

COMPOSITIONS AND METHODS FOR THE TARGETING OF PCSK9

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to, and benefit of, U.S. Provisional Application Nos. 63/492,978, filed on March 29, 2023, and 63/505,888, filed on June 2, 2023, the contents of each of which are incorporated by reference herein in their entireties.

REFERENCE TO AN ELECTRONIC SEQUENCE LISTING

[0002] The contents of the electronic sequence listing (SCRB_058_02WO_SeqList_ST26.xml; Size: 141,865,746 bytes; and Date of Creation: March 26, 2024) are herein incorporated by reference in its entirety.

BACKGROUND

[0003] In mammals, cholesterol is transported within lipoproteins via emulsification. The lipoprotein particles are classified based on their density: low-density lipoproteins (LDL), very low-density lipoproteins (VLDL), high-density lipoproteins (HDL), and chylomicrons. Surface LDL receptors are internalized during cholesterol absorption. A cell with abundant cholesterol will have its LDL receptor synthesis blocked to prevent new cholesterol in LDL particles from being taken up. Conversely, LDL receptor synthesis is promoted when a cell is deficient in cholesterol. When the process is unregulated, excess LDL particles will travel in the blood without uptake by an LDL receptor. LDL particles in the blood are oxidized and taken up by macrophages, which then become engorged and form foam cells. These foam cells can become trapped in the walls of blood vessels and contribute to atherosclerotic plaque formation, which is one of the main causes of heart attacks, strokes, and other serious medical problems.

[0004] The liver protein proprotein convertase subtilisin/kexin Type 9 (PCSK9) is a secreted, globular, auto-activating serine protease that binds to the low-density lipoprotein receptor (LDL-R) during endocytosis of LDL particles, preventing recycling of the LDL-R to the cell surface and leading to reduction of LDL-cholesterol clearance. PCSK9 binds to the LDL-R (through the EGF-A domain), preventing the conformational change of the receptor-ligand complex, which redirects the LDL-R to the lysosome instead. As the receptor for low-density lipoprotein particles (LDL) typically transports thousands of fat molecules (including cholesterol) per particle within extracellular fluid, blocking or inhibiting the function of

PCSK9 to boost LDL-R-mediated clearance of LDL cholesterol can lower LDL particle concentrations. PCSK9 is expressed mainly in the liver, the intestine, the kidney, and the central nervous system, but is also highly expressed in arterial walls such as endothelium, smooth muscle cells, and macrophages, with a local effect that can regulate vascular homeostasis and atherosclerosis.

[0005] PCSK9 is a member of the proprotein convertase (PC) family and its gene is mutated in ~ 2% to 3% of individuals with familial hypercholesterolemia (FH) (Sepideh Mikaeeli, S., et al. Functional analysis of natural PCSK9 mutants in modern and archaic humans. FEBS J. 2019 Aug 6. doi: 10.1111/febs.15036). Researchers have identified several PCSK9 mutations that cause an inherited form of high cholesterol (hypercholesterolemia). These mutations change a single amino acid in the PCSK9 protein. Researchers describe the mutations responsible for hypercholesterolemia as "gain-of-function" because they appear to enhance the activity of the PCSK9 protein or give the protein a new, atypical function (Blesa, S., et al. A New PCSK9 Gene Promoter Variant Affects Gene Expression and Causes Autosomal Dominant Hypercholesterolemia. J. Clin. Endocrinol. & Metab. 93:3577(2008)). The overactive PCSK9 protein substantially reduces the number of low-density lipoprotein receptors on the surface of liver cells. With fewer receptors to remove low-density lipoproteins from the blood, people with gain-of-function mutations in the PCSK9 gene have very high blood cholesterol levels. Autosomal dominant hypercholesterolemia (ADH) is a genetic disorder characterized by increased low-density lipoprotein (LDL)-cholesterol levels, leading to high risk of premature cardiovascular disease. Approximately 10 mutations in PCSK9 have been identified as a cause of the disease in different populations. All known mutations in PCSK9 causing hypercholesterolemia produce an increase in the enzymatic activity of this protease (Bleasa, S., 2008). In addition, mutations in PCSK9 can lead to autosomal dominant familial hypobetalipoproteinemia, which can lead to hepatic steatosis, cirrhosis, and other disorders.

[0006] The advent of CRISPR/Cas systems, and the programmable nature of these minimal systems, has facilitated their use as a versatile technology for genomic manipulation and engineering. However, current methods of generating PCSK9 protective variants and loss-of-function mutants *in vivo* have been ineffective due to the large number of cells that need to be modified to modulate cholesterol levels. Other concerns involve off-target effects, genome instability, or oncogenic modifications that may be caused by genome editing, as well as a lack of safe delivery modalities for repressor protein systems. Additionally, in certain disease

indications, gene silencing, or repression, is preferable to gene editing. The ability to render CRISPR nucleases such as Cas9 and CasX catalytically-inactive has been demonstrated (WO2020247882A1 and US20200087641A1, incorporated by reference herein), which makes these systems an attractive platform for the generation of fusion proteins with repressor domains capable of gene silencing. While certain repressor systems have been described, there remains a need for additional gene repressor systems that have been optimized and/or offer improvements over earlier generations of gene repressor systems, such as those based on Cas9, for utilization in a variety of therapeutic, diagnostic, and research applications. Thus, there remains a need for improved compositions and methods to regulate PCSK9.

SUMMARY

[0007] The present disclosure provides systems comprising or encoding fusion proteins comprising DNA binding and linked repressor domains used in the repression and/or epigenetic modification of proprotein convertase subtilisin/kexin Type 9 (*PCSK9*) gene target nucleic acid sequences. In some aspects, the fusion protein comprises a DNA binding protein comprising a catalytically-dead CRISPR protein, such as a Class 2, Type V CRISPR protein, and a guide nucleic acid comprising a targeting sequence complementary to the *PCSK9* gene target nucleic acid sequence. The proteins and guide nucleic acids can be modified for passive entry into target cells and are useful in a variety of methods for repression of *PCSK9*, which methods are also provided. The present disclosure also provides vectors and lipid nanoparticles (LNP) encoding or encapsulating the fusion proteins and guide nucleic acids components for the delivery of the systems to cells for the transcriptional repression of the *PCSK9* target nucleic acid sequence. The present disclosure also provides methods for modifying a *PCSK9* gene target nucleic acid sequence. The present disclosure also provides methods for treating subjects having a *PCSK9*-related disease. In some embodiments, the compositions and methods have utility in subjects having a metabolic disorder such as, but not limited to, familial hypercholesterolemia, familial hypobetalipoproteinemia, or elevated cholesterol levels.

[0008] In other aspects, provided herein are systems comprising *PCSK9* repressor systems, or vectors comprising or encoding *PCSK9* repressor systems for use in the manufacture of a medicament for the treatment of a *PCSK9*-related disease in a subject in need thereof.

[0009] Further features and advantages of certain embodiments of the present disclosure will become more fully apparent in the following description of embodiments and drawings thereof, and from the claims.

INCORPORATION BY REFERENCE

[0010] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference. The contents of WO 2020/247882, WO 2020/247883, WO 2021/113772, WO2022/120095, WO 2022/125843, WO 2022/261150, WO 2022/261149, WO 2023/049872, WO 2023/049742 and WO 2023/240162, which disclose CasX variants and gRNA variants, and methods of delivering same, and WO2021/142342 which discloses compositions and methods of modifying PCSK9 are hereby incorporated by reference in their entirety.

BRIEF DESCRIPTION OF THE DRAWINGS

[0011] Understanding of the features and advantages of the present disclosure will be facilitated by reference to the following figures:

[0012] FIG. 1 illustrates schematics of various configurations of repressor fusion proteins with the DNMT3A ADD domain incorporated. “D3A ADD”, “D3A CD”, and “D3L ID” denote the ADD domain of DNMT3A, the catalytic domain of DNMT3A, and the interaction domain of DNMT3L, respectively. L1-L3 are linkers. NLS is the nuclear localization signal.

[0013] FIG. 2 is a bar graph showing the quantification of secreted PCSK9 levels at 6, 18, and 36 days post-transfection in Huh7 cells lipofected with mRNA encoding for CasX 676, dXR1, or LTRP5-ADD-ZIM3 when paired with the indicated targeting gRNAs, as described in Example 1. Secreted PCSK9 levels were normalized to total cell count. Naïve, untreated cells served as an experimental control.

[0014] FIG. 3 is a bar plot showing the quantification of normalized secreted PCSK9 levels at 4 days post-transfection in HepG2 cells lipofected with mRNA encoding for CasX 676, dXR1, or LTRP5-ADD-ZIM3 when paired with the indicated targeting gRNAs, as described in Example 2. Secreted PCSK9 levels were normalized to total cell count. Naïve, untreated cells served as experimental controls.

[0015] FIG. 4 is a bar plot showing the quantification of normalized secreted PCSK9 levels at 4 days post-transfection in Huh7 cells lipofected with mRNA encoding for CasX 676, dXR1, or LTRP5-ADD-ZIM3 when paired with the indicated targeting gRNAs, as described in Example 2. Secreted PCSK9 levels were normalized to total cell count. Naïve, untreated cells served as experimental controls.

[0016] FIG. 5 is a bar plot showing the quantification of normalized secreted PCSK9 levels at 4 days post-transfection in Hep3B cells lipofected with mRNA encoding for CasX 676, dXR1, or LTRP5-ADD-ZIM3 when paired with the indicated targeting gRNAs, as described in Example 2. Secreted PCSK9 levels were normalized to total cell count. Naïve, untreated cells served as experimental controls.

[0017] FIG. 6 is a bar plot showing the quantification of secreted PCSK9 levels at 4, 14, and 27 days post-transfection in Huh7 cells lipofected with mRNA encoding for CasX 676, dXR1, or LTRP5-ADD-ZIM3 when paired with the indicated targeting gRNAs, as described in Example 2. Quantification of secreted PCSK9 levels is shown as relative to the secreted levels detected in the naïve, untreated cells at the day 4 timepoint.

[0018] FIG. 7 is a schematic of gRNA scaffold variant 174 (SEQ ID NO: 1744), as described in Example 3. Structural motifs are highlighted.

[0019] FIG. 8 is a schematic of gRNA scaffold variant 235 (SEQ ID NO: 1745), as described in Example 3. Highlighted structural motifs are the same as in FIG. 7. The differences between variant 174 and variant 235 lie in the extended stem motif and several single-nucleotide changes (indicated with asterisks). Variant 316 maintains the shorter extended stem from variant 174 but harbors the four substitutions found in scaffold 235.

[0020] FIG. 9 is a schematic of gRNA scaffold variant 316 (SEQ ID NO: 1746), as described in Example 3. Highlighted structural motifs are the same as in FIG. 7. Variant 316 maintains the shorter extended stem from variant 174 (FIG. 7) but harbors the four substitutions found in scaffold 235 (FIG. 8).

[0021] FIG. 10 is a schematic illustrating versions 1-3 of chemical modifications made to gRNA scaffold variant 235, as described in Example 3. Structural motifs are highlighted. Standard ribonucleotides are depicted as open circles, and 2'OMe-modified ribonucleotides are depicted as black circles. Phosphorothioate bonds are indicated with * below or beside the bond. For the v2 profile, the addition of three 3' uracils (3'UUU) is annotated with "U"s in the relevant circles.

[0022] FIG. 11 is a schematic illustrating versions 4-6 of chemical modifications made to gRNA scaffold variant 235, as described in Example 3. Structural motifs are highlighted. Standard ribonucleotides are depicted as open circles, and 2'OMe-modified ribonucleotides are depicted as black circles. Phosphorothioate bonds are indicated with * below or beside the bond.

[0023] FIG. 12 is a schematic illustrating versions 7-9 of chemical modifications made to gRNA scaffold variant 316, as described in Example 3. Structural motifs are highlighted. Standard ribonucleotides are depicted as open circles, and 2'OMe-modified ribonucleotides are depicted as black circles. Phosphorothioate bonds are indicated with * below or beside the bond.

[0024] FIG. 13 is a schematic illustrating versions 1-3 of chemical modifications made to gRNA scaffold variant 316, as described in Example 3. Structural motifs are highlighted. Standard ribonucleotides are depicted as open circles, and 2'OMe-modified ribonucleotides are depicted as black circles. Phosphorothioate bonds are indicated with * below or beside the bond.

[0025] FIG. 14 is a schematic illustrating versions 4-6 of chemical modifications made to gRNA scaffold variant 316, as described in Example 3. Structural motifs are highlighted. Standard ribonucleotides are depicted as open circles, and 2'OMe-modified ribonucleotides are depicted as black circles. Phosphorothioate bonds are indicated with * below or beside the bond.

[0026] FIG. 15 is a plot displaying a correlation between indel rate (depicted as edit fraction) at the *PCSK9* locus as measured by next-generation sequencing (NGS) (x-axis) and secreted PCSK9 levels (ng/mL) detected by enzyme-linked immunosorbent assay (ELISA) (y-axis) in HepG2 cells lipofected with CasX 491 mRNA and *PCSK9*-targeting gRNAs containing the indicated scaffold variant and spacer combination, as described in Example 3.

[0027] FIG. 16 is a plot illustrating the quantification of percent knockout of *B2M* in HepG2 cells co-transfected with 100 ng of CasX 491 mRNA and with the indicated doses of end-modified (v1) or unmodified (v0) *B2M*-targeting gRNAs with spacer 7.37, as described in Example 3. Editing level was determined by flow cytometry as the population of cells with loss of surface presentation of the HLA complex due to successful editing at the *B2M* locus.

[0028] FIG. 17 is a plot depicting the results of an editing assay measured as indel rate detected by NGS at the human *B2M* locus in HepG2 cells treated with the indicated doses of

LNPs formulated with CasX 491 mRNA and the indicated *B2M*-targeting gRNA, as described in Example 3.

[0029] FIG. 18 is a plot illustrating the quantification of percent knockout of B2M in HepG2 cells treated with the indicated doses of LNPs formulated with CasX 491 mRNA and the indicated *B2M*-targeting gRNA, as described in Example 3. Editing level was determined by flow cytometry as population of cells that did not have surface presentation of the HLA complex due to successful editing at the *B2M* locus.

[0030] FIG. 19 is a plot depicting the results of an editing assay measured as indel rate detected by NGS at the mouse *ROSA26* locus in Hepa1-6 cells treated with the indicated doses of LNPs formulated with CasX 676 mRNA #2 and the indicated *ROSA26*-targeting gRNA with either the v1 or v5 modification profile, as described in Example 3.

[0031] FIG. 20 is a plot illustrating the quantification of percent editing measured as indel rate detected by NGS at the *ROSA26* locus in mice treated with LNPs formulated with CasX 676 mRNA #2 and the indicated chemically-modified *ROSA26*-targeting gRNA, as described in Example 3.

[0032] FIG. 21 is a bar graph showing the results of the editing assay measured as indel rate detected by NGS at the mouse *PCSK9* locus in mice treated with LNPs formulated with CasX 676 mRNA #1 and the indicated chemically-modified *PCSK9*-targeting gRNA, as described in Example 3. Untreated mice served as experimental control.

[0033] FIG. 22 is a diagram of the secondary structure of guide RNA scaffold 235 (SEQ ID NO: 1745), noting the regions with CpG motifs, as described in Example 6. CpG motifs in (1) the pseudoknot stem, (2) the scaffold stem, (3) the extended stem bubble, (4) the extended step, and (5) the extended stem loop are labeled on the structure.

[0034] FIG. 23 is a diagram of the CpG-reducing mutations that were introduced into each of the five regions in the coding sequence of the guide RNA scaffold, as described in Example 6.

[0035] FIG. 24 provides the results of an editing experiment in which AAV vectors with various CpG-reduced or CpG-depleted guide RNA scaffolds were used to edit the B2M locus in induced neurons, as described in Example 6. The AAV vectors were administered at a multiplicity of infection (MOI) of 4e3. The bars show the mean $\pm$ the SD of two replicates per sample. “No Tx” indicates a non-transduced control, and “NT” indicates a control with a non-targeting spacer.

[0036] FIG. 25 provides the results of an editing experiment in which AAV vectors with various CpG-reduced or CpG-depleted guide RNA scaffolds were used to edit the B2M locus in induced neurons, as described in Example 6. The AAV vectors were administered at an MOI of 3×10^3 . The bars show the mean $\pm$ the SD of two replicates per sample. “No Tx” indicates a non-transduced control.

[0037] FIG. 26 provides the results of an editing experiment in which AAV vectors with various CpG-reduced or CpG-depleted guide RNA scaffolds were used to edit the B2M locus in induced neurons, as described in Example 6. The AAV vectors were administered at an MOI of 1×10^3 . The bars show the mean $\pm$ the SD of two replicates per sample. “No Tx” indicates a non-transduced control.

[0038] FIG. 27 provides the results of an editing experiment in which AAV vectors with various CpG-reduced or CpG-depleted guide RNA scaffolds were used to edit the B2M locus in induced neurons, as described in Example 6. The AAV vectors were administered at an MOI of 3×10^2 . The bars show the mean $\pm$ the SD of two replicates per sample. “No Tx” indicates a non-transduced control.

[0039] FIG. 28 illustrates the schematics of five configurations of long-term repressor protein (LTRP, also referred to herein as “repressor fusion proteins”) fusion proteins with repressor molecules linked to catalytically-dead CasX. D3A and D3L denote DNA methyltransferase 3 alpha (DNMT3A) and DNMT3A-like protein (DNMT3L), respectively. L1-L4 are linkers. NLS is the nuclear localization signal.

[0040] FIG. 29 illustrates the schematics of the various LTRP #5 architectures, where the additional DNMT3A domains were incorporated, as described in Example 7. The additional DNMT3A domains were the ADD domain of DNMT3A (“D3A ADD”) and the PWWP domain (Pro-Trp-Trp-Pro motif) of DNMT3A (“D3A PWWP”). “D3A endo” encodes for an endogenous sequence that occurs between DNMT3A PWWP and ADD domains. “D3A CD” and “D3L ID” denote the catalytic domain of DNMT3A and the interaction domain of DNMT3L respectively. “L1-L3” are linkers. “NLS” is the nuclear localization signal. See Table 13 for LTRP sequences.

[0041] FIG. 30 presents the results of a time-course experiment comparing B2M repression activities (represented as percentage of HLA-negative cells) of the indicated LTRP -ZIM3 and its variants with *B2M*-targeting gRNA using spacer 7.37, as described in Example 7. Data are presented as mean with standard deviation, N=3. CD = catalytic domain of DNMT3A.

[0042] FIG. 31 presents the results of the same time-course experiment shown in FIG. 30, but shows B2M repression activities of the indicated LTRP -ZIM3 variants with *B2M*-targeting gRNA using spacer 7.160, as described in Example 7. Data are presented as mean with standard deviation, N=3.

[0043] FIG. 32 presents the results of the same time-course experiment shown in FIG. 30, but shows B2M repression activities of the indicated LTRP -ZIM3 variants with *B2M*-targeting gRNA using spacer 7.165, as described in Example 7. Data are presented as mean with standard deviation, N=3.

[0044] FIG. 33 presents the results of the same time-course experiment shown in FIG. 30, but shows B2M repression activities of the indicated LTRP -ZIM3 variants with a non-targeting gRNA, as described in Example 7. Data are presented as mean with standard deviation, N=3.

[0045] FIG. 34 is a violin plot of percent CpG methylation for CpG sites downstream of the transcription start site of the *VEGFA* locus for each indicated LTRP -ZIM3 variant for the three *B2M*-targeting gRNA and non-targeting gRNA, as described in Example 7.

[0046] FIG. 35 is a scatterplot showing the relative activity (average percentage of HLA-negative cells at day 21 for spacer 7.160) versus specificity (percentage of off-target CpG methylation at the *VEGFA* locus quantified at day 7 for spacer 7.160) for the indicated LTRP5-ZIM3 variants, as described in Example 7.

[0047] FIG. 36 illustrates the schematics of the general architectures of the LTRP molecules with the ADD domain for LTRP configuration #1, #4, and #5 tested in Example 8. “D3A ADD”, “D3A CD” and “D3L ID” denote the ADD domain of DNMT3A, the catalytic domain of DNMT3A, and the interaction domain of DNMT3L respectively, as described in Example 8. “L1-L4” are linkers. “NLS” is the nuclear localization signal. See Table 11 for LTRP sequences.

[0048] FIG. 37 presents the results of a time-course experiment comparing B2M repression activities (represented as percentage of HLA-negative cells) of LTRPs with the ZIM3-KRAB domain having configuration #1, #4, or #5 with or without the DNMT3A ADD domain when paired with the *B2M*-targeting gRNA with spacer 7.160, as described in Example 8. Data are presented as mean with standard deviation, N=3. “NT” is a gRNA with a non-targeting spacer.

[0049] FIG. 38 is a plot showing the results of the same time-course experiment shown in FIG. 37, but illustrates B2M repression activities for LTRP #5 with the ZNF10 or ZIM3-

KRAB domain, with or without the DNMT3A ADD domain, paired with the *B2M*-targeting gRNA with spacer 7.160, as described in Example 8. Data are presented as mean with standard deviation, N=3. “NT” is a gRNA with a non-targeting spacer.

[0050] FIG. 39 is a plot showing the results of the same time-course experiment shown in FIG. 37, but illustrates B2M repression activities for LTRP5-ZIM3 with or without the DNMT3A ADD domain paired with a *B2M*-targeting gRNA with the indicated spacers, as described in Example 8. Data are presented as mean with standard deviation, N=3. “NT” is a gRNA with a non-targeting spacer.

[0051] FIG. 40 is a plot illustrating the results of B2M repression activities on day 27 post-transfection for LTRPs with either the ZNF10 or ZIM3-KRAB domain having configuration #1 with or without the DNMT3A ADD domain for the indicated gRNAs, as described in Example 8. Data are presented as mean with standard deviation, N=3. “NT” is a gRNA with a non-targeting spacer.

[0052] FIG. 41 is a plot illustrating the results of B2M repression activities on day 27 post-transfection for LTRPs with either the ZNF10 or ZIM3-KRAB domain having configuration #4 with or without the DNMT3A ADD domain for the indicated gRNAs, as described in Example 8. Data are presented as mean with standard deviation, N=3. “NT” is a gRNA with a non-targeting spacer.

[0053] FIG. 42 is a plot illustrating the results of B2M repression activities on day 27 post-transfection for LTRPs with either the ZNF10 or ZIM3-KRAB domain having configuration #5 with or without the DNMT3A ADD domain for the indicated gRNAs, as described in Example 8. Data are presented as mean with standard deviation, N=3. “NT” is a gRNA with a non-targeting spacer.

[0054] FIG. 43 is a plot illustrating the results of bisulfite sequencing used to determine off-target methylation at the *VEGFA* locus on day 5 post-transfection for LTRPs with either the ZNF10 or ZIM3-KRAB domain having configuration #1 with or without the DNMT3A ADD domain for the indicated gRNAs, as described in Example 8. Data are presented as mean percentage of CpG methylation for CpG sites near the *VEGFA* locus; standard error of the mean is also presented; N=3. “NT” is a gRNA with a non-targeting spacer.

[0055] FIG. 44 is a plot illustrating the results of bisulfite sequencing used to determine off-target methylation at the *VEGFA* locus on day 5 post-transfection for LTRPs with either the ZNF10 or ZIM3-KRAB domain having configuration #4 with or without the DNMT3A ADD domain for the indicated gRNAs, as described in Example 8. Data are presented as

mean percentage of CpG methylation for CpG sites near the *VEGFA* locus; standard error of the mean is also presented; N=3. “NT” is a gRNA with a non-targeting spacer.

[0056] FIG. 45 is a plot illustrating the results of bisulfite sequencing used to determine off-target methylation at the *VEGFA* locus on day 5 post-transfection for LTRPs with either the ZNF10 or ZIM3-KRAB domain having configuration #5 with or without the DNMT3A ADD domain for the indicated gRNAs, as described in Example 8. Data are presented as mean percentage of CpG methylation for CpG sites near the *VEGFA* locus; standard error of the mean is also presented; N=3. “NT” is a gRNA with a non-targeting spacer.

[0057] FIG. 46 is a dot plot showing the relative activity (average percentage of HLA-negative cells at day 27) versus specificity (percentage of off-target CpG methylation at the *VEGFA* locus quantified at day 5) for the LTRP molecules with the ZIM3-KRAB domain having configurations #1, #4, and #5, for *B2M*-targeting gRNA with spacer 7.160, as described in Example 8.

[0058] FIG. 47 is a dot plot showing the relative activity (average percentage of HLA-negative cells at day 27) versus specificity (percentage of off-target CpG methylation at the *VEGFA* locus quantified at day 5) for the LTRP molecules with the ZNF10-KRAB domain having configurations #1, #4, and #5, for *B2M*-targeting gRNA with spacer 7.160, as described in Example 8.

[0059] FIG. 48 is a dot plot showing the relative activity (average percentage of HLA-negative cells at day 27) versus specificity (percentage of off-target CpG methylation at the *VEGFA* locus quantified at day 5) for the LTRP molecules with the ZIM3-KRAB domain having configurations #1, #4, and #5, for *B2M*-targeting gRNA with spacer 7.37, as described in Example 8.

[0060] FIG. 49 is a dot plot showing the relative activity (average percentage of HLA-negative cells at day 27) versus specificity (percentage of off-target CpG methylation at the *VEGFA* locus quantified at day 5) for the LTRP molecules with the ZNF10-KRAB domain having configurations #1, #4, and #5, for *B2M*-targeting gRNA with spacer 7.37, as described in Example 8.

[0061] FIG. 50 is a dot plot showing the relative activity (average percentage of HLA-negative cells at day 27) versus specificity (percentage of off-target CpG methylation at the *VEGFA* locus quantified at day 5) for the LTRP molecules with the ZIM3-KRAB domain having configurations #1, #4, and #5, for *B2M*-targeting gRNA with spacer 7.165, as described in Example 8.

[0062] FIG. 51 is a dot plot showing the relative activity (average percentage of HLA-negative cells at day 27) versus specificity (percentage of off-target CpG methylation at the VEGFA locus quantified at day 5) for the LTRP molecules with the ZNF10-KRAB domain having configurations #1, #4, and #5, for B2M-targeting gRNA with spacer 7.165, as described in Example 8.

[0063] FIG. 52 is a plot illustrating percentage of HEK293T cells, transfected with a plasmid encoding the indicated CasX or LTRP:gRNA construct, that expressed B2M six days post-treatment with the DNMT1 inhibitor 5-azadC at varying concentrations, as described in Example 9.

[0064] FIG. 53 is a plot that juxtaposes the quantification of B2M repression in HEK293T cells transfected with a plasmid encoding the indicated CasX or LTRP:gRNA construct and cultured for 58 days, with the quantification of B2M reactivation upon treatment of transfected cells with 5-azadC, as described in Example 9.

[0065] FIG. 54 illustrates the schematic of an LTRP5 molecule without the ADD domain of the DNMT3A, as described in Example 11. “D3A CD” and “D3L ID” denote the catalytic domain of DNMT3A and the interaction domain of DNMT3L respectively. “L1”, “L2”, “L3A”, and “L3B” are linkers. “NLS” is the nuclear localization signal. “RD1” denotes a repressor domain.

[0066] FIG. 55 is a bar graph illustrating the results of a time-course experiment comparing levels of B2M repression (represented as average percentage of HLA-negative cells) in HEK293T cells transfected with the LTRP5 variant plasmids containing linker sets 1-11, as described in Example 13. Data for each timepoint (day 8, day 15, and day 45) are superimposed and presented as mean with standard deviation, N =3. A non-targeting spacer (NT) was included as experimental control.

[0067] FIG. 56 is a bar graph illustrating the results of a time-course experiment comparing the level of repression of target 1 (represented as percentage of total cells with knockdown of target 1) of LTRP5 variants containing linker sets 1-11 measured in HEK293T, as described in Example 13. Data for each timepoint (day 8, day 15, and day 45) are superimposed and presented as mean with standard deviation, N =3.

[0068] FIG. 57 is a bar graph illustrating the results of a time-course experiment comparing the level of repression of target 2 (represented as percentage of total cells with knockdown of target 2) of LTRP5 variants containing linker sets 1-11 measured in HEK293T, as described in Example 13. Data for each timepoint (day 8, day 15, and day 45) are superimposed and

presented as mean with standard deviation, N =3. A non-targeting spacer (NT) was included as experimental control.

[069] FIG. 58 is a bar graph illustrating the results of a time-course experiment comparing levels of B2M repression (represented as average percentage of HLA-negative cells) in HEK293T cells transfected with the LTRP5 variant plasmids containing linker sets 12-28, as described in Example 13. Data for each timepoint (day 7 and day 17) are superimposed and presented as mean with standard deviation, N =3. A non-targeting spacer (NT) was included as experimental control. A non-targeting spacer (NT) was included as experimental control.

[070] FIG. 59 is a bar graph illustrating the results of a time-course experiment comparing the level of repression of target 1 (represented as percentage of total cells with knockdown of target 1) of LTRP5 variants containing linker sets 12-28 measured in HEK293T, as described in Example 13. Data for each timepoint (day 7 and day 17) are superimposed and presented as mean with standard deviation, N =3. A non-targeting spacer (NT) was included as experimental control.

[071] FIG. 60 is a bar graph illustrating the results of a time-course experiment comparing the level of repression of target 2 (represented as percentage of total cells with knockdown of target 2) of LTRP5 variants containing linker sets 12-28 measured in HEK293T, as described in Example 13. Data for each timepoint (day 7 and day 17) are superimposed and presented as mean with standard deviation, N =3. A non-targeting spacer (NT) was included as experimental control.

[072] FIG. 61 illustrates the schematics of various configurations of LTRP molecules with the DNMT3A ADD domain. “D3A ADD”, “D3A CD”, and “D3L ID” denote the ADD domain of DNMT3A, the catalytic domain of DNMT3A, and the interaction domain of DNMT3L respectively. “L1”, “L2”, “L3A”, “L3B”, and “L4” are linkers. “NLS” is the nuclear localization signal. “RD1” denotes a repressor domain, and “RD1a” and “RD1b” represent repressor domain variants.

[073] FIG. 62 is a violin plot with each point representing mean % methylation at individual CpGs. Median methylation is indicated by dashed lines, with upper and lower quartiles indicated by dotted lines. Transcriptional start site (TSS)-proximal DNA methylation was measured from homogenized liver-extracted gDNA by amplicon enzymatic methylation sequencing (EM-seq) from N=3 mice sacrificed at days 7, 14, and 42 post-treatment, as described in example 14.

[074] FIG. 63 is a violin plot with each point representing mean % methylation at individual CpGs. Median methylation is indicated by dashed lines, with upper and lower quartiles indicated by dotted lines. Transcriptional start site (TSS)-proximal DNA methylation was measured from homogenized liver-extracted gDNA by amplicon enzymatic methylation sequencing (EM-seq) from N=3 mice sacrificed at day 7 post-treatment, as described in example 14.

DETAILED DESCRIPTION

[075] While exemplary embodiments have been shown and described herein, it will be apparent to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the spirit of the disclosure. Thus, it is understood that various alternatives to the embodiments specifically described herein may be employed in practicing the embodiments of the disclosure.

[076] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present embodiments, suitable methods and materials are described below. In case of conflict, the patent specification, including definitions, will control. In addition, the materials, methods, and examples disclosed herein are illustrative only and not intended to be limiting. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the spirit of the disclosure and the claims.

Definitions

[077] As used in this specification and the appended claims, the singular forms “a,” “an,” and “the” include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to “a host cell” includes two or more such host cells, reference to “a dCasX protein” includes one or more dCasX protein(s), reference to “a nucleic acid sequence” includes one or more nucleic acid sequences, and the like.

[078] As used herein, the term “about” is understood by persons of ordinary skill in the art and may vary to some extent depending upon the context in which it is used. If there are uses of the term which are not clear to persons of ordinary skill in the art given the context in

which the term “about” is used, “about” will mean up to plus or minus 10% of the particular term.

[079] As will be understood by one skilled in the art, for any and all purposes, all ranges disclosed herein also include any and all possible subranges and combinations of subranges thereof. Furthermore, as will be understood by one skilled in the art, a range includes each individual member. Thus, for example, a group having 1-3 members refers to groups having 1, 2, or 3 members. Similarly, a group having 1-5 members refers to groups having 1, 2, 3, 4, or 5 members, and so forth.

[080] The term “combinations thereof” includes every possible combination of elements to which the term refers.

[081] The term “exemplary” as used herein, refers to an example or illustration, and is not intended to imply any preference or value.

[082] The term “CasX protein,” as used herein, refers to a family of proteins, including *e.g.*, all naturally-occurring CasX proteins (“reference CasX”), as well as CasX proteins with sequence modifications, such as a dCasX, possessing one or more improved characteristics relative to a CasX protein from which it was derived, described more fully, herein below.

[083] The terms “polynucleotide” and “nucleic acid,” used interchangeably herein, refer to a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides. Thus, terms “polynucleotide” and “nucleic acid” encompass single-stranded DNA; double-stranded DNA; multi-stranded DNA; single-stranded RNA; double-stranded RNA; multi-stranded RNA; genomic DNA; cDNA; DNA-RNA hybrids; and a polymer comprising purine and pyrimidine bases or other natural, chemically or biochemically modified, non-natural, or derivatized nucleotide bases.

[084] “Hybridizable” or “complementary” are used interchangeably to mean that a nucleic acid (*e.g.*, RNA, DNA) comprises a sequence of nucleotides that enables it to non-covalently bind, *i.e.*, form Watson-Crick base pairs and/or G/U base pairs, “anneal”, or “hybridize,” to another nucleic acid in a sequence-specific, antiparallel, manner (*i.e.*, a nucleic acid specifically binds to a complementary nucleic acid) under the appropriate *in vitro* and/or *in vivo* conditions of temperature and solution ionic strength. It is understood that the sequence of a polynucleotide need not be 100% complementary to that of its target nucleic acid to be specifically hybridizable; it can have at least about 70%, at least about 80%, or at least about 90%, or at least about 95% sequence identity and still hybridize to the target nucleic acid. Moreover, a polynucleotide may hybridize over one or more segments such that intervening

or adjacent segments are not involved in the hybridization event (*e.g.*, a loop structure or hairpin structure, a 'bulge', 'bubble' and the like). Thus, the skilled artisan will understand that while individual bases within a sequence may not be complementary to another sequence, the sequence as a whole is still considered to be complementary.

[085] A "gene," for the purposes of the present disclosure, includes a DNA region encoding a gene product (*e.g.*, a protein, RNA), as well as all DNA regions which regulate the production of the gene product, whether or not such regulatory sequences are adjacent to coding and/or transcribed sequences. Accordingly, a gene may include accessory element sequences including, but not necessarily limited to, promoter sequences, terminators, translational regulatory sequences such as ribosome binding sites and internal ribosome entry sites, enhancers, silencers, insulators, boundary elements, replication origins, matrix attachment sites and locus control regions. Coding sequences encode a gene product upon transcription or transcription and translation; the coding sequences of the disclosure may comprise fragments and need not contain a full-length open reading frame. A gene can include both the strand that is transcribed as well as the complementary strand containing the anticodons.

[086] The term "downstream" refers to a nucleotide sequence that is located 3' to a reference nucleotide sequence. In certain embodiments, downstream nucleotide sequences relate to sequences that follow the starting point of transcription. For example, the translation initiation codon of a gene is located downstream of the start site of transcription.

[087] The term "upstream" refers to a nucleotide sequence that is located 5' to a reference nucleotide sequence. In certain embodiments, upstream nucleotide sequences relate to sequences that are located on the 5' side of a coding region or starting point of transcription. For example, most promoters are located upstream of the start site of transcription.

[088] The term "adjacent to" with respect to polynucleotide or amino acid sequences refers to sequences that are next to, or adjoining each other in a polynucleotide or polypeptide. The skilled artisan will appreciate that two sequences can be considered to be adjacent to each other and still encompass a limited amount of intervening sequence, *e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 nucleotides or amino acids.

[089] The term "regulatory element" is used interchangeably herein with the term "regulatory sequence," and is intended to include promoters, enhancers, and other expression regulatory elements. It will be understood that the choice of the appropriate regulatory element will depend on the encoded component to be expressed (*e.g.*, protein or RNA) or

whether the nucleic acid comprises multiple components that require different polymerases or are not intended to be expressed as a fusion protein.

[090] The term “accessory element” is used interchangeably herein with the term “accessory sequence,” and is intended to include coding and non-coding sequences that enhance expression, trafficking of the nucleic acid, or the function of an mRNA or protein and include, *inter alia*, polyadenylation signals (poly(A) signal), enhancer elements, introns, posttranscriptional regulatory elements (PTREs), nuclear localization signals (NLS), deaminases, DNA glycosylase inhibitors, additional promoters, factors that stimulate CRISPR-mediated homology-directed repair (e.g., in cis or in trans), self-cleaving sequences, and fusion domains, for example a fusion domain fused to a CRISPR protein. It will be understood that the choice of the appropriate accessory element or elements will depend on the encoded component to be expressed (e.g., protein or RNA) or whether the nucleic acid comprises multiple components that require different polymerases or are not intended to be expressed as a fusion protein.

[091] The term “promoter” refers to a DNA sequence that contains a transcription start site and additional sequences to facilitate polymerase binding and transcription. Exemplary eukaryotic promoters include elements such as a TATA box, and/or B recognition element (BRE) and assists or promotes the transcription and expression of an associated transcribable polynucleotide sequence and/or gene (or transgene). A promoter can be synthetically produced or can be derived from a known or naturally occurring promoter sequence or another promoter sequence. A promoter can be proximal or distal to the gene to be transcribed. A promoter can also include a chimeric promoter comprising a combination of two or more heterologous sequences to confer certain properties. A promoter of the present disclosure can include variants of promoter sequences that are similar in composition, but not identical to, other promoter sequence(s) known or provided herein. A promoter can be classified according to criteria relating to the pattern of expression of an associated coding or transcribable sequence or gene operably linked to the promoter, such as constitutive, developmental, tissue-specific, inducible, etc. A promoter can also be classified according to its strength. As used in the context of a promoter, “strength” refers to the rate of transcription of the gene controlled by the promoter. A “strong” promoter means the rate of transcription is high, while a “weak” promoter means the rate of transcription is relatively low.

[092] A promoter of the disclosure can be a Polymerase II (Pol II) promoter. Polymerase II transcribes all protein coding and many non-coding genes. A representative Pol II promoter

includes a core promoter, which is a sequence of about 100 base pairs surrounding the transcription start site, and serves as a binding platform for the Pol II polymerase and associated general transcription factors. The promoter may contain one or more core promoter elements such as the TATA box, BRE, Initiator (INR), motif ten element (MTE), downstream core promoter element (DPE), downstream core element (DCE), although core promoters lacking these elements are known in the art. All Pol III promoters are envisaged as within the scope of the instant disclosure.

[093] A promoter of the disclosure can be a Polymerase III (Pol III) promoter. Pol III transcribes DNA to synthesize small ribosomal RNAs such as the 5S rRNA, tRNAs, and other small RNAs. Representative Pol III promoters use internal control sequences (sequences within the transcribed section of the gene) to support transcription, although upstream elements such as the TATA box are also sometimes used. All Pol III promoters are envisaged as within the scope of the instant disclosure.

[094] The term “enhancer” refers to regulatory DNA sequences that, when bound by specific proteins called transcription factors, regulate the expression of an associated gene. Enhancers may be located in the intron of the gene, or 5’ or 3’ of the coding sequence of the gene. Enhancers may be proximal to the gene (*i.e.*, within a few tens or hundreds of base pairs (bp) of the promoter), or may be located distal to the gene (*i.e.*, thousands of bp, hundreds of thousands of bp, or even millions of bp away from the promoter). A single gene may be regulated by more than one enhancer, all of which are envisaged as within the scope of the instant disclosure.

[095] As used herein, a “post-transcriptional regulatory element (PTRE),” such as a hepatitis PTRE, refers to a DNA sequence that, when transcribed creates a tertiary structure capable of exhibiting post-transcriptional activity to enhance or promote expression of an associated gene operably linked thereto.

[096] “Operably linked” means with reference to a juxtaposition of two or more components (such as sequence elements), in which the components are arranged such that both components function normally and allow the possibility that at least one of the components can mediate a function that is exerted upon at least one of the other components; e.g., a promoter and an encoding sequence. The skilled artisan will appreciate that the two components need not be physically linked to be operably linked.

[097] In the context of the present disclosure and with respect to a gene, “repress”, “repression”, “repressing”, “transcriptional repression”, “inhibition of gene expression”,

“downregulation”, and “silencing” are used interchangeably herein to refer to the inhibition or blocking of transcription of a gene or a portion thereof. Accordingly, repression of a gene can result in a decrease in production of a gene product. Examples of gene repression processes which decrease transcription include, but are not limited to, those which inhibit formation of a transcription initiation complex, those which decrease transcription initiation rate, those which decrease transcription elongation rate, those which decrease processivity of transcription and those which antagonize transcriptional activation (by, for example, blocking the binding of a transcriptional activator). Gene repression can constitute, for example, prevention of activation as well as inhibition of expression below an existing level. Transcriptional repression includes both reversible and irreversible inactivation of gene transcription; the latter can result from epigenetic modification of the gene.

[098] “Repressor” or “repressor domain” are used interchangeably to refer to polypeptide factors that act as regulatory elements on DNA that inhibit, repress, or block transcription of DNA, resulting in repression of gene expression. In the context of the present disclosure, the linking of a repressor domain to DNA binding protein that can, when bound to the target nucleic acid, prevent transcription from a promoter or otherwise inhibit the expression of a gene. Without wishing to be bound by theory, it is thought that transcriptional repressors can function by a variety of mechanisms, including physically blocking RNA polymerase passage by steric hindrance, altering the polymerase's post-translational modification state, modifying the epigenetic state of the nascent RNA, changing the epigenetic state of the DNA through methylation, changing the epigenetic state of the DNA through histone deacetylation or modulating nucleosome remodeling, or preventing enhancer-promoter interactions, thereby leading to gene silencing or a reduction in the level of gene expression.

[099] “Long-term repressor fusion protein” or “LTRP” is used interchangeably herein with “repressor fusion protein” and refers to a fusion protein comprising a DNA binding protein (or DNA binding domain of a protein) fused to one or more domains capable of repressing transcription of a target nucleic acid sequence. Optionally, the long-term repressor fusion proteins of the disclosure may contain additional elements, such as linkers between any of the domains of the fusion protein, nuclear localization signals, nuclear export signals, as well as additional protein domains that confer additional activities upon the long-term repressor fusion protein.

[0100] As used herein a “LTRP: gRNA system” is a system for transcriptional repression and comprises a long-term repressor fusion protein comprising a catalytically-dead CRISPR

protein and one or more linked repressor domains, and a guide nucleic acid (gRNA) that binds to the catalytically-dead CRISPR protein. For clarity, the system also includes any encoding DNA, RNA or vectors and the like that can be used to produce the long-term repressor fusion proteins and gRNA components of the system.

[0101] As used herein, a "catalytically dead DNA binding protein" refers to a protein, or domain of a protein, capable of binding to DNA but is unable to nick or cleave the DNA. As used herein a "catalytically-dead CRISPR protein" refers to a CRISPR protein that lacks endonuclease activity. The skilled artisan will appreciate that a CRISPR protein can be catalytically-dead, and still able to carry out additional protein functions, such as DNA binding. Similarly, a "catalytically-dead CasX" refers to a CasX protein that lacks endonuclease activity but is still able to carry out additional protein functions, such as DNA binding.

[0102] "Recombinant," as used herein, means that a particular nucleic acid (DNA or RNA) is the product of various combinations of cloning, restriction, and/or ligation steps resulting in a construct having a structural coding or non-coding sequence distinguishable from endogenous nucleic acids found in natural systems. Generally, DNA sequences encoding the structural coding sequence can be assembled from cDNA fragments and short oligonucleotide linkers, or from a series of synthetic oligonucleotides, to provide a synthetic nucleic acid which is capable of being expressed from a recombinant transcriptional unit contained in a cell or in a cell-free transcription and translation system. Such sequences can be provided in the form of an open reading frame uninterrupted by internal non-translated sequences, or introns, which are typically present in eukaryotic genes. Genomic DNA comprising the relevant sequences can also be used in the formation of a recombinant gene or transcriptional unit. Sequences of non-translated DNA may be present 5' or 3' from the open reading frame, where such sequences do not interfere with manipulation or expression of the coding regions, and may indeed act to modulate production of a desired product by various mechanisms (see "enhancers" and "promoters", above).

[0103] The term "recombinant polynucleotide" or "recombinant nucleic acid" refers to one which is not naturally occurring, *e.g.*, is made by the artificial combination of two otherwise separated segments of sequence through human intervention. This artificial combination is often accomplished by either chemical synthesis means, or by the artificial manipulation of isolated segments of nucleic acids, *e.g.*, by genetic engineering techniques. Such is usually done to replace a codon with a redundant codon encoding the same or a conservative amino

acid, while typically introducing or removing a sequence recognition site. Alternatively, it is performed to join together nucleic acid segments of desired functions to generate a desired combination of functions. This artificial combination is often accomplished by either chemical synthesis means, or by the artificial manipulation of isolated segments of nucleic acids, *e.g.*, by genetic engineering techniques.

[0104] Similarly, the term “recombinant polypeptide” or “recombinant protein” refers to a polypeptide or protein which is not naturally occurring, *e.g.*, is made by the artificial combination of two otherwise separated segments of amino sequence through human intervention. Thus, *e.g.*, a protein that comprises a heterologous amino acid sequence is recombinant.

[0105] As used herein, “lipid nanoparticle” or “LNP” refers to a particle having at least one dimension on the order of nanometers (*e.g.*, 1-1,000 nm) comprising one or more lipids (*e.g.*, cationic lipids, non-cationic lipids, helper phospholipids, and PEG-modified lipids), as well as cholesterol. Specific components of LNP are described more fully, below. Lipid nanoparticles can be included in a formulation that can be used to deliver an active agent or therapeutic agent, such as a nucleic acid (*e.g.*, mRNA) to a target site of interest (*e.g.*, cell, tissue, organ, tumor, and the like). The lipid nanoparticles of the disclosure can comprise a nucleic acid. Such lipid nanoparticles typically comprise neutral lipids, charged lipids, steroids and polymer conjugated lipids. The active agent or therapeutic agent, such as a nucleic acid, may be encapsulated in the lipid portion of the lipid nanoparticle or an aqueous space enveloped by some or all of the lipid portion of the lipid nanoparticle, thereby protecting it from enzymatic degradation or other undesirable effects induced by the mechanisms of the host organism or cells *e.g.* an adverse immune response.

[0106] As used herein, “lipid encapsulated” refers to a lipid nanoparticle that provides an active agent or therapeutic agent, such as a nucleic acid (*e.g.*, mRNA, gRNA or both mRNA and gRNA), with full encapsulation, partial encapsulation, or both. The nucleic acid (*e.g.*, mRNA) can be fully encapsulated in the lipid nanoparticle.

[0107] As used herein, “lipoprotein”, such as VLDL, LDL and HDL, refers to a group of proteins found in the serum, plasma and lymph that function in lipid transport. The chemical composition of each lipoprotein differs, for example, in that the HDL has a higher proportion of protein versus lipid, whereas the VLDL has a lower proportion of protein versus lipid.

[0108] As used herein, “atherosclerosis” means a hardening of the arteries affecting large and medium-sized arteries and is characterized by the presence of fatty deposits. The fatty

deposits are called “atheromas” or “plaques,” which consist mainly of cholesterol and other fats, calcium and scar tissue, and damage the lining of arteries.

[0109] As used herein, “coronary heart disease (CHD)” means a narrowing of the small blood vessels that supply blood and oxygen to the heart, which is often a result of atherosclerosis.

[0110] As used herein, “dyslipidemia” refers to a disorder of lipid and/or lipoprotein metabolism, including lipid and/or lipoprotein overproduction or deficiency. Dyslipidemias can be manifested by elevation of lipids such as chylomicron, cholesterol and triglycerides as well as lipoproteins such as low-density lipoprotein (LDL) cholesterol.

[0111] As used herein, “high density lipoprotein-C” or “HDL-C” means cholesterol associated with high-density lipoprotein particles. Concentration of HDL-C in serum (or plasma) is typically quantified in mg/dL or nmol/L. “Serum HDL-C” and “plasma HDL-C” mean HDL-C in serum and plasma, respectively.

[0112] As used herein, “low density lipoprotein-cholesterol (LDL-C)” means cholesterol carried in low density lipoprotein particles. Concentration of LDL-C in serum (or plasma) is typically quantified in mg/dL or nmol/L. “Serum LDL-C” and “plasma LDL-C” mean LDL-C in the serum and plasma, respectively.

[0113] As used herein, “hypercholesterolemia” means a condition characterized by elevated cholesterol or circulating (plasma) cholesterol, LDL-cholesterol and VLDL-cholesterol, as per the guidelines of the Expert Panel Report of the National Cholesterol Educational Program (NCEP) of Detection, Evaluation of Treatment of high cholesterol in adults (*see*, Arch. Int. Med. 148: 36 (1988)).

[0114] As used herein, “hyperlipidemia” or “hyperlipemia” is a condition characterized by elevated serum lipids or circulating (plasma) lipids. This condition manifests an abnormally high concentration of fats. The lipid fractions in the circulating blood are cholesterol, low-density lipoproteins, very low density lipoproteins, chylomicrons and triglycerides. The Fredrickson classification of hyperlipidemias is based on the pattern of TG and cholesterol-rich lipoprotein particles, as measured by electrophoresis or ultracentrifugation and is commonly used to characterize primary causes of hyperlipidemias such as hypertriglyceridemia.

[0115] As used herein, “triglyceride” or “TG” means a lipid or neutral fat consisting of glycerol combined with three fatty acid molecules.

[0116] As used herein, “hypertriglyceridemia” means a condition characterized by elevated triglyceride levels. Its etiology includes primary (*i.e.* genetic causes) and secondary (other underlying causes such as diabetes, metabolic syndrome/insulin resistance, obesity, physical inactivity, cigarette smoking, excess alcohol and a diet very high in carbohydrates) factors or, most often, a combination of both

[0117] As used herein, “diabetes mellitus” or “diabetes” is a syndrome characterized by disordered metabolism and abnormally high blood sugar (hyperglycemia) resulting from insufficient levels of insulin or reduced insulin sensitivity. The characteristic symptoms are excessive urine production (polyuria) due to high blood glucose levels, excessive thirst and increased fluid intake (polydipsia) attempting to compensate for increased urination, blurred vision due to high blood glucose effects on the eye's optics, unexplained weight loss, and lethargy.

[0118] As used herein, “diabetic dyslipidemia” or “type 2 diabetes with dyslipidemia” means a condition characterized by Type 2 diabetes, reduced HDL-C, elevated triglycerides (TG), and elevated small, dense LDL particles.

[0119] As used herein, the term “contacting” means establishing a physical connection between two or more entities. For example, contacting a target nucleic acid with a guide nucleic acid means that the target nucleic acid and the guide nucleic acid are made to share a physical connection; *e.g.*, can hybridize if the sequences share sequence similarity.

[0120] “Dissociation constant”, or “ K_d ”, are used interchangeably and mean the affinity between a ligand “L” and a protein “P”; *i.e.*, how tightly a ligand binds to a particular protein. It can be calculated using the formula $K_d = [L][P]/[LP]$, where [P], [L] and [LP] represent molar concentrations of the protein, ligand and complex, respectively.

[0121] The term “knock-down” as used herein refers to reduction in the expression of a gene or its gene product(s). As a result of a gene knock-down, the protein activity or function may be attenuated or the protein levels may be reduced or eliminated.

[0122] A polynucleotide or polypeptide has a certain percent “sequence similarity” or “sequence identity” to another polynucleotide or polypeptide, meaning that, when aligned, that percentage of bases or amino acids are the same, and in the same relative position, when comparing the two sequences. Sequence similarity (sometimes referred to as percent similarity, percent identity, or homology) can be determined in a number of different manners. To determine sequence similarity, sequences can be aligned using the methods and computer programs that are known in the art, including BLAST, available over the world

wide web at ncbi.nlm.nih.gov/BLAST. Percent complementarity between particular stretches of nucleic acid sequences within nucleic acids can be determined using any convenient method. Example methods include BLAST programs (basic local alignment search tools) and PowerBLAST programs (Altschul et al., J. Mol. Biol., 1990, 215, 403-410; Zhang and Madden, Genome Res., 1997, 7, 649-656) or by using the Gap program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, Madison Wis.), *e.g.*, using default settings, which uses the algorithm of Smith and Waterman (Adv. Appl. Math., 1981, 2, 482-489).

[0123] The terms “polypeptide,” and “protein” are used interchangeably herein, and refer to a polymeric form of amino acids of any length, which can include coded and non-coded amino acids, chemically or biochemically modified or derivatized amino acids, and polypeptides having modified peptide backbones. The term includes fusion proteins, including, but not limited to, fusion proteins with a heterologous amino acid sequence.

[0124] A “vector” or “expression vector” is a replicon, such as plasmid, phage, virus, or cosmid, which can include another DNA segment, *i.e.*, an expression cassette, so as to bring about the replication or expression of the other DNA segment in a cell.

[0125] The term “naturally-occurring” or “unmodified” or “wild type” as used herein as applied to a nucleic acid, a polypeptide, a cell, or an organism, refers to a nucleic acid, polypeptide, cell, or organism that is found in nature.

[0126] As used herein, a “mutation” refers to an insertion, deletion, substitution, duplication, or inversion of one or more amino acids or nucleotides as compared to a wild-type or reference amino acid sequence or to a wild-type or reference nucleotide sequence.

[0127] As used herein the term “isolated” is meant to describe a polynucleotide, a polypeptide, or a cell that is in an environment different from that in which the polynucleotide, the polypeptide, or the cell naturally occurs. An isolated genetically modified host cell may be present in a mixed population of genetically modified host cells.

[0128] A “host cell,” as used herein, denotes a eukaryotic cell, a prokaryotic cell, or a cell from a multicellular organism (*e.g.*, a cell line) cultured as a unicellular entity, which eukaryotic or prokaryotic cells are used as recipients for a nucleic acid (*e.g.*, an AAV vector), and include the progeny of the original cell which has been genetically modified by the nucleic acid. It is understood that the progeny of a single cell may not necessarily be completely identical in morphology or in genomic or total DNA complement as the original parent, due to natural, accidental, or deliberate mutation. A “recombinant host cell” (also

referred to as a “genetically modified host cell”) is a host cell into which has been introduced a heterologous nucleic acid, *e.g.*, an AAV vector.

[0129] The term "conservative amino acid substitution" refers to the interchangeability in proteins of amino acid residues having similar side chains. For example, a group of amino acids having aliphatic side chains consists of glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains consists of serine and threonine; a group of amino acids having amide-containing side chains consists of asparagine and glutamine; a group of amino acids having aromatic side chains consists of phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains consists of lysine, arginine, and histidine; and a group of amino acids having sulfur-containing side chains consists of cysteine and methionine. Exemplary conservative amino acid substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine.

[0130] As used herein, “treatment” or “treating,” are used interchangeably herein and refer to an approach for obtaining beneficial or desired results, including but not limited to a therapeutic benefit and/or a prophylactic benefit. By therapeutic benefit is meant eradication or amelioration of the underlying disorder or disease being treated. A therapeutic benefit can also be achieved with the eradication or amelioration of one or more of the symptoms or an improvement in one or more clinical parameters associated with the underlying disease such that an improvement is observed in the subject, notwithstanding that the subject may still be afflicted with the underlying disorder.

[0131] The terms “therapeutically effective amount” and “therapeutically effective dose”, as used herein, refer to an amount of a drug or a biologic, alone or as a part of a composition, that is capable of having any detectable, beneficial effect on any symptom, aspect, measured parameter or characteristics of a disease state or condition when administered in one or repeated doses to a subject such as a human or an experimental animal. Such effect need not be absolute to be beneficial.

[0132] As used herein, “administering” means a method of giving a dosage of a compound (*e.g.*, a composition of the disclosure) or a composition (*e.g.*, a pharmaceutical composition) to a subject.

[0133] A “subject” is a mammal. Mammals include, but are not limited to, domesticated animals, non-human primates, humans, dogs, rabbits, mice, rats and other rodents.

[0134] The term “low-density lipoprotein (LDL)” refers to one of the five major groups of lipoprotein, from least dense (lower weight-volume ratio particles) to most dense (larger weight-volume ratio particles): chylomicrons, very low-density lipoproteins (VLDL), low-density lipoproteins (LDL), intermediate-density lipoproteins (IDL), and high-density lipoproteins (HDL). Lipoproteins transfer lipids (fats) around the body in the extracellular fluid thereby facilitating the transfer of fats to the cells body via receptor-mediated endocytosis. An LDL particle is about 220-275 angstroms in diameter.

[0135] “Low-density lipoprotein (LDL) receptor” refers to a receptor protein of 839 amino acids (after removal of 21-amino acid signal peptide) that mediates the endocytosis of cholesterol-rich LDL particles. It is a cell-surface receptor that recognizes the apoprotein B100 and apoE protein found in chylomicron remnants and VLDL remnants (IDL) resulting in the binding and endocytosis of LDL-cholesterol. This process occurs in all nucleated cells, but mainly in the liver which removes approximately 70% of LDL from the circulation. The human LDLR gene is described in part in the NCBI database ([ncbi.nlm.nih.gov](https://www.ncbi.nlm.nih.gov)) as Reference Sequence NG_009060.1, which is incorporated by reference herein.

[0136] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

I. General Methods

[0137] The practice of the present invention employs, unless otherwise indicated, conventional techniques of immunology, biochemistry, chemistry, molecular biology, microbiology, cell biology, genomics and recombinant DNA, which can be found in such standard textbooks as *Molecular Cloning: A Laboratory Manual*, 3rd Ed. (Sambrook et al., Harbor Laboratory Press 2001); *Short Protocols in Molecular Biology*, 4th Ed. (Ausubel et al. eds., John Wiley & Sons 1999); *Protein Methods* (Bollag et al., John Wiley & Sons 1996); *Nonviral Vectors for Gene Therapy* (Wagner et al. eds., Academic Press 1999); *Viral Vectors* (Kaplift & Loewy eds., Academic Press 1995); *Immunology Methods Manual* (I. Lefkovits ed., Academic Press 1997); and *Cell and Tissue Culture: Laboratory Procedures in Biotechnology* (Doyle & Griffiths, John Wiley & Sons 1998), the disclosures of which are incorporated herein by reference.

[0138] Where a range of values is provided, it is understood that endpoints are included and that each intervening value, to the tenth of the unit of the lower limit unless the context

clearly dictates otherwise, between the upper and lower limit of that range and any other stated or intervening value in that stated range, is encompassed. The upper and lower limits of these smaller ranges may independently be included in the smaller ranges, and are also encompassed, 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.

[0139] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. All publications mentioned herein are incorporated herein by reference to disclose and describe the methods and/or materials in connection with which the publications are cited.

[0140] It will be appreciated that certain features of the disclosure, which are, for clarity, described in the context of separate embodiments, may also be provided in combination in a single embodiment. In other cases, various features of the disclosure, which are, for brevity, described in the context of a single embodiment, may also be provided separately or in any suitable sub-combination. It is intended that all combinations of the embodiments pertaining to the disclosure are specifically embraced by the present disclosure and are disclosed herein just as if each and every combination was individually and explicitly disclosed. In addition, all sub-combinations of the various embodiments and elements thereof are also specifically embraced by the present disclosure and are disclosed herein just as if each and every such sub-combination was individually and explicitly disclosed herein.

II. Systems for Epigenetic Modification and Repression of *PCSK9* Genes

[0141] In a first aspect, the present disclosure provides systems comprising or encoding a long-term repressor fusion protein ("LTRP") comprising a DNA binding protein and repressor domains capable of binding a target nucleic acid sequence of a *PCSK9* gene targeted for transcriptional repression, silencing, and/or epigenetic modification. The present disclosure also provides systems comprising or encoding the LTRP. In some cases, the system is designed to repress transcription of a *PCSK9* gene in eukaryotic cells having a mutation.

[0142] As used herein, a "system" is used interchangeably with "composition." The disclosure also provides nucleic acids encoding the systems provided herein. Also provided herein are methods of making the systems, as well as methods of using the systems, including methods of gene repression and/or epigenetic modification and methods of treatment of

PCSK9-related diseases. For clarity, the term “system” also includes any encoding DNA, RNA or vectors and the like that can be used to produce the repressor fusion proteins and gRNA components of the system.

[0143] In some embodiments, the DNA binding proteins for use in the long-term repressor fusion proteins of the disclosure comprise zinc finger (ZF) or TALE (transcription-activator-like effector) proteins, or DNA binding domains thereof, also referred to herein as a DNA binding protein, that bind but do not cleave the target nucleic acid. The DNA binding domain of a TALE is comprised of a tandem array of 33–34 amino acid (aa)-long customizable monomers that theoretically can be assembled to recognize any genetic sequence following a one-repeat-binds-one-base-pair recognition code (Jain, S., et al. TALEN outperforms Cas9 in editing heterochromatin target sites. Nat. Commun. 12:606 (2021)). The specificity of TALEs for binding DNA arises from two polymorphic amino acids, the so-called repeat variable diresidues (RVDs) located at positions 12 and 13 of a repeated unit. By re-arranging the repeats, the DNA binding specificities of TALE can be changed at will. Zinc finger proteins are transcription factors, where each finger recognizes 3–4 bases of DNA. By mixing and matching these finger modules, the ZFs can be customized for the sequence to be targeted. Exemplary ZFs that are capable of binding the *PCSK9* gene are described in WO2018049009A2.

[0144] In some embodiments, the DNA binding protein for use in the long-term repressor fusion protein is a catalytically-dead Class 1 or Class 2 CRISPR protein. Catalytically-dead CRISPR proteins are also referred to in the art as “catalytically inactive” CRISPR proteins. In one embodiment, the Class 2, Type II protein is a catalytically-dead Cas9. In another embodiment, the Class 2 CRISPR protein is selected from the group consisting of a Type II, Type V, or Type VI protein. In one embodiment, the Class 2 Type V protein is selected from the group consisting of Cas12a (Cpf1), Cas12b (C2c1), Cas12c (C2c3), Cas12d (CasY), Cas12e (CasX), Cas12f, Cas12g, Cas12h, Cas12i, Cas12j, Cas12k, Cas14, and/or CasΦ, in each case rendered catalytically-dead by specific mutations, as described herein. The CRISPR-based systems further comprise a guide nucleic acid, for example a guide ribonucleic acid (gRNA) with a targeting sequence complementary to the target sequence of a *PCSK9* gene to be bound and repressed by a complex of the fusion protein (the CRISPR protein and linked repressor domains) and the gRNA.

[0145] In some embodiments, the present disclosure provides systems comprising or encoding a long-term repressor fusion protein comprising a catalytically-dead CasX CRISPR

nuclease protein and linked repressor domains and a guide ribonucleic acid (gRNA) comprising a targeting sequence complementary to a target nucleic acid sequence of a *PCSK9* gene targeted for transcriptional repression, silencing, or downregulation, as well as nucleic acids encoding the long-term repressor fusion protein and/or gRNA. In some embodiments, the system comprises a long-term repressor fusion protein and a gRNA of the disclosure as a gene repressor pair ("LTRP:gRNA system") that are capable of forming a ribonucleoprotein (RNP) complex and binding a *PCSK9* target nucleic acid. In other cases, the disclosure provides systems of nucleic acids encoding the long-term repressor fusion protein and the gRNA. In still other cases, the disclosure provides systems of a gRNA and an mRNA encoding the long-term repressor fusion protein for use in certain particle formulations (*e.g.*, an LNP) described herein.

[0146] Also provided herein are methods of making long-term repressor fusion proteins and gRNA, as well as methods of using the LTRP:gRNA systems, including methods of gene repression and/or epigenetic modification of a *PCSK9* gene and methods of treatment of a *PCSK9*-related disease or disorder. The DNA binding proteins (*e.g.*, dCasX) and linked repressor domain(s), and the gRNA components, of the LTRP:gRNA systems and their features, as well as the delivery modalities and the methods of using the systems for the repression, down-regulation or silencing of a *PCSK9* gene are described more fully, below.

[0147] In some embodiments, the disclosure provides systems specifically designed to repress or silence transcription of the *PCSK9* gene. In some cases, the system is designed to repress transcription of the *PCSK9* gene in eukaryotic cells having a gain of function mutation. In some cases, the system is designed to repress transcription of the wild-type *PCSK9* gene in eukaryotic cells. In the alternative, the system is designed to repress transcription of a mutant allele of the *PCSK9* gene in eukaryotic cells. Generally, any portion of the *PCSK9* gene can be targeted using the programmable systems and methods provided herein, described more fully, herein.

[0148] The *PCSK9* gene encodes proprotein convertase subtilisin/kexin Type 9 ("PCSK9"), a protein that binds to the receptor for low-density lipoprotein particles (LDL) for transport of LDL into the cell. The *PCSK9* gene encompasses the sequence that spans chr1:55,039,476-55,064,853 of the human genome (GRCh38/hg38) (the notation refers to the chromosome 1 (chr1), starting at the 55,039,476 bp to 55,064,853 bp on chromosome 1 (Homo sapiens Updated Annotation Release 109.20190905, GRCh38.p13) (NCBI). The human *PCSK9* gene is described in part in the NCBI database (ncbi.nlm.nih.gov) as Reference Sequence

NG_009061.1, which is incorporated by reference herein. The *PCSK9* locus has 12 exons that produces an mRNA of 3636 bp encoding a 692-amino acid protein that, following its synthesis, undergoes an autocatalytic cleavage reaction that clips off the prodomain, resulting in an activated protein having 540 amino acids. The prodomain remains attached to the catalytic and resistin-like domains, likely because the prodomain serves as a chaperone and facilitates folding and secretion (Seidah, NG et al., Proc Natl Acad Sci USA 100(3):928 (2003)). The secretory proprotein convertase neural apoptosis-regulated convertase 1 (NARC-1): liver regeneration and neuronal differentiation (Seidah NG, et al.). This protein, also called neural apoptosis regulated convertase, is a serine protease belonging to the protease K subfamily of subtilases.

[0149] The human PCSK9 gene (HGNC:20001) encodes a protein (Q8NBP7) having the sequence

```
MGTVSSRRSWWPLPLLLLLLLLLLGPAGARAQEDGDYEELVLALRSEEDGLAEAPEHGTATFHR
CAKDPWRLPGTYVVVLKEETHLSQSERTARRLQAQAARRGYLTKILHVFHGLLPGLVKMSGDLLE
LALKLPHVDYIEEDSSVFAQSIPWNLERITPPRYRADEYQPPDGGSLEVEYLLDTSIQSDHREIEG
RVMVTD FENVPEEDGTRFHRQASKCD SHGTHLAGVVS GRDAGVAKGASMRSLRVLNQCQKGTVSGT
LIGLEFIRKSQQLVQPVGPLVLLPLAGGYSRVLNAACQRLARAGVVLVTAAGNFRDDACLYSPASA
PEVITVGATNAQDQPVTLGTLGTNFGRCVDLFAPGEDII GASSDCSTCFVSQSGTSQAAAHVAGIA
AMMLSAEPELT LAELRQRLIHFS AKDVINEAWFPEDQRVLT PNLVAALPPSTHGAGWQLFCRTVWS
AHS GPTRMATAVARCAPDEELLSCSSFSRSGKRRGERMEA QGGKLV CRAHNAFGGEGVYAIARCC L
LPQANCSVHTAPPAEASMGTRVHCHQQGHVLTGCS SHWEVEDLGTHKPPVLRPRGQPNQCVGHREA
SIHASCC HAPGLECKVKEHGIPAPQEQVTVACEEGWTLTGCSALPGTSHVLGAYAVDNTCVVRSRD
VSTTGSTSEGA VTAVAICCRSRHLAQASQELQ (SEQ ID NO: 1823).
```

III. Catalytically-dead Proteins for Use in the Repressor Systems

[0150] In some embodiments, the DNA binding proteins for use in the long-term repressor fusion proteins of the disclosure are zinc finger (ZF) or TALE (transcription-activator-like effector) proteins that can bind, but not cleave a *PCSK9* target nucleic acid.

[0151] In some embodiments, the DNA binding protein is a catalytically-dead Class 1 or Class 2 CRISPR protein. In some embodiments, the DNA binding protein is a Class 2, Type II CRISPR protein. In one embodiment, the Class 2, Type II protein is a catalytically-dead Cas9. In another embodiment, the Class 2 CRISPR protein is selected from the group consisting of a Type II, Type V, or Type VI protein. In one embodiment, the Class 2 CRISPR Type V protein is selected from the group consisting of Cas12a (Cpf1), Cas12b (C2c1), Cas12c (C2c3), Cas12d (CasY), Cas12e (CasX), Cas12f, Cas12g, Cas12h, Cas12i, Cas12j,

Cas12k, Cas14, and/or Cas Φ , in each case rendered catalytically-dead by specific mutations, as described herein. In another embodiment, the Class 2 CRISPR Type V protein is a catalytically-dead CasX protein.

[0152] CasX proteins of the disclosure, including dCasX, comprise the following domains: a non-target strand binding (NTSB) domain, a target strand loading (TSL) domain, a helical I domain, a helical II domain, an oligonucleotide binding domain (OBD), and a RuvC domain, and, in some cases, domains can be further divided into subdomains, as listed in Table 1.

[0153] In the context of the present disclosure, the CasX for use in the systems are catalytically-dead (dCasX); achieved by mutations introduced at select locations in the RuvC sequence, described below.

a. Reference CasX Proteins

[0154] The disclosure provides naturally-occurring CasX proteins (referred to herein as a “reference CasX protein”), which were subsequently modified to create the engineered dCasX of the disclosure. For example, reference CasX proteins can be isolated from naturally occurring prokaryotes, such as *Deltaproteobacteria*, *Planctomycetes*, or *Candidatus* *Sungbacteria* species. A reference CasX protein (interchangeably referred to herein as a reference CasX polypeptide) is a Class 2, Type V CRISPR/Cas endonuclease belonging to the CasX (interchangeably referred to as Cas12e) family of proteins that interacts with a guide RNA to form a ribonucleoprotein (RNP) complex.

[0155] In some cases, a reference CasX protein is isolated or derived from *Deltaproteobacter*, and comprises a sequence of SEQ ID NO: 1.

[0156] In some cases, a reference CasX protein is isolated or derived from *Planctomycetes*, and comprises a sequence of SEQ ID NO: 2.

[0157] In some cases, a reference CasX protein is isolated or derived from *Candidatus* *Sungbacteria*, and comprises a sequence of SEQ ID NO: 3.

b. Catalytically-dead Class 1 or Class 2 CRISPR Proteins

[0158] In the long-term repressor fusion proteins and gene repressor systems comprising same of the disclosure, the catalytically-dead Class 1 or Class 2 CRISPR protein is catalytically-dead in that it is unable to cleave DNA, but retains the ability to bind a target nucleic acid when complexed with a guide RNA (gRNA). The present disclosure provides catalytically-dead variants of Class 1 or Class 2 CRISPR proteins, wherein the catalytically-dead variants comprise multiple modifications in select domains. The present disclosure provides catalytically-dead CasX variants (interchangeably referred to herein as “dCasX

variant” or “dCasX variant protein”), wherein the catalytically-dead CasX variants comprise multiple modifications in the RuvC domain relative to the catalytically-dead versions of reference CasX proteins comprising sequences of SEQ ID NOS:1-3 (described, *supra*). In some embodiments, a catalytically-dead reference CasX protein comprises substitutions at residues 672, 769 and/or 935 with reference to SEQ ID NO: 1. In some embodiments, a catalytically-dead reference CasX protein comprises substitutions of D672A, E769A and/or D935A with reference to SEQ ID NO: 1. In other embodiments, a catalytically-dead reference CasX protein comprises substitutions at amino acids 659, 756 and/or 922 with reference to SEQ ID NO: 2. In some embodiments, a catalytically-dead reference CasX protein comprises D659A, E756A and/or D922A substitutions with reference to of SEQ ID NO: 2. An exemplary RuvC domain of the dCasX of the disclosure comprises amino acids 661-824 and 935-986 of SEQ ID NO: 1, or amino acids 648-812 and 922-978 of SEQ ID NO: 2, with one or more amino acid modifications relative to said RuvC cleavage domain sequence, wherein the dCasX variant exhibits one or more improved characteristics compared to the reference dCasX. In further embodiments, a catalytically-dead CasX variant protein comprises deletions of all or part of the RuvC domain of the reference CasX protein. It will be understood that the same foregoing substitutions or deletions can similarly be introduced into CasX variants known in the art, resulting in a dCasX variant (*see, e.g.*, WO2022120095A1 and US11,560,555, incorporated by reference herein, for exemplary sequences).

[0159] In some embodiments, the long-term repressor fusion protein comprising a dCasX variant with linked repressor domains exhibits at least one improved characteristic compared to a long-term repressor fusion protein comprising the reference dCasX protein with comparable linked repressor domains. All dCasX variants that improve one or more functions or characteristics of the long-term repressor fusion protein comprising the dCasX variant protein with linked repressor domains compared to the long-term repressor fusion protein comprising the reference dCasX protein are envisaged as being within the scope of the disclosure. In some embodiments, the modification is a mutation in one or more amino acids of the reference dCasX other than those rendering the dCasX catalytically-dead. For example, dCasX variants can comprise one or more amino acid substitutions, insertions, deletions, or swapped domains, or any combinations thereof, relative to a reference dCasX protein sequence. Any amino acid can be substituted for any other amino acid in the substitutions described herein. The substitution can be a conservative substitution (*e.g.*, a basic amino acid

is substituted for another basic amino acid). The substitution can be a non-conservative substitution (*e.g.*, a basic amino acid is substituted for an acidic amino acid or vice versa). For example, a proline in a reference dCasX protein can be substituted for any of arginine, histidine, lysine, aspartic acid, glutamic acid, serine, threonine, asparagine, glutamine, cysteine, glycine, alanine, isoleucine, leucine, methionine, phenylalanine, tryptophan, tyrosine or valine to generate a dCasX variant protein of the disclosure. In some embodiments, the dCasX variant exhibits an improved characteristic compared to a reference dCasX. Exemplary improved characteristics of the dCasX variant embodiments include, but are not limited to improved folding of the variant, increased binding affinity to the target nucleic acid, improved ability to utilize a greater spectrum of PAM sequences in the transcriptional repression and/or binding of target nucleic acid, improved unwinding of the target DNA, increased target strand loading, increased binding of the non-target strand of DNA, improved protein stability, increased ability to complex with gRNA, increased binding affinity to the gRNA, improved protein:gRNA (RNP) complex stability, and, with linked repressor domains and when complexed as an RNP, increased repressor activity, improved repressor specificity for the target nucleic acid, decreased off-target repression, increased percentage of a eukaryotic genome that can be efficiently repressed and/or epigenetically modified. In some embodiments, an improved characteristic of the dCasX variant is at least about 1.1 to about 100,000-fold improved relative to the reference dCasX protein. In some embodiments, an improved characteristic of the dCasX variant is at least about 1.1 to about 10,000-fold improved, at least about 1.1 to about 1,000-fold improved, at least about 1.1 to about 500-fold improved, at least about 1.1 to about 400-fold improved, at least about 1.1 to about 300-fold improved, at least about 1.1 to about 200-fold improved, at least about 1.1 to about 100-fold improved, at least about 1.1 to about 50-fold improved, at least about 1.1 to about 40-fold improved, at least about 1.1 to about 30-fold improved, at least about 1.1 to about 20-fold improved, at least about 1.1 to about 10-fold improved, at least about 1.1 to about 9-fold improved, at least about 1.1 to about 8-fold improved, at least about 1.1 to about 7-fold improved, at least about 1.1 to about 6-fold improved, at least about 1.1 to about 5-fold improved, at least about 1.1 to about 4-fold improved, at least about 1.1 to about 3-fold improved, at least about 1.1 to about 2-fold improved, at least about 1.1 to about 1.5-fold improved, at least about 1.5 to about 3-fold improved, at least about 1.5 to about 4-fold improved, at least about 1.5 to about 5-fold improved, at least about 1.5 to about 10-fold improved, at least about 5 to about 10-fold improved, at least about 10 to about 20-fold

improved, at least 10 to about 30-fold improved, at least 10 to about 50-fold improved or at least 10 to about 100-fold improved relative to the reference dCasX protein. In some embodiments, an improved characteristic of the dCasX variant is at least about 10 to about 1000-fold improved relative to the reference dCasX protein. Additional disclosure on improved characteristics is described herein, below.

[0160] In other embodiments, the modification is a substitution of one or more domains of the reference dCasX with one or more domains from a different CasX. In some embodiments, insertion includes the insertion of a part or all of a domain from a different CasX protein. Mutations can be placed in any one or more domains of the dCasX variant, and may include, for example, deletion of part or all of one or more domains, or one or more amino acid substitutions, deletions, or insertions in any domain. The domains of dCasX proteins include the non-target strand binding (NTSB) domain, the target strand loading (TSL) domain, the helical I domain, the helical II domain, the oligonucleotide binding domain (OBD), and the RuvC DNA cleavage domain, which can further comprise subdomains, described below.

[0161] In some embodiments, the dCasX variant protein comprises between 800 and 1100 amino acids or between 900 and 1000 amino acids.

[0162] The long-term repressor fusion protein comprising dCasX variant and linked repressor domains of the disclosure have an enhanced ability to efficiently bind target nucleic acid, when complexed with a gRNA as an RNP, utilizing and binding to a PAM TC motif, including PAM sequences selected from TTC, ATC, GTC, or CTC, compared to an RNP of a fusion protein comprising the reference dCasX protein and a gRNA in a comparable assay system. In the foregoing, the PAM sequence is located at least 1 nucleotide 5' to the non-target strand of the protospacer having identity with the targeting sequence of the gRNA.

[0163] In some embodiments, an RNP comprising the long-term repressor fusion protein comprising the dCasX variant with linked repressor domains and a gRNA of the disclosure, at a concentration of 20 pM or less, is capable of binding a double stranded DNA target with an efficiency of at least 70%, at least 80%, at least 85%, at least 90% or at least 95%. In one embodiment, an RNP of a long-term repressor fusion protein comprising a dCasX variant with linked repressor domains and a gRNA variant exhibits greater binding of a target sequence in the target nucleic acid compared to an RNP comprising a comparable long-term repressor fusion protein comprising a reference dCasX protein with linked repressor domains and a gRNA in a comparable assay system, wherein the PAM sequence of the target nucleic acid is TTC. In another embodiment, an RNP of a long-term repressor fusion protein

comprising dCasX variant with linked repressor domains and gRNA variant exhibits greater binding affinity of a target sequence in the target nucleic acid compared to an RNP comprising a comparable long-term repressor fusion protein comprising a reference dCasX protein with linked repressor domains and a gRNA in a comparable assay system, wherein the PAM sequence of the target nucleic acid is ATC. In another embodiment, an RNP of a long-term repressor fusion protein comprising a dCasX variant with linked repressor domains and gRNA variant exhibits greater binding affinity of a target sequence in the target nucleic acid compared to an RNP comprising a comparable long-term repressor fusion protein comprising a reference dCasX protein with linked repressor domains and a gRNA in a comparable assay system, wherein the PAM sequence of the target nucleic acid is CTC. In another embodiment, an RNP of a long-term repressor fusion protein comprising a dCasX variant with linked repressor domains and gRNA variant exhibits greater binding affinity of a target sequence in the target nucleic acid compared to an RNP a comparable long-term repressor fusion protein comprising a reference dCasX protein with linked repressor domains and a gRNA in a comparable assay system, wherein the PAM sequence of the target nucleic acid is GTC. In other embodiments, an RNP of a repressor fusion protein comprising the dCasX variant with linked repressor domains and a gRNA exhibits greater binding affinity for a target sequence in the target nucleic acid compared to an RNP comprising a comparable repressor fusion protein comprising the reference dCasX protein with linked repressor domains and the gRNA in a comparable assay system, wherein the PAM sequence of the target nucleic acid is GTC, TTC, ATC or CTC. In the foregoing embodiments, the increased binding affinity for the one or more PAM sequences is at least 1.5-fold greater or more compared to the binding affinity of an RNP of any one of the reference dCasX proteins (modified from SEQ ID NOS: 1-3) with linked repressor domains and the gRNA of SEQ ID NOS: 1731-1743 of Table 8 for the PAM sequences.

c. dCasX Variant Proteins with Domains from Multiple Source Proteins

[0164] In certain embodiments, the disclosure provides a chimeric dCasX variant protein for use in the long-term repressor fusion proteins.

[0165] As used herein, a "chimeric dCasX" protein refers to both a catalytically-dead CasX protein containing at least two domains from different sources, as well as catalytically-dead CasX protein containing at least one domain that itself is chimeric. Accordingly, in some embodiments, a chimeric dCasX protein is one that includes at least two domains isolated or derived from different sources, such as from two different naturally occurring CasX proteins,

(e.g., from two different CasX reference proteins). In other embodiments, the chimeric dCasX protein is one that contains at least one domain that is a chimeric domain, e.g. in some embodiments, part of a domain comprises a substitution from a different CasX protein (from a reference CasX protein, or another dCasX protein).

[0166] In some embodiments, the at least one chimeric domain can be any of the NTSB, TSL, helical I, helical II, OBD or RuvC domains as described herein. In the case of split or non-contiguous domains such as helical I, RuvC and OBD, a portion of the non-contiguous domain can be replaced with the corresponding portion from any other source. In some embodiments, the helical I-II domain (sometimes referred to as helical I-a) of the dCasX derived from SEQ ID NO: 2 is replaced with the corresponding helical I-II sequence from SEQ ID NO: 1, resulting in a chimeric dCasX protein.

[0167] In some embodiments, the helical I-II domain and NTSB domain of the dCasX derived from SEQ ID NO: 2 is replaced with the corresponding helical I-II and NTSB sequences from SEQ ID NO: 1, resulting in a chimeric dCasX protein.

[0168] A chimeric dCasX variant protein may comprise an NTSB, TSL, helical I-I, helical I-II, helical II, OBD-I, and OBD-II domains from a CasX protein of SEQ ID NO: 2, and a RuvC-I and/or RuvC-II domain from a CasX protein of SEQ ID NO: 1, or vice versa, in which mutations or other sequence alterations are introduced to create the catalytically-dead variant with improved properties of the variant, relative to the reference dCasX protein. As an example of the foregoing, the chimeric RuvC domain comprises amino acids 660 to 823 of SEQ ID NO: 1 and amino acids 921 to 978 of SEQ ID NO: 2. As an alternative example of the foregoing, a chimeric RuvC domain comprises amino acids 647 to 810 of SEQ ID NO: 2 and amino acids 934 to 986 of SEQ ID NO: 1. In a particular embodiment, a dCasX for use in the long-term repressor fusion protein comprises an NTSB domain and helical I-II domain from SEQ ID NO: 1 and a helical I-I domain from SEQ ID NO: 2; the latter being a chimeric domain, it being understood that the dCasX variants have additional amino acid changes at select locations (relative to the reference sequence), and the resulting chimeric dCasX proteins have improved characteristics relative to the reference dCasX proteins. Sequences of Table 2 having the NTSB domain and helical I-II domain from SEQ ID NO: 1 and a helical I-I domain from SEQ ID NO: 2 include dCasX 491 (SEQ ID NO: 4), 515 (SEQ ID NO: 6), 516 (SEQ ID NO: 7), 518-520 (SEQ ID NOS: 9-11), 522-527 (SEQ ID NOS: 12-17), 532 (SEQ ID NO: 22), 593 (SEQ ID NO: 25), 676 (SEQ ID NO: 28), with a L169K substitution in the NTSB domain), and 812 (SEQ ID NO: 29). Coordinates of CasX domains in the reference

CasX proteins of SEQ ID NO: 1 and SEQ ID NO: 2 are provided in Table 1 below. The skilled artisan will understand that the domain boundaries indicated in Table 1 below are approximate, and that protein fragments whose boundaries differ from those given in the table below by 1, 2, or 3 amino acids may have the same activity as the domains described below.

Table 1: Domain coordinates in Reference CasX proteins

Domain Name	Coordinates in SEQ ID NO: 1*	Coordinates in SEQ ID NO: 2*
OBD-I	1-55	1-57
helical I-I	56-99	58-101
NTSB	100-190	102-191
helical I-II	191-331	192-332
helical II	332-508	333-500
OBD-II	509-659	501-646
RuvC-I	660-823	647-810
TSL	824-933	811-920
RuvC-II	934-986	921-978

* amino acid position

[0169] In some embodiments, a dCasX variant protein utilized in the long-term repressor fusion proteins of the disclosure comprises a sequence selected from the group consisting of SEQ ID NOS: 4-29 as set forth in Table 2, wherein the sequence comprises a RuvC domain comprising one or more mutations that inactivates the cleavage activity of the RuvC domain, In other embodiments, a dCasX variant protein utilized in the long-term repressor fusion proteins of the disclosure comprises a sequence at least about 70% identical, at least about 75% identical, at least about 80% identical, at least about 81% identical, at least about 82% identical, at least about 83% identical, at least about 84% identical, at least about 85% identical, at least about 86% identical, at least about 86% identical, at least about 87% identical, at least about 88% identical, at least about 89% identical, at least about 89% identical, at least about 90% identical, at least about 91% identical, at least about 92% identical, at least about 93% identical, at least about 94% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, or at least about 99.5% identical to a sequence of SEQ ID NOS: 4-29 as set forth in Table 2. In some embodiments, the dCasX variant protein comprises a RuvC domain comprising one or more mutations that inactivates the cleavage activity of the RuvC domain In some embodiments, the long-term repressor fusion protein comprising the dCasX retains the ability to form an RNP with a gRNA. In a particular

embodiment, the dCasX variant protein utilized in the long-term repressor fusion protein of the gene repressor systems of the disclosure comprises a sequence of SEQ ID NO: 4 (dCasX 491). In another particular embodiment, the dCasX variant protein utilized in the long-term repressor fusion protein of the gene repressor systems of the disclosure comprises a sequence of SEQ ID NO: 6 (dCasX 515). In another particular embodiment, the dCasX variant protein utilized in the long-term repressor fusion protein of the gene repressor systems of the disclosure comprises a sequence of SEQ ID NO: 29 (dCasX 812).

Table 2: dCasX Variant Sequences

SEQ ID NO	dCasX	Amino Acid Sequence
4	dCasX491	QEIKRINKIRRLVKDSNTKKAGKTGPMKTLVLRVMTPLRERLENLRKKPENIPQ PISNTSRANLNKLLTDYTEMKKAILHVYWEWFQKDPVGLMSRVAQPASKKIDQNKL KPEMDEKGNLTAGFACSCGQPLFVYKLEQVSEK GKAYTNYFGRCNVAEHEKLIL LAQLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGP VGKALSDACMGTIASFSLSKYQDIIIEHQKVVGKNQKRLESRLRELAKENLEYPSVT LPPQPHTKEGVDAAYNEVIARVRMWNLNWQKLKLSRDDAKPLLRLKGFPSFPLVE RQANEVDWDMVCNVKKLINEKKEDGKVFQNLAGYKRQEALRPYLSSEEDRKKGK KFARYQLGDLHLEKKHGEDWGKVDYDEAWERIDKKVEGLSKHIKLEEEERSEDAQ SKAALTDWLRKASFVIEGLKEADKDEFRCCELKLQKWYGDRLGKPFATIAENSIL DISGFSKQYNCAFIWQKDGKVLNLYLIINYFKGGKLRFKKIKPEAFANRFTYVI NKKSGEIVPMEVNFNFDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGRV IEKTLYNRRTRQDEPALFVALTFERREVLDSNFKPMNLIGVARGENIPAVIALTD PEGCPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKAK NLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGRKFMAERQYTRMEDWLT AKLAYEGLSKTYLSKTLAQYTSKTCSCNGFTITSADYDRVLEKLKKTATGWMTTIN GKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALSLL KKRFSHRPVQEKFVCLNCGFETHAAEQAALNIARSWLFLRSQYKYQTNKTTGNT DKRAFVETWQSFYRKKLKEVWKPAPV
5	dCasX514	QEIKRINKIRRLVKDSNTKKAGKTGPMKTLVLRVMTPLRERLENLRKKPENIPQ PISNTSRANLNKLLTDYTEMKKAILHVYWEWFQKDPVGLMSRVAQPASKKIDQNKL KPEMDEKGNLTAGFACSCGQPLFVYKLEQVSEK GKAYTNYFGRCNVAEHEKLIL LAQLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGP VGKALSDACMGTIASFSLSKYQDIIIEHQKVVGKNQKRLESRLRELAKENLEYPSVT LPPQPHTKEGVDAAYNEVIARVRMWNLNWQKLKLSRDDAKPLLRLKGFPSFPLVE RQANEVDWDMVCNVKKLINEKKEDGKVFQNLAGYKRQEALRPYLSSEEDRKKGK KFARYQLGDLHLEKKHGEDWGKVDYDEAWERIDKKVEGLSKHIKLEEEERSEDAQ SKAALTDWLRKASFVIEGLKEADKDEFRCCELKLQKWYGDRLGKPFATIAENSIL DISGFSKQYNCAFIWQKDGKVLNLYLIINYFKGGKLRFKKIKPEAFANRFTYVI NKKSGEIVPMEVNFNFDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGRV IEKTLYNRRTRQDEPALFVALTFERREVLDSNFKPMNLIGVARGENIPAVIALTD PEGCPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKAK NLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGRKFMAERQYTRMEDWLT AKLAYEGLSKTYLSKTLAQYTSKTCSCNGFTIHTSADYDRVLEKLKKTATGWMTTI NGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALSLL LKKRFSHRPVQEKFVCLNCGFETHAAEQAALNIARSWLFLRSQYKYQTNKTTGNT TDKRAFVETWQSFYRKKLKEVWKPAPV
6	dCasX515	QEIKRINKIRRLVKDSNTKKAGKTGPMKTLVLRVMTPLRERLENLRKKPENIPQ PISNTSRANLNKLLTDYTEMKKAILHVYWEWFQKDPVGLMSRVAQPASKKIDQNKL KPEMDEKGNLTAGFACSCGQPLFVYKLEQVSEK GKAYTNYFGRCNVAEHEKLIL LAQLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGP

SEQ ID NO	dCasX	Amino Acid Sequence
		VGKALSDACMGTIASFLSKYQDIIIEHQKVVGKNQKRLESRELAGKENLEYPSVT LPPQPHTKEGVDAVNEVIARVRMWNLNWQKLKLSRDDAKPLLRLKGFPSFPLVE RQANEVDWDMVCNVKKLINEKKEDGKVFQWQNLAGYKRQEALRPYLSSEEDRKKGK KFARYQLGDLHLEKKHGEDWGKVDYDEAWERIDKKVEGLSKHIKLEEEERSEDAQ SKAALTDWLRKASFVIEGLKEADKDEFCRCCELKLQKWYGDRLRGKPFATIEAENSIL DISGFSKQYNCAFIWQKDGKVLNLYLIINYFKGGKLRFKKIKPEAFANRFYTVI NKKSGEIVPMEVNFNFDPNLIILPLAFGKRQGREFIWNLDSLETGSLKLANGRV IEKTLNRRTRQDEPALFVALTFERREVLDSNNIKPMNLIGVARGENIPAVIALTD PEGCPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKAK NLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGRKFMAERQYTRMEDWLT AKLAYEGLPSKTYLSKTLAQYTSKTCNCGFTITSADYDRVLEKLKKTATGWMTTI NGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALS LKKRFSHRPVQEKFVCLNCGFETHAAEQAALNIARSWLFLRSQYKQYQTNKTTGN TDKRAFVETWQSFYRKKLKEVWKPAP
7	dCasX516	QEIKRINKIRRLVKDSNTKKAGKTGPMKTLVVRVMTDPLRERLENLRKKPENIPQ PISNTSRANLNKLLTDYTEMKKAHLHVVWEEFQKDPVGLMSRVAQPASKKIDQNK KPEMDEKGNLTAGFACSCGQPLFVYKLEQVSEKGAITYNFGRCNVAEHEKLIL LAQLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGP VGKALSDACMGTIASFLSKYQDIIIEHQKVVGKNQKRLESRELAGKENLEYPSVT LPPQPHTKEGVDAVNEVIARVRMWNHNLWQKLKLSRDDAKPLLRLKGFPSFPLVE RQANEVDWDMVCNVKKLINEKKEDGKVFQWQNLAGYKRQEALRPYLSSEEDRKKGK KFARYQLGDLHLEKKHGEDWGKVDYDEAWERIDKKVEGLSKHIKLEEEERSEDAQ SKAALTDWLRKASFVIEGLKEADKDEFCRCCELKLQKWYGDRLRGKPFATIEAENSIL DISGFSKQYNCAFIWQKDGKVLNLYLIINYFKGGKLRFKKIKPEAFANRFYTVI NKKSGEIVPMEVNFNFDPNLIILPLAFGKRQGREFIWNLDSLETGSLKLANGRV IEKTLNRRTRQDEPALFVALTFERREVLDSNNIKPMNLIGVARGENIPAVIALTD PEGCPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKAK NLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGRKFMAERQYTRMEDWLT AKLAYEGLSKTYLSKTLAQYTSKTCNCGFTITSADYDRVLEKLKKTATGWMTTIN GKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALS LKKRFSHRPVQEKFVCLNCGFETHAAEQAALNIARSWLFLRSQYKQYQTNKTTGN TDKRAFVETWQSFYRKKLKEVWKPAP
8	dCasX517	QEIKRINKIRRLVKDSNTKKAGKTGPMKTLVVRVMTDPLRERLENLRKKPENIPQ PISNTSRANLNKLLTDYTEMKKAHLHVVWEEFQKDPVGLMSRVAQPASKKIDQNK KPEMDEKGNLTAGFACSCGQPLFVYKLEQVSEKGAITYNFGRCNVAEHEKLIL LAQLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGA PVGKALSDACMGTIASFLSKYQDIIIEHQKVVGKNQKRLESRELAGKENLEYPSV TLPPQPHTKEGVDAVNEVIARVRMWNLNWQKLKLSRDDAKPLLRLKGFPSFPLV ERQANEVDWDMVCNVKKLINEKKEDGKVFQWQNLAGYKRQEALRPYLSSEEDRKKG KKFARYQLGDLHLEKKHGEDWGKVDYDEAWERIDKKVEGLSKHIKLEEEERSEDA QSKAALTDWLRKASFVIEGLKEADKDEFCRCCELKLQKWYGDRLRGKPFATIEAENSIL DISGFSKQYNCAFIWQKDGKVLNLYLIINYFKGGKLRFKKIKPEAFANRFYTVI INKKSGEIVPMEVNFNFDPNLIILPLAFGKRQGREFIWNLDSLETGSLKLANGRV VIEKTLNRRTRQDEPALFVALTFERREVLDSNNIKPMNLIGVARGENIPAVIALT DPEGCPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKA KNLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGRKFMAERQYTRMEDWL TAKLAYEGLSKTYLSKTLAQYTSKTCNCGFTITSADYDRVLEKLKKTATGWMTTI NGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALS LKKRFSHRPVQEKFVCLNCGFETHAAEQAALNIARSWLFLRSQYKQYQTNKTTGN TDKRAFVETWQSFYRKKLKEVWKPAP
9	dCasX518	RQEIKRINKIRRLVKDSNTKKAGKTGPMKTLVVRVMTDPLRERLENLRKKPENIP QPISNTSRANLNKLLTDYTEMKKAHLHVVWEEFQKDPVGLMSRVAQPASKKIDQNK LKPEMDEKGNLTAGFACSCGQPLFVYKLEQVSEKGAITYNFGRCNVAEHEKLIL LLAQLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASG PVGKALSDACMGTIASFLSKYQDIIIEHQKVVGKNQKRLESRELAGKENLEYPSV TLPPQPHTKEGVDAVNEVIARVRMWNLNWQKLKLSRDDAKPLLRLKGFPSFPLV

SEQ ID NO	dCasX	Amino Acid Sequence
		ERQANEVDWDMVCNVKKLINEKKEDGKVFQNLAGYKRQEALRPYLSSEEDRKKG KKFARYQLGDLHLHLEKKHGEDWGKVDYDEAWERIDKKVEGLSKHIKLEEEERSEDA QSKAALTDWLRKASFVIEGLKEADKDEFRCCELKLQKWYGDRLGKPFAlEAENSI LDISGFSKQYNCAFIWQKDGVKKLNLYLIINYFKGGKLRFKKIKPEAFEANRFYTV INKKSGEIVPMEVNFNFDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGR VIEKTLYNRRTRQDEPALFVALTFERREVLDSNFKPMNLIGVARGENIPAVIALT DPEGCPKSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKA KNLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGRKFMAERQYTRMEDWL TAKLAYEGLSKTYLSKTLAQYTSKTCNCGFTITSADYDRVLEKLLKKTATGWMTTI NGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALSL LKKRFSHRPVQEKFVCLNCGFETHAAEQAALNIARSWLFLRSQYKYQTNKTTGN TDKRAFVETWQSFYRKLLKEVWKPAV
10	dCasX519	QEIKRINKIRRLVKDSNTKKAGKTGPMKTLVVRVMTPLRERLENLRKKPENIPQ PISNTSRANLNKLLTDYTEMKKAILHVVWEEFQKDPVGLMSRVAQPASKKIDQNK KPEMDEKGNLTAGFACSCGQPLFVYKLEQVSEKQKAYTNYFGRCNVAEHEKLLI LAQLKPEKDSDEAVTYSLGKFGQALDFYSIHVTKESTHVPVKPLAQIAGNRYASGP VGKALSDACMGTIASFLSKYQDIIIEHQVVKGNQKRLESRLRELAKENLEYPST LPPQPHTKEGVDAYNEVIARVRMWNLNWQKLKLSRDDAKPLLRKLGFPSPFLVE RQANEVDWDMVCNVKKLINEKKEDGKVFQNLAGYKRQEALRPYLSSEEDRKKGK KFARYQLGDLHLHLEKKHGEDWGKVDYDEAWERIDKKVEGLSKHIKLEEEERSEDA QSKAALTDWLRKASFVIEGLKEADKDEFRCCELKLQKWYGDRLGKPFAlEAENSI LDISGFSKQYNCAFIWQKDGVKKLNLYLIINYFKGGKLRFKKIKPEAFEANRFYTVI NKKSGEIVPMEVNFNFDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGRV IEKTLYNRRTRQDEPALFVALTFERREVLDSNFKPMNLIGVARGENIPAVIALTD PEGCPKSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKA KNLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGRKFMAERQYTRMEDWL TAKLAYEGLSKTYLSKTLAQYTSKTCNCGFTITSADYDRVLEKLLKKTATGWMTTI NGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALSL LKKRFSHRPVQEKFVCLNCGFETHAAEQAALNIARSWLFLRSQYKYQTNKTTGN TDKRAFVETWQSFYRKLLKEVWKPAV
11	dCasX520	QEIKRINKIRRLVKDSNTKKAGKTGPMKTLVVRVMTPLRERLENLRKKPENIPQ PISNTSRANLNKLLTDYTEMKKAILHVVWEEFQKDPVGLMSRVAQPASKKIDQNK KPEMDEKGNLTAGFACSCGQPLFVYKLEQVSEKQKAYTNYFGRCNVAEHEKLLI LAQLKPEKDSDEAVTYSLGKFGQALDFYSIHVTKESTHVPVKPLAQIAGNRYASGP VGKALSDACMGTIASFLSKYQDIIIEHQVVKGNQKRLESRLRELAKENLEYPST LPPQPHTKEGVDAYNEVIARVRMWNLNWQKLKLSRDDAKPLLRKLGFPSPFLVE RQANEVDWDMVCNVKKLINEKKEDGKVFQNLAGYKRQEALRPYLSSEEDRKKGK KFARYQLGDLHLHLEKKHGEDWGKVDYDEAWERIDKKVEGLSKHIKLEEEERSEDA QSKAALTDWLRKASFVIEGLKEADKDEFRCCELKLQKWYGDRLGKPFAlEAENSI LDISGFSKQYNCAFIWQKDGVKKLNLYLIINYFKGGKLRFKKIKPEAFEANRFYTVI NKKSGEIVPMEVNFNFDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGRV IEKTLYNRRTRQDEPALFVALTFERREVLDSNFKPMNLIGVARGENIPAVIALTD PEGCPKSRFKDSLGNPHTILRIGESYKEKQRTTQAKKEVEQRRAGGYSRKYASKAK NLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGRKFMAERQYTRMEDWLT AKLAYEGLSKTYLSKTLAQYTSKTCNCGFTITSADYDRVLEKLLKKTATGWMTTIN GKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALSL LKKRFSHRPVQEKFVCLNCGFETHAAEQAALNIARSWLFLRSQYKYQTNKTTGNT DKRAFVETWQSFYRKLLKEVWKPAV
12	dCasX522	QEIKRINKIRRLVKDSNTKKAGKTGPMKTLVVRVMTPLRERLENLRKKPENIPQ PISNTSRANLNKLLTDYTEMKKAILHVVWEEFQKDPVGLMSRVAQPASKKIDQNK KPEMDEKGNLTAGFACSCGQPLFVYKLEQVSEKQKAYTNYFGRCNVAEHEKLLI LAQLKPEKDSDEAVTYSLGKFGQALDFYSIHVTKESTHVPVKPLAQIAGNRYASGP VGKALSDACMGTIASFLSKYQDIIIEHQVVKGNQKRLESRLRELAKENLEYPST LPPQPHTKEGVDAYNEVIARVRMWNLNWQKLKLSRDDAKPLLRKLGFPSPFLVE RQANEVDWDMVCNVKKLINEKKEDGKVFQNLAGYKRQEALRPYLSSEEDRKKGK KFARYQLGDLHLHLEKKHGEDWGKVDYDEAWERIDKKVEGLSKHIKLEEEERSEDA QSKAALTDWLRKASFVIEGLKEADKDEFRCCELKLQKWYGDRLGKPFAlEAENSI LDISGFSKQYNCAFIWQKDGVKKLNLYLIINYFKGGKLRFKKIKPEAFEANRFYTVI NKKSGEIVPMEVNFNFDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGRV IEKTLYNRRTRQDEPALFVALTFERREVLDSNFKPMNLIGVARGENIPAVIALTD PEGCPKSRFKDSLGNPHTILRIGESYKEKQRTTQAKKEVEQRRAGGYSRKYASKAK NLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGRKFMAERQYTRMEDWLT AKLAYEGLSKTYLSKTLAQYTSKTCNCGFTITSADYDRVLEKLLKKTATGWMTTIN GKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALSL LKKRFSHRPVQEKFVCLNCGFETHAAEQAALNIARSWLFLRSQYKYQTNKTTGNT DKRAFVETWQSFYRKLLKEVWKPAV

SEQ ID NO	dCasX	Amino Acid Sequence
		SKAALTDWLRKASFVIEGLKEADKDEFRCCELKLQKWYGDLRGKPFAlEAENSIL DISGFSKQYNCAFIWQKDGVKKLNLYLIINYFKGGKLRFKKIKPEAFEANRFYTVI NKKSGEIVPMEVNFNFDDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGRV IEKTLYNRRTRQDEPALFVALTFERREVLDSNIIKPMNLIGVARGENIPAVIALTD PEGCPLSRFKRSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKAK NLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGRKFMAERQYTRMEDWLT AKLAYEGLSKTYLSKTLAQYTSKTCSCGFTITSADYDRVLEKLLKKTATGWMTTIN GKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTGRSGEALSLL KKRFSHRPVQEKFVCLNCGFETHAAEQAALNIARSWLFLRSQEYKKYQTNKTTGNT DKRAFVETWQSFYRKLLKEVWKPAP
13	dCasX523	QEIKRINKIRRLVKDSNTKKAGKTYPMKTLVVRVMTPLRERLENLRKKPENIPQ PISNTSRANLNKLLTDYTEMKKAILHVYWEFQKDPVGLMSRVAQPASKKIDQNKL KPEMDEKGNLTAGFACSCQGQPLFVYKLEQVSEKGAAYTNYFGRCNVAEHEKLIIL LAQLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGP VGKALSDACMGTIASFLSKYQDIIIEHQVVKGNQKRLESRLRELAKENLEYPSVT LPPQPHTKEGVDAYNEVIARVRMWVNLNLWQKLKLSRDDAKPLLRKLGFPSPFLVE RQANEVDWDMVCNVKKLINEKKEDGKVFQWNLAGYKRQEALRPYLSSEEDRKKGK KFARYQLGDLILLHLEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKLEEEERSEDAQ SKAALTDWLRKASFVIEGLKEADKDEFRCCELKLQKWYGDLRGKPFAlEAENSIL DISGFSKQYNCAFIWQKDGVKKLNLYLIINYFKGGKLRFKKIKPEAFEANRFYTVI NKKSGEIVPMEVNFNFDDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGRV IEKTLYNRRTRQDEPALFVALTFERREVLDSNIIKPMNLIGVARGENIPAVIALTD PEGCPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKAK NLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGRKFMAERQYTRMEDWLT AKLAYEGLSKTYLSKTLAQYTSKTCSCGFTITSADYDRVLEKLLKKTATGWMTTIN GKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTGRSGEALSLL KKRFSHRPVQEKFVCLNCGFETHAAEQAALNIARSWLFLRSQEYKKYQTNKTTGNT DKRAFVETWQSFYRKLLKEVWKPAP
14	dCasX524	QEIKRINKIRRLVKDSNTKKAGKTGPMKTLVVRVMTPLRERLENLRKKPENIPQ PISNTSRANLNKLLTDYTEMKKAILHVYWEFQKDPVGLMSRVAQPASKKIDQNKL KPEMDEKGNLTAGFACSCQGQPLFVYKLEQVSEKGAAYTNYFGRCNVAEHEKLIIL LAQLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGP VGKALSDACMGTIASFLSKYQDIIIEHQVVKGNQKRLESRLRELAKENLEYPSVT LPPQPHTKEGVDAYNEVIARVRMWVNLNLWQKLKLSRDDAKPLLRKLGFPSPFLVE RQANEVDWDMVCNVKKLINEKKEDGKVFQWNLAGYKRQEALRPYLSSEEDRKKGK KFARYQLGDLILLHLEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKLEEEERSEDAQ SKAALTDWLRKASFVIEGLKEADKDEFRCCELKLQKWYGDLRGKPFAlEAENSIL DISGFSKQYNCAFIWQKDGVKKLNLYLIINYFKGGKLRFKKIKPEAFEANRFYTVI NKKSGEIVPMEVNFNFDDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGRV IEKTLYNRRTRQDEPALFVALTFERREVLDSNIIKPMNLIGVARGENIPAVIALTD PEGCPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKAK NLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGRKFMAERQYTRMEDWLT AKLAYEGLSKTYLSKTLAQYTSKTCSCGFTIHSADYDRVLEKLLKKTATGWMTTIN GKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTGRSGEALSLL KKRFSHRPVQEKFVCLNCGFETHAAEQAALNIARSWLFLRSQEYKKYQTNKTTGNT DKRAFVETWQSFYRKLLKEVWKPAP
15	dCasX525	QEIKRINKIRRLVKDSNTKKAGKTGPMKTLVVRVMTPLRERLENLRKKPENIPQ PISNTSRANLNKLLTDYTEMKKAILHVYWEFQKDPVGLMSRVAQPASKKIDQNKL KPEMDEKGNLTAGFACSCQGQPLFVYKLEQVSEKGAAYTNYFGRCNVAEHEKLIIL LAQLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGP VGKALSDACMGTIASFLSKYQDIIIEHQVVKGNQKRLESRLRELAKENLEYPSVT LPPQPHTKEGVDAYNEVIARVRMWVNLNLWQKLKLSRDDAKPLLRKLGFPSPFLVE RQANEVDWDMVCNVKKLINEKKEDGKVFQWNLAGYKRQEALRPYLSSEEDRKKGK KFARYQLGDLILLHLEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKLEEEERSEDAQ SKAALTDWLRKASFVIEGLKEADKDEFRCCELKLQKWYGDLRGKPFAlEAENSIL DISGFSKQYNCAFIWQKDGVKKLNLYLIINYFKGGKLRFKKIKPEAFEANRFYTVI

SEQ ID NO	dCasX	Amino Acid Sequence
		NKKSGEIVPMEVNFNFDDPNLIILPLAFGKRQGREFIWNDDLSLETGSLKLANGRV IEKTLYNRRTRQDEPALFVALTFERREVLDSNNIKPMNLIGVARGENIPAVIALTD PEGCPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKAK NLADDMVRNTARDLLYYAATQDAMLI FANLSRGFGRQGRKFMAERQYTRMEDWLT AKLAYEGLSKTYLSKTLAQYTSKTCSCNGFTITSADYDRVLEKLKKTATGWMTTIN GKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALSLL KKRFSHRPVQEKFVCLNCGFETHAAEQAALNIARSWLFLRSQEYKKYQTNKTTGNT DKRAFVETWQSFYRKKLKEVWKPAPV
16	dCasX526	QEIKRINKIRRLVKDSNTKKAGKTGPMKTLVVRVMTPLDRERLENLRKKPENIPQ PISNTSRANLNKLLTDYTEMKKAILHVVWEEFQKDPVGLMSRVAQPASKKIDQNK KPEMDEKGNLTAGFACSCGQPLFVYKLEQVSEKGAITYNYFGRCNVAEHEKLI LAQLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGP VGKALSDACMGTIASFLSKYQDIIIEHQKVVGKNQKRLESRELAKENLEYPST LPPQPHTKEGVDAYNEVIARVRMWNLNWQKLKLSRDDAKPLRLKGFPSFPLVE RQANEVDWDMVCNVKLLINEKKEDGKVFQWNLAGYKRQEALRPYLSSEEDRKKGK KFARYQLGDLHLLHLEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKLEEEERSEDAQ SKAALTDWLRAKASFVIEGLKEADKDFCRCELKQKWDLRGKPFATIEAENSIL DISGFSKQYNCAFIWQKDGVKLLNLYLIINYFKGGKLRFKKIKPEAFANRFTYVI NKKSGEIVPMEVNFNFDDPNLIILPLAFGKRQGREFIWNDDLSLETGSLKLANGRV IEKTLYNRRTRQDEPALFVALTFERREVLDSNNIKPMNLIGVARGENIPAVIALTD PEGCPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKAK NLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGRKFMAERQYTRMEDWLT AKLAYEGLSKTYLSKTLAQYTSKTCSCNGFTITSADYDRVLEKLKKTATGWMTTIN GKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALSLL KKRFSHRPVQEKFVCLNCGFETHAAEQAALNIARSWLFLRSQEYKKYQTNKTTGNT DKRAFVETWQSFYRKKLKEVWKPAPV
17	dCasX527	QEIKRINKIRRLVKDSNTKKAGKTGPMKTLVVRVMTPLDRERLENLRKKPENIP QPI SNTSRANLNKLLTDYTEMKKAILHVVWEEFQKDPVGLMSRVAQPASKKIDQNK LKPEMDEKGNLTAGFACSCGQPLFVYKLEQVSEKGAITYNYFGRCNVAEHEKLI LLAQLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASG PVGKALSDACMGTIASFLSKYQDIIIEHQKVVGKNQKRLESRELAKENLEYPST TLPPQPHTKEGVDAYNEVIARVRMWNLNWQKLKLSRDDAKPLRLKGFPSFPLV ERQANEVDWDMVCNVKLLINEKKEDGKVFQWNLAGYKRQEALRPYLSSEEDRKKG KKFARYQLGDLHLLHLEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKLEEEERSEDA QSKAALTDWLRAKASFVIEGLKEADKDFCRCELKQKWDLRGKPFATIEAENSIL DISGFSKQYNCAFIWQKDGVKLLNLYLIINYFKGGKLRFKKIKPEAFANRFTYVI INKKSGEIVPMEVNFNFDDPNLIILPLAFGKRQGREFIWNDDLSLETGSLKLANGR VIEKTLYNRRTRQDEPALFVALTFERREVLDSNNIKPMNLIGVARGENIPAVIALT DPEGCPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKA KNLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGRKFMAERQYTRMEDWL TAKLAYEGLSKTYLSKTLAQYTSKTCSCNGFTITSADYDRVLEKLKKTATGWMTTIN NGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALSLL LKKRFSHRPVQEKFVCLNCGFETHAAEQAALNIARSWLFLRSQEYKKYQTNKTTGNT TDKRAFVETWQSFYRKKLKEVWKPAPV
18	dCasX528	QEIKRINKIRRLVKDSNTKKAGKTGPMKTLVVRVMTPLDRERLENLRKKPENIPQ PISNTSRANLNKLLTDYTEMKKAILHVVWEEFQKDPVGLMSRVAQPASKKIDQNK KPEMDEKGNLTAGFACSCGQPLFVYKLEQVSEKGAITYNYFGRCNVAEHEKLI LAQLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASYP VGKALSDACMGTIASFLSKYQDIIIEHQKVVGKNQKRLESRELAKENLEYPST LPPQPHTKEGVDAYNEVIARVRMWNLNWQKLKLSRDDAKPLRLKGFPSFPLVE RQANEVDWDMVCNVKLLINEKKEDGKVFQWNLAGYKRQEALRPYLSSEEDRKKGK KFARYQLGDLHLLHLEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKLEEEERSEDAQ SKAALTDWLRAKASFVIEGLKEADKDFCRCELKQKWDLRGKPFATIEAENSIL DISGFSKQYNCAFIWQKDGVKLLNLYLIINYFKGGKLRFKKIKPEAFANRFTYVI NKKSGEIVPMEVNFNFDDPNLIILPLAFGKRQGREFIWNDDLSLETGSLKLANGRV IEKTLYNRRTRQDEPALFVALTFERREVLDSNNIKPMNLIGVARGENIPAVIALTD

SEQ ID NO	dCasX	Amino Acid Sequence
		PEGCPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKAK NLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQKRTFMAERQYTRMEDWLT AKLAYEGLPSKTYLSKTLAQYTSKTC SNCGFTITSADYDRVLEKLKKTATGWMTTI NGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALSL LKKRFSHRPVQEKFVCLNCGFETHAAEQ AALNIARSWLFLRSQEYKKYQTNKTTGN TDKRAFVETWQSFYRKKLKEVWKPAV
19	dCasX529	QEIKRINKIRRLVKDSNTKKAGKTGPMKTL LVRVMT PDLRERLENLRKKPENIPQ PISNTSRANLNKLLTDYTEMKKAILHVYWE EFQKDPVGLMSRVAQPASKKIDQNKL KPEMDEKGNLT TAGFACSCGQPLFVYKLEQVSEK GKAYTNYFGRCNVAEHEKLIL LAQLKPEKDSDEAVTYS LGKFGQ RALDFYSIHVTKESTHPVKPLAQIAGNRYASNP VGKALSDACMGTIASF LSKYQDIIIEHQKV VKNQKRLES LRELAKENLEYPST LPPQPHTKEGVDAYNEVIARVRMWNLN LWQKLKLSRDDAKPLLR LKGFPSFPLVE RQANEVDWDMVCNVKKLINEKKEDGKVFWQNL AGYKRQEALRPYLSSEEDRKKGK KFARYQLGDL LHLLEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKLEEEERSEDAQ SKAALTDWLRAKASFVIEGLKEADKDEF CRCELKLQKWYGD LRGKPF AIEAENSIL DISGFSKQYNCAFIWQKDGVKKLNLYLI INYFKGGKLRFKKIKPEAF EANRFYTVI NKKSGEIVPMEVNFNFDDPNLII LPLAFGKRQGREFIWNDLLSLETGSLKLANGRV IEKTLYNRRTRQDEPALFVALTFERREVLDS SNIKPMNLIGVARGENI PAVIALTD PEGCPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKAK NLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQKRTFMAERQYTRMEDWLT AKLAYEGLPSKTYLSKTLAQYTSKTC SNCGFTITSADYDRVLEKLKKTATGWMTTI NGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALSL LKKRFSHRPVQEKFVCLNCGFETHAAEQ AALNIARSWLFLRSQEYKKYQTNKTTGN TDKRAFVETWQSFYRKKLKEVWKPAV
20	dCasX530	QEIKRINKIRRLVKDSNTKKAGKTGPMKTL LVRVMT PDLRERLENLRKKPENIPQ PISNTSRANLNKLLTDYTEMKKAILHVYWE EFQKDPVGLMSRVAQPASKKIDQNKL KPEMDEKGNLT TAGFACSCGQPLFVYKLEQVSEK GKAYTNYFGRCNVAEHEKLIL LAQLKPEKDSDEAVTYS LGKFGQ RALDFYSIHVTKESTHPVKPLAQIAGNRYASGP VGKALSDACMGTIASF LSKYQDIIIEHQKV VKNQKRLES LRELAKENLEYPST LPPQPHTKEGVDAYNEVIARVRMWNLN LWQKLKLSRDDAKPLLR LKGFPSFPLVE RQANEVDWDMVCNVKKLINEKKEDGKVFWQNL AGYKRQEALRPYLSSEEDRKKGK KFARYQLGDL LHLLEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKLEEEERSEDAQ SKAALTDWLRAKASFVIEGLKEADKDEF CRCELKLQKWYGD LRGKPF AIEAENSIL DISGFSKQYNCAFIWQKDGVKKLNLYLI INYFKGWGKLRFKKIKPEAF EANRFYTV INKKSGEIVPMEVNFNFDDPNLII LPLAFGKRQGREFIWNDLLSLETGSLKLANGR VIEKTLYNRRTRQDEPALFVALTFERREVLDS SNIKPMNLIGVARGENI PAVIALT DPEGCPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKA KNLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQKRTFMAERQYTRMEDWL TAKLAYEGLPSKTYLSKTLAQYTSKTC SNCGFTITSADYDRVLEKLKKTATGWMTT INGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALS LKKRFSHRPVQEKFVCLNCGFETHAAEQ AALNIARSWLFLRSQEYKKYQTNKTTG NTDKRAFVETWQSFYRKKLKEVWKPAV
21	dCasX531	QEIKRINKIRRLVKDSNTKKAGKTGPMKTL LVRVMT PDLRERLENLRKKPENIPQ PISNTSRANLNKLLTDYTEMKKAILHVYWE EFQKDPVGLMSRVAQPASKKIDQNKL KPEMDEKGNLT TAGFACSCGQPLFVYKLEQVSEK GKAYTNYFGRCNVAEHEKLIL LAQLKPEKDSDEAVTYS LGKFGQ RALDFYSIHVTKESTHPVKPLAQIAGNRYASGP VGKALSDACMGTIASF LSKYQDIIIEHQKV VKNQKRLES LRELAKENLEYPST LPPQPHTKEGVDAYNEVIARVRMWNLN LWQKLKLSRDDAKPLLR LKGFPSFPLVE RQANEVDWDMVCNVKKLINEKKEDGKVFWQNL AGYKRQEALRPYLSSEEDRKKGK KFARYQLGDL LHLLEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKLEEEERSEDAQ SKAALTDWLRAKASFVIEGLKEADKDEF CRCELKLQKWYGD LRGKPF AIEAENSIL DISGFSKQYNCAFIWQKDGVKKLNLYLI INYFKGYGKLRFKKIKPEAF EANRFYTV INKKSGEIVPMEVNFNFDDPNLII LPLAFGKRQGREFIWNDLLSLETGSLKLANGR VIEKTLYNRRTRQDEPALFVALTFERREVLDS SNIKPMNLIGVARGENI PAVIALT DPEGCPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKA KNLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQKRTFMAERQYTRMEDWL

SEQ ID NO	dCasX	Amino Acid Sequence
		TAKLAYEGLPSKTYLSKTLAQYTSKTC SNCGFTITSADYDRVLEKLKKTATGWMTT INGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALS LLKKRFSHRPVQEKFVCLNCGFETHAAEQ AALNIARSWLFLRSQ EYKKYQTNKTTG NTDKRA FVETWQS FYRKKLKEVWKPAV
22	dCasX532	QEI KRINKIRRLVKDSNTKKAGKTRGPMKTL LVRVMT PDLRERLENLRKKPENIP QPI SNTSRANLNKLLTDYTEMKA I LHVYWEEFQKDPVGLMSRVAQPASKKIDQNK LKPEMDEKGNLT TAGFACSQCGQPLFVYKLEQVSEK GKAYTNYFGRCNVAEHEKLI LLAQLKPEKDSDEAVTYS LGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASG PVGKALSDACMGTIASFLSKYQDIIIEHQKVVGKNQKRLES LRELAKENLEYPSV TLPPQPHTKEGVDA YNEVIARVRMWNLN LWQKLKLSRDDAKP LLRLKGFPSPFLV ERQANEVDWWD MVCNVKKLINEKKEDGKVFQNL AGYKRQEALRPYLSSEEDRKKG KKFARYQLGDLLLHLEKKHGEDWGKVYDEAWERIDKKVEGLSKH I KLEEERRSEDA QSKAALTDWLRAKASFVIEGLKEADKDEFRCCELKLQKWYGD LRGKPF AIEAENSI LDISGFSKQYNCAFIWQKDG VKKLNLYLI INYFKGGKLRFKKIKPEAFEANRFYTV INKKS GEIVPMEVNFNFDDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGR VIEKTLYNRRTRQDEPALFVALT FERREVLDS SNIKPMNLIGVARGENIPAVIALT DPEGCP LSRFKDSLGNP THILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKA KNLADD MVRNTARDLLYYAVTQDAMLI FANLSRGFGRQ GKRTFMAERQYTRMEDWL TAKLAYEGLPSKTYLSKTLAQYTSKTC SNCGFTITSADYDRVLEKLKKTATGWMTT INGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALS LLKKRFSHRPVQEKFVCLNCGFETHAAEQ AALNIARSWLFLRSQ EYKKYQTNKTTG NTDKRA FVETWQS FYRKKLKEVWKPAV
23	dCasX533	QEI KRINKIRRLVKDSNTKKAGKTRGPMKTL LVRVMT PDLRERLENLRKKPENIP QPI SNTSRANLNKLLTDYTEMKA I LHVYWEEFQKDPVGLMSRVAQPASKKIDQNK LKPEMDEKGNLT TAGFACSQCGQPLFVYKLEQVSEK GKAYTNYFGRCNVAEHEKLI LLAQLKPEKDSDEAVTYS LGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASY PVGKALSDACMGTIASFLSKYQDIIIEHQKVVGKNQKRLES LRELAKENLEYPSV TLPPQPHTKEGVDA YNEVIARVRMWNLN LWQKLKLSRDDAKP LLRLKGFPSPFLV ERQANEVDWWD MVCNVKKLINEKKEDGKVFQNL AGYKRQEALRPYLSSEEDRKKG KKFARYQLGDLLLHLEKKHGEDWGKVYDEAWERIDKKVEGLSKH I KLEEERRSEDA QSKAALTDWLRAKASFVIEGLKEADKDEFRCCELKLQKWYGD LRGKPF AIEAENSI LDISGFSKQYNCAFIWQKDG VKKLNLYLI INYFKGGKLRFKKIKPEAFEANRFYTV INKKS GEIVPMEVNFNFDDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGR VIEKTLYNRRTRQDEPALFVALT FERREVLDS SNIKPMNLIGVARGENIPAVIALT DPEGCP LSRFKDSLGNP THILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKA KNLADD MVRNTARDLLYYAVTQDAMLI FANLSRGFGRQ GKRTFMAERQYTRMEDWL TAKLAYEGLPSKTYLSKTLAQYTSKTC SNCGFTITSADYDRVLEKLKKTATGWMTT INGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALS LLKKRFSHRPVQEKFVCLNCGFETHAAEQ AALNIARSWLFLRSQ EYKKYQTNKTTG NTDKRA FVETWQS FYRKKLKEVWKPAV
24	dCasX535	QEI KRINKIRRLVKDSNTKKAGKTRGPMKTL LVRVMT PDLRERLENLRKKPENIP PISNTSRANLNKLLTDYTEMKA I LHVYWEEFQKDPVGLMSRVAQPASKKIDQNK KPEMDEKGNLT TAGFACSQCGQPLFVYKLEQVSEK GKAYTNYFGRCNVAEHEKLI LLAQLKPEKDSDEAVTYS LGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASSP VGKALSDACMGTIASFLSKYQDIIIEHQKVVGKNQKRLES LRELAKENLEYPSVT LPPQPHTKEGVDA YNEVIARVRMWNLN LWQKLKLSRDDAKP LLRLKGFPSPFLVE RQANEVDWWD MVCNVKKLINEKKEDGKVFQNL AGYKRQEALRPYLSSEEDRKKG KFARYQLGDLLLHLEKKHGEDWGKVYDEAWERIDKKVEGLSKH I KLEEERRSEDAQ SKAALTDWLRAKASFVIEGLKEADKDEFRCCELKLQKWYGD LRGKPF AIEAENSI LDISGFSKQYNCAFIWQKDG VKKLNLYLI INYFKGGKLRFKKIKPEAFEANRFYTVI NKKSGEIVPMEVNFNFDDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGRV IEKTLYNRRTRQDEPALFVALT FERREVLDS SNIKPMNLIGVARGENIPAVIALTD PEGCP LSRFKDSLGNP THILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKA NLADD MVRNTARDLLYYAVTQDAMLI FANLSRGFGRQ GKRTFMAERQYTRMEDWLT AKLAYEGLPSKTYLSKTLAQYTSKTC SNCGFTITSADYDRVLEKLKKTATGWMTTI NGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALS

SEQ ID NO	dCasX	Amino Acid Sequence
		LKKRFSHRPVQEKFVCLNCGFETHAAEQAALNIARSWLFLRSQEYKKYQTNKTTGN TDKRAFVETWQSFYRKKLKEVWKPAV
25	dCasX593	QEIKRINKIRRLVKDSNTKKAGKTGPMKTL LVRVMT PDLRERLENLRKKPENI PQ PISNTSRANLNKLLTDYTEMKKAILH VYWEEFQKDPVGLMSRVAQPASKKIDQNK KPEMDEKGNLTTAGFACSQCGQPLFVYKLEQVSEK GKAYTNYFGRCNVAEHEKLI L LAQLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGP VGKALSDACMGTIASFLSKYQDIIIEHQKVVKGNQKRLES LRELAKENLEYPSVT LPPQPHTKEGV DAYNEVIARVRWWVNLNLWQKLKLSRDDAKPLLRLKGFPSPFLVE RQANEVDWDMVCNVKKLINEKKEDGKVFQWQNLAGYKRQEALRPYLSSEEDRKKGK KFARYQLGDL LLLHLEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKLEEERSEDAQ SKAALTDWLRAKASFVIEGLKEADKDEFRCCELKLQKWYGD LRGKPF AIEAENSIL DISGFSKQYNCAFIWQKDGVKKLNLYLI INYFKGGKLRFKKIKPEAFEANRFYTVI NKKSGEIVPMEVNFNFDDPNLII LPLAFGKRQGREFIWN DL SLETGSLKLANGRV IEKTLYNRRTRQDEPALFVALT FERREVL DSSNIKPMNLIGVARGENI PAVIALTD PEGCPLSRFKDSLGNP THILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKAK NLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGKRTFMAERQYTRMEDWLT AKLAYEGLPSKTYLSKTLAQYTSKTC SNCGFTITSADYDRVLEK LKKTATGWMTTI NGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDI SSWTKGRSGEALS LKKRFSHRPVQEKFVCLNCGFETHAAEQAALNIARSWLFLRSQEYKKYQTNKTTGN TDKRAFVETWQSFYRKKLKEVWKPAV
26	dCasX668	QEIKRINKIRRLVKDSNTKKAGKTRGPMKTL LVRVMT PDLRERLENLRKKPENI P QPI SNTSRANLNKLLTDYTEMKKAILH VYWEEFQKDPVGLMSRVAQPASKKIDQNK LKPEMDEKGNLTTAGFACSQCGQPLFVYKLEQVSEK GKAYTNYFGRCNVAEHEKLI L LLAQLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASS PVGKALSDACMGTIASFLSKYQDIIIEHQKVVKGNQKRLES LRELAKENLEYPSV TLPPQPHTKEGV DAYNEVIARVRMWVNLNLWQKLKLSRDDAKPLLRLKGFPSPFLV ERQANEVDWDMVCNVKKLINEKKEDGKVFQWQNLAGYKRQEALRPYLSSEEDRKKG KKFARYQLGDL LLLHLEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKLEEERSEDA QSKAALTDWLRAKASFVIEGLKEADKDEFRCCELKLQKWYGD LRGKPF AIEAENSI LDISGFSKQYNCAFIWQKDGVKKLNLYLI INYFKGGKLRFKKIKPEAFEANRFYTV INKKSGEIVPMEVNFNFDDPNLII LPLAFGKRQGREFIWN DL SLETGSLKLANGR VIEKTLYNRRTRQDEPALFVALT FERREVL DSSNIKPMNLIGVARGENI PAVIALT DPEGCPLSRFKDSLGNP THILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKA KNLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGKRTFMAERQYTRMEDWL TAKLAYEGLPSKTYLSKTLAQYTSKTC SNCGFTITSADYDRVLEK LKKTATGWMTTI INGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDI SSWTKGRSGEALS LLKKRFSHRPVQEKFVCLNCGFETHAAEQAALNIARSWLFLRSQEYKKYQTNKTTG NTDKRAFVETWQSFYRKKLKEVWKPAV
27	dCasX672	QEIKRINKIRRLVKDSNTKKAGKTGPMKTL LVRVMT PDLRERLENLRKKPENI PQ PISNTSRANLNKLLTDYTEMKKAILH VYWEEFQKDPVGLMSRVAQPASKKIDQNK KPEMDEKGNLTTAGFACSQCGQPLFVYKLEQVSEK GKAYTNYFGRCNVAEHEKLI K LAQLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASSP VGKALSDACMGTIASFLSKYQDIIIEHQKVVKGNQKRLES LRELAKENLEYPSVT LPPQPHTKEGV DAYNEVIARVRMWVNLNLWQKLKLSRDDAKPLLRLKGFPSPFLVE RQANEVDWDMVCNVKKLINEKKEDGKVFQWQNLAGYKRQEALRPYLSSEEDRKKGK KFARYQLGDL LLLHLEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKLEEERSEDAQ SKAALTDWLRAKASFVIEGLKEADKDEFRCCELKLQKWYGD LRGKPF AIEAENSIL DISGFSKQYNCAFIWQKDGVKKLNLYLI INYFKGGKLRFKKIKPEAFEANRFYTVI NKKSGEIVPMEVNFNFDDPNLII LPLAFGKRQGREFIWN DL SLETGSLKLANGRV IEKTLYNRRTRQDEPALFVALT FERREVL DSSNIKPMNLIGVARGENI PAVIALTD PEGCPLSRFKDSLGNP THILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKAK NLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGKRTFMAERQYTRMEDWLT AKLAYEGLPSKTYLSKTLAQYTSKTC SNCGFTITSADYDRVLEK LKKTATGWMTTI NGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDI SSWTKGRSGEALS LKKRFSHRPVQEKFVCLNCGFETHAAEQAALNIARSWLFLRSQEYKKYQTNKTTGN TDKRAFVETWQSFYRKKLKEVWKPAV

SEQ ID NO	dCasX	Amino Acid Sequence
28	dCasX676	QEIKRINKIRRLVKDSNTKKAGKTGPMKTLVVRVMTPLRERLENLRKKPENIP QPI SNTSRANLNKLLTDYTEMKA I LHVYWE EFQKDPVGLMSRVAQ PASKKIDQNK LKPEMDEKGNLT TAGFAC SQCGQPLFVYKLEQVSEK GAYTNYFGRCNVAEHEKLI KLAQLKPEKDSDEAVTYS LGKFGQ RALDFYSIHVTKESTHPVKPLAQIAGNRYASS VPGKALSDACMGTIASFLSKYQDIIIEHQVVKGNQKRLESRLRELAKENLEYPSV TLPPQPHTKEGVDAYNEVIARVMWVNLNLWQKLKLSRDDAKPLRLKGFPSPFLV ERQANEVDWDMVCNVKKLINEKKEDGKVFQNLAGYKRQEALRPYLSSEEDRKKG KKFARYQLGDL LLHLEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKLEEEERSEDA QSKAALTDWLRAKASFVIEGLKEADKDEFRCCELKQKQWYGD LRGPFAIEAENSI LDISGFSKQYNCAFIWQKDGVKKLNLYLIINYFKGGKLRFKKIKPEAFEANRFYTV INKKSGEIVPMEVNFNDDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGR VIEKTLYNRRTRQDEPALFVALTFERREVLDSN IKPMNLIGVARGENIPAVIALT DPEGCP LSRFKDSLGNP THILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKA KNLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGRKFMAERQYTRMEDWL TAKLAYEGLPSKTYLSKTLAQYTSKTC SNCGFTITSADYDRVLEK LKKTATGWM T INGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALS LLKKRFSHRPVQEKFVCLNCGFETHAAEQ AALNIARSWLFLRSQ EYKKYQTNKTTG NTDKRAFVETWQSFYRKKLKEVWKPAV
29	dCasX812	QEIKRINKIRRLVKDSNTKKAGKTGPMKTLVVRVMTPLRERLENLRKKPENIPQ PISNTSRANLNKLLTDYTEMKA I LHVYWE EFQKDPVGLMSRVAQ PASKKIDQNK KPEMDEKGNLT TAGFAC SQCGQPLFVYKLEQVSEK GAYTNYFGRCNVAEHEKLI LAQLKPEKDSDEAVTYS LGKFGQ RALDFYSIHVTKESTHPVKPLAQIAGNRYASGP VGKALSDACMGTIASFLSKYQDIIIEHQVVKGNQKRLESRLRELAKENLEYPSVT LPPQPHTKEGVDAYNEVIARVMWVNLNLWQKLKLSRDDAKPLRLKGFSPFLVE RQANEVDWDMVCNVKKLINEKKEDGKVFQNLAGYKRQEALRPYLSSEEDRKKGK KFARYQLGDL LLHLEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKLEEEERSEDAQ SKAALTDWLRAKASFVIEGLKEADKDEFRCCELKQKQWYGD LRGPFAIEAENSI LDISGFSKQYNCAFIWQKDGVKKLNLYLIINYFKGGKLRFKKIKPEAFEANRFYTVI NKKSGEIVPMEVNFNDDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGRV IEKTLYNRRTRQDEPALFVALTFERREVLDSN IKPMNLIGVARGENIPAVIALTD PEGCP LSRFKDSLGNP THILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKAK NLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGRKFMAERQYTRMEDWLT AKLAYEGLPSKTYLSKTLAQYTSKTC SNCGFTITSADYDRVLEK LKKTATGWM TTI NGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALS LKKRFSHRPVQEKFVCLNCGFETHAAEQ AALNIARSWLFLRSQ EYKKYQTNKTTGN TDKRAFVETWQSFYRKKLKEVWKPAV

d. Affinity for the gRNA

[0170] In some embodiments, a long-term repressor fusion protein comprising a dCasX with linked repressor domains has improved affinity for the gRNA relative to a comparable long-term repressor fusion protein comprising a reference dCasX protein with the corresponding linked repressor domains, leading to the formation of the ribonucleoprotein complex (RNP). Increased affinity of the long-term repressor fusion protein for the gRNA may, for example, result in a lower K_d for the generation of a RNP complex, which can, in some cases, result in a more stable ribonucleoprotein complex formation. In some embodiments, the K_d of a long-term repressor fusion protein for a gRNA is increased relative to a reference dCasX protein and linked repressor domains by a factor of at least about 1.1, at

least about 1.2, at least about 1.3, at least about 1.4, at least about 1.5, at least about 1.6, at least about 1.7, at least about 1.8, at least about 1.9, at least about 2, at least about 3, at least about 4, at least about 5, at least about 6, at least about 7, at least about 8, at least about 9, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35, at least about 40, at least about 45, at least about 50, at least about 60, at least about 70, at least about 80, at least about 90, or at least about 100. In some embodiments, the long-term repressor fusion protein comprising the dCasX variant has about 1.1 to about 10-fold increased binding affinity to the gRNA compared to the corresponding repressor fusion protein comprising the catalytically-dead variant of reference CasX protein of SEQ ID NO: 2.

[0171] In some embodiments, increased affinity of the long-term repressor fusion protein for the gRNA results in increased stability of the ribonucleoprotein complex when delivered to mammalian cells, including *in vivo* delivery to a subject. This increased stability can affect the function and utility of the complex in the cells of a subject, as well as result in improved pharmacokinetic properties in blood, when delivered to a subject. In some embodiments, increased affinity of the long-term repressor fusion protein, and the resulting increased stability of the ribonucleoprotein complex, allows for a lower dose of the long-term repressor fusion protein to be delivered to the subject or cells while still having the desired activity; for example *in vivo* or *in vitro* gene repression and/or epigenetic modification. The increased ability to form RNP and keep them in stable form can be assessed using *in vitro* assays known in the art.

[0172] In some embodiments, a higher affinity (tighter binding) of a dCasX variant protein of the long-term repressor fusion protein to a gRNA allows for a greater amount of transcriptional repression and/or epigenetic modification events when both the long-term repressor fusion protein and the gRNA remain in an RNP complex. Increased transcriptional repression events can be assessed using assays described herein.

[0173] Methods of measuring long-term repressor fusion protein binding affinity for a gRNA include *in vitro* methods using purified a long-term repressor fusion protein and a gRNA. The binding affinity for the long-term repressor fusion protein can be measured by fluorescence polarization if the gRNA or long-term repressor fusion protein is tagged with a fluorophore. Alternatively, or in addition, binding affinity can be measured by biolayer interferometry, electrophoretic mobility shift assays (EMSAs), or filter binding. Additional standard techniques to quantify absolute affinities of RNA binding proteins such as the dCasX proteins of the disclosure for specific gRNAs such as reference gRNAs and variants

thereof include, but are not limited to, isothermal calorimetry (ITC), and surface plasmon resonance (SPR).

e. Improved Specificity for a Target Nucleic Acid Sequence

[0174] In some embodiments, a long-term repressor fusion protein comprising a dCasX variant protein with linked repressor domains has improved specificity for a target nucleic acid sequence that is complementary to the targeting sequence of the gRNA relative to a reference dCasX protein with linked repressor domains. As used herein, “specificity,” sometimes referred to as “target specificity,” refers to the degree to which a CRISPR/Cas system ribonucleoprotein complex binds off-target sequences that are similar, but not identical to the target nucleic acid sequence; *e.g.*, an long-term repressor fusion protein containing RNP with a higher degree of specificity would exhibit reduced off-target methylation of sequences relative to an RNP of a reference dCasX with linked repressor domains. The specificity, and the reduction of potentially deleterious off-target effects, of long-term repressor fusion proteins can contribute to achieving an acceptable therapeutic index for use in mammalian subjects. Without wishing to be bound by theory, it is possible that amino acid changes in the helical I and II domains that increase the specificity of the long-term repressor fusion protein for the target nucleic acid strand and can thereby increase the specificity of the long-term repressor fusion protein for the target nucleic acid overall. In some embodiments, amino acid changes that increase specificity of long-term repressor fusion proteins for target nucleic acid may also result in decreased affinity of long-term repressor fusion proteins for DNA, but the overall benefit and safety of the composition is enhanced.

f. Repressor fusion proteins with heterologous proteins

[0175] Also contemplated within the scope of the disclosure are long-term repressor fusion proteins comprising a heterologous protein fused to the long-term repressor fusion protein for use in the systems of the disclosure. This includes long-term repressor fusion proteins comprising N-terminal or C-terminal fusions to a heterologous protein or domain thereof. In some embodiments, the long-term repressor fusion protein is fused to one or more proteins or domains thereof that has a different activity of interest.

[0176] In some cases, a heterologous polypeptide (a fusion partner) for use with a long-term repressor fusion protein provides for subcellular localization, *i.e.*, the heterologous polypeptide contains a subcellular localization sequence (*e.g.*, a nuclear localization signal (NLS) for targeting to the nucleus, a sequence to keep the fusion protein out of the nucleus, a

nuclear export sequence (NES), a sequence to keep the fusion protein retained in the cytoplasm, a mitochondrial localization signal for targeting to the mitochondria, a chloroplast localization signal for targeting to a chloroplast, an ER retention signal, and the like).

[0177] In some cases, a long-term repressor fusion protein includes (is fused to) a nuclear localization signal (NLS). In some cases, a long-term repressor fusion protein is fused to 2 or more, 3 or more, 4 or more, or 5 or more 6 or more, 7 or more, 8 or more NLSs. In some cases, one or more NLSs (2 or more, 3 or more, 4 or more, or 5 or more NLSs) are positioned at or near (*e.g.*, within 20 amino acids of) the N-terminus and/or the C-terminus of the long-term repressor fusion protein. In some cases, one or more NLSs (2 or more, 3 or more, 4 or more, or 5 or more NLSs) are positioned at or near (*e.g.*, within 20 amino acids of) the N-terminus of the long-term repressor fusion protein. In some cases, one or more NLSs (2 or more, 3 or more, 4 or more, or 5 or more NLSs) are positioned at or near (*e.g.*, within 20 amino acids of) the C-terminus of the long-term repressor fusion protein. In some cases, one or more NLSs (3 or more, 4 or more, or 5 or more NLSs) are positioned at or near (*e.g.*, within 20 amino acids of) both the N-terminus and the C-terminus of the long-term repressor fusion protein. In some cases, a single NLS is positioned at the N-terminus and a single NLS is positioned at the C-terminus of the long-term repressor fusion protein. The person of ordinary skill in the art will understand that an NLS at or near the N- or C-terminus of a protein can be within 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 amino acids of the N- or C-terminus. In some embodiments, the NLS linked to the N-terminus of the dCasX or long-term repressor fusion protein are identical to the NLS linked to the C-terminus. In other embodiments, the NLS linked to the N-terminus of the dCasX or long-term repressor fusion protein are different to the NLS linked to the C-terminus. Representative configurations of long-term repressor fusion proteins with NLS are shown in FIG. 61. In some embodiments, NLSs suitable for use with a long-term repressor fusion protein in the systems of the disclosure comprise sequences having at least about 85%, at least about 90%, or at least about 95% identity or are identical to sequences derived from: the NLS of the SV40 virus large T-antigen, having the amino acid sequence PKKKRKV (SEQ ID NO: 30); the NLS from nucleoplasmin (*e.g.*, the nucleoplasmin bipartite NLS with the sequence KRPAATKKAGQAKKKK (SEQ ID NO: 31); the c-MYC NLS having the amino acid sequence PAAKRVKLD (SEQ ID NO: 32) or RQRRNELKRSP (SEQ ID NO: 33). In some embodiments, the NLS and a short peptide linker linked to the N-terminus of the long-term repressor fusion protein is the sequence PKKKRKVSR (SEQ ID NO: 34). In some

embodiments, the NLS and a short peptide linker linked to the N-terminus of the long-term repressor fusion protein is the sequence PKKKRKVSRVNGSGSGGG (SEQ ID NO: 3298). In some embodiments, the NLS and a short peptide linker linked to the C-terminus of the long-term repressor fusion protein is the sequence TSPKKKRKV (SEQ ID NO: 3273). In some embodiments, the NLS linked to the N-terminus of the long-term repressor fusion protein is selected from the group consisting of the N-terminal sequences as set forth in Table 3. In some embodiments, the NLS comprises a sequence selected from the group consisting of SEQ ID NOS: 30-97, 3273 and 3298. In some embodiments, the NLS comprises a sequence selected from the group consisting of SEQ ID NOS: 34-97. In some embodiments, the NLS linked to the C-terminus of the long-term repressor fusion protein is selected from the group consisting of the C-terminal sequences as set forth in Table 4. In some embodiments, NLSs suitable for use with a long-term repressor fusion protein in the systems of the disclosure include sequences having at least about 80%, at least about 90%, or at least about 95% identity or are identical to one or more sequences of Table 3 or Table 4. The person of ordinary skill in the art will appreciate that any of the NLS sequences set forth in Table 3 and Table 4 can be fused to, or proximal to either the N- or C-terminus of the long-term repressor fusion proteins described herein.

Table 3: N-terminal NLS Amino Acid Sequences

NLS Amino Acid Sequence*	SEQ ID NO
PKKKRKVSR	34
PKKKRKVGGS PKKKRKVGGS PKKKRKVGGS PKKKRKVSR	35
PKKKRKVGGS PKKKRKVGGS PKKKRKVGGS PKKKRKVGGS PKKKRKVGGS PKKKRKVSR	36
PAAKRVKLD SR	37
PAAKRVKLD GGS PAAKRVKLD SR	38
PAAKRVKLD GGS PAAKRVKLD GGS PAAKRVKLD GGS PAAKRVKLD SR	39
PAAKRVKLD GGS PAAKRVKLD GGS PAAKRVKLD GGS PAAKRVKLD GGS PAAKRVKLD GGS PAAKRVKLD SR	40
KRPAATKKAGQAKKKK SR	41
KRPAATKKAGQAKKKK GGS KRPAATKKAGQAKKKK SR	42
PAAKRVKLD GGS PKKKRKVSR	43
PAAKKKKLD GGS PKKKRKVSR	44
PAAKKKKLD SR	45

NLS Amino Acid Sequence*	SEQ ID NO
PAAKKKKLD GGSPAAKKKKLDGGSPAAKKKKLDSR	46
PAAKKKKLD GGSPAAKKKKLDGGSPAAKKKKLDGGSPAAKKKKLDSR	47
PAKRARRGYK CSR	48
PAKRARRGYK CGSPAKRARRGYKCSR	49
PRRKREE SR	50
PYRGRKE SR	51
PLRKRPRR SR	52
PLRKRPRR GSPLRKRPRRSR	53
PAAKRVKLD GGKRTADGSEFESPKKKRKVGGS	54
PAAKRVKLD GGKRTADGSEFESPKKKRKVPPPPG	55
PAAKRVKLD GGKRTADGSEFESPKKKRKVGIHGVPAAPG	56
PAAKRVKLD GGKRTADGSEFESPKKKRKVGGSGGGSPG	57
PAAKRVKLD GGKRTADGSEFESPKKKRKVPGGSGGGSPG	58
PAAKRVKLD GGKRTADGSEFESPKKKRKVAEAAAKEAAAKEAAKAPG	59
PAAKRVKLD GGSPKKKKRKVGGS	60
PAAKRVKLD PPPSPKKKKRKVPG	61
PAAKRVKLD PG	62
PAAKRVKLD GGSGGGSGGGSGGS	63
PAAKRVKLD PPP	64
PAAKRVKLD GGSGGGSGGGSPPP	65
PKKKRKV PPP	66
PKKKRKV GGGS	67

*Residues in **bold** are NLS residues, while unbolded residues are linkers.

Table 4: C-terminal NLS Amino Acid Sequences

NLS Amino Acid Sequence	SEQ ID NO
GS PKKKRKV GGSPKKKKRKVGGS PKKKRKV GGSPKKKKRKV	68
GS PKKKRKV GGSPKKKKRKVGGS PKKKRKV GGSPKKKKRKVGGS PKKKRKV GGSPKKKKRKV	69
GS PAAKRVKLD GGSPAAKRVKLD	70
GS PAAKRVKLD GGSPAAKRVKLDGGSPAAKRVKLDGGSPAAKRVKLD	71

NLS Amino Acid Sequence	SEQ ID NO
GSKRPAATKKAGQAKKKK	72
KRPAATKKAGQAKKKKGGSKRPAATKKAGQAKKKK	73
GSKLGP RKATGRWGS	74
GSKRKGS PERGERKRHWGS	75
GS P K K K R K V G S G S K R P A A T K K A G Q A K K K K L E	76
GP K R T A D S Q H S T P P K T K R K V E F E P K K K R K V	77
GGSGGGSKRTADSQHSTPPKTKRKVEFEPKKKRKV	78
AEAAAKEAAAKEAAAKAKRTADSQHSTPPKTKRKVEFEPKKKRKV	79
GPPKKKRKVGSGSKRTADSQHSTPPKTKRKVEFEPKKKRKV	80
GPAAEAAKEAAAKEAAAKAPAAKRVKLD	81
GPGGSGGGSGGGSPAAKRVKLD	82
GP PAAKRVKLD	83
VGSKRPAATKKAGQAKKKK	84
TGGGPGGGAAAGSGSPKKKRKVGSGSKRPAATKKAGQAKKKKLE	85
TGGGPGGGAAAGSGSPKKKRKVGSGS	86
PPP P K K K R K V P P P	87
GGSPKKKRKVP P P	88
PPP P K K K R K V	89
GGSPKKKRKV	90
GGSPKKKRKVGSGSGSGS	91
GGSPKKKRKVGSGSPKKKRKV	92
GGSGGGSGSPKKKRKVGSGSPKKKRKV	93
VGGSGGGSGGGSPAAKRVKLD	94
VPPP PAAKRVKLD	95
VPPPGGGSGGGSGGGSPAAKRVKLD	96
VGSPAAKRVKLD	97

[0178] In some embodiments, the one or more NLSs are linked to the long-term repressor fusion protein or to adjacent NLS with an optional linker peptide. In some embodiments, the linker peptide is selected from the group consisting of SR, GS, GP, TS, VGS, GGS, (G)_n (SEQ ID NO: 98), (GS)_n (SEQ ID NO: 99), (GSGGS)_n (SEQ ID NO: 100), (GGSGGS)_n (SEQ ID NO: 101), (GGGS)_n (SEQ ID NO: 102), GGSG (SEQ ID NO: 103), GGSGG (SEQ

ID NO: 104), GSGSG (SEQ ID NO: 105), GSGGG (SEQ ID NO: 106), GGGSG (SEQ ID NO: 107), GSSSG (SEQ ID NO: 108), GP GP (SEQ ID NO: 109), GGP, PPP, VPPP, PPAPPA (SEQ ID NO: 110), PPPG (SEQ ID NO: 111), PPPGPPP (SEQ ID NO: 112), PPP(GGGS)_n (SEQ ID NO: 113), (GGGS)_nPPP (SEQ ID NO: 114), AEAAAKEAAAKEAAKA (SEQ ID NO: 115), VPPPGGGSGGGSGGG (SEQ ID NO: 116), TGGGPGGGAAAGSGS (SEQ ID NO: 117), GGGSGGGSGGGSPPP (SEQ ID NO: 118), TPPKTKRKVEFE (SEQ ID NO: 119), GSGGGG (SEQ ID NO: 120), GSGSGG (SEQ ID NO: 121), SSGNSNANSRGPSSFSSGLVPLSLRGSH (SEQ ID NO: 122), GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSE (SEQ ID NO: 123), and GGSGGG (SEQ ID NO: 124), where n is 1 to 5.

[0179] In general, NLS (or multiple NLSs) are of sufficient strength to drive accumulation of a long-term repressor fusion protein in the nucleus of a eukaryotic cell. Detection of accumulation in the nucleus may be performed by any suitable technique. For example, a detectable marker may be fused to a long-term repressor fusion protein such that location within a cell may be visualized. Cell nuclei may also be isolated from cells, the contents of which may then be analyzed by any suitable process for detecting protein, such as immunohistochemistry, Western blot, or enzyme activity assay. Accumulation in the nucleus may also be determined indirectly.

IV. Long-term Repressor Domain Fusion Proteins

[0180] The disclosure provides systems of long-term fusion proteins comprising a DNA binding protein linked to multiple repressor domains, wherein the system is capable of binding to a *PCSK9* target nucleic acid and repressing transcription of the *PCSK9* gene, including by epigenetic modification of the target nucleic acid. Exemplary DNA binding proteins for use in the fusion proteins include zinc finger (ZF), TALE (transcription-activator-like effector) proteins, and catalytically-dead CRISPR proteins.

[0181] In some embodiments, the disclosure provides systems of long-term repressor fusion proteins comprising a catalytically-dead CRISPR protein, such as a dCasX, linked to multiple repressor domains that, when complexed with a guide ribonucleic acid (gRNA) comprising a targeting sequence complementary to a target nucleic acid sequence of *PCSK9*, wherein the system is capable of binding to a target nucleic acid of *PCSK9* and repressing or silencing transcription of a *PCSK9* gene. Examples of gene repression processes which

decrease transcription include, but are not limited to, those which inhibit formation of a transcription initiation complex, those which decrease transcription initiation rate, those which decrease transcription elongation rate, those which decrease processivity of transcription and those which antagonize transcriptional activation (by, for example, blocking the binding of a transcriptional activator). Gene repression can constitute, for example, prevention of activation as well as inhibition of expression below an existing level.

Transcriptional repression includes both reversible and irreversible inactivation of gene transcription; the latter can result from epigenetic modification of the target nucleic acid.

[0182] Amongst repressor domains that have the ability to repress, or silence genes, the Krüppel-associated box (KRAB) repressor domain is amongst the most powerful in human genome systems (Alerasool, N., et al. An efficient KRAB domain for CRISPRi applications. *Nat. Methods* 17:1093 (2020)). KRAB-like domains are present in approximately 400 human zinc finger protein-based transcription factors that upon binding of the linked dCasX to the target nucleic acid, is capable of recruiting additional repressor domains such as, but not limited to, Trim28 (also known as Kap1 or Tif1-beta) that, in turn, assembles a protein complex with chromatin regulators such as CBX5/HP1 α and SETDB1 that induce repression of transcription of the gene, but do so in a limited, temporal fashion. Representative, non-limiting examples of KRAB domains suitable for use in the systems of the disclosure include ZIM3 (SEQ ID NO: 129) and ZNF10 (SEQ ID NO: 128). The disclosure provides additional repressor domains that are from human sources, as well as repressor domains that are from non-human sources with markedly distinct sequences (referred to herein as "RD1") that have been found to result in enhanced transcriptional repression compared to ZIM3 and ZNF10 when incorporated in long-term repressor fusion protein construct embodiments, as described herein.

[0183] In some embodiments, the disclosure provides systems in which the modification imparted by use of the LTRP:gRNA system is epigenetic, and hence the silencing of the *PCSK9* gene is heritable by mechanisms other than by edited DNA replication. As used herein "epigenetic modification" means a modification to either DNA or histones associated with DNA, other than a change in the DNA sequence itself (*e.g.*, a substitution, deletion or rearrangement), wherein the modification is either a direct modification by a component of the system or is indirect by the recruitment of one or more additional cellular components, but in which the DNA target nucleic acid sequence itself is not edited to change the sequence. For example, DNA methyltransferase 3A (DNMT3A) (or its catalytic domain) directly

modifies the DNA by methylating it, whereas KRAB recruits KAP-1/TIF1 β corepressor complexes that act as potent transcriptional repressors and can further recruit factors associated with DNA methylation and formation of repressive chromatin, such as heterochromatin protein 1 (HP1), histone deacetylases and histone methyltransferases (Ying, Y., et al. The Krüppel-associated box repressor domain induces reversible and irreversible regulation of endogenous mouse genes by mediating different chromatin states. *Nucleic Acids Res.* 43(3): 1549 (2015)). Further, the catalytically inactive DNMT3L cofactor helps establish a heritable methylation pattern after DNA replication, together with endogenous DNMT1 of the cell. The ATRX-DNMT3-DNMT3L domain (ADD) of DNMT3A is known to have two key functions: 1) it allosterically regulates the catalytic activity of DNMT3A by serving as a methyltransferase auto-inhibitory domain, and 2) it specifically interacts with histone H3 tails that are unmethylated at lysine (K)4, leading to the preferential methylation of DNA bound to chromatin H3 tails that are unmethylated at K4 (Zhang, Y., et al. Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Research* 38:4246 (2010)). In some embodiments, the inclusion of the ADD domain enhances transcriptional repression of targeted genes when incorporated into LTRP, when compared to otherwise equivalent LTRP lacking the ADD domain. In other embodiments, the inclusion of the ADD domain enhances the specificity of the transcriptional repression of targeted genes when compared to otherwise equivalent LTRP lacking the ADD domain. Supporting data for the foregoing are provided in the Examples, as well as in WO2023049742A2, incorporated by reference herein.

[0184] In some embodiments, the long-term repressor fusion protein comprises a DNA binding protein linked to a first, second, and third repressor domain, wherein each of the repressor domains are different and the fusion protein is capable of binding to the *PCKS9* target nucleic acid. In some embodiments, the long-term repressor fusion protein comprises a DNA binding protein linked to a first, second, third, and fourth repressor domain, wherein each of the repressor domains are different. In any of the foregoing embodiments, the fusion protein is capable of forming an RNP with a gRNA of the system that binds to the target nucleic acid when the DNA binding domain comprises a catalytically dead CRISPR protein.

[0185] In some embodiments, the DNA binding protein comprises a TALE that can bind but not cleave the target nucleic acid. In some embodiments, the DNA binding protein comprises a zinc-finger protein modified to bind but not cleave the target nucleic acid. In some embodiments, the DNA binding protein comprises a catalytically dead CRISPR protein

that can complex with a gRNA to form an RNP that can bind but not cleave the target nucleic acid. In some embodiments, the long-term repressor fusion protein comprises a catalytically-dead CRISPR protein sequence, a first repressor domain (hereinafter referred to as "RD1"), a DNMT3A catalytic domain (herein after referred to as "DNMT3A") from the DNMT3A protein as the second domain, and a DNMT3L interaction domain (herein after referred to as "DNMT3L") from the DNMT3L protein as the third domain. In some embodiments, the long-term repressor fusion protein comprises a catalytically-dead CRISPR protein sequence, an RD1, a DNMT3A as the second domain, a DNMT3L as the third domain, and an ATRX-DNMT3-DNMT3L domain (herein after referred to as "ADD") from the DNMT3A protein as the fourth domain. In some embodiments, the long-term repressor fusion protein comprises a dCasX and further comprises a first and a second NLS and one or more linker peptides described herein. In some embodiments, the long-term repressor fusion protein is capable of forming an RNP with a gRNA that binds to the target nucleic acid. It has been discovered that the use of the foregoing domains, when configured in select orientations relative to a dCasX in a long-term repressor fusion protein, can result in pronounced epigenetic modification of a *PCSK9* target nucleic acid when complexed with a gRNA with a targeting sequence complementary to defined regions of the *PCSK9* gene, and that the combination of the repressor domains work in synchrony, resulting in an additive or synergistic effect on transcriptional silencing of the targeted gene, depending on the configuration.

[0186] Representative amino acid sequences of the components utilized in the long-term repressor fusion protein constructs are provided herein.

[0187] In some embodiments, the dCasX of the long-term repressor fusion protein comprises a sequence selected from the group consisting of SEQ ID NOS: 4-29, or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% identity thereto. In some embodiments, the dCasX of the long-term repressor fusion protein comprises a sequence selected from the group consisting of SEQ ID NOS: 4-29. In some embodiments, the dCasX of the long-term repressor fusion protein comprises the sequence of SEQ ID NO: 4, or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% identity thereto.

In some embodiments, the dCasX of the long-term repressor fusion protein comprises the sequence of SEQ ID NO: 4.

[0188] In some embodiments, the RD1 of the long-term repressor fusion protein comprises a sequence selected from the group consisting of SEQ ID NOS: 128-1726, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto. In some embodiments, the RD1 of the long-term repressor fusion protein comprises a sequence selected from the group consisting of SEQ ID NOS: 128-1726. In some embodiments, the RD1 of the long-term repressor fusion protein comprises a sequence selected from the group consisting of SEQ ID NO: 128 or SEQ ID NO: 129, or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the RD1 of the long-term repressor fusion protein comprises a sequence selected from the group consisting of SEQ ID NO: 128 or SEQ ID NO: 129. In another embodiment, the first repressor domain (RD1) of the long-term repressor fusion protein comprises a sequence selected from the group consisting of SEQ ID NOS: 130-1726, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In another embodiment, the first repressor domain (RD1) of the long-term repressor fusion protein comprises a sequence selected from the group consisting of SEQ ID NOS: 130-1726. In another embodiment, the first repressor domain (RD1) of the long-term repressor fusion protein comprises a sequence selected from the group consisting of SEQ ID NOS: 130-224 or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In another embodiment, the first repressor domain (RD1) of the long-term repressor fusion protein comprises a sequence selected from the group consisting of SEQ ID NOS: 130-224. In another embodiment, the first repressor domain (RD1) of the long-term repressor fusion protein comprises a sequence selected from the group consisting of SEQ ID NOS: 130-138 or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least

about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In another embodiment, the first repressor domain (RD1) of the long-term repressor fusion protein comprises a sequence selected from the group consisting of SEQ ID NOS: 130-138. In another embodiment, the first repressor domain (RD1) of the long-term repressor fusion protein comprises a sequence of SEQ ID NO: 135 or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In another embodiment, the first repressor domain (RD1) of the long-term repressor fusion protein comprises a sequence of SEQ ID NO: 131 or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In another embodiment, the first repressor domain (RD1) of the long-term repressor fusion protein comprises a sequence of SEQ ID NO: 135. In another embodiment, the first repressor domain (RD1) of the long-term repressor fusion protein comprises a sequence of SEQ ID NO: 130, or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In another embodiment, the first repressor domain (RD1) of the long-term repressor fusion protein comprises a sequence of SEQ ID NO: 130. In another embodiment, the first repressor domain (RD1) of the long-term repressor fusion protein comprises a sequence of SEQ ID NO: 131. In another embodiment, the first repressor domain (RD1) of the long-term repressor fusion protein comprises a sequence of SEQ ID NO: 132. In another embodiment, the first repressor domain (RD1) of the long-term repressor fusion protein comprises a sequence of SEQ ID NO: 133. In another embodiment, the first repressor domain (RD1) of the long-term repressor fusion protein comprises a sequence of SEQ ID NO: 134. In another embodiment, the first repressor domain (RD1) of the long-term repressor fusion protein comprises a sequence of SEQ ID NO: 135. In another embodiment, the first repressor domain (RD1) of the long-term repressor fusion protein comprises a sequence of SEQ ID NO: 136. In another embodiment, the first repressor domain (RD1) of the long-term repressor fusion protein comprises a sequence of SEQ ID NO: 137. In another embodiment, the first repressor domain (RD1) of the long-term repressor fusion protein comprises a sequence of SEQ ID NO: 138.

[0189] In some embodiments, the second repressor domain of the long-term repressor fusion protein is a DNMT3A, comprising a sequence of SEQ ID NO: 126, or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% identity thereto. In some embodiments, the second repressor domain of the long-term repressor fusion protein comprises a sequence of SEQ ID NO: 126.

[0190] In some embodiments, the third repressor domain of the long-term repressor fusion protein is a DNMT3L, comprising a sequence of SEQ ID NO: 127, or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto.

[0191] In some embodiments, the optional fourth repressor domain of the long-term repressor fusion protein is an ADD, comprising a sequence of SEQ ID NO: 125, or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the optional fourth repressor domain of the long-term repressor fusion protein comprises a sequence of SEQ ID NO: 125. In some embodiments, the C terminal of the ADD is linked to the N-terminal of the DNMT3A. It has been discovered that the addition of the ADD to the long-term repressor fusion proteins comprising the RD1, DNMT3A, and DNMT3L greatly enhances or increases the long-term repression and/or epigenetic modification of the target nucleic acid, as well as the specificity of the repression, in comparison to long-term repressor fusion proteins lacking the ADD. Exemplary data for the improved repression and specificity of long-term repressor fusion proteins comprising the ADD are presented in the Examples, and is described in WO2023049742A2, incorporated by reference herein. In some embodiments, the fusion protein comprises one or more linker peptides selected from the group consisting of SEQ ID NOS: 98-124, 3278-3289 (exemplary sequences are shown in Tables 5 and 58), and one or more NLS comprising a sequence selected from the group consisting of SEQ ID NOS: 30-97, 32998 and 3299. Exemplary configurations of long-term repressor fusion proteins comprising the ADD are presented in FIG. 61.

[0192] In some embodiments, the long-term repressor fusion protein comprises, from N- to C-terminus, a DNMT3A, a DNMT3L, a DNA binding protein, and an RD1. In some embodiments, the long-term repressor fusion protein comprises, from N- to C-terminus, an ADD, a DNMT3A, a DNMT3L, a DNA binding protein, and an RD1. In some embodiments, the DNA binding protein may be a zinc finger, a TALE, or a catalytically-dead CRISPR protein. In some embodiments the long-term repressor fusion protein comprises an NLS at the N terminus, the C terminus, or both. In some embodiments, the long-term repressor fusion protein comprises one or more linkers between the DNMT3A and DNMT3L, between the DNMT3L and DNA binding protein, and/or between the DNA binding protein and the RD1.

[0193] In some embodiments, the long-term repressor fusion protein comprises, from N- to C-terminus, a DNA binding protein, an RD1, a DNMT3A, and a DNMT3L. In some embodiments, the long-term repressor fusion protein comprises, from N- to C-terminus, a DNA binding protein, an RD1, and ADD, a DNMT3A, and a DNMT3L. In some embodiments, the DNA binding protein may be a zinc finger, a TALE, or a catalytically-dead CRISPR protein. In some embodiments the long-term repressor fusion protein comprises an NLS at the N terminus. In some embodiments the long-term repressor fusion protein comprises an NLS between the RD1 and the DNMT3A. In some embodiments the long-term repressor fusion protein comprises an NLS at the N terminus, the C terminus, or both. In some embodiments, the long-term repressor fusion protein comprises one or more linkers between an N terminal NLS and the DNA binding protein, between the DNA binding protein and the RD1, between the RD1 and the DNMT3A, or optionally, the ADD, and/or between the DNMT3A and the DNMT3L.

[0194] In some embodiments, the long-term repressor fusion protein comprises, from N- to C-terminus, a DNA binding protein, a DNMT3A, a DNMT3L, and an RD1. In some embodiments, the long-term repressor fusion protein comprises, from N- to C-terminus, a DNA binding protein, an ADD, a DNMT3A, a DNMT3L, and an RD1. In some embodiments, the DNA binding protein may be a zinc finger, a TALE, or a catalytically-dead CRISPR protein. In some embodiments the long-term repressor fusion protein comprises an NLS at the N terminus, the C terminus, or both. In some embodiments, the long-term repressor fusion protein comprises one or more linkers between an N terminal NLS and the DNA binding protein, between the DNA binding protein and the DNMT3A, or optionally the ADD, between the DNMT3A and the DNMT3L, and/or between the DNMT3L and the RD1.

[0195] In some embodiments, the long-term repressor fusion protein comprises, from N- to C-terminus, an RD1, a DNMT3A, a DNMT3L, and a DNA binding protein. In some embodiments, the long-term repressor fusion protein comprises, from N- to C-terminus, an RD1, an ADD, a DNMT3A, a DNMT3L, and a DNA binding protein. In some embodiments, the DNA binding protein may be a zinc finger, a TALE, or a catalytically-dead CRISPR protein. In some embodiments the long-term repressor fusion protein comprises an NLS at the N terminus, the C terminus, or both. In some embodiments, the long-term repressor fusion protein comprises one or more linkers between the RD1 and the DNMT3A, or optionally the ADD, between the DNMT3A and the DNMT3L, between the DNMT3L and the DNA binding protein, and/or between the DNA binding protein and a C terminal NLS.

[0196] In some embodiments, the long-term repressor fusion protein comprises, from N- to C-terminus, a DNMT3A, a DNMT3L, an RD1, and a DNA binding protein. In some embodiments, the long-term repressor fusion protein comprises, from N- to C-terminus, an ADD, a DNMT3A, a DNMT3L, an RD1, and a DNA binding protein. In some embodiments, the DNA binding protein may be a zinc finger, a TALE, or a catalytically-dead CRISPR protein. In some embodiments the long-term repressor fusion protein comprises an NLS at the N terminus, the C terminus, or both. In some embodiments, the long-term repressor fusion protein comprises one or more linkers between the DNMT3A and the DNMT3L, between the DNMT3L and the RD1, between the RD1 and the DNA binding protein, and/or between the DNA binding protein and a C terminal NLS.

[0197] In some embodiments, the long-term repressor fusion protein comprises, from N- to C-terminus, a DNMT3A, a DNMT3L, an RD1, a DNA binding protein, and a second RD1. In some embodiments, the long-term repressor fusion protein comprises, from N- to C-terminus, an ADD, a DNMT3A, a DNMT3L, an RD1, a DNA binding protein, and a second RD1. In one embodiment of the foregoing, the second RD1 may be identical in sequence to the first RD1. In another embodiment of the foregoing, the second RD1 may be different in sequence to the first RD1. In some embodiments, the DNA binding protein may be a zinc finger, a TALE, or a catalytically-dead CRISPR protein. In some embodiments the long-term repressor fusion protein comprises an NLS at the N terminus, the C terminus, or both. In some embodiments, the long-term repressor fusion protein comprises one or more linkers between the DNMT3A and DNMT3L, between the DNMT3L and the RD1, between the RD1 and the DNA binding protein, between the DNA binding protein and the second RD1, and/or between the second RD1 and a C terminal NLS.

[0198] In some cases, the long-term repressor fusion protein also comprises one or more NLS. In some embodiments, the long-term repressor fusion protein comprises a configuration of, N-terminal to C-terminal of NLS-ADD-DNMT3A-DNMT3L-DNA binding protein-RD1-NLS, NLS-DNA binding protein-RD1-NLS-ADD-DNMT3A-DNMT3L, NLS- DNA binding protein-ADD-DNMT3A-DNMT3L- RD1-NLS), NLS-RD1-ADD-DNMT3A-DNMT3L-DNA binding protein-NLS, NLS-ADD-DNMT3A-DNMT3L-RD1-DNA binding protein-NLS, or NLS-ADD-DNMT3A-DNMT3L-RD1-DNA binding protein-RD1-NLS. In some embodiments, the DNA binding protein may be a zinc finger, a TALE, or a catalytically-dead CRISPR protein. In some embodiments, the long-term repressor fusion protein is capable of binding to and repressing or silencing transcription of the *PCSK9* gene.

[0199] In some embodiments of the long-term repressor fusion protein, one or more linker peptides may be inserted between any two adjacent domains of the long-term repressor fusion protein. In some embodiments, the long-term repressor fusion protein comprises a configuration of, N-terminal to C-terminal, of NLS-ADD-DNMT3A-Linker2-DNMT3L-Linker1-Linker3A-DNA binding protein-Linker3B-RD1-NLS (configuration 1). In some embodiments, the long-term repressor fusion protein comprises a configuration of, N-terminal to C-terminal, of NLS-Linker3A-DNA binding protein-Linker3B-RD1-NLS-Linker1-ADD-DNMT3A-Linker2-DNMT3L (configuration 2). In some embodiments, the long-term repressor fusion protein comprises a configuration of, N-terminal to C-terminal, of NLS-Linker3A-DNA binding protein-Linker1-ADD-DNMT3A-Linker2-DNMT3L-Linker3B-RD1-NLS (configuration 3). In some embodiments, the long-term repressor fusion protein comprises a configuration of, N-terminal to C-terminal, of NLS-RD1-Linker3A-ADD-DNMT3A-Linker2-DNMT3L-Linker1-DNA binding protein-Linker3B-NLS (configuration 4). In some embodiments, the long-term repressor fusion protein comprises a configuration of, N-terminal to C-terminal, of NLS-ADD-DNMT3A-Linker2-DNMT3L-Linker3A-RD1-Linker1-DNA binding protein-Linker3B-NLS (configuration 5). In some embodiments, the DNA binding protein of the foregoing configurations may be a zinc finger, a TALE, or a catalytically-dead CRISPR protein. Schematics of the configurations portrayed in FIG. 61. In some embodiments of the LTRP of configurations 1-5, the NLS may comprise a sequence selected from the group consisting of SEQ ID NOS: 30-97, 3298 and 3299 (Tables 3 and 4), the linker sequences may comprise sequences independently selected from the group consisting of SEQ ID NOS: 98-124, 3278-3289 (representative linkers shown in Tables 5 and 58). In some embodiments, the NLS may comprise a sequence of SEQ ID NO:

30, and the linker sequences may comprise sequences independently selected from the group consisting of SEQ ID NOS: 120, and 122-124. In some embodiments, the DNA binding protein may be a dCasX sequence selected from the group consisting of SEQ ID NOS: 4-29, or a sequence having or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90% at least about 91%, at least about 92%, at least about 93% at least about 94% at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the second repressor domain is a DNMT3A domain comprising the sequence of SEQ ID NO: 126, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90% at least about 91%, at least about 92%, at least about 93% at least about 94% at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the third repressor is a DNMT3L domain comprising the sequence of SEQ ID NO: 127, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90% at least about 91%, at least about 92%, at least about 93% at least about 94% at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the fourth repressor is an ADD domain comprising the sequence of SEQ ID NO: 125, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90% at least about 91%, at least about 92%, at least about 93% at least about 94% at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the long-term repressor fusion proteins are capable of binding to and repressing or silencing the *PCSK9* gene target nucleic acid.

[0200] In some embodiments, the disclosure provides a system comprising a long-term repressor fusion protein comprising two RD1 domains, a second repressor domain, a third repressor domain, and a fourth repressor domain operably linked to a DNA binding protein; *e.g.*, a zinc finger, TALE, or catalytically-dead CRISPR protein. In some embodiments, the DNA binding protein comprises a catalytically-dead CasX comprising a sequence selected from the group consisting of SEQ ID NOS: 4-29, or sequence having at least 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90% at least about 91%, at least about 92%, at least about 93% at least about 94% at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto. In some embodiments, the sequences of the two RD1 are identical. In other embodiments, the two RD1 comprise different sequences. In some embodiments, the two RD1 are located N-

terminal to the DNA binding protein. In some embodiments, the two RD1 are located C-terminal to the DNA binding protein. In some embodiments, one RD1 is located N-terminal to the DNA binding protein and one RD1 is located C-terminal to the DNA binding protein.

[0201] In some embodiments, the long-term repressor fusion protein comprises two RD1. In some embodiments, the long-term repressor fusion protein comprises a configuration of, N-terminal to C-terminal, of NLS-ADD-DNMT3A-Linker2-DNMT3L-Linker3A-a RD1a-Linker1-DNA binding protein-Linker3B-RD1a-Linker 4-NLS (configuration 6a), wherein the RD1a sequences are identical (see FIG. 61 as a schematic of the long-term repressor fusion protein). In some embodiments, the long-term repressor fusion protein comprises a configuration of, N-terminal to C-terminal, of NLS-ADD-DNMT3A-Linker2-DNMT3L-Linker3A-a RD1a-Linker1-DNA binding protein-Linker3B-RD1b-Linker 4-NLS (configuration 6b), wherein the RD1a and RD1b sequences are different (see FIG. 61 as a schematic of the long-term repressor fusion protein). In some embodiments, the DNA binding protein may be a zinc finger, a TALE, or a catalytically-dead CRISPR protein. In some embodiments, a long-term repressor protein embodiment of the paragraph is capable of binding to and repressing or silencing the expression of the *PCSK9* gene.

[0202] In some embodiments, the disclosure provides a long-term repressor fusion protein of configuration 6a, wherein the two RD1 sequences are identical. In some embodiments of the configuration 6a long-term repressor fusion protein, the DNA binding protein comprises a dCasX of the sequence of SEQ ID NO: 4, or a sequence having at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90% at least about 91%, at least about 92%, at least about 93% at least about 94% at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto, the first and second copies of the RD1a comprise a sequence selected from the group consisting of SEQ ID NOS: 130-1726, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90% at least about 91%, at least about 92%, at least about 93% at least about 94% or at least about 95% identity thereto, the second repressor domain is a DNMT3A comprising the sequence of SEQ ID NO: 126, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90% at least about 91%, at least about 92%, at least about 93% at least about 94% at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto, the third repressor is a DNMT3L comprising the sequence of SEQ ID NOS: 127, or a sequence variant having at least about 70%, at least about 80%, at least about 85%, at least about 90% at least about

91%, at least about 92%, at least about 93% at least about 94% at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto, the fourth repressor is an ADD comprising the sequence of SEQ ID NO: 125, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90% at least about 91%, at least about 92%, at least about 93% at least about 94% at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto, the NLSs comprise sequences independently selected from the group consisting of SEQ ID NOS: 30-97 (Tables 3 and 4), the L1 linker comprises a sequence of SEQ ID NO: 123, the L2 linker comprises a sequence of SEQ ID NO: 122, the L3A linker comprises a sequence of SEQ ID NO: 124, the L3B linker comprises a sequence of SEQ ID NO: 120, and the L4 linker comprises a sequence of SEQ ID NO: 3288 or SEQ ID NO: 3289. In some embodiments of the long-term repressor fusion protein of configuration 6a, the linker sequences are independently selected from the group consisting of SEQ ID NOS: 98-124, 3278-3289 (exemplary sequences shown in Tables 5 and 58). In some embodiments of the long-term repressor fusion protein of configuration 6a, wherein the two copies of the RD1 are identical, each of the RD1 comprise a sequence selected from the group consisting of SEQ ID NOS: 130, 131, and 135, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90% at least about 91%, at least about 92%, at least about 93% at least about 94% at least about 95% identity thereto. A schematic of configuration 6a is shown in FIG. 61. In some embodiments, the long-term repressor fusion protein of configuration 6a is capable of forming an RNP with a gRNA of the disclosure and is capable of binding to and repressing or silencing the gene target nucleic acid.

[0203] In some embodiments of the configuration 6b long-term repressor fusion protein, wherein the two RD1 sequences are different, the DNA binding protein comprises a dCasX of the sequence of SEQ ID NO: 4, or a sequence having at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90% at least about 91%, at least about 92%, at least about 93% at least about 94% at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto, the first copy of the first repressor domain (RD1a) comprises a sequence selected from the group consisting of SEQ ID NOS: 130-1726, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90% at least about 91%, at least about 92%, at least about 93% at least about 94% or at least about 95% identity thereto, the second copy of the first repressor domain (RD1b) comprises a different sequence selected

from the group consisting of SEQ ID NOS: 130-1726, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90% at least about 91%, at least about 92%, at least about 93% at least about 94% or at least about 95% identity thereto, the second repressor domain is a DNMT3A comprising the sequence of SEQ ID NO: 126, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90% at least about 91%, at least about 92%, at least about 93% at least about 94% at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto, the third repressor is a DNMT3L comprising the sequence of SEQ ID NOS: 127, or a sequence variant having at least about 70%, at least about 80%, at least about 85%, at least about 90% at least about 91%, at least about 92%, at least about 93% at least about 94% at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto, the fourth repressor is an ADD comprising the sequence of SEQ ID NO: 125, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90% at least about 91%, at least about 92%, at least about 93% at least about 94% at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto, the NLSs comprises sequences independently selected from the group consisting of SEQ ID NOS: 30-97 of Tables 3 and 4, the L1 linker comprises a sequence of SEQ ID NO: 123, the L2 linker comprises a sequence of SEQ ID NO: 122, the L3A linker comprises a sequence of SEQ ID NO: 124, the L3B linker comprises a sequence of SEQ ID NO: 120, and the L4 linker comprises a sequence of SEQ ID NO: 3288 or SEQ ID NO: 3289. In some embodiments of the long-term repressor fusion protein of configuration 6b, the linker sequences are selected from the group consisting of SEQ ID NOS: 98-124, 3278-3289 (exemplary sequences shown in Tables 5 and 58). In some embodiments of the configuration 6b long-term repressor fusion protein, wherein the two copies of the RD1 are different, the RD1a comprises a sequence of SEQ ID NO: 130, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, or at least about 95% identity thereto, and the RD1b comprises a sequence of SEQ ID NO: 131, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90% at least about 91%, at least about 92%, at least about 93%, at least about 94%, or at least about 95% identity thereto. In some embodiments of the configuration 6b long-term repressor fusion protein, wherein the two copies of the RD1 are different, the RD1a comprises a sequence of SEQ ID NO: 130, or a sequence having at least about 70%, at least about 80%, at least about

85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, or at least about 95% identity thereto, and the RD1b comprises a sequence of SEQ ID NO: 135, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, or at least about 95% identity thereto. In some embodiments of the configuration 6b long-term repressor fusion protein, wherein the two copies of the RD1 are different, the RD1a comprises a sequence of SEQ ID NO: 131, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, or at least about 95% identity thereto, and the RD1b comprises a sequence of SEQ ID NO: 130, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, or at least about 95% identity thereto. In some embodiments of the configuration 6b long-term repressor fusion protein, wherein the two copies of the RD1 are different, the RD1a comprises a sequence of SEQ ID NO: 131, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, or at least about 95% identity thereto, and the RD1b comprises a sequence of SEQ ID NO: 135, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, or at least about 95% identity thereto. In some embodiments of the configuration 6b long-term repressor fusion protein, wherein the two copies of the RD1 are different, the RD1a comprises a sequence of SEQ ID NO: 135, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, or at least about 95% identity thereto, and the RD1b comprises a sequence of SEQ ID NO: 130, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, or at least about 95% identity thereto. In some embodiments of the configuration 6b long-term repressor fusion protein, wherein the two copies of the RD1 are different, the RD1a comprises a sequence of SEQ ID NO: 135, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, or at least about 95% identity thereto, and the RD1b comprises a sequence of SEQ ID NO: 131, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about

90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, or at least about 95% identity thereto. In some embodiments of the configuration 6b long-term repressor fusion protein, wherein the two copies of the RD1 are different, the RD1a and RD1b are independently selected from the group consisting of the sequences of SEQ ID NOS: 132-134 and 136-138. In some embodiments of the configuration 6b long-term repressor fusion protein, wherein the two copies of the RD1 are different, the RD1a and Rd1b are independently selected from the group consisting of the sequences of SEQ ID NOS: 130, 131 and 135. A schematic of configuration 6b is shown in FIG. 61. In some embodiments, the long-term repressor fusion protein is capable of forming an RNP with a gRNA of the disclosure and is capable of binding to and repressing or silencing the gene target nucleic acid.

Table 5: Exemplary linker amino acid sequences for long-term repressor fusion proteins

Linker position	Linker Set #	Amino acid sequence	SEQ ID NO
L1	Baseline/Original	GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTST EPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSE	123
	Set 85	ASPAAPAPASPAAPASAPAA	3278
	Set 43	GSGSGSGSGSPVPSTPPTPSPSTPPTP	3279
	Set 36	PPTPSPSPVPSTPPTNSSSTPPTPS	3280
	Set 1	ASAAAPAAASAAASAPSAAAA	3281
	Set 66	AASPAAPSAPPAAASP	3282
L2	Baseline/Original	SSGNSNANSRGPFSGLVPLSLRGS	122
L3A	Baseline/Original	GGSGGG	124
	Set 85	IRAHGD	3283
	Set 43	AFPAAPAPA	3284
	Set 36	GSGNSSGSGGS	3285
	Set 1	ASAAAPAAA	3286
	Set 66	PPTP	3287
L3B	Baseline/Original	GGSGGGS	120
L4v1	Baseline/Original	GSGS	3288
L4v2	Baseline/Original	GSGSGSG	3289

[0204] In some embodiments of the long-term repressor fusion protein comprising a dCasX and, optionally, an ADD, and which is configured as configuration 1, the long-term repressor protein comprises a sequence selected from the group consisting of SEQ ID NOS: 22836-22855, or a sequence having at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or at least 99% identity thereto. In

some embodiments of the long-term repressor fusion protein comprising a dCasX and, optionally, an ADD, and which is configured as configuration 1, the long-term repressor protein comprises a sequence selected from the group consisting of SEQ ID NOS: 22836-22855. In some embodiments of the long-term repressor fusion protein comprising a dCasX and, optionally, an ADD, and which is configured as configuration 1, the long-term repressor protein comprises a sequence selected from the group consisting of SEQ ID NOS: 22838 and 22847. In some embodiments of the long-term repressor fusion protein comprising a dCasX and, optionally, an ADD, and which is configured as configuration 1, the long-term repressor protein comprises a sequence selected from the group consisting of SEQ ID NOS: 22839 and 22848. In some embodiments of the long-term repressor fusion protein comprising a dCasX and, optionally, an ADD, and which is configured as configuration 1, the long-term repressor protein comprises a sequence selected from the group consisting of SEQ ID NOS: 22840 and 22849.

[0205] In some embodiments of the system comprising a DNA binding protein, and is configured as a configuration selected from the group consisting of configuration 1, configuration 2, configuration 3, configuration 4, configuration 5, configuration 6a, and configuration 6b, described supra, the long-term repressor fusion protein comprises a dCasX and is capable of complexing with a gRNA with a targeting sequence complementary to the *PCSK9* target nucleic acid in a cell to form an RNP and, upon binding of an RNP to the *PCSK9* target nucleic acid in a cell, the target nucleic acid is epigenetically-modified and transcription of the *PCSK9* gene is repressed. In some embodiments, transcription of the *PCSK9* gene is repressed by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or at least 99%. In some embodiments, transcription of the *PCSK9* gene of at least about 1%, at least about 2%, at least about 3%, at least about 4%, at least about 5%, at least about 6%, at least about 7%, at least about 8%, at least about 9%, or at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, or at least about 80% or more of cells in a population of cells is repressed. In some embodiments, the transcriptional repression is assayed in an *in vitro* assay, including cell-based assays, when compared to untreated cells or cells treated with a comparable system comprising a non-targeting spacer. In some embodiments, the transcriptional repression is assayed *in vivo*, in cells obtained from a subject in which the long-term repressor fusion

protein and the gRNA with a targeting sequence complementary to the *PCSK9* target nucleic acid of a gene in a cell are administered; either as a protein and gRNA, or as nucleic acids (e.g., a gRNA and an mRNA encoding the long-term repressor fusion protein), wherein the subject is selected from the group consisting of mouse, rat, pig, non-human primate, and human.

[0206] Most preferably, *PCSK9* gene repression results in complete inhibition of gene expression, such that no gene product is detectable. However, the skilled artisan will appreciate that incomplete inhibition may still be useful and desired for various applications. In some embodiments, the repression of transcription of the *PCSK9* gene by an embodiment of the system is sustained for at least about 8 hours, at least about 1 day, at least about 7 days, at least 2 weeks, at least about 3 weeks, at least about 1 month, or at least about 2 months, when assayed in an *in vitro* assay, including cell-based assays. In some embodiments, the repression of transcription of the *PCSK9* gene by the system compositions is sustained for at least about 7 days, at least 2 weeks, at least about 3 weeks, at least about 1 month, at least about 2 months, at least about 3 months, at least about 4 months, at least about 5 months, or at least about 6 months in targeted cells of a subject when the composition comprising the system of the disclosure is administered as a therapeutically effective dose. In some embodiments, the subject is selected from the group consisting of mouse, rat, pig, non-human primate, and human. In some embodiments, use of an LTRP of configurations 1, 4, 5 and 6, when used in a LTRP:gRNA system of the embodiments, results in off-target methylation or off-target activity that is less than about 10%, less than about 9%, less than about 8%, less than about 7%, less than about 6%, less than about 5%, less than about 4%, less than about 3%, less than about 2%, less than about 1%, less than 0.5%, or less than 0.1% in the cells.

[0207] In some embodiments, repression of transcription of the *PCSK9* gene in cells treated with an LTRP:gRNA system of the embodiments is heritable and is stable through one or more cell divisions. In some embodiments, repression of transcription is stable through 1, 10, 20, 30, 40, 50, 60, 70, 80, 90, or 100 cell divisions, or more.

V. mRNA compositions encoding long-term repressor fusion proteins

[0208] In another aspect, the disclosure relates to messenger RNA (mRNA) compositions comprising sequences of the individual components, as well as the full-length mRNA sequences of the long-term repressor domain fusion protein constructs of the disclosure. In some embodiments, the mRNA compositions have utility in the transcriptional repression and epigenetic modification of a gene, e.g. a *PCSK9* gene, when used to express the long-term

repressor fusion protein. In some cases, the mRNA are designed for use in certain delivery formulations; *e.g.*, nanoparticles such as synthetic nanoparticles or lipid nanoparticles (LNP). The disclosure also provides methods utilized to design the mRNA sequences of the mRNAs used in the compositions. In some cases, an mRNA encoding the long-term repressor fusion protein may be co-formulated in a nanoparticle (such as an LNP) with a gRNA comprising a targeting sequence complementary to a sequence of a *PCSK9* gene. Upon delivery to the target cell by the LNP, the long-term repressor domain fusion protein is expressed from the mRNA and can complex with the gRNA as an RNP capable of binding the *PCSK9* target nucleic acid. In other cases, the mRNA encoding the long-term repressor domain fusion protein and the gRNA may be formulated in separate nanoparticles and delivered either separately or as a mixture.

[0209] In some embodiments, the mRNA compositions of the disclosure have been modified to result in one or more improved characteristics relative to an unmodified mRNA encoding the same long-term repressor protein, and therefore may have a positive impact on the efficacy of mRNA-based delivery. Exemplary improved characteristics of the mRNA described herein include, but are not limited to improved expression upon delivery to a cell, reduced immunogenicity, increased stability, and enhanced manufacturability compared to unmodified mRNA. In some cases, the modifications to the mRNA result in an improved characteristic of at least about 1.1 to about 100,000-fold improved relative to the unmodified mRNA. In some embodiments, an improved characteristic of the modified mRNA is at least about 1.1 to about 10,000-fold improved, at least about 1.1 to about 1,000-fold improved, at least about 1.1 to about 500-fold improved, at least about 1.1 to about 400-fold improved, at least about 1.1 to about 300-fold improved, at least about 1.1 to about 200-fold improved, at least about 1.1 to about 100-fold improved, at least about 1.1 to about 50-fold improved, at least about 1.1 to about 40-fold improved, at least about 1.1 to about 30-fold improved, at least about 1.1 to about 20-fold improved, at least about 1.1 to about 10-fold improved, at least about 1.1 to about 9-fold improved, at least about 1.1 to about 8-fold improved, at least about 1.1 to about 7-fold improved, at least about 1.1 to about 6-fold improved, at least about 1.1 to about 5-fold improved, at least about 1.1 to about 4-fold improved, at least about 1.1 to about 3-fold improved, at least about 1.1 to about 2-fold improved, at least about 1.1 to about 1.5-fold improved, at least about 1.5 to about 3-fold improved, at least about 1.5 to about 4-fold improved, at least about 1.5 to about 5-fold improved, at least about 1.5 to about 10-fold improved, at least about 5 to about 10-fold improved, at least about 10 to about 20-fold

improved, at least 10 to about 30-fold improved, at least 10 to about 50-fold improved or at least 10 to about 100-fold improved relative to the unmodified mRNA. In some embodiments, an improved characteristic of the modified mRNA is at least about 10 to about 1000-fold improved relative to the unmodified mRNA.

[0210] Optimization of coding sequences and untranslated regions (UTRs) may be useful when delivering an mRNA encoding a protein of interest, as opposed to a DNA template that would be transcribed into an mRNA. DNA templates are long-lived, can replicate, and can produce many RNA transcripts over their lifetimes. For DNA templates, efficiency of transcription and pre-mRNA processing are major determinants of protein expression levels. In contrast, mRNAs generally have a much shorter half-life, on the order of hours, as they are vulnerable to degradation in the cytoplasm, and cannot produce more copies of themselves. As such, mRNA stability and translation efficiency can be key determinants of protein expression levels for mRNA-based delivery, and the specific sequences of UTRs and coding sequences that dictate mRNA stability and translation efficiency can therefore be enhanced to improve the efficacy of mRNA-based delivery.

a. 5' cap

[0211] In some embodiments of the mRNA of the disclosure, the mRNA comprises a 5' cap linked 5' to the 5' UTR of the mRNA sequence of any of the embodiments described herein. In some embodiments, the 5' cap is a 7-methylguanylate cap. In some embodiments, the 5' cap comprises m⁷G(5')ppp(5')mAG. In other embodiments, the 5' cap comprises m⁷G(5')ppp(5')(A,G(5')ppp(5')A or G(5')ppp(5')G. Exemplary caps are known in the art, and described, for example, in WO 2017/053297, the contents of which are incorporated by reference herein.

b. 5' untranslated region (UTR)

[0212] The 5' UTR of an mRNA molecule can be a key determinant of both the stability of the mRNA and how efficiently it is translated into protein. Specifically, the 5' UTR, in conjunction with the 5' cap structure, serves as a binding site and recruitment platform for the translation pre-initiation complex as well as additional regulatory proteins that may positively or negatively affect translation. Structures within the 5' UTR can enhance translation by recruiting initiation factors or other protein or RNA factors, reduce translation by physically blocking ribosome binding and scanning, and contribute to the stability of the mRNA by affecting both hydrolysis and nuclease digestion.

[0213] An exemplary 5' UTR sequence for use in the mRNA of the disclosure is provided in Table 6. Table 6 lists the RNA sequence, and the RNA sequence with N1-methyl-pseudouridine substituted in place of uridine.

Table 6: 5' UTR sequences

SEQ ID NO	Nucleic Acid Sequence*	Description
22787	AAAUAAGAGAGAAAAGAAGAGUAAGAAGAAUAUAAGA	RNA sequence of 5' UTR
3300	AAAmΨAAGAGAGAAAAGAAGAGmΨAAGAAGAAmΨAmΨAAGA	N1-methyl-pseudouridine substituted RNA sequence of 5' UTR

*‘mψ’ = N1-methyl-pseudouridine

[0214] In some embodiments, the 5' UTR comprises the sequence of SEQ ID NO: 22787, or a sequence having at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, or at least about 95% identity thereto. In some embodiments, the 5' UTR comprises the sequence of SEQ ID NO: 22787. In some embodiments, the 5' UTR consists of the sequence of SEQ ID NO: 22787. In some embodiments, the 5' UTR comprises the sequence of SEQ ID NO: 3300, or a sequence having at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, or at least about 95%, identity thereto. In some embodiments, the 5' UTR comprises the sequence of SEQ ID NO: 3300. In some embodiments, the 5' UTR consists of the sequence of SEQ ID NO: 3300.

c. 3' UTR

[0215] 3' UTR sequences can have an impact on mRNA stability and translation efficiency and thus may determine both subcellular localization and tissue-specific expression. Factors influencing these properties include microRNA binding sites, AU-rich elements that recruit an array of RNA-binding proteins, Pumilio binding elements, and other binding sites for RNA-binding proteins. While many of these interactions with the 3' UTR are known to negatively impact stability or expression, some can enhance translation. The effects of a 3' UTR sequence can be highly cell-type specific due to differential expression of microRNAs and RNA binding proteins, which provides opportunities for engineering tissue-specific expression into a therapeutic mRNA. In some embodiments, the 3' UTR for use in the mRNA of the disclosure is a mouse 3' UTR. In some embodiments, the 3' UTR is a mouse *HBA* gene 3' UTR of Table 7.

[0216] Exemplary 3' UTR sequences of the disclosure are provided in Table 7. Table 7 lists the RNA sequence, and the RNA sequence with N1-methyl-pseudouridine substituted in place of uridine.

Table 7: 3' UTR sequences

SEQ ID NO	Nucleic Acid Sequence*	Description
3125	GCUGCCUUCUGCGGGGCUUGCCUUCUGGCCAUGCCCUUCUUCUCUCC CUUGCACCGUACCUCUUGGUCUUUGAAUAAAGCCUGAGUAGGAAG	RNA sequence of 3' UTR
3310	GCmΨGCCmΨmΨCmΨGCGGGGmΨmΨGCCmΨmΨCmΨGGCCAmΨGCCmΨ ΨmΨCmΨmΨCmΨCmΨCmΨGCACmΨGmΨACmΨCmΨmΨGGmΨCmΨ ΨmΨmΨGAAmΨAAAGCCmΨGAGmΨAGGAAG	N1-methyl-pseudouridine substituted RNA sequence of 3' UTR

*‘mψ’ = N1-methyl-pseudouridine

[0217] In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO: 3125, or a sequence having at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, or at least about 95% identity thereto. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO: 3125. In some embodiments, the 3' UTR consists of the sequence of SEQ ID NO: 3125. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO: 3310, or a sequence having at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, or at least about 95% identity thereto. In some embodiments, the 3' UTR comprises the sequence of SEQ ID NO: 3310. In some embodiments, the 3' UTR consists of the sequence of SEQ ID NO: 3310.

d. Poly(A) sequence

[0218] Inclusion of a 3' poly(A) tail in an mRNA sequence can contribute to the stability and translation efficiency of the mRNA. Generally, longer poly(A) tails are associated with increased mRNA stability, thereby allowing their translation and promoting high protein expression.

[0219] In some embodiments, an mRNA of the disclosure comprises a poly(A) sequence having at least about 40, at least about 50, at least about 60, at least about 70, at least about 80, at least about 90, at least about 100, at least about 110, at least about 120, at least about 130, at least about 140, at least about 150, at least about 160, at least about 170, at least about 180, at least about 185, or at least about 190 adenine nucleotides. In some embodiments, the poly(A) sequence of an mRNA of the disclosure comprises 80 adenine nucleotides. In some embodiments, the poly(A) sequence comprises the nucleic acid sequence of

AA
 AA (SEQ ID NO: 3057). In some
 embodiments, the poly(A) sequence comprises the nucleic acid sequence of
 AA
 AAAAAAAGCAUGCAA
 AAAAAAAAAAAAAAAAAAAAAAAAAA (SEQ ID NO: 3307).

e. Sequence Modifications to the mRNA

[0220] In some embodiments, an mRNA sequence of the disclosure was modified by codon optimization of the sequence encoding the long-term repressor protein using one or more parameters to enhance expression in the target cell. Non-limiting examples of such parameters include the codon usage in human host cells (e.g., utilizing the codon adaptation index (CAI)), minimizing rare codons, utilizing codon-usage tables derived from biologics intended for use as therapeutics, utilizing an mRNA stability index, or reducing GC content. Methods of codon-optimization, and codon usage in various organisms is known in the art. See, for example, www.genscript.com/tools/codon-frequency-table.

[0221] In some embodiments, provided herein are mRNA sequences for the long-term repressor protein constructs that are codon-optimized for expression in a human cell. In other cases, various naturally-occurring or modified nucleosides may be used to produce the modified mRNA according to the present disclosure. In some embodiments, an mRNA is or comprises natural nucleosides (e.g., adenosine, guanosine, cytidine, uridine); nucleoside analogs (e.g., 2-aminoadenosine, 2-thiothymidine, inosine, pyrrolo-pyrimidine, 3-methyl adenosine, 5-methylcytidine, C-5 propynyl-cytidine, C-5 propynyl-uridine, 2-aminoadenosine, C5-bromouridine, C5-fluorouridine, C5-iodouridine, C5-propynyl-uridine, C5-propynyl-cytidine, C5-methylcytidine, 2-aminoadenosine, 7-deazaadenosine, 7-deazaguanosine, 8-oxoadenosine, 8-oxoguanosine, O(6)-methylguanine, pseudouridine, (e.g., N-1-methyl-pseudouridine), 2-thiouridine, and 2-thiocytidine); chemically modified bases; biologically modified bases (e.g., methylated bases); intercalated bases; modified sugars (e.g., 2'-fluororibose, ribose, 2'-deoxyribose, arabinose, and hexose); and/or modified phosphate groups (e.g., phosphorothioates and 5'-N-phosphoramidite linkages). In some embodiments, the mRNA comprises one or more nonstandard nucleotide residues. The nonstandard nucleotide residues may include, e.g., 5-methyl-cytidine ("5 mC"), pseudouridine ("ψU"), and/or 2-thio-uridine ("2sU"). In a particular embodiment, one or more or all of the uridine residues of the mRNA of the disclosure are replaced with 1-methyl-pseudouridine. In some

embodiments, all of the uridine residues of the mRNA of the disclosure are replaced with 1-methyl-pseudouridine. See, *e.g.*, U.S. Pat. No. 8,278,036 or WO2011012316, incorporated by reference herein, for a discussion of such residues and their incorporation into mRNA. In some embodiments, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 99%, or 100% of uridine nucleosides of the mRNA sequence are replaced with N1-methylpseudouridine.

[0222] In some embodiments, the modifications to the mRNA by codon optimization and/or substitution of modified nucleotides in the sequence results in an improved characteristic of at least about 1.1 to about 100,000-fold improved relative to the unmodified mRNA. In some embodiments, an improved characteristic of the modified mRNA is at least about 1.1 to about 10,000-fold improved, at least about 1.1 to about 1,000-fold improved, at least about 1.1 to about 500-fold improved, at least about 1.1 to about 400-fold improved, at least about 1.1 to about 300-fold improved, at least about 1.1 to about 200-fold improved, at least about 1.1 to about 100-fold improved, at least about 1.1 to about 50-fold improved, at least about 1.1 to about 40-fold improved, at least about 1.1 to about 30-fold improved, at least about 1.1 to about 20-fold improved, at least about 1.1 to about 10-fold improved, at least about 1.1 to about 9-fold improved, at least about 1.1 to about 8-fold improved, at least about 1.1 to about 7-fold improved, at least about 1.1 to about 6-fold improved, at least about 1.1 to about 5-fold improved, at least about 1.1 to about 4-fold improved, at least about 1.1 to about 3-fold improved, at least about 1.1 to about 2-fold improved, at least about 1.1 to about 1.5-fold improved, at least about 1.5 to about 3-fold improved, at least about 1.5 to about 4-fold improved, at least about 1.5 to about 5-fold improved, at least about 1.5 to about 10-fold improved, at least about 5 to about 10-fold improved, at least about 10 to about 20-fold improved, at least 10 to about 30-fold improved, at least 10 to about 50-fold improved or at least 10 to about 100-fold improved relative to the unmodified mRNA. In some embodiments, an improved characteristic of the modified mRNA is at least about 10 to about 1000-fold improved relative to the unmodified mRNA.

***f.* LTRP mRNA Component Sequences**

[0223] The present disclosure provides mRNA comprising sequences encoding the components utilized in the long-term repressor fusion proteins of the disclosure. In some embodiments, the mRNA comprises sequences encoding the DNA binding proteins, including TALE, ZF, and catalytically dead CRISPR proteins. In some embodiments, the mRNA comprises the sequence of SEQ ID NO: 3274 encoding dCasX 515 (SEQ ID NO: 6),

or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto. In some embodiments, the mRNA comprises the sequence of SEQ ID NO: 3276 encoding dCasX 812 (SEQ ID NO: 29), or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto. In some embodiments, the mRNA comprises the sequence of SEQ ID NO: 3275 encoding dCasX 491 (SEQ ID NO: 4), or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto. In some embodiments, the mRNA comprises the sequence of SEQ ID NO: 22736 encoding dCasX 676 (SEQ ID NO: 28), or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

[0224] In some embodiments, the sequence of the mRNA encoding dCasX 491 (SEQ ID NO: 4) has pseudouridine nucleosides replacing one or more, or all uridines in the sequence (SEQ ID NO: 22729). In some embodiments, the sequence of the mRNA encoding dCasX 515 (SEQ ID NO: 6) has pseudouridine nucleosides replacing one or more, or all uridines in the sequence (SEQ ID NO: 22728). In some embodiments, the sequence of the mRNA encoding dCasX 676 (SEQ ID NO: 28) has pseudouridine nucleosides replacing one or more, or all uridines in the sequence. In some embodiments, the sequence of the mRNA encoding dCasX 812 (SEQ ID NO: 29) has pseudouridine nucleosides replacing one or more, or all uridines in the sequence (SEQ ID NOS: 22730 or 22731).

[0225] In some embodiments, the disclosure provides mRNA sequences encoding the RD1 domains. In some embodiments, the sequence of the mRNA encoding the RD1 comprises a sequence selected from the group consisting of SEQ ID NOS: 3120, 22804, and 3334-6527, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, or at least about 95% identity thereto. In some embodiments, the sequence of the mRNA encoding the RD1 comprises a sequence selected from the group consisting of SEQ ID NOS: 3120, 22804, and 3334-6527. In another embodiment, the sequence of the mRNA encoding the RD1 comprises a sequence selected from the group consisting of SEQ ID NOS: 3334-3342 and 4931-4939, or a sequence having at least about 70%, at least about 80%, at least about

90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In another embodiment, the sequence of the mRNA encoding the RD1 comprises a sequence selected from the group consisting of SEQ ID NOS: 3334-3342 and 4931-4939. In another embodiment, the sequence of the mRNA encoding the RD1 comprises a sequence of SEQ ID NO: 3334 or SEQ ID NO: 4931, or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In another embodiment, the sequence of the mRNA encoding the RD1 comprises a sequence of SEQ ID NO: 3334 or SEQ ID NO: 4931. In another embodiment, the sequence of the mRNA encoding the RD1 comprises a sequence of SEQ ID NO: 3335 or SEQ ID NO: 4932, or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In another embodiment, the sequence of the mRNA encoding the RD1 comprises a sequence of SEQ ID NO: 3335 or SEQ ID NO: 4932. In another embodiment, the sequence of the mRNA encoding the RD1 comprises a sequence of SEQ ID NO: 3336 or SEQ ID NO: 4933, or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In another embodiment, the sequence of the mRNA encoding the RD1 comprises a sequence of SEQ ID NO: 3336 or SEQ ID NO: 4933. In another embodiment, the sequence of the mRNA encoding the RD1 comprises a sequence of SEQ ID NO: 3337 or SEQ ID NO: 4934, or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In another embodiment, the sequence of the mRNA encoding the RD1 comprises a sequence of SEQ ID NO: 3337 or SEQ ID NO: 4934. In another embodiment, the sequence of the mRNA encoding the RD1 comprises a sequence of SEQ ID NO: 3338 or SEQ ID NO: 4935, or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In another embodiment, the sequence

of the mRNA encoding the RD1 comprises a sequence of SEQ ID NO: 3338 or SEQ ID NO: 4935. In another embodiment, the sequence of the mRNA encoding the RD1 comprises a sequence of SEQ ID NO: 3339 or SEQ ID NO: 4936, or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In another embodiment, the sequence of the mRNA encoding the RD1 comprises a sequence of SEQ ID NO: 3339 or SEQ ID NO: 4936. In another embodiment, the sequence of the mRNA encoding the RD1 comprises a sequence of SEQ ID NO: 3340 or SEQ ID NO: 4937, or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In another embodiment, the sequence of the mRNA encoding the RD1 comprises a sequence of SEQ ID NO: 3340 or SEQ ID NO: 4937. In another embodiment, the sequence of the mRNA encoding the RD1 comprises a sequence of SEQ ID NO: 3341 or SEQ ID NO: 4938, or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In another embodiment, the sequence of the mRNA encoding the RD1 comprises a sequence of SEQ ID NO: 3341 or SEQ ID NO: 4938.

[0226] In some embodiments, the mRNA comprises a sequence encoding the second repressor domain. In some embodiments, the second repressor domain comprises DNMT3A. In some embodiments, the sequence of the mRNA encoding the DNMT3A comprises a sequence of SEQ ID NO: 3128 or SEQ ID NO: 3331, or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% identity thereto. In some embodiments, the sequence of the mRNA encoding the DNMT3A comprises a sequence of SEQ ID NO: 3128 or SEQ ID NO: 3331.

[0227] In some embodiments, the mRNA comprises a sequence encoding the third repressor domain. In some embodiments, the third repressor domain comprises DNMT3L. In some embodiments, the sequence of the mRNA encoding the DNMT3L comprises a sequence SEQ ID NO: 3119 or SEQ ID NO: 3332, or a sequence having at least about 70%,

at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% identity thereto. In some embodiments, the sequence of the mRNA encoding the DNMT3L comprises a sequence SEQ ID NO: 3119 or SEQ ID NO: 3332.

[0228] In some embodiments, the mRNA comprises a sequence encoding the fourth repressor domain. In some embodiments, the fourth repressor domain comprises the ADD. In some embodiments, the sequence of the mRNA encoding the ADD comprises a sequence of SEQ ID NO: 3296 or SEQ ID NO: 22726, or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% identity thereto. In some embodiments, the sequence of the mRNA encoding the ADD comprises a sequence of SEQ ID NO: 3296 or SEQ ID NO: 22726.

[0229] In some embodiments, the mRNA of the disclosure comprises a Kozak sequence. In some embodiments, the Kozak sequence comprises a GCCACCAUGG (SEQ ID NO: 22805). In some embodiments, an mRNA of the disclosure comprises a Kozak sequence and two bases upstream of the NLS (to produce the methionine and alanine upstream of the NLS). In some embodiments, the mRNA comprises the sequence GCCACCAUGGCC (SEQ ID NO: 22806) between the 5' UTR and the sequence encoding the NLS.

[0230] In another embodiment, the mRNA comprises an NLS. In some embodiments, the sequence encoding the NLS comprises a sequence selected from the group consisting of SEQ ID NO: 3291 and SEQ ID NO: 22807, or a sequence having at least about 70%, at least about 80%, or at least about 90% identity thereto. In another embodiment, the mRNA comprises a sequence selected from the group consisting of SEQ ID NO: 3291 and SEQ ID NO: 22807. In another embodiment, for examples those embodiments wherein the long-term repressor protein comprises more than one NLS, the mRNA comprises two or more sequences independently selected from the group consisting of SEQ ID NO: 3291 and SEQ ID NO: 22807.

[0231] In some embodiments, the mRNA comprises sequences encoding the linkers of the long-term repressor fusion protein. In some embodiments, the sequence encoding the linker comprises a sequence selected from the group consisting of SEQ ID NOS: 3312-3327 and 22808-22828, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95% identity

thereto. In some embodiments, the sequence encoding the linker comprises a sequence selected from the group consisting of SEQ ID NOS: 3312-3327 and 22808-22828. In some embodiments, for example those embodiments where the long-term repressor fusion protein comprises more than one linker, the sequences encoding the linkers are independently selected from the group consisting of SEQ ID NOS: 3312-3327 and 22808-22828, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95% identity thereto.

[0232] In some embodiments, the mRNA comprises a sequence encoding the long-term repressor fusion protein of the configurations provided herein. In some embodiments, the mRNA encodes a long-term repressor fusion protein of configuration 1, and the mRNA comprises a sequence is selected from the group consisting of SEQ ID NOS: 6529-8134, 9741-11347, 14628-16233, and 17841-19446, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% identity thereto. In some embodiments, the mRNA comprises a sequence is selected from the group consisting of SEQ ID NOS: 6529-8134, 9741-11347, 14628-16233, and 17841-19446.

[0233] In some embodiments, the mRNA comprises a sequence encoding the long-term repressor fusion protein of configuration 1, and comprises a sequence encoding the RD1 comprising a sequence of SEQ ID NO: 130. In some embodiments, the mRNA comprises a sequence of SEQ ID NOS: 6531, 9744, 14630, or 17843, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the mRNA comprises a sequence of SEQ ID NOS: 6531, 9744, 14630, or 17843. In some embodiments, the mRNA comprises a sequence encoding the long-term repressor fusion protein of configuration 1, and comprises a sequence encoding the RD1 of SEQ ID NO: 131. In some embodiments, the mRNA comprises a sequence of SEQ ID NOS: 6529, 9742, 14628, or 17841, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the mRNA comprises a sequence of SEQ ID NOS: 6529, 9742, 14628, or 17841. In some embodiments, the mRNA comprises a sequence encoding the long-term repressor fusion protein of

configuration 1, and the mRNA comprises a sequence encoding the RD1 of SEQ ID NO: 132. In some embodiments, the mRNA comprises a sequence of SEQ ID NOS: 6535, 9748, 14634, or 17847, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the mRNA comprises a sequence of SEQ ID NOS: 6535, 9748, 14634, or 17847. In some embodiments, the mRNA comprises a sequence encoding the long-term repressor fusion protein of configuration 1, and comprises a sequence encoding the RD1 of SEQ ID NO: 133. In some embodiments, the mRNA comprises a sequence of SEQ ID NOS: 6536, 9749, 14635, or 17848, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the mRNA comprises a sequence of SEQ ID NOS: 6536, 9749, 14635, or 17848. In some embodiments, the mRNA comprises a sequence encoding the long-term repressor fusion protein of configuration 1, and the mRNA comprises a sequence encoding the RD1 of SEQ ID NO: 134. In some embodiments, the mRNA comprises a sequence of SEQ ID NOS: 6534, 9747, 14633, or 17846, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the mRNA comprises a sequence of SEQ ID NOS: 6534, 9747, 14633, or 17846. In some embodiments, the mRNA comprises a sequence encoding the long-term repressor fusion protein of configuration 1, and comprises a sequence encoding the RD1 of SEQ ID NO: 135. In some embodiments, the mRNA comprises a sequence of SEQ ID NOS: 6530, 9743, 14629, or 17842, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the mRNA comprises a sequence of SEQ ID NOS: 6530, 9743, 14629, or 17842. In some embodiments, the mRNA comprises a sequence encoding the long-term repressor fusion protein of configuration 1, and comprises a sequence encoding the RD1 of SEQ ID NO: 136. In some embodiments, the mRNA comprises a sequence of SEQ ID NOS: 6537, 9750, 14636, or 17849, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about

95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the mRNA comprises a sequence of SEQ ID NOS: 6537, 9750, 14636, or 17849. In some embodiments, the mRNA comprises a sequence encoding the long-term repressor fusion protein of configuration 1, and comprises a sequence encoding the RD1 of SEQ ID NO: 137. In some embodiments, the mRNA comprises a sequence of SEQ ID NOS: 6533, 9746, 14632, or 17845, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the mRNA comprises a sequence of SEQ ID NOS: 6533, 9746, 14632, or 17845. In some embodiments, the mRNA comprises a sequence encoding the long-term repressor fusion protein of configuration 1, and comprises a sequence encoding the RD1 of SEQ ID NO: 138. In some embodiments, the mRNA comprises a sequence of SEQ ID NOS: 6532, 9745, 14631, or 17844, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the mRNA comprises a sequence of SEQ ID NOS: 6532, 9745, 14631, or 17844.

[0234] The disclosure provides an mRNA encoding a long-term repressor fusion protein of configuration 5. In some embodiments, the mRNA comprises a sequence is selected from the group consisting of SEQ ID NOS: 8135-9740, 11348-12953, 16234-17839, and 19447-21052, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the mRNA comprises a sequence is selected from the group consisting of SEQ ID NOS: 8135-9740, 11348-12953, 16234-17839, and 19447-21052.

[0235] The disclosure provides an mRNA encoding a long-term repressor fusion protein of configuration 6a, wherein the encoded RD1 are identical. In some embodiments, the mRNA comprises a sequence is selected from the group consisting of SEQ ID NOS: 12954-14553, and 21053-22652, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the mRNA comprises a sequence is selected from the group consisting of SEQ ID NOS: 12954-14553 and 21053-22652.

[0236] The disclosure provides an mRNA encoding a long-term repressor fusion protein of configuration 6b, wherein the encoded RD1 are different. In some embodiments, the mRNA comprises a e sequence is selected from the group consisting of SEQ ID NOS: 14554-14626 and 22653-22725, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the mRNA comprises a e sequence is selected from the group consisting of SEQ ID NOS: 14554-14626 and 22653-22725.

[0237] In some embodiments of the mRNA encoding a long-term repressor fusion protein of the LTRP:gRNA systems of the disclosure, upon delivery of the system to a targeted cell, the mRNA is expressed and the long-term repressor fusion protein is capable of complexing with a gRNA and binding the target DNA, resulting in repression or silencing of transcription of the *PCSK9* gene in the cell.

VI. Guide Nucleic Acids of the Systems

[0238] In another aspect, the disclosure relates to specifically designed guide ribonucleic acids (gRNA) comprising a scaffold and a linked targeting sequence complementary to (and are therefore able to hybridize with) a target nucleic acid sequence of a *PCSK9* gene. The gRNA described herein can be used with the long-term repressor proteins, and systems comprising same, to repress transcription of the *PCSK9* target nucleic acid in a eukaryotic cell. As used herein, the term “gRNA” covers naturally-occurring molecules and gRNA variants, including chimeric gRNA variants comprising domains from different gRNA. gRNAs of the disclosure comprise a scaffold and a targeting sequence complementary to a target nucleic acid of a cell linked to the 3' end of the scaffold.

[0239] In some embodiments, the systems comprising an mRNA encoding a long-term repressor fusion protein comprise a dCasX protein and one or more gRNA, upon expression of the dCasX protein in a transfected cell, form a ribonucleoprotein (RNP) complex with the gRNA, which can target and bind specific locations in the target nucleic acid sequence of the cell. The gRNA provides target specificity to the complex by including a targeting sequence (or “spacer”) having a nucleotide sequence that is complementary to a sequence of the target nucleic acid sequence while the long-term repressor fusion protein of the system provides the site-specific activity such as the binding and repression of the target sequence that is guided to a target site (*e.g.*, stabilized at a target site) within a target nucleic acid sequence by virtue of its association with the gRNA.

[0240] Embodiments of gRNAs and formulations of mRNAs and gRNAs for use in the repression and/or epigenetic modification of PCSK9 target nucleic acids are described herein, below.

a. Reference gRNA and gRNA variants

[0241] As used herein, a “reference gRNA” refers to a CRISPR guide ribonucleic acid comprising a wild-type sequence of a naturally-occurring gRNA. In some embodiments, a gRNA scaffold of the disclosure may be subjected to one or more mutagenesis methods, such as the mutagenesis methods described in WO2023235818A2, WO2022120095A1 and WO2020247882A1, incorporated by reference herein, which may include Deep Mutational Evolution (DME), deep mutational scanning (DMS), error prone PCR, cassette mutagenesis, random mutagenesis, staggered extension PCR, gene shuffling, domain swapping, or chemical modification to generate one or more gRNA variants with enhanced or varied properties relative to the gRNA scaffold that was modified. The activity of the gRNA scaffold from which a gRNA variant was derived may be used as a benchmark against which the activity of the gRNA variant is compared, thereby measuring improvements in function or other characteristics of the gRNA scaffold.

[0242] Table 8 provides the sequences of reference gRNA tracr and scaffold sequences. In some embodiments, the disclosure provides gRNA variants wherein the gRNA has a scaffold comprising a sequence having one or more nucleotide modifications relative to a reference gRNA sequence of any one of SEQ ID NOS: 1731-1743 of Table 8.

Table 8: Reference gRNA tracr and scaffold sequences

SEQ ID NO.	Nucleotide Sequence
1731	ACAUCUGGCGGUUUUAUCCAUUACUUUGGAGCCAGUCCCAGCGACUAUGUCGUAUGGACGAAG CGCUUAUUUAUCGGAGAGAAACCGAUAAGUAAAACGCAUCAAAG
1732	UACUGGCGCUUUUAUCUCAUUAUACUUUGAGAGCCAUACCAGCGACUAUGUCGUAUGGGUAAAGC GCUUAUUUAUCGGAGAGAAUCCGAUAAUAAGAAGCAUCAAAG
1733	ACAUCUGGCGGUUUUAUCCAUUACUUUGGAGCCAGUCCCAGCGACUAUGUCGUAUGGACGAAG CGCUUAUUUAUCGGAGA
1734	ACAUCUGGCGGUUUUAUCCAUUACUUUGGAGCCAGUCCCAGCGACUAUGUCGUAUGGACGAAG CGCUUAUUUAUCGG
1735	UACUGGCGCUUUUAUCUCAUUAUACUUUGAGAGCCAUACCAGCGACUAUGUCGUAUGGGUAAAGC GCUUAUUUAUCGGAGA
1736	UACUGGCGCUUUUAUCUCAUUAUACUUUGAGAGCCAUACCAGCGACUAUGUCGUAUGGGUAAAGC GCUUAUUUAUCGG
1737	GUUUACACACUCCCUCAUAGGGU
1738	GUUUACACACUCCCUCAUAGGGU
1739	UUUUACAUAACCCCUCAUAGGGAU

SEQ ID NO.	Nucleotide Sequence
1740	GUUUACACACUCCUCUCAUGGGGG
1741	CCAGCGACUAUGUCGUAUGG
1742	GCGCUUAUUUAUCGGAGAGAAAUCCGAUAAUAAGAAGC
1743	GGCGCUUUUAUCUCAUUACUUUGAGAGCCAUCACCAGCGACUAUGUCGUAUGGGUAAAGCGCUU AUUUUAUCGGA

b. gRNA Domains and their Function

[0243] The gRNAs of the systems of the disclosure comprise two segments: a targeting sequence and a protein-binding segment. The targeting segment of a gRNA includes a nucleotide sequence (referred to interchangeably as a spacer, a targeter, or a targeting sequence) that is complementary to (and therefore hybridizes with) a specific sequence (a target site) within the target nucleic acid sequence (*e.g.*, a strand of a double stranded target DNA, a target ssRNA, a target ssDNA, etc.), described more fully below. The targeting sequence of a gRNA is capable of binding to a target nucleic acid sequence, including, in the context of the present disclosure, a coding sequence, a complement of a coding sequence, a non-coding sequence, and to accessory elements. The protein-binding segment (or “activator” or “protein-binding sequence”) of the gRNA interacts with (*e.g.*, binds to) a dCasX protein as a complex, forming an RNP (described more fully, below). As used herein, “scaffold” refers to all parts to the guide with the exception of the targeting sequence, which is comprised of several regions, described more fully, below. The properties and characteristics of CasX gRNA, both wild-type and variants, are described in WO2020247882A1, US20220220508A1, and WO2022120095A1, incorporated by reference herein.

[0244] In the case of a reference gRNA, the gRNA occurs naturally as a dual guide RNA (dgRNA), wherein the targeter and the activator portions each have a duplex-forming segment that have complementarity with one another and hybridize to one another to form a double stranded duplex (dsRNA duplex for a gRNA). The term “targeter” or “targeter RNA” is used herein to refer to a crRNA-like molecule (crRNA: “CRISPR RNA”) of a CasX dual guide RNA (and therefore of a CasX single guide RNA when the “activator” and the “targeter” are linked together, *e.g.*, by intervening nucleotides). The crRNA has a 5' region that anneals with the tracrRNA followed by the nucleotides of the targeting sequence. In the case of the gRNA for use in the systems of the disclosure, the scaffolds are designed such that the activator and targeter portions are covalently linked to one another (rather than hybridizing to one another) and comprise a single molecule, and can be referred to as a

“single-molecule gRNA,” “single guide RNA,” a “single-molecule guide RNA,” a “one-molecule guide RNA”, or a “sgRNA”. The gRNA variants of the disclosure for use in the systems are all single molecule versions.

[0245] Collectively, the assembled gRNAs of the disclosure comprise distinct structured regions, or domains: the RNA triplex, the scaffold stem loop, the extended stem loop, the pseudoknot, and the targeting sequence that, in the embodiments of the disclosure is complementary to a sequence of a target nucleic acid and is located on the 3' end of the gRNA. The RNA triplex, the scaffold stem loop, the pseudoknot and the extended stem loop, together with the unstructured triplex loop that bridges portions of the triplex, together, are referred to as the “scaffold” of the gRNA. In some cases, the scaffold stem further comprises a bubble. In other cases, the scaffold further comprises a triplex loop region. In still other cases, the scaffold further comprises a 5' unstructured region. In some embodiments, the gRNA scaffolds of the disclosure for use in the LTRP:gRNA systems comprise a scaffold stem loop having the sequence of CCAGCGACUAUGUCGUAGUGG (SEQ ID NO: 1822), or a sequence with at least 1, 2, 3, 4 or 5 mismatches thereto.

[0246] Each of the structured domains contribute to establish the global RNA fold of the guide and retain functionality of the guide; particularly the ability to properly complex with the dCasX protein. For example, the guide scaffold stem interacts with the helical I domain of dCasX protein, while residues within the triplex, triplex loop, and pseudoknot stem interact with the OBD of the dCasX protein. Together, these interactions confer the ability of the guide to bind and form an RNP with the LTRP that retains stability, while the spacer (or targeting sequence) directs and defines the specificity of the RNP for binding a specific sequence of DNA.

[0247] Site-specific binding of a target nucleic acid sequence (*e.g.*, genomic DNA) by the dCasX protein of the LTRP can occur at one or more locations (*e.g.*, a sequence of a *PCSK9* target nucleic acid) determined by base-pairing complementarity between the targeting sequence of the gRNA and the target nucleic acid sequence. Thus, for example, the gRNA of the disclosure have sequences complementarity to and therefore can hybridize with the target nucleic acid that is adjacent to a sequence complementary to a TC protospacer adjacent motif (PAM) motif or a PAM sequence, such as ATC, CTC, GTC, or TTC. Because the targeting sequence of a guide sequence hybridizes with a sequence of a target nucleic acid sequence, a targeting sequence can be modified by a user to hybridize with a specific target nucleic acid sequence, so long as the location of the PAM sequence is considered. In some embodiments,

for design of a targeting sequence, the target nucleic acid comprises a PAM sequence located 5' of the targeting sequence with at least a single nucleotide separating the PAM from the first nucleotide of the target nucleic acid complementary to that of the targeting sequence. This feature distinguishes the systems described herein from c Cas9 systems, and results in the ability of the systems of the disclosure to modify different locations in the DNA sequence when compared to Cas9 systems. In some embodiments, the PAM is located on the non-targeted strand of the target region, *i.e.* the strand that is complementary to the target nucleic acid. In some embodiments, the targeting sequence of the gRNA is complementary to a *PCSK9* target nucleic acid sequence one nucleotide from an ATC PAM sequence. In some embodiments, the targeting sequence of the gRNA is complementary to a *PCSK9* target nucleic acid sequence one nucleotide from an CTC PAM sequence. In some embodiments, the targeting sequence of the gRNA is complementary to a *PCSK9* target nucleic acid sequence one nucleotide from an GTC PAM sequence. In some embodiments, the targeting sequence of the gRNA is complementary to a *PCSK9* target nucleic acid sequence one nucleotide from an TTC PAM sequence. By selection of the targeting sequences of the gRNA, defined regions of the target nucleic acid sequence or sequences bracketing a particular location within the target nucleic acid can be repressed using the LTRP:gRNA systems described herein. In some embodiments, the targeting sequence of the gRNA has between 15 and 22 consecutive nucleotides. In some embodiments, the targeting sequence has 15, 16, 17, 18, 19, 20, 21 or 22 consecutive nucleotides. In some embodiments, the targeting sequence consists of 22 consecutive nucleotides. In some embodiments, the targeting sequence consists of 21 consecutive nucleotides. In some embodiments, the targeting sequence consists of 20 consecutive nucleotides. In some embodiments, the targeting sequence consists of 19 consecutive nucleotides. In some embodiments, the targeting sequence consists of 18 consecutive nucleotides. In some embodiments, the targeting sequence consists of 17 consecutive nucleotides. In some embodiments, the targeting sequence consists of 16 consecutive nucleotides. In some embodiments, the targeting sequence consists of 15 consecutive nucleotides. By selection of the targeting sequences of the gRNA, defined regions of the *PCSK9* target nucleic acid sequence can be repressed and/or epigenetically modified using the LTRP:gRNA systems described herein.

[0248] The gene repressor systems of the present disclosure can be designed to target any region of, or proximal to, a *PCSK9* gene or region of a *PCSK9* gene for which repression of transcription is sought. When the entirety of the gene is to be repressed, designing a guide

with a targeting sequence complementary to a sequence encompassing or proximal to the transcription start site (TSS) is contemplated by the disclosure. The TSS selection occurs at different positions within the promoter region, depending on promoter sequence and initiating-substrate concentration. The core promoter serves as a binding platform for the transcription machinery, which comprises Pol II and its associated general transcription factors (GTFs) (Haberle, V. et al. Eukaryotic core promoters and the functional basis of transcription initiation (Nat Rev Mol Cell Biol. 19(10):621 (2018)). Variability in TSS selection has been proposed to involve DNA ‘scrunching’ and ‘anti-scrunching,’ the hallmarks of which are: (i) forward and reverse movement of the RNA polymerase leading edge, but not trailing edge, relative to DNA, and (ii) expansion and contraction of the transcription bubble. In some embodiments, the target nucleic acid sequence bound by an RNP of the LTRP:gRNA system is within 1.5 kilobase (kb) of a transcription start site (TSS) in the *PCSK9* gene. In some embodiments, the target nucleic acid sequence bound by an RNP of the system is within 20 basepairs (bp), 50 bp, 100 bp, 150 bp, 200 bp, 250 bp, 500 bps, 1 kb, or 1.5 kb upstream of a TSS of the *PCSK9* gene. In some embodiments, the target nucleic acid sequence bound by an RNP of the system is within 20 bp, 50 bp, 100 bp, 150 bp, 200 bp, 250 bp, 500 bp, 1 kb, or 1.5 kb downstream of a TSS of the *PCSK9* gene. In some embodiments, the target nucleic acid sequence bound by an RNP of the system is within 1.5 kb upstream to 1.5 downstream, 1kb upstream to 1 kb downstream, 500 bp upstream to 500 bp downstream, or 300 bp upstream to 300 bp downstream, or 100 bp upstream to 100 bp downstream of a TSS of the *PCSK9* gene. In some embodiments, the target nucleic acid sequence bound by an RNP of the system is within 20 bp, 50 bp, 100 bp, 150 bp, 200 bp, 250 bp, 500 bp, 1 kb, or 1.5 kb of an enhancer of the *PCSK9* gene. In some embodiments, the target nucleic acid sequence bound by an RNP of the system of the disclosure is within 1 kb 3' to a 5' untranslated region of the *PCSK9* gene. In other embodiments, the target nucleic acid sequence bound by an RNP of the system is within the open reading frame of the *PCSK9* gene, inclusive of introns (if any). In some embodiments, the targeting sequence of a gRNA of the system of the disclosure is complementary to a sequence of an exon of the *PCSK9* gene. In a particular embodiment, the targeting sequence of a gRNA of the system of the disclosure is complementary to a sequence of exon 1 of the *PCSK9* gene. In other embodiments, the targeting sequence of a gRNA of the system of the disclosure is complementary to a sequence of an intron of the *PCSK9* gene. In other embodiments, the targeting sequence of the gRNA of the system of the disclosure is complementary to a

sequence of an intron-exon junction of the *PCSK9* gene. In other embodiments, the targeting sequence of the gRNA of the system of the disclosure is complementary to a sequence of a regulatory element of the *PCSK9* gene. In other embodiments, the targeting sequence of the gRNA of the system of the disclosure is complementary to a sequence of an intergenic region of the *PCSK9* gene. In other embodiments, the targeting sequence of a gRNA of the system of the disclosure is complementary to a sequence of a junction of the exon, an intron, and/or a regulatory element of the *PCSK9* gene. In those cases where the targeting sequence is complementary to a sequence of a regulatory element, such regulatory elements include, but are not limited to promoter regions, enhancer regions, intergenic regions, 5' untranslated regions (5' UTR), 3' untranslated regions (3' UTR), conserved elements, and regions comprising cis-regulatory elements. The promoter region is intended to encompass nucleotides within 5 kb of the initiation point of the encoding sequence or, in the case of gene enhancer elements or conserved elements, can be thousands of bp, hundreds of thousands of bp, or even millions of bp away from the encoding sequence of the *PCSK9* gene. In the foregoing, the targets are those in which the encoding *PCSK9* gene of the target is intended to be repressed such that the *PCSK9* gene product is not expressed or is expressed at a lower level in a cell. In some embodiments, upon binding of the RNP of the system of the disclosure to the binding location of the target nucleic acid, the system is capable of repressing transcription of the *PCSK9* gene 5' to the binding location of the RNP. In other embodiments, upon binding of the RNP of the system to the binding location of the target nucleic acid, the system is capable of repressing transcription of the *PCSK9* gene 3' to the binding location of the RNP.

[0249] In some embodiments, the target nucleic acid comprises a PAM sequence located 5' of the targeting sequence with at least a single nucleotide separating the PAM from the first nucleotide of the targeting sequence. In some embodiments, the PAM is located on the non-targeted strand of the target region, *i.e.* the strand that is complementary to the target nucleic acid. , and are shown below as Table 9, representing targeting sequences for *PCSK9* target nucleic acid for linkage to the gRNA scaffolds of the disclosure; *e.g.*, gRNA 174 (SEQ ID NO: 1744), 235 (SEQ ID NO: 1745), 316 (SEQ ID NO: 1746), or chemically-modified versions thereof. In some embodiments, the PAM sequence is TTC.

[0250] In some embodiments, the targeting sequence of the gRNA comprises a sequence selected from the group consisting of SEQ ID NOS: 1824-2544, 2672, 2675, 2694, and 2714, as set forth in Table 9. In some embodiments, the targeting sequence of the gRNA comprises

a sequence selected from the group consisting of SEQ ID NOS: 1824-1880, 1883, 1884, 1888, 1889, 2672, 2675, 2694, and 2714. In some embodiments, the targeting sequence of the gRNA comprises a sequence selected from the group consisting of SEQ ID NOS: 1834, 1849, 1853, 1855-1858, 1860, 1862, 1863, 1867, 1869, 1870, 1872, 1874, and 1875. In some embodiments, the targeting sequence of the gRNA comprises a sequence selected from the group consisting of SEQ ID NOS: 1855, 1867 and 1869. In a particular embodiment, the targeting sequence of the gRNA consists of a sequence of SEQ ID NO: 1855. In another particular embodiment, the targeting sequence of the gRNA consists of a sequence of SEQ ID NO: 1867. In another particular embodiment, the targeting sequence of the gRNA consists of a sequence of SEQ ID NO: 1869. In some embodiments, the foregoing targeting sequences are linked to the gRNA scaffolds of gRNA 174 (SEQ ID NO: 1744), 235 (SEQ ID NO: 1745), or 316 (SEQ ID NO: 1746), or chemically-modified versions thereof.

Table 9: Targeting Sequences Specific to PCSK9 using long-term repressor fusion proteins

Targeting Sequence	RNA sequence	SEQ ID	PAM Sequence
TG-06-342	AAUUACAGGCAACAGGAAGG	1824	TTC
TG-06-343	CCCAUGUAAGAGAGGAAGU	1825	TTC
TG-06-344	CAGUUUCUGCCUCGCCGCGG	1826	TTC
TG-06-345	GCCUCGCCGCGGCACAGGUG	1827	TTC
TG-06-346	CCCACCUGUGCCGCGGCGAG	1828	TTC
TG-06-347	CUCCUUCACCCACCUGUGCC	1829	TTC
TG-06-348	AGGCAUUCACUCCUUCACCC	1830	TTC
TG-06-349	CUGUGCCUGGUGCAGUUCCC	1831	TTC
TG-06-350	GUGUCAUAAAGAAAUGCCU	1832	TTC
TG-06-351	UUAUGACACAGAACUCAUGC	1833	TTC
TG-06-001	GAGGAGGACGGCCUGGCCGA	1834	TTC
TG-06-002	ACCGCUGCGCCAAGGUGCGG	1835	TTC
TG-06-004	GCCAGGCCGUCCUCCUCGGA	1836	TTC
TG-06-005	GUGCUCGGGUGCUUCGGCCA	1837	TTC
TG-06-117	ACUUUGUUUGCAAAGACCUC	1838	TTC
TG-06-118	GAGUGAAAUGGCCUGCUCUG	1839	TTC
TG-06-119	GAGCAGGCCAUUUCACUCGG	1840	TTC
TG-06-120	CUCGGAUCUGCUGUGCAUC	1841	TTC
TG-06-121	GGAAGGGCUGUCGAUACUGG	1842	TTC
TG-06-123	UCCAGUAUCGACAGCCCUU	1843	TTC
TG-06-124	CAGUAUCGACAGCCCUUCCA	1844	TTC

TG-06-125	AGAAAGAGCAAGCCUCAUGU	1845	TTC
TG-06-128	AGAAAUCAACUGGACAAGCA	1846	TTC
TG-06-131	UGAACAUUGGUGUGUAAAAGG	1847	TTC
TG-06-132	AGAAGAUUCAAUUUGCAAAG	1848	TTC
TG-06-133	AUGGUAGGCACAAGCUCAGC	1849	TTC
TG-06-134	GAAUUCUAUGGUAGGCACAA	1850	TTC
TG-06-135	GGAAAGCUGAGCUUGUGCCU	1851	TTC
TG-06-138	AGGGAUUUAUACUACAAAGA	1852	TTC
TG-06-139	AGGAGCAGCUAGUUGGUAAG	1853	TTC
TG-06-140	AAACUUAGCCUGGACCCCCU	1854	TTC
TG-06-141	ACUGGCCUUAACCUGGCAGC	1855	TTC
TG-06-142	UUCCACUGGCCUUAACCUGG	1856	TTC
TG-06-143	GAAUCAAUCCUACUGUGGAC	1857	TTC
TG-06-144	GUGGGCAGCGAGGAGUCCAC	1858	TTC
TG-06-145	UGGGUCCACCUUGUCUCCUG	1859	TTC
TG-06-146	GAAGUCUCACUGGUCAGCAG	1860	TTC
TG-06-147	GUGUUUCCUGGGUCCACCUU	1861	TTC
TG-06-149	AGCCCAGUUAGGAUUUGGGA	1862	TTC
TG-06-150	UCCUCUGCGCGUAAUCUGA	1863	TTC
TG-06-151	CUCUGCGCGUAAUCUGACGC	1864	TTC
TG-06-152	GCCUCGCCCUCGCCAAACAG	1865	TTC
TG-06-153	GUUAAUGUUUAAUCAGAUAG	1866	TTC
TG-06-154	AGGGUGUGGGUGCUUGACGC	1867	TTC
TG-06-155	GCAGCGACGUCGAGGCGCUC	1868	TTC
TG-06-157	GGGUCUGAGCCUGGAGGAGU	1869	TTC
TG-06-158	GGAGCAGGGCGCGUGAAGGG	1870	TTC
TG-06-159	GCGCGCCCCUUCACGCGCCC	1871	TTC
TG-06-160	CGCGCCCUGCUCCUGAACUU	1872	TTC
TG-06-161	GCUCCUGCACAGUCCUCCCC	1873	TTC
TG-06-167	CACUGAAUAGCGCAGCCGCA	1874	TTC
TG-06-168	GUGGGAAGGUUCGCGGGGUU	1875	TTC
TG-06-169	CGGGGUUGGGAGACCCGGAG	1876	TTC
TG-06-170	UCGGCCUCCGGGUCUCCCAA	1877	TTC
TG-06-171	CAGUACGUUCCAGGCAUUCA	1878	TTC
TG-06-449	GCUGAAACAGAUGGAAUACU	1879	TTC
TG-06-249	AAACCAAUCGGAACCCACU	1880	TTC
TG-06-386	UGUUGCCUGUAAUUGGAAUU	1881	TTC
TG-06-122	CCUUUGUUUCUUCCAGUAU	1883	TTC
TG-06-127	UCCUCCUGCCUGGUACACAA	1884	TTC
TG-06-137	GAAUGUACCUAUAUGACGUC	1885	TTC

TG-06-431	CCCCGGCCUCCCAUCCCUAC	1886	TTC
TG-06-433	CUUGGCACGAUCUUGGGGAC	1887	TTC
TG-06-445	GAUUUGGUUUUGGAAAACAUG	1888	TTC
TG-06-446	CUCCAGGCCCUCCACCCUCC	1889	TTC
TG-06-447	CACCCCGCCCCUGUCUCGGG	1890	TTC
TG-06-355	CCAACCUCAGAAACUUGGAG	1891	TTC
TG-06-356	GAGGUUGGAGAAGGUAGCCA	1892	TTC
TG-06-359	AUCUACUCAUUCAGUCUAUG	1893	TTC
TG-06-360	CCUCAUAGACUGAAUGAGUA	1894	TTC
TG-06-361	GUCUAUGAGGGGAAGGCAAU	1895	TTC
TG-06-362	AGCCAUUGCCUCCCCUCAU	1896	TTC
TG-06-363	UCCCUCAA AAAAUGCCACUU	1897	TTC
TG-06-364	CUUCACCCACAAAGCAUCCU	1899	TTC
TG-06-365	UGAGUUAUUCUUUCCCUUCA	1900	TTC
TG-06-366	GGGCAGAGGGGACCAUCCUC	1901	TTC
TG-06-367	UCCUUCAGGGCAGAGGGGAC	1902	TTC
TG-06-368	CCUCCUGCCUUCUCCUUCA	1903	TTC
TG-06-369	GGGGCCUCCCCUCCACCACA	1904	TTC
TG-06-370	GACUUACCCUGAUUGCUGGU	1905	TTC
TG-06-371	CCCACACCCUGGCCAGCAAA	1906	TTC
TG-06-372	CAAAUGGGUUCCCCACACC	1907	TTC
TG-06-373	CCCUCUUCUCAA AUGGGUUC	1908	TTC
TG-06-374	UUUGGGCAACUUAUGGGAGG	1909	TTC
TG-06-375	UCCCUACCAAUUACCCUC	1910	TTC
TG-06-376	CAUCCCAUAUCCUCCACUGC	1911	TTC
TG-06-377	AUCUGUCGCAGAGCCCAGGG	1912	TTC
TG-06-378	GAGGGAGGUGGGGGAGGAAG	1913	TTC
TG-06-379	UCCCCACCUCUCCUAGAAC	1914	TTC
TG-06-380	UGGGCCUGGGGACAUCCAUG	1915	TTC
TG-06-381	UCAUUCUCUGGGUGCACGGU	1916	TTC
TG-06-382	CUGGGUGCACGGUAACGACC	1917	TTC
TG-06-383	CUGCUGCUCUUUUGGGAC	1918	TTC
TG-06-384	UGCUCAGGGGAGGCCUUUG	1919	TTC
TG-06-385	UCAUCAAGGCCUCCCCUGG	1920	TTC
TG-06-450	UCGGCUGAAACAGAUGGAAU	1921	TTC
TG-06-451	AUCUGUUUCAGCCGAAGAAA	1922	TTC
TG-06-452	UUUCUUCGGCUGAAACAGAU	1923	TTC
TG-06-453	GCCGAAGAAAAGAACCAGCU	1924	TTC
TG-06-454	CCCCUGCCCCUUCAGCUGGU	1925	TTC
TG-06-455	CGAGGCCCAUUGGCGUCCUU	1926	TTC

TG-06-456	CUGCCUGAGUCCUAAAGGAC	1927	TTC
TG-06-457	CACAGCCUGAGGGCCAGAAG	1928	TTC
TG-06-458	GGCCCUCAGGCUGUGGGAAG	1929	TTC
TG-06-459	UCCCGGGGCGAGACCACUAG	1930	TTC
TG-06-460	CGGGGCGAGACCACUAGCUU	1931	TTC
TG-06-461	AAGUAUUACCAGCCCAGGAC	1932	TTC
TG-06-462	GUGUCCCCCAGCUUGGAGUC	1933	TTC
TG-06-463	UCUCACUAGCUGUGGUGCUU	1934	TTC
TG-06-464	CUCAUGGUCAACAGAACUUU	1937	TTC
TG-06-465	GUUGACCAUGAGUGAACUUA	1938	TTC
TG-06-466	CCUCUGCCCCAGGAGCUGUG	1939	TTC
TG-06-467	GCCACUCACAGCUGCCUGCC	1940	TTC
TG-06-468	GUCUGAGUGUGCUGGGUGGC	1941	TTC
TG-06-469	UUGUCCCCUGAGCCUUUGAC	1942	TTC
TG-06-470	CUGAGCCUUUGACUUUCUCG	1943	TTC
TG-06-471	CGAGGGAUGUUGUGGGGUUG	1944	TTC
TG-06-472	UAUCCUGGCCACAACCCAC	1945	TTC
TG-06-473	AGUUACCACUGCUCCAAAAC	1946	TTC
TG-06-474	GGAAAUAGUGAGUACCCCAU	1947	TTC
TG-06-475	GGUGGUCAUACAGCCUCUGC	1948	TTC
TG-06-476	ACCAGAGAAGAAACAUCCUC	1949	TTC
TG-06-477	AAACUUCCACCAGAGAAGAA	1950	TTC
TG-06-478	UCUCUGGUGGAAGUUUGGAA	1951	TTC
TG-06-479	CUGGUGGAAGUUUGGAACAG	1952	TTC
TG-06-480	CACAAAUAACCAUUGAAUG	1953	TTC
TG-06-481	AUUCCACAAAUAACCAUU	1954	TTC
TG-06-482	AAAUUCGAUUCCACAAAAU	1955	TTC
TG-06-483	UUCUAAAUUCGAUUCCACA	1956	TTC
TG-06-484	CAUGACUAAAAGAAAAGGAU	1957	TTC
TG-06-485	UUUAGUCAUGAGAAACUGAG	1958	TTC
TG-06-486	AAUGCACCUGGGGUGACCUC	1959	TTC
TG-06-487	GGAAACCCAGUUCUAAUGCA	1960	TTC
TG-06-488	AGAACUGACACUCCACUGCA	1961	TTC
TG-06-489	AAGACCCAUCUGAAAAUGCC	1962	TTC
TG-06-490	GAUGGGUCUUUGAAGCAUUA	1963	TTC
TG-06-491	CAUGACACCAUCAUCGCUGA	1964	TTC
TG-06-492	GCGAUGAUGGUGUCAUGAGA	1965	TTC
TG-06-493	UGGUACCUCCUAAUCCUAGA	1966	TTC
TG-06-494	AGGAUUAGGAGGUACCAUGA	1967	TTC
TG-06-495	CAGCUCUGUAUCUUUGUUCA	1968	TTC

TG-06-496	GACGGUGGAUAAAGAACUGU	1969	TTC
TG-06-497	UUAUCCACCGUCUGAAAUGU	1970	TTC
TG-06-498	CGUCAGGUGCUACCAUCAA	1971	TTC
TG-06-499	UCCAAGCAGACUCUGAUUGC	1972	TTC
TG-06-500	AAGCAGACUCUGAUUGCCCA	1973	TTC
TG-06-501	AAAAUCCAAACAAGUUCUGA	1974	TTC
TG-06-502	ACUCUGACUGGAGUCCAAGA	1975	TTC
TG-06-503	CUGGCUCCCCACAGCCUCAU	1976	TTC
TG-06-504	UGAGGCUGUGGGGAGCCAGU	1977	TTC
TG-06-505	UAAGCAGGUAUGUCUGCCUG	1978	TTC
TG-06-506	CAGGGGUUCCUCCCCCAUGA	1979	TTC
TG-06-507	AGGGAAAUCAUAGCGGGAGA	1980	TTC
TG-06-508	CCCGCUAUGAUUCCCUUGA	1981	TTC
TG-06-509	CUUGAAGGCUGAUUAUUA	1982	TTC
TG-06-510	CUUGCCAUGUGAUUGAUCUC	1983	TTC
TG-06-511	GCCAUCAGUGAAAGCAGCUU	1984	TTC
TG-06-512	UGGAGAGCUUGCAGAAUCAU	1985	TTC
TG-06-513	GCUCAUGAUUCUGCAAGCUC	1986	TTC
TG-06-514	UUUAUAUAGCAACACAAAAG	1987	TTC
TG-06-515	UAAAGAGAGAGAAAAGUCAA	1988	TTC
TG-06-516	CUCUCUUUAAGAAGUGUUGC	1989	TTC
TG-06-517	GCCUGCAUUUGUAGAGCAUC	1990	TTC
TG-06-518	CAGCAGGCCUAUGUCAUAGG	1991	TTC
TG-06-519	GGUCCACGGUUUCCUGGCUC	1993	TTC
TG-06-520	UGGCUCCAAAGCCCAUGAUU	1994	TTC
TG-06-521	CUCAAUUUAUUCUGACUGGG	1995	TTC
TG-06-522	GACUGGGGCAUGGGGGAGGG	1996	TTC
TG-06-523	UCUGUACCUGCACCCAGCUC	1997	TTC
TG-06-524	GGCAAAAUGGGGAACUACGG	1998	TTC
TG-06-525	AAUGUGCCAUUCAGGCAAAA	1999	TTC
TG-06-526	UUAUCAUCUACUUAUCUCUG	2001	TTC
TG-06-527	UCCAUGUCAUCAUGUCCUC	2002	TTC
TG-06-528	AUGUCAUCAUGUCCUCCUU	2003	TTC
TG-06-529	UCCUUGCAUGGGGCCAGGAU	2004	TTC
TG-06-530	GCACCACCACGUAGGUGCCA	2005	TTC
TG-06-531	AUGGCCUUCUCCUGGCUUC	2006	TTC
TG-06-532	CCAGGAAGCCAGGAAGAAGG	2007	TTC
TG-06-533	UCCUGGCUUCCUGGUGAAGA	2008	TTC
TG-06-534	UGGCUUCCUGGUGAAGAUGA	2009	TTC
TG-06-535	UGGUGAAGAUGAGUGGCGAC	2010	TTC

TG-06-536	CAAAAAGGGUGGCUCACCAG	2011	TTC
TG-06-537	UGAUAGGGCUGGGCCACUGC	2012	TTC
TG-06-538	GCAAUGGGCCUACUAAGCAC	2013	TTC
TG-06-539	AUACAGAAAGCCAGUAGUUA	2014	TTC
TG-06-540	GUAUAGAAUCCCCUUAAGC	2015	TTC
TG-06-541	CUUUAAGCCUGGCCAUGCCC	2016	TTC
TG-06-542	AAUGAAGAUAGACGUACCAC	2017	TTC
TG-06-543	GUCUCCUCGUCUUUCAAUG	2018	TTC
TG-06-544	UUUGAAAGACGAGGAGACUG	2019	TTC
TG-06-545	GAGGGGACCACACAGACAGC	2020	TTC
TG-06-546	UGUGCCAAUGCAUCUGCUGC	2021	TTC
TG-06-547	GUAGGCAGCAGAUGCAUUGG	2022	TTC
TG-06-548	UUUGAAGCUGUUGCUAUCCA	2023	TTC
TG-06-549	CUCUGGAUAGCAACAGCUUC	2024	TTC
TG-06-550	UUAGGACAGGGUGGCCCUCC	2025	TTC
TG-06-551	GCCCCUCCACCGAAGAUGUG	2026	TTC
TG-06-552	GCACCACCUUCUAGCCCACC	2029	TTC
TG-06-553	AGCCCACCUCGUUCCUGGC	2030	TTC
TG-06-554	UGGCCUCUAACUUGAUGAGA	2031	TTC
TG-06-555	CACUGAUUCCACAUUGGCA	2032	TTC
TG-06-556	CCACAUGGCAGGCGGUGCUU	2033	TTC
TG-06-557	UAGCCUCCUGCAGACAGUGA	2034	TTC
TG-06-558	GCCACGGUAUGCUGAGCUGA	2035	TTC
TG-06-559	AAGGCAGGGACCCUGUCAAU	2036	TTC
TG-06-560	AUCAGUCUGCUCUACACACG	2037	TTC
TG-06-561	CCUCUGCAGCACACAGUAGG	2038	TTC
TG-06-562	CUUGAGCUUACAGUUCAGGA	2039	TTC
TG-06-563	GGAGGAGAGACUGACCAGUG	2040	TTC
TG-06-564	UCUCUCCCCAAGGCUGAGCA	2041	TTC
TG-06-565	UCGUCUCUCCCCAAGGCUGA	2042	TTC
TG-06-566	GCUUGGCCUACUCUCCUCUG	2043	TTC
TG-06-567	GUUUGCCUUCUGCUUGGCCU	2044	TTC
TG-06-568	AACUGGGCAUCAUACCAGCC	2045	TTC
TG-06-569	AGUCCUGCCCAGCCCCUUAU	2046	TTC
TG-06-570	UCUAUCUGUAAACCAGUUAU	2047	TTC
TG-06-571	UUUAAGGACUAAUAGAGGUC	2048	TTC
TG-06-572	CGACCUCAUUUAGUCCUUAU	2049	TTC
TG-06-573	GUACCUAGCAUCUGCUGAAU	2050	TTC
TG-06-574	GCAGAUGCUGGUACGGAAA	2051	TTC
TG-06-575	GCGAGUUUCCGUACCUAGCA	2052	TTC

TG-06-576	UAACCUGCCCCCACUUCAGC	2053	TTC
TG-06-577	UAUCCCCAGAGGCUUCUUA	2054	TTC
TG-06-578	GACCUAAGUGGUAACAGCCU	2055	TTC
TG-06-579	CUCUCCUUACCCUUCAGACC	2056	TTC
TG-06-580	CUCUGCCCAGUUGGAGCCGG	2057	TTC
TG-06-581	CCGUGCCAUGCGCAGGGAG	2059	TTC
TG-06-582	CCAGCCUGCUCAGUGCAAAC	2060	TTC
TG-06-583	GUUUGUCCCCUAGUCUAUUC	2061	TTC
TG-06-584	CUCCAUGUCCAUCCUGCCCA	2062	TTC
TG-06-585	GAACUUGACAUCUUGUCUGG	2063	TTC
TG-06-586	GAAUCCUUCCUAAAAGGAAU	2064	TTC
TG-06-587	UAAAAGGAAUCCUCUAUGCC	2065	TTC
TG-06-588	GCAGGCCUUGAAGUUGCCCC	2066	TTC
TG-06-589	ACGGGAUGCUCUGGGCAAAG	2067	TTC
TG-06-590	UCCGCCCGGUACCGUGGAGG	2068	TTC
TG-06-591	GCUCAGAUGGGGUGGGGCAG	2069	TTC
TG-06-592	UGUUGCCCCACAAAUGCAA	2070	TTC
TG-06-593	AAAUGGUGAGAAAAGGAAAA	2071	TTC
TG-06-594	CACCAUUUGGAAUGGCCCAG	2072	TTC
TG-06-595	AAGGUAAGUGCAGCCUGCUG	2073	TTC
TG-06-596	UCUGAUGGUGACUCUACCUA	2074	TTC
TG-06-597	UGGGGCUCCCUGCCUGGGCC	2075	TTC
TG-06-598	CUCUAGGCCUUUCUUUCCAC	2076	TTC
TG-06-599	UUCCACUGUCCCUCUGCCU	2077	TTC
TG-06-600	CUCUGCCUGGUGUGGCCUGG	2078	TTC
TG-06-601	AUAGCCAGGCCACACCAGGC	2079	TTC
TG-06-602	AUGGCCGCUCCUCCUCCUC	2080	TTC
TG-06-603	AGCAGGGACUUGCAGGUGGC	2081	TTC
TG-06-604	CCGACUGAAUUGCCACCUGC	2082	TTC
TG-06-605	GUCGGGGAAGACUCUAGAUG	2083	TTC
TG-06-606	CUCCUCAGGCCAGGUGCAUC	2084	TTC
TG-06-607	UUCUCUCCUCAGGCCAGGUG	2085	TTC
TG-06-608	AGUUGGACUGUGUUAAGUUU	2086	TTC
TG-06-609	CCCCAAGCUCUGCGUCCUGG	2087	TTC
TG-06-610	CCCCUUACCGGUUCUCGUGG	2088	TTC
TG-06-611	CAAGAUCCCACGAGAACCGG	2089	TTC
TG-06-612	CGUGGGAUCUUGGGAAGCCA	2090	TTC
TG-06-613	UCUAUGAGCUUUGUCGUUCU	2091	TTC
TG-06-614	AUGAGCUUUGUCGUUCUUGU	2092	TTC
TG-06-615	CACAGGUUGCUGUGAGACUC	2093	TTC

TG-06-616	ACUCAUUGGAGUCUCACAGC	2094	TTC
TG-06-617	AAAAGCAUCUGCACAUCCUU	2095	TTC
TG-06-618	CUUCCAAAAGCAUCUGCACA	2096	TTC
TG-06-619	CCACACCCAGCCCACCAGCU	2097	TTC
TG-06-620	GCUGCUGCCACCCUUGCCAG	2098	TTC
TG-06-621	UACAGAGGAGCUGUCCUCUC	2099	TTC
TG-06-622	UUUCCAACCAGGUGAGACCU	2100	TTC
TG-06-623	AACCAGGUGAGACCUCUUCA	2101	TTC
TG-06-624	GUGGAGCCUGGAGCCCCCU	2102	TTC
TG-06-625	CAGUCCCUGCCUGGACUGGG	2103	TTC
TG-06-626	AACCGAGGGGGUGGUAAUUU	2104	TTC
TG-06-627	CUUGCCUCUUCCGUGUCCCA	2105	TTC
TG-06-628	GUGUCCCAACAGCGGCAGCA	2106	TTC
TG-06-629	UUGUCUCAUAAGGACACUUG	2108	TTC
TG-06-630	GGCCCAAGGGAGAGGAGGGA	2109	TTC
TG-06-631	CUCCUCUCCCUUGGGCCAGA	2110	TTC
TG-06-632	UCCCCAUAGGAUGCCCAUGG	2111	TTC
TG-06-633	CCAUAAGGAUGCCCAUGGAUU	2112	TTC
TG-06-634	CGCUUCCCCCGAGUGUCUGU	2113	TTC
TG-06-635	CCCGAGUGUCUGUGGGAGGU	2114	TTC
TG-06-636	ACACCCUGCCUGCCAUUCC	2116	TTC
TG-06-637	GGAGUCAUAUCAAGGGCUUG	2117	TTC
TG-06-638	UCCAGCCAAGCCCUUGAUAU	2118	TTC
TG-06-639	UCAUUCAUGGGGAGGCCACA	2119	TTC
TG-06-640	UUCAUGGGGAGGCCACAAGC	2120	TTC
TG-06-641	UGGGGAGGCCACAAGCAUGG	2121	TTC
TG-06-642	CUGUCAGGUAUUGAGUGAGG	2122	TTC
TG-06-643	UUCACAUUUGUAUCCCCAGC	2124	TTC
TG-06-644	UUUCAUUCACAUUUGUAUCC	2125	TTC
TG-06-645	GGAUACACUGUUGGGGGAAG	2126	TTC
TG-06-646	CCCAACAGUGUAUCCAGAAG	2127	TTC
TG-06-647	UUAACAAAUGCUACUUGAC	2128	TTC
TG-06-648	GAGCUCUCAGCCUAUCCAGA	2129	TTC
TG-06-649	GGAUAGGCUGAGAGCUCUGA	2130	TTC
TG-06-650	GCUGGAUACACAGGCUCGCA	2131	TTC
TG-06-651	UGAUCCUUCUGCUGGAUACA	2132	TTC
TG-06-652	UGGAGGAAGCGCUGUUCUAG	2133	TTC
TG-06-653	AGCCAAGACCUACGGGGGCA	2134	TTC
TG-06-654	ACUGCUUGGACACCGUCCCC	2135	TTC
TG-06-655	UUCCACUGCUUGGACACCGU	2136	TTC

TG-06-656	AUCCACGUUCAUUCCACUGC	2137	TTC
TG-06-657	GCACAGGCCACACUCCCUCC	2138	TTC
TG-06-658	CAUCUGACCUGGAGGCCACC	2139	TTC
TG-06-659	CAGGCGUAUGGCAGGGCUGC	2141	TTC
TG-06-660	CUCUCACCAGGCAGCCUUGC	2142	TTC
TG-06-661	AGCCUCUGGAAAGCUCUGUG	2143	TTC
TG-06-662	UGAGGCCACCUUGCCACGCU	2144	TTC
TG-06-663	CUGUGACAGCGUGGCAAGGU	2145	TTC
TG-06-664	AAACCCCCGGGUGAGGCAGG	2146	TTC
TG-06-665	UAGGCCUCCUCUAUCUUCC	2147	TTC
TG-06-666	AGUGAGAGGCUGGAAAGGGG	2148	TTC
TG-06-667	UUCCAGUGAGAGGCUGGAAA	2149	TTC
TG-06-668	AGCCUCUCACUGGAAGGAAG	2150	TTC
TG-06-669	UGCCCGCUCACACCAGCCCA	2151	TTC
TG-06-670	AAGUGCAGGCGGUUCCUGCC	2152	TTC
TG-06-671	AUCUGCCGCUGGGGCCACAC	2153	TTC
TG-06-672	GCUGCAAGAUGAUGUGACCA	2154	TTC
TG-06-673	GGUCCACCCGUGCUUGGGA	2155	TTC
TG-06-674	UCGUCGUGUGUGAGGAGGGA	2156	TTC
TG-06-675	CUCCUCACACACGACGAGGA	2157	TTC
TG-06-676	AGUCCAUGUUCCUCGUCGUG	2158	TTC
TG-06-677	GCUCGAUGUGCGUUGUGUGG	2159	TTC
TG-06-678	CUCCCUGGACGUAAGCCACA	2160	TTC
TG-06-679	AGAUAAAGCCUGUAUUCCCCG	2161	TTC
TG-06-680	CCGGGCUGUCUGCCAAUGUA	2162	TTC
TG-06-681	AGGCUGACGGGACAACUGGA	2163	TTC
TG-06-682	CUCAGAGCUUCCAGGCUGAC	2164	TTC
TG-06-683	UGAGCACCUGUAUCCCCUGC	2165	TTC
TG-06-684	ACAUUCCAUAAGCCCAGCUCA	2166	TTC
TG-06-685	AUAGCCCAGCUCAAUAUCUA	2167	TTC
TG-06-686	UCUAGGGUGGCGAGCACUGU	2168	TTC
TG-06-687	GGUCGGAGCUCACAGCGCUC	2169	TTC
TG-06-688	CAGUGUAGCACCGGGCACAU	2170	TTC
TG-06-689	CUCAUCAAUAGGCAUCUACC	2171	TTC
TG-06-690	GUAGCAUGCUGAGCAGGUCC	2172	TTC
TG-06-691	UCUGUAGCAUGCUGAGCAGG	2173	TTC
TG-06-692	UGCCGAGACUGGAGGCCUUA	2174	TTC
TG-06-693	GACUGUGCCCAGACCUACUA	2175	TTC
TG-06-694	GGCAAGGUCGGCCCCCAGUC	2176	TTC
TG-06-695	UGUGGGAUCUCAGAGCAGCC	2177	TTC

TG-06-696	GUGGCUGCUCUGAGAUCCCA	2178	TTC
TG-06-697	CCAAAUAAGGUCUGGCUCAC	2179	TTC
TG-06-698	CUCAUAAAGACACAUCACAG	2180	TTC
TG-06-699	CCAAUCUCCACCUCUGUCUC	2181	TTC
TG-06-700	UCCCAGAGCCUUCGCUGCUG	2182	TTC
TG-06-701	CAGAGCCUUCGCUGCUGAGU	2183	TTC
TG-06-702	CAGACUCAGCAGCGAAGGCU	2184	TTC
TG-06-703	CUGCUGAGUCUGGGAAUUUG	2185	TTC
TG-06-704	GUCACAAAUUCCCAGACUCA	2186	TTC
TG-06-705	UAUGGCUUCGGUCACAAAUU	2187	TTC
TG-06-706	CAUCUCAGGACAGGUGAGCA	2188	TTC
TG-06-707	GAAGUGCCAUCUGAACAGCA	2189	TTC
TG-06-708	GAUGGCACUUCAGAAGACUC	2190	TTC
TG-06-709	GAAGACUCAGGAGACCCUGG	2191	TTC
TG-06-710	UGGCUGCAGGGGUUCCAGGG	2192	TTC
TG-06-711	GCUGGCGCAGGACUCUGCUA	2193	TTC
TG-06-712	GGCAGGUUGCUGCUAGCAGA	2194	TTC
TG-06-713	UUUGGCCUGGCUGAGAGUUU	2195	TTC
TG-06-714	GAGACCUGCGCUGGAGCGGA	2196	TTC
TG-06-715	UCUCUCCCUUCUCCAUCCAG	2197	TTC
TG-06-716	CCAUCACAGCAGGCUGGACCU	2198	TTC
TG-06-717	CCUAUACCCUAACCUUUGUC	2199	TTC
TG-06-718	CCAAGAAGCAGGCACUGGGG	2200	TTC
TG-06-719	CUGAGGUUGUGACUCGUGUG	2201	TTC
TG-06-720	GCCUCACACGAGUCACAACC	2202	TTC
TG-06-721	CCACUCCACAAGAUACCAGG	2203	TTC
TG-06-722	UGUCUCUAAAAUGAGCCGGC	2204	TTC
TG-06-723	CCUUUGAGUCUUUAAAUCUU	2205	TTC
TG-06-724	CAGCUCACCGAGUCUGCAAG	2206	TTC
TG-06-725	UGCAGACUCGGUGAGCUGUG	2207	TTC
TG-06-726	UAUGAAGGAGCAGGAUGACU	2208	TTC
TG-06-727	UGGCAGUAGCAUUGCCAGCU	2209	TTC
TG-06-728	CCAGUGCCUGGGAUGUGCUC	2211	TTC
TG-06-729	GUGUGUUAACUAUAAGGUUG	2212	TTC
TG-06-730	AAGGCGACAUUUGUGGGAGA	2214	TTC
TG-06-731	CGGUGGUCACUCUGUAUGCU	2215	TTC
TG-06-732	CGAAGUCGGUGACCAUGACC	2216	TTC
TG-06-733	AGAAUGUGCCCGAGGAGGAC	2217	TTC
TG-06-734	ACAGACAGGUAAGCACGGCC	2218	TTC
TG-06-735	CACAACUGAUGGCAGGGAGG	2219	TTC

TG-06-736	UCCAUCCAGCCACCUGCUGA	2220	TTC
TG-06-737	CAUGUGGUCCUUGUGUUCGU	2221	TTC
TG-06-738	UCGAGCAGGCCAGCAAGUGU	2222	TTC
TG-06-739	CUUGGCAGUUGAGCACGCGC	2223	TTC
TG-06-740	UGGUUAGGAGACAUUAGCUC	2224	TTC
TG-06-741	CUGGCUCUGCGGCAGAGGCU	2225	TTC
TG-06-742	UCUGCACUCGUGGCCACUGG	2226	TTC
TG-06-743	CAGUGGCCACGAGUGCAGAU	2227	TTC
TG-06-744	UCAUCUGCACUCGUGGCCAC	2228	TTC
TG-06-745	GCCUUGACCCUGGGUCAAUG	2229	TTC
TG-06-746	ACAGUUCCACACUGCUCUCCC	2230	TTC
TG-06-747	AUGUGGAGGGACUCGAGGUG	2231	TTC
TG-06-748	AUCUUCUGCUUGCAUACCUC	2232	TTC
TG-06-749	GCUUGCAUACCUCUGAGACC	2233	TTC
TG-06-750	AGACUGUGACUACAUUUAGU	2235	TTC
TG-06-751	CAACUGGACUUCAUAGGGGG	2236	TTC
TG-06-752	CUGUCCACCGGCCCCAGAG	2237	TTC
TG-06-753	UUCUGUGGCCAGACAUUGAG	2238	TTC
TG-06-754	GUGGCCAGACAUUGAGCAAG	2239	TTC
TG-06-755	UCAACUAUUUAGCAGCUACG	2240	TTC
TG-06-756	CCAGCUUGAGCUGUGCGACC	2241	TTC
TG-06-757	CAGCGAGUUCUCCAGCUUGA	2242	TTC
TG-06-758	ACUGUAAGCCCCAGCCCCCA	2243	TTC
TG-06-759	CAGCCUUUGAAACAGGCUAG	2244	TTC
TG-06-760	AAGGCUGAGAAGGCCCCAGGA	2245	TTC
TG-06-761	GCCCUAGAGACUUUCCUACAG	2246	TTC
TG-06-762	UGCCUUUAAAACUGAACAUUC	2247	TTC
TG-06-763	GCCCCAUCAGGUGACCCCUU	2248	TTC
TG-06-764	GGAACUGACCUGACUGAGCC	2249	TTC
TG-06-765	UGAAGGUCAGUGUGUGGAGG	2250	TTC
TG-06-766	GGGAAAGAACAGCCUCCCCU	2251	TTC
TG-06-767	UUCCCAGAAAGGCCUUCCCC	2252	TTC
TG-06-768	CCAGGCCUGGUGCUCUGGCC	2254	TTC
TG-06-769	CCAGGAAGUCCUCCAGAGGC	2255	TTC
TG-06-770	UGGAGAAGUCCCUUCUUUGG	2256	TTC
TG-06-771	UUGGGGUCCCAGUCAGUGUA	2257	TTC
TG-06-772	CAUACACUGACUGGGACCCC	2258	TTC
TG-06-773	UCCUCUGUAAGCAGGUGAUU	2259	TTC
TG-06-774	UCACCUCCACACUCAGUGCC	2260	TTC
TG-06-775	CCUCCACACUCAGUGCCACU	2261	TTC

TG-06-776	AGCUCCAUCCUUGGCCACGG	2262	TTC
TG-06-777	GGAUCCUGGCCUCUAAGAUG	2263	TTC
TG-06-778	CCCCUGUCAAAUGAGGUGUG	2264	TTC
TG-06-779	AGGUCCCCUUGCACAUCAGG	2265	TTC
TG-06-780	AACCCUCUGCCAGCCUGUGA	2266	TTC
TG-06-781	GGCCUCAGUGGCACUCCACU	2267	TTC
TG-06-782	UUCUGGCCUCAGUGGCACUC	2268	TTC
TG-06-783	UUAGACAAUCCUGGGAGCAA	2269	TTC
TG-06-784	AGUUUGUGUGGGUCCAACAC	2271	TTC
TG-06-785	CCCGCCACUCUCUGAGGCCU	2272	TTC
TG-06-786	CCACCCAACUUCCCCCGCCA	2273	TTC
TG-06-787	CAUUCCUCUGAGGCUCAUCC	2274	TTC
TG-06-788	UCUGAGGCUCAUCCUCUUGG	2275	TTC
TG-06-789	AUGAAAGAGAGGCACCAAGA	2276	TTC
TG-06-790	UGGAAAGGGAUAAUAAGGUU	2277	TTC
TG-06-791	UAUAUUCAGACGCUUAGACA	2278	TTC
TG-06-792	GACGCUUAGACAGAGCCAGG	2279	TTC
TG-06-793	CUGUGCCUGGCUCUGUCUAA	2280	TTC
TG-06-794	AGCUCAUGUGACCGUGCGGG	2281	TTC
TG-06-795	GCAGGAAUUGCCAGCCUGCC	2283	TTC
TG-06-796	UGCAGAAUUUGAUGUUGCAA	2284	TTC
TG-06-797	UUCUUCCCUGGGAGGCCUCA	2286	TTC
TG-06-798	UCCCUGGGAGGCCUCAGUCU	2287	TTC
TG-06-799	CUGGGAGGCCUCAGUCUGUU	2288	TTC
TG-06-800	CAGCAGCAUUUCCACUGGCU	2289	TTC
TG-06-801	CAGCCCCUCCCUUACAGGCA	2290	TTC
TG-06-802	UCCAGGCCUGGAGUUUAUU	2291	TTC
TG-06-803	GAAUAAACUCCAGGCCUGGG	2292	TTC
TG-06-804	GAAAAGCCAGCUGGUCCAGC	2293	TTC
TG-06-805	GGGACGAUGCCUGCCUCUAC	2294	TTC
TG-06-806	UGUGGCACGUGGGCUUCUUG	2295	TTC
TG-06-807	UGUGGCACGUUCCUGGAGGC	2296	TTC
TG-06-808	GCCUCCAGGAACGUGCCACA	2297	TTC
TG-06-809	UGGAGGCCGAACCCUUCUGG	2298	TTC
TG-06-810	AAAGCCAGAAGGGUUCGGCC	2299	TTC
TG-06-811	GGCUUUGGAAGGAGUCGUCA	2300	TTC
TG-06-812	UCCCCAGCCUCCCGCAUGGC	2301	TTC
TG-06-813	AGCCCUUCCUCCCCAGCCU	2302	TTC
TG-06-814	GACUCUGCGAUGAUGGAGGU	2303	TTC
TG-06-815	GGAGGGAGAGUGGCCAGGAC	2304	TTC

TG-06-816	GGUACCCGCCCCCUCCCCAU	2305	TTC
TG-06-817	CAUUAGCCUUGCUCUAGUUG	2306	TTC
TG-06-818	UCGGGUGAGAGGUCUGAGCC	2307	TTC
TG-06-819	CCUAGAUACCCGCAGUGUCC	2308	TTC
TG-06-820	UUCCCCUAGAUACCCGCAGU	2309	TTC
TG-06-821	AGACUUGAAGUGGAACAAGA	2310	TTC
TG-06-822	ACUUCAAGUCUGGAACUUCA	2311	TTC
TG-06-823	AGUCUGGAACUUCAAGUCUG	2312	TTC
TG-06-824	AGUCUGUGUGUGUGCGUGCG	2313	TTC
TG-06-825	GCCGCCAGGGAAGCCUUGAU	2314	TTC
TG-06-826	CUGGCGGCCGAAUUUAAAGG	2315	TTC
TG-06-827	CCCCGUGGCACUUAGAACCA	2316	TTC
TG-06-828	GUGUAUCAAGUGGUUCUAAAG	2317	TTC
TG-06-829	UAUCUCCCCCACUAGGAUGU	2318	TTC
TG-06-830	CAAUGCCCCUGUGGAGUUUA	2319	TTC
TG-06-831	UUAGGAGGGGACAUUUGAGU	2320	TTC
TG-06-832	AGGUCAUCACAGUUGGGGCC	2322	TTC
TG-06-833	UCUCUGGUCCCCCUGCACAG	2323	TTC
TG-06-834	AGGACACUCAGUCUGAUGAG	2324	TTC
TG-06-835	UCUGUGCACUCGCCUCCCUC	2325	TTC
TG-06-836	CUGACCCUCGGGCUGGGCAG	2326	TTC
TG-06-837	UCCAUCAAGCUCCCGAUCAA	2327	TTC
TG-06-838	CCUGGCCGAUAAGACAGCAA	2328	TTC
TG-06-839	CGAAGGUAAGCCGCCUGCAG	2329	TTC
TG-06-840	AGAAGAGAGCUUAGUGUCUG	2330	TTC
TG-06-841	CCCCUGCACCAGGCAUUGCA	2331	TTC
TG-06-842	CUGCCAAAGAUGUCAUCAAU	2332	TTC
TG-06-843	CUGAGGACCAGCGGGUACUG	2333	TTC
TG-06-844	CCCUUGUCUGUGUAAGGAGG	2334	TTC
TG-06-845	GAGCCCCAUUCUCAUUUAAU	2336	TTC
TG-06-846	UUAAGCCCUCCUCUCUCCUA	2337	TTC
TG-06-847	UGGUAGGAGAGAGGAGGGCU	2338	TTC
TG-06-848	GUGGAGGUCCCCUCACUCCC	2339	TTC
TG-06-849	CACUCCUGGAGAAACUGGAG	2340	TTC
TG-06-850	CCAGGAGUGGGAAGCGGCGG	2341	TTC
TG-06-851	GUAAGCUUACAGGGCUGGCC	2342	TTC
TG-06-852	CGCGGCUGGGGGCUGCUGGA	2343	TTC
TG-06-853	CGCUGUGGCUCUCUCCAGCA	2344	TTC
TG-06-854	UAGAGGCUGCAGGACAGUGC	2345	TTC
TG-06-855	AUGCUCACCUGGGUGUGAGC	2346	TTC

TG-06-856	ACCAAAAUGCUGCAACCAGA	2348	TTC
TG-06-857	GGCUUGAGGAGAUGGGGUCU	2349	TTC
TG-06-858	AAAUGGAGAUAGAGAGAAAU	2350	TTC
TG-06-859	AAACACAGAGUGGUUUCAAA	2351	TTC
TG-06-860	UCCAAACACAGAGUGGUUUC	2352	TTC
TG-06-861	GCCAGUGGGCAGCUCUGAAU	2353	TTC
TG-06-862	GAGCUGCCCACUGGCAGAAC	2354	TTC
TG-06-863	UUGGGCAGGAGUGAGCUCCU	2355	TTC
TG-06-864	GGGGACAAGGAGCUCACUCC	2356	TTC
TG-06-865	AAGCCCAUUCUAAAGCAGAU	2357	TTC
TG-06-866	AAAGCAGAUUCCCAUUUCCG	2358	TTC
TG-06-867	CAUUUCCGUCUUUGACUCUA	2359	TTC
TG-06-868	GUCUUUGACUCUAAGGCCCA	2360	TTC
TG-06-869	CCUUAUGCACCCACUGCCCG	2361	TTC
TG-06-870	AUACUCUGGUUCUGCCACUU	2362	TTC
TG-06-871	AUGCCCUUUGAGCCUGGGCA	2364	TTC
TG-06-872	CCUCAUCUCAGCUUCCUCCU	2365	TTC
TG-06-873	UCCUCCAUAAGAGGGAAAAA	2366	TTC
TG-06-874	GCCCCAGCCGGGGGCCUCUC	2367	TTC
TG-06-875	AACCGAGGUCUGCCCUCCUU	2368	TTC
TG-06-876	CCCUGCCGCUCUAUAGCUGU	2369	TTC
TG-06-877	UCAUACAGAGAUGUUAAGUA	2370	TTC
TG-06-878	AGCUAUCGUUACUAUUGUUU	2371	TTC
TG-06-879	UCACCCGUUGCUUUAAAAUC	2372	TTC
TG-06-880	ACACAGGGUAGCCCGUCAAU	2373	TTC
TG-06-881	UGCUAACUGCCUGGGCAUGU	2374	TTC
TG-06-882	GGUUGGGUGGAGCAUGAGAG	2375	TTC
TG-06-883	AGGCUGGAGUCCCACUGUCU	2377	TTC
TG-06-884	CCUGGACGUGCUCAGACGAG	2378	TTC
TG-06-885	CGUCUGAGCACGUCCAGGGG	2379	TTC
TG-06-886	CAAACCCUCAGUGGUUCCCC	2380	TTC
TG-06-887	GGGCACAGCAGGCUGGCAUG	2381	TTC
TG-06-888	UGAACAUCCAUCCUACAUGC	2382	TTC
TG-06-889	GGAAAGAUGAGCAUAGUCAG	2383	TTC
TG-06-890	CUCCAAGCCAGGACCCUGA	2384	TTC
TG-06-891	GGGUUCCUGGCUUGGAGGGA	2385	TTC
TG-06-892	AAAGGGGAGCGUGAUUAGAG	2386	TTC
TG-06-893	AACUGACAUGUGCCUGUGAG	2387	TTC
TG-06-894	AUGACUCAAACUGACAUGU	2388	TTC
TG-06-895	UGGAGUCAGAUUCCAUGACU	2389	TTC

TG-06-896	CUUAGCUGCAAAAUGGUGCU	2390	TTC
TG-06-897	CUAUGUCCUAGCUGCAAGGG	2391	TTC
TG-06-898	GAGCACAUGCCAGGCUCCAU	2392	TTC
TG-06-899	CACGUGUCAUCUUCACAGUG	2393	TTC
TG-06-900	CAGUGUCCCUGUGAGUAGGU	2394	TTC
TG-06-901	CUUUCCAUCUUACAAAUGAG	2395	TTC
TG-06-902	UAGCCAAUGGGUGUGUCUGA	2397	TTC
TG-06-903	GAGGUGGUUCAGACAGACCC	2398	TTC
TG-06-904	GACAGACCCCGGCUUACCCU	2399	TTC
TG-06-905	AGGGUAAGCCGGGGUCUGUC	2400	TTC
TG-06-906	GCCUGCUGGCUGCAUAGGGA	2401	TTC
TG-06-907	ACCCAGAGAUUGGAGACAGUG	2402	TTC
TG-06-908	CAGGGAAAACACCAUCUUUC	2403	TTC
TG-06-909	CUGAGAAACUAAGGCUCAGA	2404	TTC
TG-06-910	UUGAGUUGUUUCUAGGUUUC	2405	TTC
TG-06-911	AGGUUUCUAGCUCUUGCCU	2406	TTC
TG-06-912	UAGCUCUUGCCUCAGACCUU	2407	TTC
TG-06-913	CAUCUGAGCUGGCUUCCUC	2408	TTC
TG-06-331	UCUGCCCCAGGCUGCAGCUC	2409	TTC
TG-06-915	UGCUGCCAUGCCCCAGGUCU	2410	TTC
TG-06-916	AGACCUGGGGCAUGGCAGCA	2411	TTC
TG-06-917	AUGCUCUUGACUUUGCAUU	2412	TTC
TG-06-332	CCUGCUCCUGAGGGGCCGGG	2413	TTC
TG-06-919	GUGUCAGUUUGUGCCACCAC	2414	TTC
TG-06-920	GCCAUUGUGUGGACAGCAUG	2415	TTC
TG-06-921	ACUCCAAGUGGAGUGGGGUA	2416	TTC
TG-06-922	UUAUGCCAUAUUUAUGCACA	2417	TTC
TG-06-923	GCAUGCAGCUCAGGCACCCC	2418	TTC
TG-06-924	UGCCUUGCUGGUCAUGCUIAA	2419	TTC
TG-06-925	AGGGCUUGGCACAUGCCUUG	2420	TTC
TG-06-926	AGAAGGCCCGGCUGUCACAU	2421	TTC
TG-06-927	AGAAACCUGAGUGUAUAACU	2422	TTC
TG-06-928	UUUAAGAGUGGACUCCUAG	2423	TTC
TG-06-929	GGUCUGGAUGGUGCAUUCCU	2424	TTC
TG-06-930	UCGAGAGCAAAAGACAGUCU	2425	TTC
TG-06-931	AGGUGCUCUUCCUGGCACUG	2426	TTC
TG-06-932	CCCUCCUGGAGCAUCUAUUU	2427	TTC
TG-06-933	UCCACUAAAAUAGAUGCUC	2428	TTC
TG-06-934	GUACUCUAUAUUAUGGUUU	2430	TTC
TG-06-935	UCCACCUUGUCAGCGAGACC	2431	TTC

TG-06-936	CUGCCUCAUGGCCUGUUUCG	2432	TTC
TG-06-937	AGGCAAAGCAAAGGCCCCCA	2433	TTC
TG-06-938	UGGUACUCCCUGCACAUCCC	2434	TTC
TG-06-939	CACCUCUCUGGUUUCUGGU	2435	TTC
TG-06-940	GAUCCAGGCCAGACCACACU	2436	TTC
TG-06-941	CAGCUCAGCAGCUAGGCCAC	2437	TTC
TG-06-942	GCCCAAGGUCCUUGUCAUCU	2438	TTC
TG-06-943	GAGGGAUGAGUUUGGCACUG	2440	TTC
TG-06-944	UGCCUGGUUUGAGGGCCGUG	2441	TTC
TG-06-945	GCACCAUGGCUGACCAAACA	2442	TTC
TG-06-946	GUGAGGCCAAAGUGCAGACG	2443	TTC
TG-06-947	CCCUCCGACAUCCUUCUGUG	2444	TTC
TG-06-948	UUUCUCCCUCGACAUCCUU	2445	TTC
TG-06-949	UUUCCUCGGGCUCUGGCAGG	2446	TTC
TG-06-950	UCGGGCUCUGGCAGGUGACC	2447	TTC
TG-06-951	CUGGUGCUGCCUGUAGUGCU	2448	TTC
TG-06-952	GACUUGUCCCUCUCUCAGCC	2449	TTC
TG-06-953	GCCUUUCCGGGGCUGCUGGC	2450	TTC
TG-06-333	GGGGCUGCUGGCCUGGCCCU	2451	TTC
TG-06-955	AGGCAAGGAGGCUGCCCCAC	2452	TTC
TG-06-956	UGGCACCUCCACUGGGGAG	2453	TTC
TG-06-957	CCAUAUCAAACAGGUCGAGCU	2454	TTC
TG-06-958	AACAGGUCGAGCUGUGCUCG	2455	TTC
TG-06-959	GCCCACGGACAUCGGCACAU	2456	TTC
TG-06-960	GUGCCAGGCAUUCAAUCCUC	2457	TTC
TG-06-961	AUCCUCAGGUCUCCACCAAG	2458	TTC
TG-06-962	UCCCAUGGAUAGGGGAGGGG	2459	TTC
TG-06-963	CAUGGAUAGGGGAGGGGGCG	2460	TTC
TG-06-964	CCACAGUUAGCUGGAGAUGA	2461	TTC
TG-06-965	GGAUGGCAUCUAGCCAGAGG	2462	TTC
TG-06-966	UGAGCCACCUUUACUCUGCU	2463	TTC
TG-06-967	UCCUCUGUGAAGUAGGGGUG	2465	TTC
TG-06-968	CAGAGGAAGAAACCUGGAAC	2466	TTC
TG-06-335	UGCUGUGUGAGCUUGGCAGG	2467	TTC
TG-06-970	GGCUCAGUUCUGCUGUGUG	2468	TTC
TG-06-971	GAGCCAGCCCAAUCUGCGUU	2469	TTC
TG-06-972	UACUUCACCCGGCUGGGCUC	2470	TTC
TG-06-973	CCCGGCUGGGCUCCUCAUUU	2471	TTC
TG-06-974	CAGCCUCACUGUUAACCGUA	2472	TTC
TG-06-336	CCUUCCCAGCCUCACUGUUA	2473	TTC

TG-06-976	UGGUCUGUGUUCCCCUUCCC	2474	TTC
TG-06-977	GCCAUCACUCACCGAGCUUC	2475	TTC
TG-06-978	AUGCCUGCAGGCAUCGUUCU	2476	TTC
TG-06-337	GUUAUCACCCAGGCCUGAUU	2477	TTC
TG-06-338	CUGGCCUGGCGGAGAUGCUU	2478	TTC
TG-06-339	AAGGCAUGGUCGGGGGAGAG	2479	TTC
TG-06-341	AGUUACAAAAGCAAAACAGG	2480	TTC
TG-06-983	UAGUAUAAGCAAGGCAUCAC	2481	TTC
TG-06-984	ACAAACUUACCCAGUUCCUA	2482	TTC
TG-06-985	CAUCACCAGUACUUGCCCAA	2483	TTC
TG-06-986	UAGGAUAGGCACAAGGAGUA	2484	TTC
TG-06-987	UUCCUAGGAUAGGCACAAGG	2485	TTC
TG-06-988	UGCAGCUCCCCUGACCACCC	2486	TTC
TG-06-989	UUUCUCCCCAGUACCUGGCA	2487	TTC
TG-06-990	CCCCAGUACCUGGCACACAG	2488	TTC
TG-06-991	GUAAAUACUCGGGCCAGGCU	2489	TTC
TG-06-992	GUAAUGAGCCAGUGCUAAGU	2490	TTC
TG-06-993	UGAUUUAAAAUUCAGAACAG	2491	TTC
TG-06-994	GAACAGUCUCAGAUUGGGUG	2492	TTC
TG-06-995	UAAACACAAGCAAUCCUCCU	2493	TTC
TG-06-996	AGACCCUGGGCAACAGUGAG	2494	TTC
TG-06-997	AUGUGGGACCUCCAGGCCCU	2495	TTC
TG-06-998	CCAGCCUUAGAGUGGAGGGC	2496	TTC
TG-06-999	CUCUCUCUCUAGACAGCCCA	2497	TTC
TG-06-1000	UUAUGACAUGGGCUGUCUAG	2498	TTC
TG-06-1001	GACAAUUCAGUGUCAUAUUAU	2499	TTC
TG-06-1002	AUAUAUGACACUGAAUUGUC	2500	TTC
TG-06-1003	GUGUCAUAUAUGGAAAGCAU	2501	TTC
TG-06-1004	AGAGUUGACAGAUGGCAAUG	2502	TTC
TG-06-1005	AGCCAUUGCCAUCUGUCAAC	2503	TTC
TG-06-1006	AACGGGCAGAGAUAACACAG	2504	TTC
TG-06-1007	GUGGAGGGGAAGGAGAGAGA	2505	TTC
TG-06-1008	CCUCCACAGAAUCAAGUAUC	2506	TTC
TG-06-1009	AACAUCACAGAGCUGUCAGG	2507	TTC
TG-06-1010	UGCAGAGGGUUAGCAGGUGG	2508	TTC
TG-06-1011	UCCCCGAUCCUGGUGACCA	2509	TTC
TG-06-1012	CUAUUUGCUAGUCCAGAUUU	2511	TTC
TG-06-1013	UAGUGGCCAGCGUUGCUCAC	2512	TTC
TG-06-1014	UGCUAGAUCCACAGGUCUC	2513	TTC
TG-06-1015	GCAUCCUCCUCUAAGCUUUG	2514	TTC

TG-06-1016	UCUGCAAAGCUUAGAGGAGG	2515	TTC
TG-06-1017	GGAUUCAUCUGCAAAGCUUA	2516	TTC
TG-06-1018	CUGGAGGGGCUGACCAUUUC	2517	TTC
TG-06-1019	UUUCUCUGGAGGGGCUGACC	2518	TTC
TG-06-1020	UCUUAGUACUUUACACUAAU	2519	TTC
TG-06-1021	CCCGCCCACUACGGCUGUGU	2520	TTC
TG-06-1022	UAUAUCAGAACACAGCCGUA	2521	TTC
TG-06-1023	UCAUAUAUCAGAACACAGCC	2522	TTC
TG-06-1024	GAUAUAUGAAGAAUGGCGGC	2523	TTC
TG-06-1025	AAAACAGAGUUGGAGUCUUG	2524	TTC
TG-06-1026	UAGCUUCUAAGAGUAUAAAA	2525	TTC
TG-06-1027	UGUGUGGCAAAGUUCUUAGC	2526	TTC
TG-06-1028	AAAUUUGGUUCAUGUGUGGC	2527	TTC
TG-06-1029	UCCCUAGUGUAUCUACUCC	2528	TTC
TG-06-1030	CUAGUGUAUCUACUCCUUU	2529	TTC
TG-06-1031	AAAGGAAGUAGAUACACUAG	2530	TTC
TG-06-1032	UCUGCUACCUCUUGUGUAUA	2531	TTC
TG-06-1033	AGUCUGCUUGUCCCAAGCUC	2532	TTC
TG-06-1034	GAGCUUGGAACAAGCAGACU	2533	TTC
TG-06-1035	AAGCUCAGAAGGUAAACUGCA	2534	TTC
TG-06-1036	CUGCUAACAUUCGAGGUGUGU	2535	TTC
TG-06-1037	UAAAACUGACUGUAGUUGUC	2536	TTC
TG-06-1038	UUCCCAGGAUGCCUUCAACG	2537	TTC
TG-06-1039	GCCCGUUGAAGGCAUCCUGG	2538	TTC
TG-06-1040	CAGGAUGCCUUCAACGGGCA	2539	TTC
TG-06-1041	ACGGGCAGAAUCGGCCAUCU	2540	TTC
TG-06-1042	CAAGAAGCUCUCCUUAUGUGAG	2541	TTC
TG-06-1043	UGAGAACACACAGCAUUAUU	2542	TTC
TG-06-1044	UGCUCUGUUUGCAGACCAUU	2543	TTC
TG-06-1045	AAAUGGUCUGCAAACAGAGC	2544	TTC
TG-06-126	CCUUUUAUCCUCCUGCCUG	2672	TTC
TG-06-129	UUUUACACACCAUGUUCAAG	2675	TTC
TG-06-148	GCCGGGCCCACCUUUUCAGU	2694	TTC
TG-06-1046	UCUCUUACAUGGGGGGAAAC	2714	TTC

c. gRNA Modifications

[0251] In another aspect, the disclosure relates to gRNAs (sometimes referred to as gRNA variants herein), which comprise modifications relative to a reference gRNA from which the gRNAs were derived. The gRNAs can be used in the long-term gene repressor fusion protein

systems described herein. In some embodiments, a gRNA variant for use in the systems of the disclosure comprises one or more nucleotide substitutions, insertions, deletions, or swapped or replaced domains relative to a gRNA sequence of the disclosure that improve a characteristic. Exemplary regions for modifications and swapped regions or domains include the RNA triplex, the pseudoknot, the scaffold stem loop, and the extended stem loop. In some embodiments, the gRNA variant of the disclosure comprises at least a first swapped region from a different gRNA, resulting in a chimeric gRNA. A representative example of such a chimeric gRNA is guide 316 (SEQ ID NO: 1746), in which the extended stem loop of gRNA scaffold 235 (SEQ ID NO: 1745) is replaced with the extended stem loop of gRNA scaffold 174 (SEQ ID NO: 1744), wherein the resulting 316 variant retains the ability to form an RNP with a long-term repressor fusion protein and exhibits an improved characteristic compared to the parent 235, when assessed in an *in vitro* or *in vivo* assay under comparable conditions.

[0252] All gRNAs that have one or more improved functions, characteristics, or add one or more new functions when the gRNA scaffold variant is compared to a gRNA scaffold from which it was derived, while retaining the functional properties of being able to complex with the long-term repressor fusion protein and guide the ribonucleoprotein holo RNP complex to the *PCSK9* target nucleic acid are envisaged as within the scope of the disclosure. In some embodiments, the gRNA has an improved characteristic selected from the group consisting of increased pseudoknot stem stability, increased triplex region stability, increased scaffold stem stability, extended stem stability, reduced off-target folding intermediates, increased binding affinity to a long-term repressor fusion protein, and increased repression activity when complexed with a long-term repressor fusion protein, or any combination thereof. In some cases, the improved characteristic is assessed in an *in vitro* assay, including the assays of the Examples. In other cases, the improved characteristic is assessed *in vivo*.

[0253] Table 10 provides exemplary gRNA variant scaffold sequences of the disclosure that are utilized as gRNA scaffolds or for the generation of the gRNAs for use in the LTRP:gRNA systems of the disclosure. In some embodiments, the gRNA variant scaffold for use in the systems comprises a sequence selected from the group consisting of SEQ ID NOS: 1744-1746, or a sequence having at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% sequence identity thereto, wherein the gRNA variant retains the ability to form an RNP with a long-term repressor fusion protein of the disclosure. In some embodiments, the gRNA variant scaffold for use in

the systems comprises a sequence selected from the group consisting of SEQ ID NOS: 1744-1746. In some embodiments, the gRNA variant scaffold for use in the systems comprises SEQ ID NO: 1744. In some embodiments, the gRNA variant scaffold for use in the systems comprises SEQ ID NO: 1745. In some embodiments, the gRNA variant scaffold for use in the systems comprises SEQ ID NO: 1746. It will be understood that in those embodiments wherein a vector comprises a DNA encoding sequence for a gRNA, that thymine (T) bases can be substituted for the uracil (U) bases of any of the gRNA sequence embodiments described herein. Similarly, any RNA sequence disclosed herein can be encoded by a DNA in which the uracil bases are substituted by thymine. In some embodiments, the disclosure provides gRNA variants of Table 10 that are chemically-modified, described below. In some embodiments, the gRNA comprises a scaffold comprising a sequence of SEQ ID NOS: 1744-1746, and is chemically modified.

Table 10: gRNA Scaffold Sequences

SEQ ID NO:	Scaffold variant ID	Nucleotide sequence
1744	174	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGCGACUAUGUCGUAGUGG GUAAAGCUCUCCUCUUCGGAGGGAGCAUCAAG
1745	235	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGCGACUAUGUCGUAGUGG GUAAAGCCGCUUACGGACUUCGGUCCGUAAGAGGCAUCAGAG
1746	316	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGCGACUAUGUCGUAGUGG GUAAAGCUCUCCUCUUCGGAGGGAGCAUCAGAG

[0254] Additional gRNA variants contemplated for use in the systems of the disclosure are selected from the group consisting of SEQ ID NOS: 1747-1821. In some embodiments, the gRNA comprises a scaffold comprising a sequence of SEQ ID NOS: 1747-1821, and is chemically modified.

[0255] Guide scaffolds can be made by several methods, including recombinantly or by solid-phase RNA synthesis. However, the length of the scaffold can affect the manufacturability when using solid-phase RNA synthesis, with longer lengths resulting in increased manufacturing costs, decreased purity and yield, and higher rates of synthesis failure. For use in particle formulations, such as lipid nanoparticle (LNP) formulations, solid-phase RNA synthesis of the scaffold is preferred to generate the quantities needed for commercial development. While previous experiments had identified gRNA scaffold 235 (SEQ ID NO: 1745) as having enhanced properties relative to gRNA scaffold 174 (SEQ ID NO: 1744), its increased length (in nucleotides) rendered its use for LNP formulations less desirable due to synthetic manufacturing constraints. Accordingly, alternative sequences were

sought. In some embodiments, the disclosure provides gRNA variant scaffolds having improved manufacturability compared to the gRNA scaffold from which it was derived. In some embodiments, the disclosure provides a gRNA wherein the gRNA scaffold and linked targeting sequence has a sequence that is less than about 115 nucleotides, less than about 110 nucleotides, or less than about 100 nucleotides. In a particular embodiment, the 316 gRNA scaffold (SEQ ID NO: 1746) has a shorter sequence compared to the 235 scaffold from which it was derived. The 316 gRNA scaffold was designed wherein the scaffold 235 sequence was modified by a domain swap in which the extended stem loop of scaffold 174 replaced the extended stem loop of the 235 scaffold, resulting in the chimeric gRNA scaffold 316, having the sequence

ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGCGACUAUGUCGUA
GUGGGUAAAGCUCCUCUUCGGAGGGAGCAUCAGAG (SEQ ID NO: 1746), having 89 nucleotides, compared with the 99 nucleotides of gRNA scaffold 235. The resulting 316 scaffold had the further advantage in that the extended stem loop does not contain CpG motifs; an enhanced property conferring reduced potential to elicit an immune response. In some embodiments, the shorter sequence length of the 316 scaffold confers the improvements of a higher fidelity in the ability to create the guide synthetically with the correct and complete sequence, as well as an enhanced ability to be successfully incorporated into an LNP. In some embodiments, the disclosure provides gRNA 316 variants that are chemically-modified, described below.

d. Chemically-modified gRNAs

[0256] In some embodiments, the gRNAs have one or more chemical modifications. In some embodiments, the chemical modification is the addition of a 2'O-methyl group to one or more nucleotides of the sequence. In some embodiments, the chemical modification is substitution of a phosphorothioate bond between two or more nucleotides of the sequence. . In some embodiments, the chemical modification is a substitution of phosphorothioate bonds between two or more nucleotides on each terminal end of the gRNA. In some embodiments, the gRNA comprises a substitution of phosphorothioate bonds between two or more nucleotides located 1, 2, 3 or 4 nucleotides the from the 5' terminal end, the 3' terminal, or both terminal ends of the gRNA. In some embodiments, the gRNA comprises an addition of a 2'O-methyl group to one or more nucleotides of the gRNA. In some embodiments, one or more nucleotides located 1, 2, 3, or 4 nucleotides from the 5' terminal end, the 3' terminal, or both terminal ends of the gRNA are modified by an addition of a 2'O-methyl group. In

some embodiments, the first 1, 2, or 3 nucleotides of the 5' end of the scaffold (*i.e.*, A, C, and U in the case of gRNA 174, 235, and 316) are modified by the addition of a 2'-O-methyl group and each of the modified nucleotides is linked to the adjoining nucleotide by a phosphorothioate bond. Similarly, the last 1, 2, or 3 nucleotides of the 3' end of the targeting sequence linked to the 3' end of the scaffold are similarly modified. In some embodiments, the disclosure provides gRNA with chemical modifications selected from the group consisting of the sequences of SEQ ID NOS: 2948-2956, 2958-2966, and 2968-2976, as set forth in Table 22, or a sequence having at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto. In any of the foregoing, the 20 nucleotides on the 3' end of the sequence are replaced with a targeting sequence complementary to a *PCSK9* target nucleic acid sequence, optionally the targeting sequence is selected from the group consisting of SEQ ID NOS: 1824-1880, 1883, 1884, 1888, 1889, 2672, 2675, 2694, and 2714, but includes the chemical modifications on the 3' end of the gRNA. In some embodiments, the 20 nucleotides on the 3' end of the sequence are replaced with a targeting sequence complementary to a *PCSK9* target nucleic acid sequence selected from the group consisting of SEQ ID NOS: 1834, 1849, 1853, 1855-1858, 1860, 1862, 1863, 1867, 1869, 1870, 1872, 1874, and 1875, but includes the chemical modifications on the 3' end of the gRNA. In some embodiments, the 20 nucleotides on the 3' end of the sequence are replaced with a targeting sequence complementary to a *PCSK9* target nucleic acid sequence selected from the group consisting of SEQ ID NOS: 1855, 1867 and 1869, but includes the chemical modifications on the 3' end of the gRNA. In some embodiments, the gRNA with chemical modifications comprises a sequence selected from the group consisting of the sequences of SEQ ID NOS: 2948-2956, 2958-2966, and 2968-2976 with the 20 nucleotides on the 3' end replaced with a targeting sequence selected from the group consisting of SEQ ID NOS: 1824-1880, 1883, 1884, 1888, 1889, 2672, 2675, 2694, and 2714 but having the chemical modifications on the 3' end. In some embodiments, the chemically modified gRNA comprises the sequence of SEQ ID NO: 2968 with the 20 nucleotides on the 3' end replaced with a targeting sequence selected from the group consisting of SEQ ID NOS: 1824-1880, 1883, 1884, 1888, 1889, 2672, 2675, 2694, and 2714 but having the chemical modifications on the 3' end. In some embodiments, the chemically modified gRNA comprises the sequence of SEQ ID NO: 2968 with the 20 nucleotides on the 3' end replaced with a targeting sequence selected from the group consisting of SEQ ID NOS: 1855, 1867 and 1869, but having the

chemical modifications on the 3' end. In some embodiments, the modified gRNA comprises a sequence selected from the group consisting of SEQ ID NOS: 22780- 22803. In some embodiments, the modified gRNA a sequence selected from the group consisting of SEQ ID NOS: 22788-22790. A schematic of the structure of gRNA variants 174, 235, and 316 are shown in FIGS. 7-9, respectively, and chemically modified versions of gRNA variants 235 and 316 are shown in FIGS. 10-14. In some embodiments, the gRNA with chemical modifications exhibit improved stability compared to a gRNA without chemical modifications.

e. Complex Formation with Repressor Fusion Protein

[0257] Upon delivery or expression of the components of the system in a target cell, the gRNA variant is capable of complexing with a long-term repressor fusion protein as an RNP, and binding to the target nucleic acid of the *PCSK9* gene. In some embodiments, a gRNA variant has an improved ability to form an RNP complex with a long-term repressor fusion protein when compared to a reference gRNA. Improving ribonucleoprotein complex formation may, in some embodiments, improve the efficiency with which functional RNPs are assembled. In some embodiments, greater than 90%, greater than 93%, greater than 95%, greater than 96%, greater than 97%, greater than 98% or greater than 99% of RNPs comprising a gRNA variant and its targeting sequence and a long-term repressor fusion protein of the disclosure are competent for gene repression of a target nucleic acid.

VII. Polynucleotides and Vectors

[0258] The disclosure provides polynucleotides encoding the long-term repressor fusion protein and/or gRNA that have utility in the repression and epigenetic modification of the *PCSK9* gene. The disclosure provides polynucleotides encoding mRNA encoding the long-term repressor fusion protein, e.g. DNA polynucleotides encoding the corresponding mRNA.

[0259] A long-term repressor fusion protein or an mRNA encoding the long-term repressor fusion protein of the disclosure may be prepared by *in vitro* synthesis, using conventional methods as known in the art. Various commercial synthetic apparatuses are available, for example, automated synthesizers by Applied Biosystems, Inc., Beckman, etc. By using synthesizers, naturally occurring amino acids or nucleotides (as applicable) may be substituted with unnatural amino acids or nucleotides. The particular sequence and the manner of preparation will be determined by convenience, economics, purity required, and the like. A gRNA can also be produced synthetically; for example, by use of a T7 RNA polymerase system known in the art.

[0260] The long-term repressor fusion protein and/or the gRNA may also be prepared by recombinantly producing a polynucleotide sequence coding for the long-term repressor fusion protein or gRNA of any of the embodiments described herein using standard recombinant techniques known in the art and incorporating the encoding gene into an expression vector appropriate for a host cell. For production of the encoded long-term repressor fusion protein and/or gRNA of any of the embodiments described herein, the methods include transforming an appropriate host cell with an expression vector comprising the encoding polynucleotide, and culturing the host cell under conditions causing or permitting the resulting long-term repressor fusion protein or gRNA of any of the embodiments described herein to be expressed or transcribed in the transformed host cell, which are recovered by methods described herein or by standard purification methods known in the art or as described in the Examples. Standard recombinant techniques in molecular biology are used to make the polynucleotides and expression vectors of the present disclosure.

[0261] A long-term repressor fusion protein and/or a gRNA of the disclosure may also be isolated and purified in accordance with conventional methods of recombinant synthesis. A lysate may be prepared of the expression host and the lysate purified using high performance liquid chromatography (HPLC), exclusion chromatography, gel electrophoresis, affinity chromatography, or other purification technique. For the most part, the compositions which are used will comprise 50% or more by weight of the desired product, more usually 75% or more by weight, preferably 95% or more by weight, and for therapeutic purposes, usually 99.5% or more by weight, in relation to contaminants related to the method of preparation of the product and its purification. Usually, the percentages will be based upon total protein. Thus, in some cases, a long-term repressor fusion protein or gRNA of the present disclosure is at least 80% pure, at least 85% pure, at least 90% pure, at least 95% pure, at least 98% pure, or at least 99% pure (*e.g.*, free of contaminants or other macromolecules, etc.).

[0262] Additionally, the disclosure provides vectors comprising polynucleotides encoding the long-term repressor fusion proteins and the gRNAs described herein. In some cases, the vectors are utilized for the expression and recovery of the CasX and gRNA components of the LTRP:gRNA system, when the system is delivered as a repressor fusion protein and gRNA, or an RNP. In other cases, the vectors are utilized for the delivery of the encoding polynucleotides to target cells for the repression and/or epigenetic modification of the target nucleic acid, as described more fully, below. In some embodiments, sequences encoding the long-term repressor fusion protein and a gRNA are templated on the same vector. In some

embodiments, sequences encoding the long-term repressor fusion protein and a gRNA are templated on different vectors. Suitable vectors are described, for example, in WO2023235818A2, WO2022120095A1 and WO2020247882A1, incorporated by reference herein. As described in WO2023235818A2, WO2022120095A1 and WO2020247882A1, depending on the host/vector system utilized, any of a number of suitable transcription and translation control elements, including constitutive and inducible promoters, transcription enhancer elements, transcription terminators, etc. may be used in the expression vector.

[0263] The present disclosure provides polynucleotide sequences encoding the long-term repressor fusion protein of any of the embodiments described herein. In some embodiments, the disclosure provides a DNA sequence encoding a long-term repressor fusion protein for use in a vector. In some embodiments, the disclosure provides an mRNA sequence encoding a long-term repressor fusion protein of any of the embodiments described herein for use in a particle system for delivery to a cell. In some embodiments, the disclosure provides an mRNA sequence encoding a long-term repressor fusion protein of any of the embodiments described herein for use in an LNP particle formulation for delivery to a cell. In a particular embodiment, the disclosure provides a gRNA and an mRNA sequence encoding a long-term repressor fusion protein of any of the embodiments described herein for use in an LNP particle formulation for delivery to a cell. In some embodiments, the disclosure provides an isolated polynucleotide sequence encoding a gRNA variant of any of the embodiments described herein, with a linked targeting sequence complementary to a *PCSK9* target nucleic acid sequence.

[0264] In some embodiments, the disclosure relates to methods to produce polynucleotide sequences encoding the long-term repressor fusion protein or the gRNA of any of the embodiments described herein, including variants thereof, as well as methods to express the proteins or RNA transcribed by the polynucleotide sequences. In general, the methods include producing a polynucleotide sequence coding for the long-term repressor fusion protein or the gRNA of any of the embodiments described herein and incorporating the encoding gene into an expression vector. In some embodiments, the vector is designed for transduction of cells for repression and/or epigenetic modification of the *PCSK9* target nucleic acid. Such vectors can include a retroviral vector, a lentiviral vector, an adenoviral vector, an adeno-associated viral (AAV) vector, a herpes simplex virus (HSV) vector, a plasmid, a minicircle, a nanoplasmid, a DNA vector, and an RNA vector. In other embodiments, the expression vector is designed for production of a long-term repressor

fusion protein, mRNA encoding the long-term repressor fusion protein, or gRNA in either a cell-free system or in a host cell. For production of the encoded long-term repressor fusion protein or the gRNA of any of the embodiments described herein in a host cell, the methods include transforming an appropriate host cell with an expression vector comprising the encoding polynucleotide, and culturing the host cell under conditions causing or permitting the resulting long-term repressor fusion protein or the gRNA of any of the embodiments described herein to be expressed or transcribed in the transformed host cell, thereby producing the long-term repressor fusion protein or the gRNA, which are recovered by methods described herein (*e.g.*, in the Examples, below) or by standard purification methods known in the art. Standard recombinant techniques in molecular biology are used to make the polynucleotides and expression vectors of the present disclosure.

[0265] In accordance with the disclosure, nucleic acid sequences that encode the long-term repressor fusion protein or the gRNA of any of the embodiments described herein are used to generate recombinant DNA molecules that direct the expression in appropriate host cells. Several cloning strategies are suitable for performing the present disclosure, many of which are used to generate a construct that comprises a gene coding for a long-term repressor fusion protein or gRNA of the present disclosure, or its complement. In some embodiments, the cloning strategy is used to create a gene that encodes a construct that comprises nucleotides encoding the long-term repressor fusion protein or the gRNA. In some embodiments, the gene, for example as part of a vector, is used to transform a host cell for expression of the gene, *e.g.* expression of the long-term repressor fusion protein or gRNA.

[0266] In one approach, a construct is first prepared containing the DNA sequence encoding a long-term repressor fusion protein or a gRNA. Exemplary methods for the preparation of such constructs are described in the Examples. The construct is then used to create an expression vector suitable for transforming a host cell, such as a prokaryotic or eukaryotic host cell for the expression and recovery of the protein construct, in the case of the long-term repressor fusion protein, or the gRNA. Where desired, the host cell is an *E. coli*. In other embodiments, the host cell is a eukaryotic cell. The eukaryotic host cell can be selected from Baby Hamster Kidney fibroblast (BHK) cells, human embryonic kidney 293 (HEK293), human embryonic kidney 293T (HEK293T), NS0 cells, SP2/0 cells, YO myeloma cells, P3X63 mouse myeloma cells, PER cells, PER.C6 cells, hybridoma cells, NIH3T3 cells, CV-1 (simian) in Origin with SV40 genetic material (COS), HeLa, Chinese hamster ovary (CHO), yeast cells, or other eukaryotic cells known in the art suitable for the production of

recombinant products. Exemplary methods for the creation of expression vectors, the transformation of host cells and the expression and recovery of the long-term repressor fusion protein or the gRNA are described in the Examples.

[0267] The gene encoding the long-term repressor fusion protein or the gRNA construct can be made in one or more steps, either fully synthetically or by synthesis combined with enzymatic processes, such as restriction enzyme-mediated cloning, PCR and overlap extension, including methods more fully described in the Examples. The methods disclosed herein can be used, for example, to ligate sequences of polynucleotides encoding the various components into a gene of a desired sequence. Genes encoding polypeptide compositions are assembled from oligonucleotides using standard techniques of gene synthesis.

[0268] In some embodiments, the nucleotide sequence encoding a long-term repressor fusion protein is codon optimized. This type of optimization can entail a mutation of an encoding nucleotide sequence to mimic the codon preferences of the intended host organism or cell while encoding the same protein. Thus, the codons can be changed, but the encoded protein remains unchanged. For example, if the intended target cell of the long-term repressor fusion protein was a human cell, a human codon-optimized long-term repressor fusion protein-encoding nucleotide sequence could be used. As another non-limiting example, if the intended host cell were a mouse cell, then a mouse codon-optimized long-term repressor fusion protein-encoding nucleotide sequence could be generated. The gene design can be performed using algorithms that optimize codon usage and amino acid composition appropriate for the host cell utilized in the production of the long-term repressor fusion protein or the gRNA. In one method of the disclosure, a library of polynucleotides encoding the components of the constructs is created and then assembled, as described above. The resulting genes are then assembled and the resulting genes used to transform a host cell and produce and recover the long-term repressor fusion protein or the gRNA compositions for evaluation of its properties or for use in the modification of the *PCSK9* target nucleic acid, as described herein.

[0269] In some embodiments, a nucleotide sequence encoding a gRNA is operably linked to a control element, *e.g.*, a transcriptional control element, such as a promoter. In some embodiments, a nucleotide sequence encoding a long-term repressor fusion protein is operably linked to a control element, *e.g.*, a transcriptional control element, such as a promoter. In some cases, the promoter is a constitutively active promoter. In some cases, the promoter is a regulatable promoter. In some cases, the promoter is an inducible promoter. In

some cases, the promoter is a tissue-specific promoter. In some cases, the promoter is a cell type-specific promoter. In some cases, the transcriptional control element (*e.g.*, the promoter) is functional in a targeted cell type or targeted cell population. For example, in some cases, the transcriptional control element can be functional in eukaryotic cells, *e.g.*, hepatocytes or a liver sinusoidal endothelial cell.

[0270] Non-limiting examples of Pol II promoters operably linked to the polynucleotide encoding the long-term repressor fusion protein of the disclosure include, but are not limited to EF-1alpha, EF-1alpha core promoter, Jens Tørnøe (JeT), promoters from cytomegalovirus (CMV), CMV immediate early (CMVIE), CMV enhancer, herpes simplex virus (HSV) thymidine kinase, early and late simian virus 40 (SV40), the SV40 enhancer, long terminal repeats (LTRs) from retrovirus, mouse metallothionein-I, adenovirus major late promoter (Ad MLP), CMV promoter full-length promoter, the minimal CMV promoter, the chicken β -actin promoter (CBA), CBA hybrid (CBh), chicken β -actin promoter with cytomegalovirus enhancer (CB7), chicken beta-Actin promoter and rabbit beta-Globin splice acceptor site fusion (CAG), the rous sarcoma virus (RSV) promoter, the HIV-Ltr promoter, the hPGK promoter, the HSV TK promoter, a 7SK promoter, the Mini-TK promoter, the human synapsin I (SYN) promoter which confers neuron-specific expression, beta-actin promoter, super core promoter 1 (SCP1), the Mecp2 promoter for selective expression in neurons, the minimal IL-2 promoter, the Rous sarcoma virus enhancer/promoter (single), the spleen focus-forming virus long terminal repeat (LTR) promoter, the TBG promoter, promoter from the human thyroxine-binding globulin gene (Liver specific), the PGK promoter, the human ubiquitin C promoter (UBC), the UCOE promoter (Promoter of HNRPA2B1-CBX3), the synthetic CAG promoter, the Histone H2 promoter, the Histone H3 promoter, the U1a1 small nuclear RNA promoter (226 nt), the U1a1 small nuclear RNA promoter (226 nt), the U1b2 small nuclear RNA promoter (246 nt) 26, the GUSB promoter, the CBh promoter, rhodopsin (Rho) promoter, silencing-prone spleen focus forming virus (SFFV) promoter, a human H1 promoter (H1), a POL1 promoter, the TTR minimal enhancer/promoter, the b-kinesin promoter, mouse mammary tumor virus long terminal repeat (LTR) promoter, the human eukaryotic initiation factor 4A (EIF4A1) promoter, the ROSA26 promoter, the glyceraldehyde 3-phosphate dehydrogenase (GAPDH) promoter, tRNA promoters, and truncated versions and sequence variants of the foregoing. In a particular embodiment, the Pol II promoter is EF-1alpha, wherein the promoter enhances transfection efficiency, the

transgene transcription or expression of the CRISPR nuclease, the proportion of expression-positive clones and the copy number of the episomal vector in long-term culture.

[0271] Non-limiting examples of Pol III promoters operably linked to the polynucleotide encoding the gRNA variants of the disclosure include, but are not limited to U6, mini U6, U6 truncated promoters, 7SK, and H1 variants, BiH1 (Bidirectional H1 promoter), BiU6, Bi7SK, BiH1 (Bidirectional U6, 7SK, and H1 promoters), gorilla U6, rhesus U6, human 7SK, human H1 promoters, and truncated versions and sequence variants thereof. In the foregoing embodiment, the pol III promoter enhances the transcription of the gRNA. In a particular embodiment, the Pol III promoter is U6, wherein the promoter enhances expression of the gRNA. . Experimental details and data for the use of such promoters are provided in the Examples.

[0272] Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art, as it related to controlling expression. The expression vector may also contain a ribosome binding site for translation initiation, and a transcription terminator. The expression vector may also include appropriate sequences for amplifying expression. The expression vector may also include nucleotide sequences encoding protein tags (*e.g.*, 6xHis tag, hemagglutinin tag, fluorescent protein, *etc.*) that can be fused to the long-term repressor fusion protein, thus resulting in a chimeric protein that are used for purification or detection.

[0273] Recombinant expression vectors of the disclosure can also comprise elements that facilitate robust expression of the proteins and the gRNAs of the disclosure. For example, recombinant expression vectors can include one or more of a polyadenylation signal (poly(A)), an intronic sequence or a post-transcriptional regulatory element such as a woodchuck hepatitis post-transcriptional regulatory element (WPRE). Exemplary poly(A) sequences include hGH poly(A) signal (short), HSV TK poly(A) signal, synthetic polyadenylation signals, a split poly(A) sequence with a SphI restriction site between two 60 A stretches (SEQ ID NO: 3307), SV40 poly(A) signal, β -globin poly(A) signal and the like. A person of ordinary skill in the art will be able to select suitable elements to include in the recombinant expression vectors described herein.

[0274] The polynucleotides encoding the long-term repressor fusion protein or the gRNA sequences can be individually cloned into an expression vector. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art, as it relates to controlling expression, *e.g.*, for repressing expression and/or epigenetic modification of the PCSK9 gene. The expression vector may also contain a ribosome binding site for translation

initiation and a transcription terminator. The expression vector may also include appropriate sequences for amplifying expression.

[0275] The nucleic acid sequence is inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan. Such techniques are well known in the art and well described in the scientific and patent literature. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage that may conveniently be subjected to recombinant DNA procedures, and the choice of vector will often depend on the host cell into which it is to be introduced. Thus, the vector may be an autonomously replicating vector, *i.e.*, a vector, which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, *e.g.*, a plasmid. Alternatively, the vector may be one which, when introduced into a host cell, is integrated into the host cell genome and replicated together with the chromosome(s) into which it has been integrated. Once introduced into a suitable host cell, expression of the long-term repressor fusion protein or gRNA can be determined using any nucleic acid or protein assay known in the art. For example, the presence of transcribed mRNA of the long-term repressor fusion protein can be detected and/or quantified by conventional hybridization assays (*e.g.*, Northern blot analysis), amplification procedures (*e.g.*, RT-PCR), SAGE (U.S. Pat. No. 5,695,937), and array-based technologies (see *e.g.*, U.S. Pat. Nos. 5,405,783, 5,412,087 and 5,445,934), using probes complementary to any region of the polynucleotide.

[0276] In some embodiments, a vector is created for the transcription of the long-term repressor fusion protein, and expression and recovery of the resulting encoding mRNA. In some embodiments, the mRNA is generated by *in vitro* transcription (IVT) using a PCR product or linearized plasmid DNA template and a T7 RNA polymerase, wherein the plasmid contains a T7 promoter. If using a PCR product, DNA sequences encoding candidate mRNAs will be cloned into a plasmid containing a T7 promoter, wherein the plasmid DNA template will be linearized and then used to perform IVT reactions for expression of the mRNA. Exemplary methods for the generation of such vectors and the production and recovery of the mRNA are provided in the Examples, below.

VIII. Particles for Delivery of LTRP:gRNA systems

[0277] In another aspect, the present disclosure provides particle compositions for delivery of the LTRP:gRNA systems to cells or to subjects for the transcriptional repression or silencing of the *PCSK9* gene. Particles envisaged as within the scope of the instant disclosure include, but are not limited to, nanoparticles such as synthetic nanoparticles, polymeric nanoparticles, lipid nanoparticles, viral particles and virus-like particles. Particles of the disclosure may encapsulate payloads such as gRNA variants, as described herein, optionally in combination with mRNA encoding the long-term repressor fusion proteins of any of the embodiments described herein. Alternatively, or in addition, particles of the disclosure may encapsulate payloads of gRNA variants and long-term repressor fusion proteins, for example when associated as a ribonucleoprotein (RNP) complex. In some embodiments, the particles are synthetic nanoparticles that encapsulate payloads of gRNA variants and mRNA encoding long-term repressor fusion protein of any of the embodiments described herein. In some embodiments, the synthetic nanoparticles comprise biodegradable polymeric nanoparticles (PNP). In some embodiments, materials for the creation of biodegradable polymeric nanoparticles (PNP) include polylactide, poly (lactic-co-glycolic acid) (PLGA), poly(ethyl cyanoacrylate), poly(butyl cyanoacrylate), poly(isobutyl cyanoacrylate), and poly(isohexyl cyanoacrylate), polyglutamic acid (PGA), poly (ϵ -caprolactone) (PCL), cyclodextrin, and natural polymers for instance chitosan, albumin, gelatin, and alginate, which are the most utilized polymers for the synthesis of PNP (Production and clinical development of nanoparticles for gene delivery. *Molecular Therapy-Methods & Clinical Development* 3:16023; doi:10.1038 (2016)). In other embodiments, the particles are lipid nanoparticles that encapsulate a gRNA variant and an mRNA encoding the long-term repressor fusion protein of any of the embodiments described herein, described more fully, below. In other embodiments, the particles are lipid nanoparticles (LNPs) that separately encapsulate a gRNA variant and a mRNA encoding long-term repressor fusion protein of any of the embodiments described herein in different particles, that are the co-formulated as a mixture for administration, described more fully, below. In other embodiments, the particles are lipid nanoparticles that separately encapsulate a gRNA variant and a mRNA encoding long-term repressor fusion protein of any of the embodiments, and the two types of particles are administered separately.

a. Lipid Nanoparticles (LNP)

[0278] In another aspect, the present disclosure provides lipid nanoparticles (LNP) for delivery of the LTRP:gRNA systems of the disclosure to cells or to subjects for the transcriptional repression of the *PCSK9* gene. In some embodiments, the LNPs of the disclosure are tissue- or organ-specific (*e.g.*, the liver), have excellent biocompatibility, and can deliver the systems with high efficiency, and thus can be usefully used for the transcriptional repression of the *PCSK9* gene.

[0279] The disclosure further provides LNP compositions and pharmaceutical compositions comprising a plurality of the LNP described herein.

[0280] In their native forms, nucleic acid polymers are generally unstable in biological fluids and cannot penetrate into the cytoplasm of target cells, thus requiring delivery systems. Lipid nanoparticles (LNP) have proven useful for both the protection and delivery of nucleic acids to tissues and cells. Furthermore, the use of mRNA in LNPs to encode the long-term repressor fusion protein eliminates the possibility of undesirable genome integration, as compared to DNA vectors. Moreover, mRNA efficiently transfects both mitotic and non-mitotic cells, as it does not require entry into the nucleus since it exerts its function in the cytoplasmic compartment. LNPs as a delivery platform thus offer the additional advantage of being able to co-formulate both the mRNA encoding the long-term repressor fusion protein and a gRNA into single LNP particles.

[0281] Accordingly, in various embodiments, the disclosure encompasses lipid nanoparticles and compositions that may be used for a variety of purposes, including the delivery of encapsulated or associated (*e.g.*, complexed) therapeutic agents such as nucleic acids to cells, both *in vitro* and *in vivo*. In certain embodiments, the disclosure encompasses methods of treating or preventing diseases or disorders in a subject in need thereof by contacting the subject with a lipid nanoparticle that encapsulates or is associated with a suitable therapeutic agent complexed through various physical, chemical or electrostatic interactions between one or more of the lipid components used in the compositions to make LNPs. In some embodiments, the suitable therapeutic agent comprises a LTRP:gRNA system as described herein.

[0282] In certain embodiments, the lipid nanoparticles are useful for the delivery of nucleic acids, including, *e.g.*, the mRNA encoding the long-term repressor fusion proteins and the gRNA variants of the disclosure. In some embodiments, an LNP comprises an mRNA comprising a sequence encoding a long-term repressor fusion protein selected from the group

consisting of SEQ ID NOS: 6528-14626 (unmodified mRNA), and 14627-22725 (mRNA modified with N1-methyl-pseudouridine), or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto. In some embodiments, an LNP comprises an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6528-8134, 9741-11347, 14628-16233, and 17840-19446 or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto. In some embodiments, an LNP comprises an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6529-6537, 9742-9750, 14628, 14636, and 17841-17849, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto. In some embodiments, an LNP comprises an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6529-6537, 9742-9750, 14628, 14636, and 17841-17849. In some embodiments, an LNP comprises an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6531, 9744, 14630, and 17843, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto. In some embodiments, an LNP comprises an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6531, 9744, 14630, and 17843. In some embodiments, an LNP comprises an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6529, 9742, 14628, and 17841, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto. In some embodiments, an LNP comprises an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6529, 9742, 14628, and 17841. In some embodiments, an LNP comprises an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6530, 9743,

14629, and 17842, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto. In some embodiments, an LNP comprises an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6530, 9743, 14629, and 17842. In some embodiments, an LNP comprises a gRNA comprising a scaffold sequence selected from the group consisting of SEQ ID NOS: 1744-1746, 2948-2956, 2958-2966, and 2968-2974, and a linked targeting sequence complementary to a sequence of the *PCSK9* gene (with, in the case of SEQ ID NOS: 2948-2956, 2958-2966, and 2968-2974, the 20 non-targeting nucleotides on the 3' end replaced with a targeting sequence complementary to a sequence of the *PCSK9* gene but retaining the chemical modifications on the 3' end). In some embodiments, an LNP comprises a gRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 2278-22803. In some embodiments, an LNP comprises a gRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 2278-22790. In some embodiments, the present disclosure provides LNP in which the gRNA and mRNA encoding the long-term repressor fusion protein are incorporated into single LNP particles. In other embodiments, the present disclosure provides LNP in which the gRNA and mRNA encoding the long-term repressor fusion protein are incorporated into separate populations of LNPs, which can be formulated together in varying ratios for administration.

[0283] The lipid nanoparticles and lipid nanoparticle compositions of the disclosure may be used to repress expression of a desired protein both *in vitro* and *in vivo* by contacting cells with a lipid nanoparticle comprising one or more ionizable lipids described herein, wherein the lipid nanoparticle encapsulates or is associated with a nucleic acid that is expressed to produce the desired protein (*e.g.*, a messenger RNA encoding the long-term repressor fusion protein). In some embodiments, the lipid nanoparticles and compositions may be used to repress the expression of a target gene both *in vitro* and *in vivo* by contacting cells with a lipid nanoparticle comprising one or more cationic lipids described herein, wherein the lipid nanoparticle encapsulates or is associated with one or more nucleic acids of the LTRP:gRNA systems of the disclosure. The lipid nanoparticles and compositions of embodiments of the disclosure may also be used for co-delivery of different nucleic acids (*e.g.*, mRNA and plasmid DNA) separately or in combination, such as may be useful to provide an effect

requiring colocalization of different nucleic acids (*e.g.*, mRNA encoding for a suitable gene repressing factor or enzyme and gRNA for targeting of the gene).

[0284] In some embodiments, LNPs and LNP compositions described herein include at least one cationic lipid, at least one conjugated lipid, at least one steroid or derivative thereof, at least one additional lipid, or any combination thereof. Alternatively, the lipid compositions of the disclosure can include an ionizable lipid, such as an ionizable cationic lipid, a helper lipid (usually a phospholipid), cholesterol, and a polyethylene glycol-lipid conjugate (PEG-lipid) to improve the colloidal stability in biological environments by, for example, reducing a specific absorption of plasma proteins and forming a hydration layer over the nanoparticles. Such lipid compositions can be formulated at typical mole ratios of 50:10:37-39:1.5-2.5 or 20-50:8-65:25-40:1-2.5, with variations made to adjust individual properties.

[0285] The LNPs and LNP compositions of the present disclosure are configured to protect and deliver an encapsulated payload of the systems of the disclosure to tissues and cells, both *in vitro* and *in vivo*. Various embodiments of the LNPs and LNP compositions of the present disclosure are described in further detail herein.

Cationic lipid

[0286] In some embodiments, the LNPs and LNP compositions of the present disclosure include at least one cationic lipid. The term “cationic lipid,” refers to a lipid species that has a net positive charge. In some embodiments, the cationic lipid is an ionizable cationic lipid that has a net positive charge at a selected pH, such as physiological pH. In some embodiments, the ionizable cationic lipid has a pKa less than 7 such that the LNPs and LNP compositions achieve efficient encapsulation of the payload at a relatively low pH. In some embodiments, the cationic lipid has a pKa of 5 to 8, 5.5 to 7.5, 6 to 7, or 6.5 to 7. In some embodiments, the cationic lipid may be protonated at a pH below the pKa of the cationic lipid, and it may be substantially neutral at a pH over the pKa. The LNPs and LNP compositions may be safely delivered to a target organ (for example, the liver, lung, heart, spleen, as well as to tumors) and/or cell (hepatocyte, LSEC, cardiac cell, cancer cell, etc.) *in vivo*, and after endocytosis, exhibit a positive charge to release the encapsulated payload through electrostatic interaction with an anionic protein of the endosome membrane.

[0287] Early formulations of LNP utilizing permanently cationic lipids resulted in LNPs with positive surface charge that proved toxic *in vivo*, plus were rapidly cleared by phagocytic cells. By changing to ionizable cationic lipids bearing tertiary amines, especially those with $pK_a < 7$, results in LNP achieving efficient encapsulation of nucleic acid polymers

at low pH by interacting electrostatically with the negative charges of the phosphate backbone of mRNA, that also result in largely neutral systems at physiological pH values, thus alleviating problems associated with permanently-charged cationic lipids.

[0288] As used herein, “ionizable lipid” means an amine-containing lipid which can be easily protonated, and, for example, it may be a lipid of which charge state changes depending on the surrounding pH. The ionizable lipid may be protonated (positively charged) at a pH below the pKa of a cationic lipid, and it may be substantially neutral at a pH over the pKa. In one example, the LNP may comprise a protonated ionizable lipid and/or an ionizable lipid showing neutrality. In some embodiments, the LNP has a pKa of 5 to 8, 5.5 to 7.5, 6 to 7, or 6.5 to 7. The pKa of the LNP can influence in vivo stability and release of the nucleic acid payload of the LNP in the target cell or organ. In some embodiments, the LNP having the foregoing pKa ranges may be safely delivered to a target organ (for example, the liver, lung, heart, spleen, as well as to tumors) and/or target cell (hepatocyte, LSEC, cardiac cell, cancer cell, etc.) in vivo, and after endocytosis, exhibit a positive charge to release the encapsulated payload through electrostatic interaction with an anionic protein of the endosome membrane.

[0289] The ionizable lipid is an ionizable compound having characteristics similar to lipids generally, and through electrostatic interaction with a nucleic acid (for example, an mRNA of the disclosure), may play a role of encapsulating the nucleic acid payloads within the LNP with high efficiency.

[0290] According to the type of the amine and the tail group comprised in the ionizable lipid, (i) the nucleic acid encapsulation efficiency, (ii) PDI (polydispersity index) and/or (iii) the nucleic acid delivery efficiency to tissue and/or cells constituting an organ (for example, hepatocytes or liver sinusoidal endothelial cells in the liver) of the LNP may be different. In certain embodiments, the ionizable lipid is an ionizable cationic lipid, and comprises from about 46 mol % to about 66 mol % of the total lipid present in the particle.

[0291] The LNP comprising an ionizable lipid comprising an amine may have one or more kinds of the following characteristics: (1) the ability to encapsulate a nucleic acid with high efficiency; (2) uniform size of prepared particles (or having a low PDI value); and/or (3) excellent nucleic acid delivery efficiency to organs such as liver, lung, heart, spleen, bone marrow, as well as to tumors, and/or cells constituting such organs (for example, hepatocytes, LSEC, cardiac cells, cancer cells, etc.).

[0292] In particular embodiments, the cationic lipid form plays an important role both in nucleic acid encapsulation through electrostatic interactions and intracellular release by disrupting endosomal membranes. The nucleic acid payloads are encapsulated within the LNP by the ionic interactions they form with the positively charged cationic lipid. Non-limiting examples of cationic lipid components utilized in the LNP of the disclosure are selected from DLin-MC3-DMA (heptatriaconta-6,9,28,31-tetraen-19-yl4-(dimethylamino)butanoate), DLin- KC2-DMA (2,2-dilinoleyl-4-(2-dimethylaminoethyl)-[1,3]-dioxolane), and TNT (1,3,5-triazine-2,4,6-trione) and TT (N1,N3,N5-tris(2-aminoethyl)benzene-1,3,5-tricarboxamide). Non-limiting examples of helper lipids utilized in the LNP of the disclosure are selected from DSPC (1,2-distearoyl-sn-glycero-3-phosphocholine), POPC (2-Oleoyl-1- palmitoyl-sn-glycero-3-phosphocholine) and DOPE (1,2-Dioleoyl-sn-glycero-3-phosphoethanolamine), 1,2-dioleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) DOPG, 1,2-Dimyristoyl-sn-glycero-3-phosphoethanolamine (DMPE), 1,2-dilauroyl-sn-glycero-3-phosphocholine (DLPC), sphingolipid, and ceramide. Cholesterol and PEG-DMG ((R)-2,3- bis(octadecyloxy)propyl-1-(methoxy polyethylene glycol 2000) carbamate), PEG-DSG (1,2-Distearoyl-rac-glycero-3-methylpolyoxyethylene glycol 2000), or DSPE-PEG2k (1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[amino(polyethylene glycol)-2000]), are components utilized in the LNP of the disclosure for the stability, circulation, and size of the LNP.

[0293] In some embodiments, the cationic lipid in the LNP of the disclosure comprises a tertiary amine. In some embodiments, the tertiary amine includes alkyl chains connected to N of the tertiary amine with ether linkages. In some embodiments, the alkyl chains comprise C12-C30 alkyl chains having 0 to 3 double bonds. In some embodiments, the alkyl chains comprise C16-C22 alkyl chains. In some embodiments, the alkyl chains comprise C18 alkyl chains. A number of cationic lipids and related analogs have been described in U.S. Patent Publication Nos. 20060083780, 20060240554, 20110117125, 20190336608, 20190381180 and 20200121809; U.S. Pat. Nos. 5,208,036; 5,264,618; 5,279,833; 5,283,185; 5,753,613; 5,785,992; 9,738,593; 10,106,490; 10,166,298; 10,221,127; and 11,219,634; and PCT Publication No. WO 96/10390, the disclosures of which are herein incorporated by reference in their entirety.

[0294] In some embodiments, the cationic lipid in the LNP of the disclosure may comprise, for example, one or more ionizable cationic lipids wherein the ionizable cationic lipid is a dialkyl lipid. In other embodiments, the ionizable cationic lipid is a trialkyl lipid.

[0295] In some embodiments, the cationic lipid in the LNP of the disclosure is selected from 1,2-dilinoleyloxy-N,N-dimethylaminopropane (DLinDMA), 1,2-dilinolenyloxy-N,N-dimethylaminopropane (DLenDMA), 2,2-dilinoleyl-4-(2-dimethylaminoethyl)-[1,3]-dioxolane (DLin-K-C2-DMA), 2,2-dilinoleyl-4-(3-dimethylaminopropyl)-[1,3]-dioxolane (DLin-K-C3-DMA), 2,2-dilinoleyl-4-(4-dimethylaminobutyl)-[1,3]-dioxolane (DLin-K-C4-DMA), 2,2-dilinoleyl-5-dimethylaminomethyl-[1,3]-dioxane (DLin-K6-DMA), 2,2-dilinoleyl-4-N-methylpepiazino-[1,3]-dioxolane (DLin-K-MPZ), 2,2-dilinoleyl-4-dimethylaminomethyl-[1,3]-dioxolane (DLin-K-DMA), 1,2-dilinoleylcarbamoyloxy-3-dimethylaminopropane (DLin-C-DAP), 1,2-dilinoleyloxy-3-(dimethylamino)acetoxyp propane (DLin-DAC), 1,2-dilinoleyloxy-3-morpholinopropane (DLin-MA), 1,2-dilinoleoyl-3-dimethylaminopropane (DLinDAP), 1,2-dilinoleylthio-3-dimethylaminopropane (DLin-S-DMA), 1-linoleoyl-2-linoleyloxy-3-dimethylaminopropane (DLin-2-DMAP), 1,2-dilinoleyloxy-3-trimethylaminopropane chloride salt (DLin-TMA.Cl), 1,2-dilinoleoyl-3-trimethylaminopropane chloride salt (DLin-TAP.Cl), 1,2-dilinoleyloxy-3-(N-methylpiperazino)propane (DLin-MPZ), 3-(N,N-dilinoleylamino)-1,2-propanediol (DLinAP), 3-(N,N-dioleylamino)-1,2-propanedio (DOAP), 1,2-dilinoleyloxo-3-(2-N,N-dimethylamino)ethoxyp propane (DLin-EG-DMA), N,N-dioleyl-N,N-dimethylammonium chloride (DODAC), 1,2-diolelyoxy-N,N-dimethylaminopropane (DODMA), 1,2-distearlyoxy-N,N-dimethylaminopropane (DSDMA), N-(1-(2,3-diolelyoxy)propyl)-N,N,N-trimethylammonium chloride (DOTMA), N,N-distearyl-N,N-dimethylammonium bromide (DDAB), N-(1-(2,3-dioleoyloxy)propyl)-N,N,N-trimethylammonium chloride (DOTAP), 3-(N—(N',N'-dimethylaminoethane)-carbamoyl)cholesterol (DC-Chol), N-(1,2-dimyristyloxyprop-3-yl)-N,N-dimethyl-N-hydroxyethyl ammonium bromide (DMRIE), 2,3-diolelyoxy-N-[2(spermine-carboxamido)ethyl]-N,N-dimethyl-1-propanaminiumtrifluoroacetate (DOSPA), dioctadecylamidoglycyl spermine (DOGS), 3-dimethylamino-2-(cholest-5-en-3-beta-oxybutan-4-oxy)-1-(cis,cis-9,12-octadecadienoxy)propane (CLinDMA), 2-[5'-(cholest-5-en-3-beta-oxy)-3'-oxapentoxy]-3-dimethyl-1-(cis,cis-9',1'-2'-octadecadienoxy)propane (CpLinDMA), N,N-dimethyl-3,4-diolelyoxybenzylamine (DMOBA), 1,2-N,N'-dioleylcarbamyl-3-dimethylaminopropane (DOcarbDAP), 1,2-N,N'-dilinoyleylcarbamyl-3-dimethylaminopropane (DLincarbDAP), and any combination of the foregoing.

[0296] In some embodiments, the cationic lipid in the LNP of the disclosure is selected from heptatriaconta-6,9,28,31-tetraen-19-yl4-(dimethylamino)butanoate (DLin-MC3-DMA),

2,2-dilinoleyl-4-(2-dimethylaminoethyl)-[1,3]-dioxolane (DLin- KC2-DMA), (1,3,5-triazinane-2,4,6-trione) (TNT), N1,N3,N5-tris(2-aminoethyl)benzene-1,3,5-tricarboxamide (TT), and any combination of the foregoing.

[0297] In some embodiments, the N/P ratio (nitrogen from the cationic/ionizable lipid and phosphate from the nucleic acid) in the LNP of the disclosure is in the range of is about 3:1 to 7:1, or about 4:1 to 6:1, or is 3:1, or is 4:1, or is 5:1, or is 6:1, or is 7:1.

Conjugated Lipid

[0298] In some embodiments, the LNPs and LNP compositions of the present disclosure include at least one conjugated lipid. In some embodiments, the conjugated lipid may be selected from a polyethyleneglycol (PEG)-lipid conjugate, a polyamide (ATTA)-lipid conjugate, a cationic-polymer-lipid conjugate (CPL), and any combination of the foregoing. In some cases, conjugated lipids can inhibit aggregation of the LNPs of the disclosure.

[0299] In some embodiments, the conjugated lipid of the LNP of the disclosure comprises a pegylated lipid. The terms “polyethyleneglycol (PEG)-lipid conjugate,” “pegylated lipid” “lipid-PEG conjugate”, “lipid-PEG”, “PEG-lipid”, “PEG-lipid”, or “lipid-PEG” are used interchangeably herein and refer to a lipid attached to a polyethylene glycol (PEG) polymer which is a hydrophilic polymer. The pegylated lipid contributes to the stability of the LNPs and LNP compositions and reduces aggregation of the LNPs.

[0300] As the PEG-lipid can form the surface lipid, the size of the LNP can be readily varied by varying the proportion of surface (PEG) lipid to the core (ionizable cationic) lipids. In some embodiments, the PEG-lipid of the LNP of the disclosure can be varied from ~1 to 5 mol% to modify particle properties such as size, stability, and circulation time.

[0301] The lipid-PEG conjugate contributes to the particle stability in serum of the nanoparticle within the LNP, and plays a role of preventing aggregation between nanoparticles. In addition, the lipid-PEG conjugate may protect nucleic acids, such as mRNAs encoding the long-term repressor fusion proteins of the disclosure, or gRNAs of the disclosure, from degrading enzymes during in vivo delivery of the nucleic acids and enhance the stability of the nucleic acids in vivo and increase the half-life of the delivered nucleic acids encapsulated in the nanoparticle. Examples of PEG-lipid conjugates include, but are not limited to, PEG-DAG conjugates, PEG-DAA conjugates, and mixtures thereof. In certain embodiments, the PEG-lipid conjugate is selected from the group consisting of a PEG-diacylglycerol (PEG-DAG) conjugate, a PEG-dialkylloxypropyl (PEG-DAA) conjugate, a PEG-phospholipid conjugate, a PEG-ceramide (PEG-Cer) conjugate, and a mixture thereof.

[0302] In some embodiments, the pegylated lipid of the LNP of the disclosure is selected from a PEG-ceramide, a PEG-diacylglycerol, a PEG-dialkylxypropyl, a PEG-dialkylxypropylcarbamate, a PEG-phosphatidylethanolamine, a PEG-phospholipid, a PEG-succinate diacylglycerol, and any combination of the foregoing.

[0303] In some embodiments, the pegylated lipid of the LNP of the disclosure is a PEG-dialkylxypropyl. In some embodiments, the pegylated lipid is selected from PEG-didecylxypropyl (C10), PEG-dilaurylxypropyl (C12), PEG-dimyristylxypropyl (C14), PEG-dipalmitylxypropyl (C16), PEG-distearylxypropyl (C18), and any combination of the foregoing.

[0304] In other embodiments, the lipid-PEG conjugate of the LNP of the disclosure may be PEG bound to phospholipid such as phosphatidylethanolamine (PEG-PE), PEG conjugated to ceramide (PEG-CER, ceramide-PEG conjugate, ceramide-PEG, cholesterol or PEG conjugated to derivative thereof, PEG-c-DOMG, PEG-DMG, PEG-DLPE, PEG-DMPE, PEG-DPPC, PEG-DSPE(DSPE-PEG), and a mixture thereof, and for example, may be C16-PEG2000 ceramide (N-palmitoyl-sphingosine-1-{succinyl[methoxy(polyethylene glycol)2000]}), DMG-PEG 2000, 14:0 PEG2000 PE.

[0305] In some embodiments, the pegylated lipid of the LNP of the disclosure is selected from 1-(monomethoxy-polyethyleneglycol)-2,3-dimyristoylglycerol, 4-O-(2',3'-di(tetradecanoyloxy)propyl-1-O-(ω-methoxy(polyethoxy)ethyl)butanedioate (PEG-S-DMG), ω-methoxy(polyethoxy)ethyl-N-(2,3-di(tetradecanoxo)propyl)carbamate, 2,3-di(tetradecanoxo)propyl-N-(ω-methoxy(polyethoxy)ethyl)carbamate, and any combination of the foregoing.

[0306] In some embodiments, the pegylated lipid of the LNP of the disclosure is selected from mPEG2000-1,2-di-O-alkyl-sn3-carbomoylglyceride (PEG-C-DOMG), 1-[8'-(1,2-dimyristoyl-3-propanoxy)-carboxamido-3',6'-dioxaoctanyl]carbomoyl-w-methyl-poly(ethylene glycol) (2 KPEG-DMG), and any combination of the foregoing.

[0307] In some embodiments, the PEG is directly attached to the lipid of the pegylated lipid. In other embodiments, the PEG is attached to the lipid of the pegylated lipid by a linker moiety selected from an ester-free linker moiety or an ester-containing linker moiety. Non-limiting examples of the ester-free linker moiety include amido (-C(O)NH-), amino (-NR-), carbonyl (-C(O)-), carbamate (-NHC(O)O-), urea (-NHC(O)NH-), disulfide (-S-S-), ether (-O-), succinyl (-O)CCH₂CH₂C(O)-, succinamidyl (-NHC(O)CH₂CH₂C(O)NH-), ether, disulfide and combinations thereof. For example, the linker may contain a carbamate linker

moiety and an amido linker moiety. Non-limiting examples of the ester-containing linker moiety include carbonate (-OC(O)O-), succinoyl, phosphate ester (-O-(O)POH-O-), sulfonate ester, and combinations thereof.

[0308] The PEG moiety of the pegylated lipid of the LNP of the disclosure described herein may have an average molecular weight ranging from about 550 daltons to about 10,000 daltons. In certain embodiments, the PEG moiety has an average molecular weight of from about 750 daltons to about 5,000 daltons, about 1,000 daltons to about 4,000 daltons, about 1,500 daltons to about 3,000 daltons, about 750 daltons to about 3,000 daltons, or about 1750 daltons to about 2,000 daltons.

[0309] In some embodiments, the conjugated lipid (*e.g.*, pegylated lipid) comprises from about 1 mol % to about 60 mol %, from about 2 mol % to about 50 mol %, from about 5 mol % to about 40 mol %, or from about 5 mol % to about 20 mol % of the total lipid present in the LNPs and/or LNP compositions. In certain embodiments, the conjugated lipid comprises from about 0.5 mol % to about 3 mol % of the total lipid present in the particle.

[0310] In additional embodiments, the conjugated lipid (*e.g.*, pegylated lipid) of the LNP of the disclosure comprises at least about 1, 2, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, or 60 mol %, or an intermediate range of any of the foregoing, of the total lipid present in the LNPs and/or LNP compositions.

[0311] For the lipid in the lipid-PEG conjugate of the LNP of the disclosure, any lipid capable of binding to polyethyleneglycol may be used without limitation, and the phospholipid and/or cholesterol which are other elements of the LNP may be also used. In some embodiments, the lipid in the lipid-PEG conjugate may be ceramide, dimyristoylglycerol (DMG), succinoyl-diacylglycerol (s-DAG), distearoylphosphatidylcholine (DSPC), distearoylphosphatidylethanolamine (DSPE), or cholesterol, but not limited thereto.

[0312] In the lipid-PEG conjugate of the LNP of the disclosure, the PEG may be directly conjugated to the lipid or linked to the lipid via a linker moiety. Any linker moiety suitable for binding PEG to the lipid may be used, and for example, includes an ester-free linker moiety and an ester-containing linker moiety. The ester-free linker moiety includes amido (-C(O)NH-), amino (-NR-), carbonyl (-C(O)-), carbamate (-NHC(O)O-), urea (-NHC(O)NH-), disulfide (-S-S-), ether (-O-), succinyl (-O)CCH₂CH₂C(O)-, succinamidyl (-NHC(O)CH₂CH₂C(O)NH-), ether, disulfide and combinations thereof (for example, a linker containing both a carbamate linker moiety and an amido linker moiety), but not limited

thereto. The ester-containing linker moiety includes for example, carbonate (-OC(O)O-), succinoyl, phosphate ester (-O-(O)POH-O-), sulfonate ester, and combinations thereof, but not limited thereto.

Steroids

[0313] In some embodiments, the LNPs and LNP compositions of the present disclosure include at least one steroid or derivative thereof. In some embodiments, the steroid comprises cholesterol. In some embodiments, the LNPs and LNP compositions comprise a cholesterol derivative selected from cholestanol, cholestanone, cholestenone, coprostanol, cholesteryl-2'-hydroxyethyl ether, cholesteryl-4'-hydroxybutyl ether, and any combination of the foregoing.

[0314] In some embodiments, the steroid (*e.g.*, cholesterol) of the LNP of the disclosure comprises from about 1 mol % to about 60 mol %, from about 2 mol % to about 50 mol %, from about 5 mol % to about 40 mol %, or from about 5 mol % to about 20 mol % of the total lipid present in the LNPs and/or LNP compositions. In other embodiments, the steroid (*e.g.*, cholesterol) of the LNP of the disclosure comprises at least about 1, 2, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, or 60 mol %, or an intermediate range of any of the foregoing, of the total lipid present in the LNPs and/or LNP compositions.

Additional Lipid

[0315] In some embodiments, the LNPs and LNP compositions of the present disclosure include at least one additional lipid. In some embodiments, the additional lipid is non-cationic lipid selected from an anionic lipid, a neutral lipid, or both. In some embodiments, the additional lipid comprises at least one phospholipid. In some embodiments, the phospholipid is selected from an anionic phospholipid, a neutral phospholipid, or both. The phospholipid of the elements of the LNPs and LNP compositions can play a role in covering and protecting a core of the LNP formed by interaction of the cationic lipid and nucleic acid in the LNP, and may facilitate cell membrane permeation and endosomal escape during intracellular delivery of the nucleic acid by binding to the phospholipid bilayer of a target cell. A phospholipid which can promote fusion of the LNP to a cell may include without limitation, any of the phospholipids selected from the group described below.

[0316] In some embodiments, the LNPs and LNP compositions comprise at least one phospholipid selected from, but not limited to, dipalmitoyl-phosphatidylcholine (DPPC), distearoyl-phosphatidylcholine (DSPC), dioleoyl-phosphatidylethanolamine (DOPE), dioleoyl-phosphatidylcholine (DOPC), dioleoyl-phosphatidylglycerol (DOPG),

palmitoyl-oleoyl-phosphatidylcholine (POPC), palmitoyl-oleoyl-phosphatidylethanolamine (POPE), palmitoyl-oleoyl-phosphatidylglycerol (POPG), dipalmitoyl-phosphatidylethanolamine (DPPE), dipalmitoyl-phosphatidylglycerol (DPPG), dimyristoyl-phosphatidylethanolamine (DMPE), distearoyl-phosphatidylethanolamine (DSPE), monomethyl-phosphatidylethanolamine, dimethyl-phosphatidylethanolamine, dielaidoyl-phosphatidylethanolamine (DEPE), stearoyl-oleoyl-phosphatidylethanolamine (SOPE), egg phosphatidylcholine (EPC), phosphatidylethanolamine (PE), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-dioleoyl-sn-glycero-3-[phospho-L-serine] (DOPS), 1,2-dioleoyl-sn-glycero-3-[phospho-L-serine], and any combination of the foregoing. In one example, the LNP comprising DOPE may be effective in mRNA delivery (excellent drug delivery efficacy).

[0317] In some embodiments, the additional lipid (*e.g.*, phospholipid) of the LNP of the disclosure comprises from about 1 mol % to about 60 mol %, from about 2 mol % to about 50 mol %, from about 5 mol % to about 40 mol %, or from about 5 mol % to about 20 mol % of the total lipid present in the LNPs and/or LNP compositions. In other embodiments, the additional lipid (*e.g.*, phospholipid) of the LNP of the disclosure comprises at least about 1, 2, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, or 60 mol %, or an intermediate range of any of the foregoing, of the total lipid present in the LNPs and/or LNP compositions.

[0318] It will be appreciated that the total lipid present in the LNPs and/or LNP compositions comprises the combination of the cationic lipid or ionizable cationic lipid, the conjugated lipid, (*e.g.*, pegylated lipid), the steroid (*e.g.*, cholesterol), and the additional lipid (*e.g.*, phospholipid).

[0319] The LNPs and/or LNP compositions may be prepared by dissolving the total lipids (or a portion thereof) in an organic solvent (*e.g.*, ethanol) followed by mixing through a micromixer with the payload (*e.g.*, nucleic acids of the systems) dissolved in an acidic buffer (*e.g.*, pH 4). At this pH the cationic lipid is positively charged and interacts with the negatively-charged nucleic acid polymers. The resulting nanostructures containing the nucleic acids are then converted to neutral LNPs when dialyzed against a neutral buffer which may then be followed by removal of the organic solvent (*e.g.*, ethanol) and exchange the LNPs into physiologically relevant buffer. The LNPs and/or LNP compositions thus formed have a distinct electron-dense nanostructured core where the cationic lipids are organized into inverted micelles around the encapsulated payload, as opposed to traditional

bilayer liposomal structures. In another embodiment, the LNP may form a bleb-like structure with nucleic acids in aqueous pockets along the non-electron dense lipid core.

b. Lipid nanoparticle properties

[0320] In some embodiments, the LNPs and/or LNP compositions comprise from about 50 mol % to about 85 mol % of the cationic lipid or ionizable cationic lipid, from about 0.5 mol % to about 10 mol % of the conjugated lipid, (*e.g.*, pegylated lipid), from about 0.5 mol % to about 10 mol % of the steroid (*e.g.*, cholesterol) and from about 5 mol % to about 50 mol % of the additional lipid (*e.g.*, phospholipid). In some embodiments, the LNPs and/or LNP compositions comprise from about 50 mol % to about 85 mol % of the cationic lipid or ionizable cationic lipid, from about 0.5 mol % to about 5 mol % of the conjugated lipid, (*e.g.*, pegylated lipid), from about 0.5 mol % to about 5 mol % of the steroid (*e.g.*, cholesterol) and from about 5 mol % to about 20 mol % of the additional lipid (*e.g.*, phospholipid).

[0321] In some embodiments, the LNPs and/or LNP compositions of the disclosure comprise cationic lipid : additional lipid (*e.g.*, phospholipid) : steroid (*e.g.*, cholesterol) : conjugated lipid, (*e.g.*, pegylated lipid) at a molar ratio of 20 to 50:10 to 30:30 to 60:0.5 to 5, at a molar ratio of 25 to 45:10 to 25:40 to 50:0.5 to 3, at a molar ratio of 25 to 45:10 to 20:40 to 55:0.5 to 3, or at a molar ratio of 25 to 45:10 to 20:40 to 55:1.0 to 1.5.

[0322] In some embodiments, the LNPs and/or LNP compositions of the disclosure have a total lipid : payload ratio (mass/mass) of from about 1 to about 100. In some embodiments, the total lipid : payload ratio is about 1 to about 50, from about 2 to about 25, from about 3 to about 20, from about 4 to about 15, or from about 5 to about 10. In some embodiments, the total lipid : payload ratio is about 5 to about 15, *e.g.*, about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or an intermediate range of any of the foregoing.

[0323] In certain embodiments, the LNPs of the disclosure comprise a total lipid: nucleic acid mass ratio of from about 5:1 to about 15:1. In some embodiments, the weight ratio of the cationic lipid and nucleic acid comprised in the LNP may be 1 to 20:1, 1 to 15:1, 1 to 10:1, 5 to 20:1, 5 to 15:1, 5 to 10:1, 7.5 to 20:1, 7.5 to 15:1, or 7.5 to 10:1.

[0324] In some embodiments, the LNP of the disclosure may comprise the cationic lipid of 20 to 50 parts by weight, the phospholipid of 10 to 30 parts by weight, cholesterol of 20 to 60 parts by weight (or 20 to 60 parts by weight), and lipid-PEG conjugate of 0.1 to 10 parts by weight (or 0.25 to 10 parts by weight, 0.5 to 5 parts by weight). Alternatively, the LNP may comprise the cationic lipid of 20 to 50 % by weight, phospholipid of 10 to 30 % by weight, cholesterol of 20 to 60 % by weight (or 30 to 60 % by weight), and lipid-PEG conjugate of

0.1 to 10 % by weight (or 0.25 to 10 % by weight, 0.5 to 5 % by weight) based on the total nanoparticle weight. As a further alternative, the LNP may comprise the cationic lipid of 25 to 50 % by weight, phospholipid of 10 to 20 % by weight, cholesterol of 35 to 55 % by weight, and lipid-PEG conjugate of 0.1 to 10 % by weight (or 0.25 to 10 % by weight, 0.5 to 5 % by weight), based on the total nanoparticle weight.

[0325] In some embodiments, the LNPs of the present disclosure have a mean diameter of from about 20 to 200 nm, 20 to 180 nm, 20 to 170 nm, 20 to 150 nm, 20 to 120 nm, 20 to 100 nm, 20 to 90 nm, 30 to 200 nm, 30 to 180 nm, 30 to 170 nm, 30 to 150 nm, 30 to 120 nm, 30 to 100 nm, 30 to 90 nm, 40 to 200 nm, 40 to 180 nm, 40 to 170 nm, 40 to 150 nm, 40 to 120 nm, 40 to 100 nm, 40 to 90 nm, 40 to 80 nm, 40 to 70 nm, 50 to 200 nm, 50 to 180 nm, 50 to 170 nm, 50 to 150 nm, 50 to 120 nm, 50 to 100 nm, 50 to 90 nm, 60 to 200 nm, 60 to 180 nm, 60 to 170 nm, 60 to 150 nm, 60 to 120 nm, 60 to 100 nm, 60 to 90 nm, 70 to 200 nm, 70 to 180 nm, 70 to 170 nm, 70 to 150 nm, 70 to 120 nm, 70 to 100 nm, 70 to 90 nm, 80 to 200 nm, 80 to 180 nm, 80 to 170 nm, 80 to 150 nm, 80 to 120 nm, 80 to 100 nm, 80 to 90 nm, 90 to 200 nm, 90 to 180 nm, 90 to 170 nm, 90 to 150 nm, 90 to 120 nm, or 90 to 100 nm, or an intermediate range of any of the foregoing.

[0326] In some embodiments, the LNPs and/or LNP compositions of the disclosure have a positive charge at acidic pH and may encapsulate the payload (*e.g.*, therapeutic agent, such as the LTRP:gRNA systems, or polynucleotides encoding same) through electrostatic charges produced by negative charges of the payload (*e.g.*, therapeutic agent). The term “encapsulation,” refers to the mixture of lipids surrounding and embedding the payload (*e.g.*, therapeutic agent) at physiological conditions, forming the LNPs. The term “encapsulation efficiency,” as used herein is the amount of payload (*e.g.*, therapeutic agent) encapsulated by the LNPs divided by the total amount of payload (*e.g.*, therapeutic agent) used to load the payload (*e.g.*, therapeutic agent) into the LNPs. The encapsulation efficiency of the LNPs and/or LNP compositions may be 70% or more, 75% or more, 80% or more, 85% or more, 90% or more, 91% or more, 92% or more, 94% or more, or 95% or more. In other embodiments, the encapsulation efficiency of the LNPs and/or LNP compositions is about 80% to 99%, about 85% to 98%, about 88% to 95%, about 90% to 95%, or the payload (*e.g.*, nucleic acids of the systems) may be fully encapsulated within the lipid portion of the LNPs compositions, and thereby protected from enzymatic degradation. In some embodiments, the payload (*e.g.*, therapeutic agent) is not substantially degraded after exposure of the LNPs and/or LNP compositions to a nuclease at 37° C. for at least about 20, 30, 45, or 60 minutes

or at least about 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, or 36 hours. In some embodiments, the payload (*e.g.*, nucleic acids of the systems) is complexed with the lipid portion of the LNPs and/or LNP compositions. The LNPs and/or LNP compositions of the present disclosure are non-toxic to mammals such as humans.

[0327] The term “fully encapsulated” indicates that the payload (*e.g.*, the nucleic acids of the system) in the LNPs and/or LNP compositions is not significantly degraded after exposure to conditions that degrade free DNA, RNA, or protein. In a fully encapsulated system, less than about 25%, more preferably less than about 10%, and most preferably less than about 5% of the payload (*e.g.*, nucleic acids of the system) in the LNPs and/or LNP compositions is degraded by conditions that would degrade 100% of a non-encapsulated payload. “Fully encapsulated” also indicates that the LNPs and/or LNP compositions are serum-stable, and do not decompose into their component parts upon in vivo administration.

[0328] In some embodiments, the amount of the LNPs and/or LNP compositions having the payload (*e.g.*, therapeutic agent), encapsulated therein is from about 30% to about 100%, from about 40% to about 100%, from about 50% to about 100%, from about 60% to about 100%, from about 70% to about 100%, from about 80% to about 100%, from about 90% to about 100%, from about 30% to about 95%, from about 40% to about 95%, from about 50% to about 95%, from about 60% to about 95%, %, from about 70% to about 95%, from about 80% to about 95%, from about 85% to about 95%, from about 90% to about 95%, from about 30% to about 90%, from about 40% to about 90%, from about 50% to about 90%, from about 60% to about 90%, from about 70% to about 90%, from about 80% to about 90%, or at least about 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or an intermediate range of any of the foregoing.

[0329] In some embodiments, the amount of the payload (*e.g.*, the nucleic acids), encapsulated within the LNPs and/or LNP compositions is from about 30% to about 100%, from about 40% to about 100%, from about 50% to about 100%, from about 60% to about 100%, from about 70% to about 100%, from about 80% to about 100%, from about 90% to about 100%, from about 30% to about 95%, from about 40% to about 95%, from about 50% to about 95%, from about 60% to about 95%, %, from about 70% to about 95%, from about 80% to about 95%, from about 85% to about 95%, from about 90% to about 95%, from about 30% to about 90%, from about 40% to about 90%, from about 50% to about 90%, from about 60% to about 90%, from about 70% to about 90%, from about 80% to about 90%, or at least

about 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or an intermediate range of any of the foregoing.

[0330] In some embodiments, the nucleic acids of the disclosure, such as the mRNA encoding the long-term repressor fusion protein, and/or the gRNA, may be provided in a solution to be mixed with a lipid solution such that the nucleic acids may be encapsulated in the lipid nanoparticles. A suitable nucleic acid solution may be any aqueous solution containing the nucleic acid to be encapsulated at various concentrations. For example, a suitable nucleic acid solution may contain the nucleic acid (or nucleic acids) at a concentration of or greater than about 0.01 mg/ml, 0.05 mg/ml, 0.06 mg/ml, 0.07 mg/ml, 0.08 mg/ml, 0.09 mg/ml, 0.1 mg/ml, 0.15 mg/ml, 0.2 mg/ml, 0.3 mg/ml, 0.4 mg/ml, 0.5 mg/ml, 0.6 mg/ml, 0.7 mg/ml, 0.8 mg/ml, 0.9 mg/ml, 1.0 mg/ml, 1.25 mg/ml, 1.5 mg/ml, 1.75 mg/ml, or 2.0 mg/ml. In some embodiments, the nucleic acid comprises an mRNA encoding an long-term repressor fusion protein, and a suitable mRNA solution may contain the mRNA at a concentration ranging from about 0.01-2.0 mg/ml, 0.01-1.5 mg/ml, 0.01-1.25 mg/ml, 0.01-1.0 mg/ml, 0.01-0.9 mg/ml, 0.01-0.8 mg/ml, 0.01-0.7 mg/ml, 0.01-0.6 mg/ml, 0.01-0.5 mg/ml, 0.01-0.4 mg/ml, 0.01-0.3 mg/ml, 0.01-0.2 mg/ml, 0.01-0.1 mg/ml, 0.05-1.0 mg/ml, 0.05-0.9 mg/ml, 0.05-0.8 mg/ml, 0.05-0.7 mg/ml, 0.05-0.6 mg/ml, 0.05-0.5 mg/ml, 0.05-0.4 mg/ml, 0.05-0.3 mg/ml, 0.05-0.2 mg/ml, 0.05-0.1 mg/ml, 0.1-1.0 mg/ml, 0.2-0.9 mg/ml, 0.3-0.8 mg/ml, 0.4-0.7 mg/ml, or 0.5-0.6 mg/ml. In some embodiments, a suitable mRNA solution may contain an mRNA at a concentration up to about 5.0 mg/ml, 4.0 mg/ml, 3.0 mg/ml, 2.0 mg/ml, 1.0 mg/ml, 0.9 mg/ml, 0.8 mg/ml, 0.7 mg/ml, 0.6 mg/ml, 0.5 mg/ml, 0.4 mg/ml, 0.3 mg/ml, 0.2 mg/ml, 0.1 mg/ml, 0.05 mg/ml, 0.04 mg/ml, 0.03 mg/ml, 0.02 mg/ml, 0.01 mg/ml, or 0.05 mg/ml. In some embodiments, a suitable gRNA solution may contain an gRNA at a concentration up to about 5.0 mg/ml, 4.0 mg/ml, 3.0 mg/ml, 2.0 mg/ml, 1.0 mg/ml, 0.9 mg/ml, 0.8 mg/ml, 0.7 mg/ml, 0.6 mg/ml, 0.5 mg/ml, 0.4 mg/ml, 0.3 mg/ml, 0.2 mg/ml, 0.1 mg/ml, 0.05 mg/ml, 0.04 mg/ml, 0.03 mg/ml, 0.02 mg/ml, 0.01 mg/ml, or 0.05 mg/ml.

[0331] In some embodiments, the LNP may have an average diameter of 20nm to 200nm, 20 to 180nm, 20nm to 170nm, 20nm to 150nm, 20nm to 120nm, 20nm to 100nm, 20nm to 90nm, 30nm to 200nm, 30 to 180nm, 30nm to 170nm, 30nm to 150nm, 30nm to 120nm, 30nm to 100nm, 30nm to 90nm, 40nm to 200nm, 40 to 180nm, 40nm to 170nm, 40nm to 150nm, 40nm to 120nm, 40nm to 100nm, 40nm to 90nm, 40nm to 80nm, 40nm to 70nm, 50nm to 200nm, 50 to 180nm, 50nm to 170nm, 50nm to 150nm, 50nm to 120nm, 50nm to

100nm, 50nm to 90nm, 60nm to 200nm, 60 to 180nm, 60nm to 170nm, 60nm to 150nm, 60nm to 120nm, 60nm to 100nm, 60nm to 90nm, 70nm to 200nm, 70 to 180nm, 70nm to 170nm, 70nm to 150nm, 70nm to 120nm, 70nm to 100nm, 70nm to 90nm, 80nm to 200nm, 80 to 180nm, 80nm to 170nm, 80nm to 150nm, 80nm to 120nm, 80nm to 100nm, 80nm to 90nm, 90nm to 200nm, 90 to 180nm, 90nm to 170nm, 90nm to 150nm, 90nm to 120nm, or 90nm to 100nm for easy introduction into liver tissue, hepatocytes and/or LSEC (liver sinusoidal endothelial cells). The LNP may be sized for easy introduction into organs or tissues, including but not limited to liver, lung, heart, spleen, as well as to tumors. When the size of the LNP is smaller than the above range, it can be difficult to maintain stability as the surface area of the LNP is excessively increased, and thus delivery to the target tissue and/or drug effect may be reduced. The LNP may specifically target liver tissue. Without wishing to be bound by theory, it is thought that one mechanism by which LNP may be used to deliver therapeutic agents is through the imitation of the metabolic behaviors of natural lipoproteins, and so LNP may be usefully delivered to a subject through the lipid metabolism processes carried out by the liver. During the delivery of therapeutic agents to hepatocytes or and/or LSEC (liver sinusoidal endothelial cells), the diameter of the fenestrae leading from the sinusoidal lumen to the hepatocytes and LSEC is about 140 nm in mammals and about 100 nm in humans, so the LNP composition for therapeutic agent delivery having LNPs with a diameter in the above ranges may have excellent delivery efficiency to hepatocytes and LSEC when compared to LNP having the diameter outside the above range.

[0332] According to one example, the LNPs of the LNP composition may comprise the cationic lipid : phospholipid : cholesterol : lipid-PEG conjugate in the range described above or at a molar ratio of 20 to 50:10 to 30:30 to 60:0.5 to 5, at a molar ratio of 25 to 45:10 to 25:40 to 50:0.5 to 3, at a molar ratio of 25 to 45:10 to 20:40 to 55:0.5 to 3, or at a molar ratio of 25 to 45:10 to 20:40 to 55:1.0 to 1.5. The LNP comprising components at a molar ratio in the above range may have excellent delivery efficiency of therapeutic agents specific to cells of target organs.

[0333] In certain aspects, the LNP exhibit a positive charge under the acidic pH condition by showing a pKa of 5 to 8, 5.5 to 7.5, 6 to 7, or 6.5 to 7, and may encapsulate a nucleic acid with high efficiency by easily forming a complex with a nucleic acid through electrostatic interaction with a therapeutic agent such as a nucleic acid showing a negative charge. In such cases, the LNP may be usefully used as a composition for intracellular or *in vivo* delivery of a therapeutic agent (for example, nucleic acid).

[0334] Herein, "encapsulate" or "encapsulation" refers to incorporation of a therapeutic agent efficient delivery, *i.e.*, by surrounding it by the particle surface and/or embedding it within the particle interior. The encapsulation efficiency means the content of the therapeutic agent encapsulated in the LNP relative the total therapeutic agent content used for preparation of the LNP.

[0335] The encapsulation of the nucleic acids of the composition in the LNP may be 70% or more, 75% or more, 80% or more, 85% or more, 90% or more, 91% or more, 92% or more, 94% or more, or 95% or more of LNP in the composition encapsulate nucleic acids. In some embodiments, the encapsulation of the nucleic acids of the composition in the LNP is such that between 80% to 99%, between 80% to 97%, between 80% to 95%, between 85% to 95%, between 87% to 95%, between 90% to 95%, between 91% or more to 95% or less, 91% or more to 94% or less, over 91% to 95% or less, 92% to 99%, between 92% to 97%, or between 92% to 95% of the LNP in the composition encapsulate nucleic acids. In some embodiments, the mRNA encoding the long-term repressor fusion protein and a gRNA of any of the embodiments of the disclosure are fully encapsulated in the LNP.

[0336] The target organs to which a nucleic acid is delivered by the LNP include, but are not limited to the liver, lung, heart, spleen, as well as to tumors. The LNP according to one example is liver tissue-specific and has excellent biocompatibility and can deliver the nucleic acids of a composition with high efficiency, and thus it can be usefully used in related technical fields such as lipid nanoparticle-mediated gene therapy. In a particular embodiment, the target cell to which the nucleic acids are delivered by the LNP according to one example may be a hepatocyte and/or LSEC *in vivo*. In other embodiments, the disclosure provides LNP formulated for delivery of the nucleic acids of the embodiments to cells *ex vivo*.

[0337] The disclosure provides a pharmaceutical composition comprising a plurality of LNPs comprising nucleic acids, such as mRNA encoding a long-term repressor fusion protein and/or a gRNA variant described herein, and a pharmaceutically acceptable carrier.

[0338] In certain embodiments, the LNP comprising the nucleic acid(s) has an electron dense core.

[0339] The disclosure provides LNP comprising one or more nucleic acids comprising: (a) an mRNA encoding the long-term repressor fusion protein, and/or a gRNA variant described herein; (b) one or more cationic lipids or ionizable cationic lipids or salts thereof comprising from about 50 mol % to about 85 mol % of the total lipid present in the LNP; (c) one or more non-cationic lipids comprising from about 13 mol % to about 49.5 mol % of the total lipid

present in the LNP; and (d) one or more conjugated lipids that inhibit aggregation of LNPs comprising from about 0.5 mol % to about 2 mol % of the total lipid present in the particle. In another embodiment, the disclosure provides LNP comprising one or more nucleic acids comprising: (a) an mRNA encoding the long-term repressor fusion protein, and/or a gRNA variant described herein; (b) one or more cationic lipids or ionizable cationic lipids or salts thereof comprising from about 22 mol % to about 85 mol % of the total lipid present in the LNP; (c) one or more non-cationic/phospholipids comprising from about 10 mol % to about 70 mol % of the total lipid present in the LNP; (d) 15 mol % to about 50 mol % sterol, and (d) 1 mol % to about 5 mol % lipid-PEG or lipid-PEG-peptide in the particle. In certain embodiments the long-term repressor fusion protein mRNA and gRNA may be present in the same nucleic acid-lipid particle, or they may be present in different nucleic acid-lipid particles.

[0340] The disclosure provides LNP comprising one or more nucleic acids comprising: (a) an mRNA encoding the long-term repressor fusion protein described herein; (b) a cationic lipid or a salt thereof comprising from about 52 mol % to about 62 mol % of the total lipid present in the LNP; (c) a mixture of a phospholipid and cholesterol or a derivative thereof comprising from about 36 mol % to about 47 mol % of the total lipid present in the LNP; and (d) a PEG-lipid conjugate comprising from about 1 mol % to about 2 mol % of the total lipid present in the LNP. In particular embodiments, the formulation is a four-component system comprising about 1.4 mol % PEG-lipid conjugate (*e.g.*, PEG2000-C-DMA), about 57.1 mol % cationic lipid (*e.g.*, DLin-K-C2-DMA) or a salt thereof, about 7.1 mol % DPPC (or DSPC), and about 34.3 mol % cholesterol (or derivative thereof). In some embodiments, the LNP comprises the mRNA encoding the CasX described herein and a gRNA.

[0341] In other embodiments, the LNP comprising one or more nucleic acids comprises: (a) an mRNA encoding the long-term repressor fusion protein and/or a gRNA of any of the embodiments described herein; (b) a cationic lipid or a salt thereof comprising from about 46.5 mol % to about 66.5 mol % of the total lipid present in the LNP; (c) cholesterol or a derivative thereof comprising from about 31.5 mol % to about 42.5 mol % of the total lipid present in the LNP; and (d) a PEG-lipid conjugate comprising from about 1 mol % to about 2 mol % of the total lipid present in the LNP. In particular embodiments, the formulation is a three-component system which is phospholipid-free and comprises about 1.5 mol % PEG-lipid conjugate (*e.g.*, PEG2000-C-DMA), about 61.5 mol % cationic lipid (*e.g.*, DLin-K-C2-DMA) or a salt thereof, and about 36.9 mol % cholesterol (or derivative thereof). In some

embodiments, the LNP comprises the mRNA encoding the CasX described herein and a gRNA.

[0342] Additional formulations are described in PCT Publication No. WO 2009/127060 and US patent publication numbers US 2011/0071208 A1 and US 2011/0076335 A1, the disclosures of which are herein incorporated by reference in their entirety.

[0343] In other embodiments, the LNP comprising one or more nucleic acids comprises: (a) an mRNA encoding the long-term repressor fusion protein and a gRNA of any of the embodiments described herein; (b) one or more cationic lipid or ionizable cationic lipids or salts thereof comprising from about 2 mol % to about 50 mol % of the total lipid present in the LNP; (c) one or more non-cationic lipid or ionizable cationic lipids comprising from about 5 mol % to about 90 mol % of the total lipid present in the LNP; and (d) one or more conjugated lipids that inhibit aggregation of particles comprising from about 0.5 mol % to about 20 mol % of the total lipid present in the LNP. In some embodiments, the LNP comprises the mRNA encoding the CasX described herein and a gRNA.

[0344] In other embodiments, the LNP comprising one or more nucleic acids comprises: (a) an mRNA encoding the long-term repressor fusion protein and a gRNA of any of the embodiments described herein; (b) a cationic lipid or a salt thereof comprising from about 30 mol % to about 50 mol % of the total lipid present in the LNP; (c) a mixture of a phospholipid and cholesterol or a derivative thereof comprising from about 47 mol % to about 69 mol % of the total lipid present in the LNP; and (d) a PEG-lipid conjugate comprising from about 1 mol % to about 3 mol % of the total lipid present in the LNP. In particular embodiments, the formulation is a four-component system which comprises about 2 mol % PEG-lipid conjugate (*e.g.*, PEG2000-C-DMA), about 40 mol % cationic lipid (*e.g.*, DLin-K-C2-DMA) or a salt thereof, about 10 mol % DPPC (or DSPC), and about 48 mol % cholesterol (or derivative thereof). In some embodiments, the LNP comprises the mRNA encoding the CasX described herein and a gRNA.

[0345] In other embodiments, the LNP comprising one or more nucleic acids comprises: (a) an mRNA encoding the long-term repressor fusion protein and a gRNA of any of the embodiments described herein; (b) one or more cationic lipid or ionizable cationic lipids or salts thereof comprising from about 50 mol % to about 65 mol % of the total lipid present in the LNP; (c) one or more non-cationic lipid or ionizable cationic lipids comprising from about 25 mol % to about 45 mol % of the total lipid present in the LNP; and (d) one or more

conjugated lipids that inhibit aggregation of particles comprising from about 5 mol % to about 10 mol % of the total lipid present in the LNP.

[0346] In other embodiments, the LNP comprising one or more nucleic acids comprises: (a) an mRNA encoding the long-term repressor fusion protein and a gRNA of any of the embodiments described herein; (b) a cationic lipid or a salt thereof comprising from about 50 mol % to about 60 mol % of the total lipid present in the LNP; (c) a mixture of a phospholipid and cholesterol or a derivative thereof comprising from about 35 mol % to about 45 mol % of the total lipid present in the LNP; and (d) a PEG-lipid conjugate comprising from about 5 mol % to about 10 mol % of the total lipid present in the LNP. In some embodiments, the LNP comprises the mRNA encoding the CasX described herein and a gRNA.

[0347] In certain embodiments, the non-cationic lipid mixture in the formulation comprises: (i) a phospholipid of from about 10 mol % to about 70 mol % of the total lipid present in the LNP; (ii) cholesterol or a derivative thereof of from about 15 mol % to about 50 mol % of the total lipid present in the LNP; and 1-5% lipid-PEG or lipid-PEG-peptide. In particular embodiments, the formulation is a four-component system which comprises about 7 mol % PEG-lipid conjugate (*e.g.*, PEG750-C-DMA), about 54 mol % cationic lipid (*e.g.*, DLin-K-C2-DMA) or a salt thereof, about 7 mol % DPPC (or DSPC), and about 32 mol % cholesterol (or derivative thereof).

[0348] In other embodiments, the LNP comprising one or more nucleic acids comprises: (a) an mRNA encoding the long-term repressor fusion protein and/or a gRNA of any of the embodiments described herein; (b) a cationic lipid or a salt thereof comprising from about 55 mol % to about 65 mol % of the total lipid present in the LNP; (c) cholesterol or a derivative thereof comprising from about 30 mol % to about 40 mol % of the total lipid present in the LNP; and (d) a PEG-lipid conjugate comprising from about 5 mol % to about 10 mol % of the total lipid present in the LNP. In particular embodiments, the formulation is a three-component system which is phospholipid-free and comprises about 7 mol % PEG-lipid conjugate (*e.g.*, PEG750-C-DMA), about 58 mol % cationic lipid (*e.g.*, DLin-K-C2-DMA) or a salt thereof, and about 35 mol % cholesterol (or derivative thereof). In some embodiments, the LNP comprises the mRNA encoding the CasX described herein and a gRNA.

[0349] In other embodiments, the LNP comprising one or more nucleic acids comprises: (a) an mRNA encoding the long-term repressor fusion protein and/or a gRNA of any of the embodiments described herein; (b) a cationic lipid or a salt thereof comprising from about 48

mol % to about 62 mol % of the total lipid present in the LNP; (c) a mixture of a phospholipid and cholesterol or a derivative thereof, wherein the phospholipid comprises about 7 mol % to about 17 mol % of the total lipid present in the LNP, and wherein the cholesterol or derivative thereof comprises about 25 mol % to about 40 mol % of the total lipid present in the LNP; and (d) a PEG-lipid conjugate comprising from about 0.5 mol % to about 3.0 mol % of the total lipid present in the LNP. In some embodiments, the LNP comprises the mRNA encoding the CasX described herein and a gRNA.

IX. Methods for Repression of PCSK9 Target Nucleic Acids

[0350] In another aspect, the present disclosure relates to methods of repressing or silencing transcription of a target nucleic acid sequence of a *PCSK9* gene in a population of cells. In some embodiments, the method comprises contacting the cells of the population with an LTRP:gRNA system of the disclosure, either *in vitro*, *ex vivo*, or *in vivo* in a subject. The programmable nature of the systems provided herein allows for the precise targeting to achieve the desired effect at one or more regions of predetermined interest in a target nucleic acid of a *PCSK9* gene. In some embodiments, it may be desirable to repress or silence a gene in the cells comprising mutations that result in a PCSK9-related disease or disorder in a subject.

[0351] In some embodiments, the method comprises introducing into the cells a long-term repressor fusion protein of the disclosure and one or more gRNAs with a targeting sequence complementary to the *PCSK9* target nucleic acid, wherein the long-term repressor fusion protein is capable of complexing with the gRNA to form an RNP, and wherein the RNP is capable of binding the target nucleic acid and repressing or silencing the transcription of the *PCSK9* gene in the cells (it being understood that additional cellular factors may be recruited and participate in the repression). In some embodiments, the method comprises introducing into the cells an mRNA encoding a long-term repressor fusion protein of the disclosure and one or more gRNAs with a targeting sequence complementary to the *PCSK9* target nucleic acid, whereupon the long-term repressor fusion protein is expressed, is capable of complexing with the gRNA to form an RNP, and wherein the RNP is capable of binding the target nucleic acid and repressing or silencing the transcription of the *PCSK9* gene in the cells. In some embodiments, the mRNA encoding the long-term repressor fusion protein and gRNA may be co-formulated in a nanoparticle for delivery to the cells of the population. In some embodiments, the mRNA encoding the long-term repressor fusion protein and gRNA

may be formulated in separate nanoparticles for delivery to the cells of the population. In some embodiments, the nanoparticle is a lipid nanoparticle (LNP), as described herein.

[0352] In some embodiments of the method of transcriptional repression of a gene, the LTRP:gRNA systems of the present disclosure can be designed to target any region of, or proximal to, a *PCSK9* gene or region of a gene for which repression of transcription is sought. When the entirety of the *PCSK9* gene is to be repressed, use of a guide with a targeting sequence complementary to a sequence encompassing or proximal to the transcription start site (TSS) is contemplated by the disclosure. The core promoter serves as a binding platform for the transcription machinery, which comprises Pol II and its associated general transcription factors (GTFs) (Haberle, V. et al. Eukaryotic core promoters and the functional basis of transcription initiation (Nat Rev Mol Cell Biol. 19(10):621 (2018)). Variability in TSS selection has been proposed to involve DNA ‘scrunching’ and ‘anti-scrunching,’ the hallmarks of which are: (i) forward and reverse movement of the RNA polymerase leading edge, but not trailing edge, relative to DNA, and (ii) expansion and contraction of the transcription bubble. In some embodiments, the targeting sequence of the gRNA of the LTRP:gRNA system is complementary to a target nucleic acid sequence located within 1 kb of a transcription start site (TSS) in the *PCSK9* gene. In some embodiments of the method, the targeting sequence of the gRNA of the LTRP:gRNA system is complementary to a target nucleic acid sequence located within 20 bp, 50 bp, 100 bp, 150 bp, 200 bp, 250 bp, 500 bp, 1 kb, or 1.5 kb upstream of a TSS of the *PCSK9* gene. In some embodiments of the method, the targeting sequence of the gRNA of the LTRP:gRNA system is complementary to a target nucleic acid sequence located within 20 bp, 50 bp, 100 bp, 150 bp, 200 bp, 250 bp, 500 bp, 1 kb, or 1.5 kb downstream of a TSS of the *PCSK9* gene. In some embodiments of the method, the targeting sequence of the gRNA of the LTRP:gRNA system is complementary to a target nucleic acid sequence located within 700 bps upstream to 700 bps downstream, 500 bps upstream to 500 bps downstream, 300 bps upstream to 300 bps downstream, or 100 bps upstream to 100 bps downstream of a TSS of the *PCSK9* gene. In some embodiments of the method, the targeting sequence of the gRNA of the LTRP:gRNA system is complementary to a target nucleic acid sequence located within 20 bp, 50 bp, 100 bp, 150 bp, 200 bp, 250 bp, 500 bp, or 1 kb of an enhancer of the *PCSK9* gene. In some embodiments of the method, the targeting sequence of the gRNA of the LTRP:gRNA system is complementary to a target nucleic acid sequence located within 1 kb 3' to a 5' untranslated region of the *PCSK9* gene. In some embodiments of the method, the targeting sequence of the

gRNA of the LTRP:gRNA system is complementary to a target nucleic acid sequence located within the open reading frame of the *PCSK9* gene. In some embodiments of the method, the targeting sequence of the gRNA of the LTRP:gRNA system is complementary to a target nucleic acid sequence of an exon of the *PCSK9* gene. In a particular embodiment, the targeting sequence of a gRNA of the system of the disclosure is complementary to a target nucleic acid sequence for exon 1 of the *PCSK9* gene. In other embodiments of the method, the targeting sequence of a gRNA of the system of the disclosure is complementary to a target nucleic acid sequence for an intron of the *PCSK9* gene. In other embodiments of the method, the targeting sequence of the gRNA of the system of the disclosure is complementary to a target nucleic acid sequence for an intron-exon junction of the *PCSK9* gene. In other embodiments of the method, the targeting sequence of the gRNA of the system of the disclosure is complementary to a target nucleic acid sequence for a regulatory element of the *PCSK9* gene. In other embodiments of the method, the targeting sequence of the gRNA of the system of the disclosure is complementary to a sequence of an intergenic region of the *PCSK9* gene. In other embodiments of the method, the targeting sequence of a gRNA of the system of the disclosure is complementary to a junction of the exon, an intron, or a regulatory element of the *PCSK9* gene. In those cases where the targeting sequence is complementary to a regulatory element, such regulatory elements include, but are not limited to promoter regions, enhancer regions, intergenic regions, 5' untranslated regions (5' UTR), 3' untranslated regions (3' UTR), conserved elements, and regions comprising cis-regulatory elements. In some embodiments of the method, the targeting sequence of a gRNA of the system of the disclosure is complementary to the gene target nucleic acid sequence within 1 kb of an enhancer of the *PCSK9* gene. In some embodiments of the method, the targeting sequence of a gRNA of the system of the disclosure is complementary to the gene target nucleic acid sequence within the 3' untranslated region of the *PCSK9* gene. The promoter region is intended to encompass nucleotides within 5 kb of the initiation point of the encoding sequence or, in the case of gene enhancer elements or conserved elements, can be thousands of bp, hundreds of thousands of bp, or even millions of bp away from the encoding sequence of the *PCSK9* gene. In the foregoing, the targets are those in which the encoding *PCSK9* gene of the target is intended to be repressed and/or epigenetically modified such that a gene product is not expressed or is expressed at a lower level in a cell. In some embodiments, upon binding of the RNP of the system of the disclosure to the binding location of the target nucleic acid, the system is capable of repressing transcription of the

PCSK9 gene 5' to the binding location of the RNP. In other embodiments, upon binding of the RNP of the system to the binding location of the target nucleic acid, the system is capable of repressing transcription of the *PCSK9* gene 3' to the binding location of the RNP.

[0353] The disclosure provides methods of transcriptional repression or silencing a *PCSK9* target gene in a population of cells. In some embodiments, the method for transcriptional repression or silencing of the *PCSK9* gene comprises contacting the population of cells with an LTRP:gRNA system comprising an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6529-14626 (unmodified mRNA), and 14628-22725 (mRNA modified with N1-methyl-pseudouridine), or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto. In some embodiments, the method comprises contacting the population of cells with an LTRP:gRNA system comprising an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6529-8134, 9742-11347, 14628-16233, and 17840-19446, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% identity thereto. In some embodiments, the method comprises contacting the population of cells with an LTRP:gRNA system comprising an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6529-8134, 9742-11347, 14628-16233, and 17840-19446. In some embodiments, the method comprises contacting the population of cells with an LTRP:gRNA system comprising an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6529-6537 and 9742-9750, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% identity thereto. In some embodiments, the method comprises contacting the population of cells with an LTRP:gRNA system comprising an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6529-6537 and 9742-9750. In some embodiments, the mRNA comprises a 5' cap sequence linked 5' to the 5' UTR of the mRNA sequence encoding the LTRP, optionally wherein the 5' cap has the nucleic acid sequence selected from the group consisting of m7G(5')ppp(5')mAG, m7G(5')ppp(5')(A,G(5'))ppp(5')A and G(5')ppp(5')G. In some embodiments, the method comprises contacting the population of cells with an LTRP:gRNA system comprising a gRNA comprising a sequence selected from the group

consisting of SEQ ID NOS: 1744-1746, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto, and the gRNA comprises a linked targeting sequence complementary to a target nucleic acid of the *PCSK9* gene. In some embodiments of the method, the LTRP:gRNA system comprises gRNA variant comprising a sequence of SEQ ID NO: 1744 and a linked targeting sequence. In some embodiments of the method, the LTRP:gRNA system comprises gRNA comprising a sequence of SEQ ID NO: 1745 and a linked targeting sequence. In some embodiments, the targeting sequence of the gRNA is selected from the group consisting of SEQ ID NOS: 1824-2544, 2672, 2675, 2694, and 2714. In some embodiments, the targeting sequence of the gRNA is selected from the group consisting of SEQ ID NOS: 1824-1880, 1883, 1884, 1888, 1889, 2672, 2675, 2694, and 2714. In some embodiments, the targeting sequence of the gRNA is selected from the group consisting of SEQ ID NOS: 1834, 1849, 1853, 1855, 1856, 1857, 1858, 1860, 1862, 1863, 1867, 1869, 1870, 1872, 1874, and 1875. In some embodiments, the targeting sequence is selected from the group consisting of SEQ ID NOS: 1855, 1867, and 1869.

[0354] In some embodiments of the method of transcriptional repression or silencing a *PCSK9* target gene in a population of cells, the LTRP:gRNA system comprises gRNA variant comprising a sequence SEQ ID NO: 1746 and a linked targeting sequence. In some embodiments of the method, the LTRP:gRNA system comprises a gRNA variant comprising one or more chemical modifications, including gRNA variants comprising a sequence selected from the group consisting of SEQ ID NOS: 2948-2956, 2958-2966, and 2968-2976, with a targeting sequence complementary to a *PCSK9* target nucleic acid substituted for the 20 nucleotides on the 3' end of the gRNA of the listed sequences. In some embodiments of the method, the LTRP:gRNA system comprises a gRNA variant comprising a sequence of SEQ ID NO: 1746 with chemical modifications, optionally comprising a sequence selected from the group consisting of SEQ ID NOS: 2968-2976, with a targeting sequence complementary to a *PCSK9* target nucleic acid substituted for the 20 nucleotides on the 3' end of the gRNA of the listed sequences. In one embodiment of the method, the LTRP:gRNA system comprises gRNA variant comprising a sequence of SEQ ID NO: 1746 with chemical modifications, optionally comprising the sequence of SEQ ID NO: 2968, with a targeting sequence complementary to a *PCSK9* target nucleic acid substituted for the 20 nucleotides on the 3' end of the gRNA of the listed sequence. In some embodiments, the modified gRNA

comprises a sequence selected from the group consisting of SEQ ID NOS: 22788- 22803. In some embodiments, the modified gRNA comprises a sequence selected from the group consisting of SEQ ID NOS: 22788-22790.

[0355] In some embodiments, the method for transcriptional repression or silencing of the *PCSK9* gene in a population of cells comprises contacting the cells with an LNP comprising a gRNA variant and an mRNA encoding the long-term repressor fusion protein. In some embodiments of the method, the LNP comprises an mRNA sequence selected from the group consisting of SEQ ID NOS: 6529-14626, and 14628-22725, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto. In some embodiments of the method, the LNP comprises an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6529-14626, and 14628-22725. In some embodiments, the method comprises contacting the population of cells with an LNP comprising an mRNA selected from the group consisting of the sequences of SEQ ID NOS: 6529-8134, 9742-11347, 14628-16233, and 17840-19446, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the method comprises contacting the population of cells with an LNP comprising an mRNA comprising a sequence selected from the group consisting of the sequences of SEQ ID NOS: 6529-8134, 9742-11347, 14628-16233, and 17840-19446. In some embodiments, the method comprises contacting the population of cells with an LNP comprising an mRNA selected from the group consisting of the sequences of SEQ ID NOS: 6529-6537 and 9742-9750, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the method comprises contacting the population of cells with an LNP comprising an mRNA comprising a sequence selected from the group consisting of the sequences of SEQ ID NOS: 6529-6537 and 9742-9750. In some embodiments, the mRNA further comprises a 5' cap sequence linked 5' to the 5' UTR of the mRNA sequence encoding the LTRP, optionally wherein the 5' cap has the nucleic acid sequence selected from the group consisting of m7G(5')ppp(5')mAG, m7G(5')ppp(5')(A,G(5')ppp(5')A and G(5')ppp(5')G. In some embodiments, the mRNA comprises a 5' UTR, 3' UTR, and/or poly(A) sequence. In some embodiments of the method,

the LNP comprises a gRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 1744-1746, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto, and the gRNA comprises a targeting sequence complementary to a target nucleic acid of the *PCSK9* gene. In some embodiments of the method, the LNP comprises gRNA variant 174 (SEQ ID NO: 1744) with a linked targeting sequence. In some embodiments of the method, the LNP comprises gRNA comprising a sequence of SEQ ID NO: 1745 and a linked targeting sequence. In some embodiments of the method, the LNP comprises gRNA comprising a sequence of SEQ ID NO: 1746 and a linked targeting sequence. In some embodiments of the method, the LNP comprises a gRNA variant with chemical modifications, optionally comprising the sequences of SEQ ID NOS: 2948-2956; 2958-2966; and 2968-2976, with a targeting sequence complementary to a *PCSK9* target nucleic acid substituted for the 20 nucleotides on the 3' end of the gRNA of the listed sequences. In some embodiments of the method, the LNP comprises a gRNA with chemical modifications comprising a sequence selected from the group consisting of SEQ ID NOS: 2968-2976, with a targeting sequence complementary to a target nucleic acid substituted for the 20 nucleotides on the 3' end of the gRNA of the listed sequences. In a particular embodiment of the method, the LNP comprises a gRNA with chemical modifications comprising the sequence of SEQ ID NOS: 2968, with a targeting sequence complementary to a target nucleic acid substituted for the 20 nucleotides on the 3' end of the gRNA of the listed sequence. In some embodiments, the targeting sequence of the gRNA is selected from the group consisting of SEQ ID NOS: 1824-2544, 2672, 2675, 2694, and 2714. In some embodiments, the targeting sequence of the gRNA is selected from the group consisting of SEQ ID NOS: 1824-1879, 1880, 1883, 1884, 1888, 1889, 2672, 2675, 2694, and 2714. In some embodiments, the targeting sequence of the gRNA is selected from the group consisting of SEQ ID NOS: 1834, 1849, 1853, 1855, 1856, 1857, 1858, 1860, 1862, 1863, 1867, 1869, 1870, 1872, 1874, and 1875. In some embodiments, the targeting sequence of the gRNA is selected from the group consisting of SEQ ID NOS: 1855, 1867, and 1869. In some embodiments of the method, the LNP comprises a chemically modified gRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 22788-22803. In some embodiments, the LNP comprises a chemically modified gRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 22788-22790.

[0356] In some embodiments of the method for transcriptional repression or silencing of a *PCSK9* gene, contacting the cells of the population with an LTRP:gRNA system of the embodiments results in repression of the *PCSK9* target nucleic acid of at least about 1%, at least about 2%, at least about 3%, at least about 4%, at least about 5%, at least about 6%, at least about 7%, at least about 8%, at least about 9%, or at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, or at least 80% or more of the cells. In some embodiments of the method, the *PCSK9* gene in the cells of the population is repressed or silenced such that expression of the PCSK9 protein is decreased by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, or at least about 90% in comparison to a cell where the PCSK9 gene has not been targeted. In some embodiments of the method, the repression of the cells of a population occurs *in vitro*. In some embodiments of the method, the repression of the cells of a population occurs *ex vivo*. In some embodiments of the method, the cell is selected from the group consisting of a rodent cell, a mouse cell, a rat cell, a primate cell, and a non-human primate cell. In some embodiments of the method, the cell is a human cell.

[0357] In some embodiments of the method, repression occurs *in vitro* inside of the cells prior to introducing the cells into a subject. In some embodiments, the cells are autologous or allogeneic with respect to the subject. In some embodiments of the method, the repression of the cell occurs *in vivo* in a subject. The systems and methods described herein can be used in a variety of cells associated with disease, *e.g.*, cells of the liver, the intestine, the kidney, the central nervous system, smooth muscle cells, macrophages, or cells of arterial walls, in which the *PCSK9* gene is to be repressed or silenced. This approach, therefore, could be used for applications in a subject with a PCSK9-related disorder such as, but not limited to autosomal dominant hypercholesterolemia (ADH), hypercholesterolemia, elevated total cholesterol levels, hyperlipidemia, elevated low-density lipoprotein (LDL) levels, elevated LDL-cholesterol levels, reduced high-density lipoprotein levels, liver steatosis, coronary heart disease, ischemia, stroke, peripheral vascular disease, thrombosis, type 2 diabetes, high elevated blood pressure, atherosclerosis, obesity, Alzheimer's disease, neurodegeneration, age-related macular degeneration (AMD), or a combination thereof.

[0358] The disclosure provides a method for reversing the repression resulting from the LTRP:gRNA systems. In some embodiments of the method, the repression is reversible by use of an inhibitor of DNMT. In some embodiments of the method, the repression is

reversible by use of a cytidine analog inhibitor of DNMT. In some embodiments, the repression is reversible by use of an inhibitor selected from the group consisting of azacytidine, decitabine, clofarabine, and zebularine. In some embodiments of the method wherein reversal of the repression is in a subject treated with a system of the disclosure, the method comprises administrations of a therapeutically effective dose of the inhibitor of DNMT.

X. Therapeutic Methods

[0359] In another aspect, the present disclosure relates to methods of treating a PCSK9-related disease or disorder in a subject in need thereof, using the LTRP:gRNA systems of the disclosure. The present disclosure provides methods of treating a PCSK9-related disorder in a subject in need thereof, including, but not limited to, autosomal dominant hypercholesterolemia (ADH), hypercholesterolemia, elevated total cholesterol levels, elevated low-density lipoprotein (LDL) levels, reduced high-density lipoprotein levels, liver steatosis, atherosclerotic cardiovascular disease, and coronary artery disease, ischemia, stroke, peripheral vascular disease, thrombosis, type 2 diabetes, high elevated blood pressure, obesity, Alzheimer's disease, neurodegeneration, age-related macular degeneration (AMD), or a combination thereof. In some embodiments, the methods of the disclosure can prevent, treat and/or ameliorate a PCSK9-related disorder of a subject by the administering to the subject of a therapeutically effective dose of a composition of the disclosure. In some embodiments, the composition administered to the subject further comprises pharmaceutically acceptable carrier, diluent or excipient.

[0360] The methods described herein can be used to treat a variety of cells in the subject associated with disease, *e.g.*, cells of the liver, the intestine, the kidney, the central nervous system, smooth muscle cells, macrophages or cells of arterial walls, in which the *PCSK9* gene is to be repressed or silenced.

[0361] In some embodiments, it may be desirable to reduce or eliminate expression of the PCSK9 protein in a subject comprising mutations, for example dominant mutations leading to hypercholesterolemia or familial or autosomal dominant hypercholesterolemia. In some cases, one or both alleles of the *PCSK9* gene of the subject comprises a mutation. In some cases, the PCSK9-related disease or disorder mutation is a gain of function mutation, including, but not limited to mutations encoding amino acid substitutions selected from the group consisting of S127R, D129G, F216L, D374H, and D374Y relative to the sequence of SEQ ID NO: 1823. In other cases, the PCSK9-related disorder mutation is a loss of function

mutation including, but not limited to mutations encoding amino acid substitutions selected from the group consisting of R46L, G106R, Y142X, N157K, R237W and C679X relative to the sequence of SEQ ID NO: 1823. In some embodiments, it may be desirable to reduce or eliminate expression of the PCSK9 protein in a subject with elevated cholesterol levels that is not the result of mutations in the *PCSK9* gene.

[0362] In some embodiments, the disclosure provides methods of treating a PCSK9-related disease or disorder in a subject in need thereof comprising repressing or silencing a *PCSK9* gene in a cell of the subject, the method comprising contacting said cells with a therapeutically effective dose of: i) a LTRP:gRNA system comprising a long-term repressor fusion protein and a gRNA of any one of the embodiments described herein; ii) a nucleic acid encoding the long-term repressor fusion protein and gRNA of any of the embodiments described herein; iii) an LNP or a synthetic nanoparticle comprising a gRNA and a mRNA encoding the long-term repressor fusion protein of any one of the embodiments described herein; iv) an LNP or a synthetic nanoparticle comprising a gRNA and a an LNP or a synthetic nanoparticle comprising an mRNA encoding the long-term repressor fusion protein of any one of the embodiments described herein; or v) combinations of two or more of (i) to (iv), wherein the target nucleic acid sequence of the cells targeted by the gRNA is repressed or silenced by the long-term repressor fusion protein.

[0363] In some embodiments, the disclosure provides a method of treating a PCSK9-related disease or disorder in a subject in need thereof comprising administering a therapeutically effective dose of an LTRP:gRNA system comprising an mRNA comprising a sequence selected from the group consisting of the sequences of SEQ ID NOS: 6529-14626, and 14628-22725, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto. In some embodiments, the disclosure provides a method of treating a PCSK9-related disease or disorder in a subject in need thereof comprising administering a therapeutically effective dose of an LTRP:gRNA system comprising an mRNA comprising a sequence selected from the group consisting of the sequences of SEQ ID NOS: 6529-14626, and 14628-22725. In some embodiments, the method comprises administering a therapeutically effective dose of an LTRP:gRNA system comprising an mRNA comprising a sequence selected from the group consisting of the sequences of SEQ ID NOS: 6529-8134, 9741-11347, 14628-16233, and 17840-19446, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about

92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the method comprises administering a therapeutically effective dose of an LTRP:gRNA system comprising an mRNA comprising a sequence selected from the group consisting of the sequences of SEQ ID NOS: 6529-8134, 9741-11347, 14628-16233, and 17840-19446. In some embodiments, the method comprises administering a therapeutically effective dose of an LTRP:gRNA system comprising an mRNA comprising a sequence selected from the group consisting of the sequences of SEQ ID NOS: 6529-6537 and 9742-9750, or a sequence having at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity thereto. In some embodiments, the method comprises administering a therapeutically effective dose of an LTRP:gRNA system comprising an mRNA comprising a sequence selected from the group consisting of the sequences of SEQ ID NOS: 6529-6537 and 9742-9750. In some embodiments, the mRNA further comprises a 5' cap sequence linked 5' to the 5' UTR of the mRNA sequence encoding the LTRP, optionally wherein the 5' cap has the nucleic acid sequence selected from the group consisting of m7G(5')ppp(5')mAG, m7G(5')ppp (5'(A,G(5')ppp(5')A and G(5')ppp(5')G. In some embodiments, the mRNA comprises a 5' UTR, a 3'UTR, a poly(A), or a combination thereof.

[0364] In some embodiments of the method of treating a PCSK9-related disease or disorder in a subject in need thereof, the method comprises administering a therapeutically effective dose of an LTRP:gRNA system comprising a gRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 1744-1746, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto, and the gRNA comprises a linked targeting sequence complementary to a target nucleic acid of the *PCSK9* gene. In some embodiments of the method, the LTRP:gRNA system comprises a gRNA variant comprising a sequence of SEQ ID NO: 1744 and a linked targeting sequence. In some embodiments of the method, the LNP comprises gRNA variant comprising a sequence of SEQ ID NO: 1745 and a linked targeting sequence. In some embodiments of the method, the LTRP:gRNA system comprises gRNA variant comprising a sequence of SEQ ID NO: 1746 and a linked targeting sequence. In some embodiments of the method, the LNP comprises a gRNA variant with chemical modifications, optionally comprising a sequence

selected from the group consisting of SEQ ID NOS: 2948-2956, 2958-2966, and 2968-2976, and a targeting sequence complementary to a *PCSK9* target nucleic acid substituted for the 20 nucleotides on the 3' end of the gRNA of the listed sequences. In some embodiments of the method, the LTRP:gRNA system comprises gRNA variant comprising a sequence of SEQ ID NO: 1746 with chemical modifications, optionally comprising a sequence selected from the group consisting of SEQ ID NOS: 2968-2976, and a targeting sequence complementary to a *PCSK9* target nucleic acid substituted for the 20 nucleotides on the 3' end of the gRNA of the listed sequences. In one embodiment of the method, the LTRP:gRNA system comprises gRNA variant comprising a sequence of SEQ ID NO: 1746 with chemical modifications comprising the sequence of SEQ ID NO: 2968, and a targeting sequence complementary to a *PCSK9* target nucleic acid substituted for the 20 nucleotides on the 3' end of the gRNA of the listed sequence. In some embodiments, the chemically modified gRNA comprises a sequence selected from the group consisting of SEQ ID NOS: 22788- 22803. In some embodiments of the method, the targeting sequence of the gRNA is selected from the group consisting of SEQ ID NOS: 1824-2544, 2672, 2675, 2694, and 2714. In some embodiments of the method, the targeting sequence of the gRNA is selected from the group consisting of SEQ ID NOS: 1824-1879, 1880, 1883, 1884, 1888, 1889, 2672, 2675, 2694, and 2714. In some embodiments of the method, the targeting sequence of the gRNA is selected from the group consisting of SEQ ID NOS: 1834, 1849, 1853, 1855, 1856, 1857, 1858, 1860, 1862, 1863, 1867, 1869, 1870, 1872, 1874, and 1875. In some embodiments of the method, the targeting sequence is selected from the group consisting of SEQ ID NOS: 1855, 1867, and 1869. In some embodiments, the chemically modified gRNA comprises a sequence selected from the group consisting of SEQ ID NOS: 22788-22790.

[0365] In some embodiments of the method of treating a PCSK9-related disease or disorder in a subject in need thereof, the method comprises administering a therapeutically effective dose of an LNP comprising an LTRP:gRNA system comprising a gRNA with a linked targeting sequence complementary to a target nucleic acid of the *PCSK9* gene and an mRNA encoding the long-term repressor fusion protein. In some embodiments of the method, the LNP comprises an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6529-14626, and 14628-22725, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto. In some embodiments of the method, the LNP comprises an mRNA comprising a sequence selected

from the group consisting of SEQ ID NOS: 6529-14626, and 14628-22725. In some embodiments of the method, the LNP comprises an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6529-8134, 9741-11347, 14628-16233, and 17840-19446, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto. In some embodiments of the method, the LNP comprises an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6529-8134, 9741-11347, 14628-16233, and 17840-19446. In some embodiments of the method, the LNP comprises an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6529-6537 and 9742-9750, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto. In some embodiments of the method, the LNP comprises an mRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 6529-6537 and 9742-9750. In some embodiments, the mRNA further comprises a 5' cap sequence linked 5' to the 5' UTR of the mRNA sequence encoding the LTRP, optionally wherein the 5' cap has the nucleic acid sequence selected from the group consisting of m7G(5')ppp(5')mAG, m7G(5')ppp(5')(A,G(5')ppp(5')A and G(5')ppp(5')G. In some embodiments, the mRNA comprises a 5' UTR, a 3' UTR, a poly(A), or a combination thereof.

[0366] In some embodiments, the method for transcriptional repression or silencing of the *PCSK9* gene comprises administering a therapeutically effective dose of an LNP to a subject, the LNP comprising an LTRP:gRNA system comprising a gRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 1744-1746, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto, and the gRNA comprises a linked targeting sequence complementary to a target nucleic acid of the *PCSK9* gene. In some embodiments of the method, the LNP comprises a gRNA variant comprising a sequence of SEQ ID NO: 1744 and a linked targeting sequence. In some embodiments of the method, the LNP comprises a gRNA variant comprising a sequence of SEQ ID NO: 1745 and a linked targeting sequence. In some embodiments of the method, the LNP comprises a gRNA variant comprising a sequence of SEQ ID NO: 1746 and a linked targeting sequence. In some embodiments of the method, the LNP comprises a gRNA variant with chemical modifications, optionally comprising a

sequence selected from the group consisting of SEQ ID NOS: 2948-2956, 2958-2966, and 2968-2976, with a targeting sequence complementary to a *PCSK9* target nucleic acid substituted for the 20 nucleotides on the 3' end of the gRNA of the listed sequences. In some embodiments of the method, the LNP comprises gRNA variant 316 with chemical modifications, comprising a sequence selected from the group consisting of SEQ ID NOS: 2968-2976, with a targeting sequence complementary to a *PCSK9* target nucleic acid substituted for the 20 nucleotides on the 3' end of the gRNA of the listed sequences. In one embodiment of the method, the LNP comprises gRNA variant with chemical modifications comprising the sequence of SEQ ID NO: 2968, and a targeting sequence complementary to a *PCSK9* target nucleic acid substituted for the 20 nucleotides on the 3' end of the gRNA of the listed sequence. In some embodiments of the method, the targeting sequence of the gRNA is selected from the group consisting of SEQ ID NOS: 1824-2544, 2672, 2675, 2694, and 2714. In some embodiments of the method, the targeting sequence of the gRNA is selected from the group consisting of SEQ ID NOS: 1824-1879, 1880, 1883, 1884, 1888, 1889, 2672, 2675, 2694, and 2714. In some embodiments of the method, the targeting sequence of the gRNA is selected from the group consisting of SEQ ID NOS: 1834, 1849, 1853, 1855, 1856, 1857, 1858, 1860, 1862, 1863, 1867, 1869, 1870, 1872, 1874, and 1875. In some embodiments of the method, the targeting sequence is selected from the group consisting of SEQ ID NOS: 1855, 1867, and 1869. In some embodiments, the LNP comprises a modified gRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 22788- 22803. In some embodiments, the LNP comprises a modified gRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 22788-22790.

[0367] In some embodiments of the method of treating a *PCSK9*- related disease or disorder in a subject in need thereof, the method results in transcriptional repression or silencing of the *PCSK9* gene in at least about 1%, at least about 2%, at least about 3%, at least about 4%, at least about 5%, at least about 6%, at least about 7%, at least about 8%, at least about 9%, or at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60% or more of the cells of the targeted organ. In some embodiments of the method, the *PCSK9* gene in the cells of the targeted organ are repressed such that expression of the *PCSK9* protein is decreased by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, or at least about 90% in comparison to an untreated cell. The cell of the treated subject can be a cell selected from the group consisting

of a rodent cell, a mouse cell, a rat cell, a primate cell, and a non-human primate cell. In some embodiments, the eukaryotic cell of the treated subject is a human cell. In some embodiments, the cell is a cell involved in the production of LDL, including but not limited to a hepatocyte, or a cell of the intestine, the kidney, the central nervous system, a smooth muscle cell, macrophage, a retinal cell, or cell of arterial walls such as the endothelium. In some embodiments, the cell is an eye cell. In some embodiments of the method of treating a PCSK9-related disorder in a subject, the subject is selected from the group consisting of mouse, rat, pig, non-human primate, and human.

[0368] A number of therapeutic strategies may be used in the methods of treatment of a subject with a PCSK9-related disorder. In some embodiments, the disclosure provides a method of treatment of a subject having a PCSK9-related disorder, the method comprising administering to the subject a LTRP:gRNA composition according to a treatment regimen comprising one or more consecutive doses using a therapeutically effective dose. In some embodiments of the treatment regimen, the therapeutically effective dose of the composition is administered as a single dose. In other embodiments of the treatment regimen, the therapeutically effective dose is administered to the subject as two or more doses over a period of at least two weeks, or at least one month, or at least two months, or at least three months, or at least four months, or at least five months, or at least six months. In some embodiments of the treatment regimen, the effective doses are administered by a route selected from the group consisting of intravenous, intraportal vein injection, intraperitoneal, intramuscular, subcutaneous, intraocular, and oral routes.

[0369] In some embodiments of the method of treatment, the administering of the therapeutically effective amount of a LTRP:gRNA modality, including an LNP comprising a gRNA and an mRNA encoding a long-term repressor fusion protein to a subject with a PCSK9-related disease or disorder leads to the prevention or amelioration of the underlying PCSK9-related disease or disorder such that an improvement is observed in the subject, notwithstanding that the subject may still be afflicted with the underlying disorder. In some embodiments, the administration of the therapeutically effective amount of the LTRP:gRNA modality leads to an improvement in at least one clinically-relevant endpoint including, but not limited to, percent change from baseline in LDL-cholesterol, decrease in plaque atheroma volume, reduction in coronary plaque, reduction in atherosclerotic cardiovascular disease (ASCVD), cardiovascular death, nonfatal myocardial infarction, ischemic stroke, nonfatal stroke, coronary revascularization, unstable angina, or visual acuity. In some embodiments,

the administration of the therapeutically effective amount of the LTRP:gRNA modality leads to an improvement in at least two clinically-relevant endpoints. In some embodiments, the subject is selected from mouse, rat, pig, dog, non-human primate, and human.

[0370] In some embodiments, the method of treatment further comprises administering a chemotherapeutic agent wherein the agent is effective in lowering LDL levels. Such agents include, but are not limited to, statins, niacin, fibrates, or anti-PCSK9 antibody drugs. In some embodiments, the method of treating a PCSK9-related disease or disorder in a subject comprises pretreating the subject with a therapeutic agent that increases hepatic LDL receptor (LDLR) expression. In some embodiments, the therapeutic agent is a PCSK9 inhibitor, such as a monoclonal antibody, nucleic acid-based agent, or a small molecule. Exemplary therapeutic agents include, but are not limited to, evolocumab, inclisiran, alirocumab, and MK-0616. Without wishing to be bound by theory or mechanism, it is believed that the pretreatment with an inhibitor of PCSK9, may lead to an increase in hepatic LDL receptor (LDLR) expression that, in turn, may facilitate the uptake of the LNP comprising the CasX:gRNA composition that is subsequently administered to the subject. By increasing the hepatic cell uptake of the LNP, it is expected that editing of the *PCSK9* gene will be enhanced such that an improvement in the PCSK9-related disorder would be attained.

[0371] Methods of obtaining samples from treated subjects for analysis to determine the effectiveness of the treatment, such as body fluids or tissues, and methods of preparation of the samples to allow for analysis are well known to those skilled in the art. Methods for analysis of RNA and protein levels are discussed above and are well known to those skilled in the art. The effects of treatment can also be assessed by measuring biomarkers associated with the target gene expression in the aforementioned fluids, tissues or organs, collected from an animal contacted with one or more compounds of the disclosure, by routine clinical methods known in the art. Biomarkers of PCSK9 disorders include, but are not limited to, PCSK9 levels, low-density lipoprotein (LDL-cholesterol), apolipoprotein B, non-HDL cholesterol, triglycerides and lipoprotein a, soluble CD40 ligand, osteopontin (OPN), osteoprotegerin (OPG), matrix metalloproteinases (MMP) and myeloperoxidase (MPOP), wherein the concentration of the marker is compared to concentrations known to be physiologically normal or in subjects not having a PCSK9 disorder.

[0372] Several mouse models expressing mutant forms of PCSK9 exist and are suitable for evaluating the methods of treatment. Transgenic mouse models of PCSK9-related disorders include knock-in mouse models having hPCSK9 (Carreras, A. *In vivo* genome and base

editing of a human PCSK9 knock-in hypercholesterolemic mouse model. *MC Biology* 17:4 (2019); Herbert B., et al. Increased secretion of lipoproteins in transgenic mice expressing human D374Y PCSK9 under physiological genetic control. *Arterioscler Thromb Vasc Biol.* 30(7):1333 (2010)).

XI. Compositions, Pharmaceutical Compositions, Kits, and Articles of Manufacture

[0373] In some embodiments, the disclosure provides compositions comprising: i) a long-term repressor fusion protein and a gRNA of the disclosure comprising a targeting sequence complementary to a sequence of an *PCSK9* gene; ii) one or more nucleic acids encoding the long-term repressor fusion protein and the gRNA of (i); iii) an LNP or synthetic nanoparticle comprising a gRNA and an mRNA encoding the long-term repressor fusion protein, or iv) an LNP or synthetic nanoparticle comprising a gRNA and an LNP or synthetic nanoparticle comprising an mRNA encoding the long-term repressor fusion protein.

[0374] In some embodiments, the disclosure provides pharmaceutical compositions comprising: i) a long-term repressor fusion protein and a gRNA of any of the embodiments of the disclosure comprising a targeting sequence complementary to a sequence of a *PCSK9* gene; ii) one or more nucleic acids encoding the long-term repressor fusion protein and the gRNA of (i); iii) an LNP or synthetic nanoparticle comprising a gRNA and an mRNA encoding a long-term repressor fusion protein, or or iv) an LNP or synthetic nanoparticle comprising a gRNA and an LNP or synthetic nanoparticle comprising an mRNA encoding the long-term repressor fusion protein, together with one or more pharmaceutically suitable excipients, buffers, diluent or carriers. In some embodiments, the pharmaceutical composition is formulated for a route of administration selected from the group consisting of intravenous, intraportal vein injection, intraperitoneal, intramuscular, subcutaneous, intraocular, and oral routes. In one embodiment, the pharmaceutical composition is in a liquid form or a frozen form. In another embodiment, the pharmaceutical composition is in a pre-filled syringe for a single injection. In another embodiment, the pharmaceutical composition is in solid form, for example the pharmaceutical composition is lyophilized.

[0375] Excipients can include a salt, an isotonic agent, a serum protein, a buffer or other pH-controlling agent, an anti-oxidant, a thickener, an uncharged polymer, a preservative or a cryoprotectant. Excipients used in compositions of the disclosure may further include an isotonic agent and a buffer or other pH-controlling agent. These excipients may be added for the attainment of preferred ranges of pH (about 6.0-8.0) and osmolarity (about 50-400

mmol/L). Examples of suitable buffers are acetate, borate, carbonate, citrate, phosphate and sulfonated organic molecule buffer. Such buffers may be present in a composition in concentrations from 0.01 to 1.0% (w/v). An isotonic agent may be selected from any of those known in the art, e.g. mannitol, dextrose, glucose and sodium chloride, or other electrolytes. In some embodiments the isotonic agent may be glucose or sodium chloride. The isotonic agents may be used in amounts that impart to the composition the same or a similar osmotic pressure as that of the biological environment into which it is introduced. The concentration of isotonic agent in the composition will depend upon the nature of the particular isotonic agent used and may range from about 0.1 to 10%. When glucose is used, it is preferably used in a concentration of from 1 to 5% w/v, more particularly 5% w/v. When the isotonic agent is sodium chloride, it is preferably employed in amounts of up to 1% w/v, in particular 0.9% w/v. The compositions of the disclosure may further contain a preservative. Examples of preservatives include polyhexamethylene-biguanidine, benzalkonium chloride, stabilized oxychloro complexes (such as those known as Purite®), phenylmercuric acetate, chlorobutanol, sorbic acid, chlorhexidine, benzyl alcohol, parabens, and thimerosal. Typically, such preservatives are present at concentrations from about 0.001 to 1.0%. Furthermore, the compositions of the disclosure may also contain a cryopreservative agent. Preferred cryopreservatives are glucose, sucrose, mannitol, lactose, trehalose, sorbitol, colloidal silicon dioxide, dextran of molecular weight preferable below 100,000 g/mol, glycerol, and polyethylene glycols of molecular weights below 100,000 g/mol or mixtures thereof. Most preferred are glucose, trehalose and polyethylene glycol. Typically, such cryopreservatives are present at concentrations from about 0.01 to 10%.

[0376] Additional pharmaceutical formulations appropriate for administration are applicable in the methods and compositions disclosed herein (*see, e.g.*, Remington's Pharmaceutical Sciences (1990) 18th ed., Mack Publishing Co., Easton, Pa.; Remington's Pharmaceutical Sciences (2023) 23rd ed., Elsevier Publishing; The Merck Index (1996) 12th ed., Merck Publishing Group, Whitehouse, N.J.; and Pharmaceutical Principles of Solid Dosage Forms, Technomic Publishing Co., Inc., Lancaster, Pa., (1993).

[0377] In other embodiments, provided herein are kits comprising the LTRP:gRNA systems, polynucleotides, vectors, and LNP formulations described herein. In other embodiments, provided herein are kits comprising a long-term repressor fusion protein and one or a plurality of CasX gRNA of any of the embodiments of the disclosure, comprising a targeting sequence complementary to a sequence of a *PCSK9* gene. In certain embodiments,

provided herein are kits comprising an mRNA encoding a long-term repressor fusion protein and one or a plurality of gRNAs of any of the embodiments of the disclosure comprising a targeting sequence complementary to a sequence of a *PCSK9* gene target nucleic acid sequence.

[0378] In some embodiments, provided herein are kits comprising an LNP formulation encapsulating an mRNA encoding a long-term repressor fusion protein and one or a plurality of gRNA of any of the embodiments of the disclosure comprising a targeting sequence complementary to a sequence of a *PCSK9* gene. In some embodiments, the mRNA further comprises a 5' cap sequence linked 5' to the 5' UTR of the mRNA sequence wherein the 5' cap has the nucleic acid sequence selected from the group consisting of m7G(5')ppp(5')mAG, m7G(5')ppp(5')(A,G(5')ppp(5')A and G(5')ppp(5')G. In some embodiments, the kit comprises a first LNP formulation comprising the mRNA, and a second LNP formulation comprising the gRNA. In some embodiments, the mRNA and gRNA are encapsulated in a single LNP formulation.

[0379] In exemplary embodiments, a kit of the disclosure comprises an mRNA comprising a sequence encoding a long-term repressor fusion protein selected from the group consisting of SEQ ID NOS: 6528-14626, and 14627-22725, and a gRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 1744-1746 and 2947-2976, with a linked targeting sequence complementary to a target nucleic acid in the *PCSK9* gene. In some embodiments, the targeting sequence is selected from the group consisting of SEQ ID NOS: 1824-1880, 1883, 1884, 1888, 1889, 2672, 2675, 2694, and 2714. In some embodiments, a kit of the disclosure comprises an mRNA comprising a sequence encoding a long-term repressor fusion protein selected from the group consisting of SEQ ID NOS: 6528-14626, and 14627-22725, and a modified gRNA comprising a sequence selected from the group consisting of SEQ ID NOS: 2278-22803. In some embodiments, a kit of the disclosure comprises an mRNA comprising a sequence encoding a long-term repressor fusion protein selected from the group consisting of SEQ ID NOS: 6528-14626, and 14627-22725, and a modified gRNA a sequence selected from the group consisting of SEQ ID NOS: 2278-22790.

[0380] In some embodiments, the kit comprises a vector comprising a sequence encoding a long-term repressor fusion protein and a gRNA of the disclosure.

[0381] In some embodiments, the kit further comprises a buffer, a nuclease inhibitor, a protease inhibitor, a liposome, a therapeutic agent, a label, a label visualization reagent,

instructions for use, or any combination of the foregoing. In some embodiments, the kit further comprises a pharmaceutically acceptable carrier, diluent or excipient.

[0382] In some embodiments, the kit further comprises a suitable container (for example a tube, vial or plate).

[0383] In some embodiments, the kit comprises appropriate control compositions for gene repression applications, and instructions for use.

[0384] In some embodiments, the disclosure provides compositions for use in the treatment a subject having an PCSK9-related disease. In some embodiments, the disclosure provides compositions of a long-term repressor fusion protein and gRNA for use in the manufacture of a medicament for use in the treatment of a disease in a subject. In some embodiments, the disclosure provides compositions of LNP comprising a long-term repressor fusion protein and gRNA for use in the manufacture of a medicament for use in the treatment of a disease in a subject.

[0385] The following Examples are merely illustrative and are not meant to limit any aspects of the present disclosure in any way.

299945236 v1

EXAMPLES

Example 1: Assessment of spacers in achieving repression of the *PCSK9* locus in human hepatocyte cells when paired with an LTRP5-ADD molecule

[0386] Experiments were performed to demonstrate that multiple spacers with the TTC recognition motif, when paired with an LTRP molecule in configuration 5 (LTRP5; diagrammed in FIG. 1) containing the ADD domain, can induce durable repression of a therapeutically-relevant endogenous locus in human cells. Specifically, an initial proof-of-concept experiment was performed in human Huh7 cells to evaluate a subset of spacers that exhibit sequence conservation to the non-human primate (NHP) genome to identify leading spacers for testing in future in vivo NHP studies.

Materials and Methods:

[0387] Computational selection of *PCSK9*-targeting spacers for experimental testing with an LTRP5 molecule containing the ADD domain:

[0388] To determine potential LTRP-specific spacers throughout the human *PCSK9* locus, a target search region was defined as starting at 5KB upstream of the transcription start site (TSS) through 5KB downstream of the transcription stop site. Spacers were determined based on the availability of TTC PAMs; consequently, a total of 1,121 TTC spacers were identified throughout the target *PCSK9* locus. These spacers were then functionally annotated by overlaying key genomic features based on their positioning, i.e., determining whether the putative spacer targeted an exon, an intron, or a candidate cis-regulatory element (cCRE), within the promoter region, and/or overlapped with a common site of genetic variation (e.g., SNPs). To narrow down and determine an initial group of spacers for experimental screening, the extracted spacers were subjected to a set of filtering criteria. Firstly, non-specific spacers were excluded by removing spacers with off-target sites that contain up to one base pair mismatch with the on-target site. Furthermore, spacers containing the following mononucleotide repeats were excluded: thymine nucleotide repeats greater than four base pairs (bp) in length or adenine, guanine, or cytosine nucleotide repeats greater than 5 bp in length. Next, from this filtered set, spacers with more than one off-target site containing mismatches in the last four nucleotides of the spacer were excluded. Lastly, spacers that targeted >2KB upstream of the TSS and >2KB downstream of the transcription stop site were excluded. This resulted in a filtered set of 722 TTC spacers (SEQ ID NO: 1824-2545). From this filtered set of 722 spacers, spacers that were TSS-proximal (within 1100 bp upstream and downstream of the TSS) were selected for experimental assessment, resulting in the identification of 67 TTC spacers. Two additional spacers, TG-06-354 and TG-06-352,

positioned beyond the 1100 bp threshold window, were also selected for inclusion. The sequences of the resulting 69 TTC spacers are shown in Table 11.

Table 11: RNA sequences of the 69 TTC spacers targeting the human *PCSK9* locus. Bolded spacers were spacers having sequence consensus between human and non-human primate genomes and were assessed in this example

Spacer ID	Spacer RNA sequence	SEQ ID NO:
TG-06-342	AAUUACAGGCAACAGGAAGG	1824
TG-06-343	CCCAUGUAAGAGAGGAAGU	1825
TG-06-344	CAGUUUCUGCCUCGCCGCGG	1826
TG-06-345	GCCUCGCCGCGGCACAGGUG	1827
TG-06-346	CCCACCUUGGCCGCGGCAG	1828
TG-06-347	CUCCUUCACCCACCUUGGCC	1829
TG-06-348	AGGCAUUCACUCCUUCACCC	1830
TG-06-349	CUGUGCCUGGUGCAGUUCUU	1831
TG-06-350	GUGUCAUAAAGAAUUGCCU	1832
TG-06-351	UUAUGACACAGAACUCAUGC	1833
TG-06-001	GAGGAGGACGGCCUGGCCGA	1834
TG-06-002	ACCGCUGCGCCAAGGUGCGG	1835
TG-06-004	GCCAGGCCGUCUCCUCGGA	1836
TG-06-005	GUGCUCGGGUGCUUCGGCCA	1837
TG-06-117	ACUUUGUUUGCAAAGACCUC	1838
TG-06-118	GAGUGAAAUGGCCUGCUCUG	1839
TG-06-119	GAGCAGGCCAUUUCACUCGG	1840
TG-06-120	CUCGGAUUCUGCUGUGCAUC	1841
TG-06-121	GGAAGGGCUGUCGAUACUGG	1842
TG-06-123	UCCCAGUAUCGACAGCCCUU	1843
TG-06-124	CAGUAUCGACAGCCCUUCCA	1844
TG-06-125	AGAAAGAGCAAGCCUCAUGU	1845
TG-06-128	AGAAAUCAACUGGACAAGCA	1846
TG-06-131	UGAACAUUGGUGUGAAAAGG	1847
TG-06-132	AGAAGAUUCAAUUGCAAAG	1848
TG-06-133	AUGGUAGGCACAAGCUCAGC	1849
TG-06-134	GAAUUCUAUGGUAGGCACAA	1850
TG-06-135	GGAAAGCUGAGCUUGUGCCU	1851
TG-06-138	AGGGAUUUAUACUACAAAGA	1852
TG-06-139	AGGAGCAGCUAGUUGGUAAG	1853
TG-06-140	AAACUUAGCCUGGACCCCUU	1854
TG-06-141	ACUGGCCUUAACCGGCAGC	1855

Spacer ID	Spacer RNA sequence	SEQ ID NO:
TG-06-142	UUCCACUGGCCUUAACCUGG	1856
TG-06-143	GAAUCAAUCCUACUGUGGAC	1857
TG-06-144	GUGGGCAGCGAGGAGUCCAC	1858
TG-06-145	UGGGUCCACCUUGUCUCCUG	1859
TG-06-146	GAAGUCUCACUGGUCAGCAG	1860
TG-06-147	GUGUUUCCUGGGUCCACCUU	1861
TG-06-149	AGCCCAGUUAGGAUUUGGGA	1862
TG-06-150	UCCCUCUGCGCGUAAUCUGA	1863
TG-06-151	CUCUGCGCGUAAUCUGACGC	1864
TG-06-152	GCCUCGCCCUCGCCAAACAG	1865
TG-06-153	GUUAAUGUUUAAUCAGAUAG	1866
TG-06-154	AGGGUGUGGGUGCUUGACGC	1867
TG-06-155	GCAGCGACGUCGAGGCGCUC	1868
TG-06-157	GGGUCUGAGCCUGGAGGAGU	1869
TG-06-158	GGAGCAGGGCGCGUGAAGGG	1870
TG-06-159	GCGCGCCCCUUCACGCGCCC	1871
TG-06-160	CGCGCCUGCUCUGAACUU	1872
TG-06-161	GCUCCUGCACAGUCCUCCCC	1873
TG-06-167	CACUGAAUAGCGCAGCCGCA	1874
TG-06-168	GUGGGAAGGUUCGCGGGGUU	1875
TG-06-169	CGGGGUUGGGAGACCCGGAG	1876
TG-06-170	UCGGCCUCCGGGUCUCCCAA	1877
TG-06-171	CAGUACGUUCCAGGCAUUCA	1878
TG-06-172	GCUGAAACAGAUGGAAUACU	1879
TG-06-249	AAACCAAUCGGAACCCACU	1880
HS-6-147	UGUUGCCUGUAAUUGGAAUU	1881
HS-6-149	CCUUCUGUUGCCUGUAAUU	1882
TG-06-122	CCUUUGUUUCUCCCAGUAU	1883
TG-06-127	UCCUCCUGCCUGGUACACAA	1884
TG-06-137	GAAUGUACCUAUAUGACGUC	1885
TG-06-188	CCCCGGCCUCCAUCCCUAC	1886
TG-06-243	CUUGGCACGAUCUUGGGGAC	1887
TG-06-250	GAUUUGGUUUGGAAAACAUG	1888
TG-06-251	CUCCAGGCCCUCCACCCUCC	1889
HS-6-159	CACCCGCCCCUGUCUCGGG	1890
TG-06-354	CCCCUGCCCCUUCAGCUGGU	1925
TG-06-352	UCCUCACCAAUACCCUC	1910

[0389] Assessment of PCSK9 secretion levels for select *PCSK9*-targeting spacers having sequence conservation with the non-human primate genome:

[0390] Of the 69 TTC spacers identified, 15 spacers that exhibited sequence conservation between human and non-human primate genomes (bolded spacers in Table 11) were initially tested to assess their effect on PCSK9 secretion levels.

[0391] mRNA encoding the following molecules were generated by in-house *in vitro* transcription (IVT), which were carried out with CleanCap® AG and N1-methyl-pseudouridine. IVT reactions were then subjected to DNase digestion and oligodT purification on-column: 1) a catalytically-active CasX 676 (as described in Example 2), 2) dXR1 (a dCasX fused to a ZIM3-KRAB domain), and 3) LTRP5-ADD-ZIM3 (LTRP5 containing the ADD domain and the ZIM3-KRAB domain). The DNA and mRNA sequences for CasX 676 are shown in Tables 19 and 20; the DNA and mRNA sequences for dXR1 are shown in Tables 12 and 13; the DNA and mRNA sequences for LTRP5-ZIM3-ADD are shown in Tables 15 and 16.

[0392] gRNAs containing NHP-conserved spacers targeting the *PCSK9* locus (bolded spacers in Table 11) were designed using gRNA scaffold 316 and chemically synthesized. Furthermore, a *B2M*-targeting gRNA was used as a non-targeting control, while spacer TG-06-138 (also known as spacer 6.138; SEQ ID NO: 1852) was used to pair with dXR1, and spacer TG-06-001 (also known as spacer 6.1; SEQ ID NO: 1834) was used to pair with CasX 676. Spacer TG-06-157 (also known as spacer 6.157; SEQ ID NO: 1869), which is not an NHP-conserved spacer, was included as a positive control given its demonstrated efficacy in sustaining repression of the *PCSK9* locus, which is shown in Example 2, below.

[0393] To assess PCSK9 secretion, seeded Huh7 cells were transfected with mRNA encoding a catalytically-active CasX 676, dXR1, or LTRP5-ADD-ZIM3 and a gRNA with scaffold 316 and spacer targeting either the *B2M* or *PCSK9* locus. Media supernatant was harvested at 6, 18, and 36 days post-transfection to assess level of PCSK9 secretion by ELISA. Levels of PCSK9 secretion were normalized to total cell count. As an additional control, PCSK9 secretion was also measured in the media supernatant harvested from wells containing untreated, naïve cells.

[0394] In a second experiment, 16 of the top spacers from the preceding experiment were lipofected into Primary human hepatocytes (PHHs) with LTRP5-ADD-ZIM3 and a gRNA with scaffold 316 and spacer targeting either the *B2M* or *PCSK9* locus, and media supernatant collected at 6 days post-lipofection.

[0395] Lonza lot 31 (1 round) or 271 (2 rounds) cells were seeded into 96 well plates and co-lipofected with a high (150 ng) or medium (75 ng) dose of mRNA + gRNA using MessengerMax reagent. Cells were maintained in culture for six days before PCSK9 knock

down was assessed both by secreted protein ELISA and transcript RT-qPCR for a total of 3 independent screening rounds. Medium dose mRNA/gRNA lipofections were harvested for amplicon bisulfite sequencing. Cells were propagated for six days before DNA extraction and amplicon bisulfite sequencing on a PCSK9 TSS-proximal region (~200 bp upstream to ~150 bp downstream) containing 34 CpG dinucleotides. Knock down was confirmed by PCSK9 transcript RT-qPCR (as described in example 16) and supernatant PCSK9 HTRF ELISA (as described in example 16). Total RNA was also extracted for RNA-seq to assess specificity.

[0396] Durability of repression was assessed in a third experiment. HuH7s were lipofected in 6 replicate wells with high dose of LTRP5-ADD-ZIM3 and a gRNA with scaffold 316 and a spacer targeting either the *B2M* or *PCSK9* locus, representing the 11 most active targeting sequences from the above experiment. Cells were propagated in culture with PCSK9 supernatant ELISA readouts every 2 weeks. Replica plates split at week 1 increased n from 6 to 12 for subsequent readouts.

Results:

[0397] Quantification of normalized PCSK9 secretion level for Huh7 cells transfected with mRNA encoding catalytically-active CasX 676, dXR1, or LTRP5-ADD-ZIM3 with an NHP-conserved gRNA targeting the *PCSK9* locus at the three timepoints is shown FIG. 2. The data demonstrate that use of most NHP-conserved spacers with LTRP5-ADD-ZIM3 resulted in sustained repression through 36 days post-transfection when compared with control conditions, i.e., naïve, untreated cells, cells treated with dXR1, and cells treated with the non-targeting control (using spacer 7.37 targeting the *B2M* locus). Specifically, in comparison to the PCSK9 secretion level observed with use of spacer 6.1 paired with CasX 676, use of TSS-proximal spacers TG-06-147, TG-06-167, TG-06-133, TG-06-146, and TG-06-154 paired with LTRP5-ADD-ZIM3 resulted in similar or further reduced level of sustained repression (FIG. 2). Interestingly, use of TG-06-352, which is positioned beyond the 1100 bp threshold window designated here as “TSS-proximal”, also resulted in effective repression (FIG. 2). Similar to the findings observed in Example 2, treatment with LTRP5-ADD-ZIM3 with spacer 6.157 resulted in sustained repression of secreted PCSK9 levels, while treatment with dXR1 and spacer 6.138 resulted in transient repression. Furthermore, treatment with any of the three mRNA molecules with spacer 7.37 targeting the *B2M* locus did not affect PCSK9 secretion (FIG. 2).

[0398] Activity levels as assessed by RT-qPCR and ELISA with the top 16 spacer sequences demonstrate that TG-06-154, TG-06-141, and TG-06-157 were among the most active spacers (Table 17). Targeting sequence TG-06-154 demonstrated the highest average

secreted protein knock down by ELISA across rounds and doses. PCSK9-targeting spacers elicited ~40-60% average CpG methylation, with the exception of targeting spacers TG-06-146 and TG-06-142 which demonstrated lower average CpG methylation (Table 17).

[0399] Specificity and durability of repression are quantified in Table 18. Treatment with LTRP5-ADD-ZIM3 and targeting sequences TG-06-133, TG-06-139, TG-06-158, and TG-06-160 demonstrated increased numbers of off-target differentially expressed genes, while targeting sequences TG-06-141, TG-06-144, TG-06-154, and TG-06-157 demonstrated no more than 1 differentially expressed off-target gene. Durability, as measured by percent change in reduction of PCSK9 from week 1 to week 9, demonstrated that PCSK9-targeting spacers maintained repression within ~30% of initial repression over 9 weeks with the exception of targeting spacers TG-06-001, TG-06-139, and TG-06-150 (Table 18).

[0400] These results demonstrate that delivery of mRNA encoding an LTRP molecule with the ADD domain with the appropriate PCSK9-targeting gRNA can result in sustained repression of an endogenous target locus in human cells. Furthermore, these experiments revealed that several human spacers having consensus sequence with the non-human primate species achieved strong phenotypic effects from targeting a therapeutically-relevant locus, supporting the potential use of these select spacers in preclinical efficacy studies utilizing non-human primate models.

Table 12: Encoding sequences of the dXR1 and LTRP1-ZIM3 mRNA molecules assessed in experiment #1 of this example*

dXR or LTRP ID	Component	DNA sequence	SEQ ID NO:
dXR1 (codon-optimize d)	5'UTR	AAATAAGAGAGAAAAGAAGAGTAAGAAGAAATATAAGAGCCACC	3047
	START codon + NLS + linker	ATGGCCCCCTAAGAAGAAGCGTAAAGTGAGCCGGGGCGGCAGCGGCG GCGGCAGCGCC	3048
	dCasX491	CAGGAGATTAAACGGATCAACAAGATCAGAAGAAGACTTGTGAAAG ACAGCAACACCAAGAAGCCGGCAAGACAGGCCCCATGAAAACCT GCTGGTTAGAGTGATGACACCCGATCTGAGAGAGCGGCTGGAAAAC CTGAGAAAGAAGCCTGAAAATATCCCCAGCCCATCAGCAATACAT CTAGAGCCAACCTGAATAAGCTGCTGACCGATTACACCGAAATGAA GAAGGCGATCCTGCATGTGTACTGGGAAGAGTTCCAGAAGGACCT GTGGGCCTGATGAGCCGGGTGGCCAGCCTGCCAGCAAGAAGATCG ATCAGAACAAGCTGAAACCTGAGATGGACGAGAAGGGCAACCTGAC CACCGCCGGCTTTGCCTGCTCTCAGTGTGGCCAGCCCTGTTCTGTG TACAAGCTGGAGCAGGTGTCTGAGAAGGGCAAGGCTTACACCACT ACTTCGGACGGTGCAATGTGGCCGAGCACGAAAAGCTGATCCTGCT GGCCCAGCTGAAGCCCCGAGAAGGATAGCGACGAAGCCGTGACATAT AGCCTGGGAAAGTTTGGGCAGAGGGCCCTGGATTTCTACAGCATTC	3049

dXR or LTRP ID	Component	DNA sequence	SEQ ID NO:
		ATGTGACCAAGGAGTCCACCCACCCCGTGAAGCCCCCTGGCCCAGAT CGCCGGAAACAGATACGCCTCCGGACCTGTGGGAAAGGCCCTGAGC GACGCATGTATGGGCACAATCGCCTCCTTCCTGTCTAAGTACCAGG ACATCATCATCGAACACCAGAAGGTGGTGAAGGGCAACCAGAAGAG ACTGGAGAGCCTGCGGGAGCTGGCCGGCAAGGAAAACCTGGAATAC CCTAGCGTGACCCTGCCACCTCAGCCTCACACCAAGGAGGGCGTTG ATGCCTACAACGAAGTGATCGCCCGGGTGCGAATGTGGGTGAACCT GAACCTGTGGCAGAAGCTGAAGCTAAGCAGAGATGATGCCAAGCCT CTGCTGAGACTGAAGGGATTCCCTTCCTTTCTCTGGTCGAGAGAC AGGCCAACGAAGTGGACTGGTGGGACATGGTGTGTAACTGAAGAA GCTGATCAACGAGAGAAAAGGAGGATGGCAAGGTGTTTTGGCAGAA CTGGCTGGCTACAAGAGACAGGAAGCCCTGAGACCATACTGAGCA GCGAGGAAGATCGGAAGAAGGGAAAGAAATTCGCTCGGTACCAGCT GGGCGACCTGCTGCTGCACCTGGAAAAGAAGCACGGCGAGGACTGG GGAAAGGTGTACGACGAGGCCTGGGAGCGGATTGACAAGAAAGTGG AAGCCTGAGCAAGCACATCAAGCTGGAAGAGGAACGGGAAGCGGA GGACGCCCAGAGCAAGGCCGCCCTGACCGACTGGCTGCGGGCTAAG GCCAGCTTCGTGATCGAGGGCCTGAAGGAGGCCGACAAGGACGAGT TCTGCAGATGCGAGCTGAAGCTGCAGAAGTGGTACGGGGACCTGCG GGGAAAGCCCTTCGCCATCGAAGCCGAGAACAGCATCCTGGACATC AGCGGCTTCAGCAAGCAGTACAACCTGTGCCTTCATCTGGCAGAAGG ACGGCGTGAAGAAGCTGAACCTGTACCTGATCATCAACTACTCAA GGGCGGCAAGCTGCGGTTCAAGAAGATCAAACCTGAAGCCTTCGAA GCCAACAGATTCTACACCGTGATCAACAAAAAGAGCGGCGAGATCG TGCCCATGGAGGTGAACCTCAACTTCGACGACCCCAACCTGATCAT CCTGCCTCTGGCCTTTGGCAAGAGACAGGGCAGAGAATTCATCTGG AACGACCTGCTGTCCCTGGAAACCGGCAGCCTGAAGCTGGCCAACG GAAGAGTGATCGAGAAGACACTGTACAACAGAAGAACCCGGCAGGA TGAGCCTGCCCTGTTCTGTGGCCCTGACCTTCGAGCGGCGGGAGGTC CTGGACTCCTCCAATATCAAACCAATGAACCTGATCGGCGTGGCAA GAGGCGAAAACATCCCCGCCGTGATCGCCCTGACCGACCCCGAGGG CTGCCCCTGAGCCGGTTTAAGGATAGCCTGGGAAACCCAAACCCAC ATCCTGAGAATCGGCGAGAGCTATAAGGAGAAGCAGCGGACCATCC AGGCCAAGAAGGAGGTGGAGCAGCGGAGAGCCGGCGGCTACAGCCG GAAGTAGCCAGCAAAGCCAAGAATCTGGCAGACGATATGGTGAGA AACACCGCTAGAGATCTGCTGTACTACGCCGTGACCCAGGATGCCA TGCTGATCTTCGCCAACCTGAGCCGGGGCTTCGGCCGGCAGGGCAA GCGGACCTTCATGGCCGAGAGACAGTACACACGGATGGAGGACTGG CTGACCGCCAAGCTGGCCTACGAGGGCCTGAGCAAGACCTACCTGT CCAAGACACTGGCCCAGTACACCTCCAAGACATGCAGCAACTGTGG GTTTACCATCACCAGCGCCGACTACGACAGGGTGCTGGAGAAGCTG AAGAAGACAGCAACAGGCTGGATGACCACAATTAACGGCAAGGAGC TGAAGGTGGAGGGCCAGATTACCTACTACAACAGATACAAGAGACA GAACGTAGTCAAGGACCTGTCCGTGAGCTGGATAGACTGAGCGAA GAATCTGTGAACAACGACATCTCCTCCTGGACAAAGGGCAGAAGCG GAGAAGCTCTGAGCCTCCTGAAGAAAAGATTCTCCCATAGACCCGT GCAGGAGAAGTTCTGTGCCTGAACTGCGGCTTCGAGACACACGCA GCCGAGCAAGCCGCCCTGAACATCGCCAGATCCTGGCTGTTCTTCGC GGAGCCAGGAGTACAAGAAATACCAGACAAACAAGACAACCGGCAA CACCGATAAGAGAGCCTTCGTGAGACCTGGCAGTCCTTTTACCGG AAGAAGCTTAAGGAGGTGTGGAACCTGCCGTG	
	Linker + buffer sequence	CGGTCTGGCGGATCTGGCGGAGGCTCCACAAGC	3050

dXR or LTRP ID	Component	DNA sequence	SEQ ID NO:
	ZIM3-KRAB	ATGAACAACCTCCAGGGCAGAGTGACCTTCGAGGACGTGACCGTGA ATTTTACACAGGGAGAGTGGCAGAGACTGAACCCCGAGCAGAGAAA CCTGTACCGGGATGTGATGCTGGAAAACTACAGCAATCTGGTGTCC GTGGGCCAGGGCGAGACCACAAAGCCTGACGTGATCCTGCGTCTGG AGCAGGGCAAGGAACCCCTGGCTGGAGGAGGAGGAGGTGCTGGGAAG CGGACGGGCCGAGAAGAACGGCGACATCGGCGGACAGATCTGGAAG CCTAAGGACGTGAAAGAAAGCCTG	3051
	Buffer sequence + NLS	ACCAGCCCCAAGAAAAAGAGAAAAGTC	3052
	Tag	GACTACAAGGATGACGATGACAAGGACTACAAGGATGACGACGACA AG	3053
	STOP codon + buffer sequence	TAATAGATAAGCGGCCGCTTAATTAA	3054
	3'UTR	GCTGCCTTCTGCGGGGCTTGCCTTCTGGCCATGCCCTTCTTCTCTC CCTTGACCTGTACCTCTTGGTCTTTGAATAAAGCCTGAGTAGGAA G	3055
	Buffer sequence	TCTAG	3056
	Poly(A) tail	AA AA	3057
LTRP1 (codon-optimize d)	5'UTR	AAATAAGAGAGAAAAGAAGAGTAAGAAGAAATATAAGAGCCACC	3047
	START codon + NLS + buffer sequence + linker	ATGGCCCCCTAAGAAGAAGCGTAAAGTGAGCCGGGTGAACGGCAGCG GCAGCGGCGGCGGC	3058
	START codon + DNMT3A catalytic domain	ATGAACCACGACCAGGAGTTCGACCCCCCTAAGGTGTACCTCCCG TCCCCGCCGAGAAGAGAAAAGCCATCCGGGTCTGAGCCTGTTCGA TGGCATCGCCACCGGTCTGCTGGTGTCTGAAGGACCTGGGCATCCAG GTGGATAGGTACATTGCCTCCGAGGTGTGCGAGGACTCCATCACC TGGGAATGGTGCCTCATCAGGGCAAGATCATGTACGTGGGCGACGT GCGGAGCGTGACACAGAAGCATATCCAGGAGTGGGGCCCTTTTCGAC CTGGTGATCGGCGGCAGCCCTTGCAATGACCTGAGCATCGTGAACC CAGCCCCGAAGGGCCTGTACGAGGGAACCGGCAGACTGTTCTTCGA GTTTTACAGACTGCTGCACGACGCCCGGCCCTAAGGAAGGCGACGAC CGGCCCTTCTTTTGGCTGTTTCGAGAATGTGGTGGCCATGGGAGTCA GCGACAAGCGGGATATTAGCCGGTTCCTGGAGAGCAACCCCGTGAT GATCGATGCCAAGGAAGTGAGCGCCGCCACCGGGCCAGATACTTC TGGGGCAATCTGCCTGGCATGAACAGACCCCTGGCCAGCACCCTGA ACGACAAGCTGGAGCTGCAGGAGTGCCTGGAGCACGGCCGGATCGC CAAGTTCAGCAAGGTGAGAACCATCACCACCCGAAGCAACAGCATC AAACAAGGCAAGGACCAGCACTTTCCTGTGTTTCATGAACGAGAAGG AGGACATCCTGTGGTGTACCGAGATGGAGAGAGTGTTCGGGTTCCT AGTCCACTACACAGATGTCAGCAACATGTCTAGACTGGCCAGACAG AGACTGCTGGGAAGAAGCTGGTCCGTCCCTGTGATCAGACACCTGT TCGCCCCCTCTGAAGGAGTACTTCGCCTGCGTG	3059
	Linker	AGCAGCGGCAACAGCAACGCCAACGCGGGGCCCCAGCTTCTCTA GCGGCCCTGGTGCCACTGTCCCTGAGAGGGAGCCAC	3060

dXR or LTRP ID	Component	DNA sequence	SEQ ID NO:
	DNMT3L interaction domain	ATGGGCCCCATGGAGATCTACAAAACCGTGAGCGCCTGGAAGCGGC AGCCTGTGCGCGTGCTGAGCCTGTTTCGGAATATCGATAAAGTCCT GAAAAGCCTGGGATTCTTGGAGAGCGGCTCTGGCTCCGGCGGTGGC ACCTGAAGTACGTGGAGGATGTGACAAACGTGGTCAGACGGGATG TGGAGAAGTGGGGCCCCCTTCGATCTGGTGTACGGCAGCACCCAAC CCTGGGCAGCTCTTGTGACCGGTGCCCTGGCTGGTACATGTTTCAG TTCCACCGGATCCTGCAGTACGCCCTGCCGAGACAGGAGTCCCAGC GGCCATTCTTTTGATTTCATGGACAACCTTGCTGCTGACCGAGGA TGACCAGGAACTACCACTCGGTTCTTCGAGACCGAAGCCGTGACC CTGCAGGACGTGAGAGGCCGGGACTACCAGAAGCCATGCGGGTGT GGTCCAACATCCCTGGACTGAAAAGCAAGCACGCACCTCTGACCCC TAAAGAAGAGGAGTACCTGCAGGCCAGGTGCGGAGCAGAAGCAAG CTGGACGCCCTAAGGTGGATCTGCTGGTGAAGAAATTGCCTCTGC CCCTGAGAGAGTACTTCAAGTATTTAGCCAGAATAGTCTGCCCT G	3061
	Linker	GGCGGCCCAAGCAGCGGCGCCCTCCTCCAGCGGCGGCAGCCAG CCGGCTCCCAACCTCTACCGAGGAGGGCACCTCTGAGTCCGCCAC CCCCGAGAGCGGCCCTGGCACCTCCACCGAGCCCAGCGAGGGCAGC GCACCCGGCAGCCCTGCCGGCAGCCCCACCTCCACAGAGGAGGGAA CCAGCACCGAGCCCAGCGAAGGCAGCGCCCCAGGCACCAGCACCGA GCCTAGTGAGGGCGGCTCTGGCGGCGGCAGCGCC	3062
	dCasX491	CAGGAGATTAAACGGATCAACAAGATCAGAAGAAGACTTGTGAAAG ACAGCAACACCAAGAAGGCCGGCAAGACAGGCCCATGAAAACCT GCTGGTTAGAGTGATGACACCCGATCTGAGAGAGCGGCTGAAAAAC CTGAGAAAGAAGCCTGAAAATATCCCCAGCCCATCAGCAATACAT CTAGAGCCAACCTGAATAAGCTGCTGACCGATTACACCGAAATGAA GAAGGCGATCCTGCATGTGTACTGGGAAGAGTTCCAGAAGGACCT GTGGGCCCTGATGAGCCGGGTGGCCAGCCTGCCAGCAAGAAGATCG ATCAGAACAAGCTGAAACCTGAGATGGACGAGAAGGGCAACCTGAC CACCGCCGGCTTTGCCTGCTCTCAGTGTGGCCAGCCCCTGTTCTGTG TACAAGCTGGAGCAGGTGTCTGAGAAGGGCAAGGCTTACACCAACT ACTTCGGACGGTGCAATGTGGCCGAGCACGAAAAGCTGATCCTGCT GGCCCAGCTGAAGCCCAGAGAAGGATAGCGACGAAGCCGTGACATAT AGCCTGGGAAAGTTTGGGCAGAGGGCCCTGGATTTCTACAGCATTC ATGTGACCAAGGAGTCCACCCACCCCGTGAAGCCCTGGCCCAGAT CGCCGGAACAGATACGCCCTCCGGACCTGTGGGAAAGGCCCTGAGC GACGCATGTATGGGCACAATCGCCTCCTTCTGTCTAAGTACCAGG ACATCATCATCGAACACCAGAAGGTGGTGAAGGGCAACCAAGAGAG ACTGGAGAGCCTGCGGGAGCTGGCCGGCAAGGAAAACCTGGAATAC CCTAGCGTGACCCTGCCACCTCAGCCTCACACCAAGGAGGGCGTTG ATGCCTACAACGAAGTGATCGCCCGGTGCGAATGTGGGTGAACCT GAACCTGTGGCAGAAGCTGAAGCTAAGCAGAGATGATGCCAAGCCT CTGCTGAGACTGAAGGGATTCCCTTCCTTTCTCTGCTCGAGAGAC AGGCCAACGAAGTGGACTGGTGGGACATGGTGTGTAACGTGAAGAA GCTGATCAACGAGAAAAAGGAGGATGGCAAGGTGTTTTGGCAGAAT CTGGCTGGCTACAAGAGACAGGAAGCCCTGAGACCATACCTGAGCA GCGAGGAAGATCGGAAGAAGGGAAAGAAATTTCGCTCGGTACCAGCT GGGCGACCTGCTGCTGCACCTGAAAAAGAAGCACGGCGAGGACTGG GAAAAGGTGTACGACGAGGCCTGGGAGCGGATTGACAAGAAAGTGG AAGGCCTGAGCAAGCACATCAAGCTGGAAGAGGAACGGAGAAGCGA GGACGCCCAGAGCAAGGCCGCCCTGACCGACTGGCTGCGGGCTAAG GCCAGCTTCGTGATCGAGGGCCTGAAGGAGGCCGACAAGGACGAGT TCTGCAGATGCGAGCTGAAGCTGCAGAGTGGTACGGGGACCTGCG GGGAAAGCCCTTCGCCATCGAAGCCGAGAACAGCATCCTGGACATC AGCGGCTTCAGCAAGCAGTACAACCTGTCCTTCATCTGGCAGAAGG ACGGCGTGAAGAAGCTGAACCTGTACCTGATCATCACTACTTCAA GGGCGGCAAGCTGCGGTTCAAGAAGATCAAACCTGAAGCCTTCGAA	3049

[illegible]

*Components are listed in a 5' to 3' order within the constructs

Table 13: Full-length RNA sequences of dXR1 and LTRP1-ZIM3 mRNA molecules assessed in experiment #1 of this example. Modification ‘mψ’ = N1-methyl-pseudouridine

dXR or LTRP ID	RNA sequence	SEQ ID NO:
dXR1	AAAmψAAGAGAGAAAAGAAGAmψAAGAAGAAAmψAmψAAGAGCCACCAmψGGCCCCmψ ψAAGAAGAAGCGmψAAAGmψGAGCCGGGGCGGCAGCGCGCGGCAGCGCCAGGAGA mψmψAAACGGAmψCAACAAGAmψCAGAAGAAGACmψmψGmψGAAAGACAGCAACACCA AGAAGGCCGGCAAGACAGGCCCCAmψGAAAACCCmψGCMψGmψmψAGAGmψGAmψGA CACCCGAmψCmψGAGAGAGCGGCMψGGAAAACCMψGAGAAAGAAGCCmψGAAAAmψAm ψCCCCCAGCCCAmψCAGCAAmψACAmψCmψAGAGCCAACCMψGAAmψAAGCMψGCMψG ACCGAmψmψACACCGAAAmψGAAGAAGGCGAmψCCmψGCAmψGmψGmψACmψGGGAAG AGmψmψCCAGAAGGACCCmψGmψGGGCCmψGAmψGAGCCGGGmψGGCCAGCCmψGCC AGCAAGAAGAmψCGAmψCAGAACAAGCMψGAAACCMψGAGAmψGGACGAGAAGGGCAA CCmψGACCACCGCCGGCMψmψmψGCCmψGCMψCmψCAGmψGmψGGCCAGCCCCmψGmψ mψCGmψGmψACAAGCMψGGAGCAGGmψGmψCmψGAGAAGGGCAAGGCMψmψACACCAA CMψACmψmψCGGACGmψGCAAmψGmψGGCCGAGCACGAAAAGCMψGAmψCmψGCMψ GGCCAGCMψGAAGCCCCGAGAAGGAmψAGCGACGAAGCCmψGACAmψAmψAGCCmψG GGAAAGmψmψmψGGGCAGAGGGCCCMψGGAmψmψmψCmψACAGCAmψmψCAmψGmψGA CCAAGGAGmψCCACCCACCCCGmψGAAGCCCCmψGGCCAGAmψCGCCGGAAACAGAm ψACGCCmψCCGGACCMψGmψGGGAAAGGCCCMψGAGCGACGCAmψGmψAmψGGGCACA AmψCGCCmψCCmψmψCmψGmψCmψAAGmψACCAGGACAmψCAmψCmψCGAACACCA GAAGGmψGmψGAAGGGCAACCAGAAGAGAmψGGAGAGCCmψGCGGGAGCMψGGCCG GCAAGGAAAACCMψGGAAmψACCCmψAGCGmψGACCCmψGCCACCMψCAGCCmψCACA CCAAGGAGGGCGmψmψGAmψGCCmψACAACGAAGmψGAmψCGCCCGGmψGCGAAmψG mψGGGmψGAACCMψGAACCMψGmψGGCAGAAGCMψGAAGCMψAAGCAGAGAmψGAmψG CCAAGCCmψCmψGCMψGAGACmψGAAGGGAmψmψCCmψmψCCmψmψmψCCmψCmψGG mψCGAGAGACAGGCCAACGAAGmψGGACmψGGmψGGGACAmψGGmψGmψAACGmψ GAAGAAGCMψGAmψCAACGAGAAAAAGGAGGAmψGGCAAGGmψGmψmψmψGGCAGA AmψCmψGGCMψGGCMψACAAGAGACAGGAAGCCCMψGAGACCAmψACCMψGAGCAGCG AGGAAGAmψCGAAGAAGGGGAAAGAAAmψmψCGCMψCGGmψACCAGCMψGGGCGACCM ψGCMψGCMψGCACCMψGGAAAAGAAGCACGGCGAGGACmψGGGGAAGGmψGmψACGA CGAGGCCmψGGGAGCGGAmψmψGACAAGAAAGmψGGAAGGCCmψGAGCAAGCACAmψC AAGCMψGGAAGAGGAACGGAGAAGCGAGGACGCCAGAGCAAGGCCGCCCMψGACCGA CMψGGCMψGCGGGCMψAAGGCCAGCMψmψCGmψGAmψCGAGGGCCmψGAAGGAGGCCG ACAAGGACGAGmψmψCmψGCAGAmψGCGAGCMψGAAGCMψGCAGAAGmψGGmψACGGG GACCMψGCGGGGAAAGCCCMψmψCGCCAmψCGAAGCCGAGAACAGCAmψCCmψGGACA mψCAGCGGCMψmψCAGCAAGCAGmψACAACmψGmψGCCmψmψCAmψCmψGGCAGAAGG ACGGCGmψGAAGAAGCMψGAACCMψGmψACCMψGAmψCAmψCAACmψACmψmψCAAGG GCGGCAAGCMψGCGGmψmψCAAGAAGAmψCAAACCMψGAAGCCmψmψCGAAGCCACA GAmψmψCmψACACCGmψGAmψCAACAAAAAGAGCGGCGAGAmψCGmψGCCAmψGGAG GmψGAACmψmψCAACmψmψCGACGACCCCAACCMψGAmψCAmψCCmψGCCmψCmψGGC CMψmψmψGGCAAGAGACAGGGCAGAGAAmψmψCAmψCmψGGAACGACCMψGCMψGmψC CCmψGGAACCGGCAGCCmψGAAGCMψGGCCAACGGAAGAmψGAmψCGAGAAGACAC mψGmψACAACAGAAGAACC CGGAGGAmψGAGCCmψGCCCMψGmψmψCGmψGGCCCMψ GACCMψmψCGAGCGGCGGGAGGmψCCmψGGACmψCCmψCCAAmψAmψCAAACCAAmψG AACCMψGAmψCGGCGmψGGCAAGAGGCGAAAACAmψCCCCGCCGmψGAmψCGCCCMψG ACCGACCCCGAGGGCMψGCCACmψGAGCCGGmψmψmψAAGGAmψAGCCmψGGGAAAC CCAACCCACAmψCCmψGAGAAmψCGGCGAGAGCMψAmψAAGGAGAAGCAGCGGACCAm ψCCAGGCCAAGAAGGAGGmψGGAGCAGCGGAGAGCCGGCGGCMψACAGCCGGAAGmψA CGCCAGCAAAGCCAAAGAmψCmψGGCAGACGAmψAmψGmψGAGAAACACCCGCMψAGA GAmψCmψGCMψGmψACmψACGCCmψGACCCAGGAmψGCCAmψGCMψGAmψCmψmψCG CCAACCMψGAGCCGGGGCMψmψCGGCCGGCAGGGCAAGCGGACCMψCAmψGGCCGA GAGACAGmψACACACGGAmψGGAGGACmψGGCMψGACCGCCAAGCMψGGCCmψACGAG GGCCmψGAGCAAGACCMψACCMψGmψCCAAGACAmψGGCCAGmψACACCMψCCAAG	3063

dXR or LTRP ID	RNA sequence	SEQ ID NO:
	ACAmψGCAGCAACmψGmψGGGmψmψmψACCmψCACCAGCGCCGACmψACGACAGGGmψGCGmψGGAGAAGCmψGAAGAAGACAGCAACAGGCGmψGGAmψGACCACAAmψmψAACGGCAAGGAGCmψGAAGGmψGGAGGGCCAGAmψmψACCmψACmψACAACAGAmψACAAGAGACAGAACGmψAGmψCAAGGACCmψGmψCCGmψCGAGCmψGGAmψAGACmψGAGCGAAGAAmψCmψGmψGAACAACGACAmψCmψCCmψCCmψGGACAAAGGGCAGAAGCGGAGAAGCmψCmψGAGCCmψCCmψGAAGAAAAGAmψmψCmψCCCmψAGACCCGmψGCAGGAGAAmψmψCmψGmψGCCmψGAACmψGCGGCGmψmψCGAGACACACGCAGCCGAGCAAGCCGCCmψGAACAmψCGCCAGAmψCCmψGGCmψGmψmψCmψGCGGAGCCAGGAGmψACAAGAAAmψACCAGACAAACAAGACAACCGGCAACACCGAmψAAGAGAGCCmψmψCGmψCGAGACCmψGGCAGmψCCmψmψmψmψACCGAAGAAGCmψmψAAGGAGGmψGmψGGAACCGmψGCGGmψCmψGGCGGAmψCmψGGCGGAGGCGmψCCACAAGCAmψGAACAACmψCCAGGGCAGAGmψGACCmψmψCGAGGACGmψGACCmψGAAmψmψmψmψACACAGGGAGAGmψGGCAGAGACmψGAACCCCGAGCAGAGAAACmψGmψACCGGGAmψGmψGAmψGCGmψGGAAGCmψACAGCAAmψCmψGGmψGmψCCGmψGGGCCAGGGCGAGACCACAAGCCmψGACGmψGAmψCCmψGCGmψCmψGGAGCAGGGCAAGGAACCCmψGGCmψGGAGGGAGAGGmψGCGmψGGGAAGCGGACGGGCCGAGAAGAACGGCGACAmψCGGCGGACAGAmψCmψGGAGCCmψAAGGACGmψGAAAGAAAGCCmψGACCAGCCCCAAGAAAGAGAAAGmψCGACmψACAAGGAmψGACGAmψGACAAGGACmψACAAGGAmψGACGACGACAAGmψAAmψAGAmψAAGCGGCCGCGmψmψAAmψmψAAGCmψGCCmψmψCmψGCGGGCGmψmψGCCmψmψCmψGGCCAmψGCCmψmψCmψmψCmψCmψCCCmψmψGCACCmψGmψACCmψCmψmψGGmψCmψmψmψGAAmψAAAGCCmψGAGmψAGGAAGmψcmψagaaaaa aaa aaaaaaaaaaaaaaaaaaaaa	
LTRP1	AAAmψAAGAGAGAAAAGAAGAGmψAAGAAGAAAmψAmψAAGAGCCACCmψGGCCCCmψAAGAAGAAGCGmψAAAGmψGAGCCGGGmψGAACGGCAGCGGCAGCGCGCGGCGGCGAmψGAACCACGACCAGGAGmψmψCGACCCCCCmψAAGGmψGmψACCmψCCCmψCCCCGCGCAGAGAGAGAAAGCCAmψCCGGGmψCCmψGAGCCmψGmψmψCGAmψGGCmψCGCCACCmψGCGmψGGmψGGmψGAAGGACmψGGGCGmψCCAGGmψGAGmψAGGmψACA mψmψGCCmψCCGAGGmψGmψGCGAGGACmψCCAmψCACCmψGGGAAmψGCGmψGCGmψCAGGGCAAGAmψAmψGmψACGmψGGGCGACGmψGCGGAGCGmψGACACAGAAGC AmψAmψCCAGAGmψGGGGCCCGmψmψmψCGACmψGmψGAmψCGGCGGCAGCCCGmψmψGCAAmψGACCmψGAGCmψCGmψGAACCCAGCCCGGAAGGGCCCGmψGmψACGAGGGGA ACCGGCAGACmψGmψmψCmψmψCGAGmψmψmψACAGACmψGCCmψGCACGACGCCCGGGCCCGmψAAGGAAGCGCAGACCGGCCCGmψmψmψmψGGCmψGmψmψCGAGAAmψG mψGGmψGGCCAmψGGGAGmψCAGCGACAAGCGGGAmψAmψmψAGCCGCGmψmψCCmψGGAGAGCAACCCCGmψGAmψGAmψCGAmψGCCAAGGAAGmψGAGCGCCGCCACCGGGGCC AGAmψACmψmψCmψGGGGCAAmψCmψGCCmψGGCmψGAACAGACCCCGmψGGCCAGCA CCGmψGAACGACAAGCmψGGAGCmψGCAGGAGmψGCCmψGGAGCAGCGCCGGAmψCGC CAAGmψmψCAGCAAGGmψGAGAACCmψCACCACCCGAAGCAACAGCAmψCAACAAG GCAGGACCAGCACmψmψmψCCmψGmψGmψmψCmψGAACGAGAAGGAGGACAmψCCm ψGmψGGmψGmψACCGAGAmψGGAGAGAGmψGmψmψCGGGmψmψCCCAGmψCCACmψAC ACAGAmψGmψCAGCAACAmψGmψCmψAGACmψGGCCAGACAGAGACmψGCCmψGGGAAG AAGCmψGGmψCCGmψCCCGmψGmψGAmψCAGACACmψGmψmψCGCCCCmψCmψGAAGG AGmψACmψmψCGCCmψGCGmψGAGCAGCGGCAACAGCAACGCCAACAGCCGGGGCCCC AGCmψmψCmψCmψAGCGGCCCGmψGGmψGCCACmψGmψCCCGmψGAGAGGGAGCCACAmψG GGCCCCAmψGGAGAmψCmψACAAAACCGmψGAGCGCCCGmψGGAAGCGGCAGCCCGmψGmψG CGCGmψGCCmψGAGCCmψGmψmψmψCGGAAmψAmψCGAmψAAAGmψCCmψGAAAAGCCm ψGGGAmψmψCCmψGGAGAGCGGCGmψCmψGGCmψCCGGCGGmψGGCACCmψGAAGmψA CGmψGGAGGAmψGmψGACAAACGmψGmψCAGACGGGAmψGmψGGAGAGmψGGGGCC CCmψmψCGAmψCmψGGmψGmψACGGCAGCACCCAACCCCGmψGGGCAGCmψCmψmψGmψ GACCGGmψGCCCGmψGGCmψGmψACAmψGmψmψmψCAGmψmψCCACCGGAmψCCmψGC AGmψACGCCCGmψGCCGAGACAGGAGmψCCAGCGGCCAmψmψCmψmψmψmψGGAmψmψ mψmψCmψGGACAACmψmψGCCmψGCCmψGACCAGGAmψGACCAGGAAACmψACCACmψ CGGmψmψCCmψGCAGACCGAAGCCCGmψGACCmψGCAGGACGmψGAGAGGCCGGGACm ψACCAGAACGCCAmψGCGGGmψGmψGmψCCAACAmψCCCGmψGGAACmψGAAAAGCAAG CACGCACCGmψCmψGACCCCGmψAAAGAAGAGGAGmψACCmψGCAGGCCAGGmψGCGGA GCAGAAGCAAGCmψGGACGCCCGmψAAGGmψGGAmψCmψGCCmψGGmψGAAGAAmψmψG CCmψCCmψGCCCGmψGAGAGAGmψACmψmψCAAGmψAmψmψmψCAGCCAGAAmψAGmψ	3064

dXR or LTRP ID	RNA sequence	SEQ ID NO:
	CmψGCCCCmψGGGCGGCCCAAGCAGCGCGCCCCmψCCmψCCCAGCGGGCGGAGCCCCA GCCGGCmψCCCCAACmψCmψACCGAGGAGGGCACCmψCmψGAGmψCCGCCACCCCCG AGAGCGGCCmψGGCACCmψCCACCGAGCCCAGCGAGGGCAGCGCACCCGGCAGCCm ψGCCGGCAGCCCCACCmψCCACAGAGGAGGGAACCAGCACCCGAGCCCAGCGAAGGCAG CGCCCCAGGCACCAGCACCCAGCCmψAGmψGAGGGCGGmψCmψGGCGGCGGAGCGC CCAGGAGAmψmψAAACGGAmψCAACAAGAmψCAGAAGAAGAmψmψGmψGAAAGACAG CAACACCAAGAAGGCCGGCAAGACAGGCCCCAmψGAAAACCCmψGmψGGmψmψAGAG mψGAmψGACACCCGAmψCmψGAGAGAGCGGmψGGAACCCmψGAGAAAGAAGCCmψG AAAAmψAmψCCCCAGCCCCAmψCAGCAAmψACAmψCmψAGAGCCAACmψGAAmψAAG CmψGmψGACCGAmψmψACACCGAAAmψGAAGAAGGCGAmψCCmψGCAmψGmψGmψAC mψGGGAAGAAGmψmψCCAGAAGGACCmψGmψGGCCmψGAmψGAGCCGGGmψGGCCCA GCCmψGCCAGCAAGAAGAmψCAGAmψCAGAACAAGCmψGAAACCCmψGAGAmψGGACGAG AAGGGCAACCmψGACCACCGCCGGCmψmψmψGCCmψGCCmψCmψCAGmψGmψGGCCAGC CCmψGmψmψCmψGmψACAAAGCmψGGAGCAGGmψGmψCmψGAGAAGGGCAAGGmψm ψACACCAACmψACmψmψCGGACGGmψGCAAmψGmψGGCCGAGCAGAAAAGCmψGAmψ CCmψGCCmψGGCCAGCmψGAAGCCCGAGAAGGAmψAGCGACGAAGCCGmψGACAmψAm ψAGCCmψGGGAAGmψmψmψGGGCAAGAGGGCCmψGGAmψmψmψCmψACAGAmψmψC AmψGmψGACCAAGGAmψCCACCCACCCCGmψGAAGCCCGmψGGCCAGAmψCGCCGG AAACAGAmψACGCCmψCCGGACCmψGmψGGAAAGGCCmψGAGCGACGCAmψGmψAm ψGGGCACAAmψCGCCmψCCmψmψCmψGmψCmψAAGmψACCAGGACAmψCAmψCAmψC GAACACCAGAAGGmψGGmψGAAGGGCAACCAGAAGAGACmψGGAGAGCCmψGCGGGAG CmψGGCCGGCAAGGAAAACmψGGAAmψACCCmψAGCGmψGACCCmψGCCACCCmψCAG CCmψCACACCAAGGAGGGCGmψmψGAmψGCCmψACAACGAAGmψGAmψCGCCCGGGmψ GCCAAmψGmψGGGmψGAACCmψGAACCmψGmψGGCAGAAGCmψGAAGCmψAAGCAGAG AmψGAmψGCCAAGCCmψCmψGCCmψGAGACmψGAAGGGAmψmψCCmψmψCmψmψmψC CmψCmψGGmψCGAGAGACAGGCCAACGAAGmψGGAACmψGGmψGGGACAmψGGmψGmψG mψAACGmψGAAGAAGCmψGAmψCAACGAGAAAAAGGAGGAmψGGCAAGGmψGmψmψmψ mψGGCAGAAmψCmψGGCmψGGCmψACAAGAGACAGGAAGCCmψGAGCCAmψACCCmψ GAGCAGCGAGGAAGAmψCGGAAGAAGGGAAAGAAAmψmψCGCmψCGGmψACCAGCmψG GGCGACCmψGCCmψGCCmψGCACCmψGGAAAAGAAGCACGGCGAGGACmψGGGGAAAGGm ψGmψACGACGAGGCCmψGGGAGCGGAmψmψGACAAGAAAGmψGGAAGGCCmψGAGCAA GCACAmψCAAGCmψGGAAGAGGAACGGAGAAGCGAGGACGCCAGAGCAAGGCCGCCCC mψGACCGACmψGGCmψGCGGGCmψAAGGCCAGCmψmψCmψGAmψCGAGGGCCmψGAA GGAGGCCGACAAGGACGAGAmψmψCmψGCAGAmψGCGAGCmψGAAGCmψGCAGAAAGmψG GmψACGGGGACCCmψGCGGGGAAGCCmψmψCGCCAmψCGAAGCCGAGAACAGCAmψC CmψGGACAmψCAGCGGCmψmψCAGCAAGCAGmψACAACmψGmψGCCmψmψCAmψCmψG GCAGAAGGACGGCGmψGAAGAAGCmψGAACCmψGmψACCmψGAmψCAmψCAACmψACm ψmψCAAGGGCGGCAAGCmψGCGGmψmψCAAGAAGAmψCAAACCCmψGAAGCCmψmψCGA AGCCAACAGAmψmψCmψACACCmψGAmψCAACAAAAAGAGCGGGCAGAmψCGmψGCC CAmψGGAGGmψGAACmψmψCAACmψmψCGACGACCCCAACCmψGAmψCAmψCCmψGCC mψCmψGGCCmψmψmψGGCAAGAGACAGGGCAGAGAAmψmψCAmψCmψGGAACGACCmψ GCCmψGmψCCmψGGAAACCGGCAGCCmψGAAGCmψGGCCAACGGAAGAmψGAmψCGA GAAGACAmψGmψACAACAGAGAACCCGGCAGGAmψGAGCCmψGCCmψGmψmψCmψ ψGGCCmψGACCmψmψCGAGCGGGCGGAGGmψCCmψGGACmψCCmψCCAAmψAmψCAA ACCAAmψGAACCmψGAmψCGGCmψGGCAAGAGGGCAAAACAmψCCCCGCCmψGAmψ CGCCmψGACCGACCCCGAGGGCmψGCCACmψGAGCCGGmψmψmψAAGGAmψAGCCm ψGGGAACCCCAACCCACAmψCCmψGAGAAmψCGGCGAGAGCmψAmψAAGGAGAAGCAG CGGACCAmψCCAGGCCAAGAAGGAGGmψGGAGCAGCGGAGAGCCGGCGGCGmψACAGCC GGAGAmψACGCCAGCAAAGCCAAGAAmψCmψGGCAGACGAmψAmψGGmψGAGAAACAC CGCmψAGAGAmψCmψGCCmψGmψACmψACGCCmψGACCAGGAmψGCCAmψGCCmψGAm ψCmψmψCGCCAACCmψGAGCCGGGGCmψmψCGGCCGGCAGGGCAAGCGGACCCmψCA mψGGCCGAGAGACAGmψACACACGGAmψGGAGGACmψGGCmψGACCGCCAAGCmψGGC CmψACGAGGGCCmψGAGCAAGACCmψACCmψGmψCCAAGACACmψGGCCAGmψACAC CmψCCAAGACAmψGCAGCAACmψGmψGGmψmψmψACCmψCACCAGCGCCGACmψAC GACAGGGmψGCCmψGGAGAAGCmψGAAGAAGACAGCAACAGGCmψGGAmψGACCACAAm ψmψAACGGCAAGGAGCmψGAAGGmψGGAGGGCCAGAmψmψACCmψACmψACAACAGAm ψACAAGAGACAGAAGCmψAGmψCAAGGACCmψGmψCCGmψCGAGCmψGGAmψAGAmψ GAGCGAAGAAmψCmψGmψGAACAACGACAmψCmψCCmψCCmψGGACAAGGGCAGAAAG	

dXR or LTRP ID	RNA sequence	SEQ ID NO:
	CGGAGAAGCmψCmψGAGCCmψCCmψGAAGAAAAGAmψmψCmψCCCmψAGACCCGmψG CAGGAGAAAmψmψCGmψGmψGCCmψGAACmψGCGGCmψmψCGAGACACACGAGCCGA GCAAGCCGCCmψGAACAmψCGCCAGAmψCCmψGGCmψGmψmψCCmψGCGGAGCCAGG AGmψACAAGAAAmψACCAGACAAACAAGACAACCGCAACACCGAmψAAGAGAGCCmψ mψCGmψCGAGACCmψGGCAGmψCCmψmψmψmψACCAGGAAGAAGCmψmψAAGGAGmψG mψGGAAACCmψGCCGmψGCGGmψCmψGGCGGAmψCmψGGCGGAGGCmψCCACAAGCAm ψGAACAACmψCCCAGGGCAGAGmψGACmψmψCGAGGACGmψGACCGmψGAAmψmψmψ mψACACAGGGAGAGmψGGCAGAGACmψGAACCCGAGCAGAGAAACCmψGmψACCGGG AmψGmψGAmψGmψGGAAACmψACAGCAAmψCmψGGmψGmψCCGmψGGGCCAGGGCG AGACCACAAAGCCmψGACGmψGAmψCCmψGCGmψCmψGGAGCAGGGCAAGGAACCCmψ GGCmψGGAGGAGGAGGAGmψGmψGGGAAGCGGACGGGCCGAGAAGAACGGCGACAm ψCGGCGGACAGAmψCmψGGAAGCCmψAAGGACGmψGAAAGAAAGCCmψGACCAGCCCC AAGAAAAAGAGAAAAGmψCGACmψACAAGGAmψGACGAmψGACAAGGACmψACAAGGA mψGACGACGACAGmψAAmψAGAmψAAGCGGCCGmψmψAAmψmψAAGCmψGCCmψmψ CmψGCGGGGmψmψGCCmψmψCmψGGCCAmψGCCmψmψCmψmψCmψCmψCCmψmψG CACmψGmψACCmψCmψmψGmψCmψmψmψGAAmψAAAGCCmψGAGmψAGGAAGmψcm ψagaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	

Table 14: Full-length protein sequences of dXR1 and LTRP1-ZIM3 molecules assessed in experiment #1 of this example

dXR or LTRP ID	AA sequence	SEQ ID NO:
dXR1	MAPKKKRKVSRRGSGGSAQEIKRINKIRRLVKDSNTKKAGKTGPMKTLVRVMT PD LRRERLENLRKKPENIPQISNTSRANLNKLLTDYTEMKKAILHVYWEFQKDPVGLMS RVAQPASKKIDQNKLPKPEMDEKGNLTAGFACSQCQPLFVYKLEQVSEKQAYTNYF GRCNVAEHEKLI LLAQLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLA QIAGNRYASGPVGKALSDACMGTIASFLSKYQDIIIEHQKVVKGNQKRLESRLREL LAG ENLEYP SVTLP PPHTEGVDAYNEVIARVMWVNLNLWQKLKLSRDDAKPLRLKGF PSFPLVERQANEVDWDMVCNVKKLINEKKEDGKVFQNLAGYKRQEALRPYLSSEED RKKGKKFARYQLGDLHLHLEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKLEEERSE DAQSKAALTDWLRAKASFVIEGLKEADKDEFRCCELKLQKWYGDLRGKPFATAEANSI LDISGFSKQYNCAFIWQKDGVKKLNLYLIINYFKGGKLRFKKIKPEAFANRFYTVIN KKSGEIVPMEVNFNFDPNLIIPLAFGKRQGREFIWNDDLSETGSLKLANGRVIEK TLYNRRTRQDEPALFVALTFERREVLDSNIIKPMNLIGVARGENIPAVIALTDPEGCP LSRFKDSLGNPTHILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKAKNLADDMV RNTARDLLYYAVTQDAMLI FANLSRGFGRQGKRTFMAERQYTRMEDWLTAKLAYEGLS KTYLSKTLAQYTSKTC SNCGFTITSADYDRVLEKLKKTATGWMTTINGKELKVEGQIT YYNRYKRQNVVKDLSELDRLSEESVNNDISSWTKGRSGEALSLLKKRFSHRPVQEF VCLNCGFETHAAEQAALNIARSWLFLRSQEYKKYQTNKTTGNTDKRAFVETWQSFYRK KLKEVWKPAVRSRGGSGGSTSMNNSQGRVTFEDVTNFTQGEWQRLNPEQRNLRYDVM LENYSNLVSVGQGETTKPDVILRLEQGKEPWLEEEVGLSGRAEKNNGDIGGQIWPKPD VKESLTSPPKKRKV	3065
LTRP1	MAPKKKRKVSRRVNGSGSGGMNHDQEFDPKVPVPAEKRKPIRVLSLFDGIATGLL VLKDLGIQVDRIYASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIEWGFDFLVIG GSPCNDLSIVNPARKGLYEGTGRLEFFEFYRLHLDARPKGEDDRPFFWLFENVVAMGVS DKRDISRFLESNPVMI DAKEVSAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHG RIAKFSKVRTITTRSNSIKQGDQHFVPFMNEKEDILWCTEMERVFGFPVHYTDVSNM SRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSSFSSGLVPLSLRGS HMGPMETIKTVSAWKRPVRVLSLFRNIDKVLKSLGFLES GSGSGGT LKYVEDVTNV VRRDVEKWGPFDFLVYGSTQPLGSSCDRCPGWYMFQFHRI LQYALPRQESQRPFFWIFM DNLLLTEDDQETTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEE YLQAQVRSRSLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGPSSGAPPSPSGGSPA	3066

dXR or LTRP ID	AA sequence	SEQ ID NO:
	GSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGT STEPSEGGSGGSAQEIKRINKIRRRRLVKDSNTKKAGKTGPMKTLVVRVMTPLDRERL ENLRKKPENIPQPISENTSRANLNKLLTDYTEMKKAILHVYWEETFQKDPVGLMSRVAQP ASKKIDQNKLPKPEMDEKGNLTTAGFACSQCGQPLFVYKLEQVSEKGKAYTNYFGRCNV AEHEKLILLAQLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGN RYASGPVGKALSDACMGTIASFLSKYQDIIIEHQKVVGKNQKRLESRELAGKENLEY PSVTLPQPHTKEGVDAVNEVIARVMWVNLNLWQKLKLSRDDAKPLLRLKGFPSPFL VERQANEVDWDDMVCNVKKLINEKKEDGKVFQNLAGYKRQEALRPYLSSEEDRKKGK KFARYQLGDLHLEKKHGEDWGKVDYDEAWERIDKKVEGLSKHIKLEEEERSEDAQSK AALTDWLRAKASFVIEGLKEADKDEFRCCELKQKQWYGDRLRGKPFPAIEAENSILDISG FSKQYNCAFIWQKDGKVLNLYLIINYFKGGKLRFKKIKPEAFEANRFYTVINKSGE IVPMEVNFNDDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGRVIEKTLYNR RTRQDEPALFVALTFERREVLDSSNIKPMNLIGVARGENIPAVIALTDPEGCPLSRFK DSLGNPTHILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKAKNLADDMVRNTAR DLLYYAVTQDAMLI FANLSRGFGRGKRTFMAERQYTRMEDWLTAKLAYEGLSKTYLS KTLAQYTSKTCSNCGFTITSADYDRVLEKLKKTATGWMTTINGKELKVEGQITYYNRY KRQNVVKDLSELDRLEESVNNDISSWTKGRSGEALSLLKKRFSHRPVQEKFVCLNC GFETHAAEQAALNIARSWLFLRSQYKQYQTNKTTGNTDKRAVFETWQSFYRKKLKEV WKPAVRSGGSGGGSTSMNNSQGRVTFEDVTVNFQGEWQRLNPEQRNLYRDVMLENYS NLVSVQGGETTKPDVILRLEQKGEPWLEEEVLGSGRAEKNGDIGQIWKPKDVKESL TSPKKKKRV	

Table 15: Encoding DNA and RNA sequences of the LTRP5-ZIM3 and LTRP5-ZIM3-ADD mRNA molecules assessed in this example*

LTRP molec ule	Compone nt	DNA sequence	SEQ ID NO	RNA sequence	SEQ ID NO
LTRP5 -ZIM3	5'UTR	AAATAAGAGAGAAAAG AAGAGTAAGAAGAAAT ATAAGAGCCACC	3047	AAAUAAAGAGAGAAAAGAAGAGUAAGAA GAAAUAAAGAGCCACC	3115
	START codon + NLS + linker	ATGGCCCCTAAGAAGA AGCGTAAAGTGAGCCG G	3104	AUGGCCCCUAAGAAGAAGCGUAAAGUG AGCCGG	3116

LTRP molecule	Component	DNA sequence	SEQ ID NO	RNA sequence	SEQ ID NO
	START codon + DNMT3A catalytic domain	See sequence in Table 12	3059	AUGAACCACGACCAGGAGUUCGACCCC CCUAAGGUGUACCCUCCCGUCCCGCC GAGAAGAGAAAGCCCAUCCGGGUCCUG AGCCUGUUCGAUGGCAUCGCCACCGGU CUGCUGGUGCUGAAGGACCUGGGCAUC CAGGUGGAUAGGUACAUGCCUCCGAG GUGUGCGAGGACUCCAUCACCGUGGGA AUGGUGCGUCAUCAGGGCAAGAUCAUG UACGUGGGCGACGUGCGGAGCGUGACA CAGAAGCAUAUCCAGGAGUGGGGCCCU UUCGACCUGGUGAUCGGCGGCAGCCCU UGCAAUGACCUGAGCAUCGUAACCCA GCCCCGAAGGGCCUGUACGAGGGAACC GGCAGACUGUUCUUCGAGUUUACAGA CUGCUGCACGACGCCCGGCCUAAGGAA GCGGACGACCGGCCCUUCUUUGGCUG UUCGAGAAUGUGGUGGCCAUGGGAGUC AGCGACAAGCGGGAUAUUAGCCGGUUC CUGGAGAGCAACCCCGUGAUGAUCGAU GCCAAGGAAGUGAGCGCCGCCACCGG GCCAGAUACUUCUGGGGCAUUCUGCCU GGCAUGAACAGACCCUGGCCAGCACC GUGAACGACAAGCUGGAGCUGCAGGAG UGCCUGGAGCACGGCCGGAUCGCCAAG UUCAGCAAGGUGAGAACCAUCACCACC CGAAGCAACAGCAUCAAACAAGGCAAG GACCAGCACUUCCUGUGUUCGAUAAAC GAGAAGGAGGACAUCUUGUGGUUACC GAGAUGGAGAGAGUGUUCGGGUUCCCA GUCCACUACACAGAUGUCAGCAACAUG UCUAGACUGGCCAGACAGAGACUGCUG GGAAGAAGCUGGUCCGUCCUGUGAUC AGACACCUGUUCGCCCCUCUGAAGGAG UACUUCGCCUGCGUG	3117
	Linker	AGCAGCGGCAACAGCA ACGCCAACAGCCGGGG CCCCAGCTTCTCTAGC GGCTTGGTGCCACTGT CCCTGAGAGGGAGCCA C	3060	AGCAGCGGCAACAGCAACGCCAACAGC CGGGGCCCCAGCUUCUCUAGCGGCCUG GUGCCACUGUCCUGAGAGGGAGCCAC	3118

LTRP molecule	Component	DNA sequence	SEQ ID NO	RNA sequence	SEQ ID NO
	DNMT3L interaction domain	See sequence in Table 12	3061	AUGGGCCCCAUGGAGAUCAACAAAACC GUGAGCGCCUGGAAGCGGCAGCCUGUG CGCGUGCUGAGCCUGUUUCGGAUAUC GAUAAAGUCCUGAAAAGCCUGGGAUUC CUGGAGAGCGGCUCUGGCUCGCGCGGU GGCACCCUGAAGUACGUGGAGGAUGUG ACAAACGUGGUCAGACGGGAUGUGGAG AAGUGGGGCCCCUUCGAUCUGGUGUAC GGCAGCACCCAACCCUGGGCAGCUCU UGUGACCGGUGCCCUGGCUGGUACAUG UUUCAGUUCACCGGAUCCUGCAGUAC GCCCUGCCGAGACAGGAGUCCCAGCGG CCAUUCUUUUGGAUUUUAUGGACAAC UUGCUGCUGACCGAGGAUGACCAGGAA ACUACCACUCGGUUCUGCAGACCGAA GCCGUGACCCUGCAGGACGUGAGAGGC CGGGACUACCAGAACGCCAUGCGGGUG UGGUCCAACAUCCUGGACUGAAAAGC AAGCACGCACCUCUGACCCCUAAAGAA GAGGAGUACCUGCAGGCCAGGUGCGG AGCAGAAGCAAGCUGGACGCCCCUAAG GUGGAUCUGCUGGUGAAGAAUUGCCUC CUGCCCCUGAGAGAGUACUUCAGUAU UUCAGCCAGAAUAGUCUGCCCCUG	3119
	Linker	GGAGGCAGCGGCGGGCGC	3105	GGAGGCAGCGGCGGGCGGC	3105
	ZIM3-KRAB	See sequence in Table 12	3051	AUGAACAACUCCAGGGCAGAGUGACC UUCGAGGACGUGACCGUGAAUUUUAACA CAGGGAGAGUGGCAGAGACUGAACCCC GAGCAGAGAAACCUGUACCGGGAUGUG AUGCUGGAAAACUACAGCAAUCUGGUG UCCGUGGGCCAGGGCAGACCACAAAG CCUGACGUGAUCCUGCGUCUGGAGCAG GGCAAGGAACCCUGGCUGGAGGAGGAG GAGGUGCUGGGAAGCGGACGGGCCGAG AAGAACGGCGACAUCGGCGGACAGAU UGGAAGCCUAAGGACGUGAAAGAAAGC CUG	3120
	Linker	GGCGGCCCAAGCAGCG GCGCCCCCTCTCCAG CGGCGGCAGCCAGCC GGCTCCCCAACCTCTA CCGAGGAGGGCACCTC TGAGTCCGCCACCCCC GAGAGCGGCCCTGGCA CCTCCACCGAGCCCAG CGAGGGCAGCGCACCC GGCAGCCCTGCCGGCA GCCCCACCTCCACAGA GGAGGGAACCAGCACC GAGCCCAGCGAAGGCA GCGCCCCAGGCACCAG CACCGAGCCTAGTGAG	3106	GGCGGCCCAAGCAGCGGCGCCCCUCCU CCCAGCGGCGGCAGCCAGCCGGCUCU CCAACCUUACCGAGGAGGGCACCUCU GAGUCCGCCACCCCCGAGAGCGGCCCU GGCACCUCACCGAGCCCAGCGAGGGC AGCGCACCCGGCAGCCUGCCGGCAGC CCCACCUCCACAGAGGAGGGAACCAGC ACCGAGCCCAGCGAAGGCAGCGCCCCA GGCACCAGCACCGAGCCUAGUGAG	3121

LTRP molec ule	Compo nent	DNA sequence	SEQ ID NO	RNA sequence	SEQ ID NO
	dCasX491	See sequence in Table 12	3049	CAGGAGAUUAAACGGAUCAACAAGAUC AGAAGAAGACUUGUGAAAAGACAGCAAC ACCAAGAAGGCCGGCAAGACAGGCCCC AUGAAAACCCUGCUGGUUAGAGUGAUG ACACCCGAUCUGAGAGAGCGGCUGGAA AACCUGAGAAAGAAGCCUGAAAAUAUC CCCCAGCCCAUCAGCAAUACAUUAGA GCCAACCUGAAUAAGCUGCUGACCGAU UACACCGAAUAGAAGGCGAUCCUG CAUGUGUACUGGGAAGAGUCCAGAAG GACCCUGUGGGCCUGAUGAGCCGGGUG GCCCAGCCUGCCAGCAAGAAGAUUGAU CAGAAACAGCUGAAACCGAGAUUGGAC GAGAAGGGCAACCUGACCACCGCCGGC UUUGCCUGCUCUCAGUGUGGCCAGCCC CUGUUCGUGUACAAGCUGGAGCAGGUG UCUGAGAAGGGCAAGGCUUACACCAAC UACUUCGACGGUGCAUUGUGGCCGAG CACGAAAAGCUGAUCCUGCUGGCCCAG CUGAAGCCCAGAAAGGAUAGCGACGAA GCCGUGACAUUAGCCUGGGAAGUUU GGGCAGAGGGCCUGGAUUUCUACAGC AUUCAUGUGACCAAGGAGUCCACCCAC CCCGUGAAGCCCCUGGCCCAGAUCCGC GGAAACAGAUACGCCUCCGGACCUGUG GGAAAGGCCUGAGCGACGCAUGUAUG GGCACAACGCCUCCUCCUGUCUAAG UACCAGGACAUCAUCAUCGAACACCAG AAGGUGGUGAAGGGCAACCAGAAAGAGA CUGGAGAGCCUGCGGGAGCUGGCCGGC AAGGAAAACCGGGAUACCCUAGCGUG ACCCUGCCACCUCAGCCUCACACCAAG GAGGGCGUUGAUGCCUACAACGAAGUG AUCGCCC GGUGCGAAUGUGGGUGAAC CUGAACCUUGGGCAGAAGCUGAAGCUA AGCAGAGAUGAUGCCAAGCCUCUGCUG AGACUGAAGGGAUUCUUCCUUCUCCU CUGGUCGAGAGACAGGCCAACGAAGUG GACUGGUGGGACAUGGUGUGUAACGUG AAGAAGCUGAUCAACGAGAAAAAGGAG GAUGGCAAGGUGUUUUGGCAGAAUCUG GCUGGCUACAAGAGACAGGAAGCCUG AGACCAUACCUGAGCAGCGAGGAAGAU CGGAAGAAGGGAAGAAAUUCGUCUGG UACCAGCUGGGCGACCUGCUGCUGCAC CUGGAAAAGAAGCACGGCGAGGACUGG GGAAGGUGUACGACGAGGCCUGGGAG CGGAUUGACAAGAAAGUGGAAGGCCUG AGCAAGCACAUCAAGCUGGAAGAGGAA CGGAGAAGCGAGGACGCCAGAGCAAG GCCGCCUGACCGACUGGCUGCGGGCU AAGGCCAGCUUCGUAUCGAGGGCCUG AAGGAGGCCGACAAGGACGAGUUCUGC AGAUGCGAGCUGAAGCUGCAGAAGUGG UACGGGGACCUGCGGGGAAAGCCUUC GCCAUCGAAGCCGAGAACAGCAUCCUG GACAUCAGCGGCUUCAGCAAGCAGUAC AACUGUGCCUUCUUGGCAGAAGGAC	3122

LTRP molecule	Component	DNA sequence	SEQ ID NO	RNA sequence	SEQ ID NO
				GGCGUGAAGAAGCUGAACCGUACCUG AUCAUCAACUACUUC AAGGGCGCAAG CUGCGGUUCAAGAAGAUCAAACUGAA GCCUUCGAAGCCAACAGAUUCUACACC GUGAUCAACAAAAAGAGCGGCGAGAU GUGCCCAUGGAGGUGAACUUCACUUC GACGACCCCAACCGAUCAUCCUGCCU CUGGCCUUGGCAAGAGACAGGGCAGA GAAUUCUUCUGGAACGACCUGCUGUCC CUGGAAACCGGCAGCCUGAAGCUGGCC AACGGAAGAGUGAUCGAGAAGACACUG UACAACAGAAGAACCCGGCAGGAUGAG CCUGCCUGUUCGUGGCCUGACCUUC GAGCGCGGGAGGUCCUGGACUCCUCC AAUAUCAAAACCAUGAACCGUAUCGGC GUGGCAAGAGGCGAAAACAUCGCCGCC GUGAUCGCCCUGACCGACCCCGAGGGC UGCCACUGAGCCGGUUUAAGGAUAGC CUGGGAACCCAACCCACAUCUGAGA AUCGGCGAGAGCUAUAAGGAGAAGCAG CGGACCAUCCAGGCCAAGAAGGAGGUG GAGCAGCGGAGAGCCGGCGGCUACAGC CGGAAGUACGCCAGCAAAGCCAAGAAU CUGGCAGACGAUAUGGUGAGAAACACC GCUAGAGAUUCGUGUACUACGCCGUG ACCCAGGAUGCCAUUCUGAUUCUCCGCC AACCGAGCCGGGGCUUCGGCCGGCAG GGCAAGCGGACCUUCAUGGCCGAGAGA CAGUACACACGGAUGGAGGACUGGCUG ACCGCCAAGCUGGCCUACGAGGGCCUG AGCAAGACCUACCUGUCCAAGACACUG GCCAGUACACCUCCAAGACAUGCAGC AACUGUGGGUUUACCAUACCCAGCGCC GACUACGACAGGGUGCUGGAGAAGCUG AAGAAGACAGCAACAGGCUGGAUGACC ACAAUUAACGGCAAGGAGCUGAAGGUG GAGGGCCAGAUUACCUACUACACAGA UACAAGAGACAGAACGUAGUCAAGGAC CUGUCCGUCGAGCUGGAUAGACUGAGC GAAGAAUCUGUGAACACGACAUCC UCCUGGACAAAGGGCAGAAGCGGAGAA GCUCUGAGCCUCCUGAAGAAAAGAUUC UCCCAUAGACCCGUGCAGGAGAAGUUC GUGUGCCUGAACUGCGGCUUCGAGACA CACGCAGCCGAGCAAGCCGCCUGAAC AUCGCCAGAUCCUGGCUGUCCUGCGG AGCCAGGAGUACAAGAAAUACCGACA AACAAAGACAACCGCAACACCGAUAG AGAGCCUUCGUCGAGACCUGGCAGUCC UUUUACCGGAAGAAGCUUAAGGAGGUG UGGAAACCGCCGUG	
	Buffer + linker	CGGTCTGGCGGATCTG GCGGAGGCTCCACCAG C	3107	CGGUCUGGCGGAUCUGGCGGAGGCUCC ACCAGC	3123
	NLS + STOP codon +	CCCAAGAAAAAGAGAA AAGTCTAATAGATAA	3108	CCCAAGAAAAAGAGAAAAGUCUAAUAG AUAA	3124

LTRP molecule	Component	DNA sequence	SEQ ID NO	RNA sequence	SEQ ID NO
	buffer sequence				
	3'UTR	GCTGCCTTCTGCGGG CTTGCCTTCTGGCCAT GCCCTTCTTCTCTCCC TTGCACCTGTACCTCT TGGTCTTTGAATAAAG CCTGAGTAGGAAG	3055	GCUGCCUUCUGCGGGGCUUGCCUUCUG GCCAUGCCCUUCUUCUCCCCUUGCAC CUGUACCUCUUGGUCUUUGAAUAAAGC CUGAGUAGGAAG	3125
	Buffer sequence	TCTAG	3056	UCUAG	3126
	Poly(A) tail	AAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAA	3109	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	3109
LTRP5-ADD-ZIM3	5'UTR	<i>See sequence above</i>	3047	<i>See sequence above</i>	3115
	START codon + NLS + linker	<i>See sequence above</i>	3104	<i>See sequence above</i>	3116
	START codon + DNMT3A ADD domain	ATGGAACGCCTCGTCT ACGAGGTGCGGCAGAA GTGCAGAAACATCGAG GACATCTGCATCTCCT GCGGATCTCTGAACGT GACCCCTGGAGCACCCA CTGTTTCATCGGCGGCA TGTGCCAGAACTGTAA AAACTGTTTCTTGGAG TGTGCCTATCAATACG ACGATGACGGCTACCA GAGTACTGCACCATC TGTTGCGGCGGAAGAG AGGTGCTGATGTGTGG AAATAACAACCTGCTGC CGGTGCTTCTGCGTGG AATGCGTGGACCTGCT GGTGGGCCCCGGCGCC GCCAGGCCGCTATTA AGGAAGATCCTTGGAA CTGCTACATGTGCGGC CACAAGGGCACATACG GCCTGCTGAGACGGAG AGAGGACTGGCCTAGC AGACTGCAGATGTTCT TCGCCAAT	3111	AUGGAACGCCUCGUCUACGAGGUGCGG CAGAAGUGCAGAAACAUCGAGGACAUC UGCAUCUCCUGCGGAUCUCUGAACGUG ACCCUGGAGCACCCACUGUUCUACGGC GGCAUGUGCCAGAACUGUAAAAACUGU UUUCUGGAGUGUGCCUAUCAAUACGAC GAUGACGGCUACCAGAGCUACUGCACC AUCUGUUGCGGCGGAAGAGAGGUGCUG AUGUGUGGAAUAACAACUGCUGCCGG UGCUUCUGCGUGGAAUGCGUGGACCUG CUGGUGGGCCCCGGCGCCGCCAGGCC GCUAUUAAGGAAGAUCCUUGGAACUGC UACAUGUGCGGCCACAAGGGCACAUAC GGCCUGCUGAGACGGAGAGAGGACUGG CCUAGCAGACUGCAGAUUUCUUCGCC AAU	3127
	DNMT3A catalytic domain	AACCACGACCAGGAGT TCGACCCCCCTAAGGT GTACCCCTCCCGTCCCC GCCGAGAAGAGAAAGC CCATCCGGGTCCTGAG CCTGTTTCGATGGCATC GCCACCGTCTGCTGG	3112	AACCACGACCAGGAGUUCGACCCCCU AAGGUGUACCCUCCCGUCCCCGCCGAG AAGAGAAAGCCCAUCCGGGUCCUGAGC CUGUUCGAUGGCAUCGCCACCGGUCUG CUGGUGCUGAAGGACCUGGGCAUCCAG GUGGAUAGGUACAUUGCCUCCGAGGUG UGCGAGGACUCCAUACCGUGGGAAUG	3128

LTRP molecule	Component	DNA sequence	SEQ ID NO	RNA sequence	SEQ ID NO
		TGCTGAAGGACCTGGG CATCCAGGTGGATAGG TACATTGCCTCCGAGG TGTGCGAGGACTCCAT CACCGTGGGAATGGTG CGTCATCAGGGCAAGA TCATGTACGTGGGCGA CGTGCGGAGCGTGACA CAGAAGCATATCCAGG AGTGGGGCCCTTTCTGA CCTGGTGATCGGCGGC AGCCCTTGCAATGACC TGAGCATCGTGAACCC AGCCCGGAAGGGCCTG TACGAGGGAACCGGCA GACTGTTCTTCGAGTT TTACAGACTGCTGCAC GACGCCCGGCCTAAGG AAGGCGACGACCGGCC CTTCTTTTGGCTGTTC GAGAATGTGGTGGCCA TGGGAGTCAGCGACAA GCGGGATATTAGCCGG TTCCTGGAGAGCAACC CCGTGATGATCGATGC CAAGGAAGTGAGCGCC GCCCACCGGGCCAGAT ACTTCTGGGGCAATCT GCCTGGCATGAACAGA CCCCTGGCCAGCACCG TGAACGACAAGCTGGA GCTGCAGGAGTGCCTG GAGCACGGCCGGATCG CCAAGTTCAGCAAGGT GAGAACCATCACCACC CGAAGCAACAGCATCA AACAAGGCAAGGACCA GCACTTTCCTGTGTTC ATGAACGAGAAGGAGG ACATCCTGTGGTGTAC CGAGATGGAGAGAGTG TTCGGGTTCCTAGTCC ACTACACAGATGTCAG CAACATGTCTAGACTG GCCAGACAGAGACTGC TGGGAAGAAGCTGGTC CGTCCCTGTGATCAGA CACCTGTTGCCCCCTC TGAAGGAGTACTTCGC CTGCGTG		GUGCGUCAUCAGGGCAAGAUCAUGUAC GUGGGCGACGUGCGGAGCGUGACACAG AAGCAUAUCCAGGAGUGGGGCCCCUUC GACCUGGUGAUCGGCGGCAGCCCUUGC AAUGACCUGAGCAUCGUGAACCCAGCC CGGAAGGGCCUGUACGAGGGAACCGGC AGACUGUUCUUCGAGUUUUACAGACUG CUGCACGACGCCCCGCCUAAGGAAGGC GACGACCGGCCCCUUCUUUGGCGUUC GAGAAUGUGUGGCCAUGGGAGUCAGC GACAAGCGGGAUAUUAGCCGGUCCUG GAGAGCAACCCCGUGAUGAUCGAUGCC AAGGAAGUGAGCGCCGCCACCGGGCC AGAUACUUCUGGGGCAAUCUGCCUGGC AUGAACAGACCCCUGGCCAGCACCGUG AACGACAAGCUGGAGCUGCAGGAGUGC CUGGAGCACGGCCGGAUCGCCAAGUUC AGCAAGGUGAGAACCAUACCACCCGA AGCAACAGCAUCAACAAGGCAAGGAC CAGCACUUUCCUGUGUUAUGAACGAG AAGGAGGACAUCUGUGGUGUACCGAG AUGGAGAGAGUGUUCGGGUUCCAGUC CACUACACAGAUGUCAGCAACAUUCU AGACUGGCCAGACAGAGACUGCUGGGA AGAAGCUGGUCCGUCCUGUGAUCAGA CACCUGUUCGCCCCUCUGAAGGAGUAC UUCGCCUGCGUG	
	Linker	<i>See sequence above</i>	3060	<i>See sequence above</i>	3118
	DNMT3L interaction domain	<i>See sequence above</i>	3061	<i>See sequence above</i>	3119
	Linker	<i>See sequence above</i>	3105	<i>See sequence above</i>	3105

LTRP molecule	Component	DNA sequence	SEQ ID NO	RNA sequence	SEQ ID NO
	ZIM3-KRAB	<i>See sequence in Table 12</i>	3051	<i>See sequence above</i>	3120
	Linker	<i>See sequence above</i>	3106	<i>See sequence above</i>	3121
	dCasX491	<i>See sequence above</i>	3049	<i>See sequence above</i>	3122
	Buffer + linker	<i>See sequence above</i>	3107	<i>See sequence above</i>	3123
	NLS + STOP codon + buffer sequence	<i>See sequence above</i>	3108	<i>See sequence above</i>	3124
	3'UTR	<i>See sequence above</i>	3055	<i>See sequence above</i>	3125
	Buffer sequence	<i>See sequence above</i>	3056	<i>See sequence above</i>	3126
	Poly(A) tail	<i>See sequence above</i>	3109	<i>See sequence above</i>	3109

*Components are listed in a 5' to 3' order within the constructs

Table 16: Full-length RNA sequences of LTRP5-ZIM3 and LTRP5-ZIM3-ADD mRNA molecules assessed in this example. Modification ‘mψ’ = N1-methyl-pseudouridine

LTRP molecule	RNA sequence	SEQ ID NO
LTRP5-ZIM3	AAAmψAAGAGAGAAAAGAAGAGmψAAGAAGAAAmψAmψAAGAGCCACCAmψGGCCCC mψAAGAAGAAGCGmψAAAGmψGAGCCGGAmψGAACCACGACCAGGAGmψmψCGACCC CCmψAAGGmψGmψACCCmψCCCmψCCCCGCCGAGAAGAGAAAGCCCmψCCGGGm ψCCmψGAGCCmψGmψmψCGAmψGGCmψCGCCACCGGmψCmψGmψGGmψGmψGAA GGACmψGGGCmψCCAGmψGGAmψAGmψACAmψmψGCCmψCCGAGGmψGmψGCC AGGACmψCCAmψCACCmψGGGAAmψGmψGCCmψCmψCAGGGCAAGAmψCmψGm ψACGmψGGGCGACmψGCCGAGCGmψGACACAGAAGCmψAmψCCAGGAGmψGGGGC CCmψmψmψCGACCCmψGGmψGAmψCGGCGGCAGCCmψmψGCAAmψGACmψGAGCm ψCGmψGAACCCAGCCCGGAAGGGCCmψGmψACGAGGGAACCGGCAGACmψGmψmψCm ψmψCGAGmψmψmψACAGACmψGmψGCACGACGCCCGCCmψAAGGAAGGCGACG ACCGGCCmψmψmψmψGGCmψGmψmψCGAGAAmψGmψGGmψGGCCmψGGGA GmψCAGCGACAAGCGGGAmψAmψmψAGCCGmψmψCmψGGAGAGCAACCCCGmψGA mψGAmψCGAmψGCCAAGGAAGmψGAGCGCCGCCACCGGCCAGAmψACmψmψCmψG GGGCAAmψCmψGCCmψGGCmψGAACAGACCCmψGGCCAGCACCGmψGAACGACAA GCmψGGAGCmψGCAGGAGmψGCCmψGGAGCACGGCCGGAAmψCGCCAAGmψmψCAGCA AGGmψGAGAACCmψCACCACCCGAAGCAACAGCmψCAAACAAGGCAAGGACCAGC ACmψmψmψCmψGmψGmψCmψGAAACGAGAAGGAGGAGCmψCmψGmψGGmψGmψ ACCGAGAmψGGAGAGAGmψGmψmψCGGGmψmψCCAGmψCCACmψACACAGAmψGmψ CAGCAACAmψGmψCmψAGACmψGGCCAGACAGAGACmψGmψGGGAAGAAGCmψGGm ψCCGmψCCmψGmψGAmψCAGACCCmψGmψmψCGCCCCmψCmψGAAGGAGmψACmψ mψCGCCmψGCCmψGAGCAGCGGCAACAGCAACGCCAACAGCCGGGGCCCCAGCmψmψ CmψCmψAGCGGCCmψGGmψGCCAmψGmψCCmψGAGAGGGAGCCACAmψGGGCCCC AmψGGAGAmψCmψACAAAACCGmψGAGCGCCmψGGAAGCGGCAGCCmψGmψGCCCGm ψGCCmψGAGCCmψGmψmψCGGAAmψAmψCGAmψAAAGmψCCmψGAAAAGCCmψGGG AmψmψCCmψGGAGAGCGGCGmψCmψGGCmψCCGGCGGmψGGCACCCmψGAAAmψACGm ψGGAGAGmψGmψGACAAACGmψGmψCAGACGGGAmψGmψGGAGAAAGmψGGGGCCCC mψmψCGAmψCmψGGmψGmψACGGCAGCACCCCAACCCmψGGGCAGCmψCmψmψGmψG ACCGGmψGCCmψGGCmψGGmψACAmψGmψmψmψCAGmψmψCCACCGGAmψCCmψGC AGmψACGCCCGmψGCCGAGACAGGAGmψCCAGCGGCCAmψmψCmψmψmψGGAmψm ψmψmψCmψGGACAACmψmψGmψGmψGACCGAGGAmψGACCAGGAAACmψACCAC mψCGGmψmψCmψGAGACCGAAGCCGmψGACCCmψGCAGGACGmψGAGAGGCCGGG ACmψACCAGAACGCCAmψGCGGGmψGmψGGmψCCAACAmψCCmψGGAACmψGAAAAG CAAGCACGCACCCmψCmψGACCCCGmψAAAGAAGAGGAGmψACmψGCAGGCCCAGGmψ GCGGAGCAAGCAAGCAGCmψGGACGCCCGmψAAGGmψGGAmψCmψGmψGmψGAAAG AmψmψGCCmψCCmψGCCCGmψGAGAGAGmψACmψmψCAAGmψAmψmψmψCAGCCAGA AmψAGmψCmψGCCCGmψGGGAGGCGCGGCGGCGGCAmψGAACAACmψCCAGGGCA GAGmψGACmψmψCGAGGACGmψGACCGmψGAAmψmψmψmψACACAGGGAGAGmψGG CAGAGACmψGAACCCCGAGCAGAGAAACmψGmψACCGGAGmψGmψGAmψGmψGGA AAACmψACAGCAAmψCmψGGmψGmψCCGmψGGGCCAGGGCGAGACCACAAAGCCmψG ACGmψGAmψCCmψGCCmψCmψGGAGCAGGGCAAGGAACCCmψGGCmψGGAGGAGGAG GAGGmψGmψGGGAAGCGGACGGGCCGAGAAGAACGGCGACAmψCGGCGGACAGAmψ CmψGGAAAGCCmψAAGGACGmψGAAAGAAAGCCmψGGGCGGCCCAAGCAGCGGCGCCC CmψCCmψCCAGCGCGGCGAGCCAGCCGGCGmψCCCAACCGmψCmψACCAGGAGGG CACCGmψCmψGAGmψCCGCCACCCCGAGAGCGGCCCGmψGGCACCGmψCCACCGAGCCC AGCGAGGGCAGCGCACCCGGCAGCCCGmψGCCGGCAGCCCCACCGmψCCACAGAGGAGG GAACGACACCGAGCCAGCGAAGGCAGCGCCCCAGGCACCGACCGAGCCCGmψAGm ψGAGCAGGAGAmψmψAAACGGAmψCAACAAGAmψCAGAAGAAGAmψmψGmψGAAAG ACAGCAACACCAAGAGGCCGGCAAGACAGGCCCGAmψGAAAACCGmψGmψGGmψm ψAGAGmψGAmψGACACCCGAmψCmψGAGAGAGCGGCGmψGGAAAACCGmψGAGAAAGAA GCCmψGAAAAmψAmψCCCCAGCCCGAmψCAGCAAmψACAmψCmψAGAGCCAACCGmψG AAmψAAGCmψGmψGACCGAmψmψACACCGAAAmψGAAGAAGGCGGAmψCCmψGCAmψ GmψGmψGGGAAGAGAmψmψCCAGAAGGACCCCGmψGmψGGGCCmψGAmψGAGCCG GmψGGCCAGCCCGmψGCCAGCAAGAGAmψCGAmψCAGAACAAGCmψGAAACCGmψGAG AmψGGACGAGAAGGGCAACCGmψGACCACCGCGGCGmψmψmψGCCmψGmψCmψCAGm	3129

LTRP molecule	RNA sequence	SEQ ID NO
	1 ψGmψGGCCAGCCCCmψGmψmψCGmψGmψACAAGCmψGGAGCAGGmψGmψCmψGAGAA 2 GGGCAAGGmψmψACACCAACmψACmψmψCGGACGGmψGCAAmψGmψGGCCGAGCAC 3 GAAAAGCmψGAmψCmψGmψGGCCAGCmψGAAGCCCCGAGAAGGAmψAGCGACGAA 4 GCCGmψGACAmψAmψAGCCmψGGGAAAAGmψmψmψGGGCAGAGGGCCCCmψGGAmmψmψ 5 ψCmψACAGCAmψmψCAmψGmψGACCAAGGAGmψCCACCCACCCCGmψGAAGCCCCmψ 6 GGCCAGAmψCGCCGGAACAGAmψACGCCmψCCGGACCmψGmψGGGAAAGGCCmψ 7 GAGCGACGCmψGmψAmψGGGCACAAmψCGCCmψCCmψmψCCmψGmψCmψAAGmψAC 8 CAGGACAmψCAmψCAmψCGAACACCAGAAGGmψGGmψGAAGGGCAACCAGAAGAGAC 9 mψGGAGAGCCmψGCGGGAGCmψGGCCGGCAAGGAAAACmψGGAAmψACCCmψAGCG 10 mψGACCCmψGCCACCmψCAGCCmψCACACCAAGGAGGGCGmψmψGAmψGCCmψACAA 11 CGAAGmψGAmψCGCCCGGGmψGCGAAmψGmψGGGmψGAACCmψGAACCmψGmψGGCA 12 GAAGCmψGAAGCmψAAGCAGAGAmψGAmψGCCAAGCCmψCmψGmψGAGACmψGAAG 13 GGAmmψmψCCmψmψCCmψmψmψCCmψCmψGGmψCGAGAGACAGGCCAACGAAGmψGG 14 ACmψGGmψGGGACAmψGGmψGmψGmψAACGmψGAAGAAGCmψGAmψCAACGAGAAAA 15 AGGAGGAmψGGCAAGGmψGmψmψmψmψGGCAGAAmψCmψGGCmψGGCmψACAAGAGA 16 CAGGAAGCCCCmψGAGACCmψACCmψGAGCAGCGAGGAAGAmψCGGAAGAAAGGGAAA 17 GAAAmψmψCGCmψCGGmψACCAGCmψGGCGACCmψGmψGmψGCACCmψGGAAAA 18 GAAGCACGGCGAGGACmψGGGGAAGGmψGmψACGACGAGGCCmψGGGAGCGGAmψm 19 ψGACAAGAAAGmψGGAAGGCCmψGAGCAAGCACAmψCAAGCmψGGAAGAGGAACGGA 20 GAAGCGAGGACGCCAGAGCAAGGCCGCCmψGACCGACmψGGCmψGCGGGCmψAAG 21 GCCAGCmψmψCGmψGAmψCGAGGGCCmψGAAGGAGGCCGACAAGGACGAGmψmψCmψ 22 CGAGAmψGCGAGCmψGAAGCmψGCAGAAAGmψGGmψACGGGGACCmψGCGGGGAAAGC 23 CCmψmψCGCCAmψCGAAGCCGAGAACAGCAmψCCmψGGACAmψCAGCGGmψmψCAG 24 CAAGCAGmψACAACmψGmψGCCmψmψCAmψCmψGGCAGAAGGACGGCGmψGAAGAAG 25 CmψGAACCmψGmψACCmψGAmψCAmψCAACmψACmψmψCAAGGGCGGCAAGCmψGCG 26 GmψmψCAAGAAGAmψCAAACmψGAAGCCmψmψCGAAGCCAACAGAmψmψCmψACAC 27 CGmψGAmψCAACAAAAAGAGCGGCGAGAmψCGmψGCCAmψGGAGGmψGAAACmψmψC 28 AACmψmψCGACGACCCCAACCmψGAmψCAmψCCmψGCCmψCmψGGCCmψmψmψGGCA 29 AGAGACAGGGCAGAGAAAmψmψCAmψCmψGGAACGACCmψGmψGmψCCmψGGAAAC 30 CGGCAGCCmψGAAGCmψGGCCAACGGAAGAmψGAmψCGAGAAGACACmψGmψACAA 31 CAGAAGAACCCGGCAGGAmψGAGCCmψGCCmψGmψmψCGmψGGCCmψGACCmψmψ 32 CGAGCGGCGGGAGGmψCCmψGGACmψCCmψCCAAmψAmψCAAACCAAmψGAACCmψG 33 AmψCGGCGmψGGCAAGAGGGCGAAACAmψCCCCGCCGmψGAmψCGCCCGmψGACCCGAC 34 CAGAGGGCGmψGCCACmψGAGCCGGmψmψmψAAGGAmψAGCCmψGGGAACCCCAAC 35 CCACAmψCCmψGAGAAmψCGGCGAGAGCmψAmψAAGGAGAAGCAGCGGACCmψCCA 36 GGCCAAGAAAGGAGGmψGGAGCAGCGGAGAGCCGGCGGCGmψACAGCCGGAAGmψACGC 37 CAGCAAGGCCAAGAAmψCmψGGCAGACGAmψAmψGGmψGAGAAACACCGCmψAGAGA 38 mψCmψGmψGmψACmψACGCCGmψGACCCAGGAmψGCCAmψGmψGAmψCmψmψCGC 39 CAACCmψGAGCCGGGGCGmψmψCGGCCGGCAGGGCAAGCGGACCmψmψCAmψGGCCGA 40 GAGACAGmψACACACGGAmψGGAGGACmψGGCmψGACCGCCAAGCmψGGCCmψACGA 41 GGGCCmψGAGCAAGACCmψACCmψGmψCCAAGACAmψGGCCAGmψACACCmψCCA 42 AGACAmψGCAGCAACmψGmψGGGmψmψmψACCmψCACCAGCGCCGACmψACGACAG 43 GGmψGmψGGAGAAGCmψGAAGAAGACAGCAACAGGCGmψGGAmψGACCACAAmψmψA 44 ACGGCAAGGAGCmψGAAGGmψGGAGGGCCAGAmψmψACCmψACmψACAACAGAmψAC 45 AAGAGACAGAACGmψAGmψCAAGGACCmψGmψCCGmψCGAGCmψGGAmψAGACmψGA 46 GCGAAGAAmψCmψGmψGAACAACGACAmψCmψCCmψCCmψGGACAAAGGGCAGAAAGC 47 GGAGAAGCmψCmψGAGCCmψCCmψGAAGAAAAGAmψmψCmψCCAmψAGACCCGmψG 48 CAGGAGAAGmψmψCGmψGmψGCCmψGAACmψGCGGCGmψmψCGAGACACACGAGCCG 49 AGCAAGCCGCCmψGAACAmψCGCCAGAmψCCmψGGCmψGmψmψCCmψGCGGAGCCA 50 GGAGmψACAAGAAAmψACCAGACAAACAAGACAACCGGCAACACCGAmψAAGAGAGC 51 CmψmψCGmψCGAGACCmψGGCAGmψCCmψmψmψmψACCGGAAAGAGCmψmψAAGGAG 52 GmψGmψGGAAACCmψGCCGmψGCGGmψCmψGGCGGAmψCmψGGCGGAGGCGmψCCACC 53 AGCCCCAAGAAAAAGAGAAAAGmψCmψAAmψAGAmψAAGCmψGCCmψmψCmψGCGGG 54 GCmψmψGCCmψmψCmψGGCCAmψGCCmψmψCmψmψCmψCCmψmψGCACCmψG 55 mψACmψCmψmψGGmψCmψmψmψGAAmψAAAGCCmψGAGmψAGGAAAGmψCmψAGAAA 56 AA 57 AAAAAAAAAAAAAAAAAAAAAA	

LTRP molecule	RNA sequence	SEQ ID NO
LTRP5-ADD-ZIM3	AAAmlyAAGAGAGAAAAGAAGAmlyAAGAAGAAAmlyAmlyAAGAGCCACCAmlyGGCCCC mlyAAGAAGAAAGCmlyAAAGmlyGAGCCGGAmlyGGAACGCCmlyCGmlyCmlyACGAGGmlyG CGGCAGAAAGmlyGCAGAAACAmlyCGAGGACAmlyCmlyGCmlyCmlyCmlyCGCGAmlyCmly CmlyGAACmlyGACCmlyGGAGCACCCACmlyGmlyCmlyCmlyCGCGGCAmlyGmlyGCCAG AACmlyGmlyAAAAACmlyGmlymlymlymlyCmlyGGAGmlyGmlyGCCmlyAmlyCAAmlyACGACGA mlyGACGGCmlyACCAGAGCmlyACmlyGCACCmlyCmlyGmlymlyCGCGCGGAAGAGAGGmly GCmlyGAmlyGmlyGmlyGGAAAmlyACAACmlyGCmlyGCCGGmlyGCmlymlyCmlyGCmlyGGA AmlyGCmlyGGACmlyGCmlyGmlyGGGCCCGCGCGCCAGGCCGmlyAmlymlyAAG GAAGAmlyCCmlymlyGGAACmlyGCmlyACAmlyGmlyCGCGCCACAAGGGCACAmlyACGGCC mlyGCmlyGAGACGGAGAGAGGACmlyGGCCmlyAGCAGACmlyGCAGAmlyGmlymlyCmlymlyC GCCAAmlyAACACGACCAAGAGAmlymlyCGACCCCCmlyAAGmlyGmlyACCCmlyCCCmly lyCCCCCGGAGAAGAGAAAGCCAmlyCCGGmlyCCmlyGAGCCmlyGmlymlyCGAmlyGGC AmlyCGCCACCGGmlyCmlyGCmlyGmlyGCmlyGAAGGACmlyGGGCAmlyCCAGGmlyGGA mlyAGGmlyACAmlymlyGCCmlyCCGAGGmlyGmlyGCGAGGACmlyCCAmlyCACCmlyGGGA mlyGGmlyCGGmlyCmlyCAGGGCAAGAmlyCmlyGmlyACGmlyGGCGACGmlyCGCGAGCGmly lyGACACAGAAGCAmlyAmlyCCAGGAGmlyGGGGCCmlymlymlyCGACCCmlyGmlyGAmlyCG CGGCAGCCmlymlyGCAAmlyGACmlyGAGCmlyCGmlyGAACCCAGCCCGGAAGGGCC mlyGmlyACGAGGGAACCGGCAGACmlyGmlymlyCmlymlyCGAGmlymlymlyACAGACmlyGC mlyGCACGACGCCCCGGCCmlyAAGGAAGGCGACGACCGGCCmlymlyCmlymlymlyGGCCmly lyGmlymlyCGAGAAmlyGmlyGmlyGGCCAmlyGGGAGmlyCAGCGACAAGCGGGAmlyAmlymly AGCCGGmlymlyCCmlyGGAGAGCAACCCCGmlyGAmlyGAmlyCGAmlyGCCAAGGAAGmlyGA CGCCCGCCACCGGGCCAGAmlyACmlymlyCmlyGGGGCAAmlyCmlyGCCmlyGGCmlyGAA CAGACCCCGmlyGGCCAGCACCGmlyGAACGACAAGCmlyGGAGCmlyGCAGGAGmlyGCCmly GGAGCACGGCCGGAmlyCGCCAAGmlymlyCAGCAAGGmlyGAGAACCmlyCACCACCCGA AGCAACAGCAmlyCAAACAAGGCAAGGACCAGCACmlymlymlyCCmlyGmlyGmlymlyCmlyG AACGAGAAGGAGGACAmlyCCmlyGmlyGmlyGmlyACCGAGAmlyGGAGAGAGmlyGmlymlyC GGGmlymlyCCCAGmlyCCAmlyACACAGAmlyGmlyCAGCAACAmlyGmlyCmlyAGACmlyGGC CAGACAGAGACmlyGCmlyGGGAAGAAGCmlyGmlyCCGmlyCCmlyGmlyGAmlyCAGACAC CmlyGmlymlyCGCCCCmlyCmlyGAAGGAGmlyACmlymlyCGCCmlyGCmlyGAGCAGCGGCAA CAGCAACGCCAACAGCCGGGGCCCCAGCmlymlyCmlymlyAGCGGCCmlyGGmlyGCCACmly lyGmlyCCmlyGAGAGGGAGCCACAmlyGGGCCCGmlyGGAGAmlyCmlyACAAACCGmlyG AGCGCCmlyGGAAGCGGCAGCCmlyGmlyGCGCGmlyGCmlyGAGCCmlyGmlymlyCGGA mlyAmlyCGAmlyAAAGmlyCCmlyGAAAGCCmlyGGGAmlymlyCCmlyGGAGAGCGGCGmlyCmlyG CCmlyCGGCGGmlyGCCACCCmlyGAAGmlyACGmlyGGAGGAmlyGmlyGACAAACGmlyGG lyCAGACGGGAmlyGmlyGGAGAAGmlyGGGGCCCCmlymlyCGAmlyCmlyGGmlyGmlyACGGCA GCACCAACCCCGmlyGGGCAGCmlyCmlymlyGmlyGACCGGmlyGCCmlyGGCmlyGGmlyACA mlyGmlymlyCAGmlymlyCCACCGAmlyCmlyGCAGmlyACGCCmlyGCCGAGACAGGAG mlyCCCAGCGGCCAmlymlyCmlymlymlyGGAmlymlymlyCmlyGGACAACmlymlyGCmlyG CmlyGACCGAGGAmlyGACCAGGAAACmlyACCACmlyCGGmlymlyCCmlyGCAGACCGAAGC CGmlyGACCCmlyGCAGGACmlyGAGAGCCGGGACmlyACCAGAACGCCAmlyGCGGmly GmlyGGmlyCCAACmlyCCmlyGGACmlyGAAAAGCAAGCACGCACCmlyCmlyGACCCmly AAAGAAGAGGAGmlyACCmlyGCAGGCCAGGmlyCGGGAGCAGAAGCAAGCmlyGGACGC CCCGmlyAAGGmlyGGAmlyCmlyGCmlyGGmlyGAAGAAmlymlyGCCmlyCCmlyGCCCCmlyGAGA GAGmlyACmlymlyCAAGmlyAmlymlyCAGCCAGAAmlyAGmlyCmlyGCCCCmlyGGGAGGCA CGCGCGCGGCAmlyGAACAACmlyCCCAGGGCAGAGmlyGACCmlymlyCGAGGACGmlyGA CCGmlyGAAmlymlymlyACACAGGGAGAGmlyGGCAGAGACmlyGAACCCCGAGCAGAGA AACmlyGmlyACCGGGAmlyGmlyGAmlyGCmlyGGAACAACmlyACAGCAAmlyCmlyGGmlyGmly CCGmlyGGGCCAGGGCAGACCACAAAGCCmlyGACGmlyGAmlyCCmlyGCCmlyCmlyGGAG CAGGGCAAGGAACCCmlyGGCmlyGGAGGAGGAGGAGGmlyGmlyGGGAAGCGGACGGGC CGAGAAGAACGGCGACAmlyCGGCGGACAGAmlyCmlyGGAAGCCmlyAAGGACGmlyGAAA GAAAGCCmlyGGCGGGCCAAAGCAGCGCGGCCCGmlyCCmlyCCCAGCGGCGGACGCCCA GCCGGCmlyCCCCAACmlyCmlyACCGAGGAGGGCACmlyCmlyGAGmlyCCGCCACCCCG GAGAGCGGCCCGmlyGGCACCmlyCCACCGAGCCAGCGAGGGCAGCGCACCCGGCAGCC CmlyGCCGGCAGCCCCACCmlyCCACAGAGGAGGGAACAGCACCGAGCCAGCGAAGG CAGCGCCCCAGGCACAGCACCGAGCCmlyAGmlyGAGCAGGAGAmlymlyAAACGGAmlyC AACAGAmlyCAGAAGAAGACmlymlyGmlyGAAAGACAGCAACCAAGAAGGCCGGCAA GACAGGCCCAAmlyGAAAACCCmlyGCmlyGmlymlyAGAGmlyGAmlyGACACCCGAmlyCmly GAGAGAGCGGCGmlyGGAAAACmlyGAGAAAAGAGCCmlyGAAAAmlyAmlyCCCCCAGGCC	3130

LTRP molecule	RNA sequence	SEQ ID NO
	AmlyCAGCAAmlyACAmlyCmlyAGAGCCAACmlyGAAmlyAAGCmlyGCmlyGACCGAmlymlyA CACCGAAAmlyGAAGAAGGCGAmlyCCmlyGCmlyGmlyGmlyACmlyGGGAAGAmlymlyCCA GAAGGACCCmlyGmlyGGGCCmlyGAmlyGAGCCGGGmlyGGCCAGCCmlyGCCAGCAAGAA GAmlyCGAmlyCAGAACAAGCmlyGAAACmlyGAGAmlyGGACGAGAAGGGCAACmlyGAC CACCGCCGGCmlymlymlyGCCmlyGCmlyCmlyCAGmlyGmlyGGCCAGCCCCmlyGmlymlyCmly GmlyACAAGCmlyGGAGCAGGmlyGmlyCmlyGAGAAGGGCAAGGmlymlyACACCAACmlyAC mlymlyCGGACGGmlyGCAAmlyGmlyGGCCGAGCACGAAAAGCmlyGAmlyCCmlyGCmlyGGCC CAGCmlyGAAGCCCAGAAAGGAmlyAGCGACGAAGCCGmlyGACAmlyAmlyAGCCmlyGGGA AAGmlymlymlyGGGCAGAGGGCCcmlyGGAmlymlymlyCmlyACAGCAmlymlyCmlyGmlyGACC AAGGAGmlyCCACCACCCCGmlyGAAGCCCCmlyGGCCAGAmlyCGCCGGAACAGAmly ACGCCmlyCCGGACCmlyGmlyGGGAAGGGCCcmlyGAGCGACGCAmlyGmlyAmlyGGGCACA AmlyCGCCmlyCCmlymlyCCmlyGmlyCmlyAAGmlyACCAGGACAmlyCmlyCmlyCGAACACC AGAAGGmlyGGmlyGAAGGGCAACCAGAAGAGACmlyGGAGAGCCmlyGCGGGAGCmlyGGC CGGCAGGAAAACmlyGGAAmlyACCCmlyAGCGmlyGACCCmlyGCCACmlyCAGCCmlyC ACACCAAGGAGGGCCmlyGAmlyGCCmlyACAACGAAGmlyGAmlyCGCCCGGGmlyGCGA AmlyGmlyGGGmlyGAACmlyGAACmlyGmlyGGCAGAAGCmlyGAAGCmlyAAGCAGAGAmly GAmlyGCCAAGCCmlyCmlyGmlyGAGACmlyGAAGGGAmlymlyCCmlymlyCCmlymlymlyCmly CmlyGGmlyCGAGAGACAGGCCAACGAAGmlyGGACmlyGGmlyGGGACAmlyGGmlyGmlyGmly lyAACGmlyGAAGAAGCmlyGAmlyCAACGAGAAAAAGGAGGAmlyGGCAAGGmlyGmlymlymly mlyGGCAGAAmlyCmlyGGCmlyGGCmlyACAAGAGACAGGAAGCCcmlyGAGACCAmlyACCmly lyGAGCAGCGAGGAAGAmlyCGGAAGAAGGGAAGAAAmlymlyCGCmlyCGGmlyACCAGCmly lyGGGCGACCmlyGCmlyGCmlyGCACmlyGGAAAAGAAGCACGGCGAGGACmlyGGGGAAA GGmlyGmlyACGACGAGGCCcmlyGGGAGCGGAmlymlyGACAAGAAAGmlyGGAAGGCCmlyGA GCAAGCACAmlyCAAGCmlyGGAAGAGGAACGGAGAAGCGAGGACGCCCAGAGCAAGGC CGCCcmlyGACCGACmlyGGCmlyGCGGGCmlyAAGGCCAGCmlymlyCGmlyGAmlyCGAGGGC CmlyGAAGGAGGCCGACAAGGACGAGmlymlyCmlyGCAGAmlyGCAGCmlyGAAGCmlyGCA GAAGmlyGGmlyACGGGGACcmlyGCGGGGAAAGCCcmlymlyCGCCAmlyCGAAGCCGAGAA CAGCAmlyCCmlyGGACAmlyCAGCGCCmlymlyCAGCAAGCAGmlyACAACmlyGmlyGCmlymly lyCmlyCmlyGGCAGAAGGACGGCmlyGAAGAAGCmlyGAACcmlyGmlyACCmlyGAmlyCmly lyCAACmlyACmlymlyCAAGGGCGGCAAGCmlyGCGGmlymlyCAAGAAGAmlyCAAACcmlyGA AGCCmlymlyCGAAGCCAACAGAmlymlyCmlyACACCGmlyGAmlyCAACAAAAAGAGCGGCG AGAmlyCGmlyGCCAmlyGGAGGmlyGAAcmlymlyCAACmlymlyCGACGACCCCAACcmlyGA mlyCmlyCCmlyGCCmlyCmlyGGCCmlymlymlyGGCAAGAGACAGGGCAGAGAAmlymlyCmly CmlyGGAACGACcmlyGGCCAGmlyACACcmlyCCAAGACcmlyGCAGCAACmlyGmlyGGGmly mlymlyACCAmlyCACCAGCGCCGACmlyACGACAGGGmlyGCmlyGGAGAAGCmlyGAAGAAG ACAGCAACAGGCCmlyGGAmlyGACCACAAmlymlyAACGGCAAGGAGCmlyGAAGGmlyGGAG GGCCAGAmlymlyACCmlyACmlyACAACAGAmlyACAAGAGACAGAACGmlyAGmlyCAAGGA CCmlyGmlyCCGmlyCGAGCmlyGGAmlyAGACmlyGAGCGAAGAAmlyCmlyGmlyGAACAACGA CAmlyCmlyCCmlyCCmlyGGACAAAGGGCAGAAGCGGAGAAGCmlyCmlyGAGCCmlyCCmlyG AAGAAAAGAmlymlyCmlyCCAmlyAGACCCGmlyGCAGGAGAAGmlymlyCmlyGmlyGCCmly GAACmlyGCGGCmlymlyCGAGACACACGACGCCGAGCAAGCCGCCcmlyGAACAmlyCGCC AGAmlyCCmlyGGCmlyGmlymlyCCmlyGCGGAGCCAGGAGmlyACAAGAAAmlyACCAGACAA ACAAGACAACCGGCAACACCGAmlyAAGAGAGCCmlymlyCGmlyCGAGACCmlyGGCAGmly CCmlymlymlymlyACCAGGAAGAAGCmlymlyAAGGAGGmlyGmlyGGAAACcmlyGCCGmlyGCGG mlyCmlyGGCGGAmlyCmlyGGCGGAGGcmlyCCACCAGCCCAAGAAAAAGAAAAAGmlyC mlyAAmlyAGAmlyAAGCmlyGCCmlymlyCmlyGCGGGCmlymlyGCCmlymlyCmlyGGCCAmlyGC	

LTRP molecule	RNA sequence	SEQ ID NO
	CCmψmψCmψmψCmψCmψCCCmψmψGCACCmψGmψACCmψCmψmψGGmψCmψmψmψGA AmψAAAGCCmψGAGmψAGGAAGmψCmψAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAA	

Table 17: Top 16 spacers activity by ELISA, RT-qPCR, and mean percent CpG methylation (mCpG)

Spacer ID	ELISA % reduction	RT-qPCR % reduction	mean % mCpG
TG-06-001	-54.60	-53.66	59.00
TG-06-133	-26.60	-58.51	42.90
TG-06-139	-51.50	-63.03	55.90
TG-06-141	-53.70	-58.85	53.70
TG-06-142	-18.50	-31.76	28.40
TG-06-143	-27.20	-51.60	46.10
TG-06-144	-28.10	-41.68	54.70
TG-06-146	-22.10	-45.31	31.00
TG-06-149	-57.50	-42.07	56.90
TG-06-150	-56.30	-55.68	56.40
TG-06-154	-66.50	-55.28	45.50
TG-06-157	-55.60	-52.86	54.40
TG-06-158	-56.50	-57.04	49.40
TG-06-160	-42.00	-54.91	57.30
TG-06-167	-37.80	-50.29	46.00
TG-06-168	-39.90	-50.34	48.80
naive	N/A	N/A	6.70
NT	N/A	N/A	1.50

Table 18: Top 16 spacers Durability and Specificity by bisulfate sequencing and RNA-seq

Spacer ID	Δ%PCSK9 week 9 vs 1	# off-target differentially expressed genes
TG-06-154	-22.30	0
TG-06-157	-28.30	1
TG-06-001	-63.50	2
TG-06-141	-10.80	1
TG-06-149	-26.70	0
TG-06-158	0.40	376
TG-06-150	-35.50	2
TG-06-160	-23.50	18
TG-06-139	-34.90	24

TG-06-144	-21.30	0
TG-06-133	-29.40	6

Example 2: Demonstration that mRNA encoding for LTRPs containing the ADD domain can induce repression of an endogenous locus in multiple human cell lines

[0401] Experiments were performed to demonstrate that mRNA encoding for LTRPs containing the ADD domain can induce long-term repression of an endogenous target locus in various human cell lines, when delivered as mRNA co-transfected with a targeting gRNA. Materials and Methods

Generation of mRNA:

[0402] mRNA encoding the following molecules were generated by IVT following similar methods as described in Example 1: 1) a catalytically-active CasX 676, 2) dXR1, and 3) LTRP5-ADD-ZIM3. Sequences encoding these molecules were codon-optimized using a codon utilization table, in addition to using a publicly available codon optimization tool and adjusting parameters such as GC content. The DNA and mRNA sequences encoding the catalytically-active CasX 676 are shown in Table 19 and Table 20 respectively. The DNA and mRNA sequences encoding for dXR1 are shown in Table 12 and Table 13 respectively. The DNA and mRNA sequences encoding for LTRP5-ZIM3-ADD are shown in Table 15 and Table 16 respectively.

Table 19: Encoding sequences of the catalytically-active CasX 676 mRNA molecule assessed in this example*

CasX mRNA ID	Component (ID)	Description	DNA sequence	SEQ ID NO:
CasX 676 mRNA	5'UTR	TriLink	AAATAAGAGAGAAAAGAAGAGTAAGAAGAAATATAAGAGCCACC	3047
	START codon + c-MYC NLS		ATGGCCCCTGCTGCCAAGAGAGTGAAGCTGGATAGCAGA	3133

CasX mRNA ID	Component (ID)	Description	DNA sequence	SEQ ID NO:
	CasX 676		CAGGAGATCAAGCGGATTAATAAAAATTCGGAGAAG ACTGGTGAAGGATTCTAACACAAAGAAGGCTGGCA AGACACGGGGCCCTATGAAGACACTGCTGGTGAGA GTGATGACACCCGACCTGAGAGAAAAGACTGGAAAA CCTGAGAAAAGAAGCCTGAGAATATCCCCAGCCCA TCAGCAACACAAGCCGGGCCAACCTGAATAAGCTG CTGACCGACTACACCGAAATGAAGAAGGCCATCCT GCACGTGTATTGGGAAGAGTTCAGAAAAGACCAG TCGGCCTGATGAGCAGAGTGGCTCAGCCTGCCAGC AAGAAGATCGATCAGAACAAGCTGAAGCCCCGAAAT GGACGAGAAGGGGAACCTGACAACCGCCGGCTTTG CCTGTAGCCAGTGCGGCCAGCCCTGTTTGTGTAC AACTGGAACAGGTGAGCGAAAAGGGCAAGGCTTA CACGAATTACTTCGGCAGATGCAACGTGGCCGAGC ACGAGAAGCTGATCAAGCTGGCCAGCTGAAGCCT GAGAAGGATAGCGATGAGGCAGTGACATATTCCCT GGGCAAGTTCGGACAGCGGGCCCTGGATTTTATT CCATTTCATGTGACCAAGGAATCCACCCACCCCGTC AAGCCTCTTGCCCAAATTGCCGGCAACAGATACGC CTCCAGCCCCGTGGGCAAGGCCCTGAGCGACGCCCT GTATGGGCACCATCGCCAGCTTCCTGTCTAAGTAC CAGGACATTATCATCGAGCACCAGAAGGTGGTGAA GGGCAACCAGAAGAGACTGGAGAGCCTGCGCGAGC TGGCCGGCAAGGAAAACCTGGAGTATCCTAGCGTG ACCCTGCCTCCTCAGCCTCATACAAAGGAGGGCGT GGATGCCTACAACGAAGTGATCGCCCGGGTGCGGA TGTGGGTGAACCTGAATCTGTGGCAGAAGCTGAAG CTGTCTAGAGACGACGCCAAGCCCTGCTGAGACT GAAGGCTTCCCCAGCTTCCCTCTGGTGGAGAGAC AGGCAAATGAAGTGGACTGGTGGGACATGGTGTGT AACGTGAAGAAGCTGATCAATGAGAAGAAGGAGGA CGGCAAAGTGTTCTGGCAGAATCTGGCCGGCTACA AGCGTCAGGAGGCCCTGCGGCCCTACCTGAGCAGC GAGGAAGACAGAAAAGAAGGGCAAGAAGTTCGCCCG GTATCAGCTGGGGGACCTGCTGCTGCACCTCGAGA AGAAGCACGGCGAAGACTGGGGGAAGGTGTACGAT GAGGCCTGGGAGCGGATCGATAAGAAGGTGGAGGG CCTGAGCAAGCACATCAAGCTGGAGGAGGAACGGA GATCTGAGGACGCCAGAGCAAGGCCGCCCTGACC GACTGGCTGAGAGCCAAGGCCAGCTTCGTCAATCGA GGGGCTGAAGGAGGCCGACAAGGACGAGTCTGCC GGTGCGAACTGAAGCTGCAGAAGTGGTACGGAGAT CTGAGAGGCAAACCTTTCGCCATCGAGGCCGAGAA CAGCATCCTGGACATCAGCGGCTTCAGCAAGCAGT ACAACCTGCGCCTTTATTTGGCAGAAGGACGGAGTG AAGAAGCTGAACCTGTACCTGATCATCAACTATTT CAAGGCGGGCAAGCTGAGATTCAAGAAGATCAAGC CTGAAGCCTTCGAGGCCAACAGATTCTACACCGTG ATTAACAAGAAAAGCGGAGAGATCGTGCCAATGGA AGTGAACCTCAACTTCGACGACCCTAACCTGATCA TCCTGCCCCCTGGCATTTGGCAAGCGGCAGGGCAGA GAGTTCATCTGGAACGACCTGCTGTCTCTGGAGAC CGGCAGCCTGAAGCTGGCCAACGGCAGAGTGATCG AGAAGACACTGTACAACAGACGAACCAGACAAGAC GAGCCCGCCCTGTTGTGGCCCTGACCTTCGAGAG AAGAGAGGTGCTGGACAGCAGCAATATCAAGCCTA TGAACCTGATCGGCGTGGACCGGGCGAGAACATC CCTGCCGTGATCGCCCTTACCGACCCCGAGGGATG	3134

CasX mRNA ID	Component (ID)	Description	DNA sequence	SEQ ID NO:
			CCCTCTGAGCCGGTTTAAAGACAGCCTGGGCAACC CTACCCACATCCTGAGAATTGGCGAGTCTTACAAG GAGAAGCAGAGAACCATCCAGGCCAAGAAGGAGGT GGAGCAGCGGCGGGCTGGCGGCTACTCCCGAAGT ACGCCAGCAAGGCCAAGAACCTGGCCGACGACATG GTTAGAAATACCGCCAGAGACCTCTGTACTACGC TGTGACCCAGGACGCCATGCTGATCTTCGAGAACC TGAGCAGAGGCTTCGGCAGACAGGGCAAGAGAACC TTCATGGCCGAGAGACAGTACACCCGGATGGAGGA CTGGCTGACCGCCAAGCTGGCCTACGAGGGCCTGC CCTCTAAGACCTACCTGTCCAAGACCTTGGCACAG TACACCAGCAAGACATGCTCTAACTGCGGCTTCAC AATCACGAGCGCCGACTACGACCGGGTCTGGAGA AACTGAAGAAGACCGCCACAGGCTGGATGACCACC ATTAACGGCAAGGAGCTGAAGGTGGAGGGCCAGAT CACCTACTACAACAGGTACAAACGGCAGAACGTGG TGAAGGACCTGAGCGTGGAAGTGGATAGACTGAGC GAGGAAAGCGTAAACAATGACATCAGCAGCTGGAC CAAGGGCCGGAGCGGCGAGGCCCTGAGCCTGCTGA AGAAGAGATTCTCCACAGACCAGTGCAGGAGAAG TTCGTGTGTCTGAACTGCGGCTTCGAGACCCACGC CGACGAGCAAGCCGCCCTGAACATCGCCCGGTCTT GGCTTTTCTGCGGAGCCAGGAGTACAAGAAGTAC CAGACAAACAAGACCACAGGCAACACAGACAAGAG AGCCTTCGTGAGACCTGGCAGAGCTTCTACAGAA AGAAGCTGAAGGAGGTGTGGAAGCCTGCCGTG	
	c-MYC NLS + STOP codons		GGAAGCCCCGCTGCCAAGAGAGTGAAGCTGGACTA ATAGATAA	3135
	3'UTR	Mouse HBA	GCTGCCTTCTGCGGGGCTTGCCCTTCTGGCCATGCC CTTCTTCTCTCCCTTGACCTGTACCTCTTGGTCT TTGAATAAAGCCTGAGTAGGAAG	3055
	XbaI restriction site (partial)		TCTAG	3056
	Poly(A) tail		AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAA	3057

*Components are listed in a 5' to 3' order within the constructs

Table 20: Full-length RNA sequences of catalytically-active CasX 676 mRNA molecule assessed in this example. Modification ‘mψ’ = N1-methyl-pseudouridine

CasX mRNA	RNA sequence	SEQ ID NO:
CasX 676 mRNA	AAAmψAAGAGAGAAAAGAAGAGmψAAGAAGAAAmψAmψAAGAGCCACCAmψGGCCCCmψG CmψGCCAAGAGAGmψGAAGCmψGGAmψAGCAGACAGGAGAmψCAAGCGGAmψmψAAmψAA AAmψmψCGGAGAAAGACmψGGmψGAAGGAmψmψCmψAACACAAAGAAGGCmψGGCAAGACA CGGGGCCCmψAmψGAAGACACmψGCmψGGmψGAGAGmψGAmψGACACCCGACCmψGAGAG AAAGACmψGGAAGAACmψGAGAAAGAAGCCmψGAGAAmψAmψCCCCAGCCCAmψCAGCA ACACAAGCCGGGCCAACmψGAAGmψAAGCmψGCmψGACCGACmψACACCGAAAmψGAAGA AGGCCAmψCCmψGCACGmψGmψAmψmψGGGAAGAGmψmψCCAGAAAGACCCAGmψCGGCC	3136

CasX mRNA	RNA sequence	SEQ ID NO:
	mψGAmψGAGCAGAGmψGGCmψCAGCCmψGCCAGCAAGAAGAmψCGAmψCAGAACAAGCmψ GAAGCCCGAAAmψGGACGAGAAGGGGAACmψGACAACCGCCGGCmψmψmψGCCmψGmψA GCCAGmψGCGGCCAGCCCGmψGmψmψmψGmψGmψACAAACmψGGAACAGGmψGAGCGAAA AGGGCAAGGmψmψACACGAAmψmψACmψmψCGGCAGAmψGCAACGmψGGCCGAGCAGCA GAAGCmψGAmψCAAGCmψGGCCAGCmψGAAGCCmψGAGAAGGAmψAGCGAmψGAGGCAG mψGACAmψAmψmψCCmψGGGCAAGmψmψCGGACAGCGGGCCmψGGAmmψmψmψmψAm ψmψCCAmψmψCAmψGmψGACCAAGGAmmψCCACCCACCCCGmψCAAGCCmψCmψmψGCCC AAAmψmψGCCGGCAACAGAmψACGCCmψCCAGCCCCGmψGGGCAAGGCCmψGAGCGACG CCmψGmψAmψGGGCACCAmψCGCCAGCmψmψCCmψGmψCmψAAGmψACCAGGACAmψmψA mψCAmψCGAGCACCAAGGmψGGmψGAAGGGCAACCAGAAGAGACmψGGAGAGCCmψGC GCCAGCmψGGCCGGCAAGGAAAACmψGGAGmψAmψCCmψAGCGmψGACCCmψGCCmψCC mψCAGCCmψCAmψACAAAGGAGGGCGmψGGAmψGCCmψACAACGAAGmψGAmψCGCCCG GmψGCGGAmψGmψGGGmψGAACmψGAmmψCmψGmψGGCAGAAGCmψGAAGCmψGmψCmψ AGAGACGACGCCAAGCCCCmψGmψGAGACmψGAAGGGCmψmψCCCCAGCmψmψCCmψC mψGGmψGGAGAGACAGGCAAAmψGAAmψGGACmψGmψGGGACAmψGGmψGmψAAC GmψGAAGAAGCmψGAmψCAAmψGAGAAGAAGGAGGACGGCAAAmψGmψmψCmψGGCAGA AmψCmψGGCCGGCmψACAAGCCmψCAGGAGGCCmψGCGGCCmψACCmψGAGCAGCGAG AAGACAGAAAGAAGGGCAAGAGmψmψCGCCCGmψAmψCAGCmψGGGGGACCGmψGcmψ GCMψGCACCGmψCGAGAAGAAGCAGCGCGAAGACmψGGGGGAAGGmψGmψACGAmψGAGGC CmψGGGAGCGGAmψCGAmψAAGAAGGmψGGAGGGCCmψGAGCAAGCACAmψCAAGCmψGG AGGAGGAACGGAGAmψCmψGAGGACGCCAGAGCAAGGCCCGCCmψGACCGACmψGGCmψ GAGAGCCAAGGCCAGCmψmψCGmψCAmψCGAGGGGCGmψGAAGGAGGCCGACAAGGACGAG mψmψCmψGCCGGmψGCGAACmψGAAGCmψGCAGAAGmψGGmψACGGAGAmψCmψGAGAGG CAAACCGmψmψCGCCAmψCGAGGCCGAGAACAGCAmψCCmψGGACAmψCAGCGCGCmψmψ CAGCAAGCAGmψACAACmψGCGCCmψmψmψAmψmψmψGGCAGAAGGACGGAGmψGAAGAA GCmψGAACCGmψGmψACCmψGAmψCAmψCAACmψAmψmψmψCAAGGGCGGCAAGCmψGAGA mψmψCAAGAAGAmψCAAGCCmψGAAGCCmψmψCGAGGCCAACAGAmψmψCmψACACCGmψ GAmψmψAACAAGAAAAGCGGAGAGAmψCGmψGCCAAmψGGAAmψGAACmψmψCAACmψm ψCGACAGCCCGmψAACCGmψGAmψCAmψCCmψGCCCGmψGGCAGmψmψmψGGCAAGCGGAGG GCAGAGAGmψmψCAmψCmψGGAACGACCGmψGCMψGmψCmψCmψGGAGACCGGACCGCmψG AAGCmψGGCCAACGGCAGAGmψGAmψCGAGAAGACACmψGmψACAACAGACGAACCAGAC AAGACGAGCCCGCCCGmψGmψmψmψGmψGGCCCGmψGACmψmψCGAGAGAAGAGAGmψGC mψGGACAGCAGCAAmψAmψCAAGCCmψAmψGAACCGmψGAmψCGGCmψGGACCGGGGGCGA GAACAmψCCCGmψGCCCGmψGAmψCGCCCGmψmψACCGACCCCGAGGGAmψGCCCGmψCmψGAG CGCGAGGCCCGmψGAGCCmψGCGCAACCGmψACCCACAmψCCmψGAGAAmψmψGGCGA GmψCCmψACAAGGAGAAGCAGAGAACCmψCCAGGCCAAGAAGGAGGmψGGAGCAGCGGC GGCGmψGGCGGCGmψACmψCCCGGAAGmψACGCCAGCAAGGCCAAGAACCmψGGCCGACGA CAmψGGmψmψAGAAAAmψACCGCCAGAGACCGmψCCmψGmψACmψACCGmψGmψGACCAGG ACGCCAmψGCCmψGAmψCmψmψCGAGAACCmψGAGCAGAGGCGmψmψCGGCAGACAGGGCAA GAGAACCmψmψCAmψGGCCGAGAGACAGmψACACCCGGAmψGGAGGACmψGGCmψGACCG CCAAGCmψGGCCmψACGAGGGCCmψGCCCGmψCmψAAGACCmψACCmψGmψCCAAGACCmψ mψGGCACAGmψACACCAGCAAGACAmψGCMψCmψAACmψGCGGCmψmψCACAAmψCACGA GCCCGGACmψACGACCGGGmψGCMψGGAGAAACmψGAAGAAGACCGCCACAGGCGmψGGAm ψGACCACCGmψAACGGCAAGGAGCmψGAAGGmψGGAGGGCCAGAmψCACmψACmψAC AACAGGmψACAACGGCAGAACCGmψGGmψGAAGGACCGmψGAGCGmψGGAACmψGGAmψAG ACmψGAGCGAGGAAAGCGmψAAACAAmψGACAmψCAGCAGCmψGGACCAAGGGCCGGAGC GGCGAGGCCCGmψGAGCCmψGCMψGAAGAAGAGAmψmψCmψCCACAGACCAAGmψCGAGGA GAAGmψmψCGmψGmψCmψGAAmψGCGGCGmψmψCGAGACCCACGCCGACGAGCAAGC CGCCCGmψGAACAmψCGCCCGGmψCmψmψGGCmψmψmψmψCCmψGCGGAGCCAGGAGmψAC AAGAAGmψACCAGACAAACAGACCACAGGCAACACAGACAAGAGAGCCmψmψCGmψCGA GACCGmψGGCAGAGCmψmψCmψACAGAAAGAAGCmψGAAGGAGGmψGmψGGAAGCCmψGCC GmψGGGAAGCCCCCGmψGCCAAGAGAGmψGAAGCmψGGACmψAAmψAGAmψAAGCmψGCC mψmψCmψGCGGGCGmψmψGCCmψmψCmψGGCCAmψGCCCGmψmψCmψmψCmψCCCGmψm ψGCACCGmψGmψACCmψCmψmψGGmψCmψmψmψGAAmψAAAGCCmψGAGmψAGGAAGmψCm ψAGAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA	

Synthesis of gRNAs:

[0403] gRNAs targeting the human *PCSK9* locus were designed using gRNA scaffold 316 and chemically synthesized. Furthermore, a *B2M*-targeting gRNA was used as an experimental control. Sequences of targeting spacers as assessed in this example as listed in Table 21.

Table 21: Sequences of spacers assessed in this example

Spacer ID	Target	Targeting spacer sequence (RNA)	SEQ ID NO
7.37	human <i>B2M</i>	GGCCGAGAUGUCUCGCUCCG	3137
6.1	human <i>PCSK9</i>	GAGGAGGACGGCCUGGCCGA	1834
6.125	human <i>PCSK9</i>	AGAAAGAGCAAGCCUCAUGU	1845
6.138	human <i>PCSK9</i>	AGGGAUUUAUACUACAAAGA	1852
6.153	human <i>PCSK9</i>	GUUAAUGUUUAAUCAGAUAG	1866
6.157	human <i>PCSK9</i>	GGGUCUGAGCCUGGAGGAGU	1869
6.172	human <i>PCSK9</i>	GCUGAAACAGAUGGAAUACU	1879
6.181	human <i>PCSK9</i>	UCAUCUGCACUCGUGGCCAC	2228

Transfection of mRNA and gRNA into HepG2 cells, Hep3B cells, and Huh7 cells and ELISA to assess PCSK9 secretion:

[0404] The following three human hepatocyte cancer cell lines were used in this experiment: HepG2 cells, Hep3B cells, and Huh7 cells. ~15,000 cells of each cell line were seeded per well; the next day, seeded cells were transfected with mRNA encoding a catalytically-active CasX 676, dXR1, or LTRP5-ADD-ZIM3 and a gRNA with scaffold 316 and spacer targeting either the *B2M* or *PCSK9* locus (see Table 21 for specific spacers and sequences). Media supernatant was harvested at 4 days post-transfection to assess level of PCSK9 secretion by ELISA, and levels of PCSK9 secretion were normalized to total cell count and illustrated in FIGS. 3-5. Culturing of treated Huh7 cells continued, and media supernatant was harvested at 14 and 27 days post-transfection for measuring PCSK9 secretion by ELISA. As an additional experimental control, PCSK9 secretion was also measured in the media supernatant harvested from wells containing untreated, naïve cells.

Results:

[0405] HepG2 cells, Hep3B cells, and Huh7 cells were transfected with mRNA encoding catalytically-active CasX 676, dXR1, or LTRP5-ADD-ZIM3 with a gRNA targeting either the *B2M* or the *PCSK9* locus, and secreted PCSK9 levels were measured. Quantification of normalized PCSK9 secretion levels for each condition at 4 days post-transfection is depicted in FIGS. 3-5. The data demonstrate that the most efficient knockdown of PCSK9 secretion by

CasX 676, dXR1, or LTRP5-ADD-ZIM3 was observed in Huh7 cells, while HepG2 cells did not exhibit as efficient of a knockdown of secreted PCSK9 levels (FIGS. 3-5). Meanwhile, Hep3B cells demonstrated low PCSK9 secretion levels overall, illustrating that, of the cell lines used, the Hep3B cell line is the least amenable to treatment to induce and demonstrate PCSK9 repression (FIGS. 3-5).

[0406] Culturing of treated Huh7 cells continued up to day 27 post-transfection, and PCSK9 secretion was measured at day 14 and day 27. The bar plot in FIG. 6 shows the quantification results of PCSK9 repression at day 4, day 14, and day 27 timepoints displayed as PCSK9 knockdown relative to the levels detected for the naïve control at the day 4 timepoint. The data demonstrate that treatment of Huh7 cells with LTRP5-ADD-ZIM3 with gRNAs having spacers 6.138 and 6.157 resulted in the most effective repression of PCSK9 secretion, and this repression was sustained through day 27 post-transfection (FIG. 6). Similarly, sustained knockdown was observed when Huh7 cells were treated with catalytically-active CasX 676 with spacer 6.1. While treatment with dXR1 and spacer 6.138 resulted in an initial strong repression at day 4, this repressive effect was transient as PCSK9 secretion levels returned to baseline levels at day 14 and day 27 post-transfection (FIG. 6). As anticipated, treatment with any of the three mRNA molecules with spacer 7.37 targeting the *B2M* locus did not affect PCSK9 secretion levels during this time course experiment.

[0407] These experiments demonstrate that use of LTRP molecules with the ADD domain with the appropriate targeting spacer can result in long-term silencing of an endogenous target locus in various human cell lines. These findings also show that LTRP molecules with the ADD domain can be co-delivered as mRNA with a targeting gRNA to cells to induce repression.

Example 3: Design and assessment of modified gRNAs in improving editing when delivered together with CasX mRNA in vitro and in vivo

[0408] Experiments were performed to identify new gRNA variant sequences and demonstrate that chemical modifications of these gRNA variants enhance the editing efficiency of the CasX:gRNA system when delivered in vitro in conjunction with CasX mRNA.

Materials and Methods:

Synthesis of gRNAs:

[0409] All gRNAs tested in this example were chemically-synthesized and were derived from gRNA scaffolds 174, 235, and 316. The sequences of gRNA scaffolds 174, 235, and

316 and their chemical modification profiles are listed in Table 22. The sequences of the resulting gRNAs, including spacers targeting *PCSK9*, *B2M*, or *ROSA26*, and their chemical modification profiles assayed in this example are listed in Table 23. A schematic of the structure of gRNA scaffold variants 174, 235, and 316 are shown in FIGS. 7-9, respectively, and the sites of chemical modifications of the gRNA variants are shown schematically in FIGS. 10, 11, 12, 13, and 14.

Table 22: Sequences of gRNA scaffolds with their different chemical modification profiles (denoted by version number), where “NNNNNNNNNNNNNNNNNNNNNNNN” is a spacer placeholder. Chemical modifications: * = phosphorothioate bond; m = 2’OMe modification

gRNA scaffold (version)	gRNA sequence	SEQ ID NO:
174 (v0)	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGCGACUAUGUCGUAGUGGGUAAAGCUCUCCUCUUCGGAGGGAGCAUCAAGNNNNNNNNNNNNNNNNNNNNNNNN	2947
174 (v1)	mA*mC*mU*GGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGCGACUAUGUCGUAGUGGGUAAAGCUCUCCUCUUCGGAGGGAGCAUCAAGNNNNNNNNNNNNNNNNNNNNNNNNmN*mN*mN	2948
174 (v2)	mA*mC*mU*GGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGCGACUAUGUCGUAGUGGGUAAAGCUCUCCUCUUCGGAGGGAGCAUCAAGNNNNNNNNNNNNNNNNNNNNNNNN*mU*mU*mU	2949
174 (v3)	mA*mC*mU*mGmGmCmGmCmUmUmUmUmAmUmCmUmGmAmUUACUUUGmAmGmAmGmCmCmAmUmCmAmCmCAGCGAmCmUAUmGmUmCmGUAGUGmGmGmUmAmAmAmGmCmUmCmCmCmUmUmCmGmGmAmGmGmGmAmGmCmA mUmCmAAAGNNNNNNNNNNNNNNNNNNNNNNmN*mN*mN	2950
174 (v4)	mA*mC*mU*mGmGmCmGmCUUUUmAmUmCmUmGmAmUUACUUUGmAmGmAmGmCmCmAmUmCmAmCmCAGCGAmCmUAUmGmUmCmGUAGUGmGmGmUmAmAmAmGmCmUmCmCmCmUmUmUmCmGmGmAmGmGmAmGmCmAmUCA AAGNNNNNNNNNNNNNNNNNNNNNNmN*mN*mN	2951
174 (v5)	mA*mC*mU*GGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGCGAmCmUAUmGmUmCmGUAGUGGGUAAAmGmCmUmCmCmUmCmUmUmCmGmGmAmGmGmGmAmGmCAUCAAGNNNNNNNNNNNNNNNNNNNNNNmN*mN*mN	2952
174 (v6)	mA*mC*mU*GGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGCGACUAUGUCGUAGUGGGUAAAmGmCmUmCmCmUmCmUmUmCmGmGmAmGmGmGmAmGmCAUCAAGNNNNNNNNNNNNNNNNNNNNNNmN*mN*mN	2953
174 (v7)	mA*mC*mU*GGmCGmCmUUUUAmUmCUGAUUACUUUGmAmGAGCCmAmUmCmAmCCAGCmGmAmCmUAUmGmUmCmGUAGUGGmGmUmAmAmAmGmCmUmCmCmCmUmCmUmCmGmGmAmGmGmGmAmGmCmAmUCAAGNNNNNNNNNNNNNNNNNNNNNNmN*mN*mN	2954
174 (v8)	mA*mC*mU*GGCGCUUUUAUCUGAUUACUUUGAGAGCCmAmUmCmAmCCA GmCmAmCmUmUmGmUmCmGUAGUGGmGmUmAmAAmGmCmUmCmCmCmUmCmUmUmCmGmGmAmGmGmAmGmCmAmUCAAGNNNNNNNNNNNNNNNNNNNNNNmN*mN*mN	2955
174 (v9)	mA*mC*mU*GGmCmGCmUUUUAmUmCUGAUUACUUUGmAmGAGCCAUCAC CAGCmGmAmCmUAUmGmUmCmGUAGUGGGUAAAmGmCmUmCmCmUmCmUmUmCmGmGmAmGmGmGmAmGmCAUCAAGNNNNNNNNNNNNNNNNNNNNNN*mN*mN*mN	2956
235 (v0)	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGCGACUAUGUCGUAGUGGGUAAAGCCGUUACGACUUCGGUCCGUAAGAGGCAUCAGAGN NNNNNNNNNNNNNNNNNNNNNNNNN	2957

gRNA scaffold (version)	gRNA sequence	SEQ ID NO:
235 (v1)	mA*mC*mU*GGCGCUUCAUCUGAUUACUCUGAGCGCCAUCACCAGCGAC UAUGUCGUAGUGGGUAAAGCCGCUUACGGACUUCGGUCCGUAAGAGGCAU CAGAGNNNNNNNNNNNNNNNNNNmN*mN*mN	2958
235 (v2)	mA*mC*mU*GGCGCUUCAUCUGAUUACUCUGAGCGCCAUCACCAGCGAC UAUGUCGUAGUGGGUAAAGCCGCUUACGGACUUCGGUCCGUAAGAGGCAU CAGAGNNNNNNNNNNNNNNNNNN*mU*mU*mU	2959
235 (v3)	mA*mC*mU*mGmGmCmGmCmUmUmCmUmAmUmCmUmGmAmUUACUCUGmA mGmCmGmCmCmAmUmCmAmCmCAGCGAmCmUAUmGmUmCmGUAGUGmGmG mUmAmAmAmGmCmCmGmCmUmUmAmCmGmGmAmCmUmUmCmGmGmUmCmC mGmUmAmAmGmAmGmGmCmAmUmCmAGAGNNNNNNNNNNNNNNNNmN*m N*mN	2960
235 (v4)	mA*mC*mU*mGmGmCmGmCUUCUmAmUmCmUmGmAmUUACUCUGmAmGmC mGmCmCmAmUmCmAmCmCAGCGAmCmUAUmGmUmCmGUAGUGmGmGmUmA mAmAmGmCmCmGmCmUmUmAmCmGmGmAmCmUmUmCmGmGmUmCmCmGmU mAmAmGmAmGmGmCmAmUCAGAGNNNNNNNNNNNNNNNNmN*mN*mN	2961
235 (v5)	mA*mC*mU*GGCGCUUCAUCUGAUUACUCUGAGCGCCAUCACCAGCGAm CmUAUmGmUmCmGUAGUGGGUmAmAmAmGmCmCmGmCmUmUmAmCmGmGm AmCmUmUmCmGmGmUmCmCmGmUmAmAmGmAmGmCAUCAGAGNNNNNNNN NNNNNNNNNNmN*mN*mN	2962
235 (v6)	mA*mC*mU*GGCGCUUCAUCUGAUUACUCUGAGCGCCAUCACCAGCGAC UAUGUCGUAGUGGGUmAmAmAmGmCmCmGmCmUmUmAmCmGmGmAmCmUm UmCmGmGmUmCmCmGmUmAmAmGmAmGmCAUCAGAGNNNNNNNNNNNNNN NNNNmN*mN*mN	2963
235 (v7)	mA*mC*mU*GGmCGmCmUUCUmAmUmCUGAUUACUCUGmAmGCGCCmAmUm CmAmCCAGCmGmAmCmUAUmGmUmCmGUAGUGGmGmUmAmAmGmCmCmG mCmUmUmAmCmGmGmAmCmUmUmCmGmGmUmCmCmGmUmAmAmGmAmGmG mCmAmUCAGAGNNNNNNNNNNNNNNNN*mN*mN*mN	2964
235 (v8)	mA*mC*mU*GGCGCUUCAUCUGAUUACUCUGAGCGCCmAmUmCmAmCCA GmGmAmCmUAUmGmUmCmGUAGUGGmGmUmAmAmGmCmCmGmCmUmUm AmCmGmGmAmCmUmUmCmGmGmUmCmCmGmUmAmAmGmAmGmCmAmUC AGAGNNNNNNNNNNNNNNNN*mN*mN*mN	2965
235 (v9)	mA*mC*mU*GGmCGmCmUUCUmAmUmCUGAUUACUCUGmAmGCGCCAUCAC CAGCmGmAmCmUAUmGmUmCmGUAGUGGGUAAmGmCmCmGmCmUmUmAm CmGmGmAmCmUmUmCmGmGmUmCmCmGmUmAmAmGmAmGmCAUCAGAG NNNNNNNNNNNNNNNN*mN*mN*mN	2966
316 (v0)	ACUGGCGCUUCAUCUGAUUACUCUGAGCGCCAUCACCAGCGACUAUGUC GUAGUGGGUAAAGCUCCUUCGGAGGGAGCAUCAGAGNNNNNNNNNNNNNN NNNNNNNNNN	2967
316 (v1)	mA*mC*mU*GGCGCUUCAUCUGAUUACUCUGAGCGCCAUCACCAGCGAC UAUGUCGUAGUGGGUAAAGCUCCUUCGGAGGGAGCAUCAGAGNNNNNN NNNNNNNNNNNN*mN*mN*mN	2968
316 (v2)	mA*mC*mU*GGCGCUUCAUCUGAUUACUCUGAGCGCCAUCACCAGCGAC UAUGUCGUAGUGGGUAAAGCUCCUUCGGAGGGAGCAUCAGAGNNNNNN NNNNNNNNNNNNNN*mU*mU*mU	2969
316 (v3)	mA*mC*mU*mGmGmCmGmCmUmUmCmUmAmUmCmUmGmAmUUACUCUGmA mGmCmGmCmCmAmUmCmAmCmCAGCGAmCmUAUmGmUmCmGUAGUGmGmG mUmAmAmAmGmCmUmCmCmUmCmUmUmCmGmGmAmGmGmAmGmCmA mUmCmAGAGNNNNNNNNNNNNNNNN*mN*mN*mN	2970
316 (v4)	mA*mC*mU*mGmGmCmGmCUUCUmAmUmCmUmGmAmUUACUCUGmAmGmC mGmCmCmAmUmCmAmCmCAGCGAmCmUAUmGmUmCmGUAGUGmGmGmUmA mAmAmGmCmUmCmCmUmCmUmUmCmGmGmAmGmGmAmGmCmAmUCA GAGNNNNNNNNNNNNNNNN*mN*mN*mN	2971
316 (v5)	mA*mC*mU*GGCGCUUCAUCUGAUUACUCUGAGCGCCAUCACCAGCGAm CmUAUmGmUmCmGUAGUGGGUAAmGmCmUmCmCmUmCmUmUmCmGmGm mAmGmGmAmGmCAUCAGAGNNNNNNNNNNNNNNNN*mN*mN*mN	2972
316 (v6)	mA*mC*mU*GGCGCUUCAUCUGAUUACUCUGAGCGCCAUCACCAGCGAC UAUGUCGUAGUGGGUAAmGmCmUmCmCmUmCmUmUmCmGmGmAmGmG mGmAmGmCAUCAGAGNNNNNNNNNNNNNNNN*mN*mN*mN	2973

gRNA scaffold (version)	gRNA sequence	SEQ ID NO:
316 (v7)	mA*mC*mU*GGmCGmCmUUCUAmUmCUGAUUACUCUGmAmGCGCCmAmUmCmAmCCAGCmGmAmCmUAUmGmUmCmGUAGUGGmGmUmAmAmGmCmUmCmCmUmCmUmCmGmGmAmGmGmGmAmGmCmAmUCAGAGNNNNNNNNNNNNNNNNNNNNNN*mN*mN*mN	2974
316 (v8)	mA*mC*mU*GGCGCUUCUAUCUGAUUACUCUGAGCGCCmAmUmCmAmCCAGCmGmAmCmUAUmGmUmCmGUAGUGGmGmUmAmAmGmCmUmCmCmUmCmUmCmGmGmAmGmGmGmAmGmCmAmUCAGAGNNNNNNNNNNNNNNNNNNNNNN*mN*mN*mN	2975
316 (v9)	mA*mC*mU*GGmCGmCmUUCUAmUmCUGAUUACUCUGmAmGCGCCAUCACAGCmGmAmCmUAUmGmUmCmGUAGUGGGUAAAmGmCmUmCmCmUmCmUmCmUmCmGmGmAmGmGmGmAmGmCAUCAGAGNNNNNNNNNNNNNNNNNNNNNN*mN*mN*mN	2976

Table 23: Sequences of gRNAs with their different chemical modification profiles (denoted by version number) assayed in this example. Chemical modifications: * = phosphorothioate bond; m = 2'OMe modification

gRNA ID (scaffold variant-spacer)	Target	gRNA sequence	SEQ ID NO:
174-6.7 (v0)	human <i>PCSK9</i>	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGCGACUAUGUCGUAGUGGGUAAAGCUCCUCUUCGGAGGGAGCAUCAAAGUCCUGGCUUCCUGGUGAAGA	3145
174-6.7 (v1)	human <i>PCSK9</i>	mA*mC*mU*GGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGCGACUAUGUCGUAGUGGGUAAAGCUCCUCUUCGGAGGGAGCAUCAAAGUCCUGGCUUCCUGGUGAmA*mG*mA	3074
174-6.8 (v0)	human <i>PCSK9</i>	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGCGACUAUGUCGUAGUGGGUAAAGCUCCUCUUCGGAGGGAGCAUCAAAGUCCUGGCUUCCUGGUGAAGAUGA	3146
174-6.8 (v1)	human <i>PCSK9</i>	mA*mC*mU*GGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGCGACUAUGUCGUAGUGGGUAAAGCUCCUCUUCGGAGGGAGCAUCAAAGUCCUGGCUUCCUGGUGAAGAmU*mG*mA	3147
174-7.9 (v0)	human <i>B2M</i>	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGCGACUAUGUCGUAGUGGGUAAAGCUCCUCUUCGGAGGGAGCAUCAAAGGUGUAGUACAAGAGAUAGAA	3148
174-7.9 (v1)	human <i>B2M</i>	mA*mC*mU*GGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGCGACUAUGUCGUAGUGGGUAAAGCUCCUCUUCGGAGGGAGCAUCAAAGGUGUAGUACAAGAGAUAmG*mA*mA	3149
316-6.7 (v0)	human <i>PCSK9</i>	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGCGACUAUGUCGUAGUGGGUAAAGCUCCUCUUCGGAGGGAGCAUCAGAGUCCUGGCUUCCUGGUGAAGA	3150
316-6.7 (v1')	human <i>PCSK9</i>	mA*mC*mU*GGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGCGACUAUGUCGUAGUGGGUAAAGCUCCUCUUCGGAGGGAGCAUCAGAGUCCUGGCUUCCUGGUGAmA*mG*mA	3151
316-6.8 (v0)	human <i>PCSK9</i>	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGCGACUAUGUCGUAGUGGGUAAAGCUCCUCUUCGGAGGGAGCAUCAGAGUCCUGGCUUCCUGGUGAAGAUGA	3152

gRNA ID (scaffold variant-spacer)	Target	gRNA sequence	SEQ ID NO:
316-6.8 (v1')	human <i>PCSK9</i>	mA*mC*mU*GGCGCUUCUAUCUGAUUACUCUGAGCGCCAUC ACCAGCGACUAUGUCGUAGUGGGUAAAGCUCUCCUCUUCGGA GGGAGCAUCAGAGUGGCUUCCUGGUGAAGAmU*mG*mA	3153
316-7.9 (v0)	human <i>B2M</i>	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCUCCUCUUCGAGGGGAGC AUCAGAGGUGUAGUACAAGAGAUAGAA	3154
316-7.9 (v1')	human <i>B2M</i>	mA*mC*mU*GGCGCUUCUAUCUGAUUACUCUGAGCGCCAUC ACCAGCGACUAUGUCGUAGUGGGUAAAGCUCUCCUCUUCGGA GGGAGCAUCAGAGGUGUAGUACAAGAGAUAmG*mA*mA	3155
174-7.37 (v0)	human <i>B2M</i>	ACUGGCGCUUUUAUCUgAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAgUGGGUAAAGCUCUCCUCUUCGAGGGGAGC AUCAAAGGGCCGAGAUgUCUCGCUC	3156
174-7.37 (v1*)	human <i>B2M</i>	mA*mC*mU*GGCGCUUUUAUCUgAUUACUUUGAGAGCCAUC ACCAGCGACUAUGUCGUAgUGGGUAAAGCUCUCCUCUUCGGA GGGAGCAUCAAGGGCCGAGAUgUCUCG*mC*mU*mC	3157
235-6.7 (v0)	human <i>PCSK9</i>	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCCGCUUACGGACUUCGGUC CGUAAGAGGCAUCAGAGUCCUGGCUUCCUGGUGAAGA	3158
235-6.7 (v1)	human <i>PCSK9</i>	mA*mC*mU*GGCGCUUCUAUCUGAUUACUCUGAGCGCCAUC ACCAGCGACUAUGUCGUAGUGGGUAAAGCCGCUUACGGACU UCGGUCCGUAAGAGGCAUCAGAGUCCUGGCUUCCUGGUGAm A*mG*mA	3159
235-6.7 (v2)	human <i>PCSK9</i>	mA*mC*mU*GGCGCUUCUAUCUGAUUACUCUGAGCGCCAUC ACCAGCGACUAUGUCGUAGUGGGUAAAGCCGCUUACGGACU UCGGUCCGUAAGAGGCAUCAGAGUCCUGGCUUCCUGGUGAA GAU*mU*mU*mU	3160
235-6.7 (v3)	human <i>PCSK9</i>	mA*mC*mU*mGmGmCmGmCmUmUmCmUmAmUmCmUmGmAmU UACUCUGmAmGmCmGmCmAmUmCmAmCmCAGCGAmCmUA UmGmUmCmGUAGUGmGmGmUmAmAmAmGmCmGmCmUmUm AmCmGmGmAmCmUmUmCmGmGmUmCmCmGmUmAmAmGmAmG mGmCmAmUmCmAGAGUCCUGGCUUCCUGGUGAmA*mG*mA	3161
235-6.7 (v4)	human <i>PCSK9</i>	mA*mC*mU*mGmGmCmGmCUUCUmAmUmCmUmGmAmUUACU CUGmAmGmCmGmCmAmUmCmAmCmCAGCGAmCmUAUmGm UmCmGUAGUGmGmGmUmAmAmAmGmCmCmGmCmUmUmAmCm GmGmAmCmUmUmCmGmGmUmCmCmGmUmAmAmGmAmGmC mAmUCAGAGUCCUGGCUUCCUGGUGAmA*mG*mA	3162
235-6.7 (v5)	human <i>PCSK9</i>	mA*mC*mU*GGCGCUUCUAUCUGAUUACUCUGAGCGCCAUC ACCAGCGAmCmUAUmGmUmCmGUAGUGGGUmAmAmAmGmCm CmGmCmUmUmAmCmGmGmAmCmUmUmCmGmGmUmCmCmGmU mAmAmGmAmGmGmCAUCAGAGUCCUGGCUUCCUGGUGAmA* mG*mA	3163
235-6.7 (v6)	human <i>PCSK9</i>	mA*mC*mU*GGCGCUUCUAUCUGAUUACUCUGAGCGCCAUC ACCAGCGACUAUGUCGUAGUGGGUmAmAmAmGmCmGmCm UmUmAmCmGmGmAmCmUmUmCmGmGmUmCmCmGmUmAmAmG mAmGmGmCAUCAGAGUCCUGGCUUCCUGGUGAmA*mG*mA	3164

gRNA ID (scaffold variant-spacer)	Target	gRNA sequence	SEQ ID NO:
235-6.8 (v0)	human <i>PCSK9</i>	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCCGCUUACGGACUUCGGUC CGUAAGAGGCAUCAGAGUGGCUUCCUGGUGAAGAUGA	3165
235-6.8 (v1)	human <i>PCSK9</i>	mA*mC*mU*GGCGCUUCUAUCUGAUUACUCUGAGCGCCAUC ACCAGCGACUAUGUCGUAGUGGGUAAAGCCGCUUACGGACU UCGGUCCGUAAGAGGCAUCAGAGUGGCUUCCUGGUGAAGAm U*mG*mA	3166
235-6.8 (v2)	human <i>PCSK9</i>	mA*mC*mU*GGCGCUUCUAUCUGAUUACUCUGAGCGCCAUC ACCAGCGACUAUGUCGUAGUGGGUAAAGCCGCUUACGGACU UCGGUCCGUAAGAGGCAUCAGAGUGGCUUCCUGGUGAAGAU GA*mU*mU*mU	3167
235-6.8 (v3)	human <i>PCSK9</i>	mA*mC*mU*mGmGmCmGmCmUmCmUmAmUmCmUmGmAmU UACUCUGmAmGmCmGmCmAmUmCmAmCmCAGCGAmCmUA UmGmUmCmGUAGUGmGmGmUmAmAmAmGmCmCmGmCmUmUm AmCmGmGmAmCmUmUmCmGmGmUmCmCmGmUmAmAmGmAmG mGmCmAmUmCmAGAGUGGCUUCCUGGUGAAGAmU*mG*mA	3168
235-6.8 (v4)	human <i>PCSK9</i>	mA*mC*mU*mGmGmCmGmCUUCUmAmUmCmUmGmAmUUACU CUGmAmGmCmGmCmAmUmCmAmCmCAGCGAmCmUAUmGm UmCmGUAGUGmGmGmUmAmAmAmGmCmCmGmCmUmUmAmCm GmGmAmCmUmUmCmGmGmUmCmCmGmUmAmAmGmAmGmCm mAmUCAGAGUGGCUUCCUGGUGAAGAmU*mG*mA	3169
235-6.8 (v5)	human <i>PCSK9</i>	mA*mC*mU*GGCGCUUCUAUCUGAUUACUCUGAGCGCCAUC ACCAGCGAmCmUAUmGmUmCmGUAGUGGGUmAmAmAmGmCm CmGmCmUmUmAmCmGmAmCmUmUmCmGmGmUmCmCmGmU mAmAmGmAmGmCmCAUCAGAGUGGCUUCCUGGUGAAGAmU* mG*mA	3170
235-6.8 (v6)	human <i>PCSK9</i>	mA*mC*mU*GGCGCUUCUAUCUGAUUACUCUGAGCGCCAUC ACCAGCGACUAUGUCGUAGUGGGUmAmAmAmGmCmCmGmCm UmUmAmCmGmGmAmCmUmUmCmGmGmUmCmCmGmUmAmAmG mAmGmGmCAUCAGAGUGGCUUCCUGGUGAAGAmU*mG*mA	3171
316-27.107 (v0)	mouse <i>PCSK9</i>	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCCCUUUCGGAGGGAGC AUCAGAGCUGGCUUCUUGGUGAAGAUG	3172
316-27.107 (v1)	mouse <i>PCSK9</i>	mA*mC*mU*GGCGCUUCUAUCUGAUUACUCUGAGCGCCAUC ACCAGCGACUAUGUCGUAGUGGGUAAAGCUCCCUUUCGGA GGGAGCAUCAGAGCUGGCUUCUUGGUGAAG*mA*mU*mG	3173
316-27.107 (v7)	mouse <i>PCSK9</i>	mA*mC*mU*GGmCmGmCmUUCUAmUmCUGAUUACUCUGmAmG CGCCmAmUmCmAmCCAGCmGmAmCmUAUmGmUmCmGUAGUG GmGmUmAmAmGmCmUmCmCmUmCmUmUmCmGmGmAmGm GmGmAmGmCmAmUCAGAGCUGGCUUCUUGGUGAAG*mA*mU *mG	3174
316-27.107 (v8)	mouse <i>PCSK9</i>	mA*mC*mU*GGCGCUUCUAUCUGAUUACUCUGAGCGCmAm UmCmAmCCAGCmGmAmCmUAUmGmUmCmGUAGUGGmGmUmA mAmAmGmCmUmCmCmUmCmUmUmCmGmGmAmGmGmAmG mCmAmUCAGAGCUGGCUUCUUGGUGAAG*mA*mU*mG	3175
316-27.107	mouse	mA*mC*mU*GGmCmGmCmUUCUAmUmCUGAUUACUCUGmAmG	3176

gRNA ID (scaffold variant-spacer)	Target	gRNA sequence	SEQ ID NO:
(v9*)	PCSK9	CGCCAUCACCAGCmGmAmCmUAUmGmUmCmGUAGUGGGUAA AmGmCmUmCmCmCmUmCmUmUmCmGmGmAmGmGmAmGmC AUCAGAGCUGGCCUUCUUGGUGAA*mG*mA*mU*mG	
174-35.2 (v0)	<i>ROSA26</i>	ACUGGCGCUUUUAUCUGAUUACUUUGAGAGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCUCCUUCGGAGGGAGC AUCAAAGAGAAGAUGGGCGGGAGUCUU	3177
174-35.2 (v2)	<i>ROSA26</i>	mA*mC*mU*GGCGCUUUUAUCUGAUUACUUUGAGAGCCAUC ACCAGCGACUAUGUCGUAGUGGGUAAAGCUCUCCUUCGGA GGGAGCAUCAAGAGAAGAUGGGCGGGAGUCUU*mU*mU*m U	3178
316-35.2 (v0)	<i>ROSA26</i>	ACUGGCGCUUCUAUCUGAUUACUCUGAGCGCCAUCACCAGC GACUAUGUCGUAGUGGGUAAAGCUCUCCUUCGGAGGGAGC AUCAGAGAGAAGAUGGGCGGGAGUCUU	3179
316-35.2 (v1)	<i>ROSA26</i>	mA*mC*mU*GGCGCUUCUAUCUGAUUACUCUGAGCGCCAUC ACCAGCGACUAUGUCGUAGUGGGUAAAGCUCUCCUUCGGA GGGAGCAUCAGAGAGAAGAUGGGCGGGAGU*mC*mU*mU	3180
316-35.2 (v5)	<i>ROSA26</i>	mA*mC*mU*GGCGCUUCUAUCUGAUUACUCUGAGCGCCAUC ACCAGCGAmCmUAUmGmUmCmGUAGUGGGUAAAmGmCmUmC mCmCmUmCmUmUmCmGmGmAmGmGmAmGmCAUCAGAGAG AAGAUGGGCGGGAGU*mC*mU*mU	3181

Note that gRNAs annotated with a v1' design contain one less phosphorothioate bond on the 3' end of the gRNA. gRNAs annotated with v1* contain one extra phosphorothioate bond on the 3' end of the gRNA. gRNAs annotated with a v9* contain an extra phosphorothioate bond on the 3' end of the gRNA.

Biochemical characterization of gRNA activity:

[0410] Target DNA oligonucleotides with fluorescent moieties on the 5' ends were purchased commercially (sequences listed in Table 24). Double-stranded DNA (dsDNA) targets were formed by mixing the oligos in a 1:1 ratio in 1x cleavage buffer (20 mM Tris HCl pH 7.5, 150 mM NaCl, 1 mM TCEP, 5% glycerol, 10 mM MgCl₂), following by heating to 95°C for 10 minutes, and then allowing the solution to cool to room temperature. CasX ribonucleoproteins (RNPs) were reconstituted with CasX 491 and the indicated gRNAs at a final concentration of 1 μM with 1.2-fold excess of the indicated gRNA in 1x cleavage buffer. RNPs were allowed to form at 37°C for 10 minutes.

[0411] The effects of various structural and chemical modifications to the gRNA scaffold on the cleavage rate of CasX 491 RNPs were determined. Cleavage reactions were prepared with final RNP concentrations of 200 nM and final target concentrations of 10 nM, and reactions were carried out at 16°C and initiated by the addition of the labeled target DNA

substrate (Table 24). Aliquots of reactions were taken at 0.25, 0.5, 1, 2, 5, and 10 minutes and quenched by adding an equal volume of 95% formamide with 20 mM EDTA. Samples were denatured at 95°C for 10 minutes and resolved on a 10% urea-PAGE gel. Gels were imaged on a TyphoonTM laser-scanner platform and quantified using ImageQuantTM TL 8.2 image analysis software (CytivaTM). The apparent first-order rate constant of non-target strand cleavage ($k_{\text{cleave-}}$) was determined for each CasX:gRNA combination.

[0412] To determine the competent fraction formed by each gRNA, cleavage reactions were prepared with final RNP concentrations of 100 nM and final target concentrations of 100 nM. Reactions were carried out at 37°C and initiated by the addition of the labeled target substrate (Table 24). Aliquots were taken at 0.5, 1, 2, 5, 10, and 30 minutes and quenched by adding an equal volume of 95% formamide with 25 mM EDTA. Samples were denatured by heating at 95°C for 10 minutes and resolved on a 10% urea-PAGE gel. Gels were imaged and quantified as above. CasX was assumed to act as a single-turnover enzyme under the assayed conditions, as indicated by the observation that sub-stoichiometric amounts of enzyme would fail to cleave a greater-than-stoichiometric amount of target substrate even under extended time-scales, and instead would approach a plateau that scaled with the amount of enzyme present. Thus, the fraction of target substrate cleaved over long time-scales by an equimolar amount of RNP would be indicative of the fraction of RNP that was properly formed and active for cleavage. The cleavage traces were fitted with a biphasic rate model, as the cleavage reaction clearly deviated from monophasic under this concentration regime. The plateau of each fit was determined and reported as the active fraction for each RNP in Table 27.

Table 24: Sequences of target DNA substrate oligonucleotides with fluorescent moieties on the 5' ends used for biochemical characterization of gRNA activity. /700/ = IRDye700; /800/ = IRDye800

DNA substrate	Sequence
6.7/6.8 target top strand (SEQ ID NO: 3182)	/700/catgtcttccatggccttcttcctggcttctggtgaagatgagtggcgacctgctggag
6.7/6.8 target bottom strand (SEQ ID NO: 3183)	/800/ctccagcaggtcgccactcatcttcaccaggaagccaggaagaaggccatggaagacatg

In vitro transcription of CasX mRNA:

[0413] DNA templates encoding for CasX 491 (see Table 25 for encoding sequences) used for in vitro transcription were generated by PCR using forward primers containing a T7 promoter, followed by agarose gel extraction of the appropriately sized DNA. 25 ng/μL final concentration of template DNA was used in each in vitro transcription reaction that was carried out following the manufacturer's recommended protocol with slight modifications. Following in vitro transcription reaction incubation for 2-3 hours at 37°C, which were carried out with CleanCap® AG and N1-methyl-pseudouridine, DNase digestion of template DNA and column-based purification using the Zymo RNA miniprep kit were performed. The poly(A) tail was added using *E. coli* PolyA Polymerase following the manufacturer's protocol, followed by column-based purification as stated above. Poly(A) tailed in vitro transcribed RNA was eluted in RNase free water, analyzed on an Agilent TapeStation for integrity, and flash frozen prior to storage at -80°C.

Table 25: Encoding sequences of the CasX mRNA molecules assessed in this example*

CasX 491 mRNA ID	Component (ID)	DNA sequence	SEQ ID NO:
CasX 491 mRNA #1	5'UTR	GACCGGCCGCCACC	3082
	START codon + c-MYC NLS + linker	ATGGCCCCAGCGGCCAAACGGGTGAAGCTGGACTCTAGA	3184
	CasX 491	CAAGAGATCAAGAGAATCAACAAGATCAGAAGGAGACTG GTCAAGGACAGCAACACAAAGAAGGCCGGCAAGACAGGC CCCATGAAAACCTGCTCGTCAGAGTGATGACCCCTGAC CTGAGAGAGCGGCTGGAAAACCTGAGAAAGAAGCCCGAG AACATCCCTCAGCCTATCAGCAACACCAGCAGGGCCAAC CTGAACAAGCTGCTGACCGACTACACCGAGATGAAGAAA GCCATCCTGCACGTGTACTGGGAAGAGTTCCAGAAAGAC CCCGTGGGCTGATGAGCAGAGTTGCTCAGCCTGCCAGC AAGAAGATCGACCAGAACAAGCTGAAGCCCAGATGGAC GAGAAGGGCAATCTGACCACAGCCGGCTTTGCCTGCTCT CAGTGTGGCCAGCCTCTGTTTCGTGTACAAGCTGGAACAG GTGTCCGAGAAAGGCAAGGCCTACACCACTACTTCGGC AGATGTAACGTGGCCGAGCACGAGAAGCTGATTCTGCTG GCCAGCTGAAACCTGAGAAGGACTCTGATGAGGCCGTG ACCTACAGCCTGGGCAAGTTTGGACAGAGAGCCCTGGAC TTCTACAGCATCCACGTGACCAAAGAAAGCACACACCCC GTGAAGCCCCTGGCTCAGATCGCCGGCAATAGATACGCC TCTGGACCTGTGGGCAAAGCCCTGTCCGATGCCTGCATG GGAACAATCGCCAGCTTCCTGAGCAAGTACCAGGACATC ATCATCGAGCACCAGAAGGTGGTCAAGGGCAACCAGAAG AGACTGGAAGCCTGAGGGAGCTGGCCGGCAAAGAGAAC CTGGAATACCCAGCGTGACCCTGCCTCCTCAGCCTCAC ACAAAAGAAGGCGTGGACGCCTAACCGAAGTGATCGCC AGAGTGAGAATGTGGGTCAACCTGAACCTGTGGCAGAAG CTGAAACTGTCCAGGGACGACGCCAAGCCTCTGCTGAGA CTGAAGGGCTTCCTAGCTTCCTCTGGTGAAAGACAG	3185

CasX 491 mRNA ID	Component (ID)	DNA sequence	SEQ ID NO:
		GCCAATGAAGTGGATTGGTGGGACATGGTCTGCAACGTG AAGAAGCTGATCAACGAGAAGAAAGAGGATGGCAAGGTT TTCTGGCAGAACCTGGCCGGCTACAAGAGACAAGAAGCC CTGAGGCCCTTACCTGAGCAGCGAAGAGGACCGGAAGAAG GGCAAGAAGTTCGCCAGATACCAGCTGGGCGACCTGCTG CTGCACCTGGAAAAGAAGCACGGCGAGGACTGGGGCAAA GTGTACGATGAGGCCTGGGAGAGAATCGACAAGAAGGTG GAAGGCCTGAGCAAGCACATTAAGCTGGAAGAGGAAAGA AGGAGCGAGGACGCCCAATCTAAAGCCGCTCTGACCGAT TGGCTGAGAGCCAAGGCCAGCTTTGTGATCGAGGGCCTG AAAGAGGCCGACAAGGACGAGTTCTGCAGATGCGAGCTG AAGCTGCAGAAGTGGTACGGCGATCTGAGAGGCAAGCCC TTCGCCATTGAGGCCGAGAACAGCATCCTGGACATCAGC GGCTTCAGCAAGCAGTACAACCTGCGCCTTCATTTGGCAG AAAGACGGCGTCAAGAACTGAACCTGTACCTGATCATC AATTACTTCAAAGGCGGCAAGCTGCGGTTCAAGAAGATC AAACCCGAGGCCTTCGAGGCTAACAGATTCTACACCGTG ATCAACAAAAAGTCCGGCGAGATCGTGCCCATGGAAGTG AACTTCAACTTCGACGACCCCAACCTGATTATCCTGCCT CTGGCCTTCGGCAAGAGACAGGGCAGAGAGTTCATCTGG AACGATCTGCTGAGCCTGGAAACCGGCTCTCTGAAGCTG GCCAATGGCAGAGTGATCGAGAAAACCTGTACAACAGG AGAACCAGACAGGACGAGCCTGCTCTGTTTGTGGCCCTG ACCTTCGAGAGAAGAGAGGTGCTGGACAGCAGCAACATC AAGCCCATGAACCTGATCGGCGTGGACCGGGGCGAGAAT ATCCCTGCTGTGATCGCCCTGACAGACCCTGAAGGATGC CCACTGAGCAGATTCAAGGACTCCCTGGGCAACCCCTACA CACATCCTGAGAATCGGCGAGAGCTACAAAGAGAAGCAG AGGACAATCCAGGCCAAGAAAGAGGTGGAACAGCGCAGA GCCGGCGGATACTCTAGGAAGTACGCCAGCAAGGCCAAG AATCTGGCCGACGACATGGTCCGAAACACCGCCAGAGAT CTGCTGTACTACGCCGTGACACAGGACGCCATGCTGATC TTCGAGAACTCTGAGCAGAGGCTTCGGCCGGCAGGGCAAG AGAACCTTTATGGCCGAGAGGCAGTACACCAGAATGGAA GATTGGCTCACAGCTAAACTGGCCTACGAGGGACTGAGC AAGACCTACCTGTCCAAAACACTGGCCCAGTATACCTCC AAGACCTGCAGCAATTGCGGCTTCACCATCACCAGCGCC GACTACGACAGAGTGCTGGAAAAGCTCAAGAAAACCGCC ACCGGCTGGATGACCACCATCAACGGCAAAGAGCTGAAG GTTGAGGGCCAGATCACCTACTACAACAGGTACAAGAGG CAGAACGTCGTGAAGGATCTGAGCGTGGAACCTGGACAGA CTGAGCGAAGAGAGCGTGAACAACGACATCAGCAGCTGG ACAAAGGGCAGATCAGGCGAGGCTCTGAGCCTGCTGAAG AAGAGGTTTAGCCACAGACCTGTGCAAGAGAAGTTCGTG TGCCTGAACCTGCGGCTTCGAGACACACGCCGATGAACAG GCTGCCCTGAACATTGCCAGAAGCTGGCTGTTCTCTGAGA AGCCAAGAGTACAAGAAGTACCAGACCAACAAGACCACC GGCAACACCGACAAGAGGGCCTTTGTGGAACCTGGCAG AGCTTCTACAGAAAAAAGCTGAAAGAAGTCTGGAAGCCCC GCCGTG	
	Linker + c-MYC NLS	GGCTCCCCCGCCGCGAAGCGAGTGAAACTGGACGGATCCGGC	3186
	P2A mScarlet + STOP codon	GCAACAACTTCTCTCTGCTGAAACAAGCCGGAGATGTC GAAGAGAATCCTGGACCGATGGTATCAAAGGGTGAGGCG GTCATTAAGGAATTCATGCGGTTCAAAGTTCACATGGAG GGGAGTATGAATGGTCACGAATTTGAGATCGAGGGTGAG GGAGAGGGTAGACCTTACGAAGGCACCCAGACAGCAAAA CTCAAAGTCACTAAGGGCGGCCACTCCCTTTCTCATGG	3187

CasX 491 mRNA ID	Component (ID)	DNA sequence	SEQ ID NO:
		GATATACTGAGTCCTCAATTTATGTATGGCTCCCGGGCC TTTATCAAGCATCCAGCCGACATACCCGATTACTATAAA CAGTCATTCCCCGAGGGATTCAAATGGGAGAGAGTGATG AATTTTGAAGACGGGGGTGCCGTGACCGTAACTCAAGAT ACCTCTTTGGAAGATGGTACCCTCATTTACAAGGTTAAG CTTAGAGGAACTAATTTCCACAGACGGACCCGTTATG CAGAAAAAACCATGGGTTGGGAGGCTAGCACCGAACGG CTGTATCCAGAGGATGGAGTTTGTAAAGGTGATATTAAG ATGGCCCTGCGCCTGAAAGACGGAGGGCGGTACCTCGCA GATTTCAAACTACCTACAAAGCTAAGAAACCGGTCCAA ATGCCGGGTGCTTATAACGTCGATAGGAAGCTGGATATA ACATCACACAACGAAGACTATACGGTTGTGCAACAGTAT GAGCGATCTGAGGGGAGGCACAGCACGGGCGGGATGGAT GAGCTTTACAAGTGA	
CasX 676 mRNA #2	5'UTR	AAATAAGAGAGAAAAGAAGAGTAAGAAGAAATATAAGAG CCACC	3047
	START codon + c-MYC NLS	ATGGCCCCCTGCTGCCAAGAGAGTGAAGCTGGATAGCAGA	3133
	CasX 676	<i>See sequence in Table 19</i>	3134
	c-MYC NLS + STOP codons	GGAAGCCCCGCTGCCAAGAGAGTGAAGCTGGACTAATAG ATAA	3135
	3'UTR	GCTGCCTTCTGCGGGGCTTGCCTTCTGGCCATGCCCTTC TTCTCTCCCTTGACCTGTACCTCTTGGTCTTTGAATAA AGCCTGAGTAGGAAG	3055
	XbaI restriction site (partial)	TCTAG	3056
	Poly(A) tail	AA AA AA	3057

*Components are listed in a 5' to 3' order within the constructs

In vitro delivery of gRNA and CasX mRNA via transfection:

[0414] Editing at the *PCSK9* locus and consequential effects on secreted PCSK9 levels were assessed for conditions using CasX 491 mRNA co-delivered with a *PCSK9*-targeting gRNA with scaffold variant 174 compared to conditions where a *PCSK9*-targeting gRNA with scaffold variant 316 was used. 100 ng of in vitro transcribed mRNA coding for CasX 491 with a P2A and mScarlet fluorescent protein was transfected into HepG2 cells with version 1 (v1) of gRNAs 174-6.7, 174-6.8, 316-6.7, and 316-6.8 (see Table 23) using lipofectamine. After a media change, the following were harvested at 28 hours post-transfection: 1) transfected cells were harvested for editing assessment at the *PCSK9* locus by NGS (next generation sequencing); and 2) media supernatant was harvested to measure secreted PCSK9 protein levels by ELISA. For editing analysis by NGS, amplicons were amplified from 200 ng of extracted gDNA with a set of primers targeting the *PCSK9* locus

and processed by NGS (described below). Secreted PCSK9 levels in the media supernatant were also analyzed using a fluorescence resonance energy transfer-based immunoassay from CISBio following the manufacturer's instructions. Here, a gRNA using scaffold 174 with spacer 7.37 (v0; see Table 23), which targeted the endogenous *B2M* (beta-2-microglobulin) locus, served as the non-targeting (NT) control. These results are shown in FIG. 15.

[0415] To compare the editing potency of unmodified version 0 (v0) and chemically-modified version 1 (v1) of *B2M*-targeting gRNAs, ~6E4 HepG2 hepatocytes were seeded per well of a 96-well plate. 24 hours later, seeded cells were co-transfected using lipofectamine with 100 ng of in vitro transcribed mRNA coding for CasX 491 and different doses (1, 5, or 50 ng) of either v0 or v1 of the *B2M*-targeting gRNA containing scaffold variant 174 and spacer 7.37 (see Table 23). Six days post-transfection, cells were harvested for B2M protein expression analysis via immunostaining of the B2M-dependent HLA protein, followed by flow cytometry using the Attune™ NxT flow cytometer. These results are shown in FIG. 16.

[0416] V1 through v6 variants of chemically-modified *PCSK9*-targeting gRNAs (Table 23) were assessed for their effects on editing potency and consequential effects on secreted PCSK9 levels in vitro. Briefly, 100 ng of in vitro transcribed mRNA coding for CasX variant 491 and a P2A and mScarlet fluorescent protein was transfected into HepG2 cells with 50 ng of the indicated chemically-modified gRNA using lipofectamine. After a media change, the following were harvested at 28 hours post-transfection: 1) transfected cells for editing assessment at the *PCSK9* locus by NGS as described above; and 2) media supernatant to measure secreted PCSK9 protein levels by ELISA, as described above. Here, a *B2M*-targeting gRNA was used as a non-targeting control. These results are shown in Table 28.

NGS processing and analysis:

[0417] Genomic DNA (gDNA) from harvested cells were extracted using the Zymo Quick-DNA Miniprep Plus kit following the manufacturer's instructions. Target amplicons were formed by amplifying regions of interest from ~50-200 ng of extracted gDNA with a set of primers targeting the human *PCSK9* locus. These gene-specific primers contained an additional sequence at the 5' ends to introduce Illumina reads 1 and 2 sequences. Further, they contained a 16-nucleotide random sequence that functioned as a unique molecular identifier (UMI). The quality and quantification of the amplicon was assessed using a Fragment Analyzer DNA analyzer kit (Agilent, dsDNA 35-1500 bp). Amplicons were sequenced on the Illumina MiSeq™ according to the manufacturer's instructions. Raw fastq sequencing files were processed by trimming for quality and adapter sequences and merging read 1 and read 2 into a single insert sequence; insert sequences were then analyzed by the

CRISPResso2 (v 2.0.29) program. The percentage of reads modified in a window around the 3' end of the spacer was determined. The activity of the CasX molecule was quantified as the total percent of reads that contain insertions, substitutions, and/or deletions anywhere within this window for each.

Formulations of lipid nanoparticles (LNPs):

[0418] CasX mRNA and gRNA were encapsulated into LNPs using GenVoy-ILM™ lipids on the Precision NanoSystems Inc. (PNI) Ignite™ Benchtop System and following the manufacturer's guidelines. GenVoy-ILM™ lipids are manufactured by PNI, with a composition of ionizable lipid:DSPC:cholesterol:stabilizer at 50:10:37.5:2.5 mol%.

[0419] Briefly, to formulate LNPs, equal mass ratios of CasX mRNA and gRNA were diluted in PNI Formulation Buffer, pH 4.0. GenVoy-ILM™ was diluted 1:1 in anhydrous ethanol. mRNA/gRNA co-formulations were performed using a predetermined N/P ratio. The RNA and lipids were run through a PNI laminar flow cartridge at a predetermined flow rate ratio (RNA:Genvoy-ILM™) on the PNI Ignite™ Benchtop System. After formulation, the LNPs were diluted in PBS, pH 7.4, to decrease the ethanol concentration and increase the pH, which increases the stability of the particles. Buffer exchange of the mRNA/sgRNA-LNPs was achieved by overnight dialysis into PBS, pH 7.4, at 4°C using 10k Slide-A-Lyzer™ Dialysis Cassettes (Thermo Scientific™). Following dialysis, the mRNA/gRNA-LNPs were concentrated to > 0.5 mg/mL using 100 kDa Amicon®-Ultra Centrifugal Filters (Millipore) and then filter-sterilized. Formulated LNPs were analyzed on a Stunner (Unchained Labs) to determine their diameter and polydispersity index (PDI). Encapsulation efficiency and RNA concentration was determined by RiboGreen™ assay using Invitrogen's Quant-iT™ Ribogreen™ RNA assay kit. LNPs were used in various experiments as described herein to deliver CasX mRNA and gRNA to target cells and tissue.

Delivery of LNPs encapsulating CasX mRNA and targeting gRNAs in vitro:

[0420] ~50,000 HepG2 cells, cultured in DMEM/F-12 media containing 10% FBS and 1% PenStrep, were seeded per well in a 96-well plate. The next day, seeded cells were treated with varying concentrations of LNPs, which were prepared in six 2-fold serial dilutions starting at 250 ng. These LNPs were formulated to encapsulate CasX 491 mRNA and a *B2M*-targeting gRNA incorporating either scaffold variant 174 or 316 with spacer 7.9 (v1; see Table 23). Media was changed 24 hours after LNP treatment, and cells were cultured for six additional days prior to harvesting for gDNA extraction for editing assessment at the *B2M* locus by NGS and B2M protein expression analysis via HLA immunostaining, followed by flow cytometry using the Attune NxT flow cytometer. Briefly, for editing assessment,

amplicons were amplified from 200 ng of extracted gDNA with primers targeting the human *B2M* locus and processed by NGS using methods as described below. The results of these assays are shown in FIGS. 17 and 18.

[0421] ~20,000 mouse Hepa1-6 hepatocytes were seeded per well in a 96-well plate. The following day, seeded cells were treated with varying concentrations of LNPs, which were prepared in eight 2-fold serial dilutions starting at 1000 ng. These LNPs were formulated to encapsulate CasX 676 mRNA #2 (see Table 25) and a *ROSA26*-targeting gRNA incorporating scaffold variant 316 with spacer 35.2 (v1 or v5; see Table 23). Media was changed 24 hours post-treatment with LNPs, and cells were cultured for seven additional days prior to harvesting for gDNA extraction for editing assessment at the *ROSA26* locus by NGS. Briefly, amplicons were amplified from extracted gDNA with primers targeting the mouse *ROSA26* locus and processed by NGS using similar methods as described below. The results of this experiment are shown in FIG. 19.

Delivery of LNPs encapsulating CasX mRNA and targeting gRNA in vivo:

[0422] To assess the effects of using v1 and v5 of scaffold 316 in vivo, CasX 676 mRNA #2 (see Table 25) and a *ROSA26*-targeting gRNA using scaffold 316 with spacer 35.2 (v1 or v5; see Table 23) were encapsulated within the same LNP using a 1:1 mass ratio for mRNA:gRNA. Formulated LNPs were buffer-exchanged to PBS for in vivo injection. Briefly, LNPs were administered intravenously through the retro-orbit sinus into 4-week old C57BL/6 mice. Mice were observed for five minutes after injection to ensure recovery from anesthesia before being placed into their home cage. Naïve, uninjected animals served as experimental controls. Six days post-administration, mice were euthanized, and the liver tissue was harvested for gDNA extraction using the Zymo Research Quick DNA/RNA Miniprep kit following the manufacturer's instructions. Target amplicons were then amplified from the extracted gDNA with a set of primers targeting the mouse *ROSA26* locus and processed using similar methods as described below for editing assessment by NGS. The results of this experiment are shown in FIG. 20.

[0423] To compare the effects of using v7, v8, and v9 of scaffold 316 on editing at the *PCSK9* locus *in vivo*, CasX 676 mRNA #1 (see Table 26 for sequences) and a *PCSK9*-targeting gRNA using scaffold 316 with spacer 27.107 (v1, v7, v8, or v9; see Table 23), were encapsulated within the same LNP using a 1:1 mass ratio for mRNA:gRNA for each gRNA. LNPs were administered retro-orbitally into 6-week old C57BL/6 mice, as described above, and mice were euthanized seven days post-injection to harvest liver tissue for gDNA extraction for editing assessment by NGS at the *PCSK9* locus. The results of this experiment

are shown in FIG. 21.

Table 26: Encoding sequences of CasX 676 mRNA #1 molecule

CasX ID	Component (ID)	Description	DNA sequence	SEQ ID NO:
CasX 676 mRNA #1	5'UTR	hHBA	ACTCTTCTGGTCCCCACAGACTCAGAGAGAA CCCGCCACC	3188
	START codon + c-MYC NLS		ATGGCCCCTGCTGCCAAGAGAGTGAAGCTGG ATAGCAGA	3133
	CasX 676		See sequence in Table 19	3134
	c-MYC NLS + STOP codons		GGAAGCCCCGCTGCCAAGAGAGTGAAGCTGG ACTAATAGATAA	3135
	3'UTR	hHBA	GCTGGAGCCTCGGTGGCCATGCTTCTTGCCC CTTGGGCCTCCCCCAGCCCCCTCTCCCTT CCTGCACCCGTACCCCGTGGTCTTTGAATA AAGTCTGAGTGGGCGGCA	3189
	Poly(A) tail		AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	3057

*Components are listed in a 5' to 3' order within the constructs

Results:

Assessing the effects of various chemical modifications on gRNA activity:

[0424] Several studies involving Cas9 have demonstrated that chemical modifications of the gRNA resulted in significantly improved editing activity when delivered with Cas9 mRNA. Following delivery of Cas9 mRNA and gRNA into target cells, unprotected gRNA is susceptible to degradation during the mRNA translation process. Addition of chemical modifications such as 2'O-methyl (2'OMe) groups and phosphorothioate bonds can reduce the susceptibility of the gRNA to cellular RNases, but also have the potential to disrupt folding of the gRNA and its interactions with the CRISPR-Cas protein. Given the lack of structural similarity between CasX and Cas9, as well as their respective gRNAs, appropriate chemical modification profiles must be designed and validated *de novo*. Using published structures of wild-type CasX from *Deltaproteobacteria* (PDB codes 6NY1, 6NY2, and 6NY3) as reference, residues that appeared potentially amenable to modification were selected. However, the published structures were of a wild type CasX ortholog and gRNA distinct from the species used as the basis for the engineered variants presented here, and they also lacked the resolution to confidently determine interactions between protein side-chains and the RNA backbone. These limitations introduced a significant amount of ambiguity into determining which nucleotides might be safely modified. As a result, six profiles of chemical

modifications (denoted as versions) were designed for initial testing, and these six profiles are illustrated in FIGS. 10 and 11. The v1 profile was designed as a simple end-protected structure, where the first and last three nucleotides were modified with 2'OMe and phosphorothioate bonds. In the v2 profile, 3'UUU tail was added to mimic the termination sequence used in cellular transcription systems and to move the modified nucleotides outside of the region of the spacer involved in target recognition. The v3 profile included the end protection as in v1, as well as the addition of 2'OMe modifications at all nucleotides identified to be potentially modifiable based on structural analysis. The v4 profile was modeled based on v3, but with all the modifications in the triplex region removed, as this structure was predicted to be more sensitive to any perturbation of the RNA helical structure and backbone flexibility. The v5 profile maintained chemical modifications in the scaffold stem and extended stem regions, while the v6 profile harbored modifications only in the extended stem. The extended stem is a region that would become fully exposed to solvent in the RNP and is amenable to replacement by other hairpin structures, and therefore presumably relatively insensitive to chemical modifications.

[0425] The minimally modified v1 gRNA was initially assessed against an unmodified gRNA (v0) to determine the potential benefit of such chemical modifications on editing when the gRNA was co-delivered with CasX mRNA to target cells. Modified (v1) and unmodified (v0) *B2M*-targeting gRNAs with spacer 7.37 were co-transfected with CasX mRNA into HepG2 cells, and editing at the *B2M* locus was measured by loss of surface presentation of the B2M-dependent HLA complex, as detected by flow cytometry (FIG. 16). The data demonstrate that use of the v1 gRNA resulted in substantially greater loss of B2M expression compared to the levels seen with v0 gRNA across the various doses, thereby confirming that end modifications of the gRNA increased CasX-mediated editing activity upon delivery of the CasX mRNA and gRNA.

[0426] The broader set of gRNA chemical modification profiles were assessed using *PCSK9*-targeting gRNAs using scaffold variant 235 and spacers 6.7 and 6.8 to determine whether the additional chemical modifications would be able to support the formation of active RNPs. In vitro cleavage assays described above were performed to determine k_{cleave} and fraction competence for these engineered gRNAs harboring the various chemical modification profiles. The results from these in vitro cleavage assays are shown in Table 27. The data demonstrate that gRNAs with the v3 profiles exhibited no activity, an indication that the addition of some chemical modifications significantly interfered with RNP formation or activity. Adding v4 chemical modifications resulted in a reasonable cleavage rate in the

excess RNP condition, but exhibited very low fraction competence. The difference between v3 and v4 modifications confirmed that modifications in the triplex region prevented the formation of any active RNP, either due to the inability of the gRNA to fold properly or a disruption in the gRNA-protein interactions. The reduced fraction competence resulting from appending v4 modifications suggest that while the gRNA was able to successfully assemble with the CasX protein to form a cleavage-competent RNP, a large majority of the gRNA was misfolded, or that the appended chemical modifications reduced the affinity of the gRNA for the CasX protein and impeded the efficiency of RNP formation. Application of the v5 or v6 profiles resulted in competent fractions that were comparable to, but slightly lower than, those obtained for reactions using the v1 and v2 modifications. While the k_{cleave} values were relatively consistent between v5 and v6 gRNAs, both v5 and v6 gRNAs achieved nearly half of the k_{cleave} values for v1 and v2 gRNAs. The reduced k_{cleave} value for v6 gRNA was particularly surprising, given the lack of expected interaction between the gRNA and CasX protein in the modified extended stem. However, for both v5 and v6 gRNAs, it is possible that the reduced flexibility of the gRNA, resulting from the 2'OMe modifications, inhibited structural changes in the RNP required for efficient cleavage, or that the modified initial base-pairs of the hairpin involved in CasX protein interaction had been negatively impacted by the inclusion of the 2'OMe groups.

Table 27: Parameters of cleavage activity assessed for CasX RNPs with the various PCSK9-targeting gRNAs using scaffold 235 and harboring the indicated chemical modification profile, denoted by version number

gRNA (scaffold variant-spacer, version no.)	k_{cleave} (min ⁻¹)	Fraction competence
235-6.7, v1	0.901	0.398
235-6.8, v1	1.36	0.398
235-6.7, v2	0.454	0.386
235-6.8, v2	2.03	0.361
235-6.7, v3	0	0
235-6.8, v3	0	0
235-6.7, v4	0.434	0.031
235-6.8, v4	0.257	0.005

235-6.7, v5	0.506	0.313
235-6.8, v5	0.680	0.388
235-6.7, v6	0.462	0.346
235-6.8, v6	0.715	0.325

[0427] The chemically-modified *PCSK9*-targeting gRNAs based on scaffold 235 were subsequently assessed for editing in a cell-based assay. CasX mRNA and chemically modified *PCSK9*-targeting gRNAs were co-transfected into HepG2 cells using lipofectamine. Editing levels were measured by indel rate at the *PCSK9* locus by NGS and secreted PCSK9 levels by ELISA, and the data are displayed in Table 28. The data demonstrate that use of v3 and v4 gRNAs resulted in minimal editing activity at the *PCSK9* locus, consistent with findings from the biochemical in vitro cleavage assays shown in Table 27. Meanwhile, use of v5 and v6 gRNAs resulted in editing levels, measured by indel rate and PCSK9 secretion, that were slightly lower than the levels attained with use of v1 and v2 gRNAs (Table 28). Specifically, the results show that use of v1 and v2 gRNAs, which harbored end modifications, resulted in ~80-85% editing at the *PCSK9* locus, indicating that adding chemical modifications to the gRNA ends was sufficient to achieve efficient editing with CasX. While the data demonstrate that use of v5 and v6 gRNAs resulted in efficient editing in vitro, near-saturating levels of editing were observed with use of the v1 gRNA in this experiment where a single dose of the gRNA was transfected. As a result, the use of a single dose rendered it challenging to assess clearly the effects of the chemical modifications on editing under guide-limiting conditions. Therefore, profiles v1 and v5 were chosen for further testing, as v1 contains the simplest modification profile, and v5 is the most heavily modified profile whose application demonstrated robust activity in vitro (Tables 27 and 28).

Table 28: Editing levels measured by indel rate at *PCSK9* locus by NGS and secreted *PCSK9* levels by ELISA in HepG2 cells co-transfected with CasX 491 mRNA and various chemically-modified *PCSK9*-targeting gRNAs using scaffold 235 and either spacer 6.7 or 6.8

Experimental condition	Indel rate (edit fraction)		Secreted PCSK9 (ng/mL)	
	Mean	Stdev	Mean	Stdev
CasX mRNA only	0.0021	0.003	52	14
235-6.7, v1	0.83	0.0058	18	5.7
235-6.7, v2	0.79	0.0071	21	4
235-6.7, v3	0.024	0.02	48	19
235-6.7, v4	0.12	0.006	34	5.5
235-6.7, v5	0.73	0.023	21	9
235-6.7, v6	0.75	0.0069	22	8.8
235-6.8, v1	0.85	0.017	16	4.4
235-6.8, v2	0.83	0.0028	20	1.5
235-6.8, v3	0.023	0.0027	39	2.7
235-6.8, v4	0.088	0.0086	42	10
235-6.8, v5	0.77	0.017	19	1.6
235-6.8, v6	0.78	0.014	24	6.9
Non-targeting ctrl	0.0019	0.0026	42	12

[0428] The v1 and v5 profiles were further tested in another cell-based assay to assess their effects on editing efficiency. LNPs were formulated to co-encapsulate CasX 676 mRNA #2 and v1 and v5 chemically-modified *ROSA26*-targeting gRNAs using the newly-designed gRNA scaffold 316 (described further in the following sub-section). The “v5” profile was modified slightly for application to the 316 scaffold. Three 2' OMe modifications in the non-base-paired region immediately 5' of the extended stem were removed to restrict modifications to the two stemloop regions. Hepa1-6 hepatocytes were treated with the resulting LNPs at various doses and harvested eight days post-treatment to assess editing at the *ROSA26* locus, measured as indel rate detected by NGS (FIG. 19). The data demonstrate that treatment with LNPs delivering the v5 *ROSA26*-targeting gRNA resulted in markedly

lower editing levels across the range of doses compared to the levels achieved with the v1 counterpart (FIG. 19). There are several possible explanations for the differences in relative activity observed with use of v5 gRNA in FIG. 19 relative to that observed in Table 28. The first and most likely possible explanation is that the single dose used to achieve editing shown in Table 28 was too high to measure differences in activity accurately between use of v5 gRNA and v1 gRNA. It is also possible that the removal of the modifications outside the stemloop motifs in the 316 version of v5 negatively impacted guide activity. While it is possible that these modifications provide stability benefits that outweigh an activity cost imparted by the stemloop modifications, this seems unlikely given that increasing levels of modification have so far resulted in decreased activity. A final possible explanation is that the modifications in the v5 profile might negatively impact LNP formulation or behavior through differential interactions between the modified nucleotide backbone and the ionizable lipid of the LNP, potentially resulting in less efficient gRNA encapsulation or in less efficient gRNA release following internalization.

[0429] LNPs co-encapsulating the CasX mRNA #2 and v1 and v5 chemically-modified *ROSA26*-targeting gRNAs based on scaffold 316 were further tested *in vivo*. FIG. 20 shows the results of the editing assay as percent editing measured as indel rate at the *ROSA26* locus. The data demonstrate that use of the v5 gRNA resulted in ~5-fold lower editing compared to that achieved with use of the v1 gRNA, under more relevant testing conditions of *in vivo* LNP delivery. These findings support the reduced cleavage rate observed biochemically for the v5 gRNA in Table 27, an indication that the v5 modifications have interfered with some aspect of CasX activity. Given the consistent decrease in activity detected in v5 and v6 profiles (Table 27), the reduced editing may be attributed to modifications in the extended stem region. Although the extended stem of the gRNA has minimal interactions with the CasX protein, it is possible that addition of 2'OMe groups at the first base-pair disrupted either the CasX protein-gRNA interactions or the complex RNA fold where the extended stem meets the pseudoknot and triplex regions. More specifically, inclusion of the 2'OMe groups might have adversely affected the basal base-pairs of the gRNA extended stem and residues R49, K50, and K51 of the CasX protein. Finally, structural studies of CasX have suggested that flexibility of the gRNA is required for efficient DNA cleavage (Liu J, *et al*, CasX enzymes comprise a distinct family of RNA-guided genome editors. *Nature* 566:218-223 (2019); Tsuchida CA, *et al*, Chimeric CRISPR-CasX enzymes and guide RNAs for improved genome editing activity. *Mol Cell* 82(6): 1199-1209 (2022)). Thus, the addition of the 2'OMe groups throughout the extended stem might have enforced a more rigid A-form

helical structure and prevented the needed flexibility for the gRNA for efficient cleavage. Furthermore, it is possible that the additional modifications in the scaffold stem in the v5 and v6 profiles might be detrimental to activity, though this is currently unclear given the limited comparisons between the v5 and v6 profiles.

[0430] Additional modification profiles were designed with the goal of enhancing gRNA stability while mitigating the adverse effects on RNP cleavage activity. Using recently published structures of wild-type CasX from *Planctomycetes* (PDB codes 7WAY, 7WAZ, 7WB0, 7WB1), which has a higher homology to the engineered CasX variants being assessed, additional chemical modification profiles for gRNAs were designed and are illustrated in FIG. 12. These profiles illustrate the addition of 2'OMe groups and phosphorothioate bonds to a newly-designed gRNA scaffold variant, which is described in the ensuing sub-section. These new gRNA chemical modification profiles were designed based on the initial data demonstrating sufficient editing activity observed in Table 28 with use of the v5 gRNA that suggested that modifications to the extended stem and scaffold stem regions would not negatively impact activity. The v7 profile was designed to include 2'OMe at residues likely to be modifiable throughout the gRNA structure, which excluded the triplex region, given the dramatic negative effects of adding such modifications observed earlier with the v3 profile. More conservative profiles, v8 and v9, were also designed, as illustrated in FIG. 12. For the v8 construct, modifications were removed in the pseudoknot and triplex loop region, but were retained in the scaffold stem, extended stem, and their flanking single-stranded regions, in addition to the 5' and 3' termini. For the v9 profile, modifications were removed in the single-stranded regions flanking the stemloops, but were retained in the stemloops themselves, in addition to the pseudoknot, triplex loop, and 5' and 3' termini. The additional chemical modification profiles v7, v8, and v9 of the newly designed gRNA scaffold variant 316 (discussed further below) were assessed in vivo at the *PCSK9* locus. The results of the editing assay in vivo quantified as percent editing at the *PCSK9* locus measured as indel rate as detected NGS are illustrated in FIG. 21. Despite the fact that low editing efficiency was detected overall, the data demonstrate that use of v7, v8, and v9 gRNAs resulted in lower editing levels at the *PCSK9* locus compared to the indel rate achieved with use of the v1 gRNA (FIG. 21). Without wishing to be bound by theory, given the findings in FIGS. 19-20, showing inferior editing activity attained with the v5 gRNA, it is unsurprising that v7, v8, and v9 profiles similarly demonstrated comparatively lower editing activity. As illustrated in FIG. 12, the v7, v8, and v9 profiles include modifications throughout the extended stem region, which might have interfered with RNP activity.

Comparison of gRNA scaffold variant 174 and 316 using an in vitro cleavage assay:

[0431] Previous work had established gRNA scaffold variant 235 as a top-performing scaffold variant across multiple delivery conditions. However, the longer length of scaffold 235 (119 bp, when using a 20 bp spacer) relative to gRNAs including scaffold 174 (109 bp, when using a 20 bp spacer) increased the difficulty of solid-phase RNA synthesis, which could result in increased manufacturing costs, decreased purity and yield, and higher rates of synthesis failures. To address these issues but retain the improved activity of using scaffold variant 235, a chimeric gRNA scaffold was designed primarily on the basis of the scaffold 235 sequence, but the extended stemloop of scaffold 235 was replaced with the shorter extended stemloop of scaffold variant 174 (FIGS. 7-9). The resulting chimeric scaffold, named scaffold 316, was synthesized in parallel with scaffold 174 and *PCSK9*-targeting spacers 6.7 and 6.8, and *B2M*-targeting spacer 7.9 harboring the v1 chemical modification profile, with 2'OMe and phosphorothioate bonds on the first and last three nucleotides of all gRNAs (see Table 23). Scaffold variant 174 was chosen as the comparator rather than variant 235 because variant 174 was the best previously characterized scaffold with a similar length as variant 316.

[0432] In vitro cleavage activity was assessed for gRNAs with scaffold 174 and 316 and spacers 6.7 and 6.8. Cleavage assays were carried out with 20-fold excess RNP over a matching dsDNA target. Cleavage rates were quantified for all four guides, and the results are shown in Table 29. The data demonstrate that in the context of spacer 6.7, use of either scaffold 174 or 316 resulted in similar cleavage rates, with scaffold 316 resulting in marginally faster cleavage than that achieved with scaffold 174. In the context of spacer 6.8, the difference in cleavage activity was more pronounced: CasX RNPs using scaffold 316 were able to cleave DNA nearly twice as quickly as CasX RNPs using scaffold 174 (Table 29).

[0433] Assays were also performed with equimolar amounts of RNP and DNA target over a longer time course to assess the fraction of expected RNP active for cleavage. As the CasX RNP is essentially single-turnover over the tested timescale, and the concentrations used are expected to be substantially higher than the K_D of the DNA-binding reaction, the amount of cleaved DNA should approximate the amount of active RNP. For either spacer 6.7 or 6.8, the active fraction of CasX RNPs incorporating scaffold 316 was 25-30% higher than for CasX RNPs using scaffold 174 (Table 29). These data suggest that a higher fraction of gRNA using scaffold 316 was properly folded for association with the CasX protein, or that the gRNA using scaffold 316 was able to associate more strongly with the CasX protein. Compared to

scaffold 174, scaffold 316 harbors mutations expected to stabilize the pseudoknot and triplex structures required for proper gRNA folding. The increased stability of these motifs in particular, which were more likely to misfold than the simple hairpins found elsewhere in the gRNA structure, might result in a slightly higher fraction of the gRNAs folding into an active conformation.

Table 29: Parameters of cleavage activity assessed for CasX RNPs with gRNAs containing scaffold variant 174 or 316 with the version 1 (v1) chemical modification profile

gRNA (scaffold variant-spacer)	k_{cleave} (min^{-1})	Fraction competence
174-6.7, v1	0.236	0.194
174-6.8, v1	0.142	0.165
316-6.7, v1	0.264	0.244
316-6.8, v1	0.272	0.213

Comparison of gRNA scaffold variant 174 and 316 in a cell-based assay:

[0434] An editing assessment using gRNA scaffold variant 174 compared to variant 316 was performed in a cell-based assay. CasX 491 mRNA and the version 1 (v1) of PCSK9-targeting gRNAs using spacers 6.7 and 6.8 were lipofected into HepG2 cells. Treated cells were harvested 28 hours post-transfection for analysis of editing levels at the PCSK9 locus by NGS and secreted PCSK9 levels by ELISA, and the data are presented in FIG. 15. The data demonstrate that use of any of the PCSK9-targeting gRNA tested resulted in efficient editing at the PCSK9 locus and substantial reduction in PCSK9 secretion compared to the non-targeting control using the B2M-targeting gRNA. The results also show that use of scaffold 316 resulted in more effective editing at the PCSK9 locus than that observed with use of scaffold 174 (~10 percentage point increase in editing rate achieved with scaffold 316 over scaffold 174). This finding is further supported by the ELISA results, such that use of scaffold 316 resulted in more effective reduction of PCSK9 secretion compared to that achieved with use of scaffold 174.

[0435] Scaffold variants 174 and 316 were also assessed in an editing assay where LNPs were formulated to co-encapsulate CasX 491 mRNA and B2M-targeting gRNA harboring either scaffold variant. HepG2 cells were treated with the resulting LNPs at various doses and harvested seven days post-treatment to assess editing at the B2M locus, measured as indel

rate detected by NGS (FIG. 17) and loss of surface presentation of the B2M-dependent HLA complex, as detected by flow cytometry (FIG. 18). The results from both assays demonstrate that treatment with LNPs to deliver the B2M-targeting gRNA using scaffold 316 resulted in higher editing potency at the B2M locus compared to LNPs delivering the gRNA using scaffold 174 at each dose (FIGS. 17 and 18). Specifically, at the highest dose of 250 ng, use of scaffold 316 resulted in an editing level that was nearly two-fold higher than the level attained with using scaffold 174. This substantial increase in editing efficacy when using scaffold 316 versus scaffold 174, compared to the comparatively modest difference in activity observed from the *in vitro* cleavage assays, might be attributed to the destabilization of gRNA structure and folding during LNP formulation. The low pH conditions and association of cationic lipids during LNP formulation could adversely affect parts of the gRNA structure and result in unfolding. Consequently, it would be necessary for the gRNA to refold quickly in the cytoplasm upon delivery, both to bind the CasX protein to form the RNP and to evade RNase degradation. The stability-increasing mutations in scaffold 316 compared to scaffold 174 might provide a substantial benefit in supporting proper gRNA refolding in the cytoplasm after LNP delivery, while the deliberate folding protocol carried out for the gRNA prior to biochemical experiments likely reduced the impact of these mutations.

Example 4: Design and assessment of codon-optimized CasX mRNA on editing efficiency when delivered together with targeting gRNAs *in vitro*

[0436] mRNA sequence and associated modifications may have a significant impact on the efficacy of mRNA-based delivery. Modified nucleotides, including those that encode the 5' cap structure, are important determinants of mRNA stability, translatability, and immunogenicity. Here, for all designed and tested CasX mRNAs, a "Cap 1" structure was used, which included a 5' m⁷G in a 5'-5' triphosphate linkage to an initiating nucleotide with a 2'OMe modification. This structure, similar to the "Cap 0" structure lacking the 2'OMe modification, promotes efficient translation, and has reduced immunogenicity compared to the "Cap 0" structure. Furthermore, the use of modified nucleobases is expected to reduce immunogenicity of the mRNA. Here, the N1-methyl-pseudouridine was used to substitute the uridine ribonucleoside for all *in vitro* transcription reactions, since published studies have demonstrated that the N1-methyl-pseudouridine substantially enhances mRNA performance and reduces mRNA immunogenicity. The modifications are expected to result in reduced immunogenicity and higher translation rates *in vivo*, potentially by avoiding activation of

RIG-I, a primary cytosolic sensor for double-stranded RNA, which is a common contaminant of *in vitro* transcribed mRNA.

[0437] Optimization of the poly(A) tail will also be explored. The poly(A) tail is required for translation and mRNA stability, with longer tails being associated with a longer mRNA half-life. Polyadenylation can be carried out post-transcriptionally with a poly(A) polymerase, but this results in variable tail lengths and adds a step to the mRNA production process. mRNA productions were conducted using plasmids containing a template 80A-tail, terminating with a Type IIS restriction site to allow for run-off transcription, as constructs with plasmids containing a template 120A-tail were unstable during propagation in *E. coli*, often resulting in clones with significant reductions in tail length. Alternate plasmids were also cloned with a SphI restriction site between two-60A stretches, since published studies have demonstrated that similar constructs were more stable during subcloning and amplification in *E. coli* and produced mRNA with equivalent activity in mammalian cells. These alternate versions will be compared for activity using an *in vitro* editing assay across a range of CasX mRNA and gRNA doses to determine the consequential effects on editing activity. The sequences of the poly(A) tails described herein are listed in Table 30.

[0438] The sequences encoding the 5' and 3' UTRs, as well as the codons used for the CasX protein-coding sequences, contribute to the effectiveness of translation. UTRs were selected from annotated human gene transcripts based on genes (e.g., those encoding the α -globin, β -globin proteins) previously characterized to have high mRNA stability, as well as genes expected or previously demonstrated to be particularly well-expressed in the liver (i.e., genes encoding for the following proteins: albumin, complement 3, and cytochrome P450 2E1). The sequences of the 5' and 3' UTRs from these various genes are listed in Table 30. For the 3' UTR, concatenations of individual 3' UTRs were also tested. These constructs were cloned into plasmids containing a T7 promoter, CasX variant 515 or 676, and a poly(A) tail. To isolate the effects of individual 5' and 3' UTRs, each UTR was cloned into a construct that contained either the 3' or 5' α -globin UTR, respectively. IVTs will be performed and purified by binding to poly(dT) beads to capture full-length transcripts. The resulting mRNAs will initially be assessed by co-transfection with a *B2M*-targeting gRNA into HepG2 cells using a range of doses. Editing efficiency will be determined by HLA-immunostaining and flow cytometry. Six days post-transfection, cells were harvested for B2M protein expression analysis via immunostaining of the B2M-dependent HLA protein, followed by flow cytometry using the Attune™ NxT flow cytometer. The best-performing

individual UTRs will be combined into various configurations, formulated into LNP, and tested in primary human hepatocytes and in mice.

[0439] Alternate codon optimizations are also being explored. In addition to the CasX codon optimization used for other delivery modalities, new versions were designed by building a codon usage table based on ribosomal protein codon usage and rebalancing CasX codon usage to match. In addition to potential improvements to the translation rate, this also effectively results in depletion of uracil bases, which may reduce immunogenicity. This codon optimization was also used for production of mRNAs. Additional codon usages have been designed using a variety of available codon optimization tools, adjusting settings as needed to achieve a range of GC content levels. These codon optimizations will be tested under a similar experimental design used for testing UTRs as described above, and the leading codon-optimized CasX candidates will be combined with leading UTR candidates to generate new CasX leads for further validation.

Table 30: List of encoding DNA sequences for the indicated elements used for the generation and optimization of CasX mRNA

Description	Encoding Sequence	SEQ ID NO:
Poly(A) tails		
A ₈₀	AA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	3057
A ₁₂₀	AA AA AAAAAAAAAAAAAAAAAAAAA	3198
A ₆₀ SphIA ₆₀	AA AAAAAAAAGCATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	3199
5' UTRs		
α-globin	ACTCTTCTGGTCCCCACAGACTCAGAGAGAACCC	3200
β-globin	ACATTGTCTTCTGACACAACCTGTGTTCACTAGCAACCTCAAACA	3201
Albumin	CTAGCTTTTCTCTTCTGTCAACCCCACACGCCTTT	3202
Cytochrome P450 2E1 (CYP2E1)	CTCCCGGGCTGGCAGCAGGGCCCCAGC	3203
Complement 3 (C3)	ACTCCTCCCCATCCTCTCCCTCTGTCCCTCTGTCCCTCTGACCCTGCACTGTCCC	3204

3' UTRs		
α -globin	GCTGGAGCCTCGGTGGCCATGCTTCTTGCCCCTTGGGCCTCCCCCAGCCC CTCCTCCCCTTCCTGCACCCGTACCCCGTGGTCTTTGAATAAAGTCTGAG TGGGCGGCA	3189
Albumin	CATCACATTTAAAAGCATCTCAGCCTACCATGAGAATAAGAGAAAGAAAAT GAAGATCAAAAGCTTATTCATCTGTTTTCTTTTCGTTGGTGTAAGCCA ACACCCGTGTCTAAAAACATAAATTTCTTTAATCATTTTGCCTCTTTTCTC TGTGCTTCAATTAATAAAAAATGGAAAGAATCTAATAGAGTGGTACAGCAC TGTTATTTTTCAAAGATGTGTTGCTATCCTGAAAATTCTGTAGGTTCTGTG GAAGTTCAGTGTTCTCTTATTCACCTTCGGTAGAGGATTTCTAGTTTC TTGTGGGCTAATTAATAAATCATTAATACTCTTCTAAGTTATGGATTATA AACATTCAAAATAATATTTTGACATTATGATAATTCTGAATAAAGAACAA AAACCA	3205
Albumin (truncated)	GCATCACATTTAAAAGCATCTCAGCCTACCATGAGAATAAGAGAAAGAAAA TGAAGATCAATAGCTTATTCATCTCTTTTTCTTTTCGTTGGTGTAAGCC AACACCCGTGTCTAAAAACATAAATTTCTTTAATCATTTTGCCTCTTTTCT CTGTGCTTCAATTAATAAAAAATGGAAAGAACCTAGATCT	3206
β -globin	GCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGGTTCCCTTGTTCCTAAG TCCAATACTAACTGGGGGATATTATGAAGGGCCTTGAGCATCTGGATTC TGCCTAATAAAAAACATTTATTTTCATTGCAA	3207
α -globin + β - globin	GCTGGAGCCTCGGTGGCCATGCTTCTTGCCCCTTGGGCCTCCCCCAGCCC CTCCTCCCCTTCCTGCACCCGTACCCCGTGGTCTTTGAATAAAGTCTGAG TGGGCGGCAGCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGGTTCCCTTG TTCCCTAAGTCCAATACTAACTGGGGGATATTATGAAGGGCCTTGAGCA TCTGGATTCTGCCAATAAAAAACATTTATTTTCATTGCAA	3208
β -globin + α - globin	GCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGGTTCCCTTGTTCCTAAG TCCAATACTAACTGGGGGATATTATGAAGGGCCTTGAGCATCTGGATTC TGCCTAATAAAAAACATTTATTTTCATTGCAAGCTGGAGCCTCGGTGGCCA TGCTTCTTGCCCCTTGGGCCTCCCCCAGCCCTCCTCCCCTTCCTGCACC CGTACCCCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGCA	3209

Example 5: LTRP Delivery of LTRP mRNA and targeting gRNA via LNPs to achieve repression of target locus in vitro

[0440] Experiments are performed to assess whether delivery of lipid nanoparticles (LNPs) encapsulating LTRP mRNA and targeting gRNA induce durable repression of a target endogenous locus in a cell-based assay.

Materials and Methods:

Generation of LTRP mRNAs:

[0441] mRNA encoding LTRP molecules are be generated by IVT, as described earlier in Example 1. Sequences encoding the LTRP molecule are codon-optimized as briefly described

in Example 1. Examples of DNA sequences encoding LTRP mRNA are listed in Table 12 and Table 31, with the corresponding mRNA sequences listed in Table 13 and Table 32.

[0442] Targeting gRNAs (e.g., targeting the endogenous *B2M* locus) are synthesized as described above in Example 2.

LNP formulations:

[0443] To formulate LNPs, equal mass ratios of XR or LTRP mRNA and gRNA are diluted in PNI Formulation Buffer, pH 4.0. GenVoy-ILM™ lipids are diluted 1:1 in anhydrous ethanol. mRNA/gRNA co-formulations are generated using a predetermined N/P ratio. The RNA and lipids are run through a PNI laminar flow cartridge at a predetermined flow rate ratio on the PNI Ignite™ Benchtop System. After formulation, the LNPs are diluted in PBS, pH 7.4, to decrease the ethanol concentration and increase the pH, which increases the stability of the particles. Buffer exchange of the mRNA/sgRNA-LNPs is achieved by overnight dialysis into PBS, pH 7.4, at 4°C using 10k Slide-A-Lyzer™ Dialysis Cassettes (Thermo Scientific™). Following dialysis, the mRNA/gRNA-LNPs is concentrated to > 0.5 mg/mL using 100 kDa Amicon®-Ultra Centrifugal Filters (Millipore) and then filter-sterilized. Formulated LNPs are analyzed on a Stunner (Unchained Labs) to determine their diameter and polydispersity index (PDI). Encapsulation efficiency and RNA concentration is determined by RiboGreen™ assay using Invitrogen's Quant-iT™ RiboGreen™ RNA assay kit.

Delivery of LNPs encapsulating LTRP mRNA and targeting gRNAs into mouse liver Hepa1-6 cells:

[0444] Hepa1-6 cells are seeded in a 96-well plate. The next day, seeded cells are treated with varying concentrations of LNPs, which are prepared in six 2-fold serial dilutions starting at 250 ng. These LNPs are formulated to encapsulate an LTRP mRNA and a *B2M*-targeting gRNA. Media is changed 24 hours after LNP treatment, and cells are cultured before being harvested at multiple timepoints (e.g., 7, 14, 21, 28, and 56 days post-treatment) for gDNA extraction for editing assessment at the *B2M* locus by NGS and for bisulfite sequencing to assess off-target methylation at the *VEGFA* locus.

Bisulfite sequencing to assess LTRP specificity measured by off-target methylation levels at target locus:

[0445] To determine off-target methylation levels at the *B2M* locus, gDNA from harvested cells are extracted using the Zymo Quick-DNA Miniprep Plus kit following the manufacturer's instructions. The extracted gDNA will then subjected to bisulfite conversion using the EZ DNA Methylation™ Kit (Zymo) following the manufacturer's protocol,

converting any non-methylated cytosine into uracil. The resulting bisulfite-treated DNA will subsequently be sequenced using next-generation sequencing (NGS) to determine the levels of off-target methylation at the *B2M* and *VEGFA* loci.

NGS processing and analysis:

[0446] Target amplicons are amplified from 100 ng bisulfite-treated DNA via PCR with a set of primers specific to the bisulfite-converted target locations of interest (human *B2M* and *VEGFA* loci). These gene-specific primers contained an additional sequence at the 5' end to introduce an Illumina™ adapter. Amplified DNA products are purified with the Cytiva Sera-Mag Select DNA cleanup kit. Quality and quantification of the amplicon is assessed using a Fragment Analyzer DNA Analysis kit (Agilent, dsDNA 35-1500 bp). Amplicons are sequenced on the Illumina™ Miseq™ according to the manufacturer's instructions. Raw fastq files from sequencing are processed using Bismark Bisulfite Read Mapper and Methylation caller. PCR amplification of the bisulfite-treated DNA would convert all uracil nucleotides into thymine, and sequencing of the PCR product would determine the rate of cytosine-to-thymine conversion as a readout of the level of potential off-target methylation at the *B2M* and *VEGFA* loci mediated by each LTRP molecule.

[0447] The results from this experiment are expected to show that LTRP mRNA and targeting gRNA can be co-encapsulated within LNPs to be delivered to target cells to induce heritable silencing of a target endogenous locus.

Table 31: Encoding sequences of the LTRP1-ZIM3 and LTRP5-ZIM3 mRNA molecules assessed in experiment #2 of this example*

LTRP ID	Component	DNA sequence	SEQ ID NO:
LTRP 1 - ZIM3-KRA B	5'UTR	GACCGGCCGCCACC	3082
	START codon + NLS + linker	ATGGCCCCAAAGAAGAAGCGGAAGGTCTCTAGAGTTAACGGATCAG GCTCTGGAGGTGGA	3083
	START codon + DNMT3A catalytic domain	ATGAACCATGACCAGGAATTTGACCCCCAAAGGTTTACCCACCTG TGCCAGCTGAGAAGAGGAAGCCCATCCGCGTGCTGTCTCTTTGA TGGGATTGCTACAGGGCTCCTGGTGCTGAAGGACCTGGGCATCCAA GTGGACCGCTACATCGCCTCCGAGGTGTGTGAGGACTCCATCACGG TGGGCATGGTGCGGCACCAGGGAAGATCATGTACGTCGCGGACGT CCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGAC CTGGTGATTGGAGGCAGTCCCTGCAACGACCTCTCCATTGTCAACC CTGCCCCGAAGGGACTTTATGAGGGTACTGGCCGCCTCTTCTTTGA GTTCTACCGCCTCCTGCATGATGCGCGGCCCAAGGAGGGAGATGAT CGCCCCCTTCTTGGCTCTTTGAGAAATGTGGTGGCCATGGGCGTTA GTGACAAGAGGGACATCTCGCGATTCTTGAGTCTAACCCCGTGAT GATTGACGCCAAAGAAGTGTCTGCTGCACACAGGGCCCGTTACTTC TGGGGTAACCTTCTTGGCATGAACAGGCCTTTGGCATCCACTGTGA ATGATAAGCTGGAGCTGCAAGAGTGTCTGGAGCACGGCAGAATAGC CAAGTTCAGCAAAGTGAGGACCATTACCACCAGGTCAAACCTCTATA	3084

LTRP ID	Component	DNA sequence	SEQ ID NO:
		AAGCAGGGCAAAGACCAGCATTTCCTCCCGTCTTCATGAACGAGAAGG AGGACATCCTGTGGTGCACTGAAATGGAAAGGGTGTGGCTTCCC CGTCCACTACACAGACGTGTCCAACATGAGCCGCTTGGCGAGGCAG AGACTGCTGGGCCGGTCGTGGAGCGTGCCGGTCATCCGCCACCTCT TCGCTCCGCTGAAGGAATATTTTGCTTGTGTG	
	Linker	TCTAGCGGCAATAGTAACGCTAACAGCCGCGGGCCGAGCTTCAGCA GCGGCCTGGTGCCGTTAAGCTTGCGCGGCAGCCAT	3085
	DNMT3L interaction domain	ATGGGCCCTATGGAGATATACAAGACAGTGTCTGCATGGAAGAGAC AGCCAGTGCGGGTACTGAGCCTCTTCAGAAACATCGACAAGGTACT AAAGAGTTTGGGCTTCTTGAAAGCGGTTCTGGTTCTGGGGGAGGA ACGCTGAAGTACGTGGAAGATGTCAAAATGTCTGTGAGGAGGGACG TGGAGAAATGGGGCCCCCTTTGACCTGGTGTACGGCTCGACGCAGCC CCTAGGCAGCTCTTGTGATCGCTGTCCCGGCTGGTACATGTTCCAG TTCCACCGGATCCTGCAGTATGCGCTGCCTCGCCAGGAGAGTCAGC GGCCCTTCTTCTGGATATTCATGGACAATCTGCTGCTGACTGAGGA TGACCAAGAGACAACACCCGCTTCCTTCAGACAGAGGCTGTGACC CTCCAGGATGTCCGTGGCAGAGACTACCAGAATGCTATGCGGGTGT GGAGCAACATTCCAGGGCTGAAGAGCAAGCATGCGCCCTGACCCC AAAGGAAGAAGAGTATCTGCAAGCCCAAGTCAGAAGCAGGAGCAAG CTGGACGCCCCGAAAGTTGACCTCCTGGTGAAGAACTGCCTTCTCC CGCTGAGAGAGTACTTCAAGTATTTTCTCAAACTCACTTCCTCT T	3086
	Linker	GGAGGGCCGAGCTCTGGCGCACCCCCACCAAGTGGAGGGTCTCCTG CCGGGTCCCCAACATCTACTGAAGAAGGCACCAGCGAATCCGCAAC GCCCAGTCAGGCCCTGGTACCTCCACAGAACCATCTGAAGGTAGT GCGCCTGGTTCCCCAGCTGGAAGCCCTACTTCCACCGAAGAAGGCA CGTCAACCGAACCAAGTGAAGGATCTGCCCCCTGGGACCAGCACTGA ACCATCTGAG	3087
	Linker + buffer	GGCGGTTCCGGCGGAGGAAGCGCT	3088
	dCasX491	CAAGAGATCAAGAGAATCAACAAGATCAGAAGGAGACTGGTCAAGG ACAGCAACACAAAGAAGGCCGGCAAGACAGGCCCCATGAAAACCT GCTCGTCAGAGTGATGACCCCTGACCTGAGAGAGCGGCTGGAAAAC CTGAGAAAAGAAGCCCGAGAACATCCCTCAGCCTATCAGCAACACCA GCAGGGCCAACCTGAACAAGCTGCTGACCGACTACACCGAGATGAA GAAAGCCATCCTGCACGTGTACTGGGAAGAGTTCCAGAAAGACCCC GTGGGCCCTGATGAGCAGAGTTGCTCAGCCTGCCAGCAAGAAGATCG ACCAGAACAAGCTGAAGCCCGAGATGGACGAGAAGGGCAATCTGAC CACAGCCGGCTTTGCCTGCTCTCAGTGTGGCCAGCCTCTGTTCTGTG TACAAGCTGGAACAGGTGTCCGAGAAAGGCAAGGCCTACACCAACT ACTTCGGCAGATGTAACGTGGCCGAGCACGAGAAGCTGATTCTGCT GGCCCAGCTGAAACCTGAGAAGGACTCTGATGAGGCCGTGACCTAC AGCCTGGGCAAGTTTGGACAGAGAGCCCTGGACTTCTACAGCATCC ACGTGACCAAGAAAGCACACACCCCGTGAAGCCCTGGCTCAGAT CGCCGGCAATAGATACGCTCTGGACCTGTGGGCAAGCCCTGTCC GATGCCTGCATGGGAACAATCGCCAGCTTCCTGAGCAAGTACCAGG ACATCATCATCGAGCACCAGAAGGTGGTCAAGGGCAACCAGAAGAG ACTGGAAAGCCTGAGGGAGCTGGCCGGCAAGAGAACCTGGAATAC CCCAGCGTGACCCTGCCTCCTCAGCCTCACACAAAAGAAGGCGTGG ACGCCTACAACGAAGTGATCGCCAGAGTGAGAATGTGGGTCAACCT GAACCTGTGGCAGAAGCTGAAACTGTCCAGGGACGACGCAAGCCT CTGCTGAGACTGAAGGGCTTCCCTAGCTTCCCTCTGGTGGAAAGAC AGGCCAATGAAGTGGATTGGTGGGACATGGTCTGCAACGTGAAGAA GCTGATCAACGAGAAGAAAGAGGATGGCAAGGTTTCTGGCAGAAC CTGGCCGGCTACAAGAGACAAGAAGCCCTGAGGCCTTACCTGAGCA GCGAAGAGGACCGGAAGAAGGGCAAGAAGTTCCGCCAGATACCAGCT GGGCGACCTGCTGCTGCACCTGGAAGAAGACGCGCGAGGACTGG GGCAAAGTGATGATGAGGCCTGGGAGAGAATCGACAAGAAGGTGG	3089

LTRP ID	Component	DNA sequence	SEQ ID NO:
		AAGGCCTGAGCAAGCACATTAAGCTGGAAGAGGAAAGAAGGAGCGA GGACGCCCAATCTAAAGCCGCTCTGACCGATTGGCTGAGAGCCAAG GCCAGCTTTGTGATCGAGGGCCTGAAAGAGGCCGACAAGGACGAGT TCTGCAGATGCGAGCTGAAGCTGCAGAAGTGGTACGGCGATCTGAG AGGCAAGCCCTTCGCCATTGAGGCCGAGAACAGCATCCTGGACATC AGCGGCTTCAGCAAGCAGTACAACCTGCGCCTTCATTTGGCAGAAAG ACGGCGTCAAGAACTGAACCTGTACCTGATCATCAATTACTTCAA AGGCGGCAAGCTGCGGTTCAAGAAGATCAAACCCGAGGCCTTCGAG GCTAACAGATTCTACACCGTGATCAACAAAAGTCCGGCGAGATCG TGCCCATGGAAGTGAACCTCAACTTCGACGACCCCAACCTGATTAT CCTGCCTCTGGCCTTCGGCAAGAGACAGGGCAGAGAGTTTATCTGG AACGATCTGCTGAGCCTGGAAACCGGCTCTCTGAAGCTGGCCAATG GCAGAGTGATCGAGAAAACCTGTACAACAGGAGAACCAGACAGGA CGAGCCTGCTCTGTTTGTGGCCCTGACCTTCGAGAGAAGAGAGGTG CTGGACAGCAGCAACATCAAGCCCATGAACCTGATCGGCGTGGCCC GGGGCGAGAATATCCCTGCTGTGATCGCCCTGACAGACCTGAAGG ATGCCCACTGAGCAGATTCAAGGACTCCCTGGGCAACCTACACAC ATCCTGAGAATCGGCGAGAGCTACAAAGAGAAGCAGAGGACAATCC AGGCCAAGAAAGAGGTGGAACAGAGAAGAGCCGGCGGATACTCTAG GAAGTACGCCAGCAAGGCCAAGAATCTGGCCGACGACATGGTCCGA AACACCGCCAGAGATCTGCTGTACTACGCCGTGACACAGGACGCCA TGCTGATCTTCGCGAATCTGAGCAGAGGCTTCGGCCGGCAGGGCAA GAGAACCTTTATGGCCGAGAGGCAGTACACCAGAATGGAAGATTGG CTCACAGCTAACTGGCCTACGAGGGACTGAGCAAGACCTACCTGT CCAAAACACTGGCCAGTATACCTCCAAGACCTGCAGCAATTGCGG CTTCACCATCACCAGCGCCGACTACGACAGAGTGCTGGAAGAGCTC AAGAAAACCGCCACCGGCTGGATGACCACCATCAACGGCAAAGAGC TGAAGGTTGAGGGCCAGATCACCTACTACAACAGGTACAAGAGGCA GAACGTCGTGAAGGATCTGAGCGTGGAACTGGACAGACTGAGCGAA GAGAGCGTGAACAACGACATCAGCAGCTGGACAAAGGGCAGATCAG GCGAGGCTCTGAGCCTGCTGAAGAAGAGGTTTAGCCACAGACCTGT GCAAGAGAAGTTCTGTGCTGAACCTGCGGCTTCGAGACACACGCC GCTGAACAGGCTGCCCTGAACATTGCCAGAAGCTGGCTGTTCTCTGA GAAGCCAAGAGTACAAGAAGTACCAGACCAACAAGACCACCGGCAA CACCGACAAGAGGGCCTTTGTGGAACCTGGCAGAGCTTCTACAGA AAAAAGCTGAAAGAAGTCTGGAAGCCCGCCGTG	
	Linker + buffer	CGATCGGGCGGTTCCGGCGGAGGTTCCACTAGT	3090
	ZIM3-KRAB	ATGAACAATTCCAGGGAAGAGTGACCTTCGAGGATGTCACTGTGA ACTTCACCCAGGGGGAGTGGCAGCGGCTGAATCCCGAACAGAGAAA CTTGTACAGGGATGTGATGCTGGAGAATTACAGCAACCTTGTCTCT GTGGGACAAGGGGAAACCACCAACCCGATGTGATCTTGAGGTTGG AACAAGGAAAGGAGCCATGGTTGGAGGAAGAGGAAGTGTGGGAAG TGGCCGTGCAGAAAAAATGGGGACATTGGAGGGCAGATTGGAAG CCAAAGGATGTGAAAGAGAGTCTC	3091
	Buffer + NLS	ACTAGTCCAAAAAAGAAGAGAAAGGTA	3092
	Tag	GATTACAAAGATGACGATGACAAAGACTACAAGGATGATGATGATA AG	3093
	Buffer	GGATCCGGCTG	3094
	Poly(A) tail	AA AA AA	3095
LTRP 5 - ZIM3-	5'UTR	GACCGCCGCCACC	3082
	START codon +	ATGGCCCCAAGAAGAAGCGGAAGGTCTCTAGA	3096

LTRP ID	Component	DNA sequence	SEQ ID NO:
KRA B	NLS + buffer		
	START codon + DNMT3A catalytic domain	ATGAACCATGACCAGGAATTTGACCCCCCAAAGGTTTACCCACCTG TGCCAGCTGAGAAGAGGAAGCCCATCCGCGTGCTGTCTCTTTGA TGGGATTGCTACAGGGCTCCTGGTGCTGAAGGACCTGGGCATCCAA GTGGACCGCTACATCGCCTCCGAGGTGTGTGAGGACTCCATCACGG TGGGCATGGTGCGGCACCAGGGAAGATCATGTACGTCGGGGACGT CCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGAC CTGGTGATTGGAGGCAGTCCCTGCAACGACCTCTCCATTGTCAACC CTGCCCCGAAGGGACTTTATGAGGGTACTGGCCGCTCTTCTTTGA GTTCTACCGCCTCCTGCATGATGCGCGGCCCAAGGAGGGAGATGAT CGCCCCCTTCTTGCTCTTTGAGAATGTGGTGCCATGGGCGTTA GTGACAAGAGGGACATCTCGCGATTTCTTGAGTCTAACCCCGTGAT GATTGACGCCAAAGAAGTGTCTGCTGCACACAGGGCCCGTTACTTC TGGGGTAACCTTCTTGCCATGAACAGGCCTTTGGCATCCACTGTGA ATGATAAGCTGGAGCTGCAAGAGTGTCTGGAGCACGGCAGAATAGC CAAGTTCAGCAAAGTGAAGACCATTAACACCAGGTCAAACCTATA AAGCAGGCAAGAGACCAGCATTTCCCGTCTTCATGAACGAGAAGG AGGACATCCTGTGGTGCACTGAAATGAAAGGGTGTGGCTTCCC CGTCCACTACACAGACGTGTCCAACATGAGCCGCTTGGCGAGGCAG AGACTGCTGGGCCGGTCTGGAGCGTGCCGGTCATCCGCCACCTCT TCGCTCCGCTGAAGGAATATTTTGCTTGTGTG	3084
	Linker	TCTAGCGGCAATAGTAACGCTAACAGCCGCGGGCCGAGCTTCAGCA GCGGCCCTGGTGCCGTTAAGCTTGCGCGGCAGCCAT	3085
	DNMT3L interaction domain	ATGGGGCCCTATGGAGATATACAAGACAGTGTCTGCATGGAAGAGAC AGCCAGTGCGGGTACTGAGCCTCTTCAGAAACATCGACAAGGTACT AAAGAGTTTGGGCTTCTTGGAAGCGGTTCTGGTTCTGGGGGAGGA ACGCTGAAGTACGTGGAAGATGTCACAAATGTCGTGAGGAGGGACG TGGAGAAATGGGGCCCTTTGACCTGGTGTACGGCTCGACGCAGCC CCTAGGCAGCTCTTGTGATCGCTGTCCCGGCTGGTACATGTTCCAG TTCCACCGGATCCTGCAGTATGCGCTGCCTCGCCAGGAGAGTCAGC GGCCCTTCTTCTGGATATTATGGACAATCTGCTGCTGACTGAGGA TGACCAAGAGACAACCTACCCGCTTCCTTCAGACAGAGGCTGTGACC CTCCAGGATGTCCGTGGCAGAGACTACCAGAATGCTATGCGGGTGT GGAGCAACATTCCAGGGCTGAAGAGCAAGCATGCGCCCTGACCCC AAAGGAAGAGAGTATCTGCAAGCCCAAGTCAGAAGCAGGAGCAAG CTGGACGCCCCGAAAGTTGACCTCCTGGTGAAGAACTGCCTTCTCC CGCTGAGAGAGTACTTCAAGTATTTTCTCAAACTCACTTCCTCT T	3086
	Linker	GGCGGTTCCGGCGGAGGA	3097
	ZIM3-KRAB	ATGAACAATTTCCAGGGAAGAGTGACCTTCGAGGATGTCACTGTGA ACTTCACCCAGGGGGAGTGGCAGCGGCTGAATCCCGAACAGAGAAA CTTGATACAGGGATGTGATGCTGGAGAATTACAGCAACCTTGTCTCT GTGGGACAAGGGGAAACCACCAACCCGATGTGATCTTGAGGTTGG AACAAGGAAAGGAGCCATGGTTGGAGGAAGAGGAAGTGTGGGAAG TGGCCGTGCAGAAAAAATGGGGACATTGGAGGGCAGATTTGGAAG CCAAAGGATGTGAAAGAGAGTCTC	3091
	Linker	GGAGGGCCGAGCTCTGGCGCACCCCCACCAAGTGGAGGGTCTCCTG CCGGGTCCCCAACATCTACTGAAGAAGGCACCAGCGAATCCGCAAC GCCCAGTCAAGGCCCTGGTACCTCCACAGAACCATCTGAAGGTAGT GCGCCTGGTTCCCCAGCTGGAAGCCCTACTTCACCGAAGAAGGCA CGTCAACCGAACCAAGTGAAGGATCTGCCCCTGGGACCAGCACTGA ACCATCTGAG	3087
	dCasX491	CAAGAGATCAAGAGAATCAACAAGATCAGAAGGAGACTGGTCAAGG ACAGCAACACAAAGAAGCGCGCAAGACAGGCCCATGAAAACCT GCTCGTCAGAGTGATGACCCCTGACCTGAGAGAGCGGCTGGAAAC CTGAGAAAGAAGCCCGAGAACATCCCTCAGCCTATCAGCAACACCA GCAGGGCCAACTGAACAAGCTGCTGACCGACTACACCGAGATGAA	3089

LTRP ID	Component	DNA sequence	SEQ ID NO:
		GAAAGCCATCCTGCACGTGTAAGTGGGAAGAGTTCCAGAAAGACCCC GTGGGCCTGATGAGCAGAGTTGCTCAGCCTGCCAGCAAGAAGATCG ACGAGAACAGCTGAAGCCCGAGATGGACGAGAAGGGCAATCTGAC CACAGCCGGCTTTGCCTGCTCTCAGTGTGGCCAGCCTCTGTTCTGTG TACAAGCTGGAACAGGTGTCCGAGAAAGGCAAGGCCTACACCAACT ACTTCGGCAGATGTAACGTGGCCGAGCAGCAGAGAAGCTGATTCTGCT GGCCAGCTGAAACCTGAGAAGGACTCTGATGAGGCCGTGACCTAC AGCCTGGGCAAGTTTGGACAGAGAGCCCTGGACTTCTACAGCATCC ACGTGACCAAAGAAAGCACACACCCCGTGAAGCCCTGGCTCAGAT CGCCGGCAATAGATACGCTCTGGACCTGTGGGCAAGCCCTGTCC GATGCCCTGCATGGGAACAATCGCCAGCTTCTTGAGCAAGTACCAGG ACATCATCATCGAGCACCAGAAGGTGGTCAAGGGCAACCAGAAGAG ACTGGAAGCCTGAGGGAGCTGGCCGGCAAAGAGAACCTGGAATAC CCCAGCGTGACCCTGCCTCCTCAGCCTCACACAAAGAAGGCCTGG ACGCCTACAACGAAGTGATCGCCAGAGTGAGAATGTGGGTCAACCT GAACCTGTGGCAGAAGCTGAACTGTCCAGGGACGACGCCAAGCCT CTGCTGAGACTGAAGGGCTTCCCTAGCTTCCCTCTGGTGGAAAGAC AGGCCAATGAAGTGGATTGGTGGGACATGGTCTGCAACGTGAAGAA GCTGATCAACGAGAAGAAAGAGGATGGCAAGGTTTTCTGGCAGAAC CTGGCCGGCTACAAGAGACAAGAAGCCCTGAGGCCCTACCTGAGCA GCGAAGAGGACCGGAAGAAGGGCAAGAAGTTCGCCAGATACCAGCT GGGCGACCTGCTGCTGCACCTGAAAAAGAAGCACGGCGAGGACTGG GGCAAAGTGATCGATGAGGCCTGGGAGAGAATCGACAAGAAGGTGG AAGGCCTGAGCAAGCACATTAAGCTGGAAGAGGAAAGAAGGAGCGA GGACGCCCCAATCTAAAGCCGCTCTGACCGATTGGCTGAGAGCCAAG GCCAGCTTTGTGATCGAGGGCCTGAAAGAGGCCGACAAGGACGAGT TCTGCAGATGCGAGCTGAAGCTGCAGAAGTGGTACGGCGATCTGAG AGGCAAGCCCTTCGCCATTGAGGCCGAGAACAGCATCCTGGACATC AGCGGCTTCAGCAAGCAGTACAACCTGCGCCTTCATTTGGCAGAAAG ACGGCGTCAAGAACTGAACCTGTACCTGATCATCAATTACTTCAA AGGCGGCAAGCTGCGGTTCAAGAAGATCAAACCCGAGGCCTTCGAG GCTAACAGATTCTACACCGTGATCAACAAAAAGTCCGGCGAGATCG TGCCCATGGAAGTGAACCTCAACTTCGACGACCCCAACCTGATTAT CCTGCCCTTGGCCTTCGGCAAGAGACAGGGCAGAGAGTTTACCTGG AACGATCTGCTGAGCCTGGAAACCGCTCTCTGAAGCTGGCCATG GCAGAGTGATCGAGAAAACCTGTACAACAGGAGAACCAGACAGGA CGAGCCTGCTCTGTTTGTGGCCCTGACCTTCGAGAGAAGAGAGTG CTGGACAGCAGCAACATCAAGCCCATGAACCTGATCGGCGTGGCCC GGGGCGAGAATATCCCTGCTGTGATCGCCCTGACAGACCCTGAAGG ATGCCCACTGAGCAGATTCAAGGACTCCCTGGGCAACCCACACAC ATCCTGAGAATCGGCGAGAGCTACAAAGAGAAGCAGAGGACAATCC AGGCCAAGAAAGAGGTGGAACAGAGAAGAGCCGGCGGATACTCTAG GAAGTACGCCAGCAAGGCCAAGAATCTGGCCGACGACATGGTCCGA AACACCGCCAGAGATCTGCTGTACTACGCCGTGACACAGGACGCCA TGCTGATCTTCGCGAATCTGAGCAGAGGCTTCGGCCGGCAGGGCAA GAGAACCTTTATGGCCGAGAGGCAGTACACCAGAATGGAAGATTGG CTCACAGCTAACTGGCCTACGAGGGACTGAGCAAGACCTACCTGT CCAAAACACTGGCCAGTATACCTCCAAGACCTGCAGCAATTGCGG CTTCACCATCACCAGCGCCGACTACGACAGAGTGCTGGAAAAGCTC AAGAAAACCGCCACCGGCTGGATGACCACCATCAACGGCAAAGAGC TGAAGGTTGAGGGCCAGATCACCTACTACAACAGGTACAAGAGGCA GAACGTCGTGAAGGATCTGAGCGTGGAAGTGGACAGACTGAGCGAA GAGAGCGTGAACAACGACATCAGCAGCTGGACAAAGGGCAGATCAG GCGAGGCTCTGAGCCTGCTGAAGAAGAGGTTTAGCCACAGACCTGT GCAAGAGAAGTTCGTGTGCCTGAAGTTCGGGCTTCGAGACACACGCC GCTGAACAGGCTGCCCTGAACATTGCCAGAAGCTGGCTGTTCTCTGA GAAGCCAAGAGTACAAGAAGTACCAGACCAACAAGACCACCGGCAA CACCGACAAGAGGGCCTTTGTGGAAACCTGGCAGAGCTTCTACAGA AAAAAGCTGAAAGAAGTCTGGAAGCCCGCCGTG	

[illegible]

*Components are listed in a 5' to 3' order within the constructs

Table 32: Full-length RNA sequences of LTRP1-ZIM3 and LTRP5-ZIM3 mRNA molecules assessed in experiment #2 of this example. Modification ‘mψ’ = N1-methylpseudouridine

LTRP ID	RNA sequence	SEQ ID NO:
LTRP1 - ZIM3- KRAB	GACCGGCCGCCACCAmψGGCCCCAAAGAAGAAGCGGAAGGmψCmψCmψAGAGmψmψAACGGAmψCAGGCGmψCmψGGAGGmψGGAAmψGAACCAmψGACCAGGAAmψmψmψGACCCC CCAAAGGmψmψmψACCCACCGmψGmψGCCAGCGmψGAGAAGAGGAAGCCCCAmψCCGCGmψ GCGmψGmψCmψCmψCmψmψmψGAmψGGGAmψmψGmψACAGGGCGmψCCmψGGmψGCGmψG AAGGACCGmψGGGCAmψCCAAAGmψGGACCGCGmψACAmψCGCCmψCCGAGCGmψGmψGmψG AGGACmψCCAmψCACGGmψGGGCAmψGGmψGCGGCACCCAGGGAAAGAmψCAmψGmψAC GmψCGGGGACGmψCCGACGCGmψCACACAGAAGCAmψAmψCCAGGAGmψGGGGCCCAm ψmψCGACCGmψGGmψGAmψmψGGAGGCAGmψCCCGmψGCAACGACCGmψCmψCCAmψmψGm ψCAACCCmψGCCCCGAAGGGACmψmψmψAmψGAGGGmψACmψGGCCGCCCGmψCmψmψCmψ ψmψmψGAGmψCmψACCGCGCGmψCCCGmψCCmψGCAmψGAmψGCGCGGCCCAAGGAGGGAGmψ GAmψCGCCCCmψmψCmψmψCmψGGCGmψCmψmψmψGAGAAmψGmψGGmψGGCCAmψGGG CGmψmψAGmψGACAAGAGGGACAmψCmψCGCGAmψmψmψCmψmψGAGmψCmψAACCCC GmψGAmψGAmψmψGACGCCAAAGAAGmψGmψCmψGCGmψGCACACAGGGCCCCGmψmψAC mψmψCmψGGGGmψAACCGmψCmψCGCGAmψGAACAGGCCCGmψmψmψGGCGAmψCCACmψ GmψGAAmψGAmψAAGCGmψGGAGCGmψGCAAGAGmψGmψCmψGGAGCACGGCAGAAmψAG CCAAAGmψmψCAGCAAAAGmψGAGGACCAmψmψACCAAGCGmψCAAACTmψCmψAmψAAA GCAGGGCAAAGACCAGCAmψmψmψCCCCGmψCmψmψCAmψGAACGAGAAGGAGGACAm ψCCmψGmψGGmψGCACmψGAAAmψGGAAAGGGmψGmψmψmψGGCGmψCCCCGmψCCA CmψACACAGACGmψGmψCCAACAmψGAGCCCGmψmψGGCGAGGCAGAGCAmψGCGmψGG GCCGGmψCGmψGGAGCGmψGCCGGmψCAmψCCGCCACCGmψCmψmψCGCGmψCCGCGmψGA AGGAAmψAmψmψmψmψGCGmψmψGmψGmψGmψGmψACGCGCAAmψAGmψAACCGmψAAC AGCCGCGGGCCGAGCGmψmψCAGCAGCGGCCmψGGmψGCCGCGmψmψAAGCGmψmψGCGCGG CAGCCAmψAmψGGGCCCGmψAmψGGAGAmψAmψACAAGACAGmψGmψCmψGCAmψGGAA GAGACAGCCAGmψGCGGGGmψACmψGAGCGmψCmψmψCAGAAACAmψCGACAAAGGmψAC mψAAAGAGmψmψmψGGGCGmψmψCmψmψGGAAAGCGGGmψmψCmψGGmψmψCmψGGGGGA GGAACGCGmψGAAGmψACGmψGGAAGAmψGmψCACAAAmψGmψCGmψGAGGAGGGACGm ψGGAGCAAmψGGGGCCCGmψmψmψGACCGmψGGmψmψACGGCGmψCGACGCGGCCCGmψ AGGCAGCGmψCmψmψGmψGAmψCGCGmψGmψCCCCGGCGmψGGmψACAmψGmψmψCCAGmψ ψCCACCGGAmψCCmψGCAGmψAmψGCGCGmψGCCCGmψCGCCAGGAGAGmψCAGCGGCCCGm ψmψCmψmψCmψGGAmψAmψmψCmψGGACAAmψCmψGCGmψGCGmψGACmψGAGGAmψGA CCAAGAGACAACTmψACCCCGCGmψCCmψmψCAGACAGAGGCCCGmψGmψGACCCmψCCAGG AmψGmψCCGmψGGCAGAGCAmψACCAGAAmψGCGmψAmψGCGGGmψGmψGGAGCAACAm ψmψCCAGGGCGmψGAAGAGCAAGCAmψGCGCCCCCGmψGACCCCAAAGGAAGAGAmψAm ψCmψGCAAGCCCCAAGmψCAGAAGCAGGAGCAAGCGmψGGACGCCCCGAAAGmψmψGACC mψCCmψGGmψGAAGAACTmψGCCmψmψCmψCCCCGCGmψGAGAGAGmψACmψmψCAAGmψA mψmψmψmψmψCmψCAAACTmψCACmψmψCCmψCmψmψGGAGGGCCGAGCGmψCmψGGCG CACCCCCACCAAGmψGGAGGGCGmψCmψCCmψGCCGGCGmψCCCCAACAmψCmψACmψGAA GAAGGCACCGACGGAAmψCCGCAACGCCCGAGmψCAGGCCCGmψGGmψACCGmψCCAGAGA	3099

LTRP ID	RNA sequence	SEQ ID NO:
	ACCAmlyCmlyGAAGGmlyAGmlyGCGCCmlyGGmlymlyCCCCAGCmlyGGAAGCCcmlyACmlymly CCACCGAAGAAGGCACGmlyCAACCGAACCAAGmlyGAAGGAmlyCmlyGCCCCmlyGGGACC AGCAmlyGAACCAmlyCmlyGAGGGCGGmlymlyCCGGCGGAGGAAGCGCmlyCAAGAGAmlyC AAGAGAAmlyCAACAAGAmlyCAGAAGGAGACmlyGGmlyCAAGGACAGCAACACAAAGAAG GCCGGCAAGACAGGCCCCAmlyGAAAACCCmlyGCMlyCGmlyCAGAGmlyGAmlyGACCCcmly GACCmlyGAGAGAGCGGCMlyGGAAAACmlyGAGAAAGAAGCCCGAGAACAmlyCCcmlyCA GCCmlyAmlyCAGCAACACCAGCAGGGCCAACmlyGAACAAGCmlyGCMlyGACCGACmlyAC ACCGAGAmlyGAAGAAAGCCAmlyCCmlyGCACGmlyGmlyACmlyGGGAAGAGmlymlyCCAGAA AGACCCCGmlyGGGCCmlyGAmlyGAGCAGAGmlymlyGCMlyCAGCCmlyGCCAGCAAGAAGAm lyCGACCAAGAACAGCmlyGAAGCCCGAGAmlyGGACGAGAAGGGCAAmlyCmlyGACCACAG CCGGCMlymlymlyGCCmlyGCMlyCmlyCAGmlyGmlyGGCCAGCCmlyCmlyGmlymlyCGmlyGmlyAC AAGCmlyGGAACAGGmlyGmlyCCGAGAAAGGCAAGGCCmlyACACCAACmlyACmlymlyCGGC AGAmlyGmlyAACGmlyGGCCGAGCAGAGAAGCmlyGAmlymlyCmlyGCMlyGGCCAGCmlyGA AACmlyGAGAAGGACmlyCmlyGAmlyGAGGCCGmlyGACCmlyACAGCCmlyGGGCAAGmlymly mlyGGACAGAGAGCCcmlyGGACmlymlyCmlyACAGCAmlyCCACGmlyGACCAAGAAAGCAC ACACCCCGmlyGAAGCCCGmlyGGCmlyCAGAmlyCGCCGGCAAmlyAGAmlyACGCCmlyCmlyG GACCmlyGmlyGGGCAAGCCcmlyGmlyCCGAmlyGCCmlyGCAmlyGGGAACAAmlyCGCCAGC mlymlyCCmlyGAGCAAGmlyACCAGGACAmlyCAmlyCmlyCGAGCACCAGAAGGmlyGGmlyCA AGGGCAACCAGAAGAGACmlyGAAAGCCmlyGAGGGAGCmlyGGCCGGCAAGAGAACCM lyGGAAmlyACCCAGCGmlyGACCcmlyGCCcmlyCCmlyCAGCCmlyCACACAAAGAAGGCGm lyGGAGCCcmlyACAACGAAGmlyGAmlyCGCCAGAGmlyGAGAAmlyGmlyGGGmlyCAACcmlyG AACcmlyGmlyGGCAGAAGCmlyGAAACmlyGmlyCCAGGACGACGCCAAGCCmlyCmlyGCMly GAGACmlyGAAGGGCmlymlyCCcmlyAGCmlymlyCCcmlyCmlyGGmlyGGAAAGACAGGCCAAm lyGAAGmlyGGAcmlymlyGGmlyGGGACAmlyGGmlyCmlyGCAACGmlyGAAGAAGCmlyGAmlyCAA CGAGAAGAAAGAGGAmlyGGCAAGGmlymlymlyCmlyGGCAGAACCmlyGGCCGGCmlyACA AGAGACAAGAAGCCcmlyGAGGCCmlymlyACCmlyGAGCAGCGAAGAGGACCGGAAGAAGG GCAAGAAAGmlymlyCGCCAGAmlyACCAGCmlyGGGCGACCmlyGCMlyGCMlyGCACmlyGGAA AAGAAGCAGCGCGAGGACmlyGGGGCAAGAmlyGmlyACGAmlyGAGGCCmlyGGGAGAGAAm lyCGACAAGAAGGmlyGGAAGGCCmlyGAGCAAGCACAmlymlyAAGCmlyGGAAGAGGAAAGA AGGAGCGAGGACGCCCAAmlyCmlyAAAGCCGCMlyCmlyGACCGAmlymlyGGCmlyGAGAGCC AAGGCCAGCmlymlymlyGmlyGAmlyCGAGGGCmlyGAAAGAGGCCGACAAAGGACGAGmlymly CmlyGCAGAmlyGCGAGCmlyGAAGCmlyGCAGAAAGmlyGGmlyACGGCGAmlyCmlyGAGAGGCA AGCCcmlymlyCGCCAmlymlyGAGGCCGAGAACAGCAmlyCCmlyGGACAmlyCAGCGCmlymly CAGCAAGCAGmlyACAAcmlyGCGCCmlymlyCAmlymlymlyGGCAGAAAGACGGCGmlyCAAGA AACmlyGAACcmlyGmlyACCmlyGAmlyCAmlyCAAmlymlyACmlymlyCAAAGGCGGCAAGCmlyG CGGmlymlyCAAGAAGAmlyCAAACCCGAGGCCmlymlyCGAGGCCmlyACAGAmlymlyCmlyACA CCGmlyGAmlyCAACAAAAGmlyCCGGCGAGAmlyCGmlyGCCAmlyGGAAGmlyGAACmlymly CAACmlymlyCGACGACCCCAACcmlyGAmlymlyAmlyCCmlyGCCmlyCmlyGGCCmlymlyCGGCA AGAGACAGGGCAGAGAmlymlyCmlyCmlyGGAACGAmlyCmlyGCMlyGAGCCmlyGGAACCC GGCMlyCmlyCmlyGAAGCmlyGGCCAAmlyGGCAGAGmlyGAmlyCGAGAAAACCCmlyGmlyACA ACAGGAGAACCAGACAGGACGAGCCmlyGCMlyCmlyGmlymlymlyGmlyGGCCcmlyGACCmlymly lyCGAGAGAAGAGAGGmlyGCMlyGGACAGCAGCAACAmlyCAAGCCCAmlyGAACcmlyGAmly CGGCmlyGGCCCGGGGCGAGAAmlyAmlyCCcmlyGCMlyGmlyGAmlyCGCCcmlyGACAGACC CmlyGAAGGAmlyGCCACmlyGAGCAGAmlymlyCAAGGACmlyCCcmlyGGGCAACCCmlyACA CACAmlyCCmlyGAGAAmlyCGGCGAGAGCmlyACAAAGAGAAGCAGAGGACAAmlyCCAGGC CAAGAAAGAGGmlyGGAACAGAGAAGAGCCGGCGGAmlyACmlyCmlyAGGAAGmlyACGCCA GCAAGGCCAAGAAmlyCmlyGGCCGACGACAmlyGGmlyCCGAAACACCGCCAGAGAmlyCmly GCMlyGmlyACmlyACGCCGmlyGACACAGGACGCCAmlyGCMlyGAmlyCmlymlyCGCGAAmlyCm lyGAGCAGAGGCMlymlyCGGCCGGCAGGGCAAGAGAACCmlymlymlyAmlyGGCCGAGAGGCA GmlyACACCAGAAmlyGGAAGAmlymlyGGCmlyCACAGCmlyAAACmlyGGCCmlyACGAGGGAC mlyGAGCAAGACCcmlyACCmlyGmlyCCAAAACAcmlyGGCCCAAmlyAmlyACCmlyCCAGAAC mlyGCAGCAAmlymlyGCGGCMlymlyCACCAmlyCACCAGCGCCGACmlyACGACAGAGmlyGCM lyGGAAAAGCmlyCAAGAAAACCGCCACCGGCMlyGGAmlyGACCACCmlyCAACGGCAAAG AGCmlyGAAGGmlymlyGAGGGCCAGAmlyCACcmlyACmlyACAACAGGmlyACAAGAGGCAGA ACmlyCGmlyGAAGGAmlyCmlyGAGCGmlyGGAACmlyGGACAGACmlyGAGCGAAGAGAGCG mlyGAACAACGACAmlyCAGCAGCmlyGGACAAAGGGCAGAmlyCAGGCGAGGCMlyCmlyGAG CCMlyGCMlyGAAGAAGAGGmlymlyAGCCACAGACCmlyGmlyGCAAGAGAAAmlymlyCGmly GmlyGCCmlyGAACmlyGCGGCMlymlyCGAGACACACGCCGCMlyGAACAGGCMlyGCCcmlyGA ACAmlymlyGCCAGAAGCmlyGGCmlyGmlymlyCCmlyGAGAAGCCAAGAGmlyACAAGAAGmlyA	

[illegible]

LTRP ID	RNA sequence	SEQ ID NO:
	GAACCAmψCmψGAGCAAGAGAmψCAAGAGAAmψCAACAAGAmψCAGAAGGAGACmψGG mψCAAGGACAGCAACACAAAGAAGGCCGCGCAAGACAGGCCCCAmψGAAAACCCmψGCM ψCGmψCAGAGmψGAmψGACCCmψGACmψGAGAGAGCGGCMψGGAAAACCCmψGAGAA AGAAGCCCGAGAACAmψCCCMψCAGCCmψAmψCAGCAACACCAGCAGGGCCAAACmψG AACAAAGCmψGCMψGACCGACmψACACCGAGAmψGAAGAAAGCCAmψCCmψGCACGmψG mψACmψGGGAAGAGmψmψCCAGAAAGACCCGmψGGGCCmψGAmψGAGCAGAGmψmψG CMψCAGCCmψGCCAGCAAGAGAmψCGACCAAGAACAGCmψGAAGCCCCGAGAmψGGAC GAGAAGGGCAAmψCmψGACCACAGCCGGCmψmψmψGCCmψGCMψCmψCAGmψGmψGGC CAGCCmψCmψGmψmψCGmψGmψACAAGCmψGGAACAGGmψGmψCCGAGAAAGGCAAGG CCmψACACCAACmψACmψmψCGGCAGAmψGmψAACGmψGGCCGAGCACGAGAAGCmψG AmψmψCmψGCMψGGCCAGCmψGAAACmψGAGAAGGACmψCmψGAmψGAGGCCGmψG ACCMψACAGCCmψGGGCAAGmψmψmψGGACAGAGAGCCmψGGACmψmψCmψACAGCA mψCCACGmψGACCAAGAAAGCACACCCCCGmψGAAGCCCCmψGGCmψCAGAmψCGC CGGCAAmψAGAmψACGCCmψCmψGGACmψGmψGGCAAGGCCmψGmψCCGAmψGCC mψGCAmψGGGAACAAmψCGCCAGCmψmψCCmψGAGCAAGmψACCAGGACAmψCAmψCA mψCGAGCACCAAGAGmψGGmψCAAGGGCAACCAGAAAGAGACmψGGAAAGCCmψGAGG GAGCmψGGCCGGCAAGAGAACmψGGAAmψACCCAGCGmψGACCCmψGCCmψCCmψ CAGCCmψCACACAAAGAAGGCGmψGGACGCCmψACAACGAAGmψGAmψCGCCAGAGm ψGAGAAmψGmψGGmψCAACCMψGAACCMψGmψGGCAGAAGCmψGAAACmψGmψCCAG GGACGACGCCAAGCCmψCmψGCMψGAGACmψGAAGGGCmψmψCCCMψAGCmψmψCCCM ψCmψGGmψGGAAAGACAGGCCAAmψGAAGmψGGAmψmψGGmψGGGACAmψGGmψCmψG CAACGmψGAAGAAGCmψGAmψCAACGAGAAGAAAGAGGAmψGGCAAGGmψmψmψmψCm ψGGCAGAACCmψGGCCGGCmψACAAGAGACAAGAAGCCmψGAGGCCmψmψACCmψGA GCAGCGAAGAGGACCGGAAGAAGGGCAAGAAGmψmψCGCCAGAmψACCAGCmψGGGCG ACCMψGCCmψGCMψGCACCMψGGAAAGAAAGCACGGCGAGGACmψGGGGCAAGmψGmψ ACGAmψGAGGCCmψGGGAGAGAAmψCGACAAGAAGGmψGGAAGGCCmψGAGCAAGCAC AmψmψAAGCmψGGAAGAGGAAAGAGGAGCGAGGACGCCAAmψCmψAAAGCCGCMψC mψGACCGAmψmψGGCmψGAGAGCCAAAGGCCAGCmψmψmψGmψGAmψCGAGGCCmψGA AAGAGGCCGACAAGGACGAGmψmψCmψGCAGAmψGCGAGCmψGAAGCmψGCAGAAAGmψ GGmψACGGCGAmψCmψGAGAGGCAAGCCmψmψCGCCAmψmψGAGGCCGAGAACAGCA mψCCmψGGACAmψCAGCGGCmψmψCAGCAAGCAGmψACAACmψGCCCMψmψCAmψmψ mψGGCAGAAAGACGGCGmψCAAGAAACmψGAACCMψGmψACCmψGAmψCAmψCAAmψm ψACmψmψCAAAGGCGGCAAGCmψGCGGmψmψCAAGAAGAmψCAAAACCGAGGCCmψmψ ψCAGGCMψAACAGAmψmψCmψACACCGmψGAmψCAACAAAAAGmψCCGGCGAGAmψCG mψGCCCCAmψGGAAGmψGAAAmψmψCAACmψmψCGACGACCCCAACmψGAmψmψAmψC CmψGCCmψCmψGGCCmψmψCGGCAAGAGACAGGGCAGAGAGmψmψCAmψCmψGGAACG AmψCmψGCCmψGAGCCmψGGAAACCGGCMψCmψCmψGAAGCmψGGCCAAmψGGCAGAGm ψGAmψCGAGAAAACCCmψGmψACAACAGGAGAACCAGACAGGACGAGCCmψGCCmψCmψ GmψmψmψGmψGGCCmψGACCMψmψCGAGAGAAGAGAGGmψGCMψGGACAGCGACCAAC AmψCAAGCCCAmψGAACCMψGAmψCGGCGmψGGCCGGGGCAGAGAAmψAmψCCCMψGC mψGmψGAmψCGCCCMψGACAGACCCmψGAAGGAmψGCCACmψGAGCAGAmψmψCAAG GACmψCCCMψGGGCAACCCmψACACACAmψCmψGAGAAmψCGGCGAGAGCmψACAAA GAGAAGCAGAGGACAAmψCCAGGCCAAGAAAGAGGmψGGAACAGAGAAGAGCCGGCGG AmψACmψCmψAGGAAGmψACGCCAGCAAGGCCAAGAAmψCmψGGCCGACGACAmψGGm ψCCGAAACACCGCCAGAGAmψCmψGCMψGmψACmψACGCCmψGACACAGGACGCCAm ψGCCmψGAmψCmψmψCGCGAAmψCmψGAGCAGAGGCMψmψCGGCCGGCAGGGCAAGAGA ACCMψmψmψAmψGGCCGAGAGGCAGmψACACCAGAAmψGGAAGAmψmψGGCmψCACAG CmψAAACmψGGCCmψACGAGGGACmψGAGCAAGACCmψACCmψGmψCCAAAACACmψG GCCAGmψAmψACCmψCCAAGACCmψGCAGCAAmψmψGCGGCMψmψCACCmψCACCA GCGCCGACmψACGACAGAGmψGCMψGGAAGAGCmψCAAGAAACCGCCACCGGCMψGG AmψGACCACCAmψCAACGGCAAGAGCmψGAAGGmψmψGAGGGCCAGAmψCACCMψAC mψACAACAGGmψACAAGAGGCAGAACmψCGmψGAAGGAmψCmψGAGCGmψGGAACmψ GGACAGACmψGAGCGAAGAGAGCGmψGAACAACGACAmψCAGCAGCmψGGACAAAGGG CAGAmψCAGGCGAGGCMψCmψGAGCCmψGCMψGAAGAAGAGGmψmψmψAGCCACAGAC CmψGmψGCAAGAGAAGmψmψCGmψGmψGCCmψGAACmψGCGGCMψmψCGAGACACACG CCGCMψGAACAGGCMψGCCmψGAACAmψmψGCCAGAAGCmψGGCmψGmψmψCCmψGA AAGCCAAAGAmψACAAGAAGmψACCAGACCAACAGAACACCGGCAACACCGCAGAACG AGGGCCmψmψmψGmψGGAAACmψGGCAGAGCmψmψCmψACAGAAAAAGCmψGAAAG AAGmψCmψGGAAGCCCCGCGmψGCGAmψCGGGCGmψmψCCGGCGAGGmψmψCCACm	

LTRP ID	RNA sequence	SEQ ID NO:
	ψAGmψCCAAAAAGAAGAGAAAGGmψAGAmψmψACAAAGAmψGACGAmψGACAAAGACmψACAAGGAmψGAmψGAmψGAmψAAGGGAmψCCGGCmψGAAAAAAAAAAAAAAAAAAAAA AA AAA	

Example 6: CpG-depletion of DNA encoding the guide RNA scaffold improves CasX-mediated editing *in vitro*

[0448] Pathogen-associated molecular patterns (PAMPs), such as unmethylated CpG motifs, are small molecular motifs conserved within a class of microbes. They are recognized by toll-like receptors (TLRs) and other pattern recognition receptors in eukaryotes and often induce a non-specific immune activation. In the context of gene therapy, therapeutics containing PAMPs are often not as well-tolerated and are rapidly cleared from the patient given the strong immune response triggered, which ultimately leads to reduced therapeutic efficacy. CpG motifs are short single-stranded DNA sequences containing the dinucleotide CG. When these CpG motifs are unmethylated, they act as PAMPs and therefore stimulate the immune response. In this example, experiments were performed to deplete CpG motifs in the guide scaffold coding sequence in the context of an AAV construct encoding CasX variant 491, guide scaffold variant 235, and spacer 7.37 targeting the endogenous B2M (beta-2-microglobulin) locus, and test the effect of CpG-depletion in the guide scaffold on editing of the B2M locus *in vitro*.

Materials and methods:
Design of CpG-depleted guide scaffolds:

[0449] Nucleotide substitutions were rationally-designed to replace native CpG motifs within the base gRNA scaffold variant (gRNA scaffold 235) with the intent to preserve editing activity while reducing scaffold immunogenicity. It was believed that as many CpG-motifs as possible should be removed from the scaffold coding sequence in order to sufficiently reduce immunogenicity. Scaffold 235 contains a total of eight CpG elements; six of which are predicted to basepair and form complementary strands of a double-stranded secondary structure (see FIG. 22). Therefore, the six basepairing CpGs forming three pairs were mutated in concert to maintain important secondary structures. This reduced the number of independent CpG-containing regions to five (three pairs and two single CpGs) to be considered independently for CpG-removal. Specifically, mutations were designed in (1) the pseudoknot stem, (2) the scaffold stem, (3) the extended stem bubble, (4) the extended step, and (5) the extended stem loop, as diagrammed in FIG. 23 and described in detail below.

[0450] In the pseudoknot stem (region 1), the CpG pair was flipped to a GpC to minimize the alteration of the base composition and sequence. Based on previous experiments involving replacing individual base pairs, it was anticipated that this mutation was not likely to be detrimental to the structure and function of the guide RNA scaffold.

[0451] Similarly, in the scaffold stem (region 2) the CpG pair was flipped to a GpC to minimize the alteration of the base composition and sequence. It was anticipated that this mutation was likely to be detrimental to the structure and function of the guide RNA scaffold because strong sequence conservation was seen in this region in previous experiments mutating individual bases or base pairs. This strong sequence conservation is likely due to the scaffold stem loop being important in interacting with the CasX protein as well as in the formation of a triplex structural element with the pseudoknot region.

[0452] In the extended stem bubble (region 3) the single CpG was removed by one of three strategies. First, the bubble was deleted by mutating CG->C. Second, the bubble was resolved to restore ideal basepairing by mutating CG->CT. Third, the entire extended stem loop was replaced with the extended stem loop of scaffold 174. Note that, by itself, the replacement of the extended stem loop with that of scaffold 174 recapitulates scaffold 316, which has previously been shown to edit efficiently. There are no CpG motifs in the extended stem loop of scaffold 174. Therefore, replacing the extended stem loop with that of scaffold 174 also removes the CpG motif in the extended stem (region 4). Based on previous experiments showing the relative robustness of the extended stem to small changes, it was anticipated that mutating the extended stem bubble was moderately likely to be detrimental to the structure and function of the guide RNA scaffold.

[0453] In the extended stem (region 4), the CpG pair could not be flipped to GpC without generating additional CpG motifs. Therefore, the CpGs were changed to a GG and a complementary CC motif. Similar to region 3, based on the relative robustness of the extended stem to small changes, it was anticipated that this mutation was not likely to be detrimental to the structure and function of the guide RNA scaffold.

[0454] Finally, the extended stem loop (region 5) was mutated in one of three ways that were designed based on previous experiments examining the stability of the stem loop. In particular, several variations of the stem loop had previously been shown to have similar stability levels, and some of these variations of the stem loop do not contain CpGs. Based on these findings, first, the loop was replaced with a new loop with a CUUG sequence. Second, the loop was replaced with a new loop with a GAAA sequence. Since the GAAA loop replacement would generate a novel CpG adjacent to the loop, it was combined with a C->G

base swap and the corresponding G C base swap on the complementary strand, ultimately resulting in a CUUCGG->GGAAAC exchange. Third, the loop was mutated by the insertion of an A to interrupt the CpG motif and thereby increase the size of the loop from 4 to 5 bases. It was anticipated that randomly mutating the extended stem loop would likely have detrimental effects on secondary structure stability and hence on editing. However, relying on previously confirmed sequences was believed to have a lower risk associated with a replacement.

[0455] To generate guide RNA scaffolds encoded by DNA with reduced CpG levels, the mutations described above were combined in various configurations. Table 33, below, summarizes combinations of the mutations that were used. In Table 33, a 0 indicates that no mutation was introduced to a given region, a 1, 2, or 3 indicates that a mutation was introduced in that region, as diagrammed in FIG. 23, and n/a indicates not applicable. Specifically, for region 1, the pseudoknot stem, a 1 indicates that a CG->GC mutation was introduced. For region 2, the scaffold stem, a 1 indicates that a CG->GC mutation was introduced. For region 3, the extended stem bubble, a 1 indicates that the bubble was removed by the deletion of the G and A bases that form the bubble, a 2 indicates that the bubble was resolved by a CG->CU mutation that allows for basepairing between the A and U bases, and a 3 indicates that the extended stem loop was replaced with the extended step loop from guide scaffold 174. For region 4, the extended stem, a 1 indicates that a CG->GC mutation was introduced. For region 5, the extended stem loop, a 1 indicates that the loop was replaced from UUCG->CUUG, a 2 indicates that the loop was replaced along with a basepair adjacent to the loop, from CUUCGG->GGAAAC, and a 3 indicates that an A was inserted between the C and the G.

Table 33: Summary of mutations for CpG-reduction and depletion in guide scaffold 235

Scaffold ID	Region 1 (Pseudoknot stem)	Region 2 (Scaffold stem)	Region 3 (Extended stem bubble)	Region 4 (Extended stem)	Region 5 (Extended stem loop)
320	1	0	0	1	0
321	1	0	1	1	0
322	1	0	2	1	0
323	1	0	3	n/a	0
324	1	0	1	1	1
325	1	0	2	1	1
326	1	0	3	n/a	1
327	1	0	1	1	2

328	1	0	2	1	2
329	1	0	3	n/a	2
330	1	0	1	1	3
331	1	0	2	1	3
332	1	0	3	n/a	3
334	1	1	2	1	1
335	1	1	3	n/a	1
336	1	1	1	1	2
337	1	1	2	1	2
338	1	1	3	n/a	2
339	1	1	1	1	3
340	1	1	2	1	3
341	1	1	3	n/a	3
235	0	0	0	0	0

[0456] Table 34, below, lists the DNA sequences of the designed CpG-reduced or depleted guide scaffolds.

Table 34: DNA sequences encoding CpG-reduced or depleted guide RNA scaffolds

Scaffold ID	DNA sequence	SEQ ID NO
320	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGCGACTATGTCGTAG TGGGTAAAGCCGCTTAGGGACTTCGGTCCCTAAGAGGCATCAGAG	3210
321	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGCGACTATGTCGTAG TGGGTAAAGCCCTTAGGGACTTCGGTCCCTAAGGGCATCAGAG	3211
322	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGCGACTATGTCGTAG TGGGTAAAGCCTCTTAGGGACTTCGGTCCCTAAGAGGCATCAGAG	3212
323	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGCGACTATGTCGTAG TGGGTAAAGCTCCCTCTTCGGAGGGAGCATCAGAG	3213
324	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGCGACTATGTCGTAG TGGGTAAAGCCCTTAGGGACCTTGGTCCCTAAGGGCATCAGAG	3214
325	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGCGACTATGTCGTAG TGGGTAAAGCCTCTTAGGGACCTTGGTCCCTAAGAGGCATCAGAG	3215
326	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGCGACTATGTCGTAG TGGGTAAAGCTCCCTCCTTGGAGGGAGCATCAGAG	3216
327	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGCGACTATGTCGTAG TGGGTAAAGCCCTTAGGGAGGAACTCCCTAAGGGCATCAGAG	3217
328	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGCGACTATGTCGTAG TGGGTAAAGCCTCTTAGGGAGGAACTCCCTAAGAGGCATCAGAG	3218
329	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGCGACTATGTCGTAG TGGGTAAAGCTCCCTGGAAACAGGGAGCATCAGAG	3219
330	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGCGACTATGTCGTAG TGGGTAAAGCCCTTAGGGACTTCAGGTCCCTAAGGGCATCAGAG	3220
331	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGCGACTATGTCGTAG TGGGTAAAGCCTCTTAGGGACTTCAGGTCCCTAAGAGGCATCAGAG	3221
332	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGCGACTATGTCGTAG TGGGTAAAGCTCCCTCTTCAGGAGGGAGCATCAGAG	3222

Scaffold ID	DNA sequence	SEQ ID NO
333	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGGCACTATGTGCTAG TGGGTAAAGCCCTTAGGGACCTTGGTCCCTAAGGGCATCAGAG	3223
334	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGGCACTATGTGCTAG TGGGTAAAGCCCTTAGGGACCTTGGTCCCTAAGGGCATCAGAG	3224
335	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGGCACTATGTGCTAG TGGGTAAAGCTCCCTCCTTGGAGGGAGCATCAGAG	3225
336	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGGCACTATGTGCTAG TGGGTAAAGCCCTTAGGGAGGAACTCCCTAAGGGCATCAGAG	3226
337	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGGCACTATGTGCTAG TGGGTAAAGCCCTTAGGGAGGAACTCCCTAAGGGCATCAGAG	3227
338	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGGCACTATGTGCTAG TGGGTAAAGCTCCCTGGAACAGGGAGCATCAGAG	3228
339	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGGCACTATGTGCTAG TGGGTAAAGCCCTTAGGGACTTCAGTCCCTAAGGGCATCAGAG	3229
340	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGGCACTATGTGCTAG TGGGTAAAGCCCTTAGGGACTTCAGTCCCTAAGGGCATCAGAG	3230
341	ACTGGGCCTTCTATCTGATTACTCTGAGGCCCATCACCAGGCACTATGTGCTAG TGGGTAAAGCTCCCTCTTCAGGAGGGAGCATCAGAG	3231

Generation of CpG-depleted AAV plasmids:

[0457] The CpG-reduced or depleted gRNA scaffolds were tested in the context of AAV vectors that were otherwise CpG-depleted, with the exception of the AAV2 ITRs. Specifically, nucleotide substitutions to replace native CpG motifs in AAV components were designed in silico based on homologous nucleotide sequences from related species for the following elements: the murine U1a snRNA (small nuclear RNA) gene promoter, the bGHpA (bovine growth hormone polyadenylation) sequence, and the human U6 promoter. The coding sequence for CasX 491 was codon-optimized for CpG depletion. All resulting sequences (Tables 34 and 35) were ordered as gene fragments with the appropriate overhangs for cloning and isothermal assembly to replace individually the corresponding elements of the existing base AAV plasmid (construct ID 183). Spacer 7.37

(GGCCGAGAUGUCUCGCUCCG; SEQ ID NO: 3137), which targets the endogenous B2M gene, was used for the experiments discussed in this example. The first time that the experiment was performed (“N=1”), a sample with the non-targeting spacer 0.0 was also included as a control (CGAGACGUAAUACGUCUCG, SEQ ID NO: 3232; see FIG. 24).

[0458] The resulting AAV constructs were generated using standard molecular cloning techniques. Cloned and sequence-validated plasmid constructs were midi-prepped for subsequent nucleofection and AAV vector production. The sequences of the additional components of AAV constructs, with the exception of sequences encoding the gRNAs (Table 33), are listed in Table 35.

Table 35: Sequences of AAV elements (5'-3' in AAV construct)

Element	DNA sequence	SEQ ID NO:
AAV2 5' ITR	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCCG GCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGA GTGGCCAACTCCATCACTAGGGGTTTCCT	3233
CpG-depleted Ula promoter	AATGGAGGTGGTACTATGTAGATGAGAATTCAGGAGCAAACCTGGGAAAAGC AACTGCTTCCAAATATTTGTGATTTTTACAGTGATTTTGGAAAACTCT TAGCCTACCAATTCTTCTAAGTGTTTTAAATGTGGGAGCCAGTACACATG AAGTTATAGAGTGTTTTAATGAGGCTTAAATATTTACTGTAACATGAAAT GCTACACATATCATGCTGTTTCAAGGCTCTGTGGCCATGCAACTCATACT	3234
CpG-depleted cMycNLS- CasX491- cMycNLS	ATGGCTCCAGCTGCCAAGAGAGTGAAGCTGGACTCTAGACAAGAGATCAAG AGGATCAACAAGATTAGAAGAAGGCTGGTCAAGGACAGCAACACCAAGAAG GCAGGCAAGACAGGCCCATGAAGACCCCTGCTTGTCAGAGTGATGACCCCT GACCTGAGAGAGAGACTGGAAAACCTGAGAAAGAAGCCTGAGAACATCCCT CAGCCTATCTCCAACACCAGCAGAGCCAACCTGAACAAGCTGCTGACAGAC TACACAGAGATGAAGAAAGCCATCCTGCATGTGTACTGGGAAGAGTTCCAG AAAGACCCCTGTGGGCCTGATGAGCAGAGTGGCCAGCCTGCCAGCAAGAAG ATTGACCAGAACAACTGAAGCCAGAGATGGATGAGAAGGGCAACCTGACC ACAGCTGGCTTTGCCTGCTCTCAGTGTTGGCCAGCCTCTGTTTGTGTACAAG CTGGAACAAGTGTCTGAGAAAGGCAAGGCCTACACCACTACTTTGGCAGA TGCAATGTGGCTGAGCATGAGAAGCTGATCCTGCTGGCCAGCTGAAGCCT GAAAAGGACTCTGATGAGGCTGTGACCTACAGCCTGGGCAAGTTTGGACAG AGAGCCCTGGACTTCTACAGCATCCATGTGACCAAAGAAAGCACACACCCT GTGAAGCCCTGGCTCAGATTGCTGGCAACAGATATGCCTCTGGACCTGTG GGCAAAGCCCTGTCTGATGCCTGCATGGGAACCATTGCCAGCTTTTCTGAGC AAGTACCAGGACATCATCATTGAGCACCAGAAAGTGGTCAAGGGGAACAG AAGAGGCTGGAATCTCTGAGAGAGCTGGCTGGCAAAGAGAACCTGGATAC CCCTCTGTGACCCCTGCCTCCTCAGCCTCACACCAAAGAGGGTGTGTATGCC TACAATGAAGTGATTGCCAGAGTCAGGATGTGGGTCAACCTGAATCTGTGG CAGAAGCTGAAGCTGAGCAGGGATGATGCCAAGCCTCTGCTGAGACTGAAG GGCTTCCCTAGCTTCCCCTGCTGGTGGAAAGGCAGGCTAATGAAGTGGATTGG TGGGACATGGTCTGCAATGTGAAGAACTGATCAATGAGAAGAAAGAGGAT GGCAAGGTTTTCTGGCAGAACTGGCTGGCTACAAGAGGCAAGAAGCCCTG AGGCCTTACCTGTCCTCTGAGGAAGATAGGAAGAAGGGGAAGAAGTTTGGC AGATACCAGCTGGGAGATCTGCTGCTGCACCTGGAAAAGAAACATGGGGAA GATTGGGGCAAAGTGTATGATGAGGCCTGGGAGAGAATTGACAAGAAGGTG GAAGGCCTGAGCAAGCACATTAAGCTGGAAGAGGAAAGAAGATCTGAGGAT GCCCAGAGCAAGGCTGCCCTGACTGATTGGCTGAGAGCCAAGGCCAGCTTT GTGATTGAGGGCTGAAAGAGGCTGACAAGGATGAGTTCTGTAGATGTGAA CTGAAGCTGCAAAAGTGGTATGGGGACCTGAGGGGCAAGCCCTTTGCCATT GAGGCAGAGAACAGCATCCTGGACATCAGTGGCTTCAAGCAAGCAGTACAAC TGTGCCTTCATTTGGCAGAAGGATGGGGTTAAGAAGCTCAACCTGTACCTG ATCATCACTACTTCAAAGGTGGCAAGCTGAGGTTCAAGAAGATCAAGCCA GAGGCCTTTGAGGCCAACAGGTTCTACACAGTGATCAACAAAAAGTCTGGG GAGATTGTGCCCATGGAAGTGAACCTCAACTTTGATGATCCCAACCTCATC ATCCTGCCACTGGCCTTTGGCAAGAGACAGGGCAGAGAATTATCTGGAAT GACCTGCTGAGCCTGGAACAGGCAGCCTGAACTGGCCAATGGCAGAGTG ATTGAGAAAACCTGTACAACAGAAGGACCAGACAGGATGAGCCTGCACTG TTTGTGGCCCTGACCTTTGAGAGAAGGGAAGTGTGGATAGCAGCAACATC AAGCCCATGAACCTGATTGGAGTGGACAGAGGGGAGAATATCCCTGTGTCT ATTGCCCTGACAGACCCTGAGGGCTGTCCCTGAGCAGATTCAAGGACTCC CTGGGCAACCCACACACATTCTGAGAATTGGGGAGAGCTACAAAGAGAAG CAGAGAACCATCCAGGCCAAGAAAGAAGTGAACAGAGGAGAGCTGGAGGC TACTCTAGGAAGTATGCCAGCAAAGCCAAGAATCTGGCTGATGACATGGTC AGAAACACAGCCAGGGACCTGCTGTACTATGCTGTGACCCAGGATGCCATG CTGATCTTTGAGAATCTGAGCAGAGGATTTGGGAGACAAGGCAAGAGGACC TTCATGGCTGAAAGACAGTACACCAGGATGGAAGATTGGCTGACTGCCAAA CTGGCCTATGAGGGACTGAGCAAGACCTACCTGTCCAAGACACTGGCCAG	3235

Element	DNA sequence	SEQ ID NO:
	TATACCTCCAAGACCTGCAGCAATTGTGGCTTCACCATCACCTCTGCAGAC TATGACAGAGTGCTGGAAAAGCTCAAGAAAACAGCCACAGGCTGGATGACC ACCATCAATGGGAAAGAGCTGAAGGTTGAGGGCCAGATCACCTACTACAAC AGATACAAAAGGCAGAATGTGGTTAAGGACCTGTCTGTGGAAGTGGACAGA CTGTCTGAGGAATCTGTGAACAATGACATCAGCAGCTGGACCAAGGGCAGA TCTGGGGAAGCTCTGAGCCTGCTGAAGAAGAGATTGAGCCACAGACCTGTG CAAGAGAAGTTTGTGTGTCTGAACTGTGGCTTTGAGACACATGCTGATGAA CAGGCTGCTCTGAACATTGCCAGAAGCTGGCTGTTCTGAGAAGCCAAGAG TACAAGAAGTACCAGACCAACAAGACCACAGGCAACACTGACAAGAGGGCC TTTGTGGAAACCTGGCAGAGCTTCTACAGAAAGAACTCAAAGAAGTCTGG AAGCCTGCTGTGGGCAGCCCTGCAGCTAAAAGGGTCAAGCTGGACTGA	
CpG-depleted bGH-polyA sequence	CTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCAGT GCCTTCCTTGACCCTGGAAGGTGCCACTCCCCTGCTCCTTTCCCTAA TAAAATGAGGAAATTGCATCACATTGTCTGAGTAGGTGTCAATCTTA TTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGA AGAGAATAGCAGGCATGCTGGGGA	3236
CpG-depleted U6 promoter	GAGGGCCTATTTCCCATGATTCCCTTCATATTTGCATATACAATACA AGGCTGTTAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAA GATATTAGTACAAAATATTTGATGTAGAAAGTAATAATTTCTTGGG TAGTTTGCAGTTTAAAATTATGTTTTAAAATGGACTATCATATGC TTACCATAACTTGAAAGTATTTCTATTTCTTGGCTTTATATATCTT GTGGAAAGGAC	3237
<i>See sgRNA sequences in Table 34, above.</i>		
AAV2 3' ITR	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGC TCACTGAGGCCGGGCGACCAAGGTGCGCCGACGCCCGGGCTTTGCCCGGG CGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	3238

AAV production:

[0459] Suspension-adapted HEK293T cells, maintained in FreeStyle 293 media, were seeded in 20-30 mL of media at 1.5E6 cells/mL on the day of transfection. Endotoxin-free pAAV plasmids with the transgene flanked by ITR repeats were co-transfected with plasmids supplying the adenoviral helper genes for replication and AAV rep/cap genome using PEI Max (Polysciences) in serum-free Opti-MEM media. Three days later, cultures were centrifuged to separate the supernatant from the cell pellet, and the AAV particles were collected, concentrated, and filtered following standard procedures.

[0460] To determine the viral genome (vg) titer, 1 μ L from crude lysate viruses was digested with DNase and ProtK, followed by quantitative PCR. 5 μ L of digested virus was used in a 25 μ L qPCR reaction composed of IDT primetime master mix and a set of primer and 6'FAM/Zen/IBFQ probe (IDT) designed to amplify a 62 bp-fragment located in the AAV2-ITR. An AAV ITR plasmid was used as reference standards to calculate the titer (vg/mL) of viral samples.

AAV transduction of induced neurons in vitro:

[0461] 24 hours prior to transduction, 50,000 induced neurons per well were seeded on Matrigel-coated 96-well plates. AAVs expressing the CasX:gRNA system with various versions of the guide scaffold were then diluted in neuronal plating media and added to cells. The first time that the experiment was performed (“N=1”), cells were transduced at a multiplicity of infection (MOI) of 4e3 viral genomes (vg)/cell (see FIG. 24). Seven days post-plating, induced neurons were transduced with virus diluted in fresh feeding media. Eight days post-transduction, cells were lifted using lysis buffer, 4-well replicates were pooled per experimental condition, and genomic DNA (gDNA) was harvested and prepared for editing analysis at the B2M locus using next generation sequencing (NGS). The second time that the experiment was performed (“N=2”), cells were transduced at an MOI of 3e3 vg/cell, 1e3 vg/cell, or 3e2 vg/cell (see FIG. 25, FIG. 26, and FIG. 27). Seven days post-plating, induced neurons were transduced with virus diluted in fresh feeding media. Seven days post-transduction, cells were lifted using lysis buffer, 2-well replicates were pooled per experimental condition, and gDNA was harvested and prepared for editing analysis at the B2M locus using NGS. Samples that were not transduced with AAV were included as controls.

NGS processing and analysis:

[0462] Genomic DNA (gDNA) from harvested cells were extracted using the Zymo Quick-DNA Miniprep Plus kit following the manufacturer’s instructions. Target amplicons were formed by amplifying regions of interest from 200 ng of extracted gDNA with a set of primers specific to the human B2M gene. These gene-specific primers contained an additional sequence at the 5’ end to introduce an Illumina adapter and a 16-nucleotide unique molecule identifier. Amplified DNA products were purified with the Ampure XP DNA cleanup kit. Quality and quantification of the amplicon were assessed using a Fragment Analyzer DNA Analysis kit (Agilent, dsDNA 35-1500bp). Amplicons were sequenced on the Illumina Miseq according to the manufacturer's instructions. Raw fastq files from sequencing were quality-controlled and processed using cutadapt v2.1, flash2 v2.2.00, and CRISPResso2 v2.0.29. Each sequence was quantified for containing an insertion or deletion (indel) relative to the reference sequence, in a window around the 3’ end of the spacer (30 bp window centered at –3 bp from 3’ end of spacer). CasX activity was quantified as the total percent of reads that contain insertions, substitutions, and/or deletions anywhere within this window for each sample.

Results:

[0463] Mutations were introduced into the guide scaffold 235 in order to reduce the CpG content of the DNA sequence coding the guide scaffold. Surprisingly, compared to scaffold 235, all of the CpG-reduced and CpG-depleted scaffold variants produced higher levels of editing in induced neurons. This was the case with two independent repeats of the experiment (with the results from the first repeat of the experiment shown in FIG. 24, and the results of the second repeat of the experiment shown in FIGS. 25-27), and across multiple MOIs (see FIGS. 26-27). The enhanced level of editing was surprising because the goal of reducing CpG content was to simply preserve editing activity while reducing immunogenicity. Instead, the mutations enhanced editing activity, rather than merely preserving it.

[0464] Notably, scaffold 320 showed a significant increase in potency over scaffold 235. Scaffold 320 includes mutations to only two regions of the scaffold; in the pseudoknot stem and the extended stem (regions 1 and 4). Further, some combinations of mutations produced worse editing than scaffold 320. However, even the CpG-reduced scaffolds that performed worse than scaffold 320, such as scaffolds 331 and 334, performed similar to or better than scaffold 235.

[0465] Based on these results, without wishing to be bound by theory, it is believed that the boost in potency seen in many of the CpG-reduced and CpG-depleted scaffolds is likely caused by one of the mutations present in all CpG-reduced scaffolds (i.e., region 1 and/or 4). Since the mutation to region 4 is not present in the scaffolds with the extended stem loop replacement (i.e., the third mutation to region 3) and these scaffolds show a similar improvement in potency over 235 as 320 did, it is believed that the beneficial effect is likely caused by the mutation in region 1 (pseudoknot stem), which is present in all of the tested scaffolds. Further experiments will be performed to test the effect of the individual mutations in the pseudoknot stem (region 1) and the extended stem (region 4) separately.

[0466] Further, the N=1 data as presented in FIG. 24 indicate that all the new scaffolds carrying the mutation in region 2 (scaffold stem) edited at a slightly lower level than their respective counterparts without this mutation. This suggests that mutating this position in the scaffold stem may have a small deleterious effect on editing potency. This will be examined in additional experiments.

[0467] The results described here demonstrate that introducing mutations that reduced the CpG content of the DNA encoding the guide RNA scaffold resulted in improvements in gene editing relative to guide scaffold 235.

Example 7: Demonstration that inclusion of the ADD domain from DNMT3A enhances activity and specificity of LTRP molecules

[0468] In addition to its C-terminal methyltransferase domain, DNMT3A contains two N-terminal domains that regulate its function and recruitment to chromatin: the ADD domain and the PWWP domain. The PWWP domain reportedly interacts with methylated histone tails, including H3K36me3. The ADD domain is known to have two key functions: 1) it allosterically regulates the catalytic activity of DNMT3A by serving as a methyltransferase auto-inhibitory domain, and 2) it recognizes unmethylated H3K4 (H3K4me0). The interaction of the ADD domain with the H3K4me0 mark unveils the catalytic site of DNMT3A, thereby recruiting an active DNMT3A to chromatin to implement de novo methylation at these sites.

[0469] Given these functions of the ADD domain, experiments were performed to assess whether the incorporation of the ADD domain into the LTRP5 construct, described previously in Example 1, would result in improved long-term repression of the target locus and reduced off-target methylation. The effect of incorporating the PWWP domain along with the ADD domain on LTRP activity and specificity was also assessed.

Materials and Methods:

Generation of LTRP constructs and plasmid cloning:

[0470] Plasmid constructs encoding for variants of the LTRP5 construct with the ZIM3-KRAB domain (LTRP5.A; see FIG. 28 for LTRP5 configuration) were built using standard molecular cloning techniques. The resulting constructs comprised of sequences encoding for one of the following four alternative variations of LTRP5-ZIM3, where the additional DNMT3A domains were incorporated: 1) LTRP5-ZIM3 + ADD; 2) LTRP5-ZIM3 + ADD + PWWP; 3) LTRP5-ZIM3 + ADD without the DNMT3A catalytic domain; and 4) LTRP5-ZIM3 + ADD + PWWP without the DNMT3A catalytic domain. The sequences of key elements within the LTRP5-ZIM3 molecule and its variants are listed in Table 36, with the full-length protein sequence for each LTRP5-ZIM3 and its variants listed in Table 37. FIG. 29 is a schematic that illustrates the various LTRP5 architectures assayed in this example. Sequences encoding the LTRP molecules also contained a 2x FLAG tag. Plasmids also harbored constructs encoding for the gRNA scaffold variant 174 having either a spacer targeting the endogenous *B2M* locus or a non-targeting control (spacer sequences listed in Table 38).

Table 36: Sequences of LTRP components (e.g., additional domains fused to dCasX) to generate LTRP5 variants illustrated in FIG. 29

Component	Protein sequence	SEQ ID NO
ZIM3 KRAB domain	MNNSQGRVTFEDVTVNFTQGEWQRLNPEQRNLYRDVMLENYSNLVSVGGQGETTKPDVILRLREQKEPWLLEEEVLGSGRAEKNIGDGGQIWKPDKVKESL	3240
DNMT3A catalytic domain (CD)	NHDQEFDPKVPYPPVPAEKRRKPIRVLSLFDGIATGLLVKDLGIQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCN DLSIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWL FENVVAMGVSDKRDISRFLESNPVMI DAKEVSAHRARYFWGNLPGMNRPLASTVND KLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWC TEMERVFGFPVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACV	126
DNMT3L interaction domain	MGPMEIYKTVSAWKRPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQESQRPFFWI FMDNLLLTEDDQETTTTRFLQTEAVTLQDVRGRDYQNA MRVWSNIPGLKSKHAPLTPKEEYLAQVRSRSLDAPKVDLLVKNCLLP LREYFKYFSQNSLPL	127
dCasX491	QEIKRINKIRRLVKDSNTKKAGKTGPMKTLVVRVMTPLRERLENLRKKPENIPQPI SNTSRANLNKLLTDYTEMKKAILHVYWEFQKDPVGLMSRVAQPASKKIDQNKLPKPEMDEKGNLTAGFACSCQCGQPLFVYKLEQVSEKGYATNYFGRCNVAEHEKLILLQAQLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGPVGKALSDACMGTIASFLSKYQDIIIEHQKVVGKNQKRLESRELAKENLEYPVTLPPQPHTKEGVDAYNEVIARVRMWVNLNLWQKLKLSRDDAKPLRLKGFPSFPLVERQANEVDWDMVCNVKKLINEKKEDGKVFQWQNLAGYKRQEALRPYLSSEEDRKKGKKFARYQLGDLLEHLEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKLEEEERSEDAQSKAALTDWLRKASFVIEGLKEADKDEFRCCELKQKWYGDRLRGKPFAlEAE NSILDISGFSKQYNCAFIWQKDGVKKLNLYLI INYFKGGKLRFKKIKPEAFANRFYTVINKKSGEIVPMEVNFNDDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGRVIEKTLYNRRTRQDEPALFVALTFERREVLDSNIPKPNLIGVARGENIPAVIALTDPEGCPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKAKNLADDMVRNTARDLLYYAVTQDAMLIFANLSRGFGRQGRKFMAERQYTRMEDWLTAKLAYEGLSKTYLSKTLAQYTSKTCSNCGFTITSADYDRVLEKLKKTATGWMTTINGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVNNDISSWTKGRSGEALSLLKKRF SHRPVQEKFVCLNCGFETHAAEQALNIARSWLFLRSQEYKKYQTNKTTGNTDKRAFVETWQSFYRKKLKEVWKPAV	4
Linker 1	GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSE	123
Linker 2	SSGNSNANSRGPSFSSGLVPLSLRGSH	122
Linker 3A'	GGSGGG	124
Linker 3B	GGSGGGS	120
Linker 4	GGSGGGG	121
NLS A	PKKKRKV	30
NLS B		
DNMT3A ADD domain	ERLVYEVQRQKCRNIEDICISCGSLNVTLEHPLFIGMCCQNCNCFLECAYYQDDDGYSYCTICCGGREVLNCGNNCCRCFCVECDLLVGPAAQAAIKEDPWNCYMCGHKGTGYGLLRREDWPSRLQMFFAN	125

Component	Protein sequence	SEQ ID NO
DNMT3A PWWP domain	TKAADDEFEYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEG TRWVMWFGDGKFSVVCVEKLMPLSSFCFAFHQATYNKQPMYRKAIYEV LQ VASSRAGKLFPA CHDSDES DSGKAVEVQNKMIEWALGGFQPSGPKGLEP PEEEKNPYKEV	3252
Endogenous sequence between DNMT3A PWWP and ADD domains (endo)	YTD MWVEPEAAAYAPPPAKKPRKSTTEKPKVKEIIDERTR	3253

Table 37: Protein sequences of LTRP5 variants assayed in this example

LTRP ID	Protein sequence	SEQ ID NO
LTRP5-ZIM3	MAPKKKKRVS RMNHDQEFDPKVPVPAEKRRKPIRVLSLFDGIATGLLV LKDLGIQVDRIY IASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKH IQEWG PFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDR PFFWLFENVVAMGVSDKRDI SRFLESNPVMI DAEKVSAAHRARYFWGNLP GMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFP VFMNEKEDI LWCTEMERVFGFPVHYTDVSNMSRLARQRL LGRSWSVPVIR HLFAPLKEYFACVSSGNANSRGP SFSSGLVPLSLRGSHMGPM EIKTV SAWKRQPV RVLSLFRNIDKVLKSLGFLES GSGSGGGT LKYVEDVTNVVRR DVEKWGPF DLVYGSTQPLGSSCDRC PGWYMFQFHRI LQYALPRQESQRPF FWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQ NAMRVWSNIPGL KSKHAPLTPKEE EYLQAQVRSRSLDAPKVDLLVKNCLLP LREYFKYFSQ NSLPLGGSGGGMNNSQGRVTFEDVTNFTQGEWQRLNPEQRNL YRDVMLE HYSNLVSVGQGETTKPDVILRL EQKEPWLEEEV LGSGRAEKNGDIGGQ IWKPKDVKESLGGPSSGAPPPSGSPAGSPTSTEEGTSESATPESGPGTS TEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEQEI KRINKI RRRLVKDSNTKKAGKTGPMKTL LVRVMT PDLRERLENLRKKPENIPQPI S NTSRANLNKLLTDYTEMKKAILHVYWE EFQKDPVGLMSRVAQ PASKKIDQ NKLKPEMDEKGNLTTAGFACSQCQGLFVYKLEQVSEK GKAYTNYFGRCN VAEHEKL ILLAQLKPEKDSDEAVTYS LGKFGQRALDFYSIHVTKESTHPV KPLAQIAGNRYASGPVGKALSDACMGTIASF LSKYQDIIIEHQKVVKGNQ KRLESIRELAGKENLEYPSVTLPPQPHTKEGV DAYNEVIARVRMWN LNL WQKLKLSRDDAKPLRLKGFPSFPLVERQANEVDWDMVCNVKKLINEKK EDGKVFWQNLAGYKRQEALRPYLSSEEDRKKGKKFARYQLGDLLLHLEKK HGEDWGKVYDEAWERIDKKVEGLSKHIKLEEEERSEDAQSKAALTDWLRA KASFVIEGLKEADKDEFRCCELK LQKWYGDLRGKPF AIEAENSILDISGF SKQYNCAFIWQKDGVKKLNL YLIINYFKGGKLRFKKIKPEAF EANRFYTV INKKSGEIVPMEVNFNFDDPNL IILPLAFGKRQGREFIWNDLLSLETGSL KLANGRVIEKTLYNRRTRQDEPALEFVALTFERREVLDSSNIKPMNLIGVA RGENIPAVIALTDPEGCP LSRFKDSLGNP THILRIGESYKEKQRTIQAKK EVEQRRAGGYSRKYASKAKNLADDMVRNTARDLLYYAVTQDAMLI FANLS RGFGRQ GKRTFMAERQYTRMEDWLTAKLAYEGLSKTYLSKTLAQYTSKTC SNCGFTITSADYDRVLEKLKKTATGWM TTINGKELKVEGQITYYNRYKRQ NVVKDLSVELDRLSEESVNNDI SSWTKGRSGEALSLLKKRFSHRPVQKEF VCLNCGFETHAAEQ AALNIARSWLFLRSQEYK KYQTNKTTGNTDKRAFVE TWQSFYRKKLKEVWKPAVRSGSGSGGSTSPKKKRKV	3131

LTRP ID	Protein sequence	SEQ ID NO
LTRP5-ZIM3 + ADD	MAPKKKKRVSRMERLVYEVVRQKCRNIEDICISCGSLNVTLEHPLFIGGMC QNCNCFLECAQYDDDDGYQSYCTICCGGREVLNCGNNCCRCFCVECV LLVGPAAQAAIKEDPWNCYMGHKGTYGLLRREDWPSRLQMFANND QEFDPKVPYPPVPAEKRKPIRVLSLFDGIATGLLVKDLGLIQVDRIASE VCEDSITVGMVRHQGKIMYVGDVRSVTQKHIEWGPFDLVIGGSPCNDLS IVNPARKGLYEGTGRLFFEFYRLLDHARPKEGDDRPFFWLFENVVAMGVS DKRDISRFLESNPVMIDAKEVSAHRARYFWGNLPGMNRPLASTVNDKLE LQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFVFMNEKEDILWCTEM ERVFGFPVHYTDVSNMSRLARQLLGRSWSVPVIRHLFAPLKEYFACVSS GNSNANSRGPSPFSSGLVPLSLRGSHMGPMIYKTVSAWKRPVRLVLSLFR NIDKVLKSLGFLSESGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGST QPLGSSCDRCPGWYMFQFHRIHQYALPRQESQRPFFWI FMDNLLLTEDDQ ETTTTRFLQTEAVTLQDVRGRDYQAMRVWSNIPGLKSKHAPLTPKEEEYL QAQVRSRSLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGSGGGMNS QGRVTFEDVTVNFTQGEWQRLNPEQRNLYRDVMLENYSNLVSVGQGETTK PDVILRLLEQGKEPWLEEEVLGSGRAEKNNGDIGGQIWKPKDVKESLGGPS SGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGS PTSTEEGTSTEPSEGSAPGSTEPSEQEI KRINKIRRRLVKDSNTKKAGK TGPMTLLVRVMTPLRERLENLRKKPENIPQPI SNTSRANLNKLLTDYT EMKKAILHVYWEFFQKDPVGLMSRVAQPAKKIDQNKLPKEMDEKGNLTT AGFACSQCGQPLFVYKLEQVSEKQAYTNYFGRCNVAEHEKILLQAQLKP EKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGP VGKALSDACMGTIASFLSKYQDIIIEHQKVVKGNQKRLESRELAKENL EYPSVTLPQPHTKEGVDAAYNEVIARVMWVNLNLWQKLLSRDDAKPLL RLKGFPSFPLVERQANEVDWDMVCNVKKLINEKKEDGKVFQNLAGYKR QEALRPYLSSEDRKKGKKFARYQLGDLHLEKKHGEDWGKDYDEAWER IDKKVEGLSKHILEEERRSEDAQSKAALTDWLRAKASFVIEGLKEADKD EFCRCLEKLQKWYGDRLGKPFADAEANSILDISGFSKQYNCAFIWQKDG VKKLNLILIINFKGGLRFFKKIKPEAFANRFYTVINKKSGEIVPMEVNF NFDDPNLII LPLAFGKRQGREFIWDLNLSLETGSLKLANGRVIEKTLYNR RTRQDEPALFVALTFERREVLDSSNIKPMNLIGVARGENIPAVIALTDPE GCPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYA SKAKNLADDMVRNTARDLLYYAVTQDAMLIFANLSRGFGRGKRTFMAER QYTRMEDWLTAKLAYEGLSKTYLSKTLAQYTSKTCNSCGFTITSADYDRV LEKLKKTATGWMTTINGKELKVEGQITYYNRYKRQNVVKDLSVELDRLE ESVNNDISSWTGKRSGEALSLLKKRFSHRPVQKQFVCLNCGFETHAEQA ALNIARSWLFLRSQYKQYQTNKTTGNTDKRAVETWQSFYRKKLKEVWK PAVRSGSGGGGSTSPKKRKV	3132
LTRP5-ZIM3 + ADD + PWWP	MAPKKKKRVSRMTKAADDEPEYEDGRGFGIGELVWGKLRGFSWWPGRIVS WWMTRGRSRAAEGTRWVMWFGDGKFSVVCVEKLMLPSSFCFAHQATYNKQ PMYRKAIYEVQVASSRAGKLPACHDSDES DSGKAVEVQNKQMI EWALG GFQPSGPKGLEPPEEKNPYKEVYTDMMWVEPEAAAYAPPPAKKPRKSTT EKPKVKEIIDERTRERLVYEVVRQKCRNIEDICISCGSLNVTLEHPLFIGG MCQNCNCFLECAQYDDDDGYQSYCTICCGGREVLNCGNNCCRCFCVECV VDLLVGPAAQAAIKEDPWNCYMGHKGTYGLLRREDWPSRLQMFANND HDQEFDPKVPYPPVPAEKRKPIRVLSLFDGIATGLLVKDLGLIQVDRIASE SEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIEWGPFDLVIGGSPCND LSIVNPARKGLYEGTGRLFFEFYRLLDHARPKEGDDRPFFWLFENVVAMG VSDKRDISRFLESNPVMIDAKEVSAHRARYFWGNLPGMNRPLASTVNDK LELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFVFMNEKEDILWCT EMERVFGFPVHYTDVSNMSRLARQLLGRSWSVPVIRHLFAPLKEYFACV SSGNSNANSRGPSPFSSGLVPLSLRGSHMGPMIYKTVSAWKRPVRLVLSL FRNIDKVLKSLGFLSESGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYG STQPLGSSCDRCPGWYMFQFHRIHQYALPRQESQRPFFWI FMDNLLLTED DQETTTTRFLQTEAVTLQDVRGRDYQAMRVWSNIPGLKSKHAPLTPKEEE YLQAQVRSRSLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGSGGGMN NSQGRVTFEDVTVNFTQGEWQRLNPEQRNLYRDVMLENYSNLVSVGQGET TKPDVILRLLEQGKEPWLEEEVLGSGRAEKNNGDIGGQIWKPKDVKESLGG	3254

LTRP ID	Protein sequence	SEQ ID NO
	PSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPA GSPTSTEEGTSTEPSEGSAPGTSTEPSEQEIKRINKIRRLVKDSNTKKA GKTGPMKTLVVRVMTPLRERLENLRKKPENI PQPISNTSRANLNKLLTD YTEMKKAILHVYWEFFQKDPVGLMSRVAQPASKKIDQNKLPKPEMDEKGNL TTAGFACSQCGQPLFVYKLEQVSEKGAAYTNYFGRCNVAEHEKLI LLAQL KPEKDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYAS GPVGKALSDACMGTIASFLSKYQDIIIEHQKVVKGNQKRLESRELAKGE NLEYPSVTLPQPHTKEGVDAAYNEVIARVRMWNLNWQKCLKLSRDDAKP LLRLKGFPSFPLVERQANEVDWDMVCNVKKLINEKKEDGKVFQNLGAY KRQEALRPYLSSEEDRKKGKKFARYQLGDL LHLLEKKHGEDWGKVYDEAW ERIDKKVEGLSKHIKLEEEERSEDAQSKAALTDWLRAKASFVIEGLKEAD KDEFRCCELKQKWDLRGKPF AIEAENSILDISGFSKQYNCAFIWQKD GVKKLNLYLIINYFKGGKLRFKKIKPEAFANRFYTVINKKSGEIVPMEV NFNFDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGRVIEKTLY NRRTRQDEPALFVALTFERREVLDSNPKMNLIGVARGENIPAVIALTD PEGCPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRK YASKAKNLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGRKFMA ERQYTRMEDWLTAKLAYEGLSKTYLSKTLAQYTSKTC SNCGFTITSADYD RVLEKLLKKTATGWMTTINGKELKVEGQITYYNNRYKRQNVVKDLSVELDRL SEESVNNDISSWTGRSGEALSLLKKRFSHRPVQEKFVCLNCGFETHAAE QAALNIARSWLFLRSQYKQYQTNKTTGNTDKRA FVETWQSFYRKKLKEV WKPAVRSGGSGGGSTSPKKKRKV	
LTRP5- ZIM3 + ADD - CD	MAPKKKRKVS RMRERLVYEVQRKCRNIEDICISCGSLNVTLEHPLFIGGMC QNCKNCFLECAQYDDGYQSYCTICCGGREVLMCGNNCCRCFCVECV LLVGPGAAQAAIKEDPWNCYMGHKGTYGLLRREDWPSRLQMFFANSSG NSNANSRGPSFSSGLVPLSLRGSHMGPMIYKTVSAWKRQPVRLSLFRN IDKVLKSLGFLSGLSGSGGGTLYVEDVTNVVRRDVEKWGPFDLVYGSTQ PLGSSCDRCPGWYMFQFHRILQYALPRQESQRPFFWIFMDNLLLTEDDQE TTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEYLO AQVRSRSLDAPKVDLLVKNCLPLREYFKYFSQNSLPLGSGGGGMNNSQ GRVTFEDVTNFTQGEWQRLNPEQRNLYRDVMLENYSNLVSVGQGETTKP DVILRLQKQKEPWLEEEVLGSGRAEKNGDIGQIWKPKDVKESLGGPSS GAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSP TSTEEGTSTEPSEGSAPGTSTEPSEQEIKRINKIRRLVKDSNTKKAGKT GPMKTLVVRVMTPLRERLENLRKKPENI PQPISNTSRANLNKLLTDYTE MKKAILHVYWEFFQKDPVGLMSRVAQPASKKIDQNKLPKPEMDEKGNLTTA GFACSQCGQPLFVYKLEQVSEKGAAYTNYFGRCNVAEHEKLI LLAQLKPE KDSDEAVTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGPV GKALSDACMGTIASFLSKYQDIIIEHQKVVKGNQKRLESRELAKENLE YPSVTLPQPHTKEGVDAAYNEVIARVRMWNLNWQKCLKLSRDDAKPLLR LKGFPSPFPLVERQANEVDWDMVCNVKKLINEKKEDGKVFQNLGAYKRQ EALRPYLSSEEDRKKGKKFARYQLGDL LHLLEKKHGEDWGKVYDEAWERI DKKVEGLSKHIKLEEEERSEDAQSKAALTDWLRAKASFVIEGLKEADKDE FCRCELKQKWDLRGKPF AIEAENSILDISGFSKQYNCAFIWQKDGVK KLNL YLIINYFKGGKLRFKKIKPEAFANRFYTVINKKSGEIVPMEVNFN FDDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGRVIEKTYNRR TRQDEPALFVALTFERREVLDSNPKMNLIGVARGENIPAVIALTDPEG CPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYAS KAKNLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGRKFMAERQ YTRMEDWLTAKLAYEGLSKTYLSKTLAQYTSKTC SNCGFTITSADYDRVL EKLKKTATGWMTTINGKELKVEGQITYYNNRYKRQNVVKDLSVELDRLSEE SVNNDISSWTGRSGEALSLLKKRFSHRPVQEKFVCLNCGFETHAAEQAA LNIARSWLFLRSQYKQYQTNKTTGNTDKRA FVETWQSFYRKKLKEVWKPA AVRS GGSGGGSTSPKKKRKV DKKDDDKYKDDDDK	3255
LTRP5- ZIM3 + ADD + PWWP	MAPKKKRKVS RMTKAADDEPEYEDGRGFGIGELVWGKLRGFSWWPGRIVS WWMTRGRSRAAEGTRWVMWFGDGKFSVVCVEKLMPLSSFCSAFHQATYNKQ PMYRKATYEV LQVASSRAGKLF PACHDSDES DSGKAVEVQNKQMI EWALG GFQPSGPKGLEPPEEEKNPYKEVYTDMMWVEPEAAAYAPPPPAKKPRKSTT	3256

LTRP ID	Protein sequence	SEQ ID NO
- CD	EKPKVKEIIDERTRERLVYEVQRQKCRNIEDICISCGSLNVTLEHPLFIGG MCQNCKNCFLECAQYDDDGYSYCTICCGGREVLNCGNNCCRCFCVEC VDLLVGPGAAQAAIKEDPWNCYMGHKGTYGLRRREDWPSRLQMFFANS SGNNSNANSRGPSFSSGLVPLSLRGSHMGPMETKYTVSAWKRQPVRLSLF RNIDKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRDVEKWGPFDLVYGS TQPLGSSCDRCPGWYMFQFHRIQYALPRQESQRPFWWIFMDNLLLTEDD QETTTTFLQTEAVTLQDVGRDYQNAMRVWSNIPGLKSKHAPLTPKEEY LQAQVRSRSLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGSGGGMNN SQGRVTFEDVTVNFTQGEWQRLNPEQRNLYRDVMLENYSNLVSVGQGETT KPDVILRLQKGKEPWLLEEEVLGSGRAEKNIGDGGQIWKPKDVKESLGGP SSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAG SPTSTEEGTSTEPSEGSAPGTSTEPSEGEIKRINKIRRRLVKDSNTKKAG KTGPMTLLVRVMTPLRLERLENLRKKPENIPQPISTNSRANLNKLLTDY TEMKKAILHVVWEEFQKDPVGLMSRVAQPAKKIDQNKLPKPEMDEKGNLT TAGFACSQCGQPLFVYKLEQVSEKQKAYTNYFGRCNVAEHEKLILLAQLK PEKDSDEAVTYSLGKFGQALDFYSIHVTKESTHPVKPLAQIAGNRYASG PVGKALSDACMGTIASFLSKYQDIIIEHQKVVKGNGKRLESRELAKEN LEYPSVTLPPQPHTKEGVDAYNEVIARVMWVNLNLWQKLKLSRDDAKPL LRLKGFPSPFLVERQANEVDWDMVCNVKKLINEKKEDGKVFQNLAGYK RQEARPYLSSEEDRKKGKKFARYQLGDLHLHLEKKHGEDWGKVYDEAWE RIDKKVEGLSKHIKLEEEERRSEDAQSKAALTDWLRAKASFVIEGLKEADK DEFRCCELKLQKWYGLRGKPFPAIEAENSILDISGFSKQYNCAFIWQKDG VKKLNLYLIINYFKGGKLRFKKIKPEAFANRFTVINKKSGEIVPMEVN FNFDDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGRVIEKTLYN RRTRQDEPALFVALTFERREVLDSNIPMNILGVARGENIPAVIALTDP EGCPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKY ASKAKNLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFGRQGRKRTFMAE RQYTRMEDWLTAKLAYEGLSKTYLSKTLAQYTSKTCSNCGFTITSADYDR VLEKLKKTATGWMTTINGKELKVEGQITYYNRYKRQNVVKDLSVELDRLS EESVNNDISSWTKGRSGEALSLLKKRFSHRPVQEKFVCLNCGFETHAAEQ AALNIARSWLFLRSQYKQYQTNKTTGNTDKRAVFVETWQSFYRKKLKEVW KPAVRSGSGGGGSTSPKKKRKV	

Table 38: Sequences of spacers used in constructs

Spacer ID	Target gene	Sequence	SEQ ID NO
0.0	Non-targeting	CGAGACGUAUUACGUCUCG	3232
7.37	<i>B2M</i>	GGCCGAGAUGUCUCGCUCCG	3137
7.160	<i>B2M</i>	UAAACAUCACGAGACUCUAA	3113
7.165	<i>B2M</i>	UCCCUAUGUCCUUGCUGUUU	3114

Transfection of HEK293T cells:

[0471] Seeded HEK293T cells were transiently transfected with 100 ng of LTRP5 variant plasmids, each containing an LTRP:gRNA construct encoding for LTRP5-ZIM3 or one of its alternative variations (FIG. 29; Table 37 for sequences), with the gRNA having either non-targeting spacer 0.0 or a *B2M*-targeting spacer (Table 38 for spacer sequences). Spacers 7.160 and 7.165 have been shown to be repress the *B2M* locus when used with LTRPs, but not when used with dXRs made up of a dCasX fused to the ZIM3 KRAB domain (data not

shown). Each construct was tested in triplicate. 24 hours post-transfection, cells were selected with 1 µg/mL puromycin for three days. Cells were harvested for repression analysis at day 5, day 12, day 21, and day 51 post-transfection. Briefly, repression analysis was conducted by analyzing B2M protein expression via HLA immunostaining followed by flow cytometry, as described in Example 4. In addition, HEK293T cells transiently transfected with LTRP5 variant plasmids and a *B2M*-targeting gRNA or non-targeting gRNA were harvested at seven days post-transfection for gDNA extraction for bisulfite sequencing to assess off-target methylation at the *VEGFA* locus, which was performed as described in Example 5.

Results:

[0472] The effects of incorporating the ADD domain with or without the PWWP domain into the LTRP5 molecule on increasing long-term repression of the target *B2M* locus and reducing off-target methylation were assessed. Variations of the LTRP5-ZIM3 molecule were evaluated with either a *B2M*-targeting gRNA (with spacer 7.37 and LTRP-specific spacers 7.160 and 7.165) or a non-targeting gRNA, and the results are depicted in the plots in FIGS. 30-33. FIG. 30 shows that use of spacer 7.37 resulted in saturating levels of repression activity when paired with LTRP5-ZIM3, LTRP5-ZIM3 + ADD, and LTRP5-ZIM3 + ADD + PWWP, rendering it more challenging to assess activity differences among the LTRP5 variants. However, the differences in repression activity among the LTRP5 variants were more pronounced when using spacers 7.160 and 7.165 (FIGS. 31 and 32). The data demonstrate that incorporation of the ADD domain resulted in a significant increase in long-term repression when paired with the two LTRP-specific spacers compared to the repression levels achieved with the other LTRP5-ZIM3 molecules. Meanwhile, incorporation of both ADD and PWWP domains did not result in improved repression of the *B2M* locus, especially compared to the baseline LTRP5-ZIM3 molecule. As anticipated, the two LTRP5 variants without the DNMT3A catalytic domain exhibited poor long-term repression. Furthermore, the results depicted in FIG. 33 indicate that addition of the ADD domain appeared to result in increased specificity, given the lower percentage of HLA-negative cells observed, relative to the baseline LTRP5-ZIM3 molecule.

[0374] Off-target CpG methylation at the *VEGFA* locus potentially mediated by the LTRP5 variants was assessed using bisulfite sequencing. FIG. 34 depicts the results from bisulfite sequencing, specifically showing the percentage of CpG methylation around the *VEGFA* locus. The results demonstrate that for all the *B2M*-targeting gRNAs, as well as the non-targeting gRNA, incorporation of the ADD domain into the LTRP5-ZIM3 molecule dramatically reduced the level of off-target methylation at the *VEGFA* locus (FIG. 34). FIG.

35 is a scatterplot mapping the activity-specificity profiles for the LTRP5-ZIM3 variants investigated in this example, where activity was measured as the average percentage of HLA-negative cells at day 21 when paired with spacer 7.160, and specificity was represented by the percentage of off-target CpG methylation at the *VEGFA* locus quantified at day 7 when paired with spacer 7.160. The scatterplot clearly shows that addition of the ADD domain significantly increased the activity of the LTRP5 molecule relative to the baseline ELX5 molecule without the ADD domain (FIG. 35).

[0473] The experiments demonstrate that inclusion of the DNMT3A ADD domain, but not inclusion of both the ADD and PWWP domains, improved repression activity and specificity of LTRP molecules. This enhancement of activity and specificity was observed with multiple gRNAs, demonstrating the significance of the incorporation of the ADD domain into LTRPs.

Example 8: Demonstration that inclusion of the ADD domain from DNMT3A into LTRPs enhances on-target activity and decreases off-target methylation

[0474] Experiments were performed to assess the effects of incorporating the ADD domain into LTRP molecules having configurations 1, 4, and 5 (FIG. 1), on long-term repression of the target locus and off-target methylation.

Materials and Methods:

Generation of LTRP constructs and plasmid cloning:

[0475] Plasmid constructs encoding for LTRP molecules having configurations 1, 4, and 5 with the ZNF10-KRAB or ZIM3-KRAB domain and the DNMT3A ADD domain were built using standard molecular cloning techniques. Sequences of the resulting LTRP molecules are listed in Table 39, which also shows the abbreviated construct names for a particular LTRP molecule (e.g., LTRP1.A, 1.B). FIG. 36 is a schematic that illustrates the general architectures of LTRP molecules with the ADD domain incorporated for LTRP configuration 1, 4, and 5. Sequences encoding the LTRP molecules also contained a 2x FLAG tag. Plasmids also harbored sequences encoding gRNA scaffold 174 having either a spacer targeting the endogenous *B2M* locus or a non-targeting control (spacer sequences listed in Table 38).

Table 39: Protein sequences of the various LTRP1, 4, and 5 variants assayed in this example

LTRP #	Domains	Protein Sequence	SEQ ID NO
LTRP1	ZNF10-KRAB, DNMT3A ADD, DNMT3A CD, DNMT3L Interaction (LTRP1.D)	MAPKKKRRKVS RMERLVYEVQRQCRNIEDICISCGSLNVTLEHPLFIGGMCQNCKNCFLECAQQYDDDDGYQSYCTICCGGR EVL MCGN NNCRCFCVCEVDLLVGPAAQAAIKEDPWNCY MCGHKGT YGLLRREDWPSRLQMFFANNHDQEFDPKVPYP PVPAEKRKPIRVLSLFDGIATGLLVKDLGIQVDRIASE VCEDSITVGMVRHQGKIMYVGDVRSVTQKHQEWGPFDLV IGGSPCNDLSIVNPARKGLYEGTGRLLFFEFYRLLDHARPK EGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKE VSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRI AKFSKVRTITTRSNSIKQKQDQHFVFMNEKEDI LWCTEM ERVFGFPVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAP LKEYFACVSSGNSNANSRGPSSFSSGLVPLSLRGSHMGPM E IYKTVSAWKRPVRVLSLFRNIDKVLKSLGFLES GSGSGG GTLKYVEDVTNVVRDVEKWGPFDLVYGSTQPLGSSCDRC PGWYMFQFHRI LQYALPRQESQRPF FFI FMDNLLLTEDDQ ETTRFLQT EAVTLQDVRGRDYQNAMRVWSNIPGLKS KHA PLTPKEEY LQAQVRSRSLDAPKVDLLVKNCLLPLEYF KYFSQNSLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSESA TPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGGSGGSAQEIKRINKIRRLVKDSNTKKA GKTGPMKTL LVRVMT PDLRERLENLRKKPENIPQPI SNTS RANLNKLLTDYTEMKKAILHVYWE EFQKDPVGLMSRVAQP ASKKIDQNKLP EMDEKGNLT TAGFACSQCGQPLFVYKLE QVSEKGKAYTNYFGRCNVAEHEKLILLALQKPEKDSDEAV TYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYAS GPVGKALSDACMGTIASF LSKYQDIIIEHQVVKGNQKRL ESLRELAKENLEYP SVTLPPQPHTKEGV DAYNEVIARVR MWVNLNLWQKLKLSRDDAKPLRLKGFPSFPLVERQANEV DWWDMVCNVKKLINEKKEDGKVFQNL LAGYKRQEALRPYL SSEEDRKKGKKFARYQLGDLLHLLEKKHGEDWGKVYDEAW ERIDKKVEGLSKHIKLEEEERSEDAQSKAALTDWLRKAS FVIEGLKEADKDEFRCCELKQLQWYGD LRGKPF AIEAENS ILDISGFSKQYNCAFIWQKDG VKKLNL YLI INYFKGKLR FKKIKPEAF EANRFYTVINKKS GEIVPMEVNFNFDDPNLI ILPLAFGKRQGREFIWNDLLSLETGSLKLANGRVIEKTL Y NRRTRQDEPALFVALTFERREVLDSN IKPMNLIGVARGE NIPAVIALTDPEGCPLSRFKDSLGNP THILRIGESYKEKQ RTIQAKKEVEQRAGGYSRKYASKAKNLADDMVRNTARDL LYYAVTQDAMLI FANLSRGFGRQGRKTFMAERQYTRMEDW LTAKLAYEGLSKTYLSKTLAQYTSKTC SNCGFTITSADYD RVLEKLKKTATGWM TTINGKELKVEGQITYYNYKRQNVV KDLSVELDRLSEESVNNDISSWTKGRSGEALSLLKKRFSH RPVQEK FVCLNCGFETHAAEQ AALNIARSWLFLRSQEYKK YQTNKTTGNTDKRAFVETWQS FYRKKLKEVWKPAVRSGGS GGGSTSM DAKSLTAW SRTLVTFKDV FVDF TREEWKLLDTA QQIVYRNVMLENYKNLVSLGYQLTKPDVILRLLEKGEEPT S PKKKRRKV	3257
	ZIM3-KRAB, DNMT3A ADD, DNMT3A CD, DNMT3L Interaction (LTRP1.C)	MAPKKKRRKVS RMERLVYEVQRQCRNIEDICISCGSLNVTLEHPLFIGGMCQNCKNCFLECAQQYDDDDGYQSYCTICCGGR EVL MCGN NNCRCFCVCEVDLLVGPAAQAAIKEDPWNCY MCGHKGT YGLLRREDWPSRLQMFFANNHDQEFDPKVPYP PVPAEKRKPIRVLSLFDGIATGLLVKDLGIQVDRIASE VCEDSITVGMVRHQGKIMYVGDVRSVTQKHQEWGPFDLV IGGSPCNDLSIVNPARKGLYEGTGRLLFFEFYRLLDHARPK EGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKE	3258

LTRP #	Domains	Protein Sequence	SEQ ID NO
		VSAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRI AKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEM ERVFGFPVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAP LKEYFACVSSGNSNANSRGPFSFSSGLVPLSLRGSHMGPM IYKTVSAWKRPVRVLSLFRNIDKVLKSLGFLESGSGSGG GTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRC PGWYMFQFHRILQYALPRQESQRPFFWI FMDNLLLTEDDQ ETTTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHA PLTPKEEEYLQAQVRSRSLDAPKVDLLVKNCLLPLREYF KYFSQNSLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSESA TPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAGGSAQEI KRINKIRRRLVKDSNTKKA GKTGPMKTL LVRVMT PDLRERLENLRKKPENIPQPI SNTS RANLNKLLTDYTEMKKAILHVYWEEFQKDPVGLMSRVAQP ASKKIDQNKLPKEMDEKGNLT TAGFACSQCQGPLFVYKLE QVSEK GKAYTNYFGRCNVAEHEKLILLAQLKPEKDSDEAV TYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYAS GPVGKALSDACMGTIASFLSKYQDIIIEHQKVVKGNQKRL ESLRELAGKENLEYPSTLPPQPHTKEGVDAYNEVIARVR MWVNLNLWQKLKLSRDDAKPLLRLKGFPSPFLVERQANEV DWWDMVCNVKKLINEKKEDGKVFQNLAGYKRQEALRPYL SSEEDRKKGKKFARYQLGDLLLHLEKKHGEDWGKVYDEAW ERIDKKVEGLSKHIKLEEERRSEDAQSKAALTDWLRKAS FVIEGLKEADKDEFRCCELKLQKWYGDLRGKPF AIEAENS ILDISGFSKQYNCAFIWQKDGVKKLNLYLIINYFKGGKLR FKKIKPEAFEANRFYTVINKKS GEIVPMEVNFNDDPNLI ILPLAFGKRQGREFIWNDLLSLETGSLKLANGRVIEKTL NRRTRQDEPALFVALTFERREVLDSNIPKMNLI GVARGE NIPAVIALTDPEGCP LSRFKDSLGNP THILRIGESYKEQ RTIQAKKEVEQRRAGGYSRKYASKAKNLADDMVRNTARL LYYAVTQDAMLI FANLSRGFGRQ GKRTFMAERQYTRMEDW LTAKLAYEGLSKTYLSKTLAQYTSKTC SNCGFTITSADYD RVLEKLKKTATGWM TTINGKELKVEGQITYYNRYKRQNVV KDLSVELDRLSEESVNNDISSWTKGRSGEALSLLKKRFSH RPVQEKFVCLNCGFETHAAEQ AALNIARSWFLRSQYKK YQTNKTTGNTDKRA FVETWQSFYRKKLKEVWKPAVRSGGS GGGSTSMNNSQGRVTFEDVTVNFTQGEWQRLNPEQRNLYR DVMLENYSNLVSVGQGETTKPDVILRLEQKKEPWLEEEEV LGS GRAEKN GDI GGQIWKP KDVKESLTS PKKKRKV	
	ZNF10-KRAB, DNMT3A CD, DNMT3L Interaction (LTRP1.B)	MAPKKKRVSRVNGSGSGGMNHDQEFDPKVPVPPVPAEK RKPIRVLSLFDGIATGLLVLDLGIQVDRYIASEVCEDSI TVGMVRHQGKIMYVGDVRSVTQKHIEWGPFDLVI GGS PCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPKGDDRP FFWLFENVVAMGVSDKRDISRFLSNPVMIDAKEVSAHR ARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSK VRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGF PVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFA CVSSGNSNANSRGPFSFSSGLVPLSLRGSHMGPMIYKTVS AWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGT LKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRC PGWYMFQFHRILQYALPRQESQRPFFWI FMDNLLLTEDDQ ETTTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKH APLTPKEEEYLQAQVRSRSLDAPKVDLLVKNCLLPLREYF KYFSQNSLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSES ATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGSAGGSAQEI KRINKIRRRLVKDSNTKKA GKTGPMKTL LVRVMT PDLRERLENLRKKPENIPQPI SNTS RANLNKLLTDYTEMKKAILHVYWEEFQKDPVGLMSRVAQP ASKKIDQNKLPKEMDEKGNLT TAGFACSQCQGPLFVYKLE	3259

LTRP #	Domains	Protein Sequence	SEQ ID NO
		QNKLPKPEMDEKGNLTTAGFACSQCGQPLFVYKLEQVSEKG KAYTNYFGRCNVAEHEKLILLAQLKPEKDSDEAVTYSLGK FGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGPVGKA LSDACMGTIASFLSKYQDIIIEHQKVVKGNQKRLESLREL AGKENLEYPSVTLPPQPHTKEGVDAYNEVIARVRMWNLN LWQKLKLSRDDAKPLRLKGFPSFPLVERQANEVDWDMV CNVKKLINEKKEDGKVFQNLAGYKRQEALRPYLSSEEDR KKGKKFARYQLGDLLHLEKKHGEDWGKVYDEAWERIDKK VEGLSKHIKLEEERRSEDAQSKAALTDWLRKASFVIEGL KEADKDEFRCCELKLQKWYGDLRGKPFALAEAEENSILDISG FSKQYNCAFIWQKDGVKKLNLYLIINYFKGGKLRFKKIKP EAFEANRFYTVINKKSGEIVPMEVNFNFDDPNLIILPLAF GKRQGREFIWNDLLSLETGSLKLANGRVIEKTLYNRRTRQ DEPALFVALTFERREVLDSNIKPMNLIGVARGENIPAVI ALTDPEGCPLSRFKDSLGNPTHILRIGESYKEKQRTIQAK KEVEQRRAGGYSRKYASKAKNLADDMVRNTARDLLYYAVT QDAMLIIFANLSRGFGRQGRKTFMAERQYTRMEDWLTAKLA YEGLSKTYLSKTLAQYTSKTCNCGFTITSADYDRVLEKL KKTATGWMTTINGKELKVEGQITYYNRYKRQNVVKDLSVE LDRLSEESVNNDISSWTKGRSGEALSLLKKRFSHRPVQEK FVCLNCGFETHAAEQAALNIARSWLFLRSQYKYYQTNTKT TGNTDKRAFVETWQSFYRKKLKEVWKPAVRSGGSGGGSTS MDAKSLTAWSRITLVTFKDVFDFTREEWKLLDTAQQIVYR NVMLENYKNLVSLGYQLTKPDVILRLEKGEEPTSPKKKRK V	
	ZIM3-KRAB, DNMT3A CD, DNMT3L Interaction (LTRP1.A)	See sequence in Table 14	3066
LTRP4	ZNF10-KRAB, DNMT3A ADD, DNMT3A CD, DNMT3L Interaction (LTRP4.D)	MAPKKRKVSMDAKSLTAWSRITLVTFKDVFDFTREEWK LLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKG EPPGGSGGGSAMERLVYEVQRQCRNIEDICISCGSLNVTL EHPLFIGGMCQNCNCFLECAQYQDDDGYSYCTICCGGR EVLNCGNNCCRCFCVECDLLVGPAAQAAIKEDPWNCY MCGHKGTYGLLRREDWPSRLQMFANNHDQEFDPKVPY PVPAEKRPPIRVLSLFDGIATGLLVLDLGIQVDRIASE VCEDSITVGMVRHQGKIMYVGDVRSVTQKHQEWGPFDLV IGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPK EGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKE VSAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRI AKFSKVRTITTRSNSIKQGKDQHFVFMNEKEDILWCTEM ERVFGFPVHYTDVSNMSRLARQRLGRSWSPVIRHLFAP LKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMGPM IYKTVSAWKRQPVRLSLFRNIDKVLKSLGFLES GSGSGG GTLKYVEDVTNVVRDVEKWGPFDLVYGSTQPLGSSCDRC PGWYMFQFHRI LQYALPRQESQRPFFWI FMDNLLLTEDDQ ETTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHA PLTPKEEYLAQVRSRSLDAPKVDLLVKNCLLPLREYF KYFSQNSLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSESA TPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEQEIKRINKIRRLVKDSNTKKAGKTGPMKT LLVRVMTPLDLRERLENLRKKPENIPQPI SNTSRANLNKLL TDYTEMKAILHVYWEEFQKDPVGLMSRVAQPAKKIDQN KLKPEMDEKGNLTTAGFACSQCGQPLFVYKLEQVSEKGKA YTNYFGRCNVAEHEKLILLAQLKPEKDSDEAVTYSLGKFG QRALDFYSIHVTKESTHPVKPLAQIAGNRYASGPVGKALS DACMGTIASFLSKYQDIIIEHQKVVKGNQKRLESLREL	3260

LTRP #	Domains	Protein Sequence	SEQ ID NO
		KENLEYPSVTLPQPHTKEGVDAAYNEVIARVRMWNLNWLWQKLKLSRDDAKPLRLKGFPSFPLVERQANEVDWDMVCNVKKLINEKKEDGKVFQWQNLQYKQKQALRPYLSSEEDRKKGKKFARYQLGDLHLEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKLEEEERSEDAQSKAALTDWLRAKASFVIEGLKEADKDEFRCCELKQKQWYGDRLGKPFQAEAEENSILDISGFSKQYNCAFIWQKDGKVKLNLYLIINYFKGGKLRFKKIKPEAFANRFYTVINKKSGEIVPMEVNFNFDDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGRVIEKTLYNRRTRQDEPALFVALTFERREVLDSNNIKPMNLIGVARGENIPAVIALTDPEGCPLSRFKDSLGNPHTILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYASKAKNLADDMVRNTARDLLYYAVTQDAMLIIFANLSRGFGRQGRKTFMAERQYTRMEDWLTAKLAYEGLSKTYLSKTLAQYTSKTCSCNGFTITSADYDRVLEKLLKTATGWMTTINGKELKVEGQITYYNRYKQNVVKDLSVELDLSEESVNNDISSWTKGRSGEALSLLKKRFSHRPVQKEFVCLNCGFETHAAEQAALNIARSWLFLRSQYKQYQTNKTTGNTDKRAFVETWQSFYRKKLKEVWKPVRSGSGSGGSTSPKKRKV	
	ZIM3-KRAB, DNMT3A ADD, DNMT3A CD, DNMT3L Interaction (LTRP4.C)	MAPKKKRVSRMNSQGRVTFEDVTVNFTQGEWQRLNPEQRNLYRDVMLENYSNLVSVGGQGETTKPDVILRLEQGGKEPWL EEEVLGSGRAEKNGDIGGQIWKPKDVKESLGGSGGGSAMERLVYEVQRQKCRNIEDICISGSLNVTLEHPLFIGGMCQNCKNCFLECAQYQDDDGYSYCTICCGGREVLNCGNNCCRCFCVECVDLLVGPGAAQAAIKEDPWNCYMCYHKGTYGLLRREDWPSRLQMFANNDQEFDPKVPYPPVPAEKRPPIRVLSLFDGIATGLLVKDLGIQVDRIYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIEQWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLEFFEFYRLLHDARPKEGDDRPFFWLFE NVVAMGVSDKRDISRFLESNPVIMIDAKEVSAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQKQDQHFVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPFSFSSGLVPLSLRGSHMGPMETKTVSAWKRPVRVLSLFRNIDKVLKSLGFLESGLSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQESQRPFFWIFMDNLLLTEDDQETTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEYQLAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEQEIKRINKIRRRLVKDSNTKKAGKTGPMKTLVVRVMTPLDRELENLRKKPENIPQISNSTRANLNKLLTDYTEMKKAILHVYWEFQKDPVGLMSRVAQPAKKIDQNKLPKEMDEKGNLTTAGFACSQCQGPLFVYKLEQVSEKQAYTNYFGRCNVAEHEKLILLAQLKPEKSDAETYSLGKFGQALDFYSIHVTKESTHPVKPLAQIAGNRYASGPVGKALSDACMGTIASFLSKYQDIIIEHQKVVKGQKRLSRLRELAKENLEYPSVTLPQPHTKEGVDAAYNEVIARVRMWNLNWLWQKLKLSRDDAKPLRLKGFPSFPLVERQANEVDWDMVCNVKKLINEKKEDGKVFQWQNLQYKQKQALRPYLSSEEDRKKGKKFARYQLGDLHLEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKLEEEERSEDAQSKAALTDWLRAKASFVIEGLKEADKDEFRCCELKQKQWYGDRLGKPFQAEAEENSILDISGFSKQYNCAFIWQKDGKVKLNLYLIINYFKGGKLRFKKIKPEAFANRFYTVINKKSGEIVPMEVNFNFDDPNLIILPLAFGKRQGREFIWNDLLSLETGSLKLANGRVIEKTLYNRRTRQDEPALFVALTFEREVLDSNNIKPMNLIGVARGENIPAVIALTDPEGCPLSRFK	3261

LTRP #	Domains	Protein Sequence	SEQ ID NO
		DSLGNPTHILRIGESYKEKQRTIQAKKEVEQRRAGGYSRK YASKAKNLADDMVRNTARDLLYYAVTQDAMLIIFANLSRGF GRQGRKFMAERQYTRMEDWLTAKLAYEGLSKTYLSKTLA QYTSKTCSNCGFTITSADYDRVLEKLKKTATGWMTTINGK ELKVEGQITYYNYKRQNVVKDLSVELDRLSEESVNNDIS SWTKGRSGEALSLKKRFSHRPVQEKFVCLNCGFETHAAE QAAALNIARSWLFLRSQYKQYQTNKTTGNTDKRAFVETWQ SFYRKKLKEVWKPAVRSGSGGGGSTSPKKKRKV	
	ZNF10-KRAB, DNMT3A CD, DNMT3L Interaction (LTRP4.B)	MAPKKKRKVS RMDAKSLTAW SRTLVT FKVDFVDF TREEWK LLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKG EEPGGSGGGSAMNHDQEFDPKVPYPPVPAEKRKPIRVLSL FDGIATGLLVLDLGIQVDRIASEVCEDSITVGMVRHQG KIMYVGDVRSVTQKHIEWGPFDLVIGGSPCNDLSIVNPA RKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVV AMGVSDKRDISRFLSNPVMIDAKEVSAHRARYFWGNLP GMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNS IKQKGDQHFVPFMNEKEDI LWCTEMERVF GFVPVHYTDVSN MSRLARQRLGLGRSWSVPVIRHLFAPLKEYFACVSSGNSNA NSRGPSFSSGLVPLSLRGSHMGPMIYKTVSAWKRPVRV LSLFRNIDKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRR DVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRIQLQA LPRQESQRPFFWI FMDNLLLTEDDQETTTREFLQTEAVTLQ DVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEYLLQAQVR SRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGPSS GAPPSPSGSPAGSPTSTEEGTSESATPESGPGTSTEPSEG SAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEQEIKR INKIRRRRLVKDSNTKKAGKTGPMKTL LVVRVMT PDLRERLE NLRKKPENIPQPI SNTSRANLNKLLTDYTEMKKAILHVYW EEFAQKDPVGLMSRVAQPAASKIDQNKLPMEDEKGNLTTA GFACSQCGQPLFVYKLEQVSEKGA YTN YFGRCNVAEHK LILLAQLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTKES THPVKPLAQIAGNRYASGPVGKALSDACMGTIASFLSKYQ DIIIEHQKVVGKNQKRLES LRELAKENLEYPSVTLPQP HTKEGVDA YNEVIARVMVNLNLWQKLKLSRDDAKPLLR LKGFPSFPLVERQANEVDWDMVCNVKLINEKKEDGKVF WQNLAGYKRQEALRPYLSSEEDRKKGKKFARYQLGDLHLH LEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKLEEERSE DAQSKAALTDWLRKASFVIEGLKEADKDEFRCCELKLQK WYGDLRGKPF AIEAENSILDISGFSKQYNCAFIWQKDGVK KLNLYLIINYFKGGKLRFKIKPEAFEANRFYTVINKKSG EIVPMEVNFNFDPNLIILPLAFGKRQGREFIWNDLLSLE TGSCLKANGRVIEKTLYNRRTRQDEPALFVALTFERREVL DSSNIKPMNLIGVARGENIPAVIALTDPEGCPLSRFKDSL GNPTHILRIGESYKEKQRTIQAKKEVEQRRAGGYSRKYAS KAKNLADDMVRNTARDLLYYAVTQDAMLIIFANLSRGFGRQ GKRTFMAERQYTRMEDWLTAKLAYEGLSKTYLSKTLAQT SKTCSNCGFTITSADYDRVLEKLKKTATGWMTTINGKELK VEGQITYYNYKRQNVVKDLSVELDRLSEESVNNDISSWT KGRSGEALSLKKRFSHRPVQEKFVCLNCGFETHAAEQAA LNIARSWLFLRSQYKQYQTNKTTGNTDKRAFVETWQSFY RKKLKEVWKPAVRSGSGGGGSTSPKKKRKV	3262
	ZIM3-KRAB, DNMT3A CD, DNMT3L Interaction (LTRP4.A)	MAPKKKRKVS RVNGSGSGGGGMNHDQEFDPKVPYPPVPAEK RKPIRVLSLFDGIATGLLVLDLGIQVDRIASEVCEDSI TVGMVRHQGKIMYVGDVRSVTQKHIEWGPFDLVIGGSPC NDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRP FFWLFENVVAMGVSDKRDISRFLSNPVMIDAKEVSAHR ARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKV RTITTRSNSIKQKGDQHFVPFMNEKEDI LWCTEMERVF GF	3263

LTRP #	Domains	Protein Sequence	SEQ ID NO
		PVHYTDVSNMSRLARQRLLGSRWSVPVIRHLFAPLKEYFA CVSSGNSNANSRGPSFSSGLVPLSLRGSHMGPMIYKTVS AWKRQPVVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYV EDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMF QFHRILQYALPRQESQRPFFWI FMDNLLLTEDDQETTRF LQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKE EEYLQAQVRSRSLDAPKVDLLVKNCLLPLREYFKYFSQN SLPLGGSGGGMDAKSLTAWSRTLVTFKDVFDFTREEWKL LDTAQQIVYRNVMLENYKNLVS LGYQLTKPDVILRLEKGE EPGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGT STEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEP SEQEIKRINKIRRLVKDSNTKKAGKTGPMKTL LVRVMT DLRERLENLRKKPENIPQPI SNTSRANLNKLLTDYTEMKK AILHVYWEFQKDPVGLMSRVAQPASKKIDQNKLPemde KGNLT TAGFACSQCGQPLFVYKLEQVSEK KAYTNYFGRC NVAEHEKLILLAQLKPEKDSDEAVTYS LGKFGQRALDFYS IHVTKESTHPVKPLAQIAGNRYASGPVGKALSDACMGTTA SFLSKYQDIIIEHQKVVGKGNQKRLESRELAGKENLEYPS VTLPPQPHTKEGVDA YNEVIARVRMWVNLNLWQKLKLSRD DAKPLLRLKGFPSFPLVERQANEVDWDMVCNVKKLINEK KEDGKVFWQNLAGYKRQEALRPYLSSEEDRKKGKKFARYQ LGDLLLHLEKKHGEDWGKVYDEAWERIDKKVEGLSKHIKL EEERRSEDAQSKAALTDWLRKAS FVIEGLKEADKDEFCE CELKQLQWYGD LRGKPF AIEAENSILDISGFSKQYNCAFI WQKDGVKKLNLYLI INYFKGGKLRFKKIKPEAF EANRFYT VINKKSGEIVPMEVNFNFDDPNLII LPLAFGKRQGREFIW NDLLSLETGSLKLANGRVIEKTLYNRRTRQDEPALFVALT FERREVLDS SNIKPMNLIGVARGENIPAVIALTDPEGCP SRFKDSLGNP THILRIGESYKEKQRTIQAKKEVEQRAGG YSRKYASKAKNLADDMVRNTARDLLYAVTQDAMLI FANL SRGFGRQGRKFMAERQYTRMEDWLTAKLAYEGLSKTYLS KTLAQYTSKTC SNCGFTITSADYDRVLEKLKKTATGWM TINGKELKVEGQITYYNRYKRQNVVKDLSVELDRLSEESVN NDISSWTKGRSGEALSLLKKRFSHRPVQEKFVCLNCGFET HAAEQ AALNIARSWLF LRSQYKKYQTNKTTGNTDKRAFV ETWQS FYRKKLKEVWKPAVRSGGSGGGSTSPKKKRKV	
LTRP5	ZNF10-KRAB, DNMT3A ADD, DNMT3A CD, DNMT3L Interaction (LTRP5.D)	MAPKKRKRKVRMNSQGRVTFEDVTVNFTQGEWQRLNPEQ RNLYRDVMLENYSNLVSVGQGETTKPDVILRLEQGKEPWL EEEEVLGSGRAEKNGDIGGQIWKP KDVKESLGGSGGGSAM NHDQEFDPKVPVPVPAEKRP IRVLSLFDGIATGLLVLK DLGIQVDRIASEVCEDSITVGMVRHQKIMYVGDVRSVT QKHIQEWGPFDLVIGGSPCNDLSIVN PARKGLYEGRLEF FEFYRLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISR FLESNPVMIDAKEVSAHRARYFWGNLPGMNRPLASTVND KLELQECLEHGRIAKFSKVRTITTRSNSIKQGDQHFFVF MNEKEDILWC TEMERVFGFPVHYTDVSNMSRLARQRLLG SWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLV PLSLRGSHMGPMIYKTVSAWKRPVVRVLSLFRNIDKVLK SLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVY GSTQPLGSSCDRCPGWYMFQFHRILQYALPRQESQRPFFW IFMDNLLLTEDDQETTTTRFLQTEAVTLQDVRGRDYQNAMR VWSNIPGLKSKHAPLTPKEEYLQAQVRSRSLDAPKVDL LVKNCLLPLREYFKYFSQNSLPLGGPSSGAPPPSGGSPAG SPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTS TEEGTSTEPSEGSAPGTSTEPSEQEIKRINKIRRLVKDS NTKKAGKTGPMKTL LVRVMT PDLRERLENLRKKPENIPQP ISNTSRANLNKLLTDYTEMKKAILHVYWEFQKDPVGLMS RVAQPASKKIDQNKLPemdeKGNLT TAGFACSQCGQPLF	3264

LTRP #	Domains	Protein Sequence	SEQ ID NO
		VYKLEQVSEK GKAYTNYFGRCNVAEHEKLI LLAQLKPEKD SDEAVTYS LGKFGQ RALDFYSIHVTKESTHPVKPLAQIAG NRYASGPVGKALSDACMGTIASFLSKYQDII IEHQKVVKG NQKRLESLRELAKENLEYP SVTLPPQPHTKEGVDAYNEV IARVRMWNLN LWQKLKLSRDDAKPLLRKGFPSFPLVER QANEVDWDMVCNVKKLINEKKEDGKVFQNLAGYKRQEA LRPYLSSEEDRKKGKKFARYQLGDLLLHLEKKHGEDWGKV YDEAWERIDKKVEGLSKHIKLEEEERRSEDAQSKAALTDWL RAKASFVIEGLKEADKDEFRCCELKLQKWYGDLRGKPF AI EAENSILDISGFSKQYNCAFIWQKDGVKKLNLYLI INYFK GGKLRFKKIKPEAF EANRFYTVINKKS GEIVPMEVNFND DPNLIILPLAFGKRQGREFTWNDLLSLETGSLKLANGRVI EKTLYNRRT RQDEPALFVALTFERREVLDSNIKPMNLIG VARGENIPAVIALTDPEGCP LSRFKDSLGNPTHILRIGES YKEKQRTIQAKKEVEQRRAGYSRKYASKAKNLADDMVRN TARDLLYYAVTQDAMLI FANLSRGFGRQ GKRTFMAERQYT RMEDWLTAKLAYEGLSKTYLSKTLAQYTSKTC SNCGFTIT SADYDRVLEKLKKTATGWM TTINGKELKVEGQITYYNRKY RQNVVKDLSVELDRLSEESVNNDI SSWTKGRSGEALSLLK KRFSHRPVQEK FVCLNCGFETHAAEQ AALNIARSWLFLRS QEYKKYQTNKTTGNTDKRAFVETWQS FYRKKLKEVWKP AV RSGGSGGGSTSPKKKRKV	
	ZIM3-KRAB, DNMT3A ADD, DNMT3A CD, DNMT3L Interaction (LTRP5.C)	<i>See sequence in Table 37</i>	3132
	ZNF10-KRAB, DNMT3A CD, DNMT3L Interaction (LTRP5.B)	MAPKKKRKVS RVNGSGSGGGMNHDQEFDP PKVYPPVPAEK RKPIRVLSLFDGIATGLLV LKDLGIQVDRYIASEVCDSI TVGMVRHQGKIMYVGDVRSVTQKH IQEWGPFDLVIGGSPC NDLSIVNPARKGLYEGTGRLFF EYRLLHDARPKEGDDRP FFWL FENVVAMGVSDKRDISR FLESNPVMIDAKEVSAHR ARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKV RTITTRSNSIKQKQDQHFPVFMEKEDI LWCTEMERVFGE PVHYTDVSNMSRLARQRL LGRSWSVPVIRHLFAPLKEYFA CVSSGNSNANSRGPSFS SGLVPLSLRGS HMGPM E IYKTVS AWKRQPV RVLSLFRNIDKVLKSLGFLES GSGSGGGT LKYV EDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMF QFHRI LQYALPRQESQRPF FWFIMDNLLLTEDDQETTRF LQTEAVTLQDV RGRDYQ NAMRVWSNIPGLKSKHAPLTPKE EEYLQAQVRSR SKLDAPKVDLLVKNCLLPLREYFKYFSQN SLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGP GTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTST EPSEGGSGGSAQEIKRINKIRRLV KDSNTKKAGKTGPM KTL LVRVMT PDLRERLENLRKKPENI PQPISNTSRANLNK LLTDYTEM KKA I LHVYWE EFQKDPVGLMSRVAQ PASKKID QNKLP EMDEKGNLTTAGFACSQCGQPLFVYKLEQVSEKG KAYTNYFGRCNVAEHEKLI LLAQLKPEKDSDEAVTYS LGK FGQ RALDFYSIHVTKESTHPVKPLAQIAGNRYASGPVGKA LSDACMGTIASFLSKYQDII IEHQKVVKGNQKRLESLREL AGKENLEYP SVTLPPQPHTKEGVDAYNEVIARVRMWNLN LWQKLKLSRDDAKPLLRKGFPSFPLVERQANEVDWDMV CNVKKLINEKKEDGKVFQNLAGYKRQEA LRPLYLSSEEDR KKGKKFARYQLGDLLLHLEKKHGEDWGKVYDEAWERIDKK VEGLSKHIKLEEEERRSEDAQSKAALTDWLRKASFVIEGL KEADKDEFRCCELKLQKWYGDLRGKPF AIEAENSILDISG FSKQYNCAFIWQKDGVKKLNLYLI INYFKGGKLRFKKIKP	3265

LTRP #	Domains	Protein Sequence	SEQ ID NO
		EAFEANRFYTVINKKSGETVPMENVNFDDPNLIILPLAF GKRQGREFIWNDLLSLETGSLKLANGRVIEKTLYNRRTRQ DEPALFVALTFERREVLDSNFKPMNLIGVARGENIPAVI ALTDPEGCPKSRFKDSLGNPTHILRIGESYKEKQRTIQAK KEVEQRRAGGYSRKYASKAKNLADDMVRNTARDLLYYAVT QDAMLI FANLSRGFGRQGRKFMAERQYTRMEDWLTAKLA YEGLSKTYLSKTLAQYTSKTCNCGFTTTSADYDRVLEKL KKTATGWMTTINGKELKVEGQITYYNRYKRQNVVKDLSVE LDRLSEESVNNDISSWTKGRSGEALSLLKKRFSHRPVQEK FVCLNCGFETHAAEQALNIARSWLFLRSQYKQYQTNKT TGNTDKRAFVETWQSFYRKKLKEVWKPAPVRSGGSGGGSTS SQGSVTFRDVAIDFSQEEWKWLQPAQRDLRYRRVMLENYGH LVSLGLSISKPDVVSLLLEQKPEWLGKRNVRDLFSASES STSPKKKRKV	
	ZIM3-KRAB, DNMT3A CD, DNMT3L Interaction (LTRP5.A)	See sequence in Table 37	3131

Transfection of HEK293T cells:

[0476] Seeded HEK293T cells were transiently transfected with 100 ng of LTRP variant plasmids, each containing an LTRP:gRNA construct encoding for an LTRP molecule (Table 39; FIG. 36), with the gRNA having either non-targeting spacer 0.0 or a *B2M*-targeting spacer (Table 38). Each construct was tested in triplicate. 24 hours post-transfection, cells were selected with 1 µg/mL puromycin for 3 days. Cells were harvested for repression analysis at day 8, day 13, day 20, and day 27 post-transfection. Briefly, repression analysis was conducted by analyzing B2M protein expression via HLA immunostaining followed by flow cytometry, as described in Example 4. In addition, cells were also harvested on day 5 post-transfection for gDNA extraction for bisulfite sequencing to assess off-target methylation at the non-targeted *VEGFA* locus, which was performed using similar methods as described in Example 5.

Results:

[0477] The effects of incorporating the ADD domain into the LTRP molecules having configurations 1, 4, or 5, with either a ZNF10 or ZIM-KRAB, on long-term repression of the *B2M* locus and off-target methylation were evaluated. LTRP molecules were tested with either a *B2M*-targeting gRNA or a non-targeting gRNA, and the results are depicted in the plots in FIGS. 37-47. The data demonstrate that incorporation of the ADD domain into the LTRP molecules clearly resulted in a substantial increase in B2M repression across all the time points for all LTRP configurations containing the ZIM3-KRAB when using spacer

7.160 (FIG. 37), and similar findings were observed when using spacers 7.165 and 7.37 (data not shown). FIG. 38 shows the resulting B2M repression upon use of LTRP5 containing either the ZNF10 or ZIM3-KRAB when paired with a gRNA with spacer 7.160; the data demonstrate that including the ADD domain increased durable B2M repression overall, with LTRP5-ZIM3 + ADD having a higher activity compared with that of LTRP5-ZNF10 + ADD. Similar time course findings were observed for LTRP1 and LTRP4 and the other two spacers (data not shown). FIG. 39 shows the resulting B2M repression upon use of LTRP5 containing the ZIM3-KRAB when paired with any of the three *B2M*-targeting gRNAs, and the data demonstrate that inclusion of the ADD domain resulted in higher B2M repression overall. Similar time course findings were also observed for LTRP1 and LTRP4 (data not shown).

[0478] FIGS. 40-42 show the resulting B2M repression at the day 27 time point for all the LTRP configurations and gRNAs tested. The results show that the increase in B2M repression was more prominent with use of the sub-optimal spacers 7.160 and 7.165 compared to use of spacer 7.37. Furthermore, use of LTRP1 and LTRP5, which contained the DNMT3A and DNMT3L domains on the N-terminus of the molecule, resulted in the highest increase in B2M repression upon addition of the DNMT3A ADD domain (FIGS. 40-42). Use of LTRP4, which harbored the DNMT3A/3L domains 3' to the KRAB domain and 5' to the dCasX, resulted in lower activity gains, which may be attributable to a decreased ability of the ADD domain to interact with chromatin properly.

[0479] The specificity of LTRP molecules was determined by profiling the level of CpG methylation at the *VEGFA* gene, an off-target locus, using bisulfite sequencing, and the data are illustrated in FIGS. 43-51. The data demonstrate that inclusion of the DNMT3A ADD domain resulted in a substantial decrease in off-target methylation of the *VEGFA* locus across all conditions tested (FIGS. 43-45). Notably, the increased specificity mediated by the inclusion of the ADD domain was most prominent with the LTRP1 and LTRP5 configurations, both of which harbored the DNMT3A/3L domains on the N-terminal end of the molecule. Interestingly, LTRP molecules containing the ZIM3-KRAB domain led to stronger off-target methylation of the *VEGFA* locus. Furthermore, use of LTRP4 and 5 configurations, even in the absence of an ADD domain, resulted in higher specificity compared to use of the LTRP1 configuration. Compared to LTRP1-ZIM3 and LTRP4-ZIM3 configurations, inclusion of the ADD domain into LTRP5-ZIM3 resulted in the lowest off-target methylation.

[0480] FIGS. 46-51 are a series of scatterplots mapping the activity-specificity profiles for the various LTRP molecules, where activity was measured as the average percentage of

HLA-negative cells at day 27, and specificity was determined by the percentage of off-target CpG methylation at the *VEGFA* locus at day 5. The data demonstrate that across all three *B2M*-targeting spacers tested, inclusion of the ADD domain resulted in increased on-target B2M repression and decreased off-target methylation at the *VEGFA* locus. LTRP molecules having configuration 1 and 5 configurations exhibited the greatest increases in activity and specificity at each spacer tested.

[0481] The results of the experiments discussed in this example support the findings in Example 7, in that the data demonstrate that inclusion of the DNMT3A ADD domain enhances both the strength of repression at early timepoints and the heritability of silencing across cell divisions, as well as decreases the off-target methylation incurred by the DNMT3A catalytic domain in the LTRP molecules. The data also confirm that different LTRP configurations have intrinsic differences in specificity, which can be exacerbated by use of a more potent repressor domain. This decrease in specificity can be mitigated by inclusion of the DNMT3A ADD domain, which also can lead to greater on-target repression overall. The gains in repression activity are believed to be mediated by the function of the DNMT3A ADD domain to recognize H3K4me0 and subsequent recruitment to chromatin. The gains in specificity are believed to be mediated via the function of the DNMT3A ADD domain to induce allosteric inhibition of the catalytic domain of DNMT3A in the absence of binding to H3K4me0. The results also highlight that positioning of the ADD domain in the different configurations tested is important to achieve the strongest gains in both specificity and activity of LTRP molecules.

Example 9: Demonstration that silencing of a target locus mediated by LTRP molecules is reversible using a DNMT1 inhibitor

[0482] Experiments were performed to demonstrate that durable repression of a target locus mediated by LTRP molecules is reversible, such that treatment with a DNMT1 inhibitor would remove methyl marks to reactivate expression of the target gene.

Materials and Methods:

Generation of LTRP constructs and lentiviral plasmid cloning:

[0483] Lentiviral plasmid constructs coding for an LTRP molecule were built using standard molecular cloning techniques. These constructs comprised of sequences coding for catalytically-dead CasX protein 491 (dCasX491), a KRAB domain from ZNF10 or ZIM3, and the catalytic domain and interaction domain from DNMT3A (D3A) and DNMT3L (D3L) respectively. Briefly, constructs were ordered as oligonucleotides and assembled by overlap

extension PCR followed by isothermal assembly. The resulting plasmids contained constructs positioned in varying configurations to generate an LTRP molecule. Sequences encoding the LTRP molecules also contained a 2x FLAG tag. Plasmids also harbored sequences encoding gRNA scaffold variant 174 having either a spacer targeting the endogenous *B2M* locus or a non-targeting control. These constructs were all cloned upstream of a P2A-puromycin element on the lentiviral plasmid. Cloned and sequence-validated constructs were mid-prepped and subjected to quality assessment prior to transfection in HEK293T cells.

[0484] LTRP5 containing the ZIM3-KRAB domain and CasX variant 491 were used in this experiment. A *B2M*-targeting gRNA with scaffold 174 containing spacer 7.37 (SEQ ID NO: 3137) or a non-targeting gRNA containing spacer 0.0 (SEQ ID NO: 3232) were used in this experiment.

Transfection of HEK293T cells:

[0485] HEK293T cells were transfected with 100 ng of a plasmid containing a construct encoding for either CasX 491 or LTRP5 containing the ZIM3-KRAB domain with a *B2M*-targeting gRNA or non-targeting gRNA and cultured for 58 days. These transfected HEK293T cells were subsequently re-seeded at ~30,000 cells well of a 96-well plate and were treated with 5-aza-2'-deoxycytidine (5-azadC), a DNMT1 inhibitor, at concentrations ranging from 0 μ M to 20 μ M. Six days post-treatment with 5-azadC, cells were harvested for *B2M* silencing analysis at day 5, day 12, and day 21 post-transfection. Briefly, repression analysis was conducted by analyzing *B2M* protein expression via HLA immunostaining followed by flow cytometry, as described in Example 4. Treatments for each dose of 5-azadC for each experimental condition were performed in triplicates.

Results:

[0486] The plot in FIG. 52 shows the percentage of transfected HEK293T cells treated with the indicated concentrations of 5-azadC that expressed the *B2M* protein. The data demonstrate that 5-azadC treatment of cells transfected with a plasmid encoding LTRP5-ZIM3 with the *B2M*-targeting gRNA resulted in a reactivation of the *B2M* gene (FIG. 52). Specifically, ~75% of treated cells exhibited *B2M* expression with 20 μ M 5-azadC, compared to the 25% of cells with *B2M* expression at 0 μ M concentration (FIG. 52). Furthermore, 5-azadC treatment of cells transfected with a plasmid encoding CasX 491 with the *B2M*-targeting gRNA did not exhibit reactivation of the *B2M* gene. FIG. 53 is a plot that juxtaposes *B2M* repression activity with gene reactivation upon 5-azadC treatment. The data show *B2M* repression post-transfection with either CasX 491 or LTRP5-ZIM3 with the *B2M*-targeting gRNA, resulting in ~75% repression of *B2M* expression by day 58; however, *B2M*

expression is increased upon 5-azadC treatment (FIG. 53). As anticipated, 5-azadC treatment of cells transfected with either CasX 491 or LTRP5-ZIM3 with the non-targeting gRNA did not demonstrate repression or reactivation (FIGS. 52-53).

[0487] The experiments demonstrate reversibility of LTRP-mediated repression of a target locus. By using a DNMT1 inhibitor to remove methyl marks implemented by LTRP molecules, the silenced target gene was reactivated to induce expression of the target protein.

Example 10: Exemplary sequences of LTRP fusion proteins

[0488] Table 40 provides component sequences used in exemplary full-length LTRP fusion proteins in configurations 1, 4, or 5 with the ADD domain (FIG. 1), with human ZIM3 or ZNF10 KRAB domains or one of the top nine most effective repressor domains: DOMAIN_7694, DOMAIN_10123, DOMAIN_15507, DOMAIN_17905, DOMAIN_20505, DOMAIN_26749, DOMAIN_27604, DOMAIN_29304, and DOMAIN_30173. In Table 40, the components are listed in order from N- to C-terminus for the respective LTRP configuration.

Table 40: Exemplary protein sequences of LTRP fusion proteins

LTRP #	Components	Domains	AA sequence	SEQ ID NO
LTRP1 with ADD domain	START codon + NLS + buffer sequence		MAPKKKRKVS	3269
	START codon + DNMT3A ADD domain		MERLVYEVRQKCRNIEDICISCGSLNVTLEHPLFIGGMCQNCKNCFLECAQYDDDGYSYCTICCGGREVLNCGNNCCRCFCVECVDLLVGPGAAQAAIKEDPWNCYMGHKGTYGLLRREDWPSRLQMFFAN	3270
	DNMT3A catalytic domain		NHDQEFDPKVPVPPVPAEKRPPIRVLSLFDGIATGLLVKDLGIQVDRIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFE NVVAMGVSDKRDISRFLSNPVMIDAKEV SAAHRARYFWGNLPGMNRPLASTVNDKLEQECLEHGRIAKFSKVRTITTRSNSIKQG KDQHFPVFMNEKEDIWCTEMERVFGFPVHYTDVSNMSRLARQLLGRSWSVPVIRHLFAPLKEYFACV	126
	Linker (L2)		SSGNSNANSRGPFSFSSGLVPLSLRGSH	122
	DNMT3L interaction domain		MGPMEIYKTVSAWKRQPVRLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKVGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRI LQYALPRQESQRPFWWIFMDNLLLTEDDQETTTTFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEY	127

LTRP #	Components	Domains	AA sequence	SEQ ID NO
			LQAQVRSRSLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL	
	Linker (L1)		GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSE	123
	Linker (L3A) + buffer		GGSGGGSA	3271
	dCasX491		QEIKRINKIRRRRLVKDSNTKKAGKTGPMK TLLVRVMTPLDRERLENLRKKPENIPQPI SNTSRANLNKLLTDYTEMKKAILHVWEE FQKDPVGLMSRVAQPAASKKIDQNKLP DEKGNLTAGFACSQCQGPLFVYKLEQVS EKGKAYTNYFGRCNVAEHEKLILLAQLKP EKDSDEAVTYSLGKFGQRALDFYSIHVTK ESTHPVKPLAQIAGNRYASGPVGKALSDA CMGTIASFLSKYQDIIIEHQKVVKGNQKR LESLRELAKENLEYPVTLPPQPHTEG VDAYNEVIARVMWVNLNLWQKLKLSRDD AKPLLRKLGFPSPFLVERQANEVDWDMV CNVKKLINEKKEDGKVFQNLGAYKRQEA LRPYLSSEEDRKKGGKFARYQLGDLHL EKKHGEDWGKVYDEAWERIDKKVEGLSKH IKLEEEERSEDAQSKAALTDWLRKASFV IEGLKEADKDEFRCCELKQKWDLRGK PFAIEAENSILDISGFSKQYNCAFIWQKD GVKKLNLYLIINYFKGGKLRFKKIKPEAF EANRFYTVINKKSGEIVPMEVNFNDDPN LIILPLAFGKRQGREFIWNDLLSLETGSL KLANGRVIEKTLYNRRTRQDEPALFVALT FERREVLDSNPKPMNLIGVARGENIPAV IALTDPEGCPFSRFKDSLGNPTHILRIGE SYKEKQRTIQAKKEVEQRRAGGYSRKYAS KAKNLADDMVRNTARDLLYYAVTQDAMLI FANLSRGFRQGGKRTFMAERQYTRMEDWL TAKLAYEGLSKTYLSKTLAQYTSKTCSCNC GFTITSADYDRVLEKLKKTATGWMTTING KELKVEGQITYYNRYKRONVVKDLSELD RLSEESVNNDISSWTGRSGEALSLLKKR FSHRPVQEKFVCLNCGFETHAAEQALNI ARSWLFLRSQYKKYQTNKTTGNTDKRAF VETWQSFYRKKLKEVWKPAV	4
	Buffer + linker (L3B)		RSGSGGGGSTS	3272
	Repressor domain 1	Human ZIM3	MNNSQGRVTFEDVTVNFTQGEWQRLNPEQ RNLYRDVMLENYSNLVSVGQGETTKPDVI LRLEQGKEPWLEEEVLGSGRAEKNGDIG GQIWPKDVKESL	3240
		Human ZNF10	MDAKSLTAWSTLVTFKDVFDFTREEWK LLDTAQQIVYRNVMLENYKNLVSLGYQLT KPDVILRLEKGEEP	3239
		<i>Columba livia</i> repressor domain (DOMAIN 7694)	QDVVTFKDVAIYFSPEEWVRLSAGQRELY QEVMLDNYELVTSLDRESKLLYKMDPEEE SCEGVYPYSSADSGAPDSSTASAC	130
		<i>Rattus norvegicus</i> repressor domain (DOMAIN 10123)	ALVTFEDVAVRFTQEEWALLDPSQKILYR DVMRETYRNLTSGINWECWDLEACFRSL GRNLRVQVVRKCELTNSGPCAE	131

LTRP #	Components	Domains	AA sequence	SEQ ID NO
		<i>Cebus imitator</i> repressor domain (DOMAIN 15507)	SKAPITFGDLAIYFSQEWEWLSPIQKDL YEDVMLENYRNLVSLGLSFRRPNVITLLE KGKAPWMVEPARRRRGPDGSKV	132
		Chimpanzee repressor domain (DOMAIN 17905)	EMGLLTFRDIAIEFSLEEWQCLDCAQRNL YRDVMLENYRNLVSLGLAVSKPDILTCLLE QNKESQNIKRNMVAKHPVMHSH	133
		<i>Chlorocebus sabaesus</i> repressor domain (DOMAIN 20505)	SQESVAFEDVAVYFTTKWEAIMVPAERAL YRDVMLENYEAVAFVAVPPTSKPALVSHL EQGKESCFIRPPGVLSRSDWRAG	134
		<i>Ophiophagus hannah</i> repressor domain (DOMAIN 26749)	STPVTTFEDVVVYFTAAEWVHLTNWQRDFY QAVMMETYELVASVAGDGVPMADDEEGGV ERPVMQYIPRGKRRRKTPQPRAD	135
		<i>Ailuropoda melanoleuca</i> repressor domain (DOMAIN 27604)	LLTFEDVAVSFSEEEWELLDPPQKTLYND VMQENYETVISLGLKLKNDTGNDQPISSIS ALEMQASGSKVLRKARKMVAQKT	136
		<i>Peromyscus maniculatus bairdii</i> repressor domain (DOMAIN 29304)	VTYNDVHVDITQEEWALMDPSQRNLYKDV MVETYMNLTAIGYNWENLEVEEPCQNPLK HGRHERPHTGEKPYEYNQCGKAFAQP	137
		<i>Phyllostomus discolor</i> repressor domain (DOMAIN 30173)	VAFRDVIVDFTQEEWQQLKPAQKDLYRDV MMEIYWNLVSLDLETEGDMNEPDPEKDSW EDRTSIVVEGLMRNGAQGYACEKAGIQG CRV	138
	Buffer + NLS		TSPKKKRRKV	3273
LTRP4 with ADD domain	START codon + NLS + buffer sequence		See sequence above	3269
	Repressor domain 1	Human ZIM3	See sequence above	3240
		Human ZNF10	See sequence above	3239
		<i>Columba livia</i> repressor domain (DOMAIN 7694)	See sequence above	130
		<i>Rattus norvegicus</i> repressor domain (DOMAIN 10123)	See sequence above	131
		<i>Cebus imitator</i> repressor domain (DOMAIN 15507)	See sequence above	132
		Chimpanzee repressor domain (DOMAIN 17905)	See sequence above	133
		<i>Chlorocebus sabaesus</i> repressor domain (DOMAIN 20505)	See sequence above	134
		<i>Ophiophagus hannah</i> repressor domain (DOMAIN 26749)	See sequence above	135
		<i>Ailuropoda melanoleuca</i> repressor domain (DOMAIN 27604)	See sequence above	136
		<i>Peromyscus maniculatus bairdii</i> repressor domain (DOMAIN 29304)	See sequence above	137

LTRP #	Components	Domains	AA sequence	SEQ ID NO
		<i>Phyllostomus discolor</i> repressor domain (DOMAIN 30173)	See sequence above	138
	Linker (L3A) + buffer		See sequence above	3271
	START codon + DNMT3A ADD domain		See sequence above	3270
	DNMT3A catalytic domain		See sequence above	126
	Linker (L2)		See sequence above	122
	DNMT3L interaction domain		See sequence above	127
	Linker (L1)		See sequence above	123
	dCasX491		See sequence above	4
	Buffer + linker (L3B)		See sequence above	3272
	NLS		See sequence above	30
LTRP5 with ADD domain	START codon + NLS + buffer sequence		See sequence above	3269
	START codon + DNMT3A ADD domain		See sequence above	3270
	DNMT3A catalytic domain		See sequence above	126
	Linker (L2)		See sequence above	122
	DNMT3L interaction domain		See sequence above	127
	Linker (L3A)		See sequence above	124
	Repressor domain 1	Human ZIM3	See sequence above	3240
		Human ZNF10	See sequence above	3239
		<i>Columba livia</i> repressor domain (DOMAIN 7694)	See sequence above	130
		<i>Rattus norvegicus</i> repressor domain (DOMAIN 10123)	See sequence above	131
		<i>Cebus imitator</i> repressor domain (DOMAIN 15507)	See sequence above	132
		Chimpanzee repressor domain (DOMAIN 17905)	See sequence above	133
		<i>Chlorocebus sabaeus</i> repressor domain (DOMAIN 20505)	See sequence above	134

LTRP #	Components	Domains	AA sequence	SEQ ID NO
		<i>Ophiophagus hannah</i> repressor domain (DOMAIN 26749)	See sequence above	135
		<i>Ailuropoda melanoleuca</i> repressor domain (DOMAIN 27604)	See sequence above	136
		<i>Peromyscus maniculatus bairdii</i> repressor domain (DOMAIN 29304)	See sequence above	137
		<i>Phyllostomus discolor</i> repressor domain (DOMAIN 30173)	See sequence above	138
	Linker (L1)		See sequence above	123
	dCasX491		See sequence above	4
	Buffer + linker (L3B)		See sequence above	3272
	NLS		See sequence above	30

[0489] Table 41 provides exemplary amino acid sequences of components of LTRP constructs. In Table 41, the protein domains are shown without starting methionines.

Table 41: Exemplary protein sequences of components of LTRP constructs

Component	Protein sequence SEQ ID NO
DNMT3A catalytic domain (CD)	126
DNMT3L interaction domain	127
dCasX491	4
Linker 1 (L1)	123
Linker 2 (L2)	122
Linker 3A (L3A)	124
Linker 3B (L3B)	120
NLS	30
DNMT3A ADD domain	125

Example 11: Demonstration that inclusion of a second repressor domain enhances activity of LTRP-ADD molecules while maintaining their specificity

[0490] Experiments were performed to determine whether incorporation of a second repressor domain would enhance the activity and specificity of LTRP molecules containing the ADD domain. This second repressor domain was positioned on the C-terminus of the LTRP5-ADD molecule, described previously in Example 7, resulting in the generation of an LTRP molecule with a new configuration (hereafter described as an LTRP6 molecule). The schematic in FIG. 54 shows an LTRP6 molecule having either configuration 6a or 6b. A

molecule with configuration #6a utilizes the same repressor domain at both the N-terminus and C-terminus positions relative to dCasX, and a molecule with configuration 6b uses two different repressor domains at the two positions.

[0491] In this example, experiments were performed to assess whether incorporating a second repressor domain on the C-terminus of an LTRP5-ADD-ZIM3 molecule, which resulted in the generation of an LTRP6 molecule having a configuration 6b configuration, would induce long-term repression of the target locus.

Materials and Methods:

Generation of LTRP6 constructs and plasmid cloning:

[0492] Plasmid constructs coding for an LTRP6 molecule using the ZIM3-KRAB domain as the first repressor domain were built using standard molecular cloning techniques as described. Briefly, representative members of the top 95 repressor domains (Table 42) were cloned as the second repressor domain onto the C-terminus of an LTRP5-ADD-ZIM3 construct, which was generated as described in Example 7 (with DNA and protein sequences provided in Table 37). An LTRP5-ADD-ZIM3 molecule served as an experimental control. Plasmids also harbored constructs encoding for the gRNA scaffold variant 174 having either a spacer targeting the endogenous *B2M* locus (spacer 7.165; UCCCUAUGUCCUUGCUGUUU; SEQ ID NO: 3114) or a non-targeting control. These constructs were all cloned upstream of a P2A-puromycin element on the lentiviral plasmid.

Transfection of HEK293T cells:

[0493] Seeded HEK293T cells were transiently transfected with 100 ng of LTRP plasmids, each containing an LTRP:gRNA construct encoding for an LTRP6 variant containing a ZIM3-KRAB domain as the first repressor domain and an enhanced repressor domain as the second repressor domain, with the gRNA having either non-targeting spacer or a *B2M*-targeting spacer. 24 hours post-transfection, cells were selected with 1 µg/mL puromycin for three days. Cells were harvested for repression analysis at day 4, day 8, and day 13 post-transfection. Briefly, repression analysis was conducted by analyzing B2M protein expression via HLA immunostaining followed by flow cytometry, as described in Example 4.

Results:

[0494] The effect of incorporating a second repressor domain into the LTRP5-ADD-ZIM3 molecule, thereby generating an LTRP6 construct, on increasing long-term repression of the target *B2M* locus was assessed. The results from this time-course experiment are depicted in Tables 43-45, which shows the average percentage of cells characterized as HLA-negative

(indicative of B2M repression) for each condition at 4, 8, and 13 days post-transfection. SEQ ID NOs corresponding to the domain IDs in the tables below are provided in Table 42.

Table 42: List of top 9 and top 95 most effective repressor domains

Domain ID	Description	SEQ ID NO
<i>Top 9 repressor domains</i>		
DOMAIN_7694	Columba livia repressor domain	130
DOMAIN_10123	Rattus norvegicus repressor domain	131
DOMAIN_15507	Cebus imitator repressor domain	132
DOMAIN_17905	Chimp repressor domain	133
DOMAIN_20505	Chlorocebus sabaeus repressor domain	134
DOMAIN_26749	Ophiophagus hannah repressor domain	135
DOMAIN_27604	Ailuropoda melanoleuca repressor domain	136
DOMAIN_29304	Peromyscus maniculatus bairdii repressor domain	137
DOMAIN_30173	Phyllostomus discolor repressor domain	138
<i>Remaining repressor domains in the top 95 repressor domains</i>		
DOMAIN_737	Bonobo repressor domain	139
DOMAIN_10331	Colobus angolensis palliatus repressor domain	140
DOMAIN_10948	Colobus angolensis palliatus repressor domain	141
DOMAIN_11029	Mandrillus leucophaeus repressor domain	142
DOMAIN_17358	Bos indicus x Bos taurus repressor domain	143
DOMAIN_17759	Felis catus repressor domain	144
DOMAIN_18258	Physeter macrocephalus repressor domain	145
DOMAIN_19804	Callorhinus ursinus repressor domain	146
DOMAIN_221	Bonobo repressor domain	147
DOMAIN_881	Bonobo repressor domain	148
DOMAIN_2380	Orangutan repressor domain	149
DOMAIN_2942	Gibbon repressor domain	150
DOMAIN_4687	Marmoset repressor domain	151
DOMAIN_4806	Marmoset repressor domain	152
DOMAIN_4968	Marmoset repressor domain	153
DOMAIN_5066	Marmoset repressor domain	154
DOMAIN_5290	Owl Monkey repressor domain	155
DOMAIN_5463	Owl Monkey repressor domain	156
DOMAIN_6248	Saimiri boliviensis boliviensis repressor domain	157
DOMAIN_6445	Alligator sinensis repressor domain	158
DOMAIN_6802	Pantherophis guttatus repressor domain	159
DOMAIN_6807	Xenopus laevis repressor domain	160
DOMAIN_7255	Microcaecilia unicolor repressor domain	161
DOMAIN_8503	Mus caroli repressor domain	162
DOMAIN_8790	Marmota monax repressor domain	163
DOMAIN_8853	Mesocricetus auratus repressor domain	164
DOMAIN_9114	Peromyscus maniculatus bairdii repressor domain	165
DOMAIN_9331	Peromyscus maniculatus bairdii repressor domain	166
DOMAIN_9538	Mus musculus repressor domain	167
DOMAIN_9960	Octodon degus repressor domain	168

Domain ID	Description	SEQ ID NO
DOMAIN 10277	Dipodomys ordii repressor domain	169
DOMAIN 10577	Colobus angolensis palliatus repressor domain	170
DOMAIN 11348	Chlorocebus sabaues repressor domain	171
DOMAIN 11386	Capra hircus repressor domain	172
DOMAIN 11486	Bos mutus repressor domain	173
DOMAIN 11683	Nomascus leucogenys repressor domain	174
DOMAIN 12292	Sus scrofa repressor domain	175
DOMAIN 12452	Neophocaena asiaeorientalis asiaeorientalis repressor domain	176
DOMAIN 12631	Macaca fascicularis repressor domain	177
DOMAIN 13331	Macaca fascicularis repressor domain	178
DOMAIN 13468	Phascolarctos cinereus repressor domain	179
DOMAIN 13539	Gorilla repressor domain	180
DOMAIN 14659	Acinonyx jubatus repressor domain	181
DOMAIN 14755	Cebus imitator repressor domain	182
DOMAIN 15126	Callithrix jacchus repressor domain	183
DOMAIN 16444	Acinonyx jubatus repressor domain	184
DOMAIN 16688	Lipotes vexillifer repressor domain	185
DOMAIN 16806	Sapajus apella repressor domain	186
DOMAIN 17317	Otolemur garnettii repressor domain	187
DOMAIN 17432	Otolemur garnettii repressor domain	188
DOMAIN 18137	Monodelphis domestica repressor domain	189
DOMAIN 18216	Physeter macrocephalus repressor domain	190
DOMAIN 18563	OwlMonkey repressor domain	191
DOMAIN 19229	Enhydra lutris kenyoni repressor domain	192
DOMAIN 19460	Monodelphis domestica repressor domain	193
DOMAIN 19476	OwlMonkey repressor domain	194
DOMAIN 19821	Rhinopithecus roxellana repressor domain	195
DOMAIN 19892	Ursus maritimus repressor domain	196
DOMAIN 19896	Ovis aries repressor domain	197
DOMAIN 19949	Callorhinus ursinus repressor domain	198
DOMAIN 21247	Neovison vison repressor domain	199
DOMAIN 21317	Pteropus vampyrus repressor domain	200
DOMAIN 21336	Equus caballus repressor domain	201
DOMAIN 21603	Lipotes vexillifer repressor domain	202
DOMAIN 21755	Equus caballus repressor domain	203
DOMAIN 22153	Zalophus californianus repressor domain	204
DOMAIN 22270	Bonobo repressor domain	205
DOMAIN 23394	Vicugna pacos repressor domain	206
DOMAIN 23723	Carlito syrichta repressor domain	207
DOMAIN 24125	Saimiri boliviensis boliviensis repressor domain	208
DOMAIN 24458	Lynx pardinus repressor domain	209
DOMAIN 24663	Myotis brandtii repressor domain	210
DOMAIN 25289	Ursus maritimus repressor domain	211
DOMAIN 25379	Sapajus apella repressor domain	212
DOMAIN 25405	Desmodus rotundus repressor domain	213
DOMAIN 26070	Geotrypetes seraphini repressor domain	214
DOMAIN 26322	Geotrypetes seraphini repressor domain	215

Domain ID	Description	SEQ ID NO
DOMAIN_26732	Meleagris gallopavo repressor domain	216
DOMAIN_27060	Gopherus agassizii repressor domain	217
DOMAIN_27385	Octodon degus repressor domain	218
DOMAIN_27506	Bos mutus repressor domain	219
DOMAIN_27811	Callithrix jacchus repressor domain	220
DOMAIN_28640	Colinus virginianus repressor domain	221
DOMAIN_28803	Monodelphis domestica repressor domain	222
DOMAIN_30661	Physeter macrocephalus repressor domain	223
DOMAIN_31643	Micrurus lemniscatus lemniscatus repressor domain	224

Table 43: Level of B2M repression mediated by LTRP constructs with various repressor domains quantified at 4 days post-transfection. The domain ID shown for each LTRP6 construct designates the second repressor domain used and assessed

Repressor construct	Spacer	Mean % HLA-negative cells	Standard deviation	Sample Size
LTRP5-ADD-ZIM3	NT	6.827	0.359	3
LTRP5-ADD-ZIM3	7.165	53.833	1.159	3
LTRP6_DOMAIN_20505	7.165	69.333	0.493	3
LTRP6_DOMAIN_30173	7.165	70.267	1.115	3
LTRP6_DOMAIN_15507	7.165	71.800	0.436	3
LTRP6_DOMAIN_18137	7.165	71.900	0.600	3
LTRP6_DOMAIN_24458	7.165	73.333	0.961	3
LTRP6_DOMAIN_6445	7.165	73.567	0.666	3
LTRP6_DOMAIN_17432	7.165	74.667	0.603	3
LTRP6_DOMAIN_9960	7.165	76.000	0.100	3
LTRP6_DOMAIN_13468	7.165	76.167	0.493	3
LTRP6_DOMAIN_8853	7.165	77.400	0.173	3
LTRP6_DOMAIN_6802	7.165	77.700	0.361	3
LTRP6_DOMAIN_7694	7.165	78.300	0.361	3
LTRP6_DOMAIN_29304	7.165	78.933	0.814	3
LTRP6_DOMAIN_25379	7.165	80.033	0.569	3
LTRP6_DOMAIN_22153	7.165	80.200	1.493	3
LTRP6_DOMAIN_7255	7.165	80.567	0.551	3
LTRP6_DOMAIN_26749	7.165	81.467	0.503	3
LTRP6_DOMAIN_10123	7.165	81.833	0.289	3

Table 44: Level of B2M repression mediated by LTRP constructs with various repressor domains quantified at 8 days post-transfection. The domain ID shown for each LTRP6 construct designates the second repressor domain used and assessed

Repressor construct	Spacer	Mean % HLA-negative cells	Standard deviation	Sample Size
LTRP5-ADD-ZIM3	NT	2.100	0.311	2
LTRP5-ADD-ZIM3	7.165	55.000	0.300	3
LTRP6_DOMAIN_20505	7.165	72.367	1.686	3
LTRP6_DOMAIN_30173	7.165	72.733	1.069	3
LTRP6_DOMAIN_15507	7.165	74.033	0.351	3
LTRP6_DOMAIN_18137	7.165	75.100	1.253	3
LTRP6_DOMAIN_24458	7.165	76.167	0.737	3
LTRP6_DOMAIN_6445	7.165	77.600	1.249	3
LTRP6_DOMAIN_17432	7.165	78.233	0.153	3
LTRP6_DOMAIN_9960	7.165	77.600	0.346	3
LTRP6_DOMAIN_13468	7.165	79.067	0.651	3
LTRP6_DOMAIN_8853	7.165	79.333	1.150	3
LTRP6_DOMAIN_6802	7.165	78.400	0.265	3
LTRP6_DOMAIN_7694	7.165	81.433	0.416	3
LTRP6_DOMAIN_29304	7.165	80.433	0.569	3
LTRP6_DOMAIN_25379	7.165	82.167	0.351	3
LTRP6_DOMAIN_22153	7.165	81.333	1.210	3
LTRP6_DOMAIN_7255	7.165	80.933	0.404	3
LTRP6_DOMAIN_26749	7.165	85.767	0.777	3
LTRP6_DOMAIN_10123	7.165	83.000	0.625	3

Table 45: Level of B2M repression mediated by LTRP constructs with various repressor domains quantified at 13 days post-transfection. The domain ID shown for each LTRP6 construct designates the second repressor domain used and assessed

Repressor construct	Spacer	Mean % HLA-negative cells	Standard deviation	Sample Size
LTRP5-ADD-ZIM3	NT	1.587	0.670	3
LTRP5-ADD-ZIM3	7.165	50.267	0.987	3
LTRP6_DOMAIN_20505	7.165	70.133	0.945	3
LTRP6_DOMAIN_30173	7.165	70.233	0.666	3
LTRP6_DOMAIN_15507	7.165	72.067	0.252	3
LTRP6_DOMAIN_18137	7.165	72.767	1.514	3
LTRP6_DOMAIN_24458	7.165	73.967	0.404	3
LTRP6_DOMAIN_6445	7.165	75.967	0.723	3
LTRP6_DOMAIN_17432	7.165	75.667	0.351	3
LTRP6_DOMAIN_9960	7.165	74.233	0.814	3

Repressor construct	Spacer	Mean % HLA-negative cells	Standard deviation	Sample Size
LTRP6_DOMAIN_13468	7.165	76.767	0.764	3
LTRP6_DOMAIN_8853	7.165	75.133	1.457	3
LTRP6_DOMAIN_6802	7.165	74.700	1.136	3
LTRP6_DOMAIN_7694	7.165	78.967	0.321	3
LTRP6_DOMAIN_29304	7.165	78.233	1.332	3
LTRP6_DOMAIN_25379	7.165	78.600	0.985	3
LTRP6_DOMAIN_22153	7.165	77.533	0.839	3
LTRP6_DOMAIN_7255	7.165	77.700	0.265	3
LTRP6_DOMAIN_26749	7.165	83.800	1.127	3
LTRP6_DOMAIN_10123	7.165	79.367	0.702	3

[0495] The data in Tables 43-45 show that all LTRP6 constructs with a *B2M*-targeting gRNA were able to induce higher levels of B2M repression through at least 13 days post-transfection, with a slight variation in the potency of repression depending on the repressor domain utilized as a second repressor domain. Notably, use of DOMAIN_26749 (from the *Ophiophagus hannah*) as the second repressor domain resulted in ~83.8% of B2M repression, the highest level of repression amongst the other repressor domains assessed. In comparison, the LTRP5-ADD-ZIM3 construct only achieved a level of repression at ~50.3%, which is ~40% lower than that achieved by an LTRP6 construct with the ZIM3 and DOMAIN_26749 repressor domains (Tables 43-45). The observed improvement of repression from these experiments may be explained by the synergistic effects of using a more potent repressor domain, the presence of the ADD domain, and incorporation of the second repressor domain in the LTRP6 construct configuration. Example 12 below further investigates the effects of using two enhanced repressor domains on the activity and specificity of an LTRP6 construct.

[0496] The results from these experiments demonstrate that incorporation of a second repressor domain can enhance the activity of an LTRP molecule containing the ADD domain. The results described herein show that incorporating an enhanced repressor domain on the C-terminus of an LTRP5-ADD-ZIM3 molecule, resulting in the generation of an LTRP6 construct, improves repression activity at an endogenous locus in human cells.

Example 12: Assessing the effects of using two repressor domains and their relative positions within the LTRP6 molecule on activity and specificity

[0497] Experiments were performed to determine whether variations of the LTRP configuration 6 (LTRP6) molecule, described previously in Example 11, would result in improved long-term repression of the target locus and reduced off-target methylation. Here,

the effects of using an enhanced repressor domain as both a first repressor domain and a second repressor domain within an LTRP6 molecule, as well as the relative placements of these repressor domains, were evaluated. As shown in FIG. 61, two variations of the LTRP6 configuration were tested in this example. A molecule with configuration 6a utilizes the same repressor domain at both the N-terminus and C-terminus positions relative to dCasX, and a molecule with configuration 6b uses two different repressor domains at the two positions.

Materials and Methods:

Generation of LTRP6 constructs and plasmid cloning:

[0498] Plasmid constructs encoding for variants of an LTRP6 molecule, having either configuration 6a or 6b shown in the schematic in FIG. 61, were built using standard molecular cloning techniques. In this example, the following four repressor domains were used to generate variants of the LTRP6 molecule: DOMAIN_22153, DOMAIN_7255, DOMAIN_26749, and DOMAIN_10123. These four domains were selected because they induced the highest level of B2M repression at 4 days post-transfection as shown in Example 11, when incorporated as the second repressor domain in an LTRP6 configuration (Tables 43-45). Each of the four repressor domains were tested in combination at the N-terminus and/or C-terminus positions relative to dCasX, resulting in a total of 16 combinations assessed. The encoding sequences of the 16 LTRP6 variants are shown in Table 46, with the corresponding protein sequences shown in Table 47. Plasmids also harbored constructs encoding for the gRNA scaffold variant 174 having either a spacer targeting the endogenous *B2M* locus (spacer 7.165; UCCCUAUGUCCUUGCUGUUU; SEQ ID NO: 3114) or a non-targeting control. These constructs were all cloned upstream of a P2A-puromycin element on the lentiviral plasmid.

Table 46: Encoding sequences of the 16 LTRP6 variants assessed in this example*

Component	Domain Combination #	Domains	DNA sequence
START codon + NLS + buffer sequence			ATGGCCCCAAAGAAGAAGCGGAAGGTCTCTAGA
START codon + DNMT3A ADD domain			ATGGAGCGGCTGGTGTATGAGGTGCGCCAGAAGT GCAGAAACATCGAGGACATTTGTATCTCATGTGG GAGCCTCAATGTCACCCCTGGAGCACCCACTCTTC ATTGGAGGCATGTGCCAGAACTGTAAGAACTGCT TCTTGAGTGTGCTTACCAGTATGACGACGATGG GTACCAGTCCATTGCACCATCTGCTGTGGGGGG CGTGAAGTGCTCATGTGTGGGAACAACAAGTCT GCAGGTGCTTTTGTGTCGAGTGTGTGGATCTCTT GGTGGGGCCAGGAGCTGCTCAGGCAGCCATTAAG

Component	Domain Combination #	Domains	DNA sequence
			GAAGACCCCTGGAAGTGTACATGTGCGGGCATA AGGGCACCTATGGGCTGCTGCGAAGACGGGAAGA CTGGCCTTCTCGACTCCAGATGTTCTTTGCCAAT
DNMT3A catalytic domain			AACCATGACCAGGAATTTGACCCCCAAAGGTTT ACCCACCTGTGCCAGCTGAGAAGAGGAAGCCCAT CCGCGTGCTGTCTCTCTTTGATGGGATTGCTACA GGGCTCCTGGTGCTGAAGGACCTGGGCATCCAAG TGGACCGCTACATCGCCTCCGAGGTGTGTGAGGA CTCCATCACGGTGGGCATGGTGCGGCACCAAGGA AAGATCATGTACGTCGGGGACGTCCGCAGCGTCA CACAGAAGCATATCCAGGAGTGGGGCCCATTCGA CCTGGTGATTGGAGGCAGTCCCTGCAACGACCTC TCCATTGTCAACCTGCCCCGAAGGACTTTATG AGGGTACTGGCCGCCTCTTCTTTGAGTTCTACCG CCTCCTGCATGATGCGCGGCCCAAGGAGGGAGAT GATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGG TGGCCATGGGCGTTAGTGACAAGAGGGACATCTC GCGATTTCTTGAGTCTAACCCCGTGATGATTGAC GCCAAAGAAGTGTCTGCTGCACACAGGGCCCGTT ACTTCTGGGGTAACCTTCTTGGCATGAACAGGCC TTTGGCATCCACTGTGAATGATAAGCTGGAGCTG CAAGAGTGTCTGGAGCACGGCAGAATAGCCAAGT TCAGCAAAGTGAGGACCATTACCACCAGGTCAAA CTCTATAAAGCAGGGCAAAGACCAGCATTTCCCC GTCTTCATGAACGAGAAGGAGGACATCCTGTGGT GCACTGAAATGGAAAGGGTGTTTGGCTTCCCCGT CCACTACACAGACGTGTCCAACATGAGCCGCTTG GCGAGGCAGAGACTGCTGGGCCGGTCGTGGAGCG TGCCGGTCATCCGCCACCTCTTCGCTCCGCTGAA GGAATATTTTGCTTGTGTG
Linker 2 (L2)			TCTAGCGGCAATAGTAACGCTAACAGCCGCGGGC CGAGCTTCAGCAGCGGCCTGGTGCCGTTAAGCTT GCGCGGCAGCCAT
DNMT3L interaction domain			ATGGGCCCTATGGAGATATACAAGACAGTGTCTG CATGGAAGAGACAGCCAGTGCGGGTACTGAGCCT CTTCAGAAACATCGACAAGGTACTAAAGAGTTTG GGCTTCTTGGAAGCGGTTCTGGTTCTGGGGGAG GAACGCTGAAGTACGTGGAAGATGTACAAATGT CGTGAGGAGGGACGTGGAGAAATGGGGCCCTTT GACCTGGTGTACGGCTCGACGCAGCCCTAGGCA GCTCTTGTGATCGCTGTCCCGGCTGGTACATGTT CCAGTTCACCGGATTCTGCAGTATGCGCTGCCT CGCCAGGAGAGTCAGCGGCCCTTCTTCTGGATAT TCATGGACAATCTGCTGCTGACTGAGGATGACCA AGAGACAACCTACCCGCTTCCTTCAGACAGAGGCT GTGACCCTCCAGGATGTCCGTGGCAGAGACTACC AGAATGCTATGCGGGTGTGGAGCAACATTCCAGG GCTGAAGAGCAAGCATGCGCCCTGACCCCAAAG GAAGAAGAGTATCTGCAAGCCCAAGTCAGAAGCA GGAGCAAGCTGGACGCCCCGAAAGTTGACCTCCT GGTGAAGAAGTGCCTTCTCCCGCTGAGAGAGTAC TTCAAGTATTTTCTCAAACCTCACTTCCTCTT
Linker 3A (L3A)			GGCGGTTCCGGCGGAGGA
N-terminus repressor domain	Combinations 1, 5, 9, and 13	DOMAIN _22153	GTTCGTATGCCTGTTACTTTTGAAGATGTTGCTG TACTTTTACTCAAGAAGAATGGGGGCAATTAGA TCCTGCTCAACGTACTTTATATCAAGAAGTTATG

Component	Domain Combination #	Domains	DNA sequence
			TTAGAGACTTGTGGTTTATTAGTTTCTTTAGGTT GTCCTTTACCTAAACCTGAATTATTTTATCCTTT AGATCATTCTCCTGAATTACAACTTTAAAACGT GGTTTATCTCCTAATTCTTGTCTGGTGATTCTA CTAAA
	Combinations 2, 6, 10, and 14	DOMAIN _7255	ACTCAAGTTCCTGTTACCTTTGAAGATGTTGCTG TTTATTTTGTGAAGGTGAATGGGAACTTTAGC TGAATGGCAAAAAGAATTATATCGTGAACTATG AAAGAGAATTATGAACTTTAACTTCTTTAGGTT TTTCTTCTAAAAAACCTTCTTTAATTCTAAAAAT AGAACGTGAAGAAGATCCTTGTGTGCGTGATAAA CAAGATTCTCGTGATCGTCGTCGTTTACGTTCTT GTTGG
	Combinations, 3, 7, 11, and 15	DOMAIN _26749	TCTACTCCTGTTACTTTTGAAGATGTTGTTGTTT ATTTTACTGCTGCTGAATGGGTTTCATTTAACTAA TTGGCAACGTGATTTTTATCAAGCTGTTATGATG GAACTTATGAATTAGTTGCTTCTGTGCTGGTG ATGGTGTTCTATGGCTGAAGATGAAGAAGGTGG TGTTGAACGTCTGTTTGGCAATATATTCCTCGT GGTAAACGTGTCGTAACCTCCTCAACCTCGTG CTGAT
	Combinations 4, 8, 12, and 16	DOMAIN _10123	GCTTTAGTTACTTTTGAAGATGTTGCTGTTCTGTT TACTCAAGAAGAATGGGCTTTATTAGATCCTTC TCAAAAAATTTTATATCGTGATGTTATGCGTGAG ACTTATCGTAATTTAACTTCTGTTGGTATTAATT GGGAATGTTGGGATTTAGAAGCTTGTTCGTTTC TTTAGGTCGTAATTTACGTGTTCAAGTCGTTAAA CGTAAATGTGAATTAATAATTCTGGTCCTTGTG CTGAA
Linker 1 (L1)			GGAGGGCCGAGCTCTGGCGCACCCCCACCAAGTG GAGGGTCTCCTGCCGGGTCCCCAACATCTACTGA AGAAGGCACCAGCGAATCCGCAACGCCCGAGTCA GGCCCTGGTACCTCCACAGAACCATCTGAAGGTA GTGCGCCTGGTTCCTCCAGCTGGAAGCCCTACTTC CACCGAAGAAGGCACGTCAACCGAACCAAGTGAA GGATCTGCCCCTGGGACCAGCACTGAACCATCTG AG
dCasX		dCasX491	CAAGAGATCAAGAGAATCAACAAGATCAGAAGGA GACTGGTCAAGGACAGCAACACAAAGAGGCCGG CAAGACAGGCCCATGAAAACCTGCTCGTCAGA GTGATGACCCCTGACCTGAGAGAGCGGCTGGAAA ACCTGAGAAAGAAGCCCGAGAACATCCCTCAGCC TATCAGCAACACCAGCAGGGCCAACCTGAACAAG CTGCTGACCGACTACACCGAGATGAAGAAAGCCA TCCTGCACGTGTACTGGGAAGAGTTCCAGAAAGA CCCCGTGGGCCTGATGAGCAGAGTTGCTCAGCCT GCCAGCAAGAAGATCGACCAGAACAGCTGAAGC CCGAGATGGACGAGAAGGGCAATCTGACCACAGC CGGCTTTGCCTGCTCTCAGTGTGGCCAGCCTCTG TTCGTGTACAAGCTGGAACAGGTGTCCGAGAAAG GCAAGGCTACACCAACTACTTCGGCAGATGTAA CGTGGCCGAGCACGAGAAGCTGATTCTGCTGGCC CAGCTGAAACCTGAGAAGGACTCTGATGAGGCCG TGACCTACAGCCTGGGCAAGTTTGGACAGAGAGC CCTGGACTTCTACAGCATCCACGTGACCAAAGAA AGCACACACCCCGTGAAGCCCTGGCTCAGATCG CCGGCAATAGATACGCCCTCTGGACCTGTGGGCAA

Component	Domain Combination #	Domains	DNA sequence
			AGCCCTGTCCGATGCCTGCATGGGAACAATCGCC AGCTTCCTGAGCAAGTACCAGGACATCATCATCG AGCACCAGAAGGTGGTCAAGGGCAACCAGAAGAG ACTGAAAGCCTGAGGGAGCTGGCCGGCAAAGAG AACCTGGAATACCCAGCGTGACCTGCCTCCTC AGCCTCACACAAAAGAAGGCGTGGACGCCTACAA CGAAGTGATCGCCAGAGTGAGAATGTGGGTCAAC CTGAACCTGTGGCAGAAGCTGAAACTGTCCAGGG ACGACGCCAAGCCTCTGCTGAGACTGAAGGGCTT CCCTAGCTTCCCTCTGGTGGAAAGACAGGCCAAT GAAGTGGATTGGTGGGACATGGTCTGCAACGTGA AGAAGCTGATCAACGAGAAGAAAGAGGATGGCAA GGTTTTCTGGCAGAACCTGGCCGGCTACAAGAGA CAAGAAGCCCTGAGGCCTTACCTGAGCAGCGAAG AGGACCGGAAGAAGGGCAAGAAGTTCCGCAGATA CCAGCTGGGCGACCTGCTGCTGCACCTGGAAAAG AAGCACGGCGAGGACTGGGGCAAAGTGTACGATG AGGCCTGGGAGAGAATCGACAAGAAGGTGGAAGG CCTGAGCAAGCACATTAAGCTGGAAGAGGAAAGA AGGAGCGAGGACGCCCAATCTAAAGCCGCTCTGA CCGATTGGCTGAGAGCCAAGGCCAGCTTTGTGAT CGAGGGCCTGAAAGAGGCCGACAAGGACGAGTTC TGCAGATGCGAGCTGAAGCTGCAGAAGTGGTACG GCGATCTGAGAGGCAAGCCCTTCGCCATTGAGGC CGAGAACAGCATCCTGGACATCAGCGGCTTCAGC AAGCAGTACAACCTGCGCCTTCATTTGGCAGAAAG ACGGCGTCAAGAACTGAACCTGTACCTGATCAT CAATTACTTCAAAGGCGGCAAGCTGCGGTTCAAG AAGATCAAACCCGAGGCCTTCGAGGCTAACAGAT TCTACACCGTGATCAACAAAAAGTCCGGCGAGAT CGTGCCCATGGAAGTGAACCTCAACTTCGACGAC CCCAACCTGATTATCCTGCCTCTGGCCTTCGGCA AGAGACAGGGCAGAGAGTTCATCTGGAACGATCT GCTGAGCCTGGAAACCGGCTCTCTGAAGCTGGCC AATGGCAGAGTGATCGAGAAAACCTGTACAACA GGAGAACCAGACAGGACGAGCCTGCTCTGTTTGT GGCCCTGACCTTCGAGAGAAGAGAGGTGCTGGAC AGCAGCAACATCAAGCCCATGAACCTGATCGGCG TGGCCCGGGGCGAGAATATCCCTGCTGTGATCGC CCTGACAGACCCTGAAGGATGCCCACTGAGCAGA TTCAAGGACTCCCTGGGCAACCCTACACACATCC TGAGAATCGGCGAGAGCTACAAAGAGAAGCAGAG GACAATCCAGGCCAAGAAAGAGGTGGAACAGAGA AGAGCCGGCGGATACTCTAGGAAGTACGCCAGCA AGGCCAAGAATCTGGCCGACGACATGGTCCGAAA CACCGCCAGAGATCTGCTGTACTACGCCGTGACA CAGGACGCCATGCTGATCTTCGCGAATCTGAGCA GAGGCTTCGGCCGGCAGGGCAAGAGAACCTTTAT GGCCGAGAGGCAGTACACCAGAATGGAAGATTGG CTCACAGCTAACTGGCCTACGAGGGACTGAGCA AGACCTACCTGTCCAAAACACTGGCCCAAGTATAC CTCCAAGACCTGCAGCAATTGCGGCTTCACCATC ACCAGCGCCGACTACGACAGAGTGCTGGAAAAGC TCAAGAAAACCGCCACCGGCTGGATGACCACCAT CAACGGCAAAGAGCTGAAGGTTGAGGGCCAGATC ACCTACTACAACAGGTACAAGAGGCAGAACGTCTG TGAAGGATCTGAGCGTGGAACTGGACAGACTGAG CGAAGAGAGCGTGAACAACGACATCAGCAGCTGG

Component	Domain Combination #	Domains	DNA sequence
			ACAAAGGGCAGATCAGGCGAGGCTCTGAGCCTGC TGAAGAAGAGGTTTAGCCACAGACCTGTGCAAGA GAAGTTCGTGTGCCTGAACTGCGGCTTCGAGACA CACGCCGCTGAACAGGCTGCCCTGAACATTGCCA GAAGCTGGCTGTTCCCTGAGAAGCCAAGAGTACAA GAAGTACCAGACCAACAAGACCACCGGCAACACC GACAAGAGGGCCTTTGTGGAAACCTGGCAGAGCT TCTACAGAAAAAAGCTGAAAGAAGTCTGGAAGCC CGCCGTG
Buffer sequence + Linker 3B (L3B) + buffer sequence			CGATCGGGCGGTTCCGGCGGAGGTTCCACTAGT
C-terminus repressor domain	Combinations 1-4	DOMAIN _7255	ACTCAAGTTCCTGTTACCTTTGAAGATGTTGCTG TTTATTTTTGTGAAGGTGAATGGGAACTTTAGC TGAATGGCAAAAAGAATTATATCGTGAAACTATG AAAGAGAATTATGAACTTTAACTTCTTTAGGTT TTTCTTCTAAAAAACCTTCTTTAATTCTAAAAAT AGAACGTGAAGAAGATCCTTGTGTGCGTGATAAA CAAGATTCTCGTGATCGTCGTCGTTACGTTCTT GTTGG
	Combinations 5-8	DOMAIN _22153	GTTCTGATGCCTGTTACTTTTGAAGATGTTGCTG TTACTTTTACTCAAGAAGAATGGGGCAATTAGA TCCTGCTCAACGTACTTTATATCAAGAAGTTATG TTAGAGACTTGTGGTTTATTAGTTTCTTTAGGTT GTCCTTTACCTAAACCTGAATTATTTATCCTTT AGATCATTCTCCTGAATTACAACTTTAAAACGT GGTTTATCTCCTAATTCTTGTCTGCTGATTCTA CTAAA
	Combinations 9-12	DOMAIN _10123	GCTTTAGTTACTTTTGAAGATGTTGCTGTTCTGTT TTACTCAAGAAGAATGGGCTTTATTAGATCCTTC TCAAAAATTTTATATCGTGATGTTATGCGTGAG ACTTATCGTAATTTAACTTCTGTTGGTATTAATT GGGAATGTTGGGATTTAGAAGCTTGTTCGTTTC TTTAGGTCGTAATTTACGTGTTCAAGTCGTTAAA CGTAAATGTGAATTAATAATTCTGGTCCTTGTG CTGAA
	Combinations 13-16	DOMAIN _26749	TCTACTCCTGTACTTTTGAAGATGTTGTTGTTT ATTTTACTGCTGCTGAATGGGTTTCAATTAATAA TTGGCAACGTGATTTTATCAAGCTGTTATGATG GAACTTATGAATTAGTTGCTTCTGTTGCTGGTG ATGGTGTTCTATGGCTGAAGATGAAGAAGGTGG TGTTGAACGTCCTGTTTGGCAATATATCCTCGT GGTAAACGTCGTCGTAAACTCCTCAACCTCGTG CTGAT
Linker 4v1 (L4v1)			GGAAGTGGGAGC
NLS			CCCAAGAAAAAGAGAAAAGTG

*Components are listed in a 5' to 3' order within the constructs

Table 47: Full-length protein sequences of the 16 LTRP6 variants assessed in this example

Component	Domain Combination #	Domains	AA sequence
START codon + NLS + buffer sequence			MAPKKKKRKVSR
START codon + DNMT3A ADD domain			MERLVYEVQRQKCRNIEDICISCGSLNVTLEHPLF IGGMCQCNCKNCFLECAQQYDDDGYSYCTICCGG REVLNMCNNCCRCFCVECDLLVGPAAQAIAIK EDPWNCYMCCHKGTGGLRRREDWPSRLQMFFAN
DNMT3A catalytic domain			NHDQEFDPKVPVPAEKRPPIRVLSLFDGIAT GLLVKLDLGIQVDRIASEVCEDSITVGMVRHQG KIMYVGDVRSVTQKHQEWGPFDLVIGGSPCNDL STVNPARKGLYEGTGRLLFFEFYRLLDARPEGD DRPFFWLFENVVAMGVSDKRDISRFLSNPVMID AKEVSAHRARYFWGNLPGMNRPLASTVNDKLEL QECLEHGRIAKFSKVRTITTRSNSIKQKGDQHF VFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRL ARQRLGRSWSVPVIRHLFAPLKEYFACV
Linker 2 (L2)			SSGNSNANSRGPSPSSGLVPLSLRGSH
DNMT3L interaction domain			MGPMEIYKTVSAWKRPVRVLSLFRNIDKVLKSL GFLESGSGSGGGLTKYVEDVTNVVRRDVEKWGPF DLVYGSTQPLGSSCDRCPGWYMFQFHRIQYALP RQESQRPFFWIFMDNLLLTEDDQETTTFLQTEA VTLDVGRGRDYQNAMRVWSNIPGLKSKHAPLTPK EEEYLQAQVRSRSLDAPKVDLLVKNCLPLREY FKYFSQNSLPL
Linker 3A (L3A)			GGSGGG
N-terminus repressor domain	Combinations 1, 5, 9, and 13	DOMAI N_22153	VRMPVTFEDVAVTFTQEEWGQLDPAQRTLYQEV LETGGLLVSLGCPKPELFYPLDHSPELQTLKR GLSPNSCPGDSTK
	Combinations 2, 6, 10, and 14	DOMAI N_7255	TQVPVTFEDVAVYFCEGEWETLAEWQKELYRET KENYETLTSLGFSSKKPSLISKIEREEDPCVRDK QDSRDRRLRSCW
	Combinations 3, 7, 11, and 15	DOMAI N_26749	STPVTTFEDVVVYFTAAEWVHLTNWQDFYQAVMM ETYELVASVAGDGVPMADDEEGGVERPVWQYIPR GKRRRKTPQPRAD
	Combinations 4, 8, 12, and 16	DOMAI N_10123	ALVTFEDVAVRFTQEEWALLDPSQKILYRDVMRE TYRNLTSGINWECWDLEACFRSLGRNLRVQVVK RKCELTNSGPCE
Linker 1 (L1)			GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPES GPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSE GSAPGTSTEPSE
dCasX		dCasX491	QEIKRINKIRRRRLVKDSNTKKAGKTGPMKTLVR VMTFPLRERLENLRKKPENIPQFISNTSRANLNK LLTDYTEMKKAILHVYWEFFQKDPVGLMSRVAQP ASKKIDQNKLPKPEMDEKGNLTAGFACSQCGQPL FVYKLEQVSEKGAITYNYFGRCNVAEHEKLILLA QLKPEKDSDEAVTYSLGKFGQRALDFYSIHVTK STHPVKPLAQIAGNRYASGPVGKALSDACMGTTIA SFLSKYQDIIIEHQVVKGNQKRLESRELAGEKE NLEYPSVTLPQPHTKEGVDAVNEVIARVRMWN

Component	Domain Combination #	Domains	AA sequence
			LNLWQKLKLSRDDAKPLLRLKGFPSFPLVERQAN EVDWDMVCNVKKLINEKKEDGKVFWQNLAGYKR QEALRPYLSSEEDRKKGKKFARYQLGDLLHLEK KHGEDWGKVYDEAWERIDKKVEGLSKHIKLEER RSEDAQSKAALTDWLRKASFVIEGLKEADKDEF CRCELKLQKWYGDLRGKPPFAIEAENSILDISGFS KQYNCAFIWQKDGVKKLNLYLIINYFKGGKLRFK KIKPEAFEANRFYTVINKKSGEIVPMEVNFNFDD PNLIILPLAFGKRQGREFIWNDLLSLETGSLKLA NGRVIEKTLYNRRTRQDEPALFVALTFERREVL SSNIKPMNLIGVARGENIPAVIALTDPEGCPISR FKDSLGNPTHILRIGESYKEKQRTIQAKKEVEQR RAGGYSRKYASKAKNLADDMVRNTARDLLYYAVT QDAMLIIFANLSRGFGKRGKRTFMAERQYTRMEDW LTAKLAYEGLSKTYLSKTLAQYTSKTCNCGFTI TSADYDRVLEKLKKTATGWMTINGKELKVEGQI TYNRYKRQNVVKDLSELDRLSEESVNNDISSW TKGRSGEALSLKKRFSHRPVQEKFVCLNCGFET HAAEQAALNIARSWLFLRSQYKQYQTNKTTGNT DKRAFVETWQSFYRKKLKEVWKPAV
Buffer sequence + Linker 3B (L3B) + buffer sequence			RSGGSGGGSTS
C-terminus repressor domain	Combinations 1-4	DOMAI N_7255	TQVPVTFEDVAVYFCEGEWETLAEWQKELYRETM KENYETLTSLGFSKKPSLISKIEREEDPCVRDK QDSRDRRLRSCW
	Combinations 5-8	DOMAI N_22153	VRMPVTFEDVAVTFTQEEWGQLDPAQRTLYQEV LETGGLLVSLGCPKPELFYPLDHSPELQTLKR GLSPNSCPGDSK
	Combinations 9-12	DOMAI N_10123	ALVTFEDVAVRFTQEEWALLDPSQKILYRDV TYRNLTSGINWECWDLEACFRSLGRNLRVQVVK RKCELTN SGPCAE
	Combinations 13-16	DOMAI N_26749	STPVTTFEDVVVYFTAAEWVHLTNWQRDFYQAVMM ETYELVASVAGDGVMAEDEEGGVERPVWQYIPR GKRRRKTPQPRAD
Linker 4v1 (L4v1)			GSGS
NLS			PKKKRKV

[0499] Transfection of HEK293T cells was performed following similar methods as described in Example 11. Briefly, HEK293T cells were transiently transfected with a plasmid encoding an LTRP6 variant (listed in Table 46) with a gRNA having a *B2M*-targeting spacer. 24 hours post-transfection, cells were selected with puromycin for five days. Cells were harvested for determining the level of B2M repression at 8, 12, 20, 27, 44, 59, 74, and 104 days post-transfection. An LTRP5-ZIM3 molecule and an LTRP5-ADD-ZIM3 molecule served as experimental controls. Furthermore, an LTRP6 variant using a ZIM3 domain as a

first and second repressor domain, also known as an LTRP6 with dual ZIM3, was also included in this experiment for comparison.

[0500] Bisulfite sequencing was also performed to assess off-target methylation at the *VEGFA* locus, which was performed as described in Example 5. Briefly, HEK293T cells transiently transfected with LTRP6 variant plasmids, with a *B2M*-targeting gRNA or non-targeting gRNA, were harvested at five days post-transfection for gDNA extraction for bisulfite sequencing. The following controls were also included: 1) CasX 491 paired with *B2M* spacer 7.37 (GGCCGAGAUGUCUCGCUCG; SEQ ID NO: 3137); 2) a dCas9-ZNF10-DNMT3A/3L with its corresponding *B2M* spacer; 3) LTRP1 (described in FIG. 28) with either a non-targeting or *B2M* spacer 7.165; 4) LTRP5-ZIM5 with *B2M* spacer 7.165; and 5) LTRP5-ADD-ZIM3 with either a non-targeting or *B2M* spacer 7.165.

Results:

[0501] The effects of using an enhanced repressor domain as both a first repressor domain and a second repressor domain, as well as their relative placements within an LTRP6 molecule, were assessed. The results from a time-course experiment evaluating the effects on long-term repression of the target *B2M* locus are depicted in Tables 48-55. Overall, the data demonstrate that the incorporation of a second repressor domain to the C-terminus of an LTRP5-ADD molecule (resulting in the LTRP6 configuration) substantially improved durable B2M repression, recapitulating the findings observed in Example 11. The addition of a ZIM3 domain to the C-terminus of an LTRP5-ADD-ZIM3, resulting in the generation of an LTRP6-dual-ZIM3, enhanced long-term repression when compared to the level of B2M repression achieved by the LTRP5-ADD-ZIM3 construct, at least through 59 days post-transfection, although this improvement in long-term repression was not sustained by day 104 (Tables 48-55). The enhanced repression observed with the use of a second ZIM3 domain further supports that the observed improvement in activity in the LTRP6 configuration is due to the presence of a second repressor domain linked C-terminal to the dCasX. In comparison to LTRP6-dual-ZIM3, most of the other LTRP6 constructs harboring the two enhanced repressor domains further improved repressive activity even through 104 days post-transfection. Notably, LTRP6 molecules containing DOMAIN_26749 (either at the N-terminus and/or C-terminus) were the best performing molecules in inducing durable B2M repression (>90%) through day 104, with LTRP6-dual-26749 (configuration 6a) demonstrating the highest level of repression and LTRP6-10123-26749 (configuration 6b) exhibiting the second highest repressive activity (Tables 48-55). Interestingly, in some instances, the placement of the repressor domains on the N-terminus of the dCasX compared

to the C-terminus also affected activity of the molecule. For instance, placing DOMAIN_7255 on the N-terminus of the dCasX in combination with DOMAIN_22153 on the C-terminus resulted in ~58% repression; however, the level of repression increased to 91% when DOMAIN_22153 was on the N-terminus and DOMAIN_7255 was on the C-terminus (Tables 48-55).

Table 48: Level of B2M repression mediated by LTRP constructs with various repressor domains quantified at 8 days post-transfection. The order of the domain IDs shown for each LTRP6 construct designates the first repressor domain and second repressor domain at the N-terminus and C-terminus positions respectively

Repressor construct	Spacer	Mean % HLA-negative cells	Standard deviation	Sample Size
LTRP5-ADD-ZIM3	NT	8.437	2.483	3
LTRP5-ZIM3	7.165	61.133	2.003	3
LTRP5-ADD-ZIM3	7.165	74.967	1.443	3
LTRP6-dual-ZIM3	7.165	89.833	0.416	3
LTRP6-22153-7255	7.165	92.200	0.361	3
LTRP6-dual-7255	7.165	91.267	0.321	3
LTRP6-26749-7255	7.165	92.567	0.208	3
LTRP6-10123-7255	7.165	92.100	0.529	3
LTRP6-dual-22153	7.165	88.700	0.600	3
LTRP6-7255-22153	7.165	91.500	0.100	3
LTRP6-26749-22153	7.165	93.033	0.737	3
LTRP6-10123-22153	7.165	91.967	0.702	3
LTRP6-22153-10123	7.165	91.767	1.405	3
LTRP6-7255-10123	7.165	90.800	2.128	3
LTRP6-26749-10123	7.165	94.000	1.082	3
LTRP6-dual-10123	7.165	93.100	1.136	3
LTRP6-22153-26749	7.165	92.167	1.474	3
LTRP6-7255-26749	7.165	91.567	1.365	3
LTRP6-dual-26749	7.165	92.100	1.044	3
LTRP6-10123-26749	7.165	92.367	0.802	3

Table 49: Level of B2M repression mediated by LTRP constructs with various repressor domains quantified at 12 days post-transfection. The order of the domain IDs shown for each LTRP6 construct designates the first repressor domain and second repressor domain at the N-terminus and C-terminus positions respectively

Repressor construct	Spacer	Mean % HLA-negative cells	Standard deviation	Sample Size
LTRP5-ADD-ZIM3	NT	1.970	0.442	3
LTRP5-ZIM3	7.165	50.433	1.250	3

Repressor construct	Spacer	Mean % HLA-negative cells	Standard deviation	Sample Size
LTRP5-ADD-ZIM3	7.165	75.300	1.997	3
LTRP6-dual-ZIM3	7.165	90.067	0.404	3
LTRP6-22153-7255	7.165	93.567	0.493	3
LTRP6-dual-7255	7.165	90.400	0.173	3
LTRP6-26749-7255	7.165	94.367	0.306	3
LTRP6-10123-7255	7.165	93.300	0.608	3
LTRP6-dual-22153	7.165	89.067	0.924	3
LTRP6-7255-22153	7.165	90.733	0.058	3
LTRP6-26749-22153	7.165	95.267	0.709	3
LTRP6-10123-22153	7.165	94.500	0.608	3
LTRP6-22153-10123	7.165	93.400	0.100	3
LTRP6-7255-10123	7.165	89.667	1.210	3
LTRP6-26749-10123	7.165	95.700	0.200	3
LTRP6-dual-10123	7.165	94.033	0.289	3
LTRP6-22153-26749	7.165	95.067	0.651	3
LTRP6-7255-26749	7.165	95.300	0.866	3
LTRP6-dual-26749	7.165	95.567	0.404	3
LTRP6-10123-26749	7.165	95.400	0.608	3

Table 50: Level of B2M repression mediated by LTRP constructs with various repressor domains quantified at 20 days post-transfection. The order of the domain IDs shown for each LTRP6 construct designates the first repressor domain and second repressor domain at the N-terminus and C-terminus positions respectively

Repressor construct	Spacer	Mean % HLA-negative cells	Standard deviation	Sample Size
LTRP5-ADD-ZIM3	NT	6.060	3.841	3
LTRP5-ZIM3	7.165	45.233	3.958	3
LTRP5-ADD-ZIM3	7.165	71.700	2.107	3
LTRP6-dual-ZIM3	7.165	89.767	1.124	3
LTRP6-22153-7255	7.165	92.433	0.987	3
LTRP6-dual-7255	7.165	89.233	1.528	3
LTRP6-26749-7255	7.165	93.200	0.400	3
LTRP6-10123-7255	7.165	92.133	0.929	3
LTRP6-dual-22153	7.165	87.700	0.872	3
LTRP6-7255-22153	7.165	88.633	0.814	3
LTRP6-26749-22153	7.165	94.767	0.850	3
LTRP6-10123-22153	7.165	93.400	0.794	3
LTRP6-22153-10123	7.165	91.967	1.274	3
LTRP6-7255-10123	7.165	87.233	1.960	3
LTRP6-26749-10123	7.165	94.767	0.208	3
LTRP6-dual-10123	7.165	92.367	0.404	3
LTRP6-22153-26749	7.165	94.400	1.015	3

Repressor construct	Spacer	Mean % HLA-negative cells	Standard deviation	Sample Size
LTRP6-7255-26749	7.165	93.667	1.069	3
LTRP6-dual-26749	7.165	95.367	0.681	3
LTRP6-10123-26749	7.165	94.867	0.907	3

Table 51: Level of B2M repression mediated by LTRP constructs with various repressor domains quantified at 27 days post-transfection. The order of the domain IDs shown for each LTRP6 construct designates the first repressor domain and second repressor domain at the N-terminus and C-terminus positions respectively

Repressor construct	Spacer	Mean % HLA-negative cells	Standard deviation	Sample Size
LTRP5-ADD-ZIM3	NT	2.600	1.006	3
LTRP5-ZIM3	7.165	36.733	2.608	3
LTRP5-ADD-ZIM3	7.165	69.100	1.900	3
LTRP6-dual-ZIM3	7.165	87.200	1.300	3
LTRP6-22153-7255	7.165	91.300	1.179	3
LTRP6-dual-7255	7.165	87.233	1.250	3
LTRP6-26749-7255	7.165	92.367	0.551	3
LTRP6-10123-7255	7.165	90.733	1.518	3
LTRP6-dual-22153	7.165	85.567	1.007	3
LTRP6-7255-22153	7.165	86.433	1.290	3
LTRP6-26749-22153	7.165	94.067	0.681	3
LTRP6-10123-22153	7.165	92.467	1.002	3
LTRP6-22153-10123	7.165	90.900	1.646	3
LTRP6-7255-10123	7.165	85.200	1.411	3
LTRP6-26749-10123	7.165	93.567	0.643	3
LTRP6-dual-10123	7.165	91.033	0.058	3
LTRP6-22153-26749	7.165	93.900	0.693	3
LTRP6-7255-26749	7.165	92.933	1.474	3
LTRP6-dual-26749	7.165	95.300	0.557	3
LTRP6-10123-26749	7.165	94.133	1.528	3

Table 52: Level of B2M repression mediated by LTRP constructs with various repressor domains quantified at 44 days post-transfection. The order of the domain IDs shown for each LTRP6 construct designates the first repressor domain and second repressor domain at the N-terminus and C-terminus positions respectively

Repressor construct	Spacer	Mean % HLA-negative cells	Standard deviation	Sample Size
LTRP5-ADD-ZIM3	NT	1.307	0.070	3
LTRP5-ZIM3	7.165	26.533	3.232	3
LTRP5-ADD-ZIM3	7.165	63.700	1.345	3

Repressor construct	Spacer	Mean % HLA-negative cells	Standard deviation	Sample Size
LTRP6-dual-ZIM3	7.165	82.600	1.929	3
LTRP6-22153-7255	7.165	88.933	4.110	3
LTRP6-dual-7255	7.165	83.233	2.610	3
LTRP6-26749-7255	7.165	88.833	1.604	3
LTRP6-10123-7255	7.165	86.567	1.909	3
LTRP6-dual-22153	7.165	79.867	0.153	3
LTRP6-7255-22153	7.165	80.033	3.807	3
LTRP6-26749-22153	7.165	91.933	2.159	3
LTRP6-10123-22153	7.165	89.567	0.929	3
LTRP6-22153-10123	7.165	86.900	3.119	3
LTRP6-7255-10123	7.165	77.433	2.031	3
LTRP6-26749-10123	7.165	91.133	0.611	3
LTRP6-dual-10123	7.165	87.033	0.231	3
LTRP6-22153-26749	7.165	93.000	2.100	3
LTRP6-7255-26749	7.165	89.367	3.758	3
LTRP6-dual-26749	7.165	96.167	0.737	3
LTRP6-10123-26749	7.165	94.050	0.919	2

Table 53: Level of B2M repression mediated by LTRP constructs with various repressor domains quantified at 59 days post-transfection. The order of the domain IDs shown for each LTRP6 construct designates the first repressor domain and second repressor domain at the N-terminus and C-terminus positions respectively

Repressor construct	Spacer	Mean % HLA-negative cells	Standard deviation	Sample Size
LTRP5-ADD-ZIM3	NT	0.947	0.012	3
LTRP5-ZIM3	7.165	19.667	2.627	3
LTRP5-ADD-ZIM3	7.165	60.533	1.801	3
LTRP6-dual-ZIM3	7.165	76.567	4.274	3
LTRP6-22153-7255	7.165	87.933	6.385	3
LTRP6-dual-7255	7.165	84.100	2.211	3
LTRP6-26749-7255	7.165	85.733	1.704	3
LTRP6-10123-7255	7.165	83.267	3.868	3
LTRP6-dual-22153	7.165	74.500	0.608	3
LTRP6-7255-22153	7.165	74.133	5.463	3
LTRP6-26749-22153	7.165	89.833	3.691	3
LTRP6-10123-22153	7.165	89.200	1.825	3
LTRP6-22153-10123	7.165	84.733	3.581	3
LTRP6-7255-10123	7.165	71.667	5.980	3
LTRP6-26749-10123	7.165	90.567	0.808	3
LTRP6-dual-10123	7.165	83.167	1.026	3
LTRP6-22153-26749	7.165	93.467	2.301	3
LTRP6-7255-26749	7.165	86.900	6.077	3

Repressor construct	Spacer	Mean % HLA-negative cells	Standard deviation	Sample Size
LTRP6-dual-26749	7.165	96.833	0.874	3
LTRP6-10123-26749	7.165	93.450	1.202	2

Table 54: Level of B2M repression mediated by LTRP constructs with various repressor domains quantified at 74 days post-transfection. The order of the domain IDs shown for each LTRP6 construct designates the first repressor domain and second repressor domain at the N-terminus and C-terminus positions respectively

Repressor construct	Spacer	Mean % HLA-negative cells	Standard deviation	Sample Size
LTRP5-ADD-ZIM3	NT	2.407	1.086	3
LTRP5-ZIM3	7.165	21.033	9.890	3
LTRP5-ADD-ZIM3	7.165	61.167	5.807	3
LTRP6-dual-ZIM3	7.165	68.133	7.305	3
LTRP6-22153-7255	7.165	90.033	6.144	3
LTRP6-dual-7255	7.165	87.833	2.219	3
LTRP6-26749-7255	7.165	84.700	1.539	3
LTRP6-10123-7255	7.165	83.833	3.859	3
LTRP6-dual-22153	7.165	73.000	0.721	3
LTRP6-7255-22153	7.165	71.267	7.343	3
LTRP6-26749-22153	7.165	92.533	3.062	3
LTRP6-10123-22153	7.165	92.133	2.542	3
LTRP6-22153-10123	7.165	87.533	1.069	3
LTRP6-7255-10123	7.165	69.533	5.265	3
LTRP6-26749-10123	7.165	92.667	0.833	3
LTRP6-dual-10123	7.165	80.133	2.458	3
LTRP6-22153-26749	7.165	92.433	2.804	3
LTRP6-7255-26749	7.165	87.800	8.143	3
LTRP6-dual-26749	7.165	97.133	0.643	3
LTRP6-10123-26749	7.165	94.067	1.686	3

Table 55: Level of B2M repression mediated by LTRP constructs with various repressor domains quantified at 104 days post-transfection. The order of the domain IDs shown for each LTRP6 construct designates the first repressor domain and second repressor domain at the N-terminus and C-terminus positions respectively

Repressor construct	Spacer	Mean % HLA-negative cells	Standard deviation	Sample Size
LTRP5-ADD-ZIM3	NT	0.610	0.314	3
LTRP5-ZIM3	7.165	14.270	12.499	3
LTRP5-ADD-ZIM3	7.165	57.533	7.419	3
LTRP6-dual-ZIM3	7.165	43.467	9.843	3
LTRP6-22153-7255	7.165	90.900	1.929	3

Repressor construct	Spacer	Mean % HLA-negative cells	Standard deviation	Sample Size
LTRP6-dual-7255	7.165	93.367	0.603	3
LTRP6-26749-7255	7.165	84.000	1.044	3
LTRP6-10123-7255	7.165	77.700	16.664	3
LTRP6-dual-22153	7.165	69.567	6.336	3
LTRP6-7255-22153	7.165	57.567	12.667	3
LTRP6-26749-22153	7.165	87.867	4.277	3
LTRP6-10123-22153	7.165	88.600	7.134	3
LTRP6-22153-10123	7.165	84.600	0.624	3
LTRP6-7255-10123	7.165	78.033	4.922	3
LTRP6-26749-10123	7.165	94.400	1.375	3
LTRP6-dual-10123	7.165	79.100	5.216	3
LTRP6-22153-26749	7.165	92.067	5.595	3
LTRP6-7255-26749	7.165	90.067	6.274	3
LTRP6-dual-26749	7.165	98.267	0.929	3
LTRP6-10123-26749	7.165	94.633	1.779	3

[0502] Bisulfite sequencing was performed to evaluate the degree of off-target CpG methylation at the *VEGFA* locus mediated by the 16 LTRP6 molecules at 5 days post-transfection, and the results are shown in Table 56. The data demonstrate that the majority of the LTRP6 constructs demonstrated relatively low off-target methylation levels, especially when compared with the levels achieved with constructs encoding for the dCas9-ZNF10-DNMT3A/3L, LTRP1, and LTRP5-ZIM3 controls. In fact, LTRP6 constructs containing DOMAIN_26749 at either or both positions (configuration 6a or 6b) exhibited comparable off-target methylation as the LTRP5-ADD-ZIM3 and LTRP6-dual-ZIM3 molecules (Table 56). LTRP6-dual-26749 and LTRP6-10123-26749, which exhibited the highest and second highest repressive activity, respectively (Tables 48-55), induced levels of off-target methylation of ~5.15% and ~3.84% respectively.

Table 56: Level of off-target CpG methylation at the *VEGFA* locus mediated by the LTRP6 constructs with various repressor domains quantified at day 5 post-transfection. The order of the domain IDs shown for each LTRP6 construct designates the first repressor domain and second repressor domain at the N-terminus and C-terminus positions respectively. Constructs encoding for CasX, dCas9-ZNF10-DNMT3A/3L, LTRP1, LTRP5-ZIM3, and LTRP5-ADD-ZIM3 were also included as controls

Repressor construct	Spacer	Mean % CpG methylation	Standard error
CasX 491	7.37	0.851	0.160
dCas9-ZNF10-DNMT3A/3L	7.148	34.220	2.356
LTRP1	NT	16.900	1.346

Repressor construct	Spacer	Mean % CpG methylation	Standard error
LTRP1	7.165	30.880	1.829
LTRP5-ADD-ZIM3	NT	6.887	0.403
LTRP5-ZIM5	7.165	14.690	1.338
LTRP5-ADD-ZIM3	7.165	4.841	0.508
LTRP6-dual-ZIM3	7.165	5.782	0.565
LTRP6-7255-22153	7.165	3.988	0.479
LTRP6-dual-7255	7.165	11.330	0.993
LTRP6-7255-26749	7.165	5.719	0.737
LTRP6-7255-10123	7.165	5.796	0.635
LTRP6-22153-26749	7.165	4.789	0.492
LTRP6-22153-10123	7.165	1.244	0.422
LTRP6-10123-22153	7.165	6.607	0.757
LTRP6-10123-7255	7.165	16.530	1.417
LTRP6-10123-26749	7.165	4.810	0.523
LTRP6-dual-10123	7.165	5.729	0.699
LTRP6-26749-22153	7.165	5.401	0.565
LTRP6-26749-7255	7.165	4.608	0.582
LTRP6-dual-22153	7.165	3.522	0.537
LTRP6-22153-7255	7.165	4.388	0.669
LTRP6-dual-26749	7.165	5.153	0.457
LTRP6-26749-10123	7.165	3.836	0.482

[0503] The results of the experiments demonstrate that incorporation of a second repressor domain into an LTRP molecule, resulting in an LTRP6 configuration, improves durable repression, and this improvement is further enhanced with the use of an enhanced repressor domain at both the N-terminus and C-terminus positions relative to dCasX. Furthermore, the data demonstrate that use of some, but not all, novel repressor domains can induce highly durable repression of the target locus, and that their relative placement within the LTRP molecule could also affect activity. The findings of these experiments also show that incorporation of a second repressor domain into an LTRP-ADD molecule maintains molecule specificity while improving activity.

Example 13: Optimization of linker sequences to increase overall LTRP activity

[0504] Experiments were performed to demonstrate that the repressive activity of an LTRP5 construct can be increased by optimizing the linker sequence at the following two positions: 1) between the DNMT3L interaction domain and the repressor domain and 2) between the repressor domain and the catalytically-dead CasX. These linker positions correspond to “L3A” and “L1” respectively in the schematic in FIG. 55 for the LTRP5

configuration. Here, experiments were performed to screen and identify new linker combinations that would improve LTRP activity.

Materials and Methods:

[0505] An entry vector was created to harbor a restriction enzyme site between the DNMT3L interaction domain and the repressor domain and between the repressor domain and the catalytically-dead CasX oriented in the LTRP5 configuration (FIG. 55). Furthermore, a list of 52 linkers was generated by compiling linker sequences from multiple studies. The library of linkers was synthesized as an oligonucleotide pool that was subsequently amplified and randomly cloned into the entry vector via Golden Gate assembly, resulting in a library of LTRP5 variant plasmids covering ~2700 possible linker sets. ~90 LTRP5 variant plasmids containing the linker sets were randomly selected and subjected to arrayed screening, such that each LTRP5 variant plasmid was co-transfected with a plasmid encoding for a gRNA using scaffold 316 and a targeting spacer (see Table 58 for spacer sequences) into HEK293T cells seeded in 96-well plates. Following transfection, cells were selected with puromycin and hygromycin for three days and then harvested to evaluate repression at the target locus via immunostaining of the target cell surface marker, followed by flow cytometry using the Attune™ NxT flow cytometer. For variant plasmids containing linker sets 1-28 (see Table 57 for sequences of linker combinations), three different target loci (*B2M*, target 1, and target 2) were evaluated for LTRP repression activity. For linker sets 1-11, repression was measured at 8, 15, and 45 days post-transfection. For linker sets 12-28, repression was measured at 7 and 17 days post-transfection. For plasmids containing linker sets 31-89, along with linker sets 1, 10, and 25 (Table 57), the *B2M* locus was evaluated for LTRP repression activity at 6, 13, and 24 days post-transfection. A plasmid encoding for the LTRP5 molecule (without the ADD domain) with the original set of linkers (SEQ ID NO: 123 for L1; SEQ ID NO: 124 for L3A) was used as an experimental control. A non-targeting spacer was also included as a negative control.

Table 57: AA sequences of the ~90 linker sets assessed in this example

Linker Set ID	AA sequence of first linker within set (position L3A)	AA sequence of second linker within set (position L1)	SEQ ID NO. (Position L3A)	SEQ ID NO. (Position L1)
Set 1	ASAAAPAAASAAASAPSA	ASAAAPAA	3281	3286
Set 36	PPTPSPSPVPSTPPTNSSSTPTPS	GSGNSSGSGGS	3280	3285
Set 43	GSGGSGGSGGSPVPSTPPTPSPSTPPTP	AFPAAPAPA	3279	3284
Set 66	AASPAAPSAPPAAASP	PPTP	3282	3287

Linker Set ID	AA sequence of first linker within set (position L3A)	AA sequence of second linker within set (position L1)	SEQ ID NO. (Position L3A)	SEQ ID NO. (Position L1)
Set 85	ASPAAPAPASPAAPAPSAPAA	IRAHGD	3278	3283
Set 2	LQFVPPQALKREQVSQQ	HIGGINS	22745	22748
Set 3	GGSGGGGS	GGSGGGGS	22746	22746
Set 4	AASPAAPSAPPAAASP	SPAGSPTSTEEGTSESAT	3282	22773
Set 5	PGTSTEPSEGSAPGSP	PPTP	22747	3287
Set 6	HIGGINS	GGGS	22748	22763
Set 7	AASPAAPSAPPAAASP	PPTPSPSPVP	3282	22768
Set 8	ASPAAPSP	ILGELT	22749	22774
Set 9	GSGNSSGSGG	HIGGINS	22750	22748
Set 10	PPTPSPSPVPSTPPTNSSSTPP TPS	SGSETPGTSESATPES	3280	22754
Set 11	HIGGINS	PPTPSPSPVP	22748	22768
Set 12	EAAKEAAK	PPTP	22751	3287
Set 13	PPTP	SWRVRYMVA	3287	22775
Set 14	LRTRYEADLA	GSGNSSGSGGS	22752	3285
Set 15	EAAKEAAKEAAKEAAK	PPTP	22753	3287
Set 17	SGSETPGTSESATPES	GGSGGGGS	22754	22746
Set 19	GSGNSSGSGGS	STEEGTSTEPSEGSAP	3285	22756
Set 20	PPTPSPSPVSTPPTPPTPS	EAAKEAAK	22755	22751
Set 21	STEEGTSTEPSEGSAP	PPTPSPSPVP	22756	22768
Set 23	GGSGGGGS	YGAHLM	22746	22776
Set 24	STEEGTSTEPSEGSAP	GSGNSSGSGSGSGSGNSSGS GGS	22756	22777
Set 25	PPTPSPSPVPSTPPTPPTPS	GGSGGGSGGGGS	22757	22762
Set 26	GGSGGGSGGGSEAAKEAAKEA AK	VQQKYKVSDTAATVTG	22758	22765
Set 27	ASAAAPAAA	GSGNSSGSGSGSGSGNSSGS GGS	3286	22777
Set 28	PPTP	GGSGGGGS	3287	22746
Set 31	SGSETPGTSESATPES	STEEGTSTEPSEGSAP	22754	22756
Set 32	GSGSGGGSGGSPVPSTPPTPSP STPPTP	SPAGSPTSTEEGTSESAT	3279	22773
Set 33	PPTP	GGGS	3287	22763
Set 34	ASPAAPAPA	GGGS	22759	22763
Set 37	GGSGGGGS	GGGS	22746	22763
Set 38	GGSGGGGS	SGSETPGTSESATPES	22746	22754
Set 39	GGSGGGSEAAKEAAK	STEEGTSTEPSEGSAP	22760	22756
Set 40	PPTPSPSPVPSTPPTPPTPS	GGSGGGSGGGGS	22757	22762
Set 41	GGSGGGSGGGSGGGGS	SPAGSPTSTEEGTSESAT	22761	22773
Set 42	GSGNSSGSGGS	AYVVSADEREGG	3285	22778
Set 44	SGSETPGTSESATPES	GSGNSSGSGGS	22754	3285
Set 45	GGSGGGSGGGGS	PGTSTEPSEGSAPGSP	22762	22747
Set 46	GGSGGGSGGGSGGGGS	EAAKEAAK	22761	22751
Set 47	GSGNSSGSGGS	PPTPSPSPVPSTPPTPPTPS	3285	22757
Set 48	GGSGGGSEAAKEAAK	EAAKEAAKGGSGGGGS	22760	22771
Set 49	GGGS	GGSGGGGS	22763	22746
Set 50	EAAK	GSGGGSGGSPVPSTPPTP SPSTPPTP	22764	3279
Set 51	EAAKEAAK	EAAKEAAK	22751	22751
Set 52	SGSETPGTSESATPES	PPTP	22754	3287

Linker Set ID	AA sequence of first linker within set (position L3A)	AA sequence of second linker within set (position L1)	SEQ ID NO. (Position L3A)	SEQ ID NO. (Position L1)
Set 53	GGGSGGGS	PPTP	22746	3287
Set 55	GGGSGGGS	MIGP	22762	22779
Set 56	VQQKYKVSDTAATVTG	SGSETPGTSESATPES	22765	22754
Set 57	GGGSGGGS	GSGGSGGSGGSPVPSTPPTP	22766	3279
Set 58	EAAKEAAKEAAKEAAK	STEEGTSTEPSEGSAP	22753	22756
Set 59	GGGSGGGS	EAAKEAVKGGSGGGS	22746	22780
Set 60	GGGSGGGS	PPTP	22761	3287
Set 61	EAAK	EAAKEAAKGGSGGGS	22764	22771
Set 62	GGGS	PPTPSPSPVP	22763	22768
Set 63	EAAKEAVKEAAKEAAK	HIGGINS	22767	22748
Set 64	PPTPSPSPVPSTPPTPPTPS	GGSGGGS	22757	22746
Set 65	PPTPSPSPVP	PPTP	22768	3287
Set 67	STEEGTSTEPSEGSAP	PPTP	22756	3287
Set 68	GGGSGGGS	PPTPSPSPVPSTPPTNSSST	22762	3280
Set 69	PPTPSPSPVP	PGTSTEPSEGSAPGSP	22768	22747
Set 70	ASPAAPAPA	LMPYEL	22759	22781
Set 71	PPTPSPSPVP	PPTP	22768	3287
Set 72	GGGSGGGS	ASAAAPAAA	22762	3286
Set 73	GGGSGGGS	STEEGTSTEPSEGSAP	22746	22756
Set 74	PPTPSPSPVP	HIGGINS	22768	22748
Set 75	EAAKEAAK	HIGGINS	22751	22748
Set 76	PPTPSPSPVPSTPPTPPTPS	SPTP	22757	22782
Set 77	EAAKEAAKEAAKEAAK	GGGSGGGS	22753	22742
Set 78	EAAKEAAK	PPTPSPSPVPSTPPTPPTPS	22751	22757
Set 79	PPTPSPSPVP	GGGSGGGS	22768	22761
Set 80	ITHKASSRMTFA	GGGSGGGS	22769	22762
Set 81	STEEGTSTEPSEGSAP	GCGS	22756	22783
Set 82	ASSAAPAPA	GGGS	22770	22763
Set 83	STEEGTSTEPSEGSAP	GGSGGGS	22756	22746
Set 84	SGSETPGTSESATPES	GGSGGGS	22754	22742
Set 86	EAAKEAAKGGSGGGS	EAAKEAAKGGSGGGS	22771	22771
Set 87	EAAKEAAKEAAKEAAK	NKQNIWQA	22753	22784
Set 88	GGGSGGGS	ASAAAPAAA	22762	3286
Set 89	AASPAAPSAPPASAP	PPMPSPSPVP	22772	22785

Table 58: Sequences of spacers targeting the *B2M* locus

Spacer ID	Target	Targeting spacer sequence (RNA)	SEQ ID NO
7.165	<i>B2M</i>	UCCCUAUGUCCUUGCUGUUU	3114

Results:

[0506] Arrayed screening was performed to evaluate the level of repressive activity at endogenous target loci in HEK293T cells transfected with LTRP5 variant plasmids harboring different combinations of linkers. Repression at the *B2M*, target 1, and target 2 loci was measured at 8, 15, and 45 days post-transfection for LTRP5 plasmids containing linker sets 1-

11, and the results are illustrated in FIGS. 56-58. The data demonstrate that the top performing linker combinations varied by the specific target. Of the 11 linker combinations analyzed in this experiment, use of linker sets 1 and 10 overall consistently induced the highest level of repression at both the *B2M* and target 2 loci, with linker set 10 exhibiting the second highest tested linker set at the target 1 locus. Linker 1 was subsequently used as a benchmark in the subsequent two experiments analyzing linker sets 12-28 and linker sets 31-89.

[0507] Repression at the *B2M*, target 1, and target 2 loci was measured at 7 and 17 days post-transfection for LTRP5 plasmids containing linker sets 12-28, with linker set 1 from the previous experiment as a benchmark, and the results are portrayed in FIGS. 59-60. The data show that as similarly observed in the previous experiment, the top performing linker sets varied by target. Of the 17 linker combinations analyzed in this second experiment, use of linker sets 1 and 25 resulted in the highest level of repression at two of the three target loci (*B2M* and target 1 for set 1; *B2M* and target 2 for set 25).

[0508] Repression at the *B2M* locus was measured at 6, 13, and 24 days post-transfection for LTRP5 plasmids containing linker sets 31-89, with linker sets 1, 10, and 25 from the previous two experiments as benchmarks, and the results are portrayed in Tables 59-61. The data show that use of linker sets 85, 43, 36, 10, 64, 1, 40, 25, 66, 76, and 52 resulted in at least a 50% increase in repressive activity relative to the activity level of the LTRP5 control when paired with the *B2M* spacer at 24 days post-transfection (Tables 59-61). Five of these top 11 performing linker sets were then selected for future validation based on having a sequence distinct from other top candidate combinations and from the original linker set of linkers (SEQ ID NO: 123 for L1; SEQ ID NO: 124 for L3A).

Table 59: Levels of B2M repression mediated by LTRP5 constructs with linker sets 1, 10, 25, and 31-39 quantified at 6 days post-transfection

Construct	Linker set	Spacer	Mean % HLA-negative cells	Standard deviation	Sample size
LTRP5	Original/Basel line	NT	4.607	0.289	3
LTRP5	Set 54	7.165	31.000	2.955	3
LTRP5	Set 69	7.165	33.400	1.015	3
LTRP5	Set 42	7.165	35.167	1.258	3
LTRP5	Set 89	7.165	34.733	0.603	3
LTRP5	Original/Basel line	7.165	38.667	0.379	3
LTRP5	Set 61	7.165	48.300	1.931	3

Construct	Linker set	Spacer	Mean % HLA-negative cells	Standard deviation	Sample size
LTRP5	Set 59	7.165	49.467	0.709	3
LTRP5	Set 80	7.165	48.633	0.643	3
LTRP5	Set 33	7.165	53.800	1.200	3
LTRP5	Set 37	7.165	55.167	0.416	3
LTRP5	Set 51	7.165	51.900	0.819	3
LTRP5	Set 49	7.165	52.267	0.723	3
LTRP5	Set 48	7.165	50.500	1.082	3
LTRP5	Set 62	7.165	53.367	0.493	3
LTRP5	Set 41	7.165	53.067	0.473	3
LTRP5	Set 50	7.165	53.400	0.458	3
LTRP5	Set 38	7.165	53.833	0.231	3
LTRP5	Set 78	7.165	52.633	0.379	3
LTRP5	Set 86	7.165	53.833	1.484	3
LTRP5	Set 53	7.165	55.167	0.764	3
LTRP5	Set 87	7.165	55.533	0.833	3
LTRP5	Set 46	7.165	53.033	0.153	3
LTRP5	Set 55	7.165	50.633	1.650	3
LTRP5	Set 34	7.165	57.800	0.794	3
LTRP5	Set 75	7.165	53.400	0.794	3
LTRP5	Set 39	7.165	55.467	1.002	3
LTRP5	Set 56	7.165	54.000	1.114	3
LTRP5	Set 60	7.165	57.933	1.328	3
LTRP5	Set 82	7.165	57.500	0.889	3
LTRP5	Set 73	7.165	53.433	0.907	3
LTRP5	Set 83	7.165	59.333	1.069	3
LTRP5	Set 63	7.165	55.400	0.872	3
LTRP5	Set 88	7.165	58.567	1.436	3
LTRP5	Set 68	7.165	55.633	0.231	3
LTRP5	Set 58	7.165	54.433	0.723	3
LTRP5	Set 79	7.165	56.700	0.520	3
LTRP5	Set 70	7.165	55.467	1.350	3
LTRP5	Set 32	7.165	58.467	0.723	3
LTRP5	Set 77	7.165	56.133	0.651	3
LTRP5	Set 81	7.165	60.400	0.500	3
LTRP5	Set 57	7.165	57.800	1.153	3
LTRP5	Set 44	7.165	57.933	1.617	3
LTRP5	Set 47	7.165	55.400	0.173	3
LTRP5	Set 31	7.165	59.767	0.751	3
LTRP5	Set 74	7.165	58.067	0.569	3
LTRP5	Set 65	7.165	60.300	0.954	3
LTRP5	Set 84	7.165	59.333	0.902	3
LTRP5	Set 67	7.165	60.767	1.779	3
LTRP5	Set 72	7.165	57.433	0.723	3
LTRP5	Set 71	7.165	59.467	0.379	3
LTRP5	Set 52	7.165	61.533	0.153	3
LTRP5	Set 76	7.165	62.900	0.400	3
LTRP5	Set 66	7.165	61.933	0.379	3
LTRP5	Set 25	7.165	61.033	0.416	3
LTRP5	Set 40	7.165	62.800	0.889	3
LTRP5	Set 1	7.165	58.133	0.058	3

Construct	Linker set	Spacer	Mean % HLA-negative cells	Standard deviation	Sample size
LTRP5	Set 64	7.165	57.400	1.229	3
LTRP5	Set 10	7.165	59.967	1.222	3
LTRP5	Set 36	7.165	65.333	0.902	3
LTRP5	Set 43	7.165	64.800	1.253	3
LTRP5	Set 85	7.165	64.333	1.274	3

Table 60: Levels of B2M repression mediated by LTRP5 constructs with linker sets 1, 10, 25, and 31-39 quantified at 13 days post-transfection

Construct	Linker set	Spacer	Mean % HLA-negative cells	Standard deviation	Sample size
LTRP5	Original/Basel line	NT	1.877	0.315	3
LTRP5	Set 54	7.165	2.670	0.866	3
LTRP5	Set 69	7.165	1.200	0.190	3
LTRP5	Set 42	7.165	24.000	0.794	3
LTRP5	Set 89	7.165	25.500	0.755	3
LTRP5	Original/Basel line	7.165	30.800	0.917	3
LTRP5	Set 61	7.165	31.600	0.700	3
LTRP5	Set 59	7.165	32.267	0.473	3
LTRP5	Set 80	7.165	32.767	0.231	3
LTRP5	Set 33	7.165	36.833	1.137	3
LTRP5	Set 37	7.165	35.900	1.179	3
LTRP5	Set 51	7.165	34.767	0.723	3
LTRP5	Set 49	7.165	36.167	0.929	3
LTRP5	Set 48	7.165	36.267	1.274	3
LTRP5	Set 62	7.165	34.433	0.551	3
LTRP5	Set 41	7.165	35.067	1.484	3
LTRP5	Set 50	7.165	38.067	0.874	3
LTRP5	Set 38	7.165	36.567	0.351	3
LTRP5	Set 78	7.165	35.000	0.700	3
LTRP5	Set 86	7.165	34.933	1.747	3
LTRP5	Set 53	7.165	35.833	1.234	3
LTRP5	Set 87	7.165	35.800	1.015	3
LTRP5	Set 46	7.165	36.200	0.361	3
LTRP5	Set 55	7.165	36.033	1.950	3
LTRP5	Set 34	7.165	40.733	1.168	3
LTRP5	Set 75	7.165	36.033	1.604	3
LTRP5	Set 39	7.165	38.867	1.266	3
LTRP5	Set 56	7.165	37.467	0.764	3
LTRP5	Set 60	7.165	40.467	1.750	3
LTRP5	Set 82	7.165	38.567	1.069	3
LTRP5	Set 73	7.165	35.433	1.845	3
LTRP5	Set 83	7.165	39.100	0.889	3
LTRP5	Set 63	7.165	37.433	1.002	3
LTRP5	Set 88	7.165	39.033	1.930	3
LTRP5	Set 68	7.165	37.833	0.493	3

Construct	Linker set	Spacer	Mean % HLA-negative cells	Standard deviation	Sample size
LTRP5	Set 58	7.165	38.700	0.866	3
LTRP5	Set 79	7.165	39.367	0.153	3
LTRP5	Set 70	7.165	37.467	1.701	3
LTRP5	Set 32	7.165	42.833	0.289	3
LTRP5	Set 77	7.165	38.767	0.907	3
LTRP5	Set 81	7.165	39.300	1.039	3
LTRP5	Set 57	7.165	41.233	1.779	3
LTRP5	Set 44	7.165	40.667	0.814	3
LTRP5	Set 47	7.165	39.667	0.850	3
LTRP5	Set 31	7.165	43.433	1.650	3
LTRP5	Set 74	7.165	40.467	0.802	3
LTRP5	Set 65	7.165	42.000	0.608	3
LTRP5	Set 84	7.165	41.700	0.200	3
LTRP5	Set 67	7.165	41.600	1.277	3
LTRP5	Set 72	7.165	39.400	0.781	3
LTRP5	Set 71	7.165	42.533	0.153	3
LTRP5	Set 52	7.165	44.333	0.635	3
LTRP5	Set 76	7.165	45.033	1.721	3
LTRP5	Set 66	7.165	43.233	0.833	3
LTRP5	Set 25	7.165	49.233	0.231	3
LTRP5	Set 40	7.165	47.767	1.002	3
LTRP5	Set 1	7.165	48.100	0.917	3
LTRP5	Set 64	7.165	44.133	1.097	3
LTRP5	Set 10	7.165	48.733	1.069	3
LTRP5	Set 36	7.165	48.833	1.358	3
LTRP5	Set 43	7.165	49.267	1.222	3
LTRP5	Set 85	7.165	46.933	2.312	3

Table 61: Levels of B2M repression mediated by LTRP5 constructs with linker sets 1, 10, 25, and 31-39 quantified at 24 days post-transfection

Construct	Linker set	Spacer	Mean % HLA-negative cells	Standard deviation	Sample size
LTRP5	Original/Basel line	NT	1.290	0.098	3
LTRP5	Set 54	7.165	3.670	1.105	3
LTRP5	Set 69	7.165	4.270	0.282	3
LTRP5	Set 42	7.165	22.600	1.153	3
LTRP5	Set 89	7.165	24.500	0.781	3
LTRP5	Original/Basel line	7.165	25.067	0.907	3
LTRP5	Set 61	7.165	27.100	0.781	3
LTRP5	Set 59	7.165	27.167	0.351	3
LTRP5	Set 80	7.165	28.367	1.185	3
LTRP5	Set 33	7.165	28.367	0.208	3
LTRP5	Set 37	7.165	28.833	0.751	3
LTRP5	Set 51	7.165	28.933	0.551	3
LTRP5	Set 49	7.165	28.967	1.950	3

Construct	Linker set	Spacer	Mean % HLA-negative cells	Standard deviation	Sample size
LTRP5	Set 48	7.165	29.467	0.757	3
LTRP5	Set 62	7.165	29.633	0.551	3
LTRP5	Set 41	7.165	29.767	1.457	3
LTRP5	Set 50	7.165	30.000	0.436	3
LTRP5	Set 38	7.165	30.167	1.050	3
LTRP5	Set 78	7.165	30.200	1.480	3
LTRP5	Set 86	7.165	30.367	1.041	3
LTRP5	Set 53	7.165	30.567	2.055	3
LTRP5	Set 87	7.165	30.733	1.258	3
LTRP5	Set 46	7.165	31.033	0.723	3
LTRP5	Set 55	7.165	31.167	2.413	3
LTRP5	Set 34	7.165	31.833	0.802	3
LTRP5	Set 75	7.165	31.867	2.892	3
LTRP5	Set 39	7.165	31.867	0.874	3
LTRP5	Set 56	7.165	32.267	0.681	3
LTRP5	Set 60	7.165	32.733	1.206	3
LTRP5	Set 82	7.165	32.967	1.274	3
LTRP5	Set 73	7.165	33.133	0.351	3
LTRP5	Set 83	7.165	33.333	0.569	3
LTRP5	Set 63	7.165	33.400	2.551	3
LTRP5	Set 88	7.165	33.500	1.803	3
LTRP5	Set 68	7.165	33.533	0.666	3
LTRP5	Set 58	7.165	33.633	1.266	3
LTRP5	Set 79	7.165	34.000	0.100	3
LTRP5	Set 70	7.165	34.167	2.836	3
LTRP5	Set 32	7.165	34.167	1.193	3
LTRP5	Set 77	7.165	34.467	0.814	3
LTRP5	Set 81	7.165	34.733	0.404	3
LTRP5	Set 57	7.165	34.833	2.250	3
LTRP5	Set 44	7.165	34.933	1.012	3
LTRP5	Set 47	7.165	35.167	0.802	3
LTRP5	Set 31	7.165	35.433	2.060	3
LTRP5	Set 74	7.165	35.833	0.651	3
LTRP5	Set 65	7.165	36.467	0.681	3
LTRP5	Set 84	7.165	37.167	1.069	3
LTRP5	Set 67	7.165	37.300	1.311	3
LTRP5	Set 72	7.165	37.667	0.850	3
LTRP5	Set 71	7.165	37.733	0.681	3
LTRP5	Set 52	7.165	38.767	0.929	3
LTRP5	Set 76	7.165	39.167	1.858	3
LTRP5	Set 66	7.165	39.367	0.503	3
LTRP5	Set 25	7.165	40.400	1.808	3
LTRP5	Set 40	7.165	40.633	1.629	3
LTRP5	Set 1	7.165	41.000	1.127	3
LTRP5	Set 64	7.165	41.367	1.877	3
LTRP5	Set 10	7.165	42.167	2.113	3
LTRP5	Set 36	7.165	42.333	1.250	3
LTRP5	Set 43	7.165	42.433	0.666	3
LTRP5	Set 85	7.165	42.500	2.272	3

[0509] The experiments show that the repressive activity of the LTRP molecule can be further enhanced by altering the linker sequences positioned in between key domains of the LTRP molecule. Here, optimizing the linker sequences at two positions, i.e., between the DNMT3L interaction domain and the repressor domain and between the repressor domain and the catalytically-dead CasX, resulted in substantial improvement in repressive activity of the LTRP molecule.

Example 14: Evaluation of use of the ADD domain in LTRP in a mouse model

[0510] Experiments were performed to assess durability of methylation upon treatment with a CasX:gRNA system when delivered *in vivo* in a mouse model. LTRP1 orientation molecules were evaluated for durability of methylation, mPCSK9 transcript knockdown, and mPCSK9 protein production in the presence and absence of the ADD domain.

Materials and Methods:

LNP preparation:

[0511] LTRP mRNA and gRNA were encapsulated into LNPs using ALC-0315 based lipid mix using a custom-made T-mixer micro mixing device, at a flow rate of 20 mL/min and 3:1 mixing ratio of aqueous to organic phase. The composition was the following: ionizable lipid:DSPC:cholesterol:DMG-PEG2000 at 50:10:38.5:1.5 mol%. Briefly, to formulate LNPs, equal mass ratios of LTRP mRNA and gRNA were diluted in 25 mM sodium acetate, pH 4.0. The lipid mix is made at a 10 mM concentration in anhydrous ethanol. mRNA/gRNA co-formulations were generated using a predetermined N/P ratio. The RNA and lipids were run through a custom-made T-mixer device at a predetermined flow rate ratio using syringe pump infusers. After formulation, the LNPs were dialyzed into PBS, pH 7.4, to decrease the ethanol concentration and increase the pH, which increases the stability of the particles. Buffer exchange of the mRNA/gRNA-LNPs was achieved by overnight dialysis into PBS, pH 7.4, at 4°C using 10k Slide-A-Lyzer™ Dialysis or Cassettes (Thermo Scientific™) or 12-14 kDa dialysis tubing (Repligen). Following dialysis, the mRNA/gRNA-LNPs was concentrated to > 0.2 mg/mL using 30-100 kDa Amicon®-Ultra Centrifugal Filters (Millipore) and then sterile-filtered using Acrodisc PES membrane filters. Formulated LNPs were analyzed on a Malvern Zetasizer to determine their diameter and polydispersity index (PDI). Encapsulation efficiency and RNA concentration was determined by RiboGreen™ assay using Invitrogen's Quant-iT™ RiboGreen™ RNA assay kit. The LNPs described above were used in various experiments to deliver LTRP mRNA and gRNA for delivery to target tissues *in vivo*.

Delivery of LNPs encapsulating CasX mRNA and targeting gRNA *in vivo*:

[0512] To assess the effects of using LTRP1A and LTRP1B *in vivo*, an LTRP mRNA (see Table 62) and a *mPCSK9*-targeting gRNA using scaffold 316 v1 with spacer 27.94 (SEQ ID: 3330, Table 63) were encapsulated within the same LNP using a 1:1 mass ratio for mRNA:gRNA. Formulated LNPs were buffer-exchanged to PBS for *in vivo* injection. Briefly, LNPs were administered at 3.0 mg per kg intravenously through the tail-vein into 4-week-old C57BL/6 mice. Mice were observed for five minutes after injection to ensure recovery from anesthesia before being placed into their home cage. Vehicle (PBS) injected animals served as negative experimental controls and mice injected with mRNA:gRNA encoding for LTRP1 and the mPCSK9 targeting guide above served as positive controls. Baseline was taken via fasted blood draw. Further blood draws were taken every three weeks. Seven, fourteen, and forty-two days post-administration, 3 mice from each condition were euthanized, the blood and liver tissue were harvested. Blood serum was collected for mPCSK9 ELISA, liver tissue was homogenized for mRNA extraction and for gDNA extraction using the Zymo Research Quick DNA/RNA Miniprep kit following the manufacturer's instructions.

Enzymatic methylation sequencing (EM-seq):

[0513] To determine on-target methylation levels at the *PCSK9* locus, gDNA from harvested tissues was extracted using the Zymo Quick-DNA Miniprep Plus kit following the manufacturer's instructions. The extracted gDNA was then subjected to enzymatic methylation conversion using the Enzymatic Methyl-seq Conversion Module (NEB) following the manufacturer's protocol, converting any non-methylated cytosine into uracil. The resulting treated DNA was subsequently sequenced using next-generation sequencing (NGS) to determine the levels of on-target methylation.

[0514] NGS processing and analysis:

[0515] Target amplicons were amplified from 50 ng EM-treated DNA via PCR with a set of primers specific to the EM-converted target locations of interest (mouse *PCSK9* locus). These gene-specific primers contained an additional sequence at the 5' end to introduce an IlluminaTM adapter. Amplified DNA products were purified with the Cytiva Sera-Mag Select DNA cleanup kit. Quality and quantification of the amplicon were assessed using a Fragment Analyzer DNA Analysis kit (Agilent, dsDNA 35-1500 bp). Amplicons were sequenced on the IlluminaTM MiseqTM according to the manufacturer's instructions. Raw fastq files from sequencing were processed using Bismark Bisulfite Read Mapper and Methylation caller. PCR amplification of the EM-treated DNA converted all uracil nucleotides into thymine, and

sequencing of the PCR product determined the rate of cytosine-to-thymine conversion as a readout of the level of on-target methylation at the PCSK9 locus mediated by each LTRP molecule.

qPCR assay for mPCSK9 mRNA knockdown:

[0516] Liver pieces were flash frozen after necropsy. These were then immersed in excess volume of zymo DNA/RNA shield reagent and homogenized using bead beating tubes (bullet blender instrument, with ceramic or steel beads in tubes). After homogenization, lysates were treated with proteinase k. Proteinase k-treated liver lysates had mRNA extracted using zymo quick RNA miniprep kit (cat#r1055) according to manufacturer's instructions and assayed using the taqman gene expression assay with mPCSK9 and eukaryotic 18s probes following manufacturer's instructions.

ELISA method:

[0517] Serum from animal samples were assayed using legend max mouse PCSK9 ELISA kits (biolegend cat#443207) following manufacturer's instructions.

Results:

[0518] To assess durability of methylation upon treatment with a CasX:gRNA system when delivered *in vivo* in a mouse model, mRNAs encoding LTRP1 (SEQ ID NO. 3333), LTRP1A (SEQ ID NO: 6528), or LTRP1B (SEQ ID NO: 9741) and gRNA 316v1.27.94 (SEQ ID NO: 3330) were encapsulated in LNPs and injected into C57BL/6 mice. The quantification of the resulting PCSK9 mRNA knockdown is shown in Table 64. The data demonstrate comparable levels of knockdown across all three LTRP1 orientation molecules at 7 and 14 days, with LTRP1 and LTRP1A showing a higher level of suppression at day 42. Supporting this result, methylation data from EM-seq (FIG. 62) shows high levels of methylation across all three LTRP1 orientation molecules tested, with slight decrease total percent methylation of mice treated with LTRP1 and LTRP1B at days 14 and 42.

[0519] PCSK9 secretion quantification shows similar levels of PCSK9 suppression at day 7 (week 1, Table 65) across all LTRP molecules tested. LTRP1A shows increased durability of repression through week 6 of testing, while LTRP1B shows increased durability of repression through week 8 as compared to LTRP1.

[0520] These experiments demonstrate that LTRP1 orientation molecules with and without the ADD domain delivered via LNPs as mRNA and are able to induce effective silencing *in vivo*.

Table 62: mRNA sequences of LTRP constructs tested in this example

LTRP	SEQ ID NO.	Components
LTRP1	3333	LTRP1 ZIM3, 80 poly(A), SV40, C-terminal 2xFLAG
LTRP1A	6528	LTRP1 ZIM3, split poly(A), SV40
LTRP1B	9741	LTRP1 ZIM3, ADD, split poly(A), SV40

Table 63: gRNA and targeting sequence used in this example

gRNA ID	SEQ ID NO.
27.94	3071
316v1.27.94	3330

Table 64: Liver lysate mRNA qPCR

	Days post-dose					
LTRP	7 Days		14 Days		42 Days	
	Mean	std	Mean	std	Mean	std
Vehicle (N=3)	7.14	49.81	5.24	42.63	1.00	17.10
LTRP1 (N=3)	-96.55	1.77	-96.96	1.21	-93.65	2.60
LTRP1A (N=3)	-99.04	0.66	-99.76	0.09	-93.19	4.73
LTRP1B (N=3)	-96.13	1.47	-92.39	5.65	-80.55	5.79

Table 65: Change in Secreted PCSK9 by ELISA

Percent secreted PCSK9 by serum ELISA, normalized to baseline								
Study Week	Vehicle		LTRP1		LTRP1A		LTRP1B	
	Mean	SEM	Mean	SEM	Mean	SEM	Mean	SEM
1	23.80	45.23	-94.45	0.45	-93.50	0.91	-94.30	0.60
2	-15.06	14.48	-92.20	1.96	-85.10	7.07	-93.69	0.44
5	4.98	10.38	-90.18	2.78	-95.80	0.78	-93.37	0.42
6	-11.92	27.13	-92.65	1.72	-95.81	0.79	-90.36	2.14
8	17.84	9.53	-80.96	6.72	NA	NA	-90.58	1.20

Values assayed below LOQ were set to LOQ (6.25 ng/mL), LTRP1A week 8 data was not collected.

Example 15: Efficacy and durability of LTRP with RD1 domains

[0521] Experiments were performed to assess durability of methylation upon treatment with LTRP1 orientation molecules containing RD1 domains *in vivo* in a mouse model.

Materials and Methods:

[0522] LNPs were prepared as described in example 14.

[0523] An *in vivo* experiment examined the efficacy of LTRP1 orientation molecules containing a King Cobra RD1 domain (SEQ ID NO. 3339). LTRP1A-Cobra, and LTRP1B-Cobra (SEQ ID NO. 6530, SEQ ID NO. 9743) were built by substituting the ZIM3 domain in LTRP1A and LTRP 1B for the RD1 domain. Formulated LNPs containing an LTRP mRNA (Table 66) and mPCSK9 targeting gRNA (SEQ ID: 3330, Table 66) were buffer-exchanged to PBS for *in vivo* injection. Briefly, LNPs were administered intravenously through the tail-vein into 4-week old C57BL/6 mice. Mice were observed for five minutes after injection to ensure recovery from anesthesia before being placed into their home cage. Vehicle (PBS) injected animals served as negative experimental controls and mice injected with mRNA:gRNA encoding for LTRP1A (SEQ ID: 6528) and *mPCSK9*-targeting gRNA using scaffold 316 with spacer 27.94 (SEQ ID: 3330) served as positive controls. Baseline was taken via fasted blood draw. Further blood draws were taken every three weeks. Seven and forty-two days post-administration, mice were euthanized, the blood and liver tissue were harvested. Blood serum was collected for mPCSK9 ELISA, liver tissue was homogenized for mRNA extraction and for gDNA extraction using the Zymo Research Quick DNA/RNA Miniprep kit following the manufacturer's instructions.

Enzymatic methylation sequencing (EM-seq):

[0524] To determine on-target methylation levels at the *PCSK9* locus, gDNA from harvested tissues was extracted using the Zymo Quick-DNA Miniprep Plus kit following the manufacturer's instructions. The extracted gDNA was then subjected to enzymatic methylation conversion using the Enzymatic Methyl-seq Conversion Module (NEB) following the manufacturer's protocol, converting any non-methylated cytosine into uracil. The resulting treated DNA was subsequently sequenced using next-generation sequencing (NGS) to determine the levels of on-target methylation.

NGS processing and analysis:

[0525] Target amplicons were amplified from 50 ng EM-treated DNA via PCR with a set of primers specific to the EM-converted target locations of interest (mouse *PCSK9* locus). These gene-specific primers contained an additional sequence at the 5' end to introduce an IlluminaTM adapter. Amplified DNA products were purified with the Cytiva Sera-Mag Select DNA cleanup kit. Quality and quantification of the amplicon were assessed using a Fragment Analyzer DNA Analysis kit (Agilent, dsDNA 35-1500 bp). Amplicons were sequenced on the IlluminaTM MiseqTM according to the manufacturer's instructions. Raw fastq files from sequencing were processed using Bismark Bisulfite Read Mapper and Methylation caller. PCR amplification of the EM-treated DNA converts all uracil nucleotides into thymine, and

sequencing of the PCR product determines the rate of cytosine-to-thymine conversion as a readout of the level of on-target methylation at the *PCSK9* locus mediated by each LTRP molecule.

qPCR assay for mPCSK9 mRNA knockdown:

[0526] Liver pieces were flash frozen after necropsy. These were then immersed in excess volume of Zymo DNA/RNA Shield Reagent and homogenized using bead beating tubes (Bullet Blender instrument, with ceramic or steel beads in tubes). After homogenization, lysates were treated with Proteinase K. Proteinase K-treated liver lysates had mRNA extracted using Zymo Quick RNA Miniprep Kit (cat#R1055) according to manufacturer's instructions and assayed using the TaqMan gene expression assay with mPCSK9 and Eukaryotic 18S probes following manufacturer's instructions.

ELISA method:

[0527] Serum from animal samples were assayed using LEGEND MAX mouse PCSK9 ELISA kits (BioLegend cat#443207) following manufacturer's instructions.

Results:

[0528] Transcriptional start site (TSS)-proximal DNA methylation levels measured from homogenized liver-extracted gDNA by amplicon enzymatic methylation sequencing (EM-seq) from N=3 mice sacrificed at day 7 post-treatment demonstrate similar levels of methylation across all constructs, with higher median methylation trending with the higher dose (FIG. 63).

[0529] Table 67 quantifies the serum PCSK9 ELISA results, demonstrating that all LTRP1 constructs show similar initial effect, with knockdown of serum PCSK9 at ~90% for groups dosed with either 0.75 or 1.5mpk through 42 days post dosing.

[0530] Table 68 quantifies the mPCSK9 mRNA qPCR assay results, demonstrating robust transcript reduction of greater than 80% knockdown is maintained by all 3 LTRP1 constructs at both doses. LTRP1A, LTRP1A-Cobra, and LTRP1B-Cobra appear potent and durable at the mRNA transcript level up to 42 days post-dose. Both LTRP1A and LTRP1A-Cobra show a slight uptick in mPCSK9 expression at 42 days at the lower dose (0.75 mg/kg), but it is not clear if this represents a trend.

[0531] These experiments demonstrate that LTRP1 orientation molecules can produce durable methylation leading to reduction in PCSK9 expression with RD1 domains in a mouse model.

Table 66: Sequences of mRNA and gRNA assessed

mRNA	SEQ ID NO.	gRNA	Dose mg per kg
LTRP1A	6528	316v1 27.94 (SEQ ID NO. 3330)	1.5
			0.75
LTRP1A-Cobra	6530		1.5
			0.75
LTRP1B-Cobra	9743		1.5
			0.75
PBS (neg. control)	NA	NA	NA

Table 67: Serum PCSK9 ELISA at Day 7, 14, 35, and 42*

Serum PCSK9 % change from baseline						
Vehicle (PBS)				Vehicle (PBS)		
Day	Mean	SEM	n	Mean	SEM	n
0	0	0	9	N/A	N/A	N/A
7	50.87	79.48	3	N/A	N/A	N/A
14	89.78	86.04	6	N/A	N/A	N/A
35	-19.45	80.54	6	N/A	N/A	N/A
42	111.26	6.19	2	N/A	N/A	N/A
LTRP1A-Cobra (0.75 mpk)				LTRP1A-Cobra (1.5 mpk)		
Day	Mean	SEM	n	Mean	SEM	n
0	0	0	9	0	0	9
7	-90.42	3.47	3	-87.40	2.37	3
14	-87.51	9.17	6	-87.03	5.74	6
35	-83.73	12.37	6	-87.03	5.74	6
42	-91.31	2.87	2	-93.43	1.15	2
LTRP1B-Cobra (0.75 mpk)				LTRP1B-Cobra (1.5 mpk)		
Day	Mean	SEM	n	Mean	SEM	n
0	0	0	9	0	0	9
7	-90.22	4.27	3	-87.99	4.45	3
14	-88.43	6.86	6	-91.63	2.61	6
35	-89.11	3.79	6	-91.63	2.61	6
42	-88.33	1.86	2	-90.18	2.72	2
LTRP1A (0.75 mpk)				LTRP1A (1.5 mpk)		
Day	Mean	SEM	n	Mean	SEM	n
0	0	0	9	0	0	9
7	-91.57	1.12	3	-87.94	2.11	3
14	-88.21	4.21	6	-90.87	2.15	6
35	-86.95	4.41	6	-90.87	2.15	6
42	-70.89	21.14	2	-89.01	1.36	2

* Day 7 data comes from only animals sampled at that timepoint. Day 7 and 35 data comes from different individuals that continued in-life to Day 42. Day 42 data comes from only animals sampled at that timepoint.

Table 68: Percent repression of mPCSK9 mRNA at Days 7 and 42 via qPCR

mPCSK9 transcript/18S percent repression						
	Vehicle			Vehicle		
Day	Mean	SEM	n	Mean	SEM	n
0	0	0	3	N/A	N/A	N/A
7	19.65	49.23	3	N/A	N/A	N/A
42	9.38	44.25	2	N/A	N/A	N/A
	LTRP1A-Cobra (0.75 mpk)			LTRP1A-Cobra (1.5 mpk)		
Day	Mean	SEM	n	Mean	SEM	n
0	0	0	3	0	0	3
7	-92.90	1.48	3	-96.68	2.01	3
42	-79.37	0.57	2	-87.63	6.35	2
	LTRP1B-Cobra (0.75 mpk)			LTRP1B-Cobra (1.5 mpk)		
Day	Mean	SEM	n	Mean	SEM	n
0	0	0	3	0	0	3
7	-91.54	4.95	3	-98.15	0.55	3
42	-90.86	0.57	2	-96.02	1.31	2
	LTRP1A (0.75 mpk)			LTRP1A (1.5 mpk)		
Day	Mean	SEM	n	Mean	SEM	n
0	0	0	3	0	0	3
7	-92.79	1.21	3	-95.36	1.81	3
42	-84.55	5.15	2	-91.95	4.47	2

Example 16: Functional assessments of CasX:gRNA systems targeting the human *PCSK9* locus

[0532] Experiments were performed to demonstrate that delivery of LNPs encapsulating LTRP6 mRNA and a *PCSK9*-targeting gRNA induces repression at the endogenous human *PCSK9* locus in primary human hepatocytes (PHH) and thus, induce a reduction of secreted PCSK9 and PCSK9 mRNA. Four LTRP molecules (Table 69) and guide scaffold 316v1 (SEQ ID NO: 2968) were selected for assessment in this Example.

Materials and Methods:

[0533] PCSK9 knockdown in primary human hepatocytes provided an experimental basis for lipofection of four top editing targeting sequences (Table 69) into PHH cells for assessment of PCSK9 secretion and mRNA reduction. Dose response curves were generated

from two lots of PHH cells with three rounds of triplicate wells for each condition. The highest tested dose, 2400 ng/well total encapsulated RNA, was used to measure mRNA expression changes between treated and naïve cells via RNA-seq.

Synthesis of gRNAs:

[0534] gRNAs targeting the human *PCSK9* locus were chemically synthesized as the chemically-modified v1 versions for delivery. The *PCSK9*-targeting sequences are listed in Table 69.

[0535] LTRP mRNA was generated by IVT. Briefly, constructs encoding a 5'UTR region, a codon-optimized LTRP, and a 3'UTR region were cloned into a plasmid containing a T7 promoter and poly(A) tail. The resulting plasmid was linearized prior to use for IVT reactions, which were carried out with CleanCap® AG and N1-methyl-pseudouridine.

Formulation of lipid nanoparticles (LNPs):

[0536] LTRP mRNA and targeting gRNAs were encapsulated into LNPs made from ionizable lipid mix containing ALC0315 ionizable lipid: 18:0 PC (DSPC): cholesterol: DMG-PEG2000 using an in-house custom-made T-mixer at N/P 6. Briefly, to formulate LNPs, as a split formulation (containing either mRNA or only sgRNA) were diluted at a fixed ratio for coformulation or separately for split formulation in a 25 mM sodium acetate buffer, pH 4.0. ALC315 ionizable lipid mix at a molar ratio of 50:10:38.5:1.5% of above-mentioned lipids using anhydrous ethanol. The RNA and lipid phases were run through a custom-made T-mixer at a flow rate ratio of 3:1 and a flow rate of 20 mL/min. After formulation, the LNPs were transferred and dialyzed using a 10 kDa membrane dialysis cassette (Thermo Scientific™) and buffer exchanged into 1x PBS to decrease the ethanol concentration and increase the pH to 7.4, leading to formation of a mature and stabilized particles. Following dialysis, the RNA-LNPs were buffer exchanged into 300 mM sucrose in PBS at pH 7.4 storage buffer and concentrated to appropriate concentration using 100 kDa Amicon®-Ultra Centrifugal Filters (Millipore) and sterile-filtered. Formulated LNPs were then subjected to one free-thaw cycle at -80°C and analyzed on a Stunner (Unchained Labs) to determine their average particle size (d. nm.) and polydispersity index (PDI). Encapsulation efficiency and RNA concentration was determined by RiboGreen™ assay using Invitrogen's Quant-iT™ Ribogreen™ RNA assay kit. LNPs were used in various experiments as described herein to deliver mRNA and gRNA to target cells and tissue by mixing a 1:1 mass ratio of mRNA containing LNP and gRNA containing LNP.

Delivery of LNPs encapsulating CasX mRNA and targeting gRNA into primary human hepatocytes:

[0537] Two lots (lot #271 and lot #31) of primary human hepatocytes (Lonza Biologics) were plated at density of 65K cells/well. Twenty-four hours later, plates were stamped with the indicated total encapsulated RNA dose of single formulations ALC-0315 LNP (1:1 ratio, pre-incubated with human serum overnight) at the indicated dose in triplicate wells for each condition. Doses tested for each mRNA:gRNA pair (Table 70) were 2400, 800, 270, 90, 30, 10, 0.1, and 0.01 ng/well total encapsulated RNA.

[0538] Cells were maintained in Geltrex-sandwich culture, and media supernatants collected on day 5 post-treatment were assayed for secreted PCSK9 by HTRF ELISA with CISBio Human PCSK9 HTRF ELISA Kit (Revvity). The cell culture experiment was repeated for 3 rounds.

[0539] For rounds 1 and 2, cells were immediately lysed in 100 uL DNA RNA shield™ and stored at -80C after the day 5 media collection.

[0540] Naive, CasX 515 with targeting sequence 6.1, and LTRP5 with targeting sequence 27.94 as control treatments were included on each plate to assess robustness of measurements across the experiment.

qPCR assay for *PCSK9* mRNA knockdown:

[0541] Extracted mRNA was processed and analyzed by RT-qPCR. Measurement of *PCSK9* mRNA expression is performed using a single-step RT-qPCR approach. 50 uL lysate was used to extract total RNA using Zymo 96 RNA extraction kits. Total RNA was then reverse transcribed into cDNA and PCSK9 transcript quantified in reference to a GAPDH endogenous control by quantitative PCR using TaqMan probes. Percent change in mRNA was calculated by comparing to samples treated with a non-targeting sequence.

ELISA:

[0542] Secreted PCSK9 levels in the media supernatant were analyzed using the BioLegend® ELISA MAX™ kit following the manufacturer's instructions.

RNA seq:

[0543] mRNA levels were analyzed using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB #E7760) following the manufacturer's instructions. RNA from the highest dose samples treated with the indicated mRNA-gRNA pairs (N = 2-6 wells/condition) and naive wells across plates (N=4) wells was analyzed using paired end short-read RNA sequencing. Gene expression changes versus naïve cells were computed from read counts using DEseq2. Cutoffs used were $|\log_2(\text{Fold Change})| > 1$ and Benjamini and Hochberg-adjusted p value (padj) < 0.05.

Results:

[0544] The ability of pairs of mRNA:gRNA comprising scaffold 316 and an *PCSK9*-targeting sequence to reduce secretion of PCSK9 and PCSK9 mRNA reduction was assessed. Tables 71-74 below, provide the half-maximal inhibitory concentration (IC₅₀), 90% maximum inhibition (IC₉₀), and Maximum response (Emax) results in primary human hepatocytes with each individual mRNA:sgRNA pair, rounded to the nearest hundredth. Tables 71 and 72 provide results of secreted PCSK9 reduction from the serum ELISA dose response testing in units of ng of encapsulated RNA/well and pM encapsulated RNA respectively. Tables 73 and 74 provide results of PCSK9 mRNA reduction from qPCR dose response testing in units of ng of encapsulated RNA/well and pM encapsulated RNA respectively.

[0545] LTRP5 and LTRP6-Rat-Cobra were among the most potent molecules in reducing secreted protein, with secreted PCSK9 protein reduction similar to positive control editors CasX 515 and SpyCas9, 6.154 the most potent LTRP spacer (Tables 71 and 72).

[0546] LTRP5 and LTRP6-Rat-Cobra were among the most potent molecules in reducing PCSK9 mRNA levels, with reduction similar to positive control editors CasX 515 and SpyCas9, 6.154 the most potent LTRP spacer (Tables 73 and 74).

[0547] LTRP constructs LTRP6-Rat-Cobra, LTRP6-Rat-Cobra-linker set 3, and LTRP6-2xCobra, paired with gRNA 316.6.154 (V1) showed the highest specificity, as measured by mRNA expression change (Table 75). LTRP6-Rat-Cobra-linker set 3 and LTRP6-2xCobra had no significant off-target effects as measured with cutoff of $p_{adj} < 0.05$.

[0548] The results from these experiments demonstrate that delivery of LNPs encapsulating a LTRP mRNA and *PCSK9*-targeting gRNA induced repression at the endogenous human *PCSK9* locus in primary human hepatocytes that resulted in substantial reduction in secreted PCSK9 and mRNA levels.

Table 69: Sequences of mRNA and gRNA

mRNA	SEQ ID NO.	gRNA	Target	SEQ ID NO.
LTRP5	22829	316.6.1 (V1)	PCSK9	22789
LTRP6-2xCobra	22830	316.6.154 (V1)	PCSK9	22790
LTRP6-Rat-Cobra	22831	316.6.157 (V1)	PCSK9	22788
LTRP6-Rat-Cobra-linker set 3	22832	316.6.141 (V1)	PCSK9	22835
CasX 515	22833	316.27.94 (V1)	NT	3330
SpyCas9	22834	Spy.50.1 (vInt)	PCSK9	22791

Table 70: mRNA: gRNA pairs tested

Construct	gRNA
CasX 515	316.6.1 (V1)
LTRP5	316.6.141 (V1)
LTRP5	316.6.154 (V1)
LTRP5	316.6.157 (V1)
LTRP5	316.6.154 (V1), 316.6.157 (V1)
LTRP5	316.27.94 (V1)
LTRP6-2xCobra	316.6.141 (V1)
LTRP6-2xCobra	316.6.154 (V1)
LTRP6-2xCobra	316.6.157 (V1)
LTRP6-Rat-Cobra	316.6.141 (V1)
LTRP6-Rat-Cobra	316.6.154 (V1)
LTRP6-Rat-Cobra	316.6.157 (V1)
LTRP6-Rat-Cobra-linker set 3	316.6.141 (V1)
LTRP6-Rat-Cobra-linker set 3	316.6.154 (V1)
LTRP6-Rat-Cobra-linker set 3	316.6.157 (V1)
SpyCas9	Spy.50.1 (vInt)

Table 71: Levels of PCSK9 secreted protein repression mediated by LTRP constructs with various repressor domains by total cargo mass

		PHH Lot 271					
		IC50 (ng/well)		IC90* (ng/well)		Emax* (Δ%)	
mRNA	gRNA	Mean	SEM	Mean	SEM	Mean	SEM
CasX 515	316.6.1 (V1)	27.7	5	190.1	84.2	-94.6	4
LTRP5	316.6.154 (V1)	40	4	119.9	27.1	-98.8	3.1
SpyCas9	Spy.50.1 (vInt)	39.5	3.8	226.6	50.9	-99.1	2.8
LTRP5	316.6.154 (V1), 316.6.157 (V1)	42.6	4.6	114.5	26.5	-97.8	3.6
LTRP6-Rat-Cobra	316.6.154 (V1)	57.4	5.4	217.5	43.6	-98.9	2.7
LTRP5	316.6.141 (V1)	79.9	6.3	289.7	52.8	-97.2	2.5
LTRP6-2xCobra	316.6.154 (V1)	81.8	9.2	286.6	74.6	-96.3	3.8
LTRP5	316.6.157 (V1)	96	10	491.6	122.3	- 100.8	4
LTRP6-2xCobra	316.6.141 (V1)	115.8	21.4	494.7	215.8	-99.5	7
LTRP6-Rat-Cobra	316.6.141 (V1)	131.5	14.8	552.5	146.2	- 100.8	4.4
LTRP6-Rat-Cobra-linker set 3	316.6.154 (V1)	133.3	22.8	624	252.8	-98.3	5.9

LTRP6-Rat-Cobra	316.6.157 (V1)	177.6	38	1216. 6	635	-99	8.6
LTRP6-2xCobra	316.6.157 (V1)	173.3	37.7	1054. 8	558.8	-91.5	8.5
LTRP6-Rat-Cobra-linker set 3	316.6.141 (V1)	247	56.6	933.6	530.2	-92.7	9.4
LTRP6-Rat-Cobra-linker set 3	316.6.157 (V1)	253.4	66.9	1195. 8	780.8	-69.2	8.4
mRNA	gRNA	PHH Lot 31					
CasX 515	316.6.1 (V1)	23.2	5.4	180	107.8	-93.2	5.2
LTRP5	316.6.154 (V1)	13.7	1.9	73.4	25.5	-97.4	3.4
SpyCas9	Spy.50.1 (vInt)	24.8	5.1	244.2	134.2	- 100.1	5.7
LTRP5	316.6.154 (V1), 316.6.157 (V1)	29.1	2.3	80.1	15.6	-97.2	2.7
LTRP6-Rat-Cobra	316.6.154 (V1)	24.4	4.1	153.9	63.4	- 100.7	4.6
LTRP5	316.6.141 (V1)	49.6	7.2	364.6	126.9	- 100.6	4.1
LTRP6-2xCobra	316.6.154 (V1)	72	16.6	327.5	171.8	- 100.5	8
LTRP5	316.6.157 (V1)	58	8.1	222.9	66.9	-96.2	4.9
LTRP6-2xCobra	316.6.141 (V1)	98.6	12.7	310	98	-98.8	5.8
LTRP6-Rat-Cobra	316.6.141 (V1)	89.4	9.3	284.2	71.6	-98.5	4.2
LTRP6-Rat-Cobra-linker set 3	316.6.154 (V1)	88.8	18.6	406.6	199.3	-96.2	7.1
LTRP6-Rat-Cobra	316.6.157 (V1)	101.3	18.1	634	271.4	-98.5	7
LTRP6-2xCobra	316.6.157 (V1)	181.6	36.2	718.8	335.1	-92.5	9.4
LTRP6-Rat-Cobra-linker set 3	316.6.141 (V1)	158.9	29.9	527	218	-89.1	7.8
LTRP6-Rat-Cobra-linker set 3	316.6.157 (V1)	261.2	77.3	997.8	738.1	-72.3	12.7

Table 72: Levels of PCSK9 secreted protein repression mediated by LTRP constructs with various repressor domains by mRNA molarity

		PHH Lot 271					
		IC50 (pM)		IC90* (pM)		E_{max} (Δ%)	
mRNA	gRNA	Mean	SEM	Mean	SEM	Mean	SEM
LTRP5	316.6.154 (V1)	110.4	11.2	330.7	74.9	-98.8	3.1
LTRP5	316.6.154 (V1), 316.6.157 (V1)	117.6	12.6	315.8	73	-97.8	3.6
SpyCas9	Spy.50.1 (vInt)	128.1	12.5	734.7	165.1	-99.1	2.8

LTRP6-Rat-Cobra	316.6.154 (V1)	153.8	14.4	582.4	116.8	-98.9	2.7
CasX 515	316.6.1 (V1)	138.1	25.1	947	419.5	-94.6	4
LTRP5	316.6.141 (V1)	220.4	17.5	798.8	145.5	-97.2	2.5
LTRP6-2xCobra	316.6.154 (V1)	219	24.6	767.7	199.8	-96.3	3.8
LTRP5	316.6.157 (V1)	264.6	27.6	1355.5	337.3	-100.8	4
LTRP6-2xCobra	316.6.141 (V1)	310	57.3	1325.1	578	-99.5	7
LTRP6-Rat-Cobra	316.6.141 (V1)	352.1	39.7	1479.3	391.5	-100.8	4.4
LTRP6-Rat-Cobra-linker set 3	316.6.154 (V1)	365.9	62.5	1713.1	694	-98.3	5.9
LTRP6-Rat-Cobra	316.6.157 (V1)	475.6	101.7	3257.6	1700.3	-99	8.6
LTRP6-2xCobra	316.6.157 (V1)	464.2	101	2825	1496.7	-91.5	8.5
LTRP6-Rat-Cobra-linker set 3	316.6.141 (V1)	678.1	155.5	2563	1455.4	-92.7	9.4
LTRP6-Rat-Cobra-linker set 3	316.6.157 (V1)	695.8	183.7	3282.8	2143.6	-69.2	8.4
mRNA	gRNA	PHH Lot 31					
LTRP5	316.6.154 (V1)	37.9	5.2	202.4	70.3	-97.4	3.4
LTRP5	316.6.154 (V1), 316.6.157 (V1)	80.2	6.4	221	42.9	-97.2	2.7
SpyCas9	Spy.50.1 (vInt)	80.5	16.6	791.8	435.2	-100.1	5.7
LTRP6-Rat-Cobra	316.6.154 (V1)	65.2	11.1	412.2	169.7	-100.7	4.6
CasX 515	316.6.1 (V1)	115.5	27.1	896.6	537	-93.2	5.2
LTRP5	316.6.141 (V1)	136.8	19.9	1005.3	350	-100.6	4.1
LTRP6-2xCobra	316.6.154 (V1)	192.7	44.4	877.1	460.3	-100.5	8
LTRP5	316.6.157 (V1)	160.1	22.3	614.7	184.5	-96.2	4.9
LTRP6-2xCobra	316.6.141 (V1)	264	34	830.3	262.5	-98.8	5.8
LTRP6-Rat-Cobra	316.6.141 (V1)	239.3	24.9	761	191.6	-98.5	4.2
LTRP6-Rat-Cobra-linker set 3	316.6.154 (V1)	243.9	51.1	1116.2	547.1	-96.2	7.1
LTRP6-Rat-Cobra	316.6.157 (V1)	271.3	48.5	1697.7	726.8	-98.5	7
LTRP6-2xCobra	316.6.157 (V1)	486.2	97	1925.3	897.6	-92.5	9.4

LTRP6-Rat-Cobra-linker set 3	316.6.141 (V1)	436.2	82.1	1446.8	598.5	-89.1	7.8
LTRP6-Rat-Cobra-linker set 3	316.6.157 (V1)	717.1	212.2	2739.2	2026.2	-72.3	12.7

Table 73: Levels of PCSK9 mRNA repression mediated by LTRP constructs with various repressor domains by total cargo mass

		PHH Lot 271					
		IC50 (ng/well)		IC90* (ng/well)		Emax* (Δ%)	
mRNA	gRNA	Mean	SEM	Mean	SEM	Mean	SEM
CasX 515	316.6.1 (V1)	29.6	5.7	103.3	45.2	-60.1	2.7
SpyCas9	Spy.50.1 (vInt)	50.3	8.2	161.2	53.9	-59.8	2.6
LTRP5	316.6.154 (V1)	38.2	7.8	211.2	101	-96.7	5.5
LTRP5	316.6.154 (V1), 316.6.157 (V1)	45.7	4.5	143.1	30.4	-95.1	2.9
LTRP6-Rat-Cobra	316.6.154 (V1)	92.9	13.7	357.9	124.5	-94.9	4.3
LTRP5	316.6.157 (V1)	101.3	14.5	504.4	173.5	-94.7	5
LTRP6-2xCobra	316.6.154 (V1)	123.7	19.8	528.3	199.1	-91.9	5
LTRP5	316.6.141 (V1)	112.8	13.8	354.6	104.1	-88.9	3.7
LTRP6-Rat-Cobra	316.6.141 (V1)	153.1	22.4	755	264.3	-92.9	4.5
LTRP6-Rat-Cobra-linker set 3	316.6.154 (V1)	168.6	36	793.8	405.5	-83.3	6.4
LTRP6-2xCobra	316.6.141 (V1)	159.7	47.4	695.8	486.5	-92.7	8.7
LTRP6-Rat-Cobra	316.6.157 (V1)	182.7	35.7	957	453.1	-80.3	5.8
LTRP6-2xCobra	316.6.157 (V1)	189.8	42.8	817.3	439.1	-70.7	6.2
LTRP6-Rat-Cobra-linker set 3	316.6.157 (V1)	135.1	37.9	470.1	300.4	-41.1	4.3
LTRP6-Rat-Cobra-linker set 3	316.6.141 (V1)	233.2	67.6	1054.7	750.9	-73.1	8.1
		PHH Lot 31					
mRNA	gRNA	Mean	SEM	Mean	SEM	Mean	SEM
CasX 515	316.6.1 (V1)	42.7	10.6	149.3	80.3	-54.4	3.3
SpyCas9	Spy.50.1 (vInt)	37.5	28.2	293.3	539.3	-44.4	9.4
LTRP5	316.6.154 (V1)	53.8	10.9	202.5	87.7	-97	4.6

LTRP5	316.6.154 (V1), 316.6.157 (V1)	59.7	9.2	194.2	61.6	-96.5	3.9
LTRP6-Rat-Cobra	316.6.154 (V1)	39.9	5.7	387	139.8	-96.1	3.6
LTRP5	316.6.157 (V1)	115.9	10.6	438.4	93.9	-93.4	2.5
LTRP6-2xCobra	316.6.154 (V1)	106.8	16.8	327.2	125.7	-89.6	5.2
LTRP5	316.6.141 (V1)	122	30.2	701.4	417.1	-92.1	7.3
LTRP6-Rat-Cobra	316.6.141 (V1)	115	13.2	407.9	110.7	-88.6	3.6
LTRP6-Rat-Cobra-linker set 3	316.6.154 (V1)	129.6	22.2	461.3	182.3	-83.1	5.2
LTRP6-2xCobra	316.6.141 (V1)	148.6	18.7	377.4	93.9	-89.7	4.1
LTRP6-Rat-Cobra	316.6.157 (V1)	172.1	42.4	1053.3	631.5	-87.3	8
LTRP6-2xCobra	316.6.157 (V1)	196.7	35.5	659.9	271.6	-77.4	6
LTRP6-Rat-Cobra-linker set 3	316.6.157 (V1)	254.2	69.8	728.3	523.5	-55.4	7.6
LTRP6-Rat-Cobra-linker set 3	316.6.141 (V1)	228	37	536.7	210	-73.7	5.4

Table 74: Levels of PCSK9 mRNA repression mediated by LTRP constructs with various repressor domains by mRNA molarity

		PHH Lot 271					
		IC50 (pM)		IC90* (pM)		Emax* (Δ%)	
mRNA	gRNA	Mean	SEM	Mean	SEM	Mean	SEM
LTRP5	316.6.154 (V1)	105.3	21.4	582.3	278.4	-96.7	5.5
SpyCas9	Spy.50.1 (vInt)	163.2	26.6	522.5	174.8	-59.8	2.6
LTRP5	316.6.154 (V1), 316.6.157 (V1)	126	12.5	394.7	83.8	-95.1	2.9
LTRP6-Rat-Cobra	316.6.154 (V1)	248.7	36.7	958.4	333.5	-94.9	4.3
CasX 515	316.6.1 (V1)	147.5	28.6	514.9	225	-60.1	2.7
LTRP5	316.6.154 (V1)	279.3	40	1390.8	478.5	-94.7	5
LTRP6-2xCobra	316.6.154 (V1)	331.4	53	1415.1	533.1	-91.9	5
LTRP5	316.6.154 (V1)	311	38.1	977.8	286.9	-88.9	3.7
LTRP6-Rat-Cobra	316.6.154 (V1)	409.9	60	2021.7	707.6	-92.9	4.5
LTRP6-Rat-Cobra-linker set 3	316.6.154 (V1)	462.8	99	2179.2	1113.3	-83.3	6.4

LTRP6-2xCobra	316.6.154 (V1)	427.7	126.9	1863.5	1303	-92.7	8.7
LTRP6-Rat-Cobra	316.6.154 (V1)	489.3	95.5	2562.4	1213.2	-80.3	5.8
LTRP6-2xCobra	316.6.154 (V1)	508.4	114.6	2188.9	1175.9	-70.7	6.2
LTRP6-Rat-Cobra-linker set 3	316.6.154 (V1)	370.8	103.9	1290.5	824.5	-41.1	4.3
LTRP6-Rat-Cobra-linker set 3	316.6.154 (V1)	640.1	185.5	2895.5	2061.3	-73.1	8.1
PHH Lot 31							
mRNA	gRNA	Mean	SEM	Mean	SEM	Mean	SEM
LTRP5	316.6.154 (V1)	148.3	30.1	558.3	241.9	-97	4.6
SpyCas9	Spy.50.1 (vInt)	121.5	91.3	950.9	1748.5	-44.4	9.4
LTRP5	316.6.154 (V1), 316.6.157 (V1)	164.6	25.3	535.5	169.9	-96.5	3.9
LTRP6-Rat-Cobra	316.6.154 (V1)	106.8	15.2	1036.2	374.3	-96.1	3.6
CasX 515	316.6.1 (V1)	212.8	52.8	743.7	399.9	-54.4	3.3
LTRP5	316.6.154 (V1)	319.5	29.1	1208.7	259	-93.4	2.5
LTRP6-2xCobra	316.6.154 (V1)	286.1	44.9	876.3	336.5	-89.6	5.2
LTRP5	316.6.154 (V1)	336.3	83.3	1934.1	1150.2	-92.1	7.3
LTRP6-Rat-Cobra	316.6.154 (V1)	307.8	35.5	1092.1	296.5	-88.6	3.6
LTRP6-Rat-Cobra-linker set 3	316.6.154 (V1)	355.7	61	1266.3	500.4	-83.1	5.2
LTRP6-2xCobra	316.6.154 (V1)	398	50	1010.9	251.4	-89.7	4.1
LTRP6-Rat-Cobra	316.6.154 (V1)	460.9	113.4	2820.4	1690.8	-87.3	8
LTRP6-2xCobra	316.6.154 (V1)	526.9	95.2	1767.5	727.6	-77.4	6
LTRP6-Rat-Cobra-linker set 3	316.6.154 (V1)	697.9	191.7	1999.3	1437.2	-55.4	7.6
LTRP6-Rat-Cobra-linker set 3	316.6.154 (V1)	625.9	101.5	1473.4	576.4	-73.7	5.4

Table 75: Gene expression change by RNA-seq

mRNA	gRNA	Dose (ng*)	Gene	log2(Fold Change)	p-value
CasX 515	316.6.1 (V1)	2400	TUBB3	-1.05	5.90E-08
			MARCO	-2.67	8.61E-07
			ESAM	-1.04	1.07E-06
			SUSD2	-1.14	3.07E-06

			C1QB	-1.67	2.87E-05
			KIF23	-1.19	5.82E-05
SpyCas9	Spy.50.1 (vInt)	2400	PLD4	-1.36	4.14E-07
			SLC7A7	-1.03	8.09E-07
			TSPAN15	-1.28	1.54E-06
			ANXA8	-1.28	3.39E-06
			RAC2	-1.08	4.94E-06
			CYP7A1	1.45	1.09E-05
LTRP5	316.6.154 (V1), 316.6.157 (V1)	2400	PLD4	-1.36	4.14E-07
			SLC7A7	-1.03	8.09E-07
			TSPAN15	-1.28	1.54E-06
			ANXA8	-1.28	3.39E-06
			RAC2	-1.08	4.94E-06
			CYP7A1	1.45	1.09E-05
LTRP5	316.6.154 (V1)	2400	<u>PCSK9</u>	-3.87	1.93E-64
			C1QB	-2.92	1.78E-08
			GPNMB	-1.96	3.06E-08
			CD163	-1.49	3.89E-07
			ITGAX	-2.02	5.74E-07
			LAPTM5	-1.60	1.33E-06
			RNASE1	-2.12	2.05E-06
			MSR1	-2.30	2.10E-06
			RGCC	-2.20	2.88E-06
			PLEK	-2.46	1.88E-05
			MIR6501	1.04	0.00011389
			HCK	-2.84	0.00013665
LTRP5	316.6.157 (V1)	2400	<u>PCSK9</u>	-2.97	9.32E-41
			TM4SF19	-2.59	1.26E-06
			UCP2	-2.07	3.00E-05
			MIR6501	1.09	8.01E-05
LTRP5	316.27.94 (V1)	2400	NA	NA	NA
LTRP6-Rat-Cobra	316.6.154 (V1)	2400	<u>PCSK9</u>	-3.34	5.86E-47
			DNMT3A	1.83	1.55E-08
			G6PD	1.01	4.75E-05
			SDS	1.47	0.00012577
LTRP6-Rat-Cobra-linker set 3	316.6.154 (V1)	2400	<u>PCSK9</u>	-1.45	3.00E-12
LTRP6-2xCobra	316.6.154 (V1)	2400	<u>PCSK9</u>	-2.34	2.85E-27

*Dose is reported as ng of encapsulated RNA/well

299945242

CLAIMS

What is claimed is

1. A system for transcriptional repression of a proprotein convertase subtilisin/kexin Type 9 (*PCSK9*) gene, the system comprising:
 - (a) a guide ribonucleic acid (gRNA) comprising a targeting sequence complementary to a *PCSK9* gene target nucleic acid sequence, and
 - (b) an mRNA encoding a long-term repressor fusion protein (LTRP), wherein the LTRP comprises:
 - a. a catalytically-dead CasX (dCasX);
 - b. a DNA methyltransferase (DNMT) 3A catalytic domain (DNMT3A);
 - c. a DNMT3 like interaction domain (DNMT3L); and
 - d. a first repressor domain (RD1),wherein the LTRP is capable of forming a ribonucleoprotein (RNP) with the gRNA.
2. The system of claim 1, wherein the LTRP comprises, from N- to C- terminus:
 - a. the DNMT3A;
 - b. the DNMT3L;
 - c. the dCasX; and
 - d. the RD1.
3. The system of claim 1, wherein the LTRP comprises, from N- to C- terminus:
 - a. the DNMT3A;
 - b. the DNMT3L;
 - c. the RD1; and
 - d. the dCasX.
4. The system of claim 3, wherein the LTRP further comprises an additional RD1 C-terminal to the dCasX.
5. The system of claim 4, wherein the RD1 sequences are identical.
6. The system of claim 4, wherein the RD1 sequences are different.

7. The system of any one of claims 1-6, wherein the LTRP comprises a DNA DNMT3A ATRX-DNMT3-DNMT3L domain (ADD) linked to the N-terminus of the DNMT3A.
8. The system of any one of claims 1-7, wherein the LTRP comprises one or more linker peptides.
9. The system of claim 8, wherein at least one of the one or more linker peptides comprises a sequence independently selected from the group consisting of SEQ ID NOS: 98-124, and 3278-3289.
10. The system of any one of claims 1-9, wherein the LTRP comprises one or more nuclear localization signals (NLS).
11. The system of claim 10, wherein at least one of the one or more NLS comprises a sequence selected from the group consisting of SEQ ID NOS: 30-97.
12. The system of claim 10 or claim 11, wherein at least one of the one or more NLS is a simian virus 40 (SV40) NLS.
13. The system of claim 12, wherein the SV40 NLS comprises the sequence of SEQ ID NO: 30.
14. The system of claim 10 or claim 11, wherein at least one of the one or more NLS is a c-MYC NLS.
15. The system of claim 14, wherein the c-MYC NLS comprises the sequence of SEQ ID NO: 32.
16. The system of any one of claims 1-2, or 8-15, wherein the LTRP comprises, from N- to C-terminus,
 - a. a first NLS;
 - b. the DNMT3A;
 - c. a first linker peptide;
 - d. the DNMT3L;
 - e. a second linker peptide;
 - f. a third linker peptide;
 - g. the dCasX;

- h. a fourth linker peptide;
- i. the RD1, and
- j. a second NLS.

17. The system of claim 16, wherein the LTRP comprises a sequence selected from the group consisting of SEQ ID NO: 22836, and SEQ ID NOS: 22838-22846.

18. The system of any one of claims 7-15, wherein the LTRP comprises, from N- to C-terminus,

- a. a first NLS;
- b. the ADD;
- c. the DNMT3A;
- d. a first linker peptide;
- e. the DNMT3L;
- f. a second linker peptide;
- g. a third linker peptide;
- h. the dCasX;
- i. a fourth linker peptide;
- j. the RD1, and
- k. a second NLS.

19. The system of claim 18, wherein the LTRP comprises a sequence selected from the group consisting of SEQ ID NO: 22837 and SEQ ID NOS: 22847-22855.

20. The system of any one of claims 1, 3, or 8-15, wherein the LTRP comprises, from N- to C-terminus,

- a. a first NLS;
- b. the DNMT3A;
- c. a first linker peptide;
- d. the DNMT3L;
- e. a third linker peptide;
- f. the RD1;
- g. a second linker peptide;
- h. the dCasX;
- i. a fourth linker peptide, and

- j. a second NLS.

21. The system of any one of claims 7-15, wherein the LTRP comprises, from N- to C-terminus,

- a. a first NLS;
- b. the ADD;
- c. the DNMT3A;
- d. a first linker peptide;
- e. the DNMT3L;
- f. a second linker peptide;
- g. the RD1
- h. a third linker peptide;
- i. the dCasX;
- j. a fourth linker peptide, and
- k. a second NLS.

22. The system of any one of claims 7-15, wherein the LTRP comprises, from N- to C-terminus,

- a. a first NLS;
- b. the ADD;
- c. the DNMT3A;
- d. a first linker peptide;
- e. the DNMT3L;
- f. a second linker peptide;
- g. the RD1
- h. a third linker peptide;
- i. the dCasX;
- j. a fourth linker peptide;
- k. an additional RD1;
- l. a fifth linker, and
- m. a second NLS.

23. The system of claim 22, wherein the RD1 and the additional RD1 sequences are identical.

- 24.** The system of claim 22, wherein the RD1 and the additional RD1 sequences are different.
- 25.** The system of any one of claims 1-24, wherein the dCasX comprises a sequence selected from the group consisting of SEQ ID NOS: 4-29, or a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 26.** The system of claim 25, wherein the dCasX comprises a sequence of SEQ ID NO: 4, or a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 27.** The system of claim 26, wherein the dCasX consists of a sequence of SEQ ID NO: 4.
- 28.** The system of any one of claims 1-27, wherein the RD1 comprises a sequence selected from the group consisting of SEQ ID NOS: 128-1726, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 29.** The system of any one of claims 1-27, wherein the RD1 comprises a sequence selected from the group consisting of SEQ ID NOS: 130-138, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 30.** The system of any one of claims 1-27, wherein the RD1 comprises a sequence of SEQ ID NO: 130, or sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 31.** The system of any one of claims 1-27, wherein the RD1 comprises a sequence of SEQ ID NO: 131, or sequence having at least about 70%, at least about 80%, at least about 85%, at

least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

32. The system of any one of claims 1-27, wherein the RD1 comprises a sequence of SEQ ID NO: 132, or sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

33. The system of any one of claims 1-27, wherein the RD1 comprises a sequence of SEQ ID NO: 133, or sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

34. The system of any one of claims 1-27, wherein the RD1 comprises a sequence of SEQ ID NO: 134, or sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

35. The system of any one of claims 1-27, wherein the RD1 comprises the sequence of SEQ ID NO: 135, or sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

36. The system of any one of claims 1-27, wherein the RD1 comprises a sequence of SEQ ID NO: 136, or sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

37. The system of any one of claims 1-27, wherein the RD1 comprises a sequence of SEQ ID NO: 137, or sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

38. The system of any one of claims 1-27, wherein the RD1 comprises a sequence of SEQ ID NO: 138, or sequence having at least about 70%, at least about 80%, at least about 85%, at

least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

39. The system of any one of claims 7-15, 18, 19 or 19-38, wherein the ADD comprises a sequence of SEQ ID NO: 125, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

40. The system of any one of claims 1-39, wherein the DNMT3A comprises a sequence of SEQ ID NO: 126, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

41. The system of any one of claims 1-40, wherein the DNMT3L comprises a sequence of SEQ ID NO: 127, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

42. The system of any one of claims 1-41, wherein the sequence encoding the dCasX comprises a sequence selected from the group consisting of SEQ ID NOS: 3122, and 22727, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

43. The system of any one of claims 1-41, wherein the sequence encoding the RD1 comprises a sequence selected from the group consisting of SEQ ID NOS: 3334-6527, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

44. The system of any one of claims 1-41, wherein the sequence encoding the RD1 comprises a sequence selected from the group consisting of SEQ ID NOS: 3334-3342, and 4931-4939, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

- 45.** The system of any one of claims 1-41, wherein the sequence encoding the RD1 comprises a sequence selected from the group consisting of SEQ ID NO: 3335 and SEQ ID NO: 4932, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 46.** The system of any one of claims 1-41, wherein the sequence encoding the RD1 comprises a sequence selected from the group consisting of SEQ ID NO: 3339 and SEQ ID NO: 4936, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 47.** The system of any one of claims 1-41, wherein the sequence encoding the RD1 comprises a sequence selected from the group consisting of SEQ ID NO: 3334 and SEQ ID NO: 4931, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 48.** The system of any one of claims 1-41, wherein the sequence encoding the RD1 comprises a sequence selected from the group consisting of SEQ ID NO: 3336 and SEQ ID NO: 4933, or sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 49.** The system of any one of claims 1-41, wherein the sequence encoding the RD1 comprises a sequence selected from the group consisting of SEQ ID NO: 3337 and SEQ ID NO: 4934, or sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 50.** The system of any one of claims 1-41, wherein the sequence encoding the RD1 comprises a sequence selected from the group consisting of SEQ ID NO: 3338 and SEQ ID NO: 4935, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

- 51.** The system of any one of claims 1-41, wherein the sequence encoding the RD1 comprises a sequence selected from the group consisting of SEQ ID NO: 3340 and SEQ ID NO: 4937, or sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 52.** The system of any one of claims 1-41, wherein the sequence encoding the RD1 comprises a sequence selected from the group consisting of SEQ ID NO: 3341 and SEQ ID NO: 4938, or sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 53.** The system of any one of claims 1-41, wherein the sequence encoding the RD1 comprises a sequence selected from the group consisting of SEQ ID NO: 3342 and SEQ ID NO: 4939, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 54.** The system of any one of claims 1-53, wherein the sequence encoding the DNMT3A comprises a sequence selected from the group consisting of SEQ ID NO: 3128 and SEQ ID NO: 3331, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 55.** The system of any one of claims 1-54, wherein the sequence encoding the DNMT3L comprises a sequence selected from the group consisting of SEQ ID NO: 3119 and SEQ ID NO: 3332, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 56.** The system of any one of claims 7-15, 18, 19, or 21-55, wherein the sequence encoding the ADD comprises a sequence selected from the group consisting of SEQ ID NOS: 3127, 3296, and 22726, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

57. The system of any one of claims 1, 2, 8-17 or 25-55, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NOS: 6529-8134, and 14628-16233, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

58. The system of any one of claims 1, 2, 8-17 or 25-55, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NOS: 6529-6537, and 14628-14636, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

59. The system of any one of claims 1, 2, 8-17 or 25-55, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NO: 6529 and SEQ ID NO: 14628, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

60. The system of any one of claims 1, 2, 8-17 or 25-55, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NO: 6530 and SEQ ID NO: 14629, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

61. The system of any one of claims 1, 2, 8-17 or 25-55, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NO: 6531 and SEQ ID NO: 14630, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

62. The system of any one of claims 1, 2, 8-17 or 25-55, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NO: 6532 and SEQ ID NO: 14631, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

63. The system of any one of claims 1, 2, 8-17 or 25-55, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NO: 6533 and SEQ ID NO: 14632, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

64. The system of any one of claims 1, 2, 8-17 or 25-55, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NO: 6535 and SEQ ID NO: 14632, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

65. The system of any one of claims 1, 2, 8-17 or 25-55, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NO: 6534 and SEQ ID NO: 14633, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

66. The system of any one of claims 1, 2, 8-17 or 25-55, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NO: 6536 and SEQ ID NO: 14635, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

67. The system of any one of claims 1, 2, 8-17 or 25-55, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NO: 6537 and SEQ ID NO: 4636, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

68. The system of any one of claims 1, 2, 7-15, 18, 19 or 25-56, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NOS: 9742-11347, and 17840-19446, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

- 69.** The system of any one of claims 1, 2, 7-15, 18, 19 or 25-56, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NOS: 9742-9750, and 17841-17849, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 70.** The system of any one of claims 1, 2, 7-15, 18, 19 or 25-56, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NO: 9742 and SEQ ID NO: 17841, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 71.** The system of any one of claims 1, 2, 7-15, 18, 19 or 25-56, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NO: 9743 and SEQ ID NO: 17842, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 72.** The system of any one of claims 1, 2, 7-15, 18, 19 or 25-56, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NO: 9744 and SEQ ID NO: 17843, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 73.** The system of any one of claims 1, 2, 7-15, 18, 19 or 25-56, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NO: 9745 and SEQ ID NO: 17844, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 74.** The system of any one of claims 1, 2, 7-15, 18, 19 or 25-56, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NO: 9746 and SEQ ID NO: 17845, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

75. The system of any one of claims 1, 2, 7-15, 18, 19 or 25-56, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NO: 9747 and SEQ ID NO: 17846, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

76. The system of any one of claims 1, 2, 7-15, 18, 19 or 25-56, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NO: 9748 and SEQ ID NO: 17847, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

77. The system of any one of claims 1, 2, 7-15, 18, 19 or 25-56, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NO: 9749 and SEQ ID NO: 17848, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

78. The system of any one of claims 1, 2, 7-15, 18, 19 or 25-56, wherein the mRNA encoding the LTRP comprises a sequence selected from SEQ ID NO: 9750 and SEQ ID NO: 17849, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

79. The system of any one of claims 1, 3, 8-15, 20 or 25-55, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NOS: 8135-9740 and 16234-17839, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

80. The system of any one of claims 1, 3, 8-15, 20 or 25-55, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NOS: 8135-8143 and 16234-16242, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

- 81.** The system of any one of claims 1, 3, 7-15, 21 or 25-56, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NOS: 11348-12953 and 19447-21052, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 82.** The system of any one of claims 1, 3, 7-15, 21 or 25-56, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NOS: 11348-11356 and 19447-19455, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 83.** The system of any one of claims 1, 3-5, 7-15, 22, 23 or 25-56, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NOS: 12954-14553, and 21053-22652, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 84.** The system of any one of claims 1, 3-4, 6-15, 22, or 24-56, wherein the mRNA encoding the LTRP comprises a sequence selected from the group consisting of SEQ ID NOS: 14554-14626, and 22653-22725, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.
- 85.** The system of any one of claims 1-84, wherein the mRNA is codon-optimized.
- 86.** The system of claim 85, wherein the mRNA is codon-optimized for expression in a human cell.
- 87.** The system of any one of claims 1-86, wherein the mRNA is chemically modified, optionally wherein one or more or all of the uridine residues of the mRNA are replaced with 1-methyl-pseudouridine.
- 88.** The system of any one of claims 1-87, wherein the mRNA comprises a 5' untranslated region (UTR), a 3' UTR, a poly(A) sequence, and/or a 5' cap.
- 89.** The system of claim 88, wherein the 5' UTR comprises a sequence of SEQ ID NO:

3300.

90. The system of claim 88 or claim 89, wherein the 3' UTR comprises a sequence of SEQ ID NO: 3310.

91. The system of any one of claims 88-90, wherein the poly(A) comprises a sequence of SEQ ID NO: 3057.

92. The system of any one of claims 88-91, wherein the 5' cap is linked 5' to the mRNA.

93. The system of any one of claims 87-92, wherein the mRNA comprises a sequence selected from the group consisting of SEQ ID NOS: 14628-16233, 16234-17839, 17841-19446, 19447-21052, 21053-22652, and 22653-22725, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

94. The system of any one of claims 87-92, wherein the mRNA comprises a sequence of SEQ ID NOS: 14628-14635, or 14636, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

95. The system of any one of claims 87-92, wherein the mRNA comprises a sequence of SEQ ID NO: 14628, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

96. The system of any one of claims 87-92, wherein the mRNA comprises a sequence of SEQ ID NO: 14629, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

97. The system of any one of claims 87-92, wherein the mRNA comprises a sequence of SEQ ID NO: 14630, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

98. The system of any one of claims 87-92, wherein the mRNA comprises a sequence of SEQ ID NO: 14631 or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

99. The system of any one of claims 87-92, wherein the mRNA comprises a sequence SEQ ID NO: 14632, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

100. The system of any one of claims 87-92, wherein the mRNA comprises a sequence of SEQ ID NO: 14633 or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

101. The system of any one of claims 87-92, wherein the mRNA comprises a sequence of SEQ ID NO: 14634, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

102. The system of any one of claims 87-92, wherein the mRNA comprises a sequence of SEQ ID NO: 14635, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

103. The system of any one of claims 87-92, wherein the mRNA comprises a sequence of SEQ ID NO: 14636, or a sequence having at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

104. The system of any one of claims 1-103, wherein the *PCSK9* gene target nucleic acid sequence is within 1.5 kb of a transcription start site (TSS) in the gene.

105. The system of any one of claims 1-103, wherein the *PCSK9* gene target nucleic acid sequence is within 500 bps upstream to 500 bps downstream of a TSS of the gene.

- 106.** The system of any one of claims 1-103, wherein the *PCSK9* gene target nucleic acid sequence is within 300 bps upstream to 300 bps downstream or is within 100 bps upstream to 100 bps downstream of a TSS of the gene.
- 107.** The system of any one of claims 1-103, wherein the *PCSK9* gene target nucleic acid sequence is within 100 bps upstream to 100 bps downstream of a TSS of the gene.
- 108.** The system of any one of claims 1-107, wherein the *PCSK9* gene target nucleic acid sequence is within 1 kb of an enhancer of the gene.
- 109.** The system any one of claims 1-107, wherein the *PCSK9* gene target nucleic acid sequence is within the 3' untranslated region of the *PCSK9* gene.
- 110.** The system of any one of claims 1-109, wherein the targeting sequence of the gRNA comprises a sequence selected from the group consisting of SEQ ID NOS: 1824-1881, 1883-1897, 1899-1934, 1937-1991, 1993-1999, 2001-2026, 2029-2057, 2059-2106, 2108-2114, 2116-2122, 2124-2139, 2141-2209, 2211-2212, 2214-2233, 2235-2252, 2254-2269, 2271-2281, 2283-2284, 2286-2320, 2322-2334, 2236-2346, 2348-2362, 2364-2375, 2377-2395, 2397-2428, 2430-2438, 2440-2463, 2465-2509, 2511-2544, 2672, 2675, 2694 and 2714.
- 111.** The system of any one of claims 1-109, wherein the targeting sequence of the gRNA comprises a sequence selected from the group consisting of SEQ ID NOS: 1824-1880, 1883, 1884, 1888, 1889, 2672, 2675, 2694, and 2714.
- 112.** The system of any one of claims 1-109, wherein the targeting sequence of the gRNA comprises a sequence selected from the group consisting of SEQ ID NOS: 1834, 1849, 1853, 1855-1858, 1860, 1862, 1863, 1867, 1869, 1870, 1872, 1874, and 1875.
- 113.** The system of any one of claims 1-109, wherein the targeting sequence of the gRNA comprises a sequence selected from the group consisting of SEQ ID NOS: 1855, 1867, and 1869.
- 114.** The system of any one of claims 110-113, wherein the targeting sequence of the gRNA comprises a sequence wherein 1, 2, 3, 4, or 5 nucleotide(s) are removed from the 3' end of the sequence.

115. The system of any one of claims 1-114, wherein the gRNA is a single-molecule gRNA (sgRNA).

116. The system of any one of claims 1-115, wherein the gRNA comprises a scaffold stem loop comprising the sequence of CCAGCGACUAUGUCGUAGUGG (SEQ ID NO: 1822), or a sequence having 1, 2, 3, 4, or 5 mismatches thereto.

117. The system of any one of claims 1-116, wherein the gRNA comprises a scaffold comprising a sequence selected from the group consisting of SEQ ID NOS: 1744-1746, or a sequence having at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% sequence identity thereto.

118. The system of any one of claims 1-116, wherein the gRNA comprises a scaffold comprising a sequence of SEQ ID NOS: 1746, or a sequence having at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% sequence identity thereto.

119. The system of any one of claims 1-118, wherein the gRNA comprises a sequence selected from the group consisting of SEQ ID NOS: 3074-3081, 3145-3147, 3150-3153, and 3158-3176, or a sequence having at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% sequence identity thereto.

120. The system of any one of claims 1-119, wherein the gRNA is chemically modified.

121. The system of claim 120, wherein the chemical modification to the gRNA comprises an addition of a 2'O-methyl group to one or more nucleotides of the gRNA.

122. The system of claim 121, wherein one or more nucleotides located 1, 2, 3, or 4 nucleotides from the 5' terminal end, the 3' terminal, or both terminal ends of the gRNA are modified by an addition of a 2'O-methyl group.

123. The system of any one of claims 120-122, wherein the chemical modification to the gRNA comprises a substitution of a phosphorothioate bond between two or more nucleotides of the gRNA.

- 124.** The system of claim 123, wherein the chemical modification comprises a substitution of phosphorothioate bonds between two or more nucleotides located 1, 2, 3 or 4 nucleotides from the 5' terminal end, the 3' terminal, or both terminal ends of the gRNA.
- 125.** The system of any one of claims 104-124, wherein the gRNA comprises a sequence selected from the group consisting of SEQ ID NOS: 3074-3081, 3147, 3151, 3153, 3159-3164, 3166-3171, 3173-3176 and 22788-22803, or a sequence having at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% sequence identity thereto.
- 126.** The system of any one of claims 104-124, wherein the gRNA comprises a sequence selected from the group consisting of SEQ ID NOS: 22788-22803.
- 127.** The system of any one of claims 104-124, wherein the gRNA comprises a sequence selected from the group consisting of SEQ ID NOS: 22788-22790.
- 128.** A lipid nanoparticle comprising the system of any one of claims 1-127.
- 129.** A lipid nanoparticle (LNP) comprising the mRNA of any one of claims 87-103 and the gRNA of any one of claims 120-127.
- 130.** The LNP of claim 128 or 129, wherein the LNP comprises one or more components selected from the group consisting of ionizable lipids, one or more helper phospholipids, one or more polyethylene glycol (PEG)-modified lipids, and cholesterol or a derivative thereof.
- 131.** The LNP of any one of claims 128-130, wherein the LNP comprises an ionizable lipid, a helper phospholipid, a polyethylene glycol (PEG)-modified lipid, and cholesterol or a derivative thereof.
- 132.** The LNP of any one of claims 128-131, comprising a cationic lipid comprising a pKa of 5 to 8.
- 133.** A pharmaceutical composition comprising:
- the system of any one of claims 1-127; or
 - the LNP of any one of claims 128-132,
- and one or more pharmaceutically suitable excipients.

134. A pharmaceutical composition comprising a plurality of the lipid nanoparticles of any one of claims 128-132 and a pharmaceutically acceptable carrier or diluent.

135. The pharmaceutical composition of claim 134, wherein an average diameter of lipid nanoparticles in the plurality is between about 20 nm and about 200 nm.

136. The pharmaceutical composition of claim 134 or claim 135, wherein the pharmaceutical composition is formulated for a route of administration selected from the group consisting of intravenous, intraarterial, intraportal vein injection, intraperitoneal, intramuscular, intracerebroventricular, intracisternal, intrathecal, intracranial, intralumbar, intraocular, subcutaneous, and oral routes.

137. A method of repressing transcription of a *PCSK9* gene in a population of cells, the method comprising introducing into cells of the population:

- a. the system of any one of claims 1-127;
 - b. the LNP of any one of claims 128-132;
 - c. the pharmaceutical composition of any one of claims 133-136; or
 - d. combinations of two or more of (a)-(c),
- thereby repressing transcription of the *PCSK9* gene in the population of cells.

138. The method of claim 137, wherein transcription of the *PCSK9* gene is repressed in at least about 1%, at least about 2%, at least about 3%, at least about 4%, at least about 5%, at least about 6%, at least about 7%, at least about 8%, at least about 9%, or at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80% or more of the cells of the population.

139. The method of claim 137 or claim 138, wherein transcription of the *PCSK9* gene in the cells of the population is decreased by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, or at least about 90% in comparison to an untreated cell.

140. The method of any one of claims 137-139, wherein the cells of the population are eukaryotic cells.

141. The method of claim 140, wherein the eukaryotic cells are selected from the group consisting of rodent cells, mouse cells, rat cells, and non-human primate cells.

142. The method of claim 140, wherein the eukaryotic cells are human cells.
143. The method of any one of claims 140-142, wherein the eukaryotic cells are selected from the group consisting of cells of the liver, cells of the intestine, cells of the kidney, cells of the central nervous system, smooth muscle cells, macrophages, and cells of arterial walls.
144. The method of any one of claims 137-143, wherein the repression of the *PCSK9* gene of the population of cells occurs *in vitro* or *ex vivo*.
145. The method of any one of claims 137-143, wherein the repression of the *PCSK9* gene of the population of cells occurs *in vivo* in a subject.
146. The method of claim 145, wherein the subject is selected from the group consisting of a rodent, a mouse, a rat, and a non-human primate.
147. The method of claim 145, wherein the subject is a human.
148. The method of any one of claims 137-147, wherein the repression of transcription is stable through one or more cell divisions.
149. The method of any one of claims 137-148, wherein the repression of transcription is heritable.
150. The method of any one of claims 137-149, wherein the repression is reversible.
151. The method of claim 150, wherein the repression is reversible by use of an inhibitor of DNMT.
152. The method of claim 151, wherein the inhibitor of DNMT is a cytidine analog.
153. The method of claim 151 or claim 152, wherein the inhibitor of DNMT is selected from the group consisting of azacytidine, decitabine, clofarabine, and zebularine.
154. A method of treating a *PCSK9*-related disease in a subject in need thereof, comprising administering a therapeutically effective dose of:
- the system of any one of claims 1-127;
 - the LNP of any one of claims 128-132; or
 - the pharmaceutical composition of any one of claims 133-136,
- thereby treating the *PCSK9*-related disease.

- 155.** The method of claim 154, wherein the subject is treated with a therapeutically effective dose of the LNP.
- 156.** The method of claim 154 or claim 155, wherein the LNP is administered by a route of administration selected from the group consisting of intravenous, intraarterial, intraportal vein injection, intraperitoneal, intramuscular, intracerebroventricular, intracisternal, intrathecal, intracranial, intralumbar, intraocular, subcutaneous, and oral routes.
- 157.** The method of any one of claims 154-156, wherein the *PCSK9*-related disease is selected from the group consisting of autosomal dominant hypercholesterolemia (ADH), hypercholesterolemia, elevated total cholesterol levels, hyperlipidemia, elevated low-density lipoprotein (LDL) levels, elevated LDL-cholesterol levels, reduced high-density lipoprotein levels, liver steatosis, coronary heart disease, ischemia, stroke, peripheral vascular disease, thrombosis, type 2 diabetes, high elevated blood pressure, atherosclerosis, obesity, aortic stenosis, elevated *PCSK9* levels, or a combination thereof.
- 158.** The method of any one of claims 154-157, wherein the method results in improvement in at least one clinically-relevant endpoint selected from the group consisting of change from baseline in LDL-cholesterol, decrease in plaque atheroma volume, reduction in coronary plaque, reduction in atherosclerotic cardiovascular disease (ASCVD), cardiovascular death, nonfatal myocardial infarction, ischemic stroke, nonfatal stroke, coronary revascularization, unstable angina, and visual acuity.
- 159.** The method of any one of claims 154-158, wherein the method results in improvement in at least two clinically-relevant endpoints selected from the group consisting of change from baseline in LDL-cholesterol, decrease in plaque atheroma volume, reduction in coronary plaque, reduction in atherosclerotic cardiovascular disease (ASCVD), cardiovascular death, nonfatal myocardial infarction, ischemic stroke, nonfatal stroke, coronary revascularization, unstable angina and visual acuity.
- 160.** A kit comprising the system of any one of claims 1-127, the LNP of any one of claims 128-132, the pharmaceutical composition of any one of claims 133-136, or combinations thereof, and a suitable container.
- 161.** The kit of claim 160, comprising a buffer, an excipient, a nuclease inhibitor, a protease inhibitor, a liposome, a therapeutic agent, a label, a label visualization reagent, instructions

for use, or any combination of the foregoing.

162. A composition comprising:

- a. the system of any one of claims 1-127;
- b. the LNP of any one of claims 128-132; or
- c. the pharmaceutical composition of any one of claims 133-136.

163. The composition of claim 162, for use in the manufacture of a medicament for the treatment of a *PCSK9*-related disease in a subject in need thereof.

164. The composition of claim 162, for use in the treatment of a *PCSK9*-related disease in a subject in need thereof.