5' UTR) of human Chloride Voltage-Gated Channel 5 (*CLCN5*), whose mutation causes a rare kidney disease named Dent's disease, in GFP-reporter cells that were developed for sensitively detecting genome editing activities (Lu et al. (2021)). These GFP-reporter cells express no GFP due to the disruption of the GFP reading frame by inserting the *CLCN5* sgRNA target sequence (from the *CLCN5* 5' UTR region) between the start codon and the second

codon of GFP coding sequence. GFP will only be expressed if genome editing restores the GFP reading frame by the in-frame INDELs (~1 in 3 chance). Similar GFP-positive cells were observed in cells treated with *CLCN5* gRNA/Cas9 and *CLCN5* gRNA/Cas9-pol. Targeted deep sequencing confirmed genome editing at the target site by *CLCN5* gRNA/Cas9-pol (FIG. 6), demonstrating the functionality of the Cas9-pol fusion protein. Plasmid DNA expressing *CLCN5* gRNA/Cas9 or *CLCN5* gRNA/Cas9-pol (targeting human *CLCN5* 5' UTR) was transfected into HEK293T cells and similar INDEL rates generated by Cas9 and Cas9-pol (11.6% ± 1.6%, N=3 for Cas9, 14.5% ± 0.7%, N=3 for Cas9-pol, p=0.16) were observed, suggesting that the fusion did not impair Cas9 activity.

[0215] INDEL profiles induced by Cas9 and Cas9-pol targeting *CLCN5* were analyzed, and it was found that Cas9-pol caused a significant increase of 1-base pair (bp) deletions, a decrease of >1-bp deletions (FIG. 1C), and an increase of 1-bp TIS versus 1-bp non-TIS ratio (FIG. 1D). The data supported the hypothesis that targeting polymerase I to DSBs favors the generation of short deletions over long deletions, and TIS over non-TIS.

[0216] Targeting polymerase to the DSBs are necessary for modulating TIS but not deletion size. In order to test whether targeting the polymerase to the DSBs is necessary for the observed effects, NLS-pol protein, which was a deletion mutant of Cas9-pol that contained full length pol I and the nuclear localization signals (NLS), but lacked majority of the Cas9 functional domains (part of REC and RuvC, the whole HNH and part of PAM-interacting domain of Cas9) (FIG. 2A), was made. Since NLS-pol contained the NLS and linker sequences of Cas9-pol, it was expected to fold properly and be targeted to the nuclei. However, it should not have Cas9's DNA-binding and nuclease activities due to the deletion of multiple Cas9 domains. Cas9, *CLCN5* gRNA, and NLS-pol were co-expressed in HEK293T cells, and it was found that cells treated with Cas9 and NLS-pol had 1-bp deletion percentages between those of Cas9 treated cells and Cas9-pol treated cells, and had increased DNA substitution rates in a region 20bp 5' and 20bp 3' of the predicted cleavage site for unknown mechanisms (FIG. 2B). Importantly, co-expression of Cas9 and NLS-pol did not increase TIS to the level of Cas9-pol (FIG. 2C).

[0217] To determine whether the polymerase activity is necessary for the observed effects, Cas9 was fused to various mutants and truncated pol proteins, including pol^{D705A} with the polymerase activity inactivated. Klenow fragment, Klenow^{D705A} with the polymerase activity inactivated, the 5' exonuclease domain (5Exo), the 3' exonuclease domain (3Exo), and both of the 5' exonuclease and the 3' exonuclease domains (Exo) (FIG. 2A). *CLCN5* was targeted in

HEK293T cells using these Cas9 fusion proteins, and it was found that all fusion proteins without the DNA polymerase domain failed to increase 1-bp deletions and decrease >1-bp deletions. All fusion proteins with the DNA polymerase domain, regardless of whether the polymerase activity was inactivated, increased 1-bp deletions and decreased >1-bp deletions (FIG. 2B). Since Cas9-5exo, Cas9-3exo and Cas9-exo showed very similar mutation profiles with each other (FIG. 7), they were treated as one group (Cas9-all Exo) in FIG. 2B. Thus the polymerase domain, rather than the polymerase activity was enough to perturb the ratio of 1-bp and >1-bp deletions.

[0218] When TIS was analyzed, only DNA pol I and Klenow could increase 1-bp TIS and decrease 1-bp non-TIS (FIG. 2C), consistent with the need of polymerase-mediated filling in for generating TIS. The same conclusions were made when 2-bp and 3-bp TISs were analyzed (FIG. 8). Compared with pol I, Klenow showed stronger effects on increasing 2-bp and 3-bp TISs. This is likely because the Klenow fragment lacks the 5'>3' exonuclease domain present in pol I, and the 5'>3' exonuclease domain may remove the 5' overhangs and favor the generation of deletions.

[0219] *CLCN5* was similarly targeted in human primary IMR90 cells with various fusion proteins. In these cells, all fusion proteins with the Klenow domain (with or without polymerase activity) significantly increased 1-bp deletion frequency, but decreased insertion frequency rather >1-bp deletion frequency (FIG. 2D). Again, targeting the polymerase activity to the DSBs were necessary for increasing TIS (FIG. 2E).

[0220] The relationship between the mutation profiles and the overall INDEL rates targeting the *CLCN5* locus by Cas9 were examined, and it was found that overall INDEL rates did not affect mutation profiles (FIG. 9). Thus the observed effects of fusing pol I or Klenow fragment to Cas9 on DNA mutation profiles could not be explained by possible effects on Cas9 cleavage activity. The most likely mechanism could be interfering local cellular DNA repair machineries.

[0221] Altogether, targeting polymerase to the DSBs were necessary for increasing TIS, although neither fusing with Cas9 nor the polymerase activity were necessary for increasing 1-bp deletions. However, fusing polymerase to Cas9 is still beneficiary since doing so will increase local polymerase concentration and decease possible interference on other endogenous DSBs, especially in many experimental settings controlled amount of genome editing effectors rather than overexpressed ones are used for genome editing.

[0222] DNA resection was involved in the perturbation of 1-bp versus >1-bp deletions. The observations that pol^{D705A} and Klenow^{D705A} also favored the generation of 1-bp deletions over >1-bp deletions prompted us to examine whether DNA resection was involved in these effects. We thus examined the mutation profiles targeting the *CLCN5* locus in cells with and without *RBBP8* (expressing CtIP protein) knockdown. Considering the possibility of only disrupting the expression from one allele, we use the term knockdown rather than knockout. We used the previously validated *RBBP8* sgRNA (6) to mediate the mutation of this gene. *RBBP8* sgRNA-expressing DNA were co-transfected with *CLCN5* sgRNA-expressing DNA into HEK293T cells. The sgRNA-expressing constructs also contained Cas9 or Cas9-Klenow-expressing cassettes (see supplementary Table 1 for DNA constructs used). INDEL rates of the co-transfected genes (*RBBP8* and *CLCN5*) were very similar (Supplementary Table S4), consistent with co-transfection. We observed quite different overall INDEL rates in Cas9 and Cas9-Klenow-treated cells, for unknown reasons. However, since DNA mutation profiles were independent of overall INDEL rates (Supplementary Figure S4), this INDEL rate difference is unlikely to affect our mutation profile analyses.

[0223] Knocking down *RBBP8* did not affect the percentage of 1-bp deletions, but significantly decreased the percentage of >1-bp deletions, and increased the percentage of insertions when targeting *CLCN5* in HEK293T cells (Figure 3A). These results are consistent with the suppression of DNA resection, which may cause increased insertions. These observations are similar to those from the Repair-Seq, published while our paper was under revision (54). Importantly, in cells with *RBBP8* knocked down, the Klenow fragment decreased >1-bp deletions to a smaller extent than in cells without *RBBP8* knockdown, suggesting that *RBBP8* involved in Klenow-mediated decrease of >1-bp deletions. TIS of 1-bp was not significantly changed in *RBBP8* knockdown HEK293T cells, and Klenow fragment increased 1-bp TIS (Figure 3B).

[0224] The *CLCN5* mutation profiles in IMR90 cells were also examined, with and without *RBBP8* knockdown. Knocking down *RBBP8* in IMR90 cells increased Cas9-induced 1-bp deletions, an effect similar to fusing Klenow fragment to Cas9 in normal IMR90 cells (FIG. 10). Doing so also increased 1-bp TIS and decreased 1-bp non-TIS. In *MRE11* or *RBBP8* knockdown cells, Klenow had no effects on 1-bp deletions, >1-bp deletions or TIS. These observations suggest that knocking down *MRE11* or *RBBP8* and fusing Klenow to Cas9 may interfere with similar pathways in IMR90 cells. Altogether, these data suggest that interfering

with DNA resection is at least one of the mechanisms for pol I or Klenow fragment's effects on Cas9 DNA mutation profiles.

[0225] For deletions from 1 to 30 bps targeting *CLCN5* in HEK293T cells, except for 1-bp deletions, 11-bp and 8-bp deletions were the most frequently observed (Figure 3C). The frequencies of the 11-bp deletions were unaffected by *RBBP8* knockdown (Figure 3C), suggesting that other mechanisms may also involve in the generation of these deletions. These deletions were greatly suppressed by fusing the Klenow fragment to Cas9. The frequency of 8-bp deletions was decreased by *RBBP8* knockdown, and not further decreased by the Klenow fragment. This observation further supports the concept that one way in which the Klenow fragment affects the Cas9 DNA mutation profile is via interfering with DNA resection.

[0226] Microhomology was noted around the deletion when examining the most frequently deleted sequences in Cas9-treated cells (Figure 3D, underlined). The most frequently observed 11-bp deletion had microhomology (underlined green) at the very end of the predicted cleavage site (dashed line). In this case, DNA synthesis could be initiated without the need for 3'-flap endonucleases such as XPF-ERCC1 (55,56) to remove the unmatched 3' flap during MMEJ. This possibly explains why the 11-bp deletion was most frequently seen in Cas9-treated cells. *RBBP8* knockdown had no effects on the frequency of this deletion, and there could be other unknown proteins to generate similar deletions under resection inhibition. In general, the common deletions either had microhomology around the deletion, or had one end at the predicted cleavage site, or both.

[0227] **Cas9-Klenow increased 1-bp deletions on multiple loci in multiple human cell types.** Whether pol I or Klenow's effects on DNA mutation profiles were target sequence- or cell type- specific was examined. The Cas9-Klenow fusion protein was used in subsequent experiments considering its smaller size and prominent effects on increasing 1-bp deletions and TIS. Four more loci were examined, including Duchenne muscular dystrophy (*DMD*) exon 53 and *DMD* exon 44 (537 kb away from each other), the 5' coding region of *HBB*, and an intergenic locus intragenic 1 (GRCh38.p13, chromosome 20, 32752960–32752979). *DMD* exons 53 and 44 were picked because targeting these exons with a single-cut sgRNA might restore dystrophin in DMD patients caused by exon deletion. The *HBB* 5' coding region was picked for possible application of genome editing in treating sickle cell disease. In addition, this region has been targeted with CRISPR/Cas9 to examine Cas9-indcued gene conversion in human somatic cells (Parsijani et al. "CRISPR/Cas9 increases mitotic gene conversion in human cells," *Gene Ther*, **27**, 281-296 (2020)). The intragenic intragenic 1 was picked to rule

out possible contributions of target gene product on the observed effects. In addition to HEK293T cells, various loci were targeted in human primary fibroblast IMR90 cells, human CD34+ hematopoietic stem cells or human primary myoblasts. A total of 8 loci/cells were examined (Table 5).

Table 5. Effects of Cas9-Klenow on deletions, total insertions and 1-bp TIS

Locus/cell	1-bp deletions (% of all INDELs)		>1-bp deletions (% of all INDELs)		Insertions (% of all INDELs)		1-bp TIS (% of all insertions)	
	Cas9	Cas9-Kle	Cas9	Cas9-Kle	Cas9	Cas9-Kle	Cas9	Cas9-Kle
CLCN5/293T	9.9±0.92	23.1±0.53***	55.3±0.66	39.8±1.2***	34.8±0.45	37.1±1.1*	45.0±1.39	65.4±1.44***
CLCN5/IMR90	18.1±2.00	39.3±3.59**	36.7±3.90	39.3±3.72	45.1±5.16	22.7±3.60*	41.2±8.57	70.4±5.33*
DMD53/293T	8.4±0.14	19.8±0.45***	64.4±1.3	51.6±0.37***	27.2±1.40	28.5±0.31	34.0±4.17	52.6±0.68*
DMD44/myoblasts	11.2±1.95	28.8±2.71**	75.8±2.06	57.4±2.65**	13.0±1.41	13.7±0.67	32.6±9.27	66.46±6.63*
HBB/hematopoietic cells	37.4±4.11	59.6±3.05*	56.7±4.59	36.4±2.90*	4.2±1.10	4.0±0.42	21.2±6.02	46.7±5.90*
HBB/IMR90	40.8±9.99	81.2±5.47*	55.9±10.64	14.7±6.32*	2.3±1.02	4.5±2.34	Too few	Too few
HBB/293T	13.3±0.84	46.7±1.34***	65.1±0.59	41.1±1.29***	19.6±1.03	10.3±0.16**	41.9±1.14	39.1±3.33
Intragenic 1/IMR90	14.3±2.1	33.7±1.22**	47.4±2.1	48.0±1.41	38.3±0.60	18.2±0.47***	23.3±1.93	50.3±0.96**

*, **, *** indicate $p<0.05$, 0.01 and 0.0001 respectively in two-tailed t-tests. Blue, red and black indicate increased, decreased and unchanged in the Cas9-Klenow group respectively. Cas9-Kle: Cas9-Klenow. Since Klenow fusion changed the total insertion percentages significantly in 5 loci/cells, 1-bp TIS was expressed as % of all insertions.

[0228] In all cases, targeting Klenow fragment to DSBs significantly increased 1-bp deletions (Table 5), which increased from an average of 18.80% ± 4.62% (N=8) to 41.29% ± 7.39% (N=8) in the 8 analyzed loci/cells. In *HBB*/IMR90, 81.2% ± 5.47% of all INDELs generated by Cas9-Klenow were 1-bp deletions. The increase of 1-bp deletions was accompanied by four possible phenomena: 1) decrease of >1-bp deletions (*DMD44*/myoblasts, *HBB*/hematopietic cells and *HBB*/IMR90); 2) decrease of insertions (*CLCN5*/IMR90, Intragenic 1/IMR90), 3) Decrease of both >1-bp deletions and insertions (*HBB*/293T); and 4) decrease of >1-bp deletions and increase of total insertions (*CLCN5*/293T, *DMD53*/293T). Whereas Klenow's effects on insertions varied, it caused a significant decrease of >1-bp deletion in 6 of 8 loci/cell. Targeting *HBB* locus in IMR90, HEK293T and Hematopietic cells showed different ways to account for the increase of 1-bp deletions.

[0229] In all loci/cell except for *CLCN5*/IMR90 which had no evident deletion peaks (FIG. 11), Cas9-Klenow most significantly decreased the percentages of >1-bp deletions with microhomology at the predicted cleavage site (underlined by green lines), which were usually the highest >1-bp deletion peak generated by Cas9 (FIG. 4, FIG. 11). The smaller >1-bp

deletion peaks either had no microhomology but located at the predicted cleavage site (dashed vertical line), or had microhomology away from the cleavage site (underlined by red lines). These observations suggest that those MMEJ events without the need to remove 3' flaps were most sensitive to suppression by Klenow.

[0230] Cas9-Klenow increased TIS on multiple loci in multiple human cell types. The effects of Klenow fragment on TIS in these loci/cell were exained. The focus was on 1-bp TIS, since 2- and 3-bp TIS were few in the ~50,000 reads we analyzed in most loci/cell. Since in several cases Klenow greatly changed the percentage of overall insertions, the percentages of TIS in all insertions rather than all INDELs was compared. Except for one locus/cell (*HBB*/IMR90) which had too few insertions and were not analyzed, 6 of 7 loci/cells showed a significant increase in the shares of TIS in all insertions (Table 5). Only *HBB*/293T showed similar TIS percentages between Cas9 and Cas9-klenow, and ~80% 1-bp insertions were TIS. This observation suggests that in the case of *HBB*/293T the endogenous polymerases were very efficient in filling in the 5' overhangs, explaining why Cas9-Klenow did not further increase the TIS share.

[0231] Fusing Klenow fragment to Cas9 caused a significant reduction of total insertions in three loci/cells (*CLCN5*/IMR90, *HBB*/293T and Intragenic 1/IMR90). In two of the three loci/cells (*CLCN5*/IMR90 and Intragenic 1/IMR90), non-TIS insertions contributed to 100% of the insertion decrease.

[0232] Fusing Klenow to Cas9 decreased CRISPR/Cas9-induced large DNA deletions. The observations of targeting DNA polymerase to DSBs favoring 1-bp deletions over >1-bp deletions prompted examination of whether doing so could reduce the generation of large (>500 bp) on-target deletions. Large deletions targeting *CLCN5* in HEK293T cells were analyzed. A pair of primers amplifying a region of 4862 bp, with the sgRNA target sequence in the middle of the amplicon (FIG. 4), were designed. Lentivirus-like particles containing the Cas9 RNPs or the Cas9-Klenow RNPs were generated, and similar percentages of GFP-positive cells after treating the *CLCN5* GFP-reporter cells with the two types of RNPs, suggesting equivalent genome editing activities were observed.

[0233] Then, HEK293T cells were treated with Cas9 RNPs or Cas9-Klenow RNPs. DNA was amplified 72 hours after treatment and single-molecule real-time (SMRT) sequencing (PacBio) was perfored. More reads with >0.5 kb, >1 kb and >2 kb deletions in Cas9 treated cells than in Cas9-Klenow treated cells (Table 6 and FIG. 5) were found.

Table 6. Large deletions generated by Cas9 and Cas9-Klenow targeting *CLCN5*.

	HEK293T cells			IMR90 cells		
	Cas9	**Cas9-Klenow**	**No sgRNA**	**Cas9**	**Cas9-Klenow**	**No sgRNA**
Total CCS[a]	23477	11497	12040	19147	29793	11417
CCS with >0.5 kb deletions	100	10	0	151	13	1
CCS with >1 kb deletions	76	1	0	132	2	1
CCS with >2 kb deletions	28	1	0	50	2	0
Chi-square test	P<0.0001			P<0.0001		

[a]Only those CCSs containing both the 5' and 3' index sequences (boxes in FIG. 5) were analyzed. CCS: Circular Consensus

[0234] *CLCN5* waas targeted in IMR90 cells by nucleofection of plasmid DNA expressing *CLCN5* sgRNA/Cas9 or *CLCN5* sgRNA/Cas9-Klenow. Even though Cas9-treated cells had lower short INDEL rates than Cas9-Klenow-treated cells (See Table 7, which used the same DNA samples for NGS analysis), they had more >0.5kb, >1kb and >2kb deletions compared with Cas9-Klenow treated cells.

Table 7. INDEL rates (%) of Cas9 and Cas9-Klenow in human primary cells.

Loci/cell	Cas9	Cas9-Klenow	Background
HBB/IMR90	2.6 ± 0.50	6.2 ± 0.87*	0.02
HBB/CD34+ cells	0.77 ± 0.03	1.6 ± 0.09***	0.01
CLCN5/IMR90	4.9 ± 0.50	7.8 ± 0.79*	0.09
DMD44/myoblasts	11.9 ± 1.8	12.8 ± 0.4	0
Intragenic 1/IMR90	0.77 ± 0.12	5.8 ± 0.46 ***	N/A

* and *** indicate p<0.05 and p<0.0001 respectively in t-tests (n=3).

[0235] CCS reads with >0.5 kb deletions were examined, and it was found that the deleted regions either span or involved the sgRNA target site (FIG. 5), confirming that they were on-target deletions. Only limited types of deletions were observed. Cas9 generated more types of deletions and often more CCS reads for each type of deletion than Cas9-Klenow did. Both phenomenon could be explained by Cas9 being more prone to generate large deletions than did Cas9-Klenow. Multiple CCS reads for the same type of deletion could be the result of a single deletion event in one cell or multiple deletion events in multiple cells. Our observations of the same deletion type in Cas9-treated and Cas9-Klenow-treated IMR90 cells (indicated by * in FIG. 5B) supported the possibility of multiple deletion events for the same deletion type. Alternatively, Cas9 could have induced the large deletions sooner after treatment than Cas9-Klenow, which increased the representation of the deletions.

[0236] **Cas9-Klenow increased INDEL rates in human primary cells.** Whereas in HEK293T cells Cas9 and Cas9-Klenow (or Cas9-pol) generated similar levels of INDEL rates,

in human primary cells Cas9-Klenow generated significantly higher INDEL rates in 5 of 6 loci/cells (Table 7). In general, the INDEL rates were relatively low in primary cells and the reason was unclear. The observed INDELs were confirmed as authentic INDELs, since the background INDELs of cells treated with non-targeting sgRNA were very low, and all the INDELs were around the predicted cleavage sites. 1/10 GFP-expressing plasmid DNA were included in the nucleofection experiments and >50% GFP-positive cells were observed. Thus, the low INDEL rates were not due to low nucleofection efficiency. Cell death was observed during the 72 culture time post nucleofection. It is possible that many Cas9 or Cas9-Klenow positive cells might have been lost due to the toxic effects of constantly generating DSBs, as observed in human ES cells (Happaniemie et al., "CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response," *Nat Med*, **24**, 927-930 (2018)).

[0237] **Cas9-Klenow did not increase DNA substitution rates or off-targets.** Whether targeting DNA polymerase to DSBs could increase DNA mutation rates around DSBs was examined. DNA substitution rates in a 40 bps region around the predicted cleavage sites (20 bps on each side) were examined for the following reasons: 1) MRE11 nicks 15-20 nt away from the DSBs (18), 2) pol I showed a processivity of 15–20 nucleotide (61), and 3) EvolR (the fusion protein between Cas9 nickase and error prone DNA pol I) showed a mutation window of 15-20 bp (47). Cas9-Klenow did not cause increase in DNA substitution rates (Table 4). In addition, similar DNA substitution rates were also observed in negative controls with non-targeting sgRNAs, which suggests that the observed DNA substitutions were mainly from cell heterogeneity, PCR and sequencing errors. Thus, targeting DNA polymerase to DSBs did not increase DNA substitution rates.

[0238] The effects of fusing the Klenow fragment to Cas9 on possible off-targets of Cas9 were also tested. The *HBB* 5' coding region in human IMR90 cells was targeted and INDEL rates detected at four potential off-targets predicted based on sequence similarity (Table 8; PAMs are underlined; italicized nucleotides indicate a mismatch). Off-targets in Cas9 and Cas9-Klenow treated cells using targeted NGS (<0.5%, the detection limit of NGS), which is currently one of the most sensitive off-target detection methods, were not observed. Thus, the data showed that fusing Klenow to Cas9 did not increase off-targets to a detectable level by NGS. Since Klenow fragment did not cause detectable off-targets on sequences with similarity to the authentic target, it is less likely to cause random off-targets on sequences without similarity to the authentic target.

Table 8. Sequences of off-targets HBB, HBD, OT-4, OT-5, and OT-8.

Target	SEQ ID NO	Sequence
HBB	148	GTAACGGCAGACTTCTCCTCAGG
HBD	149	*T*T*G*AC*A*GCAG*T*CTTCTCCTCAGG
OT-4	150	*A*TA*CTTAAG*GACTTCTCCTCCAT
OT-5	151	G*G*A*GG*GGCAG*G*CTTCTCCTCTGG
OT-8	152	*TC*A*CA*GGCAGACTTCTCC*A*CGGG

[0239] Klenow's effects on off-targets were examined in a third way. HEK293T derived GFP-reporter cells were used for detecting CRISPR/Cas9 induced INDELs in a *HBB* sickle mutant sequence, which has a one nucleotide difference with the *HBB* sgRNA, and is an "off-target" for the *HBB* sgRNA. These cells contain *HBB* sgRNA authentic targets in the endogenous *HBB* gene and off-targets in the integrated GFP-reporter cassette. The endogenous *HBB* gene with the perfectly matching *HBB* sgRNA was targeted, and INDEL rates in the sickle mutant sequence as an off-target were examined. Analysis showed that Cas9 and Cas9-Klenow had similar INDEL rates on the endogenous *HBB* target (21.47% ± 0.80%, N=3 for Cas9; 21.70% ± 1.10%, N=3 for Cas9-Klenow, p=0.8723), and also on the integrated sickle mutant sequence (10.80% ± 0.80%, N=3 for Cas9; 12.73% ± 1.20%, N=3 for Cas9-Klenow, p=0.2509). Altogether these experiments demonstrate that targeting Klenow fragment to DSBs did not increase off-targets or DNA mutation at DSBs.

[0240] As shown herein, targeting *E. coli* DNA pol I or the Klenow fragment to DSBs with Cas9 fusion proteins increased the ratio of small deletions versus large deletions, and the ratio of TIS versus non-TIS. Importantly, doing so suppressed the generation of on-target deletions larger than 500 bp. These effects were observed in all loci analyzed (8 loci/cell), involving 4 cell types (one cell line and three primary cell types) and 5 target sites. The effects of reducing deletion sizes and increasing TIS over non-TIS are not cell type or target site specific. In primary cells, fusing Klenow to Cas9 caused a significant increase in overall INDEL rates in 4 out of 5 cases.

[0241] DNA resection is necessary for HDR, MMEJ and SSA. The latter two alternative NHEJ DNA repair pathways will generate short and long deletions. The MRE11-RAD50-NBS1 complex is responsible for initiating DNA resection, and EXO1, BLM and DNA2 are responsible for extensive resection. An attempt waa made to suppress the generation of large deletions via counteracting DNA resection. The data provided herein suggest that interfering

with DNA resection is one of the mechanisms for Klenow's effects on deletion sizes. First, it was determined that knocking down MRE11 or CtIP, proteins involved in DNA resection, increased 1-bp deletions and decreased >1-bp deletions. Second, Klenow's effects on deletion sizes were lost under MRE11 or CtIP knockdown. These observations are consistent with a report that inhibiting the MRE11 complex causes the suppression of MMEJ (Hussmann et al. "Mapping the genetic landscape of DNA double-strand break repair. *Cell*, 184, 5653-5669 e5625 (2021)). A frequently observed 11-bp deletion in *CLCN5* was not affected by MRE11 or CtIP knockdown, but was suppressed by Cas9-Klenow fusion, suggesting that Klenow fusion also decreases MRE11-independent deletions. Fusing Klenow to Cas9 did not decrease the frequency of a 9-bp deletion when targeting *DMD* exon 53 in HEK293T cells. This deletion may be generated by a process that Klenow could not interfere with.

[0242] It was determined that pol^{D705A} and $Klenow^{D705A}$ with inactivated polymerase activity had similar effects on deletion sizes as pol I and Klenow fragment. It is possible that these proteins interfered with DNA resection via the following two non-exclusive mechanisms: 1) the addition of a bulk peptide (≥629 AA) to the C-terminus of Cas9 may prevent the recruitment of the DNA resection complex or regulatory proteins, and 2) the residual DNA binding activity of pol^{D705A} or $Klenow^{D705A}$ may interfere with the DNA resection. Since pol^{D705A} and $Klenow^{D705A}$ do not affect the percentage of TIS, they could be useful when one only needing to increase small deletions and decrease large deletions. Although counteracting DNA resection appears to be be the mechanism underlying the observations, other cellular DNA damage repair machineries cannot be entirely ruled out.

[0243] Cas9-pol and Cas9-Klenow's capability of increasing TIS/non-TIS ratio depended on local availability of the polymerase activity. This dependency is consistent with the observations that TISs are the results of filling in Cas9-generated 5' overhang ends by polymerase. In the studies provided herein, the Klenow fragment was more active than pol I in increasing 2-bp and 3-bp TIS (FIG. 8). This suggests that the 5' exonuclease domain of pol I (absent in Klenow fragment) could compete with the polymerase domain for 5' overhangs. The former removes 5' overhangs and favors deletions, whereas the latter fills in the 5' overhang to produce TIS. Removing 1-nt 5' overhangs could also be one of the mechanisms for pol^{D705A} to increase 1-bp deletions.

[0244] Cas9-pol and Cas9-Klenow only increased TIS but not non-TIS. When targeting *CLCN5* or intragenic site 1 in IMR90 cells, fusing Klenow to Cas9 greatly decreased overall insertions, however, only the non-TIS but not the TIS was decreased. In human cells, DNA

polymerase μ is necessary for generating both TIS and non-TIS, but whether other proteins are needed to generate non-TIS is unclear. That targeting Klenow fragment to Cas9 only increased TIS but not non-TIS (or in some cases, only decreased non-TIS but not TIS) suggests that TIS and non-TIS are generated via different mechanisms. Fusing pol or Klenow fragment to Cas9 promotes filling in 5' overhangs, but may have an inhibitory effects on recruiting proteins responsible for generating non-TIS.

[0245] Cas9-Klenow significantly increased overall INDEL rates in 4 of 5 cases in human primary cells but not in HEK293T cells. This could be explained by several non-exclusive mechanisms: 1) counteracting DNA resection and inhibiting homologous recombination which perfectly repairs the DNA; 2) filling in 5' overhangs before DNA ligase ligates the complementary 5' overhangs without generating INDELs; and 3) Cas9 induces P53-mediated DNA damage stress in primary cells but not in HEK293T cells, and fusing Klenow to Cas9 reduces the stress by preventing repeated futile editing. The effects of fusing Klenow to Cas9 on HDR were not examined. Doing so so may have an inhibitory effect on HDR considering its effects on DNA resection that is necessary for HDR. Targeting the Klenow domain to DSBs could be useful in genome editing applications independent on HDR.

[0246] These stuides show that fusing a Klenwo domain to an RNA-guided endocnulease, for example, a Cas9 endonuclease, could improve genome editing safety and efficiency in *in vitro* and *in vivo* applications. Fusing the Klenow domain to Cas9 decreased the generation of unpredictable large on-target DNA deletions, which have been observed by multiple groups (Owens et al. "Microhomologies are prevalent at Cas9-induced larger deletions. *Nucleic Acids Res*, 47, 7402-7417 (2019); Adikusuma et al. "Large deletions induced by Cas9 cleavage. *Nature*, 560, E8-E9 (2018); and Kosicki et al. "Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat Biotechnol*, 36, 765-771 (2018). It increased genome editing efficiency in primary cells, and did not increase DNA substitution rate or off-target rate. In addition, the effects on 1-bp deletions and TIS can be used to increase the percentage of desirable types of mutations to improve the efficiency of disrupting disease-causing genes, or restoring disrupted genes by reframing.

Sequences

SEQ ID NO:1	MS2 coat protein (MCP) DNA Sequence	ATGGCTTCTAACTTTACTCAGTTCGTTCTCGTCGACAATGG CGGAACTGGCGACGTGACTGTCGCCCCAAGCAACTTCGCT AACGGGATCGCTGAATGGATCAGCTCTAACTCGCGTTCAC AGGCTTACAAAGTAACCTGTAGCGTTCGTCAGAGCTCTGC GCAGAATCGCAAATACACCATCAAAGTCGAGGTGCCTAAA GGCGCCTGGCGTTCGTACTTAAATATGGAACTAACCATTC

		CAATTTTCGCCACGAATTCCGACTGCGAGCTTATTGTTAAG GCAATGCAAGGTCTCCTAAAAGATGGAAACCCGATTCCCT CAGCAATCGCAGCAAACTCCGGCATCTAC
SEQ ID NO:2	MS2 coat protein (MCP) Amino Acid Sequence	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQA YKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFA TNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY
SEQ ID NO:3	PP7 coat protein (PCP) DNA Sequence	tccaaaacaatagtcctctccgtaggggaggcaacacggactttgaccgaaatccagtcaaccg ctgaccgacaaatctttgaagagaaagtagggcctcttgtgggccgactgcgcttgactgcaagc ttgcgacaaaacggcgcaaagactgcctatagggtcaaccttaaactcgaccaagccgacgtgg tcgatagcggtctccctaaggttcggtatacgcaggtctggagtcatgacgtaacaatcgtagcaa acagcacagaagcctcccgaaaaagcctctacgatctgacgaaatccttggtggctacgtcaca ggtggaagacctcgttgtcaaccttgtacctctgggtcga
SEQ ID NO:3	PP7 coat protein (PCP) Amino Acid Sequence	SKTIVLSVGEATRTLTEIQSTADRQIFEEKVGPLVGRLRLTASL RQNGAKTAYRVNLKLDQADVVDSGLPKVRYTQVWSHDVTI VANSTEASRKSLYDLTKSLVATSQVEDLVVNLVPLGR
SEQ ID NO:5	lambda N RNA-binding domain (positions (1-22) DNA Sequence	ATGGATGCACAAACACGCCGCCGCGAACGTCGCGCAGAG AAACAGGCTCAATGGAAAGCAGCAAAT
SEQ ID NO:6	lambda N RNA-binding domain (positions (1-22) Amino Acid Sequence	MDAQTRRRERRAEKQAQWKAAN
SEQ ID NO:7	Com Protein DNA Sequence	atgaaatcaattcgctgtaaaaactgcaacaaactgttatttaaggcggattcctttgatcacattga aatcaggtgtccgcgttgcaaacgtcacatcataatgctgaatgcctgcgagcatcccacggaga aacattgtgggaaaagagaaaaaatcacgcattctgacgaaaccgtgcgttattgagtat
SEQ ID NO:8	Com Protein Amino Acid Sequence (GenBank AAF01130.1)	MKSIRCKNCNKLLFKADSFDHIEIRCPRCKRHIIMLNACEHPT EKHCGKREKITHSDETVRY
SEQ ID NO:9	MS2 aptamer sequence (RNA)	ACAUGAGGAUCACCCAUGU
SEQ ID NO:10	MS2 aptamer sequence (DNA)	ACATGAGGATCACCCATGT
SEQ ID NO:11	PP7 aptamer sequence (RNA)	GGAGCAGACGAUAUGGCGUCGCUCC
SEQ ID NO:12	PP7 aptamer sequence (DNA)	GGAGCAGACGATATGGCGTCGCTCC
SEQ ID NO:13	Box-B; lambda N RNA-binding domain aptamer sequence (RNA)	GGGCCCUGAAGAAGGGCCC
SEQ ID NO:14	Box-B; lambda N RNA-binding domain aptamer sequence (DNA)	GGGCCCTGAAGAAGGGCCC
SEQ ID NO:15	com aptamer RNA sequence	CUGAAUGCCUGCGAGCAUC
SEQ ID NO:16	com aptamer DNA sequence	CTGAATGCCTGCGAGCAT

SEQ ID NO:17	human beta hemoglobin (HBB) 3' UTR (DNA)	gctcgctttcttgctgtccaatttctattaaaggttcctttgttccctaagtccaactactaaactgggg gatattatgaagggccttgagcatctggattctgcctaataaaaaacatttattttcattgc
SEQ ID NO:18	human beta hemoglobin (HBB) 3' UTR (RNA)	gcucgcuuucuugcuguccaauuucuauuaaagguuccuuuguucccuaaguccaacu acuaaacugggggauauuaugaagggccuugagcaucuggauucugccuaauaaaaaa cauuuauuuucauugc
SEQ ID NO:19	HIV-1 Nucleocapsid (NC) DNA Sequence	ATACAGAAAGGCAATTTTAGGAACCAAAGAAAGACTGTTA AGTGTTTCAATTGTGGCAAAGAAGGGCACATAGCCAAAAA TTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAAATGTGGA AAGGAAGGACACCAAATGAAAGATTGTACTGAGAGACAG GCTAAT
SEQ ID NO:20	HIV-1 Nucleocapsid (NC) Amino Acid Sequence	IQKGNFRNQRKTVKCFNCGKEGHIAKNCRAPRKKGCWKCG KEGHQMKDCTERQAN
SEQ ID NO:21	HIV-1 Matrix protein (MA) DNA Sequence	atgggtgcgagagcgtcagtattaagcgggggagaattagatcgatgggaaaaaattcggttaa ggccagggggaaagaaaaaatataaattaaaacatatagtatgggcaagcagggagctagaac gattcgcagttaatcctggcctgttagaaacatcagaaggctgtagacaaatactgggacagctac aaccatcccttcagacaggatcagaagaacttagatcattatataatacagtagcaaccctctattgt gtgcatcaaaggatagagataaaagacaccaaggaagctttagacaagatagaggaagagcaa aacaaaagtaagaaaaaagcacagcaagcagcagctgacacaggacacagcaatcaggtcag ccaaaattac
SEQ ID NO:22	HIV-1 Matrix protein (MA) Amino Acid Sequence	GARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELE RFAVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATL YCVHQRIEIKDTKEALDKIEEEQNKSKKKAQQAAADTGHSN QVSQNY
SEQ ID NO:23	HIV-1 Viral Protein (VPR) DNA Sequence	ATGGAACAAGCCCCAGAAGACCAGGGACCGCAGAGGGAA CCATACAATGAATGGACACTAGAACTTTTAGAGGAACTCA AGCGGGAAGCAGTCAGACACTTTCCTAGACCATGGCTTCA TGGCTTAGGACAACATATCTATGAAACCTATGGAGATACT TGGACGGGGGTGGAAGCTATAATAAGAATTCTGCAACGAC TACTGTTTGTCCATTTCAGAATTGGGTGCCAGCATAGCCGA ATAGGCATTCTAAGACAGAGAAGAGCAAGAAATGGAGCC AGTAGATCCTAA
SEQ ID NO:24	HIV-1 Viral Protein (VPR) Amino Acid Sequence	MEQAPEDQGPQREPYNEWTLELLEELKREAVRHFPRPWLHG LGQHIYETYGDTWTGVEAIIRILQRLLFVHFRIGCQHSRIGILR QRRARNGASRS
SEQ ID NO:25	HIV-1 Negative Regulatory Factor (NEF) DNA Sequence with codon changes to enhance packaging in the virus core (G3C, V153L, and E177G mutations; underlined)	atgggtTgcaagtggtcaaaaagtagtgtgattggatggcctgctgtaagggaaagaatgagac gagctgagccagcagcagatggggtgggagcagtatctcgagacctagaaaaacatggagca atcacaagtagcaatacagcagctaacaatgctgcttgtgcctggctagaagcacaagaggagg aagaggtgggttttccagtcacacctcaggtacctttaagaccaatgacttacaaggcagctgtag atcttagccactttttaaaagaaaagggggactggaagggctaattcactcccaaagaagacaa gatatccttgatctgtggatctaccacacacaaggctacttccctgattggcagaactacacacca gggccaggggtcagatatccactgacctttggatggtgctacaagctagtaccagttgagccaga taagCtGgaagaggccaataaaggagagaacaccagcttgttacaccctgtgagcctgcatgg aatggatgaccctgGAagagaagtgttagagtggaggtttgacagccgcctagcatttcatcac gtggcccgagagctgcatccggagtacttcaagaactgc (The yellow positions are changed to code for the changes explained in seq ID. 8.
SEQ ID NO:26	HIV-1 Negative Regulatory Factor (NEF) Amino	MGCKWSKSSVIGWPAVRERMRRAEPAADGVGAVSRDLEKH GAITSSNTAANNAACAWLEAQEEEEVGFPVTPQVPLRPMTY KAAVDLSHFLKEKGGLEGLIHSQRRQDILDLWIYHTQGYFPD

	Acid Sequence with mutation to enhance packaging in the virus core (G3C, V153L, and E177G mutations; underlined)	WQNYTPGPGVRYPLTFGWCYKLVPVEPDKLEEANKGENTSL LHPVSLHGMDDPGREVLEWRFDSRLAFHHVARELHPEYFKN C
SEQ ID NO:27	DNA polymerase I (DNA pol I)	VQIPQNPLILVDGSSYLYRAYHAFPPLTNSAGEPTGAMYGVLNMLR SLIMQYKPTHAAVVFDAKGKTFRDELFEHYKSHRPPMPDDLRAQIE PLHAMVKAMGLPLLAVSGVEADDVIGTLAREAEKAGRPVLISTGD KDMAQLVTPNITLINTMTNTILGPEEVVNKYGVPPELIIDFLALMGD SSDNIPGVPGVGEKTAQALLQGLGGLDTLYAEPEKIAGLSFRGAKT MAAKLEQNKEVAYLSYQLATIKTDVELELTCEQLEVQQPAAEELLG LFKKYEFKRWTADVEAGKWLQAKGAKPAAKPQETSVADEAPEVT ATVISYDNYVTILDEETLKAWIAKLEKAPVFAFDTETDSLDNISANL VGLSFAIEPGVAAYIPVAHDYLDAPDQISRERALELLKPLLEDEKAL KVGQNLKYDRGILANYGIELRGIAFDTMLESYILNSVAGRHDMDSL AERWLKHKTITFEEIAGKGKNQLTFNQIALEEAGRYAAEDADVTLQ LHLKMWPDLQKHKGPLNVFENIEMPLVPVLSRIERNGVKIDPKVLH NHSEELTLRLAELEKKAHEIAGEEFNLSSTKQLQTILFEKQGIKPLKK TPGGAPSTSEEVLEELALDYPLPKVILEYRGLAKLKSTYTDKLPLMI NPKTGRVHTSYHQAVTATGRLSSTDPNLQNIPVRNEEGRRIRQAFIA PEDYVIVSADYSQIELRIMAHLSRDKGLLTAFAEGKDIHRATAAEVF GLPLETVTSEQRRSAKAINFGLIYGMSAFGLARQLNIPRKEAQKYM DLYFERYPGVLEYMERTRAQAKEQGYVETLDGRRLYLPDIKSSNG ARRAAAERAAINAPMQGTAADIIKRAMIAVDAWLQAEQPRVRMIM QVHDELVFEVHKDDVDAVAKQIHQLMENCTRLDVPLLVEVGSGEN WDQAH
SEQ ID NO:28	Klenow fragment of DNA Pol I	APEVTATVISYDNYVTILDEETLKAWIAKLEKAPVFAFDTETDSLDN ISANLVGLSFAIEPGVAAYIPVAHDYLDAPDQISRERALELLKPLLED EKALKVGQNLKYDRGILANYGIELRGIAFDTMLESYILNSVAGRHD MDSLAERWLKHKTITFEEIAGKGKNQLTFNQIALEEAGRYAAEDAD VTLQLHLKMWPDLQKHKGPLNVFENIEMPLVPVLSRIERNGVKIDP KVLHNHSEELTLRLAELEKKAHEIAGEEFNLSSTKQLQTILFEKQGIK PLKKTPGGAPSTSEEVLEELALDYPLPKVILEYRGLAKLKSTYTDKL PLMINPKTGRVHTSYHQAVTATGRLSSTDPNLQNIPVRNEEGRRIRQ AFIAPEDYVIVSADYSQIELRIMAHLSRDKGLLTAFAEGKDIHRATA AEVFGLPLETVTSEQRRSAKAINFGLIYGMSAFGLARQLNIPRKEAQ KYMDLYFERYPGVLEYMERTRAQAKEQGYVETLDGRRLYLPDIKS SNGARRAAAERAAINAPMQGTAADIIKRAMIAVDAWLQAEQPRVR MIMQVHDELVFEVHKDDVDAVAKQIHQLMENCTRLDVPLLVEVGS GENWDQAH
SEQ ID NO:29	Cas9	MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAADKK YSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGAL LFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFF HRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDS TDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTY NQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFG NLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQ YADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLK EDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENE DILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGW GRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKE

		DIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMG RHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEH PVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQ SFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNA KLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILD SRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYH HAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQ EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIAR KKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGIT IMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRML ASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAE NIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGL YETRIDLSQLGGDKRPAATKKAGQAKKKK
SEQ ID NO: 30	Cpf1	MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDY KKAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQ KDFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLILWL KQSKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGFHENRKNV YSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKD LAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTII GGKFVNGENTKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSD TESKSFVIDKLEDDSDVVTTMQSFYEQIAAFKTVEEKSIKETLSLLFD DLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIA PKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQC RFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAE DDVKAIKDLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEE CYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNK EPDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVY KLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKG YEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFY REVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRP NLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQSIPKKITHPAKE AIANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANK FNDEINLLLKEKANDVHILSIDRGERHLAYYTLVDGKGNIIKQDTFNI IGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQV VHEIAKLVIEYNAIVVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEK LNYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAG FTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSF DYKNFGDKAAKGKWTIASFGSRLINFRNSDKNHNWDTREVYPTKE LEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNS KTGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLK GLMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRNN

What is claimed is:

1. A mammalian expression plasmid comprising a eukaryotic promoter operably linked to a non-viral nucleic acid sequence, wherein the non-viral nucleic acid sequence comprises:

    (i) a nucleic acid sequence encoding a fusion protein, wherein the fusion protein comprises:

        (a) a polypeptide comprises a DNA polymerase domain; and

        (b) a CRISPR-associated endonuclease coding sequence; and

    (ii) a guide RNA (gRNA) coding sequence, wherein the gRNA coding sequence comprises at least one aptamer coding sequence.

2. The mammalian expression plasmid of claim 1, wherein the CRISPR-associated endonuclease coding sequence encodes a Cas9 protein.

3. The mammalian expression plasmid of claim 1, wherein the polypeptide comprising the DNA polymerase domain is an *E. coli* DNA polymerase I (DNA PolI).

4. The mammalian expression plasmid of claim 1, wherein the polypeptide comprising the DNA polymerase domain is a Klenow fragment of DNA PolI.

5. The mammalian expression plasmid of claim 1, wherein the polypeptide comprising the DNA polymerase domain has reduced polymerase activity.

6. The mammalian expression plasmid of claim 1, wherein the at least one aptamer coding sequence encodes an aptamer sequence bound specifically by an ABP selected from the group consisting of MS2 coat protein, PP7 coat protein, lambda N RNA-binding domain, or Com protein.

7. The mammalian expression plasmid of claim 1, wherein the aptamer is an MS2 aptamer sequence or a com aptamer sequence.

8. The mammalian expression plasmid of claim 1, wherein the sgRNA coding sequence comprises at least one aptamer inserted into the tetraloop or the ST2 loop of the sgRNA coding sequence.

9. The mammalian expression plasmid of claim 8, wherein the the sgRNA coding comprises at least one com aptamer inserted into the ST2 loop of the gRNA coding sequence.

10. A lentiviral packaging system comprising:

    a) a packaging plasmid comprising a eukaryotic promoter operably linked to a Gag nucleotide sequence, wherein the Gag nucleotide sequence comprises a nucleocapsid (NC) coding sequence and a matrix protein (MA) coding sequence, wherein one or both of the NC coding sequence or the MA coding sequence comprises at least one non-viral aptamer-binding protein (ABP) nucleotide sequence, and wherein the packaging plasmid does not encode a functional integrase protein;

    b) at least one mammalian expression plasmid of any one of claims 1-9; and

    c) an envelope plasmid comprising an envelope glycoprotein coding sequence.

11. The lentiviral packaging system of claim 10, wherein the packaging plasmid further comprises a Rev nucleotide sequence and a Tat nucleotide sequence.

12. The lentiviral packaging system of claim 10, further comprising a second packaging plasmid comprising a Rev nucleotide sequence.

13. The lentiviral packaging system of claim 10, wherein the at least one non-viral ABP nucleotide sequence encodes MS2 coat protein, PP7 coat protein, lambda N peptide, or Com protein.

14. A lentiviral particle comprising:

    A) a fusion protein comprising a nucleocapsid (NC) protein or a matrix (MA) protein wherein the NC protein or MA protein comprises at least one non-viral aptamer binding protein (ABP); and

    B) a ribonucleotide protein (RNP) complex comprising:

(i) a nucleic acid sequence encoding a fusion protein, wherein the fusion protein comprises:

(a) a polypeptide comprising a DNA polymerase domain; and

(b) a CRISPR-associated endonuclease coding sequence; and

(ii) a guide RNA (gRNA) coding sequence, wherein the gRNA coding sequence comprises at least one aptamer coding sequence wherein the lentivirus-like particle does not comprise a functional integrase protein.

15. The lentiviral particle of claim 14, wherein the wherein the CRISPR-associated endonuclease coding sequence encodes a Cas9 protein.

16. The lentiviral particle of claim 14, wherein the the polypeptide comprising a DNA polymerase domain is an *E. coli* DNA polymerase I (DNA PolI).

17. The lentiviral particle of claim 14, wherein the polypeptide comprising a DNA polymerase domain is a Klenow fragment of DNA PolI).

18. The lentiviral particle of claim 14, wherein the polypeptide comprising a DNA polymerase domain has reduced polymerase activity.

19. A method of producing a lentiviral particle, the method comprising:

a) transfecting a plurality of eukaryotic cells with the packaging plasmid, the at least one mammalian expression plasmid, and the envelope plasmid of the system of any one of claims 10-13; and

b) culturing the transfected eukaryotic cells for sufficient time for lentiviral particles to be produced.

20. The method of claim 19, wherein the lentiviral particle comprises a ribonucleotide protein (RNP) complex comprising:

(i) nucleic acid sequence encoding a fusion protein, wherein the fusion protein comprises:

(a) a polypeptide comprising a DNA polymerase domain; and

(b) a CRISPR-associated endonuclease coding sequence; and

(ii) a guide RNA.

21. The method of claim 20, wherein the plurality of eukaryotic cells are mammalian cells.

22. A lentiviral particle made by the method of claim 19.

23. A method of modifying a genomic target sequence in a cell, the method comprising transducing a plurality of eukaryotic cells with a plurality of viral particles, wherein the plurality of viral particles comprise a lentivirus-like particle according to any one of claims 14-18, wherein the RNP complex binds to the genomic target sequence in genomic DNA of the cell, wherein the CRISPR-associated endonuclease cleaves the genomic target sequence to create a double-stranded break, thereby modifying the genomic target sequence.

24. The method of claim 23, wherein non-homologous end joining (NHEJ) increases in the cell, as compared to a cell modified by a CRISPR-associated endonuclease that is not fused to a DNA polymerase domain.

25. The method of claim 23, wherein the ratio of NHEJ to non-NHEJ increases in the cell.

26. The method of claim 25, wherein the non-NHEJ end joining is microhomology-mediated end joining (MMEJ) and/or single-stranded annealing (SSA).

27. The method of claim 23, wherein the number of on-target deletions is decreased in the cell.

28. The method of claim 27, wherein the number of on-target deletions greater than 500 base pairs in size is decreased in the cell.

29. The method of claim 23, wherein the ratio of on-target one base pair (1-bp) deletions to on-target deletions greater than 1-bp is increased in the cell.

30. The method of claim 29, wherein the ratio of on-target one base pair (1-bp) deletions to deletions greater than 500 base pairs is increased in the cell.

31. The method of claim 23, wherein the number of templated insertions (TIS) increases in the cell.

32. The method of claim 31, wherein the ratio of TIS to non-TIS increases in the cell.

33. The method of claim 23, wherein the plurality of eukaryotic cells are mammalian cells.

34. The method of claim 23, wherein the plurality of eukaryotic cells are cells present in subject.

35. The method of claim 34, wherein the subject is a human subject.

36. The method of claim 35, wherein the subject is injected with the plurality of viral particles.

37. A cell containing the plasmid of claim 1.

38. A cell containing the lentiviral packaging system of claim 10.

39. A cell containing the lentiviral particle of claim 14.

40. A cell modified using the method of claim 23.

41. A method for treating a disease in a subject comprising:

a) obtaining cells from the subject;

b) modifying the cells of the subject using the method of any one of claims 23-36; and

c) administering the modified cells to the subject.

42. The method of claim 41, wherein the disease is cancer.

43. The method of claim 41, wherein the disease is Duchenne muscular dystrophy.

44. The method of claim 41, wherein the cells are T cells.

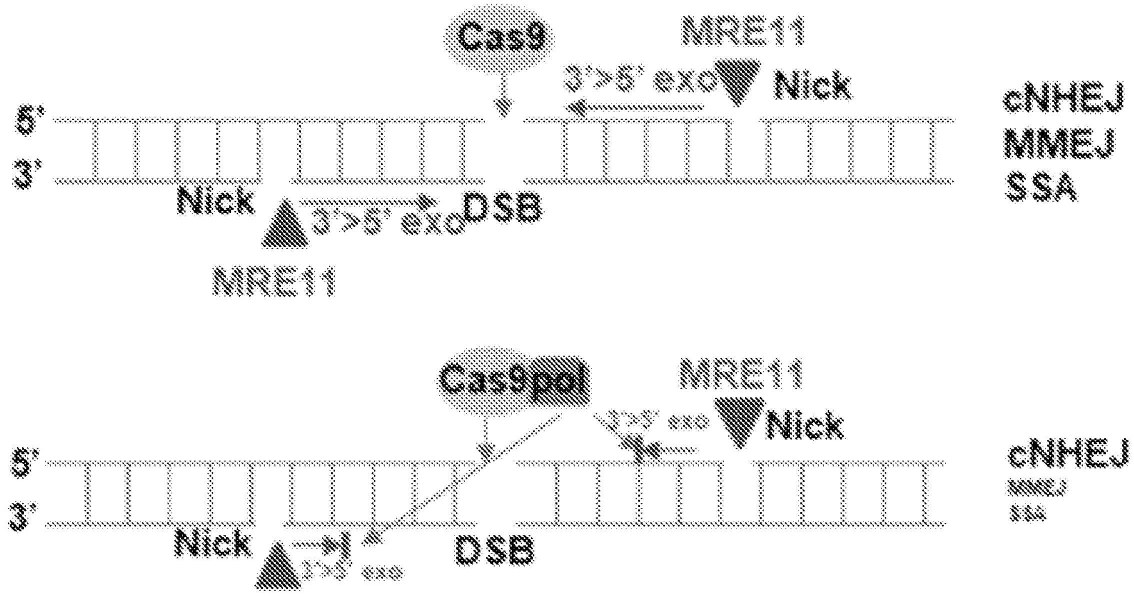


FIG. 1A

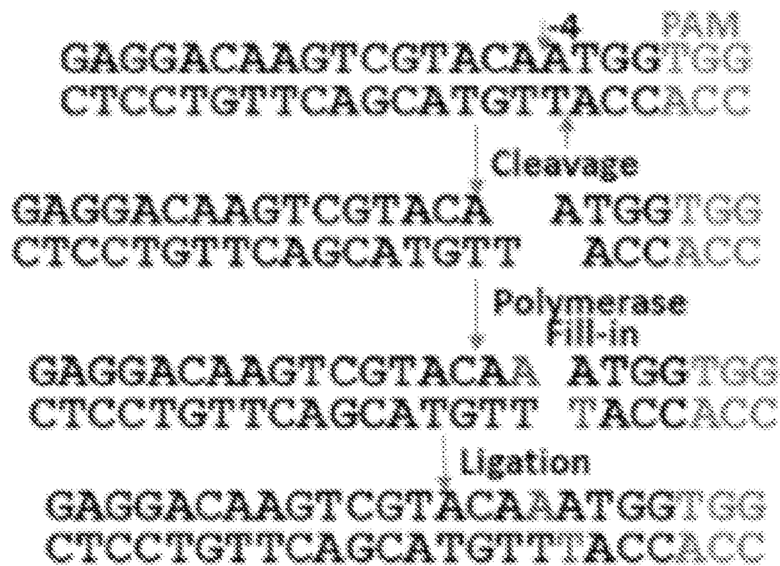


FIG. 1B

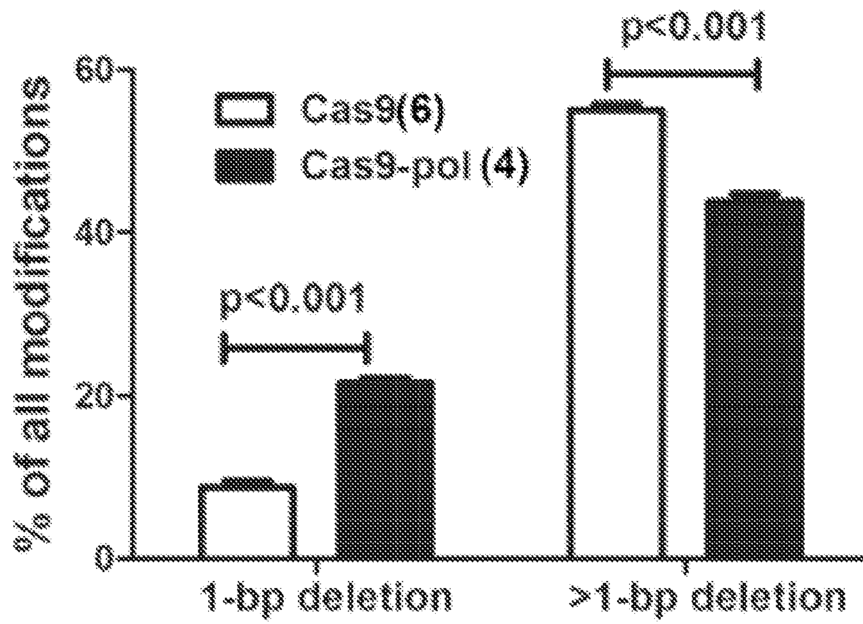


FIG. 1C

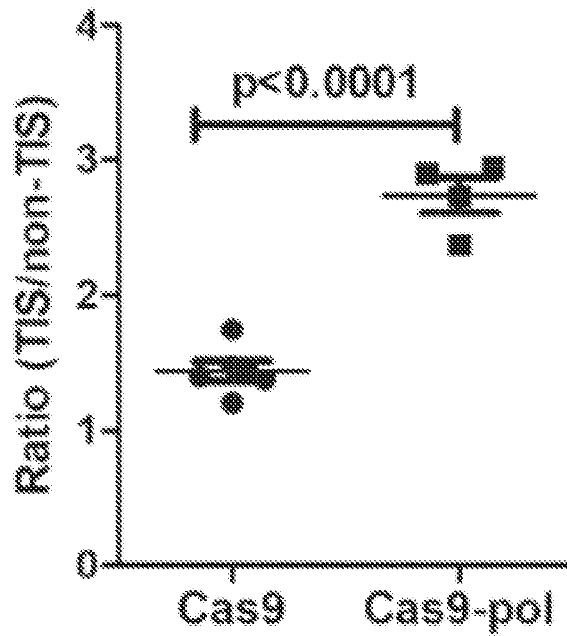


FIG. 1D

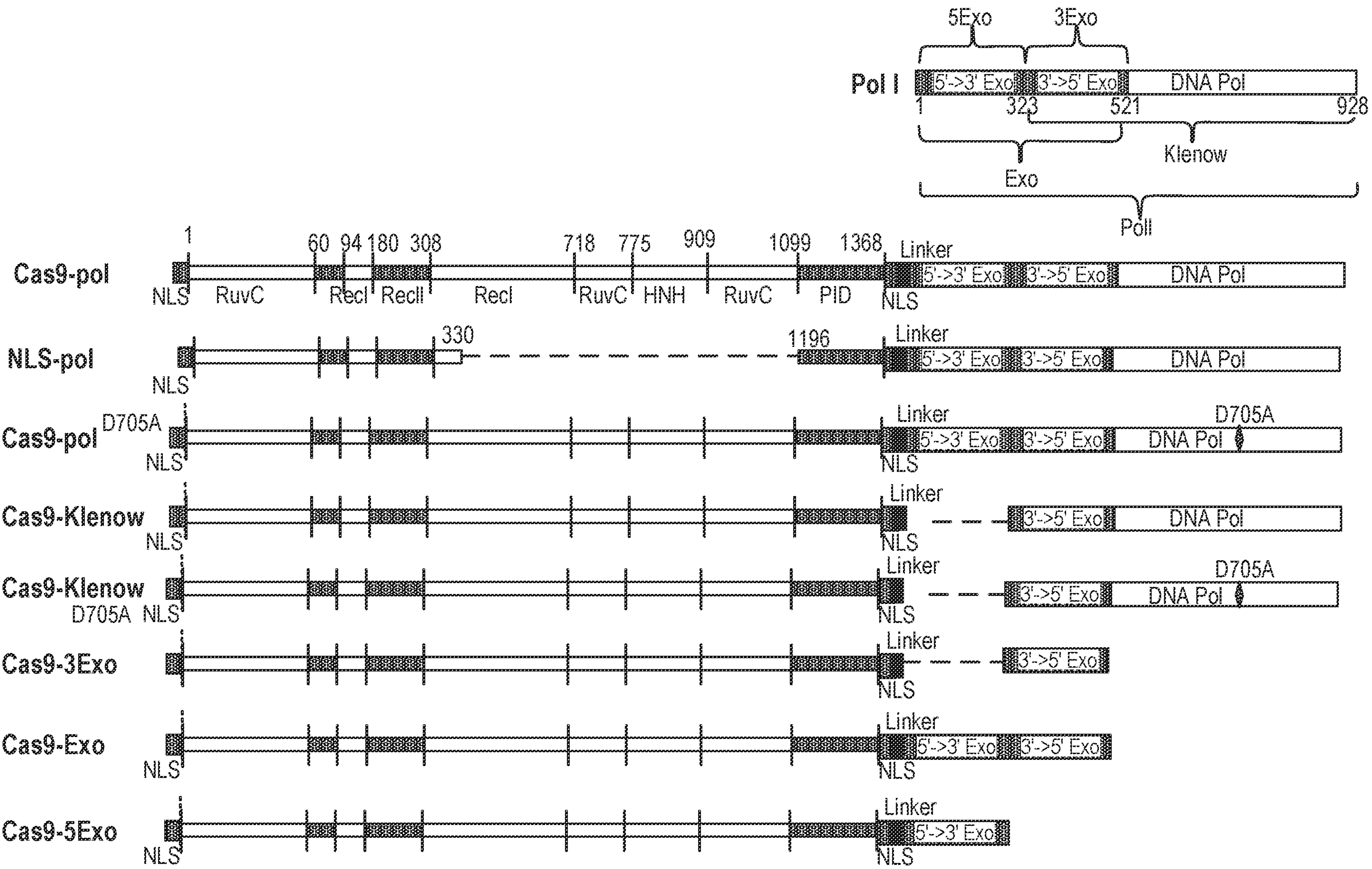


FIG. 2A

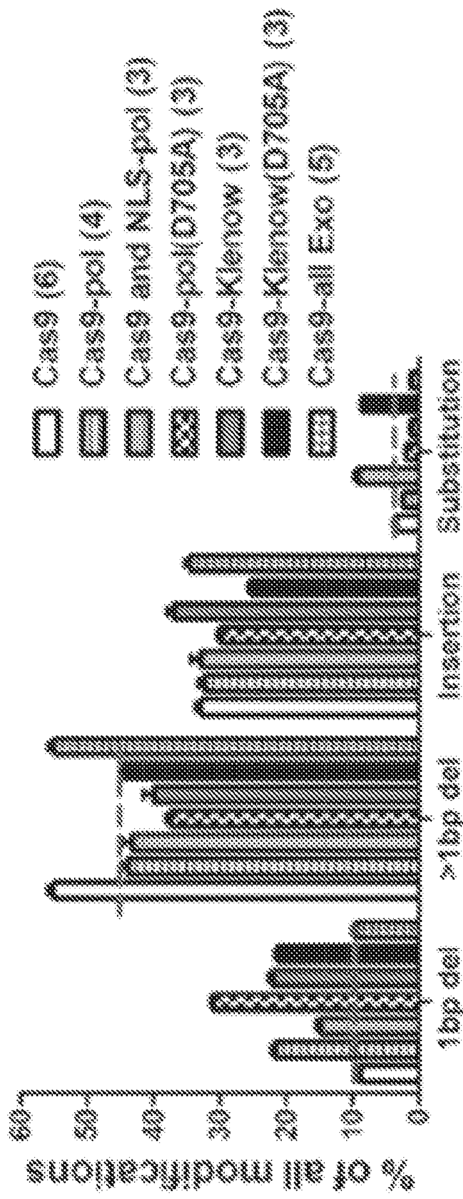


FIG. 2B

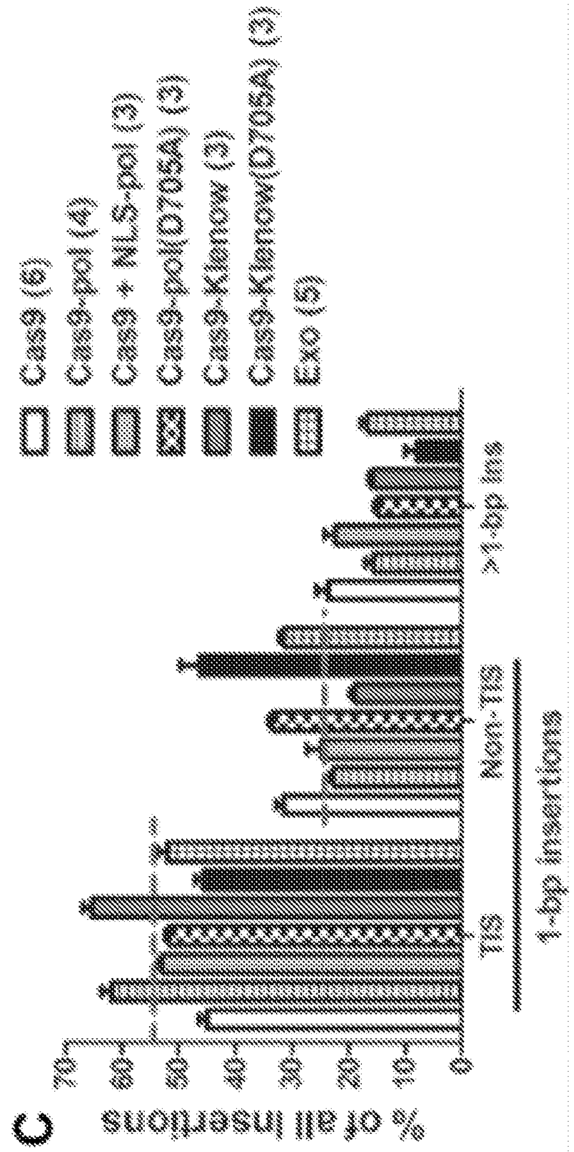


FIG. 2C

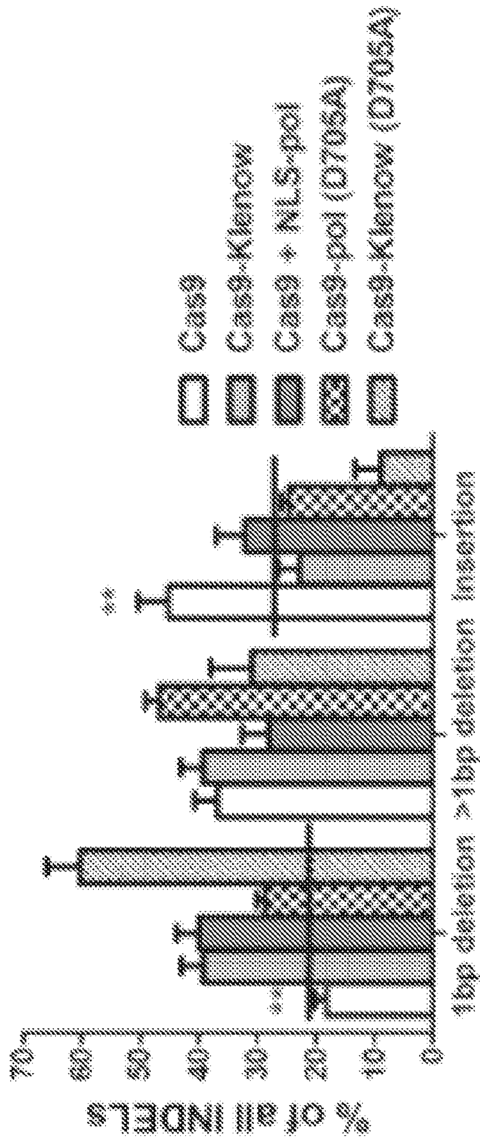


FIG 2D

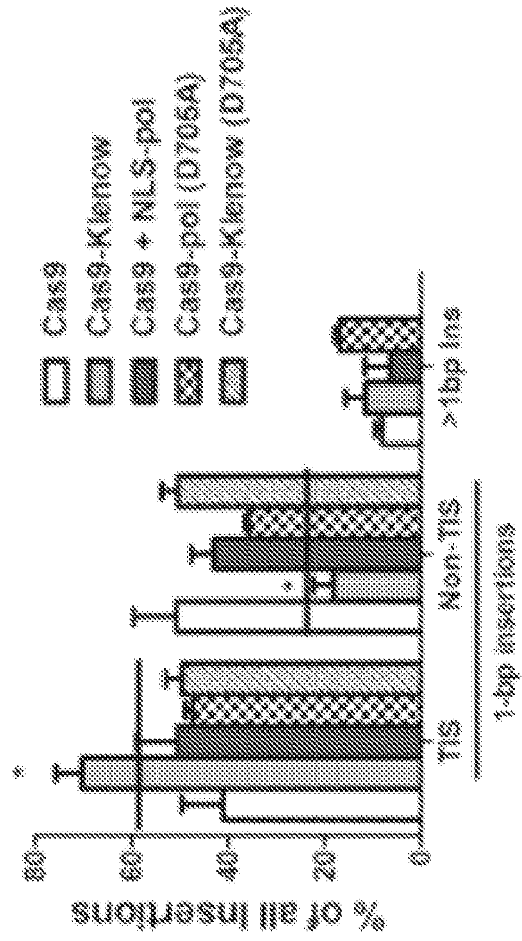


FIG. 2E

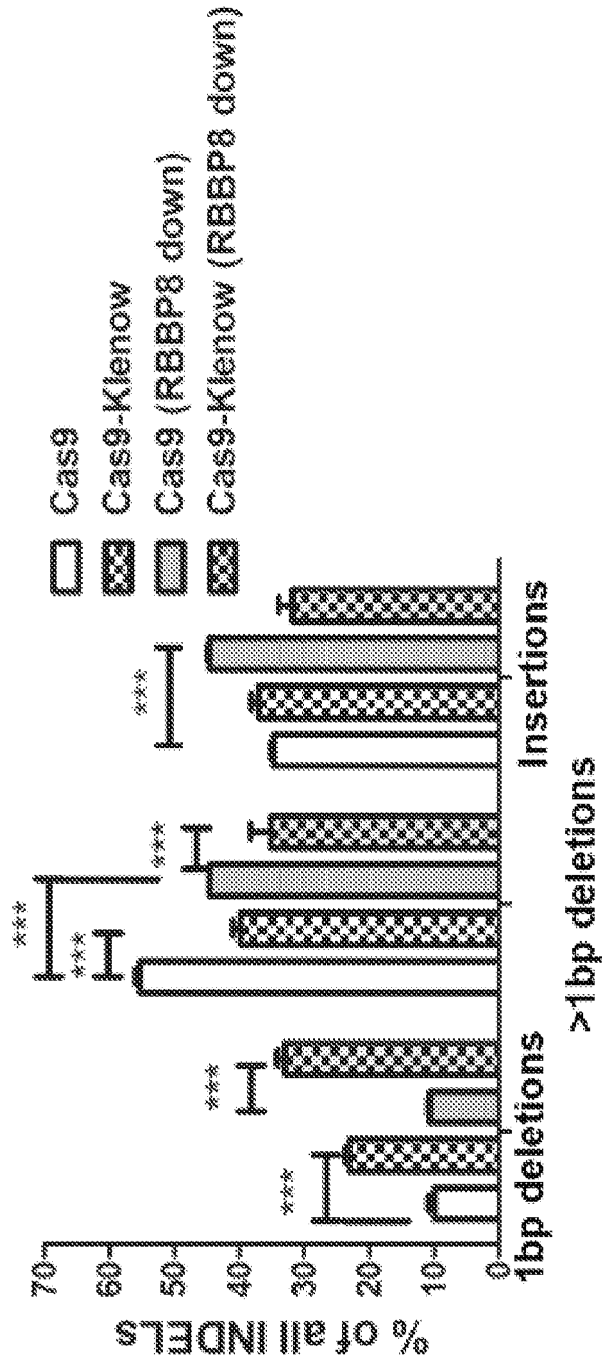


FIG. 3A

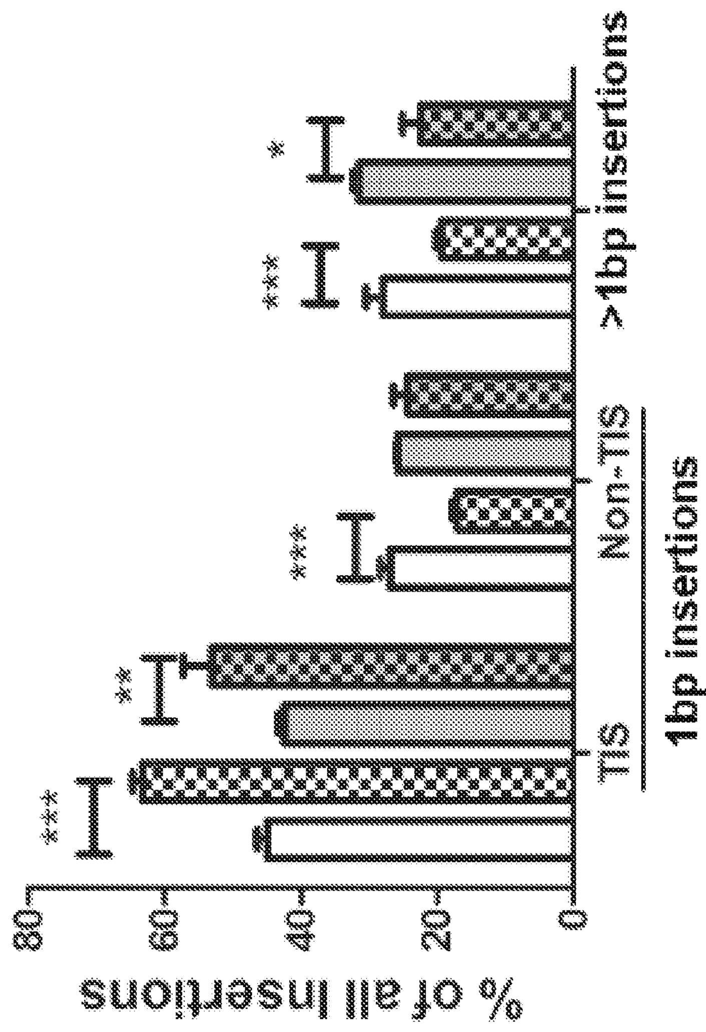


FIG. 3B

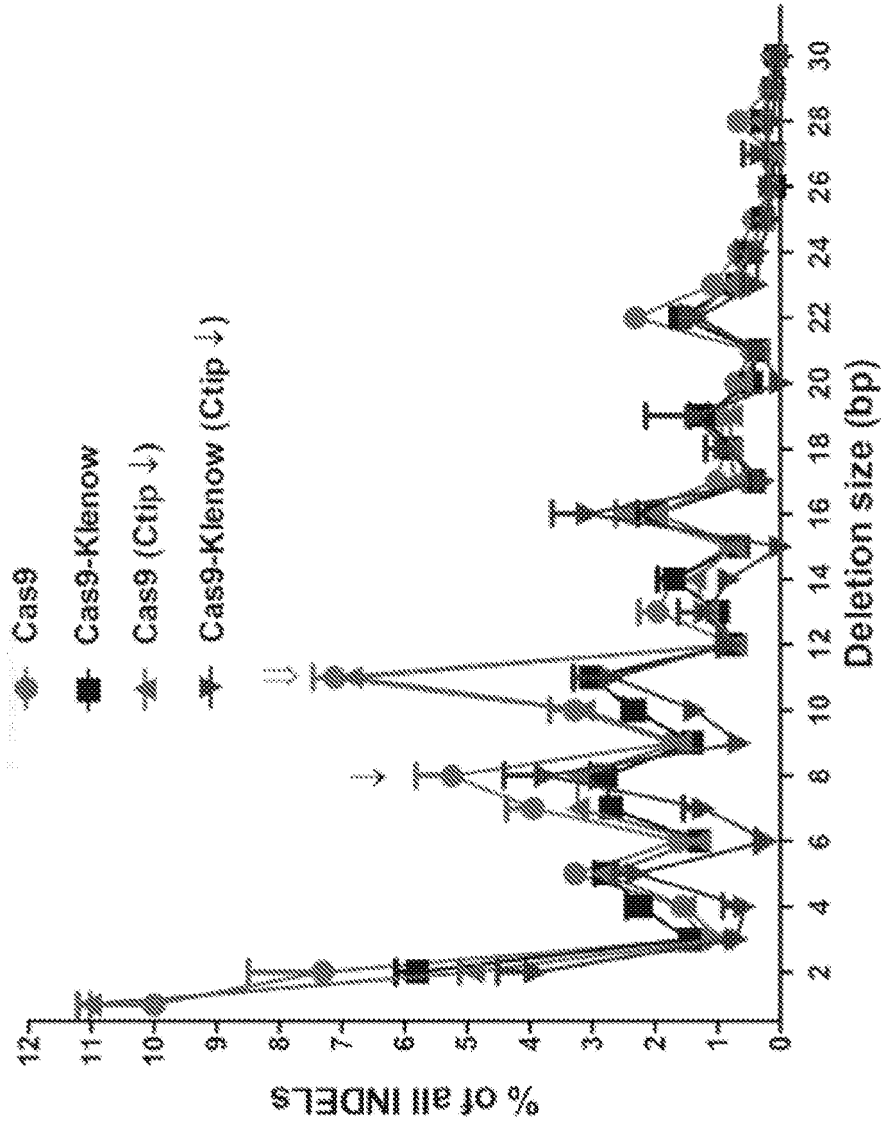


FIG. 3C

sgRNA PAM

	Sequence	% of all modifications	
Wild type	A C A G A G G A C A A G T C G T A C A A T G G T G G A G G A A T A G G T T C T T -		
11bp deletion	A C A G A G G A C A A G T C G T A C A A - - - - - - - - - - - T A G G T T C T T -	**5.5%**	(512 reads)
11bp deletion	A C A G A G G A C A A G T - - - - - - - - - - - G G A G G A A T A G G T T C T T -	**1.8%**	(167 reads)
8 bp deletion	A C A G A G G A C A A G T C G T - - - - - - - - G G A G G A A T A G G T T C T T -	**1.4%**	(135 reads)
8 bp deletion	A C A G A G G A C A A G - - - - - - - - T G G T G G A G G A A T A G G T T C T T -	**1.3%**	(125 reads)
16 bp deletion	A C A G A G G - - - - - - - - - - - - - - - - T G G A G G A A T A G G T T C T T -	**2.4%**	(226 reads)

^ ^ ^ ^ ^ ^ * * ^ ^ ^ ^ * *

FIG. 3D

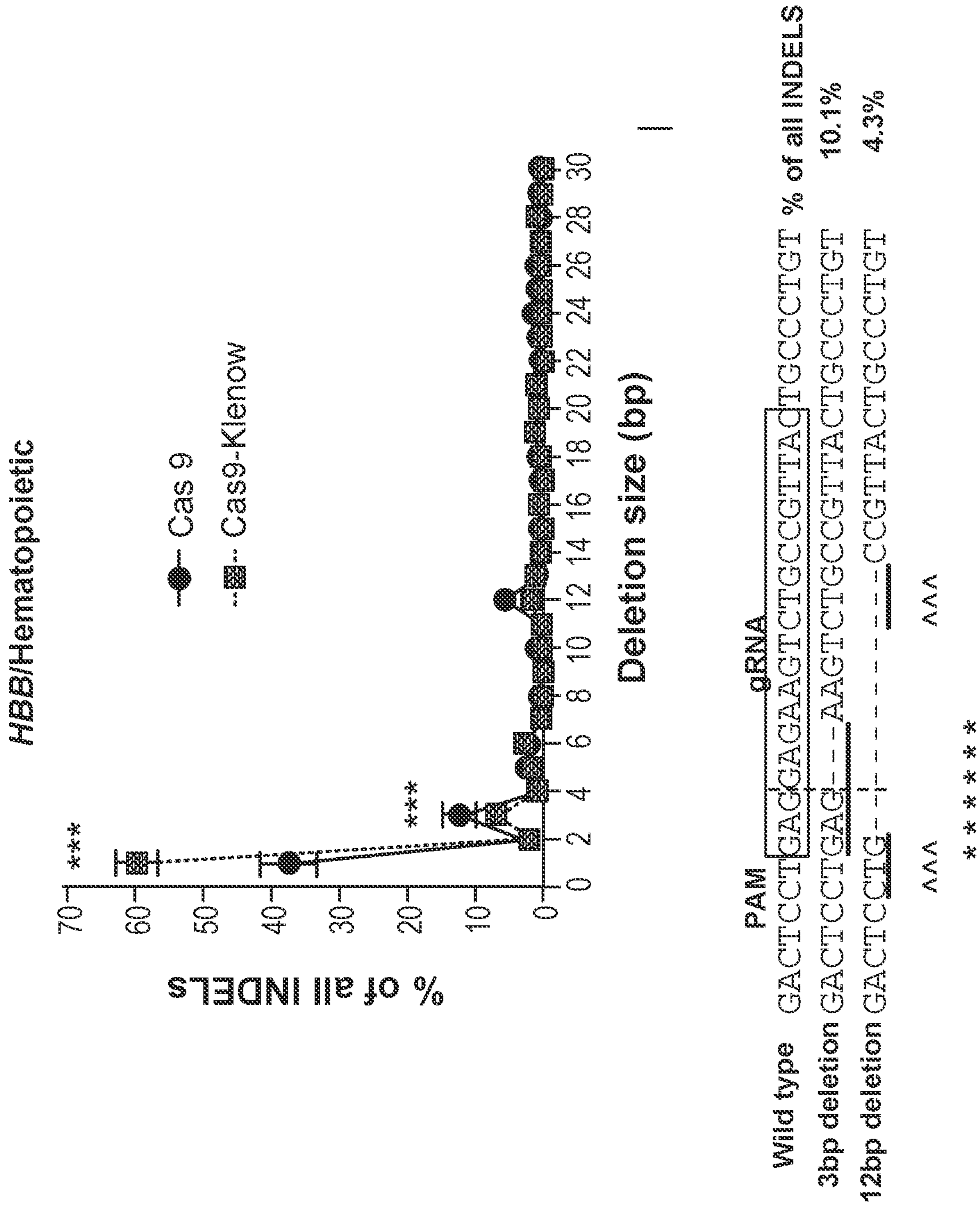


FIG. 4A

DMD53/293T

% of all INDELs

Deletion size (bp)

Cas 9

Cas9-Klenow

***

***

	PAM / gRNA	% of all INDELS
Wild type	GAAGTACAAGAACACCTTCAGAACCGGAGGCAACAGTTGA	
11bp deletion	GAAGTACAA-----------GAACCGGAGGCAACAGTTGA	13.2%
9bp deletion	GAAGTACAAGAACACC---------GGAGGCAACAGTTGA	8.5%
6bp deletion	GAAGTACAAGAACACCT------CCGGAGGCAACAGTTGA	2.5%

^ ^ ^ ^ ^ ^

* * * * * * * *

FIG. 4B

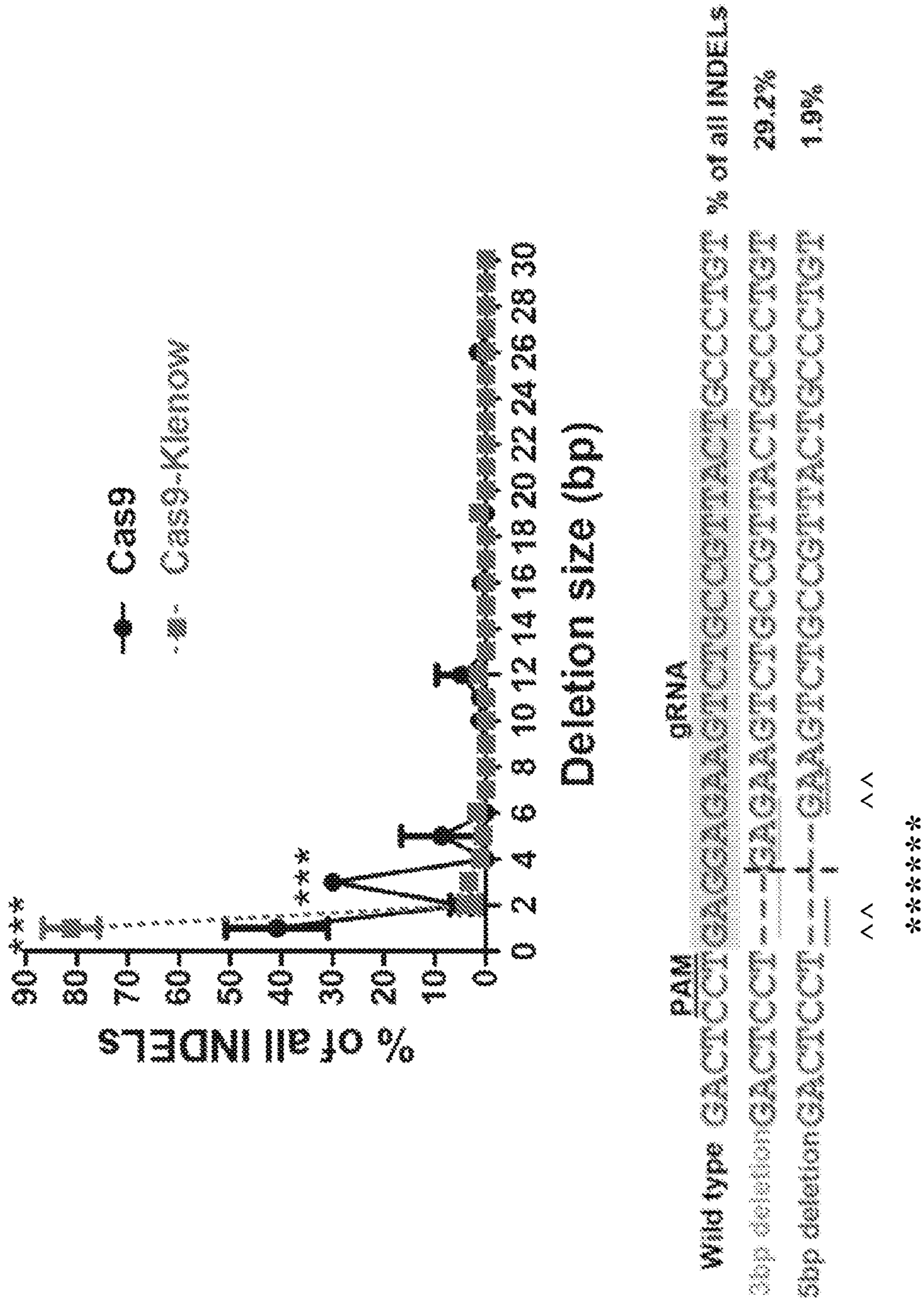
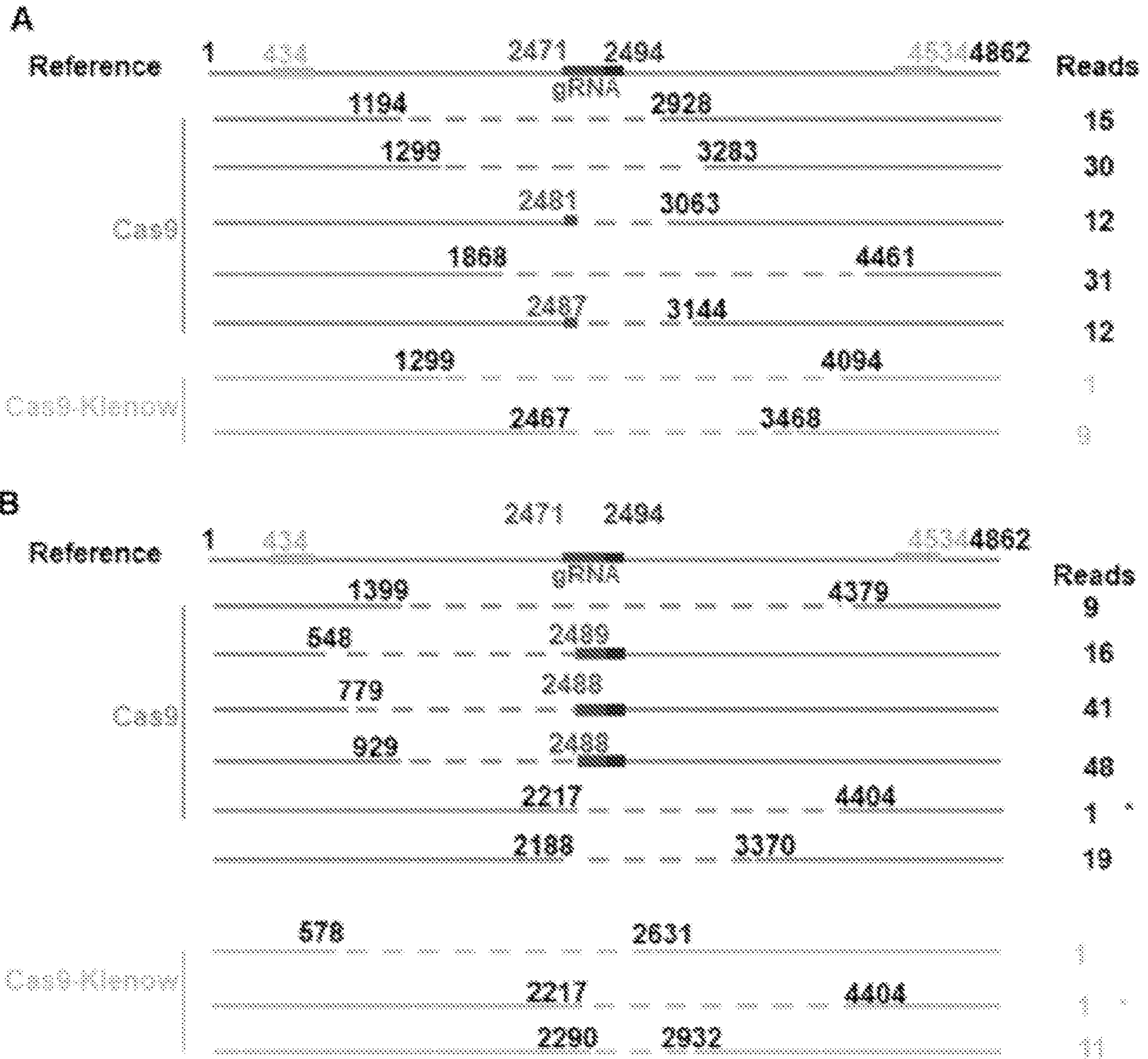


FIG. 4C

DMD44/Myoblast

FIG. 4D



FIGS. 5A-5B

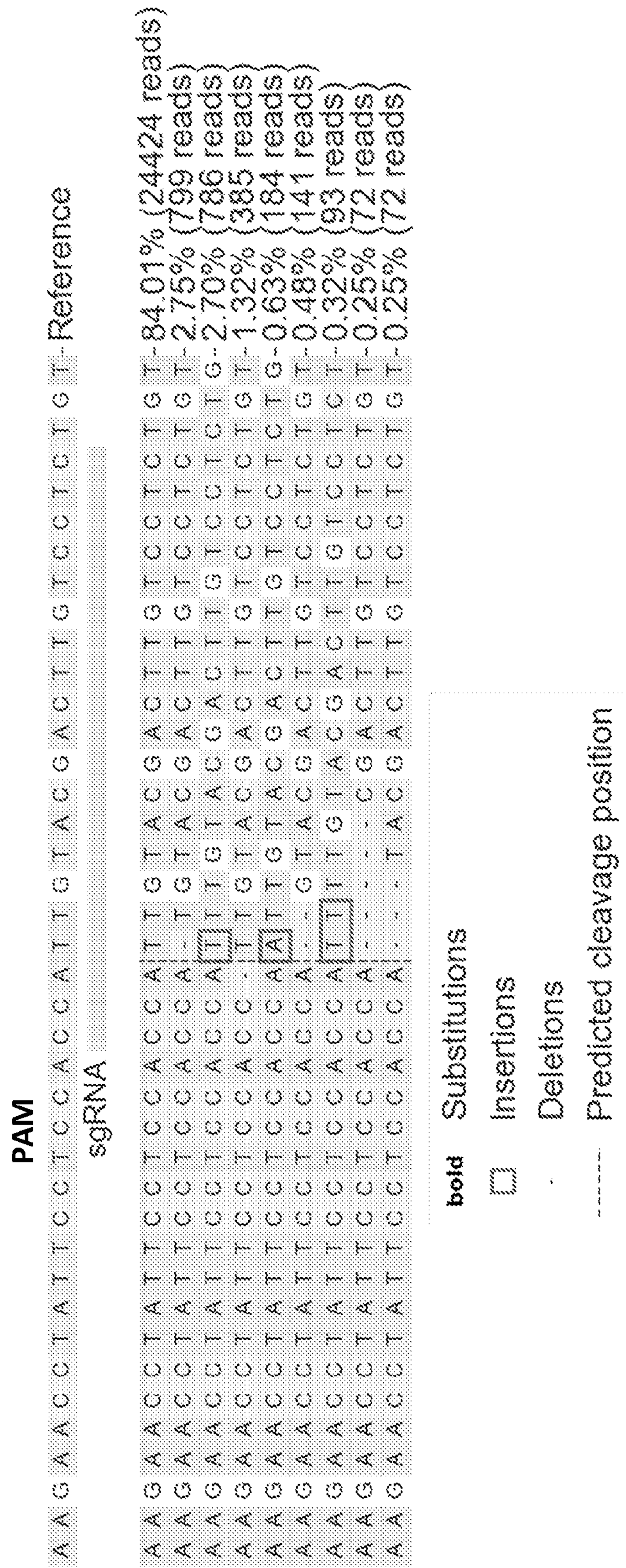


FIG. 6

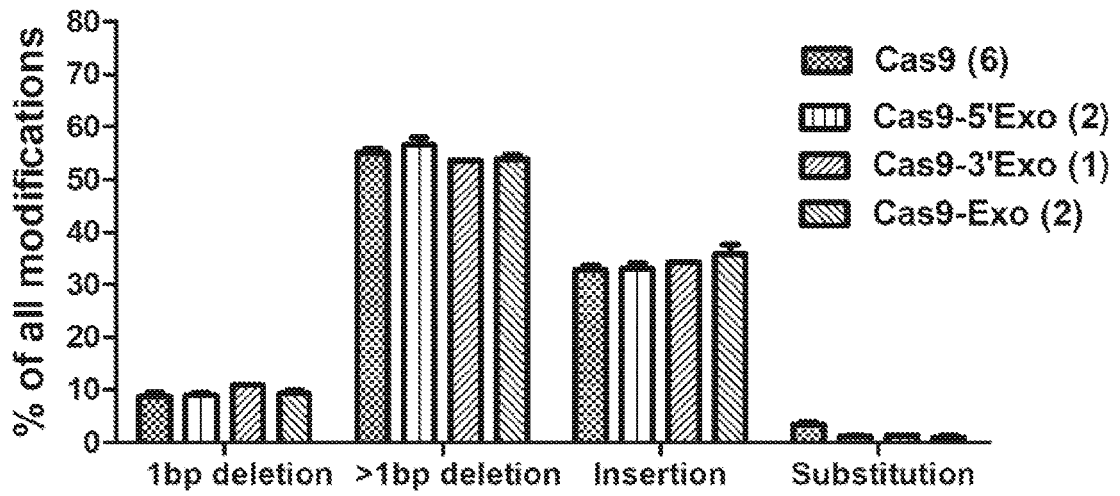


FIG. 7

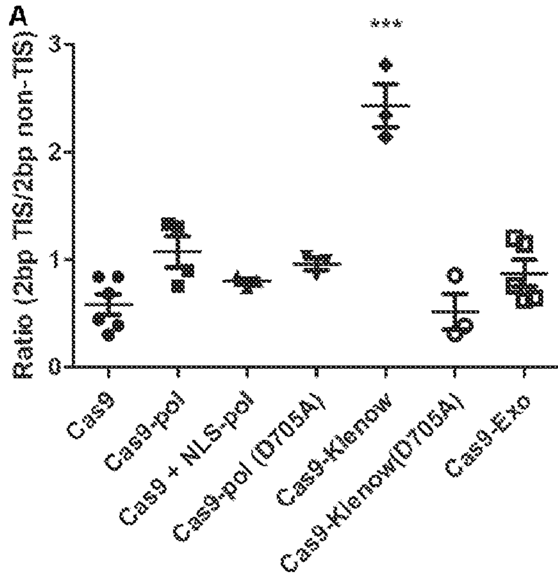


FIG. 8A

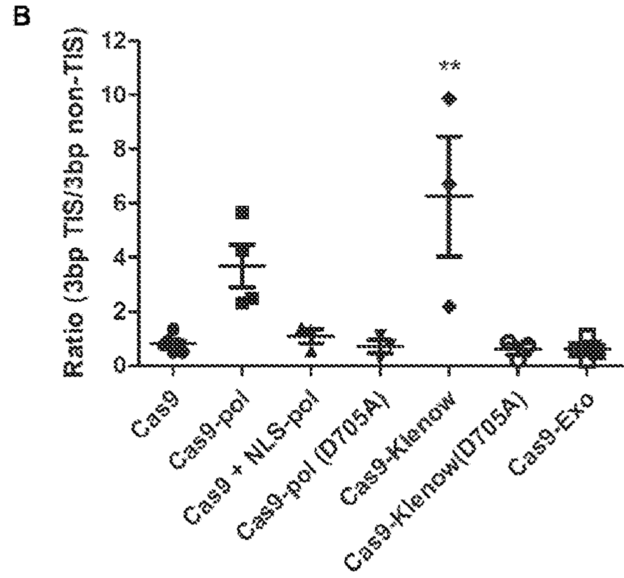


FIG. 8B

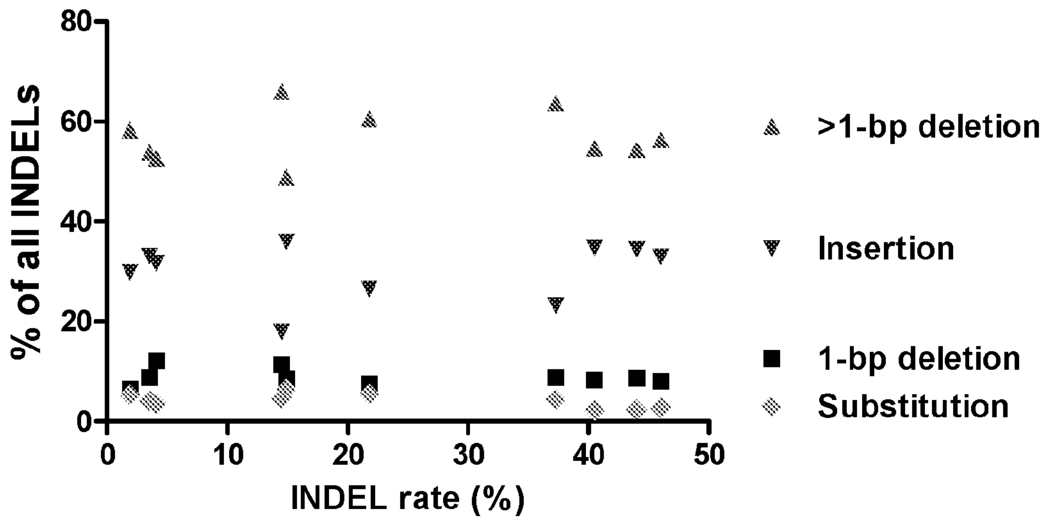


FIG. 9

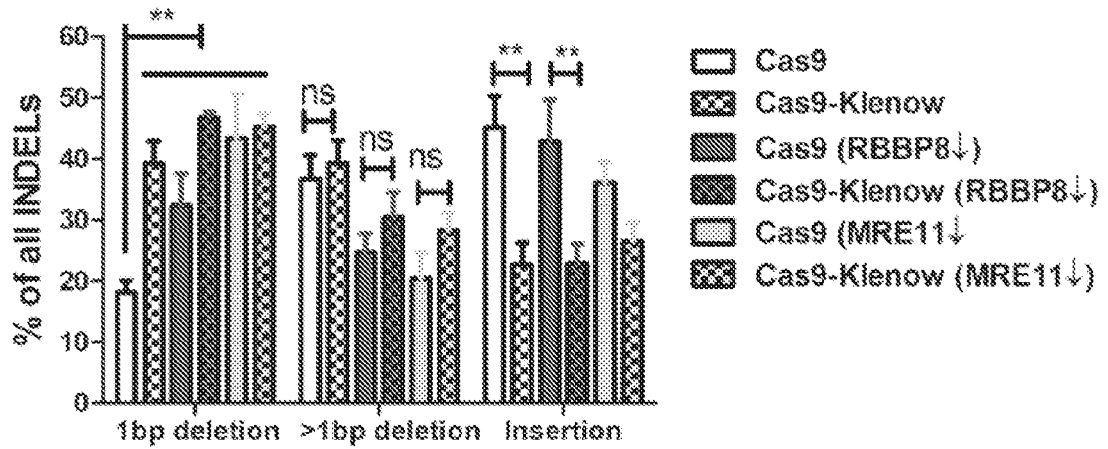


FIG. 10A

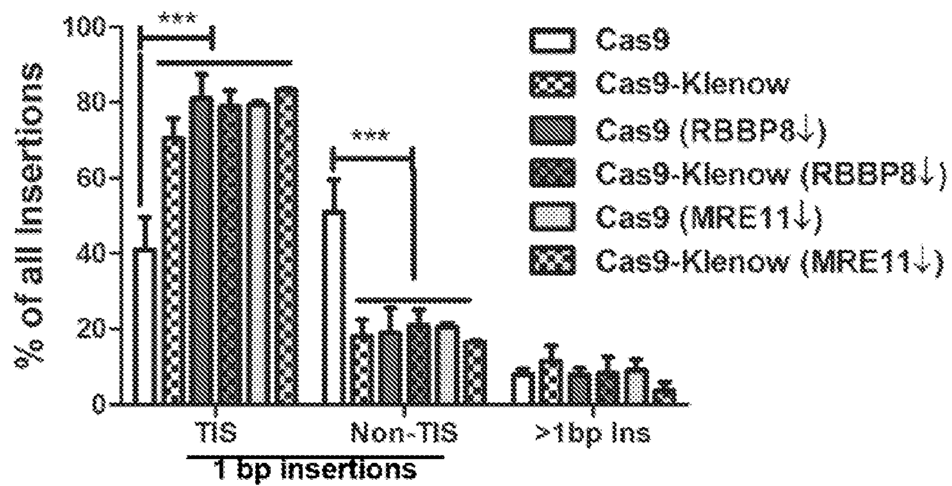


FIG. 10B

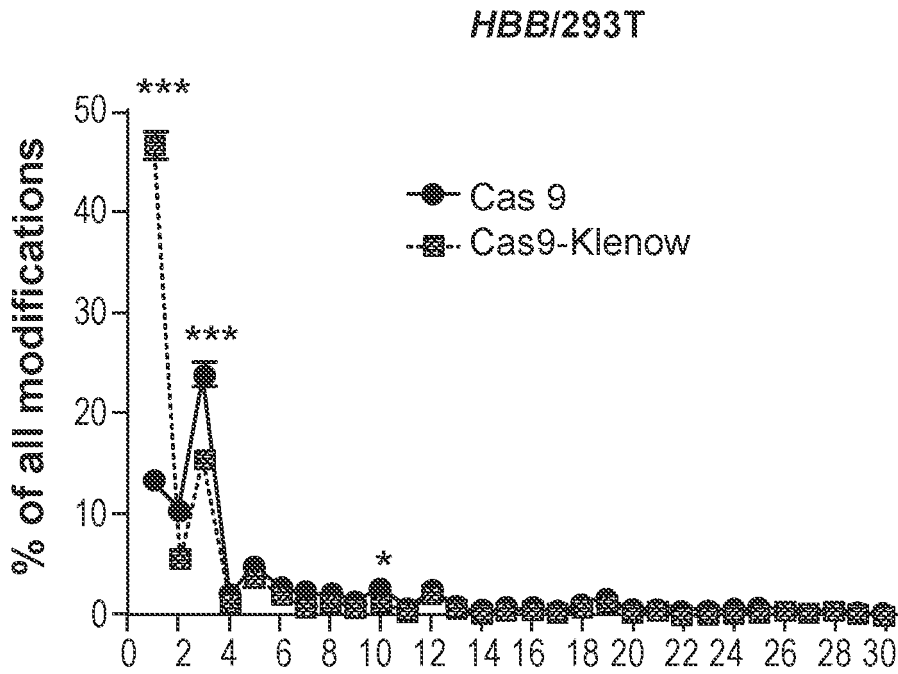
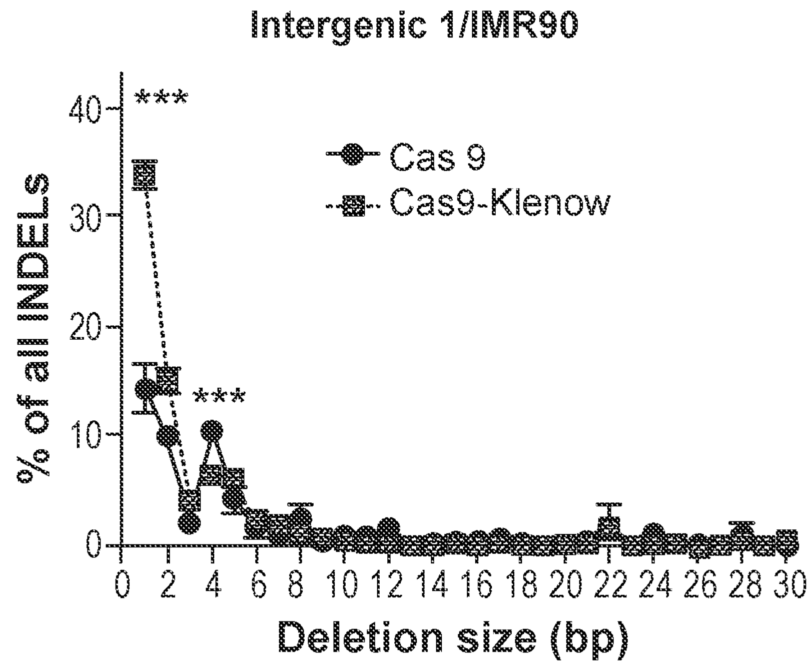
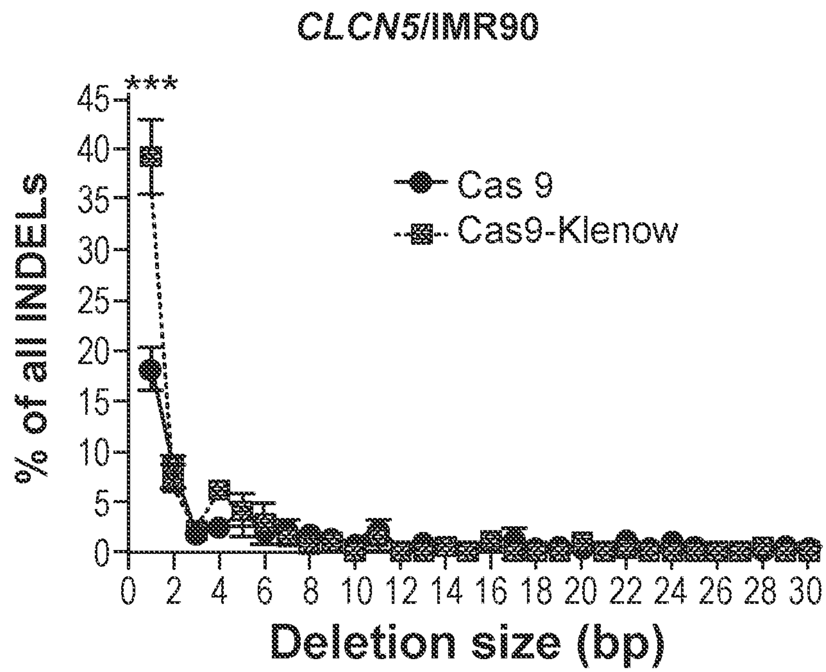


FIG. 11A



	sgRNA / PAM	
Wild type	TGAGGGTGAGGGATGAGATAATGATGAGTCAGGGCTTCAG	% of all INDELS
4bp deletion	TGAGGGTGAGGGATGAGATAATGA----TCAGGGCTTCAG	3.46%

FIG. 11B



Cas9 treated gRNA PAM

		% of all INDELS
Wild type	CAATACAGAGGACAAGTCGTACAATGGTGGAGGAATAGGTTCTTCAA	
5bp deletion	CAATACAGAGGACAAGTCG-----TGGTGGAGGAATAGGTTCTTCAA	**3.0%**
5bp deletion	CAATACAGAGGACAAGTCGTACAA-----GAGGAATAGGTTCTTCAA	**2.3%**

FIG. 11C

This XML file does not appear to have any style information associated with it. The document tree is shown below.


```xml
<ST26SequenceListing dtdVersion="V1_3" fileName="095199-1375708.xml" softwareName="WIPO Sequence" softwareVersion="2.2.0" productionDate="2023-03-07">
  <ApplicationIdentification>
    <IPOfficeCode>US</IPOfficeCode>
    <ApplicationNumberText/>
    <FilingDate/>
  </ApplicationIdentification>
  <ApplicantFileReference> 095199-1375708 </ApplicantFileReference>
  <EarliestPriorityApplicationIdentification>
    <IPOfficeCode>US</IPOfficeCode>
    <ApplicationNumberText>63/317,720</ApplicationNumberText>
    <FilingDate>2022-03-08</FilingDate>
  </EarliestPriorityApplicationIdentification>
  <ApplicantName languageCode="en">Wake Forest University Health Sciences</ApplicantName>
  <InventionTitle languageCode="en">COMPOSITIONS, SYSTEMS AND METHODS FOR EUKARYOTIC GENE EDITING</InventionTitle>
  <SequenceTotalQuantity>152</SequenceTotalQuantity>
  <SequenceData sequenceIDNumber="1">
    <INSDSeq>
      <INSDSeq_length>351</INSDSeq_length>
      <INSDSeq_moltype>DNA</INSDSeq_moltype>
      <INSDSeq_division>PAT</INSDSeq_division>
      <INSDSeq_feature-table>
        <INSDFeature>
          <INSDFeature_key>source</INSDFeature_key>
          <INSDFeature_location>1..351</INSDFeature_location>
          <INSDFeature_quals>
            <INSDQualifier>
              <INSDQualifier_name>mol_type</INSDQualifier_name>
              <INSDQualifier_value>other DNA</INSDQualifier_value>
            </INSDQualifier>
            <INSDQualifier id="q2">
              <INSDQualifier_name>organism</INSDQualifier_name>
              <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
            </INSDQualifier>
          </INSDFeature_quals>
        </INSDFeature>
      </INSDSeq_feature-table>
      <INSDSeq_sequence>atggcttctaactttactcagttcgttctcgtcgacaatggcggaactggcgacgtgactgtcgccccaagcaacttcgctaacgggatcgctgaatggatcagctctaactcgcgttcacaggcttaca
    </INSDSeq>
  </SequenceData>
  <SequenceData sequenceIDNumber="2">
    <INSDSeq>
      <INSDSeq_length>117</INSDSeq_length>
      <INSDSeq_moltype>AA</INSDSeq_moltype>
      <INSDSeq_division>PAT</INSDSeq_division>
      <INSDSeq_feature-table>
        <INSDFeature>
          <INSDFeature_key>source</INSDFeature_key>
          <INSDFeature_location>1..117</INSDFeature_location>
          <INSDFeature_quals>
            <INSDQualifier>
              <INSDQualifier_name>mol_type</INSDQualifier_name>
              <INSDQualifier_value>protein</INSDQualifier_value>
            </INSDQualifier>
            <INSDQualifier id="q4">
              <INSDQualifier_name>organism</INSDQualifier_name>
              <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
            </INSDQualifier>
          </INSDFeature_quals>
        </INSDFeature>
      </INSDSeq_feature-table>
      <INSDSeq_sequence>MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY</INSDSeq_se
    </INSDSeq>
  </SequenceData>
  <SequenceData sequenceIDNumber="3">
    <INSDSeq>
      <INSDSeq_length>363</INSDSeq_length>
      <INSDSeq_moltype>DNA</INSDSeq_moltype>
      <INSDSeq_division>PAT</INSDSeq_division>
      <INSDSeq_feature-table>
        <INSDFeature>
          <INSDFeature_key>source</INSDFeature_key>
          <INSDFeature_location>1..363</INSDFeature_location>
          <INSDFeature_quals>
            <INSDQualifier>
              <INSDQualifier_name>mol_type</INSDQualifier_name>
              <INSDQualifier_value>other DNA</INSDQualifier_value>
            </INSDQualifier>
            <INSDQualifier id="q6">
              <INSDQualifier_name>organism</INSDQualifier_name>
              <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
            </INSDQualifier>
          </INSDFeature_quals>
        </INSDFeature>
      </INSDSeq_feature-table>
      <INSDSeq_sequence>tccaaaacaatagtcctctccgtaggggaggcaacacggactttgaccgaaatccagtcaaccgctgaccgacaaatctttgaagagaaagtagggcctcttgtgggccgactgcgcttgactgcaagct
    </INSDSeq>
  </SequenceData>
  <SequenceData sequenceIDNumber="4">
    <INSDSeq>
      <INSDSeq_length>121</INSDSeq_length>
      <INSDSeq_moltype>AA</INSDSeq_moltype>
      <INSDSeq_division>PAT</INSDSeq_division>
      <INSDSeq_feature-table>
        <INSDFeature>
```

        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..121</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>protein</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q8">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>SKTIVLSVGEATRTLTEIQSTADRQIFEEKVGPLVGRLRLTASLRQNGAKTAYRVNLKLDQADVVDSGLPKVRYTQVWSHDVTIVANSTEASRKSLYDLTKSLVATSQVEDLVVNLVPLGR</INSDSe
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="5">
  <INSDSeq>
    <INSDSeq_length>66</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..66</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q10">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>atggatgcacaaacacgccgccgcgaacgtcgcgcagagaaacaggctcaatggaaagcagcaaat</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="6">
  <INSDSeq>
    <INSDSeq_length>22</INSDSeq_length>
    <INSDSeq_moltype>AA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..22</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>protein</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q12">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>MDAQTRRRERRAEKQAQWKAAN</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="7">
  <INSDSeq>
    <INSDSeq_length>193</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..193</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q14">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>atgaaatcaattcgctgtaaaaactgcaacaaactgttatttaaggcggattcctttgatcacattgaaatcaggtgtccgcgttgcaaacgtcacatcataatgctgaatgcctgcgagcatcccacg
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="8">
  <INSDSeq>
    <INSDSeq_length>62</INSDSeq_length>
    <INSDSeq_moltype>AA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..62</INSDFeature_location>

```
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>protein</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q16">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>MKSIRCKNCNKLLFKADSFDHIEIRCPRCKRHIIMLNACEHPTEKHCGKREKITHSDETVRY</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="9">
  <INSDSeq>
    <INSDSeq_length>19</INSDSeq_length>
    <INSDSeq_moltype>RNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..19</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other RNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q18">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>acatgaggatcacccatgt</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="10">
  <INSDSeq>
    <INSDSeq_length>19</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..19</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q20">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>acatgaggatcacccatgt</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="11">
  <INSDSeq>
    <INSDSeq_length>25</INSDSeq_length>
    <INSDSeq_moltype>RNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..25</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other RNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q22">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>ggagcagacgatatggcgtcgctcc</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="12">
  <INSDSeq>
    <INSDSeq_length>25</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..25</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
```

```
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q24">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>ggagcagacgatatggcgtcgctcc</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="13">
<INSDSeq>
  <INSDSeq_length>19</INSDSeq_length>
  <INSDSeq_moltype>RNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..19</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other RNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q26">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>gggccctgaagaagggccc</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="14">
<INSDSeq>
  <INSDSeq_length>19</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..19</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q28">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>gggccctgaagaagggccc</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="15">
<INSDSeq>
  <INSDSeq_length>19</INSDSeq_length>
  <INSDSeq_moltype>RNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..19</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other RNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q30">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>ctgaatgcctgcgagcatc</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="16">
<INSDSeq>
  <INSDSeq_length>18</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..18</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
```

```
          </INSDQualifier>
          <INSDQualifier id="q32">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>ctgaatgcctgcgagcat</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="17">
  <INSDSeq>
    <INSDSeq_length>132</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..132</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q34">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>gctcgctttcttgctgtccaatttctattaaaggttcctttgttccctaagtccaactactaaactgggggatattatgaagggccttgagcatctggattctgcctaataaaaaacatttattttcat
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="18">
  <INSDSeq>
    <INSDSeq_length>132</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..132</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q36">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>gctcgctttcttgctgtccaatttctattaaaggttcctttgttccctaagtccaactactaaactgggggatattatgaagggccttgagcatctggattctgcctaataaaaaacatttattttcat
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="19">
  <INSDSeq>
    <INSDSeq_length>165</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..165</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q38">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>atacagaaaggcaattttaggaaccaaagaaagactgttaagtgtttcaattgtggcaaagaagggcacatagccaaaaattgcagggcccctaggaaaaagggctgttggaaatgtggaaaggaagga
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="20">
  <INSDSeq>
    <INSDSeq_length>55</INSDSeq_length>
    <INSDSeq_moltype>AA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..55</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>protein</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q40">
```

<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>IQKGNFRNQRKTVKCFNCGKEGHIAKNCRAPRKKGCWKCGKEGHQMKDCTERQAN</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="21">
<INSDSeq>
<INSDSeq_length>396</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..396</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q42">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>atgggtgcgagagcgtcagtattaagcgggggagaattagatcgatgggaaaaaattcggttaaggccagggggaaagaaaaaatataaattaaaacatatagtatgggcaagcagggagctagaacga
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="22">
<INSDSeq>
<INSDSeq_length>131</INSDSeq_length>
<INSDSeq_moltype>AA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..131</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>protein</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q44">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>GARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEIKDTKEALDKIEEEQNKSKKKAQQAAADTGHSNQVSQ
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="23">
<INSDSeq>
<INSDSeq_length>291</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..291</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q46">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>atggaacaagccccagaagaccagggaccgcagagggaaccatacaatgaatggacactagaacttttagaggaactcaagcgggaagcagtcagacactttcctagaccatggcttcatggcttagga
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="24">
<INSDSeq>
<INSDSeq_length>96</INSDSeq_length>
<INSDSeq_moltype>AA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..96</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>protein</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q48">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>

```
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>MEQAPEDQGPQREPYNEWTLELLEELKREAVRHFPRPWLHGLGQHIYETYGDTWTGVEAIIRILQRLLFVHFRIGCQHSRIGILRQRRARNGASRS</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="25">
<INSDSeq>
  <INSDSeq_length>618</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..618</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q50">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>atgggttgcaagtggtcaaaaagtagtgtgattggatggcctgctgtaagggaaagaatgagacgagctgagccagcagcagatggggtgggagcagtatctcgagacctagaaaaacatggagcaatca
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="26">
<INSDSeq>
  <INSDSeq_length>206</INSDSeq_length>
  <INSDSeq_moltype>AA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..206</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>protein</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q52">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>MGCKWSKSSVIGWPAVRERMRRAEPAADGVGAVSRDLEKHGAITSSNTAANNAACAWLEAQEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDLWIYHTQGYFPDWQNYTP
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="27">
<INSDSeq>
  <INSDSeq_length>927</INSDSeq_length>
  <INSDSeq_moltype>AA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..927</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>protein</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q54">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>VQIPQNPLILVDGSSYLYRAYHAFPPLTNSAGEPTGAMYGVLNMLRSLIMQYKPTHAAVVFDAKGKTFRDELFEHYKSHRPPMPDDLRAQIEPLHAMVKAMGLPLLAVSGVEADDVIGTLAREAEKAGR
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="28">
<INSDSeq>
  <INSDSeq_length>612</INSDSeq_length>
  <INSDSeq_moltype>AA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..612</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>protein</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q56">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
```

</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>APEVTATVISYDNYVTILDEETLKAWIAKLEKAPVFAFDTETDSLDNISANLVGLSFAIEPGVAAYIPVAHDYLDAPDQISRERALELLKPLLEDEKALKVGQNLKYDRGILANYGIELRGIAFDTMLES
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="29">
<INSDSeq>
<INSDSeq_length>1423</INSDSeq_length>
<INSDSeq_moltype>AA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..1423</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>protein</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q58">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMA
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="30">
<INSDSeq>
<INSDSeq_length>1300</INSDSeq_length>
<INSDSeq_moltype>AA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..1300</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>protein</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q60">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDA
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="31">
<INSDSeq>
<INSDSeq_length/>
<INSDSeq_moltype/>
<INSDSeq_division/>
<INSDSeq_sequence>000</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="32">
<INSDSeq>
<INSDSeq_length>23</INSDSeq_length>
<INSDSeq_moltype>AA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..23</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>protein</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q64">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>GSSETPGTSESATPEGGSGGSGS</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="33">
<INSDSeq>
<INSDSeq_length>203</INSDSeq_length>
<INSDSeq_moltype>AA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..203</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>protein</INSDQualifier_value>

</INSDQualifier>
<INSDQualifier id="q66">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>APEVTATVISYDNYVTILDEETLKAWIAKLEKAPVFAFDTETDSLDNISANLVGLSFAIEPGVAAYIPVAHDYLDAPDQISRERALELLKPLLEDEKALKVGQNLKYDRGILANYGIELRGIAFDTMLES
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="34">
<INSDSeq>
<INSDSeq_length>658</INSDSeq_length>
<INSDSeq_moltype>AA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..658</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>protein</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q68">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>VQIPQNPLILVDGSSYLYRAYHAFPPLTNSAGEPTGAMYGVLNMLRSLIMQYKPTHAAVVFDAKGKTFRDELFEHYKSHRPPMPDDLRAQIEPLHAMVKAMGLPLLAVSGVEADDVIGTLAREAEKAGRI
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="35">
<INSDSeq>
<INSDSeq_length>324</INSDSeq_length>
<INSDSeq_moltype>AA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..324</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>protein</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q70">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>VQIPQNPLILVDGSSYLYRAYHAFPPLTNSAGEPTGAMYGVLNMLRSLIMQYKPTHAAVVFDAKGKTFRDELFEHYKSHRPPMPDDLRAQIEPLHAMVKAMGLPLLAVSGVEADDVIGTLAREAEKAGRI
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="36">
<INSDSeq>
<INSDSeq_length>2859</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..2859</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q72">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>ggatcctctgagacaccaggcacctccgagtctgccacccctgagggaggcagcggaggcagcggctccgtgcagatcccacagaacccccctgatcctggtggacggcagctcctacctgtatcgggcct
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="37">
<INSDSeq>
<INSDSeq_length>2373</INSDSeq_length>
<INSDSeq_moltype>AA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..2373</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>protein</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q74">

<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMA
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="38">
<INSDSeq>
<INSDSeq_length>24</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..24</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q76">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>accggaggacaagtcgtacaatgg</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="39">
<INSDSeq>
<INSDSeq_length>24</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..24</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q78">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>aaacccattgtacgacttgtcctc</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="40">
<INSDSeq>
<INSDSeq_length>42</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..42</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q80">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>gtctgccacccctgagggaggagcaccagaggtgacagccac</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="41">
<INSDSeq>
<INSDSeq_length>21</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..21</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q82">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>

</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>ttgtgcagcaccttagggtcg</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="42">
<INSDSeq>
<INSDSeq_length>6160</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..6160</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q84">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>atggactataaggaccacgacggagactacaaggatcatgatattgattacaaagacgatgacgataagatggccccaaagaagaagcggaaggtcggtatccacggagtcccagcagccgacaagaag
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="43">
<INSDSeq>
<INSDSeq_length>2052</INSDSeq_length>
<INSDSeq_moltype>AA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..2052</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>protein</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q86">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEM
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="44">
<INSDSeq>
<INSDSeq_length>42</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..42</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q88">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>gtctgccacccctgagggaggagcaccagaggtgacagccac</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="45">
<INSDSeq>
<INSDSeq_length>41</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..41</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q90">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>

      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>ctaggaattttctagactagtgcttctgcagatctggccac</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="46">
  <INSDSeq>
    <INSDSeq_length>4932</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..4932</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q92">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>atggactataaggaccacgacggagactacaaggatcatgatattgattacaaagacgatgacgataagatggccccaaagaagaagcggaaggtcggtatccacggagtcccagcagccgacaagaag
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="47">
  <INSDSeq>
    <INSDSeq_length>1643</INSDSeq_length>
    <INSDSeq_moltype>AA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..1643</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>protein</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q94">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEM
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="48">
  <INSDSeq>
    <INSDSeq_length>6315</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..6315</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q96">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>atggactataaggaccacgacggagactacaaggatcatgatattgattacaaagacgatgacgataagatggccccaaagaagaagcggaaggtcggtatccacggagtcccagcagccgacaagaag
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="49">
  <INSDSeq>
    <INSDSeq_length>2104</INSDSeq_length>
    <INSDSeq_moltype>AA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..2104</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>protein</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q98">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>

<INSDSeq_sequence>MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEM
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="50">
<INSDSeq>
<INSDSeq_length>24</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..24</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q100">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>cagagctgatcatcgatttcctgg</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="51">
<INSDSeq>
<INSDSeq_length>38</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..38</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q102">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>aagaaagcgagctagcccacggtggctgtcacctctgg</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="52">
<INSDSeq>
<INSDSeq_length>5313</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..5313</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q104">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>atggactataaggaccacgacggagactacaaggatcatgatattgattacaaagacgatgacgataagatggccccaaagaagaagcggaaggtcggtatccacggagtcccagcagccgacaagaag
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="53">
<INSDSeq>
<INSDSeq_length>1770</INSDSeq_length>
<INSDSeq_moltype>AA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..1770</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>protein</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q106">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEM
</INSDSeq>

</SequenceData>
<SequenceData sequenceIDNumber="54">
<INSDSeq>
<INSDSeq_length>23</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..23</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q108">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>gccaagctgaagtctacatatac</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="55">
<INSDSeq>
<INSDSeq_length>22</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..22</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q110">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>ctgagagtaggcggcggacacg</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="56">
<INSDSeq>
<INSDSeq_length>22</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..22</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q112">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>cgtgtccgccgcctactctcag</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="57">
<INSDSeq>
<INSDSeq_length>20</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..20</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q114">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>tccatttcaggtgtcgtgag</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="58">

<INSDSeq>
<INSDSeq_length>18</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..18</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q116">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>tccagctcgaccaggatg</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="59">
<INSDSeq>
<INSDSeq_length>21</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..21</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q118">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>agccagggctgggcataaaag</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="60">
<INSDSeq>
<INSDSeq_length>25</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..25</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q120">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>tgggaaaatagaccaataggcagag</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="61">
<INSDSeq>
<INSDSeq_length>22</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..22</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q122">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>tcctgttgttcatcatcctagc</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="62">
<INSDSeq>
<INSDSeq_length>20</INSDSeq_length>

```
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..20</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q124">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>tccagccattgtgttgaatc</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="63">
  <INSDSeq>
    <INSDSeq_length>20</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..20</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q126">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>gtttaagggcccgccttttg</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="64">
  <INSDSeq>
    <INSDSeq_length>20</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..20</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q128">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>tgtcttacctctcggtgcct</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="65">
  <INSDSeq>
    <INSDSeq_length>20</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..20</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q130">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>gacccaggtttcctgagctg</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="66">
  <INSDSeq>
    <INSDSeq_length>20</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
```

```
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..20</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q132">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>ttcagagcttcctcccaagc</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="67">
  <INSDSeq>
    <INSDSeq_length>20</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..20</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q134">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>ccctatacctgggctccgtt</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="68">
  <INSDSeq>
    <INSDSeq_length>24</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..24</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q136">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>gaaagggcctctctctttgtaatg</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="69">
  <INSDSeq>
    <INSDSeq_length>23</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..23</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q138">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>aagctctacaagggcagagaatg</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="70">
  <INSDSeq>
    <INSDSeq_length>22</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
```

<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..22</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q140">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>tcaaagctcccagattcacgtt</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="71">
<INSDSeq>
<INSDSeq_length>20</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..20</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q142">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>ctgctctttgcctgttggag</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="72">
<INSDSeq>
<INSDSeq_length>20</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..20</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q144">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>gctaaagctggaaggctgtg</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="73">
<INSDSeq>
<INSDSeq_length>23</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..23</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q146">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>aaaaggcagggcagagtcgactg</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="74">
<INSDSeq>
<INSDSeq_length>23</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..23</INSDFeature_location>

```
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q148">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>ggtaggaaaacagcccaagggac</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="75">
  <INSDSeq>
    <INSDSeq_length>21</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..21</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q150">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>ccatcacccttcagaacctga</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="76">
  <INSDSeq>
    <INSDSeq_length>21</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..21</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q152">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>tcagtggctaacagaagctga</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="77">
  <INSDSeq>
    <INSDSeq_length>25</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..25</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q154">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>gtaggattctaatcactgcctgctc</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="78">
  <INSDSeq>
    <INSDSeq_length>25</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..25</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
```

```
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q156">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>ctttctgcacctcctgatagccttg</INSDSeq_sequence>
 </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="79">
 <INSDSeq>
  <INSDSeq_length>20</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..20</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q158">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>gccaagtgtgaatgtgcaca</INSDSeq_sequence>
 </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="80">
 <INSDSeq>
  <INSDSeq_length>21</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..21</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q160">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>cctcttaggctatgaccaggg</INSDSeq_sequence>
 </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="81">
 <INSDSeq>
  <INSDSeq_length>20</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..20</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q162">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>tcattgggaggccgaacatc</INSDSeq_sequence>
 </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="82">
 <INSDSeq>
  <INSDSeq_length>20</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..20</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
```

```
        &lt;/INSDQualifier&gt;
        &lt;INSDQualifier id="q164"&gt;
          &lt;INSDQualifier_name&gt;organism&lt;/INSDQualifier_name&gt;
          &lt;INSDQualifier_value&gt;Synthetic construct&lt;/INSDQualifier_value&gt;
        &lt;/INSDQualifier&gt;
      &lt;/INSDFeature_quals&gt;
    &lt;/INSDFeature&gt;
  &lt;/INSDSeq_feature-table&gt;
  &lt;INSDSeq_sequence&gt;aagggctgaaggatgatgca&lt;/INSDSeq_sequence&gt;
&lt;/INSDSeq&gt;
&lt;/SequenceData&gt;
&lt;SequenceData sequenceIDNumber="83"&gt;
  &lt;INSDSeq&gt;
    &lt;INSDSeq_length&gt;20&lt;/INSDSeq_length&gt;
    &lt;INSDSeq_moltype&gt;DNA&lt;/INSDSeq_moltype&gt;
    &lt;INSDSeq_division&gt;PAT&lt;/INSDSeq_division&gt;
    &lt;INSDSeq_feature-table&gt;
      &lt;INSDFeature&gt;
        &lt;INSDFeature_key&gt;source&lt;/INSDFeature_key&gt;
        &lt;INSDFeature_location&gt;1..20&lt;/INSDFeature_location&gt;
        &lt;INSDFeature_quals&gt;
          &lt;INSDQualifier&gt;
            &lt;INSDQualifier_name&gt;mol_type&lt;/INSDQualifier_name&gt;
            &lt;INSDQualifier_value&gt;other DNA&lt;/INSDQualifier_value&gt;
          &lt;/INSDQualifier&gt;
          &lt;INSDQualifier id="q166"&gt;
            &lt;INSDQualifier_name&gt;organism&lt;/INSDQualifier_name&gt;
            &lt;INSDQualifier_value&gt;Synthetic construct&lt;/INSDQualifier_value&gt;
          &lt;/INSDQualifier&gt;
        &lt;/INSDFeature_quals&gt;
      &lt;/INSDFeature&gt;
    &lt;/INSDSeq_feature-table&gt;
    &lt;INSDSeq_sequence&gt;ctggcattctgggagcttca&lt;/INSDSeq_sequence&gt;
  &lt;/INSDSeq&gt;
&lt;/SequenceData&gt;
&lt;SequenceData sequenceIDNumber="84"&gt;
  &lt;INSDSeq&gt;
    &lt;INSDSeq_length&gt;20&lt;/INSDSeq_length&gt;
    &lt;INSDSeq_moltype&gt;DNA&lt;/INSDSeq_moltype&gt;
    &lt;INSDSeq_division&gt;PAT&lt;/INSDSeq_division&gt;
    &lt;INSDSeq_feature-table&gt;
      &lt;INSDFeature&gt;
        &lt;INSDFeature_key&gt;source&lt;/INSDFeature_key&gt;
        &lt;INSDFeature_location&gt;1..20&lt;/INSDFeature_location&gt;
        &lt;INSDFeature_quals&gt;
          &lt;INSDQualifier&gt;
            &lt;INSDQualifier_name&gt;mol_type&lt;/INSDQualifier_name&gt;
            &lt;INSDQualifier_value&gt;other DNA&lt;/INSDQualifier_value&gt;
          &lt;/INSDQualifier&gt;
          &lt;INSDQualifier id="q168"&gt;
            &lt;INSDQualifier_name&gt;organism&lt;/INSDQualifier_name&gt;
            &lt;INSDQualifier_value&gt;Synthetic construct&lt;/INSDQualifier_value&gt;
          &lt;/INSDQualifier&gt;
        &lt;/INSDFeature_quals&gt;
      &lt;/INSDFeature&gt;
    &lt;/INSDSeq_feature-table&gt;
    &lt;INSDSeq_sequence&gt;gagacctgtgggaagcgaaa&lt;/INSDSeq_sequence&gt;
  &lt;/INSDSeq&gt;
&lt;/SequenceData&gt;
&lt;SequenceData sequenceIDNumber="85"&gt;
  &lt;INSDSeq&gt;
    &lt;INSDSeq_length&gt;20&lt;/INSDSeq_length&gt;
    &lt;INSDSeq_moltype&gt;DNA&lt;/INSDSeq_moltype&gt;
    &lt;INSDSeq_division&gt;PAT&lt;/INSDSeq_division&gt;
    &lt;INSDSeq_feature-table&gt;
      &lt;INSDFeature&gt;
        &lt;INSDFeature_key&gt;source&lt;/INSDFeature_key&gt;
        &lt;INSDFeature_location&gt;1..20&lt;/INSDFeature_location&gt;
        &lt;INSDFeature_quals&gt;
          &lt;INSDQualifier&gt;
            &lt;INSDQualifier_name&gt;mol_type&lt;/INSDQualifier_name&gt;
            &lt;INSDQualifier_value&gt;other DNA&lt;/INSDQualifier_value&gt;
          &lt;/INSDQualifier&gt;
          &lt;INSDQualifier id="q170"&gt;
            &lt;INSDQualifier_name&gt;organism&lt;/INSDQualifier_name&gt;
            &lt;INSDQualifier_value&gt;Synthetic construct&lt;/INSDQualifier_value&gt;
          &lt;/INSDQualifier&gt;
        &lt;/INSDFeature_quals&gt;
      &lt;/INSDFeature&gt;
    &lt;/INSDSeq_feature-table&gt;
    &lt;INSDSeq_sequence&gt;gtctgaggtcacacagtggg&lt;/INSDSeq_sequence&gt;
  &lt;/INSDSeq&gt;
&lt;/SequenceData&gt;
&lt;SequenceData sequenceIDNumber="86"&gt;
  &lt;INSDSeq&gt;
    &lt;INSDSeq_length&gt;20&lt;/INSDSeq_length&gt;
    &lt;INSDSeq_moltype&gt;DNA&lt;/INSDSeq_moltype&gt;
    &lt;INSDSeq_division&gt;PAT&lt;/INSDSeq_division&gt;
    &lt;INSDSeq_feature-table&gt;
      &lt;INSDFeature&gt;
        &lt;INSDFeature_key&gt;source&lt;/INSDFeature_key&gt;
        &lt;INSDFeature_location&gt;1..20&lt;/INSDFeature_location&gt;
        &lt;INSDFeature_quals&gt;
          &lt;INSDQualifier&gt;
            &lt;INSDQualifier_name&gt;mol_type&lt;/INSDQualifier_name&gt;
            &lt;INSDQualifier_value&gt;other DNA&lt;/INSDQualifier_value&gt;
          &lt;/INSDQualifier&gt;
          &lt;INSDQualifier id="q172"&gt;
```

<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>ctgagagcagggaccacatc</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="87">
<INSDSeq>
<INSDSeq_length>24</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..24</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q174">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>accggatgagataatgatgagtca</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="88">
<INSDSeq>
<INSDSeq_length>24</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..24</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q176">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>aaactgactcatcattatctcatc</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="89">
<INSDSeq>
<INSDSeq_length>24</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..24</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q178">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>accgccattgttcaatatcgtccg</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="90">
<INSDSeq>
<INSDSeq_length>24</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..24</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q180">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>

```
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>aaaccggacgatattgaacaatgg</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="91">
  <INSDSeq>
    <INSDSeq_length>26</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..26</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q182">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>caccgtttagcatgttcccaattctc</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="92">
  <INSDSeq>
    <INSDSeq_length>26</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..26</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q184">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>aaacgagaattgggaacatgctaaac</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="93">
  <INSDSeq>
    <INSDSeq_length>24</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..24</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q186">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>accggtaacggcagacttctcctc</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="94">
  <INSDSeq>
    <INSDSeq_length>24</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..24</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q188">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
```

```
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>aaacgaggagaagtctgccgttac</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="95">
  <INSDSeq>
    <INSDSeq_length>23</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..23</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q190">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>accgttgcccaaagattccccag</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="96">
  <INSDSeq>
    <INSDSeq_length>23</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..23</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q192">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>aaacctggggaatctttgggcaa</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="97">
  <INSDSeq>
    <INSDSeq_length>24</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..24</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q194">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>accgtgactgagatctgagtgctc</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="98">
  <INSDSeq>
    <INSDSeq_length>24</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..24</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q196">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
```

<INSDSeq_sequence>aaacgagcactcagatctcagtca</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="99">
<INSDSeq>
<INSDSeq_length>24</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..24</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q198">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>accgactgttgcctccggttctga</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="100">
<INSDSeq>
<INSDSeq_length>24</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..24</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q200">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>aaactcagaaccggaggcaacagt</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="101">
<INSDSeq>
<INSDSeq_length>24</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..24</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q202">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>accggaggacaagtcgtacaatgg</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="102">
<INSDSeq>
<INSDSeq_length>24</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..24</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q204">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>aaacccattgtacgacttgtcctc</INSDSeq_sequence>
</INSDSeq>

</SequenceData>
<SequenceData sequenceIDNumber="103">
<INSDSeq>
<INSDSeq_length>20</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..20</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q206">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>gaggacaagtcgtacaatgg</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="104">
<INSDSeq>
<INSDSeq_length>20</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..20</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q208">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>actgttgcctccggttctga</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="105">
<INSDSeq>
<INSDSeq_length>21</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..21</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q210">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>tttagcatgttcccaattctc</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="106">
<INSDSeq>
<INSDSeq_length>20</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..20</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q212">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>gtaacggcagacttctcctc</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="107">

```
  <INSDSeq>
    <INSDSeq_length>20</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..20</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q214">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>gatgagataatgatgagtca</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="108">
  <INSDSeq>
    <INSDSeq_length>20</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..20</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q216">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>ccattgttcaatatcgtccg</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="109">
  <INSDSeq>
    <INSDSeq_length>20</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..20</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q218">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>tgactgagatctgagtgctc</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="110">
  <INSDSeq>
    <INSDSeq_length>19</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..19</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q220">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>ttgcccaaagattccccag</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="111">
  <INSDSeq>
    <INSDSeq_length>23</INSDSeq_length>
```

```
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..23</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q222">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>gaggacaagtcgtacaatggtgg</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="112">
  <INSDSeq>
    <INSDSeq_length>23</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..23</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q224">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>ctcctgttcagcatgttaccacc</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="113">
  <INSDSeq>
    <INSDSeq_length>39</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..39</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q226">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>acagaggacaagtcgtacaatggtggaggaatagttctt</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="114">
  <INSDSeq>
    <INSDSeq_length>28</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..28</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q228">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>acagaggacaagtcgtacaatagttctt</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="115">
  <INSDSeq>
    <INSDSeq_length>28</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
```

```
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..28</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q230">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>acagaggacaagtggaggaatagttctt</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="116">
  <INSDSeq>
    <INSDSeq_length>31</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..31</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q232">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>acagaggacaagtcgtggaggaatagttctt</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="117">
  <INSDSeq>
    <INSDSeq_length>28</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..28</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q234">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>acagaggacaagtggaggaatagttctt</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="118">
  <INSDSeq>
    <INSDSeq_length>26</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..26</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q236">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>acagaggtggtggaggaatagttctt</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="119">
  <INSDSeq>
    <INSDSeq_length>35</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
```

        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..35</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q238">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>gactcctgaggagaagtctgccgttactgccctgt</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="120">
  <INSDSeq>
    <INSDSeq_length>23</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..23</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q240">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>gactcctgccgttactgccctgt</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="121">
  <INSDSeq>
    <INSDSeq_length>22</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..22</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q294">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>gactcctgccgttactgcctgt</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="122">
  <INSDSeq>
    <INSDSeq_length>40</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..40</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q296">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>gaagtacaagaacaccttcagaaccggaggcaacagttga</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="123">
  <INSDSeq>
    <INSDSeq_length>29</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..29</INSDFeature_location>

<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q298">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>gaagtacaagaaccggaggcaacagttga</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="124">
<INSDSeq>
<INSDSeq_length>31</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..31</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q300">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>gaagtacaagaacaccggaggcaacagttga</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="125">
<INSDSeq>
<INSDSeq_length>34</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..34</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q302">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>gaagtacaagaacacctccggaggcaacagttga</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="126">
<INSDSeq>
<INSDSeq_length>35</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..35</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q304">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>gactcctgaggagaagtctgccgttactgccctgt</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="127">
<INSDSeq>
<INSDSeq_length>32</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..32</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>

```
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q306">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>gactcctgagaagtctgccgttactgccctgt</INSDSeq_sequence>
 </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="128">
 <INSDSeq>
  <INSDSeq_length>30</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..30</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q308">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>gactcctgaagtctgccgttactgccctgt</INSDSeq_sequence>
 </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="129">
 <INSDSeq>
  <INSDSeq_length>54</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..54</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q310">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>cagtttctcagaaagacacaaattcctgagaattgggaacatgctaaatacaaa</INSDSeq_sequence>
 </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="130">
 <INSDSeq>
  <INSDSeq_length>44</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..44</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q312">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>cagtttctcagaaagacacaaattgggaacatgctaaatacaaa</INSDSeq_sequence>
 </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="131">
 <INSDSeq>
  <INSDSeq_length>47</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..47</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
```

        </INSDQualifier>
        <INSDQualifier id="q314">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>cagtttctcagaaagacacaaattcctgggaacatgctaaatacaaa</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="132">
<INSDSeq>
  <INSDSeq_length>47</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..47</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q316">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>cagtttctcagaaagacacaaataattgggaacatgctaaatacaaa</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="133">
<INSDSeq>
  <INSDSeq_length>40</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..40</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q318">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>aagaacctattcctccaccattgtacgacttgtcctctgt</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="134">
<INSDSeq>
  <INSDSeq_length>39</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..39</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q320">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>aagaacctattcctccaccatgtacgacttgtcctctgt</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="135">
<INSDSeq>
  <INSDSeq_length>41</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..41</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q322">

```
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>aagaacctattcctccaccatttgtacgacttgtcctctgt</INSDSeq_sequence>
 </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="136">
 <INSDSeq>
  <INSDSeq_length>39</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..39</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q324">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>aagaacctattcctccaccttgtacgacttgtcctctgt</INSDSeq_sequence>
 </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="137">
 <INSDSeq>
  <INSDSeq_length>41</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..41</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q326">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>aagaacctattcctccaccaattgtacgacttgtcctctgt</INSDSeq_sequence>
 </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="138">
 <INSDSeq>
  <INSDSeq_length>35</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..35</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q328">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_quals>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>aagaacctattcctccagtacgacttgtcctctgt</INSDSeq_sequence>
 </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="139">
 <INSDSeq>
  <INSDSeq_length>42</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..42</INSDFeature_location>
      <INSDFeature_quals>
        <INSDQualifier>
          <INSDQualifier_name>mol_type</INSDQualifier_name>
          <INSDQualifier_value>other DNA</INSDQualifier_value>
        </INSDQualifier>
        <INSDQualifier id="q330">
          <INSDQualifier_name>organism</INSDQualifier_name>
          <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
```

```
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>aagaacctattcctccaccattttgtacgacttgtcctctgt</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="140">
  <INSDSeq>
    <INSDSeq_length>35</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..35</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q332">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>aagaacctattcctccaccacgacttgtcctctgt</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="141">
  <INSDSeq>
    <INSDSeq_length>37</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..37</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q334">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>aagaacctattcctccaccatacgacttgtcctctgt</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="142">
  <INSDSeq>
    <INSDSeq_length>33</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..33</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q336">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>cctgactcctgaggagaagtctgccgttactgc</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="143">
  <INSDSeq>
    <INSDSeq_length>29</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..29</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q338">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
```

</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>cctgactcctgagaagtctgccgtactgc</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="144">
<INSDSeq>
<INSDSeq_length>28</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..28</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q340">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>cctgactcgagaagtctgccgttactgc</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="145">
<INSDSeq>
<INSDSeq_length>23</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..23</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q342">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>cctgagaagtctgccgttactgc</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="146">
<INSDSeq>
<INSDSeq_length>40</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..40</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q344">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>tgagggtgagggatgagataatgatgagtcagggcttcag</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="147">
<INSDSeq>
<INSDSeq_length>36</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..36</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q346">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>

<INSDSeq_sequence>tgagggtgagggatgagataatgatcagggcttcag</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="148">
<INSDSeq>
<INSDSeq_length>23</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..23</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q348">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>gtaacggcagacttctcctcagg</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="149">
<INSDSeq>
<INSDSeq_length>23</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..23</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q350">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>ttgacagcagtcttctcctcagg</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="150">
<INSDSeq>
<INSDSeq_length>23</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..23</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q352">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>atacttaaggacttctcctccat</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="151">
<INSDSeq>
<INSDSeq_length>23</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..23</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q354">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>ggaggggcaggcttctcctctgg</INSDSeq_sequence>
</INSDSeq>

</SequenceData>
<SequenceData sequenceIDNumber="152">
<INSDSeq>
<INSDSeq_length>23</INSDSeq_length>
<INSDSeq_moltype>DNA</INSDSeq_moltype>
<INSDSeq_division>PAT</INSDSeq_division>
<INSDSeq_feature-table>
<INSDFeature>
<INSDFeature_key>source</INSDFeature_key>
<INSDFeature_location>1..23</INSDFeature_location>
<INSDFeature_quals>
<INSDQualifier>
<INSDQualifier_name>mol_type</INSDQualifier_name>
<INSDQualifier_value>other DNA</INSDQualifier_value>
</INSDQualifier>
<INSDQualifier id="q356">
<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Synthetic construct</INSDQualifier_value>
</INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>tcacaggcagacttctccacggg</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
</ST26SequenceListing>