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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
09 November 2023 (09.11.2023)

WIPO | PCT

(10) International Publication Number
WO 2023/215711 A1

(51) International Patent Classification:
C07K 14/47 (2006.01) *C12N 9/10* (2006.01)
A61K 38/00 (2006.01) *C12N 9/22* (2006.01)
A61P 3/06 (2006.01)

(21) International Application Number:
PCT/US2023/066436

(22) International Filing Date:
01 May 2023 (01.05.2023)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

63/337,164	01 May 2022 (01.05.2022)	US
63/337,167	01 May 2022 (01.05.2022)	US
63/355,083	23 June 2022 (23.06.2022)	US

(71) Applicant: **CHROMA MEDICINE, INC.** [US/US]; 201 Brookline Avenue, Suite 1101, Boston, MA 02215 (US).

(72) Inventors: **ABUBUCKER, Noorussahar**; c/o Chroma Medicine, Inc., 201 Brookline Avenue, Suite 1101, Boston, MA 02215 (US). **FRIEDLAND, Ari**; c/o Chroma Medicine, Inc., 201 Brookline Avenue, Suite 1101, Boston, MA 02215 (US). **MAEDER, Morgan**; c/o Chroma Medicine, Inc., 201 Brookline Avenue, Suite 1101, Boston, MA 02215 (US). **MYER, Vic**; c/o Chroma Medicine, Inc., 201 Brookline Avenue, Suite 1101, Boston, MA 02215 (US). **TREMBLAY, Frederic**; c/o Chroma Medicine, Inc., 201 Brookline Avenue, Suite 1101, Boston, MA 02215 (US). **MORRISON, Mary, Shirley**; c/o Chroma Medicine, Inc., 201 Brookline Avenue, Suite 1101, Boston, MA 02215 (US).

(74) Agent: **WEEKS, Anne, E.** et al.; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210-2206 (US).

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

(54) Title: COMPOSITIONS AND METHODS FOR EPIGENETIC REGULATION OF PCSK9 EXPRESSION

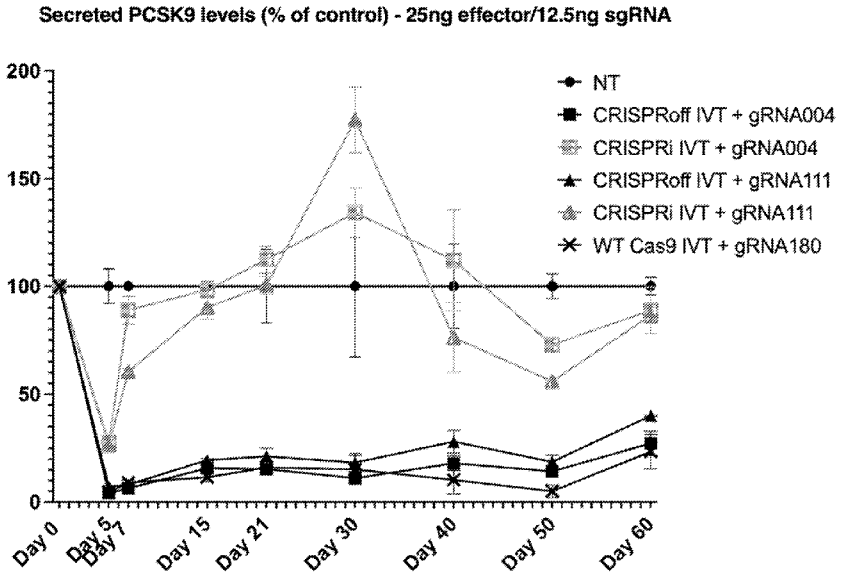


FIG. 5

(57) Abstract: This application relates to compositions and methods comprising epigenetic editors for epigenetic modification of *PCSK9*, as well as nucleic acids and vectors encoding the same. Also disclosed are cells epigenetically modified by the epigenetic editors.

WO 2023/215711 A1

[Continued on next page]

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

Published:

— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*

COMPOSITIONS AND METHODS FOR EPIGENETIC REGULATION OF *PCSK9* EXPRESSION

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit under 35 U.S.C. § 119(e) of U.S. Provisional Application No. 63/337,164, filed May 1, 2022, entitled "COMPOSITIONS AND METHODS FOR EPIGENETIC REGULATION OF PCSK9 EXPRESSION," U.S. Provisional Application No. 63/337,167, filed May 1, 2022, entitled "COMPOSITIONS AND METHODS FOR EPIGENETIC REGULATION OF PCSK9 EXPRESSION," and U.S. Provisional Application No. 63/355,083, filed June 23, 2022, entitled "COMPOSITIONS AND METHODS FOR EPIGENETIC REGULATION OF PCSK9 EXPRESSION," the entire disclosure of each of which is hereby incorporated by reference in its entirety.

REFERENCE TO AN ELECTRONIC SEQUENCE LISTING

[0002] The contents of the electronic sequence listing (C169870034WO00-SEQ-AXW.xml; Size: 1,831,551 bytes; and Date of Creation: May 1, 2023) is herein incorporated by reference in its entirety.

BACKGROUND

[0003] Genome editing has been considered a promising therapeutic approach for the treatment of genetic disease for over a decade. However, manipulation on the DNA level using traditional genetic editors remains risky given the potential for undesired double-strand DNA breaks, heterogenous repair (including large and small insertions and deletions at the intended site), and toxicity. In contrast, targeted epigenetic modification offers the potential to alter gene expression without leading to double-strand break-induced genotoxicity.

[0004] One promising candidate for epigenetic silencing is the proprotein convertase subtilisin/kexin type 9 (*PCSK9*) gene. *PCSK9* is a key target in the treatment of heart disease, the leading cause of mortality worldwide ((Berberich et al., *Nature Rev Cardiol.* (2019) 16(1):9-20). The human *PCSK9* gene, located on chromosome 1, has approximately 94% and 80% homology with its cynomolgus and mouse counterparts, respectively. The gene has CpG islands in the promoter region and is distal from other genes and cis-regulatory features. The PCSK9 protein is produced predominantly by the liver.

[0005] In humans, PCSK9 plays a key role in regulating the circulating level of low-density lipoprotein (LDL) particles as a result of its binding to the LDL receptor (LDLR). LDLR

reduces the circulating concentration of LDL particles by mediating their endocytosis and degradation in the cell. In the absence of PCSK9, or if the interaction of PCSK9 with LDLR is blocked, the rate of recycling of LDLR to the cell surface is increased and recycled LDLR proteins continue to remove LDL particles from the extracellular fluid (Tombling et al., *Atherosclerosis* (2021) 330:52-60). By contrast, when the endocytosed LDLR is bound to PCSK9, LDLR is degraded along with its passenger LDL particle. Clinical and genetic studies have established that circulating LDL causes atherosclerotic cardiovascular disease (Ference et al., *Eur Heart J* (2017) 38:2459-72). In addition, loss-of-function mutations in *PCSK9* are associated with low LDL levels (Zhao et al., *Am J Hum Genet.* (2006) 79(3):514-23). Genetic or pharmacologic reduction of PCSK9 decreases cardiovascular events (Ference et al., *N Engl J Med.* (2016) 375(22):2144-53; Sabatine et al., *N Engl J Med.* (2017) 376(18):1713-22). Lowering PCSK9 expression can help to increase the recycling of LDLR, which would lead to lower blood LDL particle concentrations.

[0006] In view of the critical role of PCSK9 in the pathogenesis of hypercholesterolemia and cardiovascular disease, there is a need for new and improved therapies that target the expression of *PCSK9*.

SUMMARY

[0007] The present disclosure provides systems and compositions for epigenetic modification ("epigenetic editors" or "epigenetic editing systems" herein), and methods of using the same to generate epigenetic modification at *PCSK9*, including in host cells and organisms.

[0008] In some aspects, the present disclosure provides a system for repressing transcription of a human *PCSK9* gene in a human cell, optionally a human hepatocyte, comprising

a) one or more fusion proteins that collectively comprise

a DNA methyltransferase (DNMT) domain and/or a domain that recruits a DNMT, optionally wherein the DNMT domain and/or the recruiter domain comprise a DNMT3A domain and/or a DNMT3L domain, and optionally wherein the recruited DNMT is DNMT3A, and

a transcriptional repressor domain,

each domain being linked to a DNA-binding domain that binds to a target region in the human *PCSK9* gene; or

b) one or more nucleic acid molecules encoding the one or more fusion proteins.

[0009] In some embodiments, the DNA-binding domain binds to a target sequence in SEQ ID NO: 1488 or 1489. In certain embodiments, the DNA-binding domain targets the fusion protein(s) to one or more sequences in the *PCSK9* gene selected from SEQ ID NOs: 700-747 and 1036-1261.

[0010] In some embodiments, the DNA-binding domain comprises a dead CRISPR Cas (dCas) domain, a ZFP domain, or a TALE domain. For example, the DNA-binding domain may comprise a dCas9 domain, and the system may further comprise (i) one or more guide RNAs (e.g., comprising any one of SEQ ID NOs: 1262-1487), or (ii) nucleic acid molecules coding for the one or more guide RNAs. In certain embodiments, the dCas domain comprises a dCas9 sequence, such as a sequence with at least 90% identity to SEQ ID NO: 12 or 13.

[0011] In some embodiments, the fusion protein comprises a dead CRISPR Cas (dCas) domain and the system comprises one or more PCSK9-binding guide RNAs (gRNAs) provided herein. In some embodiments, the system comprises a single gRNA. In some embodiments, the system comprises 2 gRNAs. In some embodiments, the system comprises 3 gRNAs. In some embodiments, the system comprises 4 gRNAs. In some embodiments, the system comprises 5 or more gRNAs. In some embodiments, the system comprises a gRNA selected from the gRNAs provided in Table 2. In some embodiments, the system comprises a gRNA selected from the gRNAs provided in Table 7. In some embodiments, the system comprises a gRNA selected from the gRNAs provided in Table 8. In some embodiments, the system comprises a sgRNA selected from the gRNAs provided in Table 10. In some embodiments, the system comprises a gRNA selected from the gRNAs provided in Table 12.

[0012] In some embodiments, the system comprises a gRNA comprising the gRNA Targeting Sequence of gRNA009 of Table 10, or a gRNA binding the same target domain sequence as gRNA009. In some embodiments, the system comprises a gRNA comprising the gRNA Targeting Sequence of gRNA003 of Table 10, or a gRNA binding the same target domain sequence as gRNA003. In some embodiments, the system comprises a gRNA comprising the gRNA Targeting Sequence of gRNA093 of Table 10, or a gRNA binding the same target domain sequence as gRNA093. In some embodiments, the system comprises a gRNA comprising the gRNA Targeting Sequence of gRNA011 of Table 10, or a gRNA binding the same target domain sequence as gRNA011. In some embodiments, the system comprises a gRNA comprising the gRNA Targeting Sequence of gRNA007 of Table 10, or a gRNA binding the same target domain sequence as gRNA007. In some embodiments, the

system comprises a gRNA comprising the gRNA Targeting Sequence of gRNA077 of Table 10, or a gRNA binding the same target domain sequence as gRNA077. In some embodiments, the system comprises a gRNA comprising the gRNA Targeting Sequence of gRNA113 of Table 10, or a gRNA binding the same target domain sequence as gRNA113. In some embodiments, the system comprises a gRNA comprising the gRNA Targeting Sequence of gRNA004 of Table 10, or a gRNA binding the same target domain sequence as gRNA004. In some embodiments, the system comprises a gRNA comprising the gRNA Targeting Sequence of gRNA008 of Table 10, or a gRNA binding the same target domain sequence as gRNA008. In some embodiments, the system comprises a gRNA comprising the gRNA Targeting Sequence of gRNA012 of Table 10, or a gRNA binding the same target domain sequence as gRNA012. In some embodiments, the system comprises a gRNA comprising the gRNA Targeting Sequence of gRNA111 of Table 10, or a gRNA binding the same target domain sequence as gRNA111. In some embodiments, the system comprises a gRNA comprising the gRNA Targeting Sequence of gRNA005 of Table 10, or a gRNA binding the same target domain sequence as gRNA005. In some embodiments, the system comprises a gRNA comprising the gRNA Targeting Sequence of gRNA013 of Table 10, or a gRNA binding the same target domain sequence as gRNA013.

[0013] In some embodiments, the system comprises Fusion Protein 9, variant 1, (Example 12) and gRNA g041. In some embodiments, the system comprises Fusion Protein 9 variant 2 (Example 12) and gRNA g041. In some embodiments, the system comprises Fusion Protein 10 (Example 12) and gRNA g041. In some embodiments, the system comprises Fusion Protein 11 (Example 12) and gRNA g041. In some embodiments, the system comprises Fusion Protein 12 (Example 12) and gRNA g041. In some embodiments, the system comprises Fusion Protein 13 (Example 12) and gRNA g041. In some embodiments, the system comprises Fusion Protein 14 (Example 12) and gRNA g041. In some embodiments, the system comprises Fusion Protein 15 (Example 12) and gRNA g041. In some embodiments, the system comprises Fusion Protein 9 variant 1 (Example 12) and gRNA g049. In some embodiments, the system comprises Fusion Protein 9 variant 2 (Example 12) and gRNA g049. In some embodiments, the system comprises Fusion Protein 10 (Example 12) and gRNA g049. In some embodiments, the system comprises Fusion Protein 11 (Example 12) and gRNA g049. In some embodiments, the system comprises Fusion Protein 12 (Example 12) and gRNA g049. In some embodiments, the system comprises Fusion Protein 13 (Example 12) and gRNA g049. In some embodiments, the system comprises Fusion Protein 14 (Example 12) and gRNA g049. In some embodiments,

the system comprises Fusion Protein 15 (Example 12) and gRNA g049. In some embodiments, the system comprises Fusion Protein 9 variant 1 (Example 12) and gRNAs g041 and g049. In some embodiments, the system comprises Fusion Protein 9 variant 2 (Example 12) and gRNAs g041 and g049. In some embodiments, the system comprises Fusion Protein 10 (Example 12) and gRNAs g041 and g049. In some embodiments, the system comprises Fusion Protein 11 (Example 12) and gRNAs g041 and g049. In some embodiments, the system comprises Fusion Protein 12 (Example 12) and gRNAs g041 and g049. In some embodiments, the system comprises Fusion Protein 13 (Example 12) and gRNAs g041 and g049. In some embodiments, the system comprises Fusion Protein 14 (Example 12) and gRNAs g041 and g049. In some embodiments, the system comprises Fusion Protein 15 (Example 12) and gRNAs g041 and g049.

[0014] In some embodiments, the DNA-binding domain comprises a ZFP domain that targets a nucleotide sequence selected from SEQ ID NOs: 700-747. In certain embodiments, the ZFP domain comprises, in order, the F1-F6 amino acid sequences of any one of ZF001 through ZF048 as shown in **Table 1**.

[0015] In some embodiments, the DNMT3A domain comprises a sequence with at least 90% identity to SEQ ID NO: 574 or 575.

[0016] The DNMT3L domain may comprise, e.g., a sequence with at least 90% identity to a sequence selected from SEQ ID NOs: 578-581. In some embodiments, the DNMT3L domain comprises a sequence with at least 90% identity to a sequence selected from SEQ ID NOs: 582-603. In some embodiments, the DNMT3L domain comprises a sequence with at least 90% identity to a sequence selected from SEQ ID NOs: 601-603.

[0017] In some embodiments, the transcriptional repressor domain comprises a sequence with at least 90% identity to a sequence selected from SEQ ID NOs: 33-570. In certain embodiments, the transcriptional repressor domain is a KRAB domain derived from KOX1, ZIM3, ZFP28, or ZN627. The KRAB domain may comprise, e.g., a sequence with at least 90% identity to a sequence selected from SEQ ID NOs: 89, 116, 245, and 255. In some embodiments, the transcriptional repressor domain comprises a fusion of the N- and C-terminal regions of ZIM3 and KOX1 KRAB, and optionally comprises the amino acid sequence of SEQ ID NO: 571 or 572. In certain embodiments, the transcriptional repressor domain is derived from KAP1, MECP2, HP1a/CBX5, HP1b, CBX8, CDYL2, TOX, TOX3, TOX4, EED, EZH2, RBBP4, RCOR1, or SCML2.

[0018] In some embodiments, the system comprises

a) a fusion protein comprising the DNMT3A domain, the DNMT3L domain, the transcriptional repressor domain, and the DNA-binding domain,

optionally wherein one or both of the DNMT3A domain and the DNMT3L domain are human, and

optionally wherein the DNA-binding domain is a dead CRISPR Cas domain or a ZFP domain; or

b) a nucleic acid molecule encoding the fusion protein.

[0019] In certain embodiments, the fusion protein comprises, from N-terminus to C-terminus, the DNMT3A domain, a first peptide linker, the DNMT3L domain, a second peptide linker, the DNA-binding domain, a third peptide linker, and the transcriptional repressor domain. For example, the fusion protein may comprise, from N-terminus to C-terminus, the DNMT3A domain, the first peptide linker, the DNMT3L domain, the second peptide linker, a first nuclear localization signal (NLS), the DNA-binding domain, a second NLS, the third peptide linker, and the transcriptional repressor domain. The fusion protein may comprise, from N-terminus to C-terminus, a first NLS, the DNMT3A domain, the first peptide linker, the DNMT3L domain, the second peptide linker, the DNA-binding domain, the third peptide linker, the transcriptional repressor domain, and a second NLS. The fusion protein may comprise, from N-terminus to C-terminus, first and second NLSs, the DNMT3A domain, the first peptide linker, the DNMT3L domain, the second peptide linker, the DNA-binding domain, the third peptide linker, the transcriptional repressor domain, and third and fourth NLSs. In particular embodiments, the transcriptional repressor domain is a KRAB domain, such as a human KOX1, ZFP28, ZN627, or ZIM3 KRAB domain. In particular embodiments, one or both of the second and third peptide linkers are XTEN linkers, which may be selected from XTEN80 (e.g., SEQ ID NO: 643) and XTEN16 (e.g., SEQ ID NO: 638), e.g., wherein the second peptide linker is XTEN80, and the third peptide linker is XTEN16.

[0020] In some embodiments, the fusion protein may comprise, from N-terminus to C-terminus, a human DNMT3A domain, a first peptide linker, a human DNMT3L domain, an XTEN80 peptide linker, a first NLS, a dSpCas9 domain, a second NLS, an XTEN16 peptide linker, and a human KOX1 KRAB domain. In certain embodiments, the fusion protein comprises SEQ ID NO: 658 or a sequence at least 90% identical thereto. In certain embodiments, the fusion protein comprises SEQ ID NO: 1495 or a sequence at least 90% identical thereto.

[0021] In some embodiments, the fusion protein comprises, from N-terminus to C-terminus, a human DNMT3A domain, a first peptide linker, a human DNMT3L domain, an XTEN80 peptide linker, a first NLS, a ZFP domain, a second NLS, an XTEN16 linker, and a human KOX1 KRAB domain. In certain embodiments, the fusion protein comprises SEQ ID NO: 659 or a sequence at least 90% identical thereto, optionally wherein the ZFP comprises, in order, the F1-F6 amino acid sequences of any one of ZF001 through ZF048 as shown in **Table 1**. In certain embodiments, the fusion protein comprises SEQ ID NO: 1496 or a sequence at least 90% identical thereto, optionally wherein the ZFP comprises, in order, the F1-F6 amino acid sequences of any one of ZF001 through ZF048 as shown in **Table 1**.

[0022] In some embodiments, the fusion protein comprises, from N-terminus to C-terminus, first and second NLSs, a human DNMT3A domain, a first peptide linker, a human DNMT3L domain, an XTEN80 peptide linker, a dSpCas9 domain, an XTEN16 peptide linker, a human KOX1 KRAB domain, and third and fourth NLSs. In particular embodiments, the fusion protein may comprise the amino acid sequence of SEQ ID NO: 660 or a sequence at least 90% identical thereto.

[0023] In some embodiments, the fusion protein comprises, from N-terminus to C-terminus, first and second NLSs, a human DNMT3A domain, a first peptide linker, a human DNMT3L domain, an XTEN80 peptide linker, a ZFP domain, an XTEN16 peptide linker, a human KOX1 KRAB domain, and third and fourth NLSs.

[0024] In some embodiments, the fusion protein comprises, from N-terminus to C-terminus, first and second NLSs, a human DNMT3A domain, a first peptide linker, a human DNMT3L domain, an XTEN80 peptide linker, a dSpCas9 domain, an XTEN16 peptide linker, a human ZFP28 KRAB domain, and third and fourth NLSs. In particular embodiments, the fusion protein may comprise the amino acid sequence of SEQ ID NO: 661 or a sequence at least 90% identical thereto.

[0025] In some embodiments, the fusion protein comprises, from N-terminus to C-terminus, first and second NLSs, a human DNMT3A domain, a first peptide linker, a human DNMT3L domain, an XTEN80 peptide linker, a ZFP domain, an XTEN16 peptide linker, a human ZFP28 KRAB domain, and third and fourth NLSs.

[0026] In some embodiments, the fusion protein comprises, from N-terminus to C-terminus, first and second NLSs, a human DNMT3A domain, a first peptide linker, a human DNMT3L domain, an XTEN80 peptide linker, a dSpCas9 domain, an XTEN16 peptide linker, a human ZN627 KRAB domain, and third and fourth NLSs. In particular

embodiments, the fusion protein may comprise the amino acid sequence of SEQ ID NO: 662 or a sequence at least 90% identical thereto.

[0027] In some embodiments, the fusion protein comprises, from N-terminus to C-terminus, first and second NLSs, a human DNMT3A domain, a first peptide linker, a human DNMT3L domain, an XTEN80 peptide linker, a ZFP domain, an XTEN16 peptide linker, a human ZN627 KRAB domain, and third and fourth NLSs.

[0028] In some embodiments, the fusion protein comprises, from N-terminus to C-terminus, first and second NLSs, a human DNMT3A domain, a first peptide linker, a human DNMT3L domain, an XTEN80 peptide linker, a dSpCas9 domain, an XTEN16 peptide linker, a human ZIM3 KRAB domain, and third and fourth NLSs. In particular embodiments, the fusion protein may comprise the amino acid sequence of SEQ ID NO: 663 or a sequence at least 90% identical thereto.

[0029] In some embodiments, the fusion protein comprises, from N-terminus to C-terminus, first and second NLSs, a human DNMT3A domain, a first peptide linker, a human DNMT3L domain, an XTEN80 peptide linker, a ZFP domain, an XTEN16 peptide linker, a human ZIM3 KRAB domain, and third and fourth NLSs.

[0030] In some embodiments, at least one of the NLSs in a fusion protein described herein is an SV40 NLS (e.g., SEQ ID NO: 644).

[0031] In some embodiments, the system comprises:

a) a first fusion protein comprising a first DNA-binding domain and comprising or recruiting the DNMT3A domain,

a second fusion protein comprising a second DNA-binding domain and comprising or recruiting the DNMT3L domain, and

a third fusion protein comprising a third DNA-binding domain and comprising or recruiting the transcriptional repressor domain; or

b) one or more nucleic acid molecules encoding the fusion proteins.

[0032] The present disclosure also provides a human cell comprising a system described herein, or progeny of the cell. In some embodiments, the cell is a hepatocyte.

[0033] The present disclosure also provides a pharmaceutical composition comprising a system described herein and a pharmaceutically acceptable excipient. In some embodiments, the composition comprises lipid nanoparticles (LNPs) comprising the system, and/or the DNA-binding domain is a dCas domain and the LNPs further comprise one or more gRNAs.

[0034] The present disclosure also provides a method of treating a patient in need thereof, comprising administering a system or pharmaceutical composition described herein to the

patient (e.g., intravenously). In some embodiments, the patient has heart disease; has elevated low-density lipoprotein cholesterol (LDL-C) or hypercholesterolemia; is at risk of developing myocardial infarction, stroke, or unstable angina; and/or has primary hyperlipidemia (e.g., heterozygous familial hypercholesterolemia (HeFH), or homozygous familial hypercholesterolemia (HoFH)).

[0035] The present disclosure also provides a system or pharmaceutical composition described herein for use in treating a patient in need thereof, e.g., in a method described herein.

[0036] The present disclosure also provides use of a system described herein in the manufacture of a medicament for treating a patient in need thereof, e.g., in a method described herein.

[0037] The present disclosure also provides articles and kits comprising the systems described herein.

[0038] Other features, objectives, and advantages of the disclosed methods and compositions are apparent in the detailed description that follows. It should be understood, however, that the detailed description, while indicating embodiments and embodiments of the disclosed methods and compositions, is given by way of illustration only, not limitation. Various changes and modifications within the scope of the disclosure will become apparent to those skilled in the art from the detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

[0039] **FIG. 1** is a diagram showing the predicted binding position of ZF proteins and computationally designed gRNAs on the *PCSK9* gene.

[0040] **FIG. 2** is a scatter plot showing the relative PCSK9 expression (y-axis) at day 7 in cells treated with CRISPR-off (DNMT3A-3L-dCas9-KRAB). The genomic distance from the gRNA target site to the *PCSK9* TSS is shown on the x-axis.

[0041] **FIG. 3** is a diagram showing the overlap of the top 40 gRNAs with the *PCSK9* gene.

[0042] **FIG. 4A** is a bar graph showing the level of secreted PCSK9 at day 7 and day 28 following treatment with the indicated gRNA. Dashed line shows silencing achieved by wildtype (WT) Cas9.

[0043] **FIG. 4B** is a scatter plot showing the correlation of *PCSK9* mRNA expression and PCSK9 protein secretion in cells following treatment with gRNAs. CRISPRi (dCas9-KRAB) represents a dCas9-KRAB fusion protein.

[0044] **FIG. 5** is a line graph showing the silencing of *PCSK9* following treatment with CRISPRi (dCas9-KRAB), CRISPR-off (DNMT3A-3L-dCas9-KRAB) and the indicated gRNAs.

[0045] **FIG. 6** is a bar graph showing PCSK9 secretion in cells treated with CRISPRoff and simvastatin, compared to cells treated with the CRISPRoff system alone.

[0046] **FIG. 7** is a bar graph showing the reduction of PCSK9 secretion in Huh7 hepatoma cells treated with CRISPRoff and the given gRNA.

[0047] **FIG. 8** is a scatter plot showing the activity and toxicity of 247 *PCSK9*-targeting ZF proteins. Relative PCSK9 expression is shown on the x-axis and corresponding cell counts relative to the pUC and off-target controls are shown on the y-axis. The diagonal line represents a 1:1 correlation between relative PCSK9 expression and cell count.

[0048] **FIG. 9** is a scatter plot showing relative PCSK9 expression (y-axis) by cells treated with a ZF-off (DNMT3A-3L-ZF-KRAB) construct and the corresponding targeted genomic distance relative to the *PCSK9* transcription start site (TSS) (x-axis).

[0049] **FIG. 10** is a diagram showing the entire human *PCSK9* gene locus flanked with 35.5 kb and 7 kb of upstream and downstream genomic regions (67.5 kb), respectively, that was introduced into and expressed in a transgenic mouse. This transgenic mouse line expresses human PCSK9 under the control of its own (human) endogenous promoter.

[0050] **FIG. 11A** shows schematic illustrations of fusion protein constructs with variant NLS configurations. **FIG. 11B** shows schematic illustrations of additional fusion protein constructs with variant KRAB domains.

[0051] **FIGs. 12A-12B** are graphs showing the percentage of PCSK9 protein levels measured after treatment with fusion protein constructs with various NLS placements in HeLa cells using 6.25 ng RNA (**FIG. 12A**) or 2.5ng RNA (**FIG. 12B**). Human and murine DNMT3L sequences are indicated as h3L and m3L, respectively.

[0052] **FIG. 13** is a graph showing that constructs with 2X NLSs are 3X more efficient than CRISPR-off in silencing *mPcsk9* in Hepa1-6 cells.

[0053] **FIG. 14A** is a graph showing that constructs with 2X NLSs are more efficient than CRISPR-off in silencing *mPcsk9* in Huh7 cells. **FIGs. 14B-14C** are graphs showing that constructs with 2X NLSs are also more efficient than CRISPR-off in silencing *mPcsk9* in Huh7 cells at different doses both at day 5 (**FIG. 14B**) and day 15 (**FIG. 14C**).

[0054] **FIG. 15** is a graph showing that, in Huh7 cells, in a CRISPR-off-like format in which dCas9 is replaced with a zinc finger, 2X NLS offers improvements across multiple ZFs.

[0055] **FIG. 16** is a set of graphs showing that methylation of the CTLA4 promoter with a bacterial DNMT protein can induce epigenetic silencing of the locus.

[0056] **FIG. 17** is a set of graphs showing methylation profiles at the VIM3 locus of cells treated with different constructs carrying bacterial DNA methyltransferases fused to dCas9, day 30. Samples treated with *M. SssI* are methylated by 20%.

[0057] **FIG. 18** is a set of graphs showing methylation profiles by hybridization capture at the CLTA locus of cells comparing *M. SssI* to murine DNMT3A/3L in dCas9 fusions, day 29.

[0058] **FIGs. 19A-19D** are a set of graphs showing alternative KRAB domains tested for epi-silencing activity against CRISPR-off when using 0.5 ng effector DNA using CLTA-GFP as a marker (**FIG. 19A**), 3 ng effector DNA using GFP as a marker (**FIG. 19B**), and 0.5 ng effector DNA using GFP as a marker (**FIG. 19C**). **FIG 19D** shows results after 30 days using varying nanogram amounts of effector DNA.

DETAILED DESCRIPTION

[0059] The present disclosure provides epigenetic editors for regulating expression of the *PCSK9* gene. By altering expression of *PCSK9*, the systems, compositions and methods described herein may be used for treating conditions such as hypercholesterolemia (e.g., heterozygous familial hypercholesterolemia (HeFH), homozygous familial hypercholesterolemia (HoFH), familial hypercholesterolemia (HF), or established atherosclerotic cardiovascular disease (ASCVD)), or renal insufficiency (RI). Unless otherwise stated, "*PCSK9*" refers herein to human *PCSK9*. A human *PCSK9* gene sequence can be found at Ensembl Accession No. ENSG00000169174. The present epigenetic editors have several advantages compared to other genome engineering methods, including reversibility, decreased risk of translocation, and durable, inheritable silencing.

[0060] In some embodiments, the region of the human *PCSK9* gene targeted for epigenetic regulation is about 2 kb long, and is approximately +/- 1 kb of the *PCSK9* TSS. In certain embodiments, the region has the nucleotide sequence of SEQ ID NO: 1488. In some embodiments, the targeted *PCSK9* region is about 1069 bps long, and is approximately +/- 500 bps of the *PCSK9* TSS. In certain embodiments, the region targeted has the nucleotide sequence of SEQ ID NO: 1489. The TSS of *PCSK9* is at #chr1:55039548 of Genome GRCh38.

[0061] In some embodiments, an epigenetic editor as described herein may comprise one or more fusion proteins, wherein each fusion protein comprises a DNA-binding domain

linked to one or more effector domains for epigenetic modification. In certain embodiments, where the DNA-binding domain is a polynucleotide guided DNA-binding domain, the epigenetic editor may further comprise one or more guide polynucleotides. DNA-binding domains, effector domains, and guide polynucleotides of an epigenetic editor as described herein may be selected, e.g., from those described below, in any functional combination.

[0062] The epigenetic editors described herein may be expressed in a host cell transiently, or may be integrated in a genome of the host cell; such cells and their progeny are also contemplated by the present disclosure. Both transiently expressed and integrated epigenetic editors or components thereof can effect stable epigenetic modifications. For example, after introducing to a host cell an epigenetic editor described herein, the target gene in the host cell may be stably or permanently repressed or silenced. In some embodiments, expression of the target gene is reduced or silenced for at least 1 week, at least 2 weeks, at least 3 weeks, at least 4 weeks, at least 5 weeks, at least 6 weeks, at least 7 weeks, at least 2 months, at least 3 months, at least 4 months, at least 5 months, at least 6 months, at least 1 year, at least 2 years, or for the entire lifetime of the cell or the subject carrying the cell, as compared to the level of expression in the absence of the epigenetic editor. The epigenetic modification may be inherited by the progeny of the host cells into which the epigenetic editor was introduced.

[0063] The present epigenetic editors may be introduced to a patient in need thereof (e.g., a human patient), e.g., into the patient's hepatocytes, biliary epithelial cells (cholangiocytes), stellate cells, Kupffer cells, and liver sinusoidal endothelial cells.

I. DNA-Binding Domains

[0064] An epigenetic editor described herein may comprise one or more DNA-binding domains that direct the effector domain(s) of the epigenetic editor to target sequences within or close to the *PCSK9* gene locus. A DNA-binding domain as described herein may be, e.g., a polynucleotide guided DNA-binding domain, a zinc finger protein (ZFP) domain, a transcription activator like effector (TALE) domain, a meganuclease DNA-binding domain, and the like. Examples of DNA-binding domains can be found in U.S. Patent 11,162,114, which is incorporated by refence herein in its entirety.

[0065] In some embodiments, a DNA-binding domain described herein is encoded by its native coding sequence. In other embodiments, the DNA-binding domain is encoded by a nucleotide sequence that has been codon-optimized for optimal expression in human cells.

A. Polynucleotide Guided DNA-Binding Domains

[0066] In some embodiments, a DNA-binding domain herein may be a protein domain directed by a guide nucleic acid sequence (e.g., a guide RNA sequence) to a target site in the *PCSK9* gene locus. In certain embodiments, the protein domain may be derived from a CRISPR-associated nuclease, such as a Class I or II CRISPR-associated nuclease. In some embodiments, the protein domain may be derived from a Cas nuclease such as a Type II, Type IIA, Type IIB, Type IIC, Type V, or Type VI Cas nuclease. In certain embodiments, the protein domain may be derived from a Class II Cas nuclease selected from Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9, Cas10, Cas14a, Cas14b, Cas14c, CasX, CasY, CasPhi, C2c4, C2c8, C2c9, C2c10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx1S, Csf1, Csf2, CsO, Csf4, and homologues and modified versions thereof. "Derived from" is used to mean that the protein domain comprises the full polypeptide sequence of the parent protein, or comprises a variant thereof (e.g., with amino acid residue deletions, insertions, and/or substitutions). The variant retains the desired function of the parent protein (e.g., the ability to form a complex with the guide nucleic acid sequence and the target DNA).

[0067] In some embodiments, the CRISPR-associated protein domain may be a Cas9 domain described herein. Cas9 may, for example, refer to a polypeptide with at least about 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity and/or sequence similarity to a wildtype Cas9 polypeptide described herein. In some embodiments, said wildtype polypeptide is Cas9 from *Streptococcus pyogenes* (NCBI Ref. No. NC_002737.2 (SEQ ID NO: 1)) and/or UniProt Ref. No. Q99ZW2 (SEQ ID NO: 2). In some embodiments, said wildtype polypeptide is Cas9 from *Staphylococcus aureus* (SEQ ID NO: 3). In some embodiments, the CRISPR-associated protein domain is a Cpf1 domain or protein, or a polypeptide with at least about 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity and/or sequence similarity to a wildtype Cpf1 polypeptide described herein (e.g., Cpf1 from *Franscisella novicida* (UniProt Ref. No. U2UMQ6 or SEQ ID NO: 4). In certain embodiments, the CRISPR-associated protein domain may be a modified form of the wildtype protein comprising one or more amino acid residue changes such as a deletion, an insertion, or a substitution; a fusion or chimera; or any combination thereof.

[0068] Cas9 sequences and structures of variant Cas9 orthologs have been described for various organisms. Exemplary organisms from which a Cas9 domain herein can be derived

include, but are not limited to, *Streptococcus pyogenes*, *Streptococcus thermophilus*, *Streptococcus sp.*, *Staphylococcus aureus*, *Listeria innocua*, *Lactobacillus gasseri*, *Francisella novicida*, *Wolinella succinogenes*, *Sutterella wadsworthensis*, *Gamma proteobacterium*, *Neisseria meningitidis*, *Campylobacter jejuni*, *Pasteurella multocida*, *Fibrobacter succinogene*, *Rhodospirillum rubrum*, *Nocardiopsis dassonvillei*, *Streptomyces pristinaespiralis*, *Streptomyces viridochromogenes*, *Streptomyces viridochromogenes*, *Streptosporangium roseum*, *Alicyclobacillus acidocaldarius*, *Bacillus pseudomycoides*, *Bacillus selenitireducens*, *Exiguobacterium sibiricum*, *Lactobacillus delbrueckii*, *Lactobacillus salivarius*, *Lactobacillus buchneri*, *Treponema denticola*, *Microscilla marina*, *Burkholderiales bacterium*, *Polaromonas naphthalenivorans*, *Polaromonas sp.*, *Crocosphaera watsonii*, *Cyanothece sp.*, *Microcystis aeruginosa*, *Synechococcus sp.*, *Acetohalobium arabaticum*, *Ammonifex degensii*, *Caldicelulosiruptor becscii*, *Candidatus Desulforudis*, *Clostridium botulinum*, *Clostridium difficile*, *Finegoldia magna*, *Natranaerobius thermophilus*, *Pelotomaculum thermopropionium*, *Acidithiobacillus caldus*, *Acidithiobacillus ferrooxidans*, *Allochromatium vinosum*, *Marinobacter sp.*, *Nitrosococcus halophilus*, *Nitrosococcus watsoni*, *Pseudoalteromonas haloplanktis*, *Ktedonobacter racemifer*, *Methanohalobium evestigatum*, *Anabaena variabilis*, *Nodularia spumigena*, *Nostoc sp.*, *Arthrospira maxima*, *Arthrospira platensis*, *Arthrospira sp.*, *Lyngbya sp.*, *Microcoleus chthonoplastes*, *Oscillatoria sp.*, *Petrotoga mobilis*, *Thermosipho africanus*, *Streptococcus pasteurianus*, *Neisseria cinerea*, *Campylobacter lari*, *Parvibaculum lavamentivorans*, *Corynebacterium diphtheria*, and *Acaryochloris marina*. Cas9 sequences also include those from the organisms and loci disclosed in Chylinski et al., *RNA Biol.* (2013) 10(5):726-37.

[0069] In some embodiments, the Cas9 domain is from *Streptococcus pyogenes* (*SpCas9*). In some embodiments, the Cas9 domain is from *Staphylococcus aureus* (*SaCas9*).

[0070] Other Cas domains are also contemplated for use in the epigenetic editors herein. These include, for example, those from CasX (Cas12E) (e.g., SEQ ID NO: 5), CasY (Cas12d) (e.g., SEQ ID NO: 6), Casφ (CasPhi) (e.g., SEQ ID NO: 7), Cas12f1 (Cas14a) (e.g., SEQ ID NO: 8), Cas12f2 (Cas14b) (e.g., SEQ ID NO: 9), Cas12f3 (Cas14c) (e.g., SEQ ID NO: 10), and C2c8 (e.g., SEQ ID NO: 11).

[0071] For epigenetic editing, the nuclease-derived protein domain (e.g., a Cas9 or Cpf1 domain) may have reduced or no nuclease activity through mutations such that the protein domain does not cleave DNA or has reduced DNA-cleaving activity while retaining the ability to complex with the guide nucleic acid sequence (e.g., guide RNA) and the target

DNA. For example, the nuclease activity may be reduced by at least 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 99% compared to the wildtype domain. In some embodiments, a CRISPR-associated protein domain described herein is catalytically inactive ("dead"). Examples of such domains include, for example, dCas9 ("dead" Cas9), dCpf1, ddCpf1, dCasPhi, ddCas12a, dLbCpf1, and dFnCpf1. A dCas9 protein domain, for example, may comprise one, two, or more mutations as compared to wildtype Cas9 that abrogate its nuclease activity. The DNA cleavage domain of Cas9 is known to include two subdomains: the HNH nuclease subdomain and the RuvC1 subdomain. The HNH subdomain cleaves the strand complementary to the gRNA, whereas the RuvC1 subdomain cleaves the non-complementary strand. Mutations within these subdomains can silence the nuclease activity of Cas9. For example, the mutations D10A (in RuvC1) and H840A (in HNH) completely inactivate the nuclease activity of SpCas9. SaCas9, similarly, may be inactivated by the mutations D10A and N580A. In some embodiments, the dCas9 comprises at least one mutation in the HNH subdomain and/or the RuvC1 subdomain that reduces or abrogates nuclease activity. In some embodiments, the dCas9 only comprises a RuvC1 subdomain, or only comprises an HNH subdomain. It is to be understood that any mutation that inactivates the RuvC1 and/or the HNH domain may be included in a dCas9 herein, e.g., insertion, deletion, or single or multiple amino acid substitution in the RuvC1 domain and/or the HNH domain.

[0072] In some embodiments, a dCas9 protein herein comprises a mutation at position(s) corresponding to position D10 (e.g., D10A), H840 (e.g., H840A), or both, of a wildtype SpCas9 sequence as numbered in the sequence provided at UniProt Accession No. Q99ZW2 (SEQ ID NO: 2). In particular embodiments, the dCas9 comprises the amino acid sequence of dSpCas9 (D10A and H840A) (SEQ ID NO: 12).

[0073] In some embodiments, a dCas9 protein as described herein comprises a mutation at position(s) corresponding to position D10 (e.g., D10A), N580 (e.g., N580A), or both, of a wildtype SaCas9 sequence (e.g., SEQ ID NO: 3). In particular embodiments, the dCas9 comprises the amino acid sequence of dSaCas9 (D10A and N580A) (SEQ ID NO.: 13).

[0074] Additional suitable mutations that inactivate Cas9 will be apparent to those of skill in the art based on this disclosure and knowledge in the field and are within the scope of this disclosure. Such mutations may include, but are not limited to, D839A, N863A, and/or K603R in SpCas9. The present disclosure contemplates any mutations that reduce or abrogate the nuclease activity of any Cas9 described herein (e.g., mutations corresponding to any of the Cas9 mutations described herein).

[0075] A dCpf1 protein domain may comprise one, two, or more mutations as compared to wildtype Cpf1 that reduce or abrogate its nuclease activity. The Cpf1 protein has a RuvC-like endonuclease domain that is similar to the RuvC domain of Cas9, but does not have an HNH endonuclease domain, and the N-terminal of Cpf1 does not have the alpha-helical recognition lobe of Cas9. In some embodiments, the dCpf1 comprises one or more mutations corresponding to position D917A, E1006A, or D1255A as numbered in the sequence of the *Francisella novicida* Cpf1 protein (FnCpf1; SEQ ID NO: 4). In certain embodiments, the dCpf1 protein comprises mutations corresponding to D917A, E1006A, D1255A, D917A/E1006A, D917A/D1255A, E1006A/D1255A, or D917A/ E1006A/D1255A, or corresponding mutation(s) in any of the Cpf1 amino acid sequences described herein. In some embodiments, the dCpf1 comprises a D917A mutation. In particular embodiments, the dCpf1 comprises the amino acid sequence of dFnCpf1 (SEQ ID NO: 14).

[0076] Further nuclease inactive CRISPR-associated protein domains contemplated herein include those from, for example, dNmeCas9 (e.g., SEQ ID NO: 15), dCjCas9 (e.g., SEQ ID NO: 16), dSt1Cas9 (e.g., SEQ ID NO: 17), dSt3Cas9 (e.g., SEQ ID NO: 18), dLbCpf1 (e.g., SEQ ID NO: 19), dAsCpf1 (e.g., SEQ ID NO: 20), denAsCpf1 (e.g., SEQ ID NO: 21), dHFAsCpf1 (e.g., SEQ ID NO: 22), dRVRAsCpf1 (e.g., SEQ ID NO: 23), dRRAsCpf1 (e.g., SEQ ID NO: 24), dCasX (e.g., SEQ ID NO: 25), and dCasPhi (e.g., SEQ ID NO: 26).

[0077] In some embodiments, a Cas9 domain described herein may be a high fidelity Cas9 domain, e.g., comprising one or more mutations that decrease electrostatic interactions between the Cas9 domain and the sugar-phosphate backbone of DNA to confer increased target binding specificity. In certain embodiments, the high fidelity Cas9 domain may be nuclease inactive as described herein.

[0078] A CRISPR-associated protein domain described herein may recognize a protospacer adjacent motif (PAM) sequence in a target gene. A "PAM" sequence is typically a 2 to 6 bp DNA sequence immediately following the sequence targeted by the CRISPR-associated protein domain. The PAM sequence is required for CRISPR protein binding and cleavage but is not part of the target sequence. The CRISPR-associated protein domain may either recognize a naturally occurring or canonical PAM sequence or may have altered PAM specificity. CRISPR-associated protein domains that bind to non-canonical PAM sequences have been described in the art. For example, Cas9 domains that bind non-canonical PAM sequences have been described in Kleinstiver et al., *Nature* (2015) 523(7561):481-5 and Kleinstiver et al., *Nat Biotechnol.* (2015) 33:1293-8. Such Cas9 domains may include, for

example, those from "VRER" SpCas9, "EQR" SpCas9, "VQR" SpCas9, "SpG Cas9," "SpRYCas9," and "KKH" SaCas9. Nuclease inactive versions of these Cas9 domains are also contemplated, such as nuclease inactive VRER SpCas9 (e.g., SEQ ID NO: 27), nuclease inactive EQR SpCas9 (e.g., SEQ ID NO: 28), nuclease inactive VQR SpCas9 (e.g., SEQ ID NO: 29), nuclease inactive SpG Cas9 (e.g., SEQ ID NO: 30), nuclease inactive SpRY Cas9 (e.g., SEQ ID NO: 31), and nuclease inactive KKH SaCas9 (e.g., SEQ ID NO: 32). Another example is the Cas9 of *Francisella novicida* engineered to recognize 5'-YG-3' (where "Y" is a pyrimidine).

[0079] Additional suitable CRISPR-associated proteins, orthologs, and variants, including nuclease inactive variants and sequences, will be apparent to those of skill in the art based on this disclosure.

[0080] Guide RNAs that can be used in conjunction with the CRISPR-associated protein domains herein are further described in Section II below.

B. Zinc Finger Protein Domains

[0081] In some embodiments, the DNA-binding domain of an epigenetic editor described herein comprises a zinc finger protein (ZFP) domain (or "ZF domain" as used herein). ZFPs are proteins having at least one zinc finger, and bind to DNA in a sequence-specific manner. A "zinc finger" (ZF) or "zinc finger motif" (ZF motif) refers to a polypeptide domain comprising a beta-beta-alpha (ββα)-protein fold stabilized by a zinc ion. A ZF binds from two to four base pairs of nucleotides, typically three or four base pairs (contiguous or noncontiguous). Each ZF typically comprises approximately 30 amino acids. ZFP domains may contain multiple ZFs that make tandem contacts with their target nucleic acid sequence. A tandem array of ZFs may be engineered to generate artificial ZFPs that bind desired nucleic acid targets. ZFPs may be rationally designed by using databases comprising triplet (or quadruplet) nucleotide sequences and individual ZF amino acid sequences, in which each triplet or quadruplet nucleotide sequence is associated with one or more amino acid sequences of ZFs that bind the particular triplet or quadruplet sequence. See, e.g., U.S. Patents 6,453,242, 6,534,261, and 8,772,453.

[0082] ZFPs are widespread in eukaryotic cells, and may belong to, e.g., C2H2 class, CCHC class, PHD class, or RING class. An exemplary motif characterizing one class of these proteins (C2H2 class) is -Cys-$(X)_{2-4}$-Cys-$(X)_{12}$-His-$(X)_{3-5}$-His- (SEQ ID NO: 657), where X is any independently chosen amino acid. In some embodiments, a ZFP domain herein may comprise a ZF array comprising sequential C2H2-ZFs each contacting three or more sequential nucleotides.

[0083] A ZFP domain of an epigenetic editor described herein may include 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or more ZFs. The ZFP domain may include an array of two-finger or three-finger units, e.g., 3, 4, 5, 6, 7, 8, 9 or 10 or more units, wherein each unit binds a subsite in the target sequence. In some embodiments, a ZFP domain comprising at least three ZFs recognizes a target DNA sequence of 9 or 10 nucleotides. In some embodiments, a ZFP domain comprising at least four ZFs recognizes a target DNA sequence of 12 to 14 nucleotides. In some embodiments, a ZFP domain comprising at least six ZFs recognizes a target DNA sequence of 18 to 21 nucleotides.

[0084] In some embodiments, ZFs in a ZFP domain described herein are connected via peptide linkers. The peptide linkers may be, e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more amino acids in length. In some embodiments, a linker comprises 5 or more amino acids. In some embodiments, a linker comprises 7-17 amino acids. The linker may be flexible or rigid.

[0085] In some embodiments a zinc finger array may have the sequence:

SRPGERPFQCRICMRNFSXXXXXXXH*XX*THTGEKPFQCRICMRNFSXXXXXXXH*XX*T
H**[linker]**FQCRICMRNFSXXXXXXXH*XX*THTGEKPFQCRICMRNFSXXXXXXXH*X*
*X*TH**[linker]**PFQCRICMRNFSXXXXXXXH*XX*THTGEKPFQCRICMRNFSXXXXXX
XH*XX*THLRGS (SEQ ID NO: 650),

or a sequence at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical thereto, where "XXXXXXX" represents the amino acids of the ZF recognition helix, which confers DNA-binding specificity upon the zinc finger; each X may be independently chosen. In the above sequence, "*XX*" in italics may be TR, LR or LK, and "[linker]" represents a linker sequence. In some embodiments, the linker sequence is TGSQKP (SEQ ID NO: 651); this linker may be used when sub-sites targeted by the ZFs are adjacent. In some embodiments, the linker sequence is TGGGGSQKP (SEQ ID NO: 652); this linker may be used when there is a base between the sub-sites targeted by the zinc fingers. The two indicated linkers may be the same or different.

[0086] ZFP domains herein may contain arrays of two or more adjacent ZFs that are directly adjacent to one another (e.g., separated by a short (canonical) linker sequence), or are separated by longer, flexible or structured polypeptide sequences. In some embodiments, directly adjacent fingers bind to contiguous nucleic acid sequences, i.e., to adjacent trinucleotides/triplets. In some embodiments, adjacent fingers cross-bind between each other's respective target triplets, which may help to strengthen or enhance the recognition of the target sequence, and leads to the binding of overlapping sequences. In some

embodiments, distant ZFs within the ZFP domain may recognize (or bind to) non-contiguous nucleotide sequences.

[0087] The amino acid sequences of the ZF DNA-recognition helices of exemplary ZFP domains herein, and their *PCSK9* target sequences, are shown below in **Table 1**, where numbers within the parentheses denote SEQ ID NOs:

Table 1. ZF Sequences of Exemplary ZFP Domains Targeting *PCSK9*

ZF No.	DNA Target	F1	F2	F3	F4	F5	F6
ZF001	CCGGCAGAACT TGGAGTC (700)	DHSSLKR (748)	QSPHLQR (749)	YKHVLVN (750)	QMSNLDR (751)	QAETLKR (752)	RNYDLGI (753)
ZF002	GGTGGGGAGGA CTGTGCA (701)	QSTTLKR (754)	QAHGLTA (755)	DAANLRR (756)	RRDNLLR (757)	RNTHLAR (758)	EAHHLSR (759)
ZF003	GGTGGGGAGGA CTGTGCA (702)	QSTTLKR (760)	GAHGLAG (761)	DQTNLRR (762)	RQDNLQR (763)	KGDHLRR (764)	EAHHLSR (765)
ZF004	GATGGGGCTCT GGTGGCG (703)	RTDTLAR (766)	RKTALNR (767)	RNESLKV (768)	VKNTLTR (769)	RREHLVR (770)	ISHNLAR (771)
ZF005	GATGGGGCTCT GGTGGCG (704)	RRLTLDR (772)	RRDLLHN (773)	RNESLKV (774)	VGASLKR (775)	RREHLVR (776)	LEHNLGR (777)
ZF006	GATGGGGCTCT GGTGGCG (705)	RNHTLQR (778)	RREVLEN (779)	RHESLIV (780)	VGASLKR (781)	KKDHLHR (782)	LMHNLTR (783)
ZF007	GCAGGAGGATG AAAAGGG (706)	KRERLER (784)	QKGNLVV (785)	QKTHLAV (786)	QGGHLKR (787)	QRPHLTN (788)	QAATLQR (789)
ZF008	GGAGCCGCTGC TGCAACG (707)	RSQTLAQ (790)	QSTTLKR (791)	LRDSLKR (792)	LGHTLNR (793)	DPSVLTR (794)	QNSHLRR (795)
ZF009	GACTGTGCAGG AGCTGAA (708)	QRGNLAR (796)	VSNTLTR (797)	QRPHLTN (798)	QAATLQR (799)	RRQHLEL (800)	DRGNLTR (801)
ZF010	GGGGAGGACTG TGCAGGA (709)	QSAHLKR (802)	QDVSLVR (803)	RKQHLQL (804)	DRGNLTR (805)	RDDNLQR (806)	RMEHLPR (807)
ZF011	GGTGGGGAGGA CTGTGCA (710)	QSTTLKR (808)	QPHGLGA (809)	LVENLRR (810)	RVENLHR (811)	RREHLVR (812)	LSQGLAR (813)
ZF012	GGTGGGGAGGA CTGTGCA (711)	QSTTLKR (814)	QAHGLTA (815)	DGSNLAR (816)	RRDNLLR (817)	RREHLVR (818)	VDHHLRR (819)
ZF013	TGCGGTGGGGA GGACTGT (712)	RKQHLQL (820)	DRGNLTR (821)	RPDNLVR (822)	RIDKLGG (823)	EAHHLSR (824)	QGRSLRA (825)
ZF014	GTGGAAGGTGG CTGTGGT (713)	EAHHLSR (826)	QPHGLRA (827)	LKEHLTR (828)	MGHHLKR (829)	QRGNLLR (830)	RPDALPR (831)
ZF015	GCGGTGGAAGG TGGCTGT (714)	RRQHLQY (832)	LKEHLTR (833)	SKQKLQV (834)	QQTNLTR (835)	RREVLEN (836)	RPDGLAR (837)

ZF No.	DNA Target	F1	F2	F3	F4	F5	F6
ZF016	GCGGTGGAAGG TGGCTGT (715)	RRQHLTL (838)	DNSHLQR (839)	LPHHLQR (840)	QSNNLTR (841)	RNFILQR (842)	RKDDLKR (843)
ZF017	GCGGCGGGGGA GGACGGG (716)	RQEHLVR (844)	EGGNLMR (845)	RPDNLVR (846)	RIDKLGG (847)	RRDDLTR (848)	RKDLLHR (849)
ZF018	GCGGCGGGGGA GGACGGG (717)	RREHLVR (850)	DMGNLGR (851)	RDDNLQR (852)	RMEHLPR (853)	RKEDLAR (854)	RLDMLAR (855)
ZF019	TAGGGATGGGA GGCCGGG (718)	RREHLVR (856)	EHRGLKR (857)	RGDNLGR (858)	RSDHLSL (859)	QQAHLVR (860)	RDDNLRT (861)
ZF020	GGCGCCGCCGG CGTGGAC (719)	DDANLRR (862)	RREVLEN (863)	LKEHLTR (864)	DPSNLRR (865)	DSSVLRR (866)	ENSKLNR (867)
ZF021	GGCGCCGCCGG CGTGGAC (720)	DRANLRR (868)	RPDALSR (869)	LKEHLTR (870)	DPSNLRR (871)	DSSVLRR (872)	ENSKLNR (873)
ZF022	GTCGTTGCAGC AGCGGCT (721)	LSQTLKR (874)	RTDTLAR (875)	QMETLKR (876)	QGGTLRR (877)	HKSSLTR (878)	DRTPLQR (879)
ZF023	GATGGGGCTCT GGTGGCG (722)	RTDTLAR (880)	RKTALNR (881)	RNESLKV (882)	VKNTLTR (883)	RREHLVR (884)	LTHNLRR (885)
ZF024	TCGGTGGGCAG CGAGGAG (723)	RLDNLDR (886)	RADNLRR (887)	SHWKLHT (888)	EKGHLNR (889)	RREVLEN (890)	RADGLQL (891)
ZF025	GGGGAGGACTG TGCAGGA (724)	QSPHLKR (892)	QSTSLQR (893)	RKQHLTL (894)	DRGNLTR (895)	RPHNLLR (896)	RREHLVR (897)
ZF026	GGGGAGGACTG TGCAGGA (725)	QQAHLVR (898)	QAETLKR (899)	RKQHLTL (900)	DRGNLTR (901)	RNTNLTR (902)	RREHLVR (903)
ZF027	GCGGCGGGGGA GGACGGG (726)	RREHLVR (904)	DPSNLQR (905)	RNTNLTR (906)	RREHLVR (907)	RTDTLAR (908)	RVDDLGR (909)
ZF028	GAGGAGTGAGC CAGGCAGT (727)	QSHSLKS (910)	ESGHLKR (911)	DLSTLRR (912)	QNEHLKV (913)	RPENLNR (914)	RRDNLNR (915)
ZF029	GGTGTGGGTGC TTGACGCC (728)	ERRGLDR (916)	DRGNLTR (917)	LSQTLNR (918)	IKHHLGR (919)	RNFILQR (920)	ERHQLVR (921)
ZF030	TCTGAGCCTGG AGGAGTGA (729)	QREHLVT (922)	RIDNLGR (923)	RMSNLVR (924)	RNESLKV (925)	RADNLGR (926)	ARNTLKG (927)
ZF031	GAGGAGTGAGC CAGGCAGT (730)	QSHSLKS (928)	EKSHLTR (929)	DSPTLRR (930)	QKVHLQV (931)	RAENLAR (932)	RRDNLLR (933)
ZF032	GAGGGCCAGGG GAGAGGTT (731)	HKSSLTR (934)	RPDNLPR (935)	QGTHLRN (936)	RTHHLIT (937)	TPSKLDR (938)	RQDNLGR (939)
ZF033	GCAGGAGGACG AGGACGGC (732)	ERAKLIR (940)	DPSNLRR (941)	RQDNLGR (942)	DQGNLGR (943)	QSAHLKR (944)	QDVSLVR (945)
ZF034	GCAGGAGGACG AGGACGGC (733)	TPSKLDR (946)	LAENLRR (947)	RQDNLGR (948)	DGGNLGR (949)	QSPHLKR (950)	QSTSLQR (951)
ZF035	GCAGGAGGACG AGGACGGC (734)	TPSKLDR (952)	DSSNLRR (953)	RQDNLGR (954)	EGGNLMR (955)	QQAHLVR (956)	QAETLKR (957)

ZF No.	DNA Target	F1	F2	F3	F4	F5	F6
ZF036	TGGGGAGGACT GTGCAGGA (735)	QRPHLTN (958)	QAATLQR (959)	RKECLVV (960)	QNPHLLR (961)	QSAHLKR (962)	RSDHLSL (963)
ZF037	GTTGCAGGCGG GCGCCGCC (736)	DPSVLTR (964)	EHRGLKR (965)	VPSKLKR (966)	RDDTLVR (967)	QKETLNR (968)	VRSSLRR (969)
ZF038	GTTGCAGGCGG GCGCCGCC (737)	DPSVLTR (970)	EHRGLKR (971)	KTDHLAR (972)	DKAHLVR (973)	QKETLNR (974)	VRSSLRR (975)
ZF039	GCTGTTTGGGG AGGGCGAG (738)	RADNLGR (976)	KQVTLRN (977)	KHSNLTR (978)	RREHLVR (979)	VKSSLTR (980)	VSNSLNR (981)
ZF040	GCTGTTTGGGG AGGGCGAG (739)	RDDNLQR (982)	KNVTLTN (983)	QSAHLKR (984)	RSDHLSL (985)	VKSSLTR (986)	VSNSLNR (987)
ZF041	GCAGAGGCCGG AGGGGGTC (740)	DPSVLKR (988)	RTEHLAR (989)	QSPHLKR (990)	DQTTLRR (991)	KHSNLTR (992)	QMETLKR (993)
ZF042	TGGGGAGGGCG AGGCCGAA (741)	QNQNLAR (994)	DKSVLAR (995)	RDDNLQR (996)	KNVTLTN (997)	QSAHLKR (998)	RSDHLSL (999)
ZF043	GCGGGTGTAGG GATGGGAG (742)	KHSNLTR (1000)	RREHLTI (1001)	KKDHLHR (1002)	QTTTLKR (1003)	EEHHLTR (1004)	REDVLGR (1005)
ZF044	GCGGGTGTAGG GATGGGAG (743)	KHSNLTR (1006)	RREHLTI (1007)	KKDHLHR (1008)	QTTTLKR (1009)	EAHHLSR (1010)	RTDDLGR (1011)
ZF045	GGAGCTGGGAG CCGCTGCT (744)	LRQTLAR (1012)	VAHSLKR (1013)	DRSVLVR (1014)	QNSHLRR (1015)	VKHSLQR (1016)	QTTHLSR (1017)
ZF046	GCAAGGCGGCG GGGGAGGA (745)	QSAHLKR (1018)	QMSHLKR (1019)	RGNHLRR (1020)	LKEHLTR (1021)	RNEHLKV (1022)	QSTTLKR (1023)
ZF047	GCTGTTTGGGG AGGGCGAGG (746)	RKPHLDN (1024)	RPDVLMR (1025)	RDDNLQR (1026)	RMEHLPR (1027)	HQSSLTR (1028)	VSNSLAR (1029)
ZF048	GCTGTTTGGGG AGGGCGAGG (747)	RNIHLQT (1030)	RKDTLAR (1031)	RADNLGR (1032)	RMEHLPR (1033)	HQSSLTR (1034)	VSNSLAR (1035)

[0088] In some embodiments, the ZFP domain of the present epigenetic editor binds to a target sequence selected from any one of SEQ ID NOs: 700-747. In further embodiments, the ZFP domain comprises, in order, the F1-F6 amino acid sequences of any one of ZF001-ZF048 as shown in **Table 1**. The F1-F6 amino acid sequences may be placed within the ZF framework sequence of SEQ ID NO: 650, or within any other ZF framework known in the art.

C. TALEs

[0089] In some embodiments, the DNA-binding domain of an epigenetic editor described herein comprises a transcription activator-like effector (TALE) domain. The DNA-binding domain of a TALE comprises a highly conserved sequence of about 33-34 amino acids, with

a repeat variable di-residue (RVD) at positions 12 and 13 that is central to the recognition of specific nucleotides. TALEs can be engineered to bind practically any desired DNA sequence. Methods for programming TALEs are known in the art. For example, such methods are described in Carroll et al., *Genet Soc Amer.* (2011) 188(4):773-82; Miller et al., *Nat Biotechnol.* (2007) 25(7):778-85; Christian et al., *Genetics* (2008) 186(2):757-61; Li et al., *Nucl Acids Res.* (2010) 39(1):359-72; and Moscou et al., *Science* (2009) 326(5959):1501.

D. Other DNA-Binding Domains

[0090] Other DNA-binding domains are contemplated for the epigenetic editors described herein. In some embodiments, the DNA-binding domain comprises an argonaute protein domain, e.g., from *Natronobacterium gregoryi* (NgAgo). NgAgo is a ssDNA-guided endonuclease that is guided to its target site by 5' phosphorylated ssDNA (gDNA), where it produces double-strand breaks. In contrast to Cas9, the NgAgo-gDNA system does not require a protospacer-adjacent motif (PAM). Thus, using a nuclease inactive NgAgo (dNgAgo) can greatly expand the bases that may be targeted. The characterization and use of NgAgo have been described, e.g., in Gao et al., *Nat Biotechnol.* (2016) 34(7):768-73; Swarts et al., *Nature* (2014) 507(7491):258-61; and Swarts et al., *Nucl Acids Res.* (2015) 43(10):5120-9.

[0091] In some embodiments, the DNA-binding domain comprises an inactivated nuclease, for example, an inactivated meganuclease. Additional non-limiting examples of DNA-binding domains include tetracycline-controlled repressor (tetR) DNA-binding domains, leucine zippers, helix-loop-helix (HLH) domains, helix-turn-helix domains, β-sheet motifs, steroid receptor motifs, bZIP domains homeodomains, and AT-hooks.

II. Guide Polynucleotides

[0092] Epigenetic editors described herein that comprise a polynucleotide guided DNA-binding domain may also include a guide polynucleotide that is capable of forming a complex with the DNA-binding domain. The guide polynucleotide may comprise RNA, DNA, or a mixture of both. For example, where the polynucleotide guided DNA-binding domain is a CRISPR-associated protein domain, the guide polynucleotide may be a guide RNA (gRNA). A "guide RNA" or "gRNA" refers to a nucleic acid that is able to hybridize to a target sequence and direct binding of the CRISPR-Cas complex to the target sequence. Methods of using guide polynucleotide sequences with programmable DNA-binding proteins (e.g., CRISPR-associated protein domains) for site-specific DNA targeting (e.g., to modify a genome) are known in the art.

[0093] A guide polynucleotide sequence (e.g., a gRNA sequence) may comprises two parts: 1) a nucleotide sequence comprising a "targeting sequence" that is complementary to a target nucleic acid sequence ("target sequence"), e.g., to a nucleic acid sequence comprised in a genomic target site; and 2) a nucleotide sequence that binds a polynucleotide guided DNA-binding domain (e.g., a CRISPR-Cas protein domain). The nucleotide sequence in 1) may comprise a targeting sequence that is 100% complementary to a genomic nucleic acid sequence, e.g., a nucleic acid sequence comprised in a genomic target site, and thus may hybridize to the target nucleic acid sequence. The nucleotide sequence in 1) may be referred to as, e.g., a crispr RNA, or crRNA. The nucleotide sequence in 2) may be referred to as a scaffold sequence of a guide nucleic acid, e.g., a tracrRNA, or an activating region of a guide nucleic acid, and may comprise a stem-loop structure. Parts 1) and 2) as described above may be fused to form one single guide (e.g., a single guide RNA, or sgRNA), or may be on two separate nucleic acid molecules. In some embodiments, a guide polynucleotide comprises parts 1) and 2) connected by a linker. In some embodiments, a guide polynucleotide comprises parts 1) and 2) connected by a non-nucleic acid linker, for example, a peptide linker or a chemical linker.

[0094] Part 2 (the scaffold sequence) of a guide polynucleotide as described herein may be, for example, as described in Jinek et al., *Science* (2012) 337:816-21; U.S. Patent Publication 2016/0208288; or U.S. Patent Publication 2016/0200779. Variants of part 2) are also contemplated by the present disclosure. For example, the tetraloop and stem loop of a gRNA scaffold (tracrRNA) sequence may be modified to include RNA aptamers, which can be bound by specific protein domains. In some embodiments, such modified gRNAs can be used to facilitate the recruitment of repressive or activating domains fused to the protein-interacting RNA aptamers.

[0095] A gRNA as provided herein typically comprises a targeting domain and a binding domain. The targeting domain (also termed "targeting sequence") may comprise a nucleic acid sequence that binds to a target site, e.g., to a genomic nucleic acid molecule within a cell. The target site may be a double-stranded DNA sequence comprising a PAM sequence as well as the target sequence, which is located on the same strand as, and directly adjacent to, the PAM sequence. The targeting domain of the gRNA may comprise an RNA sequence that corresponds to the target sequence, i.e., it resembles the sequence of the target domain, sometimes with one or more mismatches, but typically comprising an RNA sequence instead of a DNA sequence. The targeting domain of the gRNA thus may base pair (in full or partial complementarity) with the sequence of the double-stranded target site that is complementary

to the target sequence, and thus with the strand complementary to the strand that comprises the PAM sequence. It will be understood that the targeting domain of the gRNA typically does not include a sequence that resembles the PAM sequence. It will further be understood that the location of the PAM may be 5' or 3' of the target sequence, depending on the nuclease employed. For example, the PAM is typically 3' of the target sequence for Cas9 nucleases, and 5' of the target sequence for Cas12a nucleases. For an illustration of the location of the PAM and the mechanism of gRNA binding to a target site, see, e.g., Figure 1 of Vanegas et al., *Fungal Biol Biotechnol.* (2019) 6:6, which is incorporated by reference herein. For additional illustration and description of the mechanism of gRNA targeting of an RNA-guided nuclease to a target site, see Fu et al., *Nat Biotechnol* (2014) 32(3):279-84 and Sternberg et al., *Nature* (2014) 507(7490):62-7, each incorporated herein by reference.

[0096] In some embodiments, the targeting domain sequence comprises between 17 and 30 nucleotides and corresponds fully to the target sequence (i.e., without any mismatch nucleotides). In some embodiments, however, the targeting domain sequence may comprise one or more, but typically not more than 4, mismatches, e.g., 1, 2, 3, or 4 mismatches. As the targeting domain is part of gRNA, which is an RNA molecule, it will typically comprise ribonucleotides, while the DNA targeting domain will comprise deoxyribonucleotides.

[0097] An exemplary illustration of a Cas9 target site, comprising a 22 nucleotide target domain, and an NGG PAM sequence, as well as of a gRNA comprising a targeting domain that fully corresponds to the target sequence (and thus base pairs with full complementarity with the DNA strand complementary to the strand comprising the target sequence and PAM) is provided below:

```
[                  target domain (DNA)          ][ PAM  ]
5'-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-G-G-3' (DNA)
3'-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-C-C-5' (DNA)
   | | | | | | | | | | | | | | | | | | | | | | | |
5'-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-[ gRNA scaffold]-3' (RNA)
[             targeting domain ( RNA)           ][  binding domain  ]
```

[0098] An exemplary illustration of a Casl2a target site, comprising a 22 nucleotide target domain, and a TTN PAM sequence, as well as of a gRNA comprising a targeting domain that fully corresponds to the target sequence (and thus base pairs with full complementarity with the DNA strand complementary to the strand comprising the target sequence and PAM) is provided below:

```
          [  PAM  ][            target domain ( DNA)            ]
          5'-T-T-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-3' (DNA)
          3'-A-A-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-5' (DNA)
                    | | | | | | | | | | | | | | | | | | | | | | | |
5'-[gRNA scaffold]-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-3' (RNA)
[ binding domain  ][             targeting domain ( RNA)          ]
```

[0099] While not wishing to be bound by theory, at least in some embodiments, it is believed that the length and complementarity of the targeting domain with the target sequence contributes to specificity of the interaction of the gRNA/Cas9 molecule complex with a target nucleic acid. In some embodiments, the targeting domain of a gRNA provided herein is 5 to 50 nucleotides in length. In some embodiments, the targeting domain is 15 to 25 nucleotides in length. In some embodiments, the targeting domain is 18 to 22 nucleotides in length. In some embodiments, the targeting domain is 19-21 nucleotides in length. In some embodiments, the targeting domain is 15 nucleotides in length. In some embodiments, the targeting domain is 16 nucleotides in length. In some embodiments, the targeting domain is 17 nucleotides in length. In some embodiments, the targeting domain is 18 nucleotides in length. In some embodiments, the targeting domain is 19 nucleotides in length. In some embodiments, the targeting domain is 20 nucleotides in length. In some embodiments, the targeting domain is 21 nucleotides in length. In some embodiments, the targeting domain is 22 nucleotides in length. In some embodiments, the targeting domain is 23 nucleotides in length. In some embodiments, the targeting domain is 24 nucleotides in length. In some embodiments, the targeting domain is 25 nucleotides in length. In certain embodiments, the targeting domain fully corresponds, without mismatch, to a target sequence provided herein, or a part thereof. In some embodiments, the targeting domain of a gRNA provided herein comprises 1 mismatch relative to a target sequence provided herein. In some embodiments, the targeting domain comprises 2 mismatches relative to the target sequence. In some embodiments, the target domain comprises 3 mismatches relative to the target sequence.

[0100] Methods for designing, selecting, and validating gRNAs are described herein and known in the art. Software tools can be used to optimize the gRNAs corresponding to a target DNA sequence, e.g., to minimize total off-target activity across the genome. For example, DNA sequence searching algorithms can be used to identify a target sequence in crRNAs of a gRNA for use with Cas9. Exemplary gRNA design tools include the ones described in Bae et al., *Bioinformatics* (2014) 30:1473-5.

[0101] Guide polynucleotides (e.g., gRNAs) described herein may be of various lengths. In some embodiments, the length of the spacer or targeting sequence depends on the

CRISPR-associated protein component of the epigenetic editor system used. For example, Cas proteins from different bacterial species have varying optimal targeting sequence lengths. Accordingly, the spacer sequence may comprise, e.g., 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or more than 50 nucleotides in length. In some embodiments, the spacer comprises 10-24, 11-20, 11-16, 18-24, 19-21, or 20 nucleotides in length. In some embodiments, a guide polynucleotide (e.g., gRNA) is from 15-100 (e.g., 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50) nucleotides in length and comprises a spacer sequence of at least 10 (e.g., 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50) contiguous nucleotides complementary to the target sequence. In some embodiments, a guide polynucleotide described herein may be truncated, e.g., by 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50 or more nucleotides.

[0102] In certain embodiments, the 3' end of the *PCSK9* target sequence is immediately adjacent to a PAM sequence (e.g., a canonical PAM sequence such as NGG for SpCas9). The degree of complementarity between the targeting sequence of the guide polynucleotide (e.g., the spacer sequence of a gRNA) and the target sequence may be at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100%. In particular embodiments, the targeting and the target sequence may be 100% complementary. In other embodiments, the targeting sequence and the target sequence may contain, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 mismatches.

[0103] A guide polynucleotide (e.g., gRNA) may be modified with, for example, chemical alterations and synthetic modifications. A modified gRNA, for instance, can include an alteration or replacement of one or both of the non-linking phosphate oxygens and/or of one or more of the linking phosphate oxygens in the phosphodiester backbone linkage, an alteration of the ribose sugar (e.g., of the 2' hydroxyl on the ribose sugar), an alteration of the phosphate moiety, modification or replacement of a naturally occurring nucleobase, modification or replacement of the ribose-phosphate backbone, modification of the 3' end and/or 5' end of the oligonucleotide, replacement of a terminal phosphate group or conjugation of a moiety, cap, or linker, or any combination thereof.

[0104] In some embodiments, one or more ribose groups of the gRNA may be modified. Examples of chemical modifications to the ribose group include, but are not limited to, 2'-O-methyl (2'-OMe), 2'-fluoro (2'-F), 2'-deoxy, 2'-O-(2-methoxyethyl) (2'-MOE), 2'-NH2, 2'-

O-allyl, 2'-O-ethylamine, 2'-O-cyanoethyl, 2'-O-acetalester, or a bicyclic nucleotide such as locked nucleic acid (LNA), 2'-(5-constrained ethyl (S-cEt)), constrained MOE, or 2'-0,4'-C-aminomethylene bridged nucleic acid (2',4'-BNANC). 2'-O-methyl modification and/or 2'-fluoro modification may increase binding affinity and/or nuclease stability of the gRNA oligonucleotides.

[0105] In some embodiments, one or more phosphate groups of the gRNA may be chemically modified. Examples of chemical modifications to a phosphate group include, but are not limited to, a phosphorothioate (PS), phosphonoacetate (PACE), thiophosphonoacetate (thioPACE), amide, triazole, phosphonate, and phosphotriester modification. In some embodiments, a guide polynucleotide described herein may comprise one, two, three, or more PS linkages at or near the 5' end and/or the 3' end; the PS linkages may be contiguous or noncontiguous.

[0106] In some embodiments, the gRNA herein comprises a mixture of ribonucleotides and deoxyribonucleotides and/or one or more PS linkages.

[0107] In some embodiments, one or more nucleobases of the gRNA may be chemically modified. Examples of chemically modified nucleobases include, but are not limited to, 2-thiouridine, 4-thiouridine, N6-methyladenosine, pseudouridine, 2,6-diaminopurine, inosine, thymidine, 5-methylcytosine, 5-substituted pyrimidine, isoguanine, isocytosine, and nucleobases with halogenated aromatic groups. Chemical modifications can be made in the spacer region, the tracr RNA region, the stem loop, or any combination thereof.

[0108] **Table 2** below lists exemplary gRNA target sequences for epigenetic modification of human *PCSK9*, as well as the coordinates of the start and end positions of the targeted site on human chromosome 1 (SEQ: SEQ ID NO). The Table also shows the distance from the start coordinate to the TSS coordinate of the *PCSK9* gene.

Table 2. Exemplary Target Sequences of gRNAs Targeting *PCSK9*

gRNA No.	Chr. 1 Strand	START	END	gRNA Target Sequence (DNA, 5' to 3')	SEQ	TSS Distance
gRNA001	+	55039960	55039979	GGTGCTAGCCTTGCGTTCCG	1036	431
gRNA002	+	55039991	55040010	CTGGCCGAAGCACCCGAGCA	1037	462
gRNA003	-	55039529	55039548	TGCGGAAACCTTCTAGGGTG	1038	0
gRNA004	-	55039528	55039547	GCGGAAACCTTCTAGGGTGT	1039	-1
gRNA005	+	55039518	55039537	TCAAGCACCCACACCCTAGA	1040	-11
gRNA006	-	55039514	55039533	GGGTGTGGGTGCTTGACGCC	1041	-15
gRNA007	-	55039513	55039532	GGTGTGGGTGCTTGACGCCT	1042	-16
gRNA008	-	55039512	55039531	GTGTGGGTGCTTGACGCCTG	1043	-17
gRNA009	-	55039610	55039629	ACTGCCTGGCTCACTCCTCC	1044	81
gRNA010	-	55039481	55039500	TCACGCCACCAGAGCCCCAT	1045	-48
gRNA011	+	55039473	55039492	ATCGTCCGATGGGGCTCTGG	1046	-56

gRNA No.	Chr. 1 Strand	START	END	gRNA Target Sequence (DNA, 5' to 3')	SEQ	TSS Distance
gRNA012	+	55039470	55039489	AGGATCGTCCGATGGGGCTC	1047	-59
gRNA013	+	55039464	55039483	TCAGATAGGATCGTCCGATG	1048	-65
gRNA014	+	55039669	55039688	GCGGCTCCCAGCTCCCAGCC	1049	140
gRNA015	-	55039678	55039697	CGGAATCCTGGCTGGGAGCT	1050	149
gRNA016	-	55039679	55039698	GCGGAATCCTGGCTGGGAGC	1051	150
gRNA017	-	55039685	55039704	GGGCGCGCGGAATCCTGGCT	1052	156
gRNA018	-	55039686	55039705	GGGGCGCGCGGAATCCTGGC	1053	157
gRNA019	+	55038610	55038629	TGAGGTCTTTGCAAACAAAG	1054	-919
gRNA020	+	55038654	55038673	CCAGCACCTAGATTCAGAGC	1055	-875
gRNA021	-	55038657	55038676	CCTGCTCTGAATCTAGGTGC	1056	-872
gRNA022	-	55038731	55038750	TCGATACTGGGAAGAAACAA	1057	-798
gRNA023	-	55038743	55038762	CTGGAAGGGCTGTCGATACT	1058	-786
gRNA024	-	55038744	55038763	TCTGGAAGGGCTGTCGATAC	1059	-785
gRNA025	-	55038757	55038776	GAGGCTTGCTCTTTCTGGAA	1060	-772
gRNA026	-	55038758	55038777	TGAGGCTTGCTCTTTCTGGA	1061	-771
gRNA027	-	55038762	55038781	GACATGAGGCTTGCTCTTTC	1062	-767
gRNA028	-	55038776	55038795	TGTACATGTGGCATGACATG	1063	-753
gRNA029	+	55038781	55038800	CATGCCACATGTACAATCTG	1064	-748
gRNA030	+	55038786	55038805	CACATGTACAATCTGAGGCC	1065	-743
gRNA031	-	55038788	55038807	CTGGCCTCAGATTGTACATG	1066	-741
gRNA032	-	55038807	55038826	AAAAGGGGAAAGAGAGCTCC	1067	-722
gRNA033	-	55038822	55038841	CCAGGCAGGAGGATGAAAAG	1068	-707
gRNA034	-	55038823	55038842	ACCAGGCAGGAGGATGAAAA	1069	-706
gRNA035	-	55038824	55038843	TACCAGGCAGGAGGATGAAA	1070	-705
gRNA036	+	55038830	55038849	CCTCCTGCCTGGTACACAAT	1071	-699
gRNA037	-	55038833	55038852	CCTATTGTGTACCAGGCAGG	1072	-696
gRNA038	-	55038836	55038855	ACACCTATTGTGTACCAGGC	1073	-693
gRNA039	+	55038840	55038859	GGTACACAATAGGTGTTTAC	1074	-689
gRNA040	-	55038840	55038859	GTAAACACCTATTGTGTACC	1075	-689
gRNA041	+	55038870	55038889	TCCAGTTGATTTCTTGAACA	1076	-659
gRNA042	-	55038874	55038893	ACCATGTTCAAGAAATCAAC	1077	-655
gRNA043	+	55038881	55038900	TCTTGAACATGGTGTGTAAA	1078	-648
gRNA044	+	55038906	55038925	TCTTTGCAAATTGAATCTTC	1079	-623
gRNA045	-	55038970	55038989	AGTTTGCAAAGACGTCATAT	1080	-559
gRNA046	-	55038998	55039017	GATTTATACTACAAAGATTC	1081	-531
gRNA047	+	55039075	55039094	AGTTGGTAAGGTCAGTGTGC	1082	-454
gRNA048	+	55039076	55039095	GTTGGTAAGGTCAGTGTGCA	1083	-453
gRNA049	+	55039093	55039112	GCAGGGTGCATAAAGGGCAG	1084	-436
gRNA050	+	55039097	55039116	GGTGCATAAAGGGCAGAGGC	1085	-432
gRNA051	+	55039100	55039119	GCATAAAGGGCAGAGGCCGG	1086	-429
gRNA052	+	55039137	55039156	TTTAGAAGGCTGCCAGGTTA	1087	-392
gRNA053	-	55039163	55039182	GCCCACCGAATTCTTTCCAC	1088	-366
gRNA054	+	55039167	55039186	AAAGAATTCGGTGGGCAGCG	1089	-362
gRNA055	-	55039196	55039215	CTTCTGAATCAATCCTACTG	1090	-333
gRNA056	+	55039221	55039240	CTGGTCAGCAGGAGACAAGG	1091	-308
gRNA057	+	55039202	55039221	GATTGATTCAGAAGTCTCAC	1092	-327
gRNA058	+	55039210	55039229	CAGAAGTCTCACTGGTCAGC	1093	-319
gRNA059	+	55039218	55039237	TCACTGGTCAGCAGGAGACA	1094	-311
gRNA060	+	55039228	55039247	GCAGGAGACAAGGTGGACCC	1095	-301
gRNA061	+	55039242	55039261	GGACCCAGGAAACACTGAAA	1096	-287
gRNA062	+	55039245	55039264	CCCAGGAAACACTGAAAAGG	1097	-284
gRNA063	+	55039246	55039265	CCAGGAAACACTGAAAAGGT	1098	-283

gRNA No.	Chr. 1 Strand	START	END	gRNA Target Sequence (DNA, 5' to 3')	SEQ	TSS Distance
gRNA064	+	55039251	55039270	AAACACTGAAAAGGTGGGCC	1099	-278
gRNA065	+	55039281	55039300	TGGAGTCTGGCATCCCACGC	1100	-248
gRNA066	+	55039282	55039301	GGAGTCTGGCATCCCACGCA	1101	-247
gRNA067	+	55039312	55039331	CGGGAGAGGAGGAGCCCCTA	1102	-217
gRNA068	+	55039318	55039337	AGGAGGAGCCCCTAGGGCGC	1103	-211
gRNA069	-	55039329	55039348	AAGGCAGGCCGGCGCCCTAG	1104	-200
gRNA070	-	55039330	55039349	GAAGGCAGGCCGGCGCCCTA	1105	-199
gRNA071	-	55039331	55039350	GGAAGGCAGGCCGGCGCCCT	1106	-198
gRNA072	-	55039340	55039359	AACTGGGCTGGAAGGCAGGC	1107	-189
gRNA073	+	55039340	55039359	GCCTGCCTTCCAGCCCAGTT	1108	-189
gRNA074	-	55039344	55039363	TCCTAACTGGGCTGGAAGGC	1109	-185
gRNA075	+	55039346	55039365	CTTCCAGCCCAGTTAGGATT	1110	-183
gRNA076	+	55039347	55039366	TTCCAGCCCAGTTAGGATTT	1111	-182
gRNA077	-	55039348	55039367	CAAATCCTAACTGGGCTGGA	1112	-181
gRNA078	-	55039352	55039371	CTCCCAAATCCTAACTGGGC	1113	-177
gRNA079	-	55039356	55039375	AAAACTCCCAAATCCTAACT	1114	-173
gRNA080	-	55039357	55039376	AAAAACTCCCAAATCCTAAC	1115	-172
gRNA081	-	55039384	55039403	AGCGTCAGATTACGCGCAGA	1116	-145
gRNA082	-	55039385	55039404	CAGCGTCAGATTACGCGCAG	1117	-144
gRNA083	+	55039387	55039406	GCGCGTAATCTGACGCTGTT	1118	-142
gRNA084	+	55039388	55039407	CGCGTAATCTGACGCTGTTT	1119	-141
gRNA085	+	55039389	55039408	GCGTAATCTGACGCTGTTTG	1120	-140
gRNA086	+	55039392	55039411	TAATCTGACGCTGTTTGGGG	1121	-137
gRNA087	+	55039393	55039412	AATCTGACGCTGTTTGGGGA	1122	-136
gRNA088	+	55039398	55039417	GACGCTGTTTGGGGAGGGCG	1123	-131
gRNA089	+	55039422	55039441	CGAAACCTGATCCTCCAGTC	1124	-107
gRNA090	+	55039423	55039442	GAAACCTGATCCTCCAGTCC	1125	-106
gRNA091	+	55039424	55039443	AAACCTGATCCTCCAGTCCG	1126	-105
gRNA092	-	55039424	55039443	CGGACTGGAGGATCAGGTTT	1127	-105
gRNA093	+	55039425	55039444	AACCTGATCCTCCAGTCCGG	1128	-104
gRNA094	-	55039430	55039449	AACCCCCGGACTGGAGGATC	1129	-99
gRNA095	-	55039436	55039455	TAACGGAACCCCCGGACTGG	1130	-93
gRNA096	-	55039439	55039458	CATTAACGGAACCCCCGGAC	1131	-90
gRNA097	-	55039444	55039463	TTAAACATTAACGGAACCCC	1132	-85
gRNA098	+	55039450	55039469	CCGTTAATGTTTAATCAGAT	1133	-79
gRNA099	-	55039453	55039472	CCTATCTGATTAAACATTAA	1134	-76
gRNA100	+	55039462	55039481	AATCAGATAGGATCGTCCGA	1135	-67
gRNA101	+	55039463	55039482	ATCAGATAGGATCGTCCGAT	1136	-66
gRNA102	+	55039493	55039512	TGGCGTGATCTGCGCGCCCC	1137	-36
gRNA103	-	55039534	55039553	GTCGCTGCGGAAACCTTCTA	1138	5
gRNA104	-	55039535	55039554	CGTCGCTGCGGAAACCTTCT	1139	6
gRNA105	+	55039538	55039557	AGGTTTCCGCAGCGACGTCG	1140	9
gRNA106	+	55039547	55039566	CAGCGACGTCGAGGCGCTCA	1141	18
gRNA107	-	55039547	55039566	TGAGCGCCTCGACGTCGCTG	1142	18
gRNA108	+	55039554	55039573	GTCGAGGCGCTCATGGTTGC	1143	25
gRNA109	+	55039557	55039576	GAGGCGCTCATGGTTGCAGG	1144	28
gRNA110	+	55039558	55039577	AGGCGCTCATGGTTGCAGGC	1145	29
gRNA111	+	55039592	55039611	AGTTCAGGGTCTGAGCCTGG	1146	63
gRNA112	+	55039577	55039596	CGGGCGCCGCCGTTCAGTTC	1147	48
gRNA113	+	55039578	55039597	GGGCGCCGCCGTTCAGTTCA	1148	49
gRNA114	-	55039586	55039605	TCAGACCCTGAACTGAACGG	1149	57
gRNA115	-	55039589	55039608	GGCTCAGACCCTGAACTGAA	1150	60

gRNA No.	Chr. 1 Strand	START	END	gRNA Target Sequence (DNA, 5' to 3')	SEQ	TSS Distance
gRNA116	+	55039589	55039608	TTCAGTTCAGGGTCTGAGCC	1151	60
gRNA117	+	55039603	55039622	TGAGCCTGGAGGAGTGAGCC	1152	74
gRNA118	+	55039615	55039634	AGTGAGCCAGGCAGTGAGAC	1153	86
gRNA119	+	55039620	55039639	GCCAGGCAGTGAGACTGGCT	1154	91
gRNA120	+	55039621	55039640	CCAGGCAGTGAGACTGGCTC	1155	92
gRNA121	-	55039624	55039643	CCCGAGCCAGTCTCACTGCC	1156	95
gRNA122	+	55039624	55039643	GGCAGTGAGACTGGCTCGGG	1157	95
gRNA123	+	55039625	55039644	GCAGTGAGACTGGCTCGGGC	1158	96
gRNA124	+	55039629	55039648	TGAGACTGGCTCGGGCGGGC	1159	100
gRNA125	+	55039630	55039649	GAGACTGGCTCGGGCGGGCC	1160	101
gRNA126	+	55039650	55039669	GGGACGCGTCGTTGCAGCAG	1161	121
gRNA127	-	55039651	55039670	GCTGCTGCAACGACGCGTCC	1162	122
gRNA128	-	55039690	55039709	TGAAGGGGCGCGCGGAATCC	1163	161
gRNA129	-	55039698	55039717	AGGGCGCGTGAAGGGGCGCG	1164	169
gRNA130	-	55039705	55039724	CAGGAGCAGGGCGCGTGAAG	1165	176
gRNA131	-	55039706	55039725	TCAGGAGCAGGGCGCGTGAA	1166	177
gRNA132	-	55039707	55039726	TTCAGGAGCAGGGCGCGTGA	1167	178
gRNA133	-	55039717	55039736	GGAGCTGAAGTTCAGGAGCA	1168	188
gRNA134	-	55039718	55039737	AGGAGCTGAAGTTCAGGAGC	1169	189
gRNA135	-	55039724	55039743	CTGTGCAGGAGCTGAAGTTC	1170	195
gRNA136	+	55039738	55039757	GCACAGTCCTCCCCACCGCA	1171	209
gRNA137	-	55039738	55039757	TGCGGTGGGGAGGACTGTGC	1172	209
gRNA138	+	55039745	55039764	CCTCCCCACCGCAAGGCTCA	1173	216
gRNA139	-	55039748	55039767	CCTTGAGCCTTGCGGTGGGG	1174	219
gRNA140	-	55039751	55039770	GCGCCTTGAGCCTTGCGGTG	1175	222
gRNA141	-	55039752	55039771	GGCGCCTTGAGCCTTGCGGT	1176	223
gRNA142	-	55039753	55039772	CGGCGCCTTGAGCCTTGCGG	1177	224
gRNA143	+	55039754	55039773	CGCAAGGCTCAAGGCGCCGC	1178	225
gRNA144	+	55039759	55039778	GGCTCAAGGCGCCGCCGGCG	1179	230
gRNA145	-	55039776	55039795	AGGCCGTGCGCGGTCCACGC	1180	247
gRNA146	+	55039778	55039797	GTGGACCGCGCACGGCCTCT	1181	249
gRNA147	-	55039786	55039805	GGAGACCTAGAGGCCGTGCG	1182	257
gRNA148	+	55039810	55039829	CAGGACAGCAACCTCTCCCC	1183	281
gRNA149	+	55039839	55039858	TGGGCACCGTCAGCTCCAGG	1184	310
gRNA150	+	55039845	55039864	CCGTCAGCTCCAGGCGGTCC	1185	316
gRNA151	-	55039848	55039867	CCAGGACCGCCTGGAGCTGA	1186	319
gRNA152	+	55039848	55039867	TCAGCTCCAGGCGGTCCTGG	1187	319
gRNA153	-	55039857	55039876	CAGCGGCCACCAGGACCGCC	1188	328
gRNA154	+	55039895	55039914	CTGCTGCTCCTGGGTCCCGC	1189	366
gRNA155	-	55039906	55039925	CACGGGCGCCCGCGGGACCC	1190	377
gRNA156	+	55039909	55039928	TCCCGCGGGCGCCCGTGCGC	1191	380
gRNA157	+	55039912	55039931	CGCGGGCGCCCGTGCGCAGG	1192	383
gRNA158	-	55039913	55039932	TCCTGCGCACGGGCGCCCGC	1193	384
gRNA159	-	55039914	55039933	CTCCTGCGCACGGGCGCCCG	1194	385
gRNA160	+	55039918	55039937	CGCCCGTGCGCAGGAGGACG	1195	389
gRNA161	+	55039922	55039941	CGTGCGCAGGAGGACGAGGA	1196	393
gRNA162	-	55039923	55039942	GTCCTCGTCCTCCTGCGCAC	1197	394
gRNA163	-	55039924	55039943	CGTCCTCGTCCTCCTGCGCA	1198	395
gRNA164	+	55039933	55039952	GGACGAGGACGGCGACTACG	1199	404
gRNA165	+	55039939	55039958	GGACGGCGACTACGAGGAGC	1200	410
gRNA166	+	55039963	55039982	GCTAGCCTTGCGTTCCGAGG	1201	434
gRNA167	+	55039967	55039986	GCCTTGCGTTCCGAGGAGGA	1202	438

gRNA No.	Chr. 1 Strand	START	END	gRNA Target Sequence (DNA, 5' to 3')	SEQ	TSS Distance
gRNA168	-	55039971	55039990	GCCGTCCTCCTCGGAACGCA	1203	442
gRNA169	+	55039972	55039991	GCGTTCCGAGGAGGACGGCC	1204	443
gRNA170	-	55039980	55039999	TTCGGCCAGGCCGTCCTCCT	1205	451
gRNA171	-	55039993	55040012	CGTGCTCGGGTGCTTCGGCC	1206	464
gRNA172	-	55040006	55040025	GTGGCTGTGGTTCCGTGCTC	1207	477
gRNA173	-	55040007	55040026	GGTGGCTGTGGTTCCGTGCT	1208	478
gRNA174	-	55040019	55040038	GCAGCGGTGGAAGGTGGCTG	1209	490
gRNA175	+	55040023	55040042	CACCTTCCACCGCTGCGCCA	1210	494
gRNA176	-	55040025	55040044	CTTGGCGCAGCGGTGGAAGG	1211	496
gRNA177	-	55040028	55040047	CACCTTGGCGCAGCGGTGGA	1212	499
gRNA178	+	55040028	55040047	TCCACCGCTGCGCCAAGGTG	1213	499
gRNA179	+	55040029	55040048	CCACCGCTGCGCCAAGGTGC	1214	500
gRNA180	-	55040032	55040051	CCCGCACCTTGGCGCAGCGG	1215	503
gRNA181	+	55040070	55040089	GGGCGAACCCGCAGCCGGGA	1216	541
gRNA182	-	55040070	55040089	TCCCGGCTGCGGGTTCGCCC	1217	541
gRNA183	-	55040080	55040099	CACCGCACCGTCCCGGCTGC	1218	551
gRNA184	-	55040081	55040100	GCACCGCACCGTCCCGGCTG	1219	552
gRNA185	-	55040087	55040106	GAAACAGCACCGCACCGTCC	1220	558
gRNA186	+	55040091	55040110	GGTGCGGTGCTGTTTCCTCT	1221	562
gRNA187	+	55040092	55040111	GTGCGGTGCTGTTTCCTCTC	1222	563
gRNA188	-	55040109	55040128	GGGGGAAACTGAGGCCCGAG	1223	580
gRNA189	+	55040119	55040138	AGTTTCCCCCCATGTAAGAG	1224	590
gRNA190	+	55040125	55040144	CCCCCATGTAAGAGAGGAAG	1225	596
gRNA191	-	55040127	55040146	CACTTCCTCTCTTACATGGG	1226	598
gRNA192	-	55040128	55040147	CCACTTCCTCTCTTACATGG	1227	599
gRNA193	+	55040133	55040152	TAAGAGAGGAAGTGGAGTGC	1228	604
gRNA194	+	55040155	55040174	GTCGCCGAGGGCTCTTCGCT	1229	626
gRNA195	-	55040162	55040181	CGTGCCAAGCGAAGAGCCCT	1230	633
gRNA196	+	55040166	55040185	CTCTTCGCTTGGCACGATCT	1231	637
gRNA197	+	55040167	55040186	TCTTCGCTTGGCACGATCTT	1232	638
gRNA198	+	55040168	55040187	CTTCGCTTGGCACGATCTTG	1233	639
gRNA199	+	55040176	55040195	GGCACGATCTTGGGGACTGC	1234	647
gRNA200	+	55040181	55040200	GATCTTGGGGACTGCAGGCA	1235	652
gRNA201	+	55040184	55040203	CTTGGGGACTGCAGGCAAGG	1236	655
gRNA202	+	55040189	55040208	GGACTGCAGGCAAGGCGGCG	1237	660
gRNA203	+	55040190	55040209	GACTGCAGGCAAGGCGGCGG	1238	661
gRNA204	+	55040193	55040212	TGCAGGCAAGGCGGCGGGGG	1239	664
gRNA205	+	55040197	55040216	GGCAAGGCGGCGGGGGAGGA	1240	668
gRNA206	+	55040198	55040217	GCAAGGCGGCGGGGGAGGAC	1241	669
gRNA207	+	55040223	55040242	GTGGGGAGCACGGTGGAGAG	1242	694
gRNA208	+	55040224	55040243	TGGGGAGCACGGTGGAGAGC	1243	695
gRNA209	+	55040225	55040244	GGGGAGCACGGTGGAGAGCG	1244	696
gRNA210	+	55040229	55040248	AGCACGGTGGAGAGCGGGGA	1245	700
gRNA211	+	55040233	55040252	CGGTGGAGAGCGGGGACGGC	1246	704
gRNA212	+	55040277	55040296	CGTGCGGCTGCGCTATTCAG	1247	748
gRNA213	+	55040278	55040297	GTGCGGCTGCGCTATTCAGT	1248	749
gRNA214	+	55040282	55040301	GGCTGCGCTATTCAGTGGGA	1249	753
gRNA215	+	55040289	55040308	CTATTCAGTGGGAAGGTTCG	1250	760
gRNA216	+	55040290	55040309	TATTCAGTGGGAAGGTTCGC	1251	761
gRNA217	+	55040291	55040310	ATTCAGTGGGAAGGTTCGCG	1252	762
gRNA218	+	55040295	55040314	AGTGGGAAGGTTCGCGGGGT	1253	766
gRNA219	+	55040296	55040315	GTGGGAAGGTTCGCGGGGTT	1254	767

gRNA No.	Chr. 1 Strand	START	END	gRNA Target Sequence (DNA, 5' to 3')	SEQ	TSS Distance
gRNA220	+	55040337	55040356	AGGGCGAGCAGAGCACTGCC	1255	808
gRNA221	+	55040382	55040401	TTTCTGCCTCGCCGCGGCAC	1256	853
gRNA222	+	55040385	55040404	CTGCCTCGCCGCGGCACAGG	1257	856
gRNA223	+	55040386	55040405	TGCCTCGCCGCGGCACAGGT	1258	857
gRNA224	-	55040391	55040410	CACCCACCTGTGCCGCGGCG	1259	862
gRNA225	-	55040396	55040415	TCCTTCACCCACCTGTGCCG	1260	867
gRNA226	+	55040405	55040424	TGGGTGAAGGAGTGAATGCC	1261	876

[0109] In some embodiments, the gRNA herein does not comprise the sequence CCCGCACCUUGGCGCAGCGG (SEQ ID NO: 1490).

[0110] Any tracr sequence known in the art is contemplated for a gRNA described herein. In some embodiments, a gRNA described herein has a tracr sequence shown in **Table 3** below, or a tracr sequence at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical to the tracr sequence shown below (SEQ: SEQ ID NO).

Table 3. Exemplary TRACR Sequences

SEQ	Sequence (5' to 3')
653	GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACU UGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU
654	GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG GCACCGAGUCGGUGCUUUU
655	GUUUAAGAGCUAAGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACU UGAAAAAGUGGCACCGAGUCGGUGCUUUUUU
656	GUUUAAGAGCUAAGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACU UGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU

[0111] In some embodiments, the gRNA herein is provided to the cell directly (e.g., through an RNP complex together with the CRISPR-associated protein domain). In some embodiments, the gRNA is provided to the cell through an expression vector (e.g., a plasmid vector or a viral vector) introduced into the cell, where the cell then expresses the gRNA from the expression vector. Methods of introducing gRNAs and expression vectors into cells are well known in the art.

III. Effector Domains

[0112] Epigenetic editors described herein include one or more effector protein domains (also "epigenetic effector domains," or "effector domains," as used herein) that effect epigenetic modification of a target gene. An epigenetic editor with one or more effector domains may modulate expression of a target gene without altering its nucleobase sequence. In some embodiments, an effector domain described herein may provide repression or

silencing of expression of a target gene such as *PCSK9*, e.g., by repressing transcription or by modifying or remodeling chromatin. Such effector domains are also referred to herein as "repression domains," "repressor domains," or "epigenetic repressor domains." Non-limiting examples of chemical modifications that may be mediated by effector domains include methylation, demethylation, acetylation, deacetylation, phosphorylation, SUMOylation and/or ubiquitination of DNA or histone residues.

[0113] In some embodiments, an effector domain of an epigenetic editor described herein may make histone tail modifications, e.g., by adding or removing active marks on histone tails.

[0114] In some embodiments, an effector domain of an epigenetic editor described herein may comprise or recruit a transcription-related protein, e.g., a transcription repressor. The transcription-related protein may be endogenous or exogenous.

[0115] In some embodiments, an effector domain of an epigenetic editor described herein may, for example, comprise a protein that directly or indirectly blocks access of a transcription factor to the gene of interest harboring the target sequence.

[0116] An effector domain may be a full-length protein or a fragment thereof that retains the epigenetic effector function (a "functional domain"). Functional domains that are capable of modulating (e.g., repressing) gene expression can be derived from a larger protein. For example, functional domains that can reduce target gene expression may be identified based on sequences of repressor proteins. Amino acid sequences of gene expression-modulating proteins may be obtained from available genome browsers, such as the UCSD genome browser or Ensembl genome browser. Protein annotation databases such as UniProt or Pfam can be used to identify functional domains within the full protein sequence. As a starting point, the largest sequence, encompassing all regions identified by different databases, may be tested for gene expression modulation activity. Various truncations then may be tested to identify the minimal functional unit.

[0117] Variants of effector domains described herein are also contemplated by the present disclosure. A variant may, for example, refer to a polypeptide with at least about 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity and/or sequence similarity to a wildtype effector domain described herein. In particular embodiments, the variant retains at least about 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% of the epigenetic effector function of the wildtype effector domain.

[0118] In some embodiments, an effector domain described herein may comprise a fusion of two or more effector domains (e.g., KOX1 KRAB and ZIM3). The effector domain may,

for example, comprise a fusion of 2, 3, 4, 5, 6, 7, 8, 9, or 10 effector domains, such as effector domains described herein. In certain embodiments, an effector domain comprises a fusion of a truncated form of an effector domain and a second effector domain. In certain embodiments, an effector domain comprises a fusion of the truncated forms of two effector domains (e.g., fusions of the N- and C-terminal portions of the two effector domains).

[0119] In some embodiments, an epigenetic editor described herein may comprise 1 effector domain, 2 effector domains, 3 effector domains, 4 effector domains, 5 effector domains, 6 effector domains, 7 effector domains, 8 effector domains, 9 effector domains, 10 effector domains, or more. In certain embodiments, the epigenetic editor comprises one or more fusion proteins (e.g., one, two, or three fusion proteins), each with one or more effector domains (e.g., one, two, or three effector domains) linked to a DNA-binding domain. In some embodiments, the effector domains may induce a combination of epigenetic modifications, e.g., transcription repression and DNA methylation, DNA methylation and histone deacetylation, DNA methylation and histone demethylation, DNA methylation and histone methylation, DNA methylation and histone phosphorylation, DNA methylation and histone ubiquitylation, DNA methylation, and histone SUMOylation.

[0120] In certain embodiments, an effector domain described herein (e.g., DNMT3A and/or DNMT3L) is encoded by a nucleotide sequence as found in the native genome (e.g., human or murine) for that effector domain. In other embodiments, an effector domain described herein is encoded by a nucleotide sequence that has been codon-optimized for optimal expression in human cells.

[0121] Effector domains described herein may include, for example, transcriptional repressors, DNA methyltransferases, and/or histone modifiers, as further detailed below.

A. Transcriptional Repressors

[0122] In some embodiments, an epigenetic effector domain described herein mediates repression of a target gene's expression (e.g., transcription). The effector domain may comprise, e.g., a Krüppel-associated box (KRAB) repressor domain, a Repressor Element Silencing Transcription Factor (REST) repressor domain, a KRAB-associated protein 1 (KAP1) domain, a MAD domain, a FKHR (forkhead in rhabdosarcoma gene) repressor domain, an EGR-1 (early growth response gene product-1) repressor domain, an ets2 repressor factor repressor domain (ERD), a MAD smSIN3 interaction domain (SID), a WRPW motif of the hairy-related basic helix-loop-helix (bHLH) repressor proteins, an HP1 alpha chromo-shadow repressor domain, an HP1 beta repressor domain, or any combination thereof. The effector domain may recruit one or more protein domains that repress

expression of the target gene, e.g., through a scaffold protein. In some embodiments, the effector domain may recruit or interact with a scaffold protein domain that recruits a PRMT protein, a HDAC protein, a SETDB1 protein, or a NuRD protein domain.

[0123] In some embodiments, the effector domain comprises a functional domain derived from a zinc finger repressor protein, such as a KRAB domain. KRAB domains are found in approximately 400 human ZFP-based transcription factors. Descriptions of KRAB domains may be found, for example, in Ecco et al., *Development* (2017) 144(15):2719-29 and Lambert et al., *Cell* (2018) 172:650-65.

[0124] In certain embodiments, the effector domain comprises a repressor domain (e.g., KRAB) derived from KOX1/ZNF10, KOX8/ZNF708, ZNF43, ZNF184, ZNF91, HPF4, HTF10, or HTF34. In some embodiments, the effector domain comprises a repressor domain (e.g., KRAB) derived from ZIM3, ZNF436, ZNF257, ZNF675, ZNF490, ZNF320, ZNF331, ZNF816, ZNF680, ZNF41, ZNF189, ZNF528, ZNF543, ZNF554, ZNF140, ZNF610, ZNF264, ZNF350, ZNF8, ZNF582, ZNF30, ZNF324, ZNF98, ZNF669, ZNF677, ZNF596, ZNF214, ZNF37, ZNF34, ZNF250, ZNF547, ZNF273, ZNF354, ZFP82, ZNF224, ZNF33, ZNF45, ZNF175, ZNF595, ZNF184, ZNF419, ZFP28-1, ZFP28-2, ZNF18, ZNF213, ZNF394, ZFP1, ZFP14, ZNF416, ZNF557, ZNF566, ZNF729, ZIM2, ZNF254, ZNF764, ZNF785, or any combination thereof. For example, the repressor domain may be a KRAB domain derived from KOX1, ZIM3, ZFP28, or ZN627. In particular embodiments, the repressor domain is a ZIM3 KRAB domain. In further embodiments, the effector domain is derived from a human protein, e.g., a human ZIM3, a human KOX1, a human ZFP28, or a human ZN627.

[0125] Sequences of exemplary effector domains that may reduce or silence target gene expression, or protein sequences that contain them, are provided in **Table 4** below (SEQ: SEQ ID NO). Further examples of repressors and transcriptional repressor domains can be found, e.g., in PCT Patent Publication WO 2021/226077 and Tycko et al., *Cell* (2020) 183(7):2020-35, each of which is incorporated herein by reference in its entirety.

Table 4. Exemplary Effector Domains That May Reduce or Silence Gene Expression

Protein	SEQ
ZIM3	33
ZNF436	34
ZNF257	35
ZNF675	36
ZNF490	37

Protein	SEQ
ZN440_HUMAN	299
ZN583_HUMAN	300
ZN441_HUMAN	301
ZNF43_HUMAN	302
CBX5_HUMAN	303

Protein	SEQ
ZNF320	38
ZNF331	39
ZNF816	40
ZNF680	41
ZNF41	42
ZNF189	43
ZNF528	44
ZNF543	45
ZNF554	46
ZNF140	47
ZNF610	48
ZNF264	49
ZNF350	50
ZNF8	51
ZNF582	52
ZNF30	53
ZNF324	54
ZNF98	55
ZNF669	56
ZNF677	57
ZNF596	58
ZNF214	59
ZNF37A	60
ZNF34	61
ZNF250	62
ZNF547	63
ZNF273	64
ZNF354A	65
ZFP82	66
ZNF224	67
ZNF33A	68
ZNF45	69
ZNF175	70
ZNF595	71
ZNF184	72
ZNF419	73
ZFP28-1	74

Protein	SEQ
ZN589_HUMAN	304
ZNF10_HUMAN	305
ZN563_HUMAN	306
ZN561_HUMAN	307
ZN136_HUMAN	308
ZN630_HUMAN	309
ZN527_HUMAN	310
ZN333_HUMAN	311
Z324B_HUMAN	312
ZN786_HUMAN	313
ZN709_HUMAN	314
ZN792_HUMAN	315
ZN599_HUMAN	316
ZN613_HUMAN	317
ZF69B_HUMAN	318
ZN799_HUMAN	319
ZN569_HUMAN	320
ZN564_HUMAN	321
ZN546_HUMAN	322
ZFP92_HUMAN	323
YAF2_HUMAN	324
ZN723_HUMAN	325
ZNF34_HUMAN	326
ZN439_HUMAN	327
ZFP57_HUMAN	328
ZNF19_HUMAN	329
ZN404_HUMAN	330
ZN274_HUMAN	331
CBX3_HUMAN	332
ZNF30_HUMAN	333
ZN250_HUMAN	334
ZN570_HUMAN	335
ZN675_HUMAN	336
ZN695_HUMAN	337
ZN548_HUMAN	338
ZN132_HUMAN	339
ZN738_HUMAN	340

Protein	SEQ
ZFP28-2	75
ZNF18	76
ZNF213	77
ZNF394	78
ZFP1	79
ZFP14	80
ZNF416	81
ZNF557	82
ZNF566	83
ZNF729	84
ZIM2	85
ZNF254	86
ZNF764	87
ZNF785	88
ZNF10 (KOX1)	89
CBX5 (chromoshadow domain)	90
RYBP (YAF2_RYBP component of PRC1)	91
YAF2 (YAF2_RYBP component of PRC1)	92
MGA (component of PRC1.6)	93
CBX1 (chromoshadow)	94
SCMH1 (SAM_1/SPM)	95
MPP8 (Chromodomain)	96
SUMO3 (Rad60-SLD)	97
HERC2 (Cyt-b5)	98
BIN1 (SH3_9)	99
PCGF2 (RING finger protein domain)	100
TOX (HMG box)	101
FOXA1 (HNF3A C-terminal domain)	102
FOXA2 (HNF3B C-terminal domain)	103
IRF2BP1 (IRF-2BP1_2 N-terminal domain)	104
IRF2BP2 (IRF-2BP1_2 N-terminal domain)	105
IRF2BPL IRF-2BP1_2 N-terminal domain	106
HOXA13 (homeodomain)	107

Protein	SEQ
ZN420_HUMAN	341
ZN626_HUMAN	342
ZN559_HUMAN	343
ZN460_HUMAN	344
ZN268_HUMAN	345
ZN304_HUMAN	346
ZIM2_HUMAN	347
ZN605_HUMAN	348
ZN844_HUMAN	349
SUMO5_HUMAN	350
ZN101_HUMAN	351
ZN783_HUMAN	352
ZN417_HUMAN	353
ZN182_HUMAN	354
ZN823_HUMAN	355
ZN177_HUMAN	356
ZN197_HUMAN	357
ZN717_HUMAN	358
ZN669_HUMAN	359
ZN256_HUMAN	360
ZN251_HUMAN	361
CBX4_HUMAN	362
PCGF2_HUMAN	363
CDY2_HUMAN	364
CDYL2_HUMAN	365
HERC2_HUMAN	366
ZN562_HUMAN	367
ZN461_HUMAN	368
Z324A_HUMAN	369
ZN766_HUMAN	370
ID2_HUMAN	371
TOX_HUMAN	372
ZN274_HUMAN	373

Protein	SEQ
HOXB13 (homeodomain)	108
HOXC13 (homeodomain)	109
HOXA11 (homeodomain)	110
HOXC11 (homeodomain)	111
HOXC10 (homeodomain)	112
HOXA10 (homeodomain)	113
HOXB9 (homeodomain)	114
HOXA9 (homeodomain)	115
ZFP28_HUMAN	116
ZN334_HUMAN	117
ZN568_HUMAN	118
ZN37A_HUMAN	119
ZN181_HUMAN	120
ZN510_HUMAN	121
ZN862_HUMAN	122
ZN140_HUMAN	123
ZN208_HUMAN	124
ZN248_HUMAN	125
ZN571_HUMAN	126
ZN699_HUMAN	127
ZN726_HUMAN	128
ZIK1_HUMAN	129
ZNF2_HUMAN	130
Z705F_HUMAN	131
ZNF14_HUMAN	132
ZN471_HUMAN	133
ZN624_HUMAN	134
ZNF84_HUMAN	135
ZNF7_HUMAN	136
ZN891_HUMAN	137
ZN337_HUMAN	138
Z705G_HUMAN	139
ZN529_HUMAN	140
ZN729_HUMAN	141
ZN419_HUMAN	142
Z705A_HUMAN	143
ZNF45_HUMAN	144

Protein	SEQ
SCMH1_HUMAN	374
ZN214_HUMAN	375
CBX7_HUMAN	376
ID1_HUMAN	377
CREM_HUMAN	378
SCX_HUMAN	379
ASCL1_HUMAN	380
ZN764_HUMAN	381
SCML2_HUMAN	382
TWST1_HUMAN	383
CREB1_HUMAN	384
TERF1_HUMAN	385
ID3_HUMAN	386
CBX8_HUMAN	387
CBX4_HUMAN	388
GSX1_HUMAN	389
NKX22_HUMAN	390
ATF1_HUMAN	391
TWST2_HUMAN	392
ZNF17_HUMAN	393
TOX3_HUMAN	394
TOX4_HUMAN	395
ZMYM3_HUMAN	396
I2BP1_HUMAN	397
RHXF1_HUMAN	398
SSX2_HUMAN	399
I2BPL_HUMAN	400
ZN680_HUMAN	401
CBX1_HUMAN	402
TRI68_HUMAN	403
HXA13_HUMAN	404
PHC3_HUMAN	405
TCF24_HUMAN	406
CBX3_HUMAN	407
HXB13_HUMAN	408
HEY1_HUMAN	409
PHC2_HUMAN	410

Protein	SEQ
ZN302_HUMAN	145
ZN486_HUMAN	146
ZN621_HUMAN	147
ZN688_HUMAN	148
ZN33A_HUMAN	149
ZN554_HUMAN	150
ZN878_HUMAN	151
ZN772_HUMAN	152
ZN224_HUMAN	153
ZN184_HUMAN	154
ZN544_HUMAN	155
ZNF57_HUMAN	156
ZN283_HUMAN	157
ZN549_HUMAN	158
ZN211_HUMAN	159
ZN615_HUMAN	160
ZN253_HUMAN	161
ZN226_HUMAN	162
ZN730_HUMAN	163
Z585A_HUMAN	164
ZN732_HUMAN	165
ZN681_HUMAN	166
ZN667_HUMAN	167
ZN649_HUMAN	168
ZN470_HUMAN	169
ZN484_HUMAN	170
ZN431_HUMAN	171
ZN382_HUMAN	172
ZN254_HUMAN	173
ZN124_HUMAN	174
ZN607_HUMAN	175
ZN317_HUMAN	176
ZN620_HUMAN	177
ZN141_HUMAN	178
ZN584_HUMAN	179
ZN540_HUMAN	180
ZN75D_HUMAN	181

Protein	SEQ
ZNF81_HUMAN	411
FIGLA_HUMAN	412
SAM11_HUMAN	413
KMT2B_HUMAN	414
HEY2_HUMAN	415
JDP2_HUMAN	416
HXC13_HUMAN	417
ASCL4_HUMAN	418
HHEX_HUMAN	419
HERC2_HUMAN	420
GSX2_HUMAN	421
BIN1_HUMAN	422
ETV7_HUMAN	423
ASCL3_HUMAN	424
PHC1_HUMAN	425
OTP_HUMAN	426
I2BP2_HUMAN	427
VGLL2_HUMAN	428
HXA11_HUMAN	429
PDLI4_HUMAN	430
ASCL2_HUMAN	431
CDX4_HUMAN	432
ZN860_HUMAN	433
LMBL4_HUMAN	434
PDIP3_HUMAN	435
NKX25_HUMAN	436
CEBPB_HUMAN	437
ISL1_HUMAN	438
CDX2_HUMAN	439
PROP1_HUMAN	440
SIN3B_HUMAN	441
SMBT1_HUMAN	442
HXC11_HUMAN	443
HXC10_HUMAN	444
PRS6A_HUMAN	445
VSX1_HUMAN	446
NKX23_HUMAN	447

Protein	SEQ
ZN555_HUMAN	182
ZN658_HUMAN	183
ZN684_HUMAN	184
RBAK_HUMAN	185
ZN829_HUMAN	186
ZN582_HUMAN	187
ZN112_HUMAN	188
ZN716_HUMAN	189
HKR1_HUMAN	190
ZN350_HUMAN	191
ZN480_HUMAN	192
ZN416_HUMAN	193
ZNF92_HUMAN	194
ZN100_HUMAN	195
ZN736_HUMAN	196
ZNF74_HUMAN	197
CBX1_HUMAN	198
ZN443_HUMAN	199
ZN195_HUMAN	200
ZN530_HUMAN	201
ZN782_HUMAN	202
ZN791_HUMAN	203
ZN331_HUMAN	204
Z354C_HUMAN	205
ZN157_HUMAN	206
ZN727_HUMAN	207
ZN550_HUMAN	208
ZN793_HUMAN	209
ZN235_HUMAN	210
ZNF8_HUMAN	211
ZN724_HUMAN	212
ZN573_HUMAN	213
ZN577_HUMAN	214
ZN789_HUMAN	215
ZN718_HUMAN	216
ZN300_HUMAN	217
ZN383_HUMAN	218

Protein	SEQ
MTG16_HUMAN	448
HMX3_HUMAN	449
HMX1_HUMAN	450
KIF22_HUMAN	451
CSTF2_HUMAN	452
CEBPE_HUMAN	453
DLX2_HUMAN	454
ZMYM3_HUMAN	455
PPARG_HUMAN	456
PRIC1_HUMAN	457
UNC4_HUMAN	458
BARX2_HUMAN	459
ALX3_HUMAN	460
TCF15_HUMAN	461
TERA_HUMAN	462
VSX2_HUMAN	463
HXD12_HUMAN	464
CDX1_HUMAN	465
TCF23_HUMAN	466
ALX1_HUMAN	467
HXA10_HUMAN	468
RX_HUMAN	469
CXXC5_HUMAN	470
SCML1_HUMAN	471
NFIL3_HUMAN	472
DLX6_HUMAN	473
MTG8_HUMAN	474
CBX8_HUMAN	475
CEBPD_HUMAN	476
SEC13_HUMAN	477
FIP1_HUMAN	478
ALX4_HUMAN	479
LHX3_HUMAN	480
PRIC2_HUMAN	481
MAGI3_HUMAN	482
NELL1_HUMAN	483
PRRX1_HUMAN	484

Protein	SEQ
ZN429_HUMAN	219
ZN677_HUMAN	220
ZN850_HUMAN	221
ZN454_HUMAN	222
ZN257_HUMAN	223
ZN264_HUMAN	224
ZFP82_HUMAN	225
ZFP14_HUMAN	226
ZN485_HUMAN	227
ZN737_HUMAN	228
ZNF44_HUMAN	229
ZN596_HUMAN	230
ZN565_HUMAN	231
ZN543_HUMAN	232
ZFP69_HUMAN	233
SUMO1_HUMAN	234
ZNF12_HUMAN	235
ZN169_HUMAN	236
ZN433_HUMAN	237
SUMO3_HUMAN	238
ZNF98_HUMAN	239
ZN175_HUMAN	240
ZN347_HUMAN	241
ZNF25_HUMAN	242
ZN519_HUMAN	243
Z585B_HUMAN	244
ZIM3_HUMAN	245
ZN517_HUMAN	246
ZN846_HUMAN	247
ZN230_HUMAN	248
ZNF66_HUMAN	249
ZFP1_HUMAN	250
ZN713_HUMAN	251
ZN816_HUMAN	252
ZN426_HUMAN	253
ZN674_HUMAN	254
ZN627_HUMAN	255

Protein	SEQ
MTG8R_HUMAN	485
RAX2_HUMAN	486
DLX3_HUMAN	487
DLX1_HUMAN	488
NKX26_HUMAN	489
NAB1_HUMAN	490
SAMD7_HUMAN	491
PITX3_HUMAN	492
WDR5_HUMAN	493
MEOX2_HUMAN	494
NAB2_HUMAN	495
DHX8_HUMAN	496
FOXA2_HUMAN	497
CBX6_HUMAN	498
EMX2_HUMAN	499
CPSF6_HUMAN	500
HXC12_HUMAN	501
KDM4B_HUMAN	502
LMBL3_HUMAN	503
PHX2A_HUMAN	504
EMX1_HUMAN	505
NC2B_HUMAN	506
DLX4_HUMAN	507
SRY_HUMAN	508
ZN777_HUMAN	509
NELL1_HUMAN	510
ZN398_HUMAN	511
GATA3_HUMAN	512
BSH_HUMAN	513
SF3B4_HUMAN	514
TEAD1_HUMAN	515
TEAD3_HUMAN	516
RGAP1_HUMAN	517
PHF1_HUMAN	518
FOXA1_HUMAN	519
GATA2_HUMAN	520
FOXO3_HUMAN	521

Protein	SEQ
ZNF20_HUMAN	256
Z587B_HUMAN	257
ZN316_HUMAN	258
ZN233_HUMAN	259
ZN611_HUMAN	260
ZN556_HUMAN	261
ZN234_HUMAN	262
ZN560_HUMAN	263
ZNF77_HUMAN	264
ZN682_HUMAN	265
ZN614_HUMAN	266
ZN785_HUMAN	267
ZN445_HUMAN	268
ZFP30_HUMAN	269
ZN225_HUMAN	270
ZN551_HUMAN	271
ZN610_HUMAN	272
ZN528_HUMAN	273
ZN284_HUMAN	274
ZN418_HUMAN	275
MPP8_HUMAN	276
ZN490_HUMAN	277
ZN805_HUMAN	278
Z780B_HUMAN	279
ZN763_HUMAN	280
ZN285_HUMAN	281
ZNF85_HUMAN	282
ZN223_HUMAN	283
ZNF90_HUMAN	284
ZN557_HUMAN	285
ZN425_HUMAN	286
ZN229_HUMAN	287
ZN606_HUMAN	288
ZN155_HUMAN	289
ZN222_HUMAN	290
ZN442_HUMAN	291
ZNF91_HUMAN	292

Protein	SEQ
ZN212_HUMAN	522
IRX4_HUMAN	523
ZBED6_HUMAN	524
LHX4_HUMAN	525
SIN3A_HUMAN	526
RBBP7_HUMAN	527
NKX61_HUMAN	528
TRI68_HUMAN	529
R51A1_HUMAN	530
MB3L1_HUMAN	531
DLX5_HUMAN	532
NOTC1_HUMAN	533
TERF2_HUMAN	534
ZN282_HUMAN	535
RGS12_HUMAN	536
ZN840_HUMAN	537
SPI2B_HUMAN_1	538
PAX7_HUMAN	539
NKX62_HUMAN	540
ASXL2_HUMAN	541
FOXO1_HUMAN	542
GATA3_HUMAN	543
GATA1_HUMAN	544
ZMYM5_HUMAN	545
ZN783_HUMAN	546
SPI2B_HUMAN_2	547
LRP1_HUMAN	548
MIXL1_HUMAN	549
SGT1_HUMAN	550
LMCD1_HUMAN	551
CEBPA_HUMAN	552
GATA2_HUMAN	553
SOX14_HUMAN	554
WTIP_HUMAN	555
PRP19_HUMAN	556
CBX6_HUMAN	557
NKX11_HUMAN	558

Protein	SEQ
ZN135_HUMAN	293
ZN778_HUMAN	294
RYBP_HUMAN	295
ZN534_HUMAN	296
ZN586_HUMAN	297
ZN567_HUMAN	298

Protein	SEQ
RBBP4_HUMAN	559
DMRT2_HUMAN	560
SMCA2_HUMAN	561
ZNF10_HUMAN	562
EED_HUMAN	563
RCOR1_HUMAN	564

[0126] A functional analog of any one of the above-listed proteins, i.e., a molecule having the same or substantially the same biological function (e.g., retaining 70% or more, 80% or more, 90% or more, 95% or more, or 98% or more) of the protein's transcription factor function) is encompassed by the present disclosure. For example, the functional analog may be an isoform or a variant of the above-listed protein, e.g., containing a portion of the above protein with or without additional amino acid residues and/or containing mutations relative to the above protein. In some embodiments, the functional analog has a sequence identity that is at least 75, 80, 85, 90, 95, 98, or 99% to one of the sequences listed in **Table 4**. Homologs, orthologs, and mutants of the above-listed proteins are also contemplated.

[0127] In certain embodiments, an epigenetic editor described herein comprises a KRAB domain derived from KOX1, ZIM3, ZFP28, or ZN627, and/or an effector domain derived from KAP1, MECP2, HP1a, HP1b, CBX8, CDYL2, TOX, TOX3, TOX4, EED, EZH2, RBBP4, RCOR1, or SCML2, optionally wherein the parental protein is a human protein. In particular embodiments, an epigenetic editor described herein comprises a domain derived from KOX1, ZIM3, ZFP28, and/or ZN627, optionally wherein the parental protein is a human protein. In certain embodiments, the epigenetic editor may comprise a KRAB domain derived from KOX1 (ZNF10), e.g., a human KOX1. In certain embodiments, the epigenetic editor may comprise a KRAB domain derived from ZIM3 (ZNF657 or ZNF264), e.g., a human ZIM3. In certain embodiments, the epigenetic editor may comprise a KRAB domain derived from ZFP28, e.g., a human ZFP28. In certain embodiments, the epigenetic editor may comprise a KRAB domain derived from ZN627, e.g., a human ZN627. In certain embodiments, an epigenetic editor described herein may comprise a CDYL2, e.g., a human CDYL2, and/or a TOX domain (e.g., a human TOX domain) in combination with a KOX1 KRAB domain (e.g., a human KOX1 KRAB domain).

[0128] In certain embodiments, an epigenetic effector described herein comprises a repressor domain derived from KOX1/ZNF10 (SEQ ID NO: 89). For example, the repressor domain may comprise the sequence of SEQ ID NO: 89, or a sequence at least 75%, 80%,

85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 89.

[0129] In certain embodiments, an epigenetic effector described herein comprises a repressor domain derived from KOX1/ZNF10, as shown in **Table 5** below:

Table 5. Exemplary Effector Domains Derived from KOX1/ZNF10

Protein	Protein Sequence
KOX1/ZNF10 KRAB 1	SEQ ID NO: 565
KOX1/ZNF10 KRAB 2	SEQ ID NO: 566
KOX1/ZNF10 KRAB 3	SEQ ID NO: 567
KOX1/ZNF10 (aa 11-72)	SEQ ID NO: 568
KOX1/ZNF10 (aa 11-108)	SEQ ID NO: 569
KOX1/ZNF10 variant	SEQ ID NO: 570
KOX1 KRAB-ZIM3 chimera	SEQ ID NO: 571
ZIM3-KOX1 KRAB chimera	SEQ ID NO: 572

[0130] In particular embodiments, the repressor domain may comprise the amino acid sequence of SEQ ID NO: 565, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 565.

[0131] In particular embodiments, the repressor domain may comprise the amino acid sequence of SEQ ID NO: 566, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 566.

[0132] In particular embodiments, the repressor domain may comprise the amino acid sequence of SEQ ID NO: 567, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 567.

[0133] In particular embodiments, the repressor domain may comprise the amino acid sequence of SEQ ID NO: 568, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 568.

[0134] In particular embodiments, the repressor domain may comprise the amino acid sequence of SEQ ID NO: 569, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 569.

[0135] In particular embodiments, the repressor domain may comprise the amino acid sequence of SEQ ID NO: 570, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 570.

[0136] In particular embodiments, the repressor domain may comprise the amino acid sequence of SEQ ID NO: 571, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 571.

[0137] In particular embodiments, the repressor domain may comprise the amino acid sequence of SEQ ID NO: 572, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 572.

B. DNA Methyltransferases

[0138] In some embodiments, an effector domain of an epigenetic editor described herein alters target gene expression through DNA modification, such as methylation. Highly methylated areas of DNA tend to be less transcriptionally active than less methylated areas. DNA methylation occurs primarily at CpG sites (shorthand for "C-phosphate-G-" or "cytosine-phosphate-guanine" sites). Many mammalian genes have promoter regions near or including CpG islands (nucleic acid regions with a high frequency of CpG dinucleotides).

[0139] An effector domain described herein may be, e.g., a DNA methyltransferase (DNMT) or a catalytic domain thereof, or may be capable of recruiting a DNA methyltransferase. DNMTs encompass enzymes that catalyze the transfer of a methyl group to a DNA nucleotide, such as canonical cytosine-5 DNMTs that catalyze the addition of methyl groups to genomic DNA (e.g., DNMT1, DNMT3A, DNMT3B, and DNMT3C). This term also encompasses non-canonical family members that do not catalyze methylation themselves but that recruit (including activate) catalytically active DNMTs; a non-limiting examples of such a DNMT is DNMT3L. *See, e.g.,* Lyko, *Nat Review* (2018) 19:81-92. Unless otherwise indicated, a DNMT domain may refer to a polypeptide domain derived from a catalytically active DNMT (e.g., DNMT1, DNMT3A, and DNMT3B) or from a catalytically inactive DNMT (e.g., DNMT3L). A DNMT may repress expression of the target gene through the recruitment of repressive regulatory proteins. In some embodiments, the methylation is at a CG (or CpG) dinucleotide sequence. In some embodiments, the methylation is at a CHG or CHH sequence, where H is any one of A, T, or C.

[0140] In some embodiments, a DNMT described herein can be an animal DNMT (e.g., a mammalian DNMT), a plant DNMT, a fungal DNMT, or a bacterial DNMT. A bacterial DNMT can be obtained from a bacterial species (e.g., a coccus bacterium, bacillus bacterium, spiral bacterium, or an intracellular, gram-positive, or gram-negative bacterium. In certain embodiments, the bacterial species is *Mycoplasmatales bacterium*, *Mycoplasma marinum*, or *Spiroplasma chinense*. In certain embodiments, the bacterial species is not *M. penetrans*, *S.*

monbiae, *H. parainfluenzae*, *A. luteus*, *H. aegyptius*, *H. haemolyticus*, *Moraxella*, *E. coli*, *T. aquaticus*, *C. crescentus*, or *C. difficile*. In certain embodiments, an epigenetic editor described herein comprises a DNMT domain comprising SEQ ID NO: 601, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 601. In certain embodiments, an epigenetic editor described herein comprises a DNMT domain comprising SEQ ID NO: 602, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 602. In certain embodiments, an epigenetic editor described herein comprises a DNMT domain comprising SEQ ID NO: 603, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 603.

[0141] In certain embodiments, DNMTs in the epigenetic editors described herein may include, e.g., DNMT1, DNMT3A, DNMT3B, and/or DNMT3C. In some embodiments, the DNMT is a mammalian (e.g., human or murine) DNMT. In particular embodiments, the DNMT is DNMT3A (e.g., human DNMT3A). In certain embodiments, an epigenetic editor described herein comprises a DNMT3A domain comprising SEQ ID NO: 574, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 574. In certain embodiments, an epigenetic editor described herein comprises a DNMT3A domain comprising SEQ ID NO: 575, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 575. In some embodiments, the DNMT3A domain may have, e.g., a mutation at position H739 (such as H739A or H739E), R771 (such as R771L) and/or R836 (such as R836A or R836Q), or any combination thereof (numbering according to SEQ ID NO: 574).

[0142] In some embodiments, an effector domain described herein may be a DNMT-like domain. As used herein a "DNMT-like domain" is a regulatory factor of DNMT that may activate or recruit other DNMT domains, but does not itself possess methylation activity. In some embodiments, the DNMT-like domain is a mammalian (e.g., human or mouse) DNMT-like domain. In certain embodiments, the DNMT-like domain is DNMT3L, which may be, for example, human DNMT3L or mouse DNMT3L. In certain embodiments, an epigenetic editor described herein comprises a DNMT3L domain comprising SEQ ID NO: 578, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 578. In certain embodiments, an epigenetic editor herein comprises a DNMT3L domain comprising SEQ ID NO: 579, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 579. In certain embodiments, an epigenetic editor described herein comprises a

DNMT3L domain comprising SEQ ID NO: 580, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 580. In certain embodiments, an epigenetic editor described herein comprises a DNMT3L domain comprising SEQ ID NO: 581, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 581. In some embodiments, the DNMT3L domain may have, e.g., a mutation corresponding to that at position D226 (such as D226V), Q268 (such as Q268K), or both (numbering according to SEQ ID NO: 578).

[0143] In certain embodiments, an epigenetic editor herein may comprise comprising both DNMT and DNMT-like effector domains. For example, the epigenetic editor may comprise a DNMT3A-3L domain, wherein DNMT3A and DNMT3L may be covalently linked. In other embodiments, an epigenetic editor described herein may comprise an effector domain that comprises only a DNMT3A domain (e.g., human DNMT3A), or only a DNMT-like domain (e.g., DNMT3L, which may be human or mouse DNMT3L).

[0144] **Table 6** below provides exemplary DNMTs that may be part of an epigenetic effector domain described herein, or from which an effector domain of an epigenetic editor described herein may be derived.

Table 6. Exemplary DNMT Sequences

Protein Name	Species	Target	Protein Sequence
DNMT1	Human	5mC	SEQ ID NO: 573
DNMT3A (h3A)	Human	5mC	SEQ ID NO: 574
DNMT3A (catalytic domain) (h3As)	Human	5mC	SEQ ID NO: 575
DNMT3B	Human	5mC	SEQ ID NO: 576
DNMT3C	Mouse	5mC	SEQ ID NO: 577
DNMT3L (h3L)	Human	5mC	SEQ ID NO: 578
DNMT3L (catalytic domain) (h3Ls)	Human	5mC	SEQ ID NO: 579
DNMT3L (m3L)	Mouse	5mC	SEQ ID NO: 580
DNMT3L (catalytic domain) (m3Ls)	Mouse	5mC	SEQ ID NO: 581
DNMT3L	*Ailuropoda melanoleuca*	5mC	SEQ ID NO: 582
DNMT3L (catalytic domain)	*Ailuropoda melanoleuca*	5mC	SEQ ID NO: 583

Protein Name	Species	Target	Protein Sequence
DNMT3L	*Carlito syrichta*	5mC	SEQ ID NO: 584
DNMT3L (catalytic domain)	*Carlito syrichta*	5mC	SEQ ID NO: 585
DNMT3L	*Meriones unguiculatus*	5mC	SEQ ID NO: 586
DNMT3L (catalytic domain)	*Meriones unguiculatus*	5mC	SEQ ID NO: 587
DNMT3L	*Ochotona princeps*	5mC	SEQ ID NO: 588
DNMT3L (catalytic domain)	*Ochotona princeps*	5mC	SEQ ID NO: 589
DNMT3L	*Neosciurus carolinensis*	5mC	SEQ ID NO: 590
DNMT3L (catalytic domain)	*Neosciurus carolinensis*	5mC	SEQ ID NO: 591
DNMT3L	*Bison bison*	5mC	SEQ ID NO: 592
DNMT3L (catalytic domain)	*Bison bison*	5mC	SEQ ID NO: 593
DNMT3L	*Equus przewalskii*	5mC	SEQ ID NO: 594
DNMT3L (catalytic domain)	*Equus przewalskii*	5mC	SEQ ID NO: 595
DNMT3L	*Mus caroli*	5mC	SEQ ID NO: 596
DNMT3L (catalytic domain)	*Mus caroli*	5mC	SEQ ID NO: 597
DNMT3L	*Pan troglodytes*	5mC	SEQ ID NO: 598
DNMT3L (catalytic domain)	*Pan troglodytes*	5mC	SEQ ID NO: 599
TRDMT1 (DNMT2)	Human	tRNA 5mC	SEQ ID NO: 600
DNA cytosine methyltransferase	*Mycoplasmatales bacterium*	5mC	SEQ ID NO: 601
DNA cytosine methyltransferase	*Mycoplasma marinum*	5mC	SEQ ID NO: 602
DNA (cytosine-5-)-methyltransferase	*Spiroplasma chinense*	5mC	SEQ ID NO: 603
M.MpeI	*Mycoplasma penetrans*	5mC	SEQ ID NO: 604
M.SssI	*Spiroplasma monobiae*	5mC	SEQ ID NO: 605
M.HpaII	*Haemophilus parainfluenzae*	5mC (CCGG)	SEQ ID NO: 606
M.AluI	*Arthrobacter luteus*	5mC (AGCT)	SEQ ID NO: 607
M.HaeIII	*Haemophilus aegyptius*	5mC (GGCC)	SEQ ID NO: 608
M.HhaI	*Haemophilus haemolyticus*	5mC (GCGC)	SEQ ID NO: 609
M.MspI	*Moraxella*	5mC (CCGG)	SEQ ID NO: 610
Masc1	*Ascobolus*	5mC	SEQ ID NO: 611
MET1	*Arabidopsis*	5mC	SEQ ID NO: 612
Masc2	*Ascobolus*	5mC	SEQ ID NO: 613

Protein Name	Species	Target	Protein Sequence
Dim-2	*Neurospora*	5mC	SEQ ID NO: 614
dDnmt2	*Drosophila*	5mC	SEQ ID NO: 615
Pmt1	*S. pombe*	5mC	SEQ ID NO: 616
DRM1	*Arabidopsis*	5mC	SEQ ID NO: 617
DRM2	*Arabidopsis*	5mC	SEQ ID NO: 618
CMT1	*Arabidopsis*	5mC	SEQ ID NO: 619
CMT2	*Arabidopsis*	5mC	SEQ ID NO: 620
CMT3	*Arabidopsis*	5mC	SEQ ID NO: 621
Rid	*Neurospora*	5mC	SEQ ID NO: 622
hsdM gene	bacteria (*E. coli*, strain 12)	m6A	SEQ ID NO: 623
hsdS gene	bacteria (*E. coli*, strain 12)	m6A	SEQ ID NO: 624
M.TaqI	Bacteria (*Thermus aquaticus*)	m6A	SEQ ID NO: 625
M.EcoDam	*E. coli*	m6A	SEQ ID NO: 626
M.CcrMI	*Caulobacter crescentus*	m6A	SEQ ID NO: 627
CamA	*Clostridioides difficile*	m6A	SEQ ID NO: 628

[0145] A functional analog of any one of the above-listed proteins, i.e., a molecule having the same or substantially the same biological function (e.g., retaining 70% or more, 80% or more, 90% or more, 95% or more, or 98% or more) of the protein's DNA methylation function or recruiting function) is encompassed by the present disclosure. For example, the functional analog may be an isoform or a variant of the above-listed protein, e.g., containing a portion of the above protein with or without additional amino acid residues and/or containing mutations relative to the above protein. In some embodiments, the functional analog has a sequence identity that is at least 75, 80, 85, 90, 95, 98, or 99% to one of the sequences listed in **Table 6**. In some embodiments, the effector domain herein comprises only the functional domain (or functional analog thereof), e.g., the catalytic domain or recruiting domain, of an above-listed protein. In some embodiments, the effector domain herein comprises one or more epigenetic effector domains selected from **Table 6**, or functional homologs, orthologs, or variants thereof.

[0146] As used herein, a DNMT domain (e.g., a DNMT3A domain or a DNMT3L domain) refers to a protein domain that is identical to the parental protein (e.g., a human or murine DNMT3A or DNMT3L) or a functional analog thereof (e.g., having a functional fragment, such as a catalytic fragment or recruiting fragment, of the parental protein; and/or having mutations that improve the activity of the DNMT protein).

[0147] An epigenetic editor herein may effect methylation at, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, or 1000 or more CpG dinucleotide sequences in the target gene or chromosome. The CpG dinucleotide sequences may be located within or near the target gene in CpG islands, or may be located in a region that is not a CpG island. A CpG island generally refers to a nucleic acid sequence or chromosome region that comprises a high frequency of CpG dinucleotides. For example, a CpG island may comprise at least 50% GC content. The CpG island may have a high observed-to-expected CpG ratio, for example, an observed-to-expected CpG ratio of at least 60%. As used herein, an observed-to-expected CpG ratio is determined by Number of CpG * (sequence length) / (Number of C * Number of G). In some embodiments, the CpG island has an observed-to-expected CpG ratio of at least 60%, 70%, 80%, 90% or more. A CpG island may be a sequence or region of, e.g., at least 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, or 800 nucleotides. In some embodiments, only 1, or less than 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, or 50 CpG dinucleotides are methylated by the epigenetic editor.

[0148] In some embodiments, an epigenetic editor herein effects methylation at a hypomethylated nucleic acid sequence, i.e., a sequence that may lack methyl groups on the 5-methyl cytosine nucleotides (e.g., in CpG) as compared to a standard control. Hypomethylation may occur, for example, in aging cells or in cancer (e.g., early stages of neoplasia) relative to a younger cell or non-cancer cell, respectively.

[0149] In some embodiments, an epigenetic editor described herein induces methylation at a hypermethylated nucleic acid sequence.

[0150] In some embodiments, methylation may be introduced by the epigenetic editor at a site other than a CpG dinucleotide. For example, the target gene sequence may be methylated at the C nucleotide of CpA, CpT, or CpC sequences. In some embodiments, an epigenetic editor comprises a DNMT3A domain and effects methylation at CpG, CpA, CpT, CpC sequences, or any combination thereof. In some embodiments, an epigenetic editor comprises a DNMT3A domain that lacks a regulatory subdomain and only maintains a catalytic domain. In some embodiments, the epigenetic editor comprising a DNMT3A catalytic domain effects methylation exclusively at CpG sequences. In some embodiments, an epigenetic editor comprising a DNMT3A domain that comprises a mutation, e.g. a R836A or R836Q mutation (numbering according to SEQ ID NO: 574), has higher methylation activity at CpA, CpC, and/or CpT sequences as compared to an epigenetic editor comprising a wildtype DNMT3A domain.

C. Histone Modifiers

[0151] In some embodiments, an effector domain of an epigenetic editor herein mediates histone modification. Histone modifications play a structural and biochemical role in gene transcription, such as by formation or disruption of the nucleosome structure that binds to the histone and prevents gene transcription. Histone modifications may include, for example, acetylation, deacetylation, methylation, phosphorylation, ubiquitination, SUMOylation and the like, e.g., at their N-terminal ends ("histone tails"). These modifications maintain or specifically convert chromatin structure, thereby controlling responses such as gene expression, DNA replication, DNA repair, and the like, which occur on chromosomal DNA. Post-translational modification of histones is an epigenetic regulatory mechanism and is considered essential for the genetic regulation of eukaryotic cells. Recent studies have revealed that chromatin remodeling factors such as SWI/SNF, RSC, NURF, NRD, and the like, which facilitate transcription factor access to DNA by modifying the nucleosome structure; histone acetyltransferases (HATs) that regulate the acetylation state of histones; and histone deacetylases (HDACs), act as important regulators.

[0152] In particular, the unstructured N-termini of histones may be modified by acetylation, deacetylation, methylation, ubiquitylation, phosphorylation, SUMOylation, ribosylation, citrullination O-GlcNAcylation, crotonylation, or any combination thereof. For example, histone acetyltransferases (HATs) utilize acetyl-CoA as a cofactor and catalyze the transfer of an acetyl group to the epsilon amino group of the lysine side chains. This neutralizes the lysine's positive charge and weakens the interactions between histones and DNA, thus opening the chromosomes for transcription factors to bind and initiate transcription. Acetylation of K14 and K9 lysines of histone H3 by histone acetyltransferase enzymes may be linked to transcriptional competence in humans. Lysine acetylation may directly or indirectly create binding sites for chromatin-modifying enzymes that regulate transcriptional activation. On the other hand, histone methylation of lysine 9 of histone H3 may be associated with heterochromatin, or transcriptionally silent chromatin.

[0153] In certain embodiments, an effector domain of an epigenetic editor described herein comprises a histone methyltransferase domain. The effector domain may comprise, for example, a DOT1L domain, a SET domain, a SUV39H1 domain, a G9a/EHMT2 protein domain, an EZH1 domain, an EZH2 domain, a SETDB1 domain, or any combination thereof. In particular embodiments, the effector domain comprises a histone-lysine-N-methyltransferase SETDB1 domain.

[0154] In some embodiments, the effector domain comprises a histone deacetylase protein domain. In certain embodiments, the effector domain comprises a HDAC family protein domain, for example, a HDAC1, HDAC3, HDAC5, HDAC7, or HDAC9 protein domain. In particular embodiments, the effector domain comprises a nucleosome remodeling and deacetylase complex (NURD), which removes acetyl groups from histones.

D. Other Effector Domains

[0155] In some embodiments, the effector domain comprises a tripartite motif containing protein (TRIM28, TIF1-beta, or KAP1). In certain embodiments, the effector domain comprises one or more KAP1 proteins. A KAP1 protein in an epigenetic editor herein may form a complex with one or more other effector domains of the epigenetic editor or one or more proteins involved in modulation of gene expression in a cellular environment. For example, KAP1 may be recruited by a KRAB domain of a transcriptional repressor. A KAP1 protein domain may interact with or recruit one or more protein complexes that reduces or silences gene expression. In some embodiments, KAP1 interacts with or recruits a histone deacetylase protein, a histone-lysine methyltransferase protein, a chromatin remodeling protein, and/or a heterochromatin protein. For example, a KAP1 protein domain may interact with or recruit a heterochromatin protein 1 (HP1) protein, a SETDB1 protein, an HDAC protein, and/or a NuRD protein complex component. In some embodiments, a KAP1 protein domain interacts with or recruits a ZFP90 protein (e.g., isoform 2 of ZFP90), and/or a FOXP3 protein. An exemplary KAP1 amino acid sequence is shown in SEQ ID NO: 629.

[0156] In some embodiments, the effector domain comprises a protein domain that interacts with or is recruited by one or more DNA epigenetic marks. For example, the effector domain may comprise a methyl CpG binding protein 2 (MECP2) protein that interacts with methylated DNA nucleotides in the target gene (which may or may not be at a CpG island of the target gene). An MECP2 protein domain in an epigenetic editor described herein may induce condensed chromatin structure, thereby reducing or silencing expression of the target gene. In some embodiments, an MECP2 protein domain in an epigenetic editor described herein may interact with a histone deacetylase (e.g., HDAC), thereby repressing or silencing expression of the target gene. In some embodiments, an MECP2 protein domain in an epigenetic editor described herein may block access of a transcription factor or transcriptional activator to the target sequence, thereby repressing or silencing expression of the target gene. An exemplary MECP2 amino acid sequence is shown in SEQ ID NO: 630.

[0157] Also contemplated as effector domains for the epigenetic editors described herein are, e.g., a chromoshadow domain, a ubiquitin-2 like Rad60 SUMO-like (Rad60-

SLD/SUMO) domain, a chromatin organization modifier domain (Chromo) domain, a Yaf2/RYBP C-terminal binding motif domain (YAF2_RYBP), a CBX family C-terminal motif domain (CBX7_C), a zinc finger C3HC4 type (RING finger) domain (ZF-C3HC4_2), a cytochrome b5 domain (Cyt-b5), a helix-loop-helix domain (HLH), a helix-hairpin-helix motif domain (e.g., HHH_3), a high mobility group box domain (HMG-box), a basic leucine zipper domain (e.g., bZIP_1 or bZIP_2), a Myb_DNA-binding domain, a homeodomain, a MYM-type zinc finger with FCS sequence domain (ZF-FCS), an interferon regulatory factor 2-binding protein zinc finger domain (IRF-2BP1_2), an SSX repressor domain (SSXRD), a B-box-type zinc finger domain (ZF-B_box), a CXXC zinc finger domain (ZF-CXXC), a regulator of chromosome condensation 1 domain (RCC1), an SRC homology 3 domain (SH3_9), a sterile alpha motif domain (SAM_1), a sterile alpha motif domain (SAM_2), a sterile alpha motif/Pointed domain (SAM_PNT), a Vestigial/Tondu family domain (Vg_Tdu), a LIM domain, an RNA recognition motif domain (RRM_1), a paired amphipathic helix domain (PAH), a proteasomal ATPase OB C-terminal domain (Prot_ATP_ID_OB), a nervy homology 2 domain (NHR2), a hinge domain of cleavage stimulation factor subunit 2 (CSTF2_hinge), a PPAR gamma N-terminal region domain (PPARgamma_N), a CDC48 N-terminal domain (CDC48_2), a WD40 repeat domain (WD40), a Fip1 motif domain (Fip1), a PDZ domain (PDZ_6), a Von Willebrand factor type C domain (VWC), a NAB conserved region 1 domain (NCD1), an S1 RNA-binding domain (S1), an HNF3 C-terminal domain (HNF_C), a Tudor domain (Tudor_2), a histone-like transcription factor (CBF/NF-Y) and archaeal histone domain (CBFD_NFYB_HMF), a zinc finger protein domain (DUF3669), an EGF-like domain (cEGF), a GATA zinc finger domain (GATA), a TEA/ATTS domain (TEA), a phorbol esters/diacylglycerol binding domain (C1-1), polycomb-like MTF2 factor 2 domain (Mtf2_C), a transactivation domain of FOXO protein family (FOXO-TAD), a homeobox KN domain (Homeobox_KN), a BED zinc finger domain (ZF-BED), a zinc finger of C3HC4-type RING domain (ZF-C3HC4_4), a RAD51 interacting motif domain (RAD51_interact), a p55-binding region of a nethyl-CpG-binding domain protein MBD (MBDa), a Notch domain, a Raf-like Ras-binding domain (RBD), a Spin/Ssty family domain (Spin-Ssty), a PHD finger domain (PHD_3), a Low-density lipoprotein receptor domain class A (Ldl_recept_a), a CS domain, a DM DNA-binding domain, and a QLQ domain.

[0158] In some embodiments, the effector domain is a protein domain comprising a YAF2_RYBP domain or homeodomain or any combination thereof. In certain embodiments, the homeodomain of the YAF2_RYBP domain is a PRD domain, an NKL domain, a HOXL

domain, or a LIM domain. In particular embodiments, the YAF2_RYBP domain may comprise a 32 amino acid Yaf2/RYBP C-terminal binding motif domain (32 aa RYBP).

[0159] In some embodiments, the effector domain comprises a protein domain selected from a group consisting of SUMO3 domain, Chromo domain from M phase phosphoprotein 8 (MPP8), chromoshadow domain from Chromobox 1 (CBX1), and SAM_1/SPM domain from Scm Polycomb Group Protein Homolog 1 (SCMH1).

[0160] In some embodiments, the effector domain comprises an HNF3 C-terminal domain (HNF_C). The HNF_C domain may be from FOXA1 or FOXA2. In certain embodiments, the HNF_C domain comprises an EH1 (engrailed homology 1) motif.

[0161] In some embodiments, the effector domain may comprise an interferon regulatory factor 2-binding protein zinc finger domain (IRF-2BP1_2), a Cyt-b5 domain from DNA repair factor HERC2 E3 ligase, a variant SH3 domain (SH3_9) from Bridging Integrator 1 (BIN1), an HMG-box domain from transcription factor TOX or ZF-C3HC4_2 RING finger domain from the polycomb component PCGF2, a Chromodomain-helicase-DNA binding protein 3 (CHD3) domain, or a ZNF783 domain.

IV. Epigenetic Editors

[0162] Provided herein are epigenetic editors (i.e., epigenetic editing systems) that direct epigenetic modification(s) to a target sequence in a gene of interest, e.g., using one or more DNA-binding domains as described herein and one or more effector domains (e.g., epigenetic repressor domains) as described herein, in any combination. The DNA-binding domain (in concert with a guide polynucleotide such as one described herein, where the DNA-binding domain is a polynucleotide guided DNA-binding domain) directs the effector domain to epigenetically modify the target sequence, resulting in gene repression or silencing that may be durable and inheritable across cell generations. In some aspects, the epigenetic editors described herein can repress or silence genes reversibly or irreversibly in cells.

[0163] In particular embodiments, an epigenetic editor described herein comprises one or more fusion proteins, each comprising (1) DNA-binding domain(s) and (2) effector domain(s). The effector domains may be on one or more fusion proteins comprised by the epigenetic editor. For example, a single fusion protein may comprise all of the effector domains with a DNA-binding domain. Alternatively, the effector domains or subsets thereof may be on separate fusion proteins, each with a DNA-binding domain (which may be the same or different). A fusion protein described herein may further comprise one or more linkers (e.g., peptide linkers), detectable tags, nuclear localization signals (NLSs), or any combination thereof. As used herein, a “fusion protein” refers to a chimeric protein in which

two or more coding sequences (e.g., for DNA-binding domain(s) and/or effector domain(s)) are covalently or non-covalently joined, directly or indirectly.

[0164] In some embodiments, an epigenetic editor described herein comprises 2, 3, 4, 5, 6, 7, 8, 9, 10, or more effector (e.g., repression/repressor) domains, which may be identical or different. In certain embodiments, two or more of said effector domains function synergistically. Combinations of effector domains may comprise DNA methylation domains, histone deacetylation domains, histone methylation domains, and/or scaffold domains that recruit any of the above. For example, an epigenetic editor described herein may comprise one or more transcriptional repressor domains (e.g., a KRAB domain such as KOX1, ZIM3, ZFP28, or ZN627 KRAB) in combination with one or more DNA methylation domains (e.g., a DNMT domain) and/or recruiter domain (e.g., a DNMT3L domain). Such an epigenetic editor may comprise, for instance, a KRAB domain, a DNMT3A domain, and a DNMT3L domain. In some embodiments, the epigenetic editor further comprises an additional effector domain (e.g., a KAP1, MECP2, HP1b, CBX8, CDYL2, TOX, TOX3, TOX4, EED, RBBP4, RCOR1, or SCML2 domain). In some embodiments, the additional effector domain is a CDYL2, TOX, TOX3, TOX4, or HP1a domain. For example, an epigenetic editor described herein may comprise a CDYL2 and/or a TOX domain in combination with a KRAB domain (e.g., a KOX1 KRAB domain).

A. Linkers

[0165] A fusion protein as described herein may comprise one or more linkers that connect components of the epigenetic editor. A linker may be a peptide or non-peptide linker.

[0166] In some embodiments, one or more linkers utilized in an epigenetic editor provided herein is a peptide linker, i.e., a linker comprising a peptide moiety. A peptide linker can be any length applicable to the epigenetic editor fusion proteins described herein. In some embodiments, the linker can comprise a peptide between 1 and 200 (e.g., between 1 and 80) amino acids. In some embodiments, the linker comprises from 1 to 5, 1 to 10, 1 to 20, 1 to 30, 1 to 40, 1 to 50, 1 to 60, 1 to 80, 1 to 100, 1 to 150, 1 to 200, 5 to 10, 5 to 20, 5 to 30, 5 to 40, 5 to 60, 5 to 80, 5 to 100, 5 to 150, 5 to 200, 10 to 20, 10 to 30, 10 to 40, 10 to 50, 10 to 60, 10 to 80, 10 to 100, 10 to 150, 10 to 200, 20 to 30, 20 to 40, 20 to 50, 20 to 60, 20 to 80, 20 to 100, 20 to 150, 20 to 200, 30 to 40, 30 to 50, 30 to 60, 30 to 80, 30 to 100, 30 to 150, 30 to 200, 40 to 50, 40 to 60, 40 to 80, 40 to 100, 40 to 150, 40 to 200, 50 to 60 50 to 80, 50 to 100, 50 to 150, 50 to 200, 60 to 80, 60 to 100, 60 to 150, 60 to 200, 80 to 100, 80 to 150, 80 to 200, 100 to 150, 100 to 200, or 150 to 200 amino acids in length. Longer or

shorter linkers are also contemplated. In some embodiments, the peptide linker is 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 25, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 amino acids in length. For example, the peptide linker may be 4, 5, 16, 20, 24, 27, 32, 40, 64, 92, or 104 amino acids in length. The peptide linker may be a flexible or rigid linker. In particular embodiments, the peptide linker comprises the amino acid sequence of any one of SEQ ID NOs: 631-637 and 664-665 or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical thereto.

[0167] In certain embodiments, the peptide linker is an XTEN linker. Such a linker may comprise part of the XTEN sequence (Schellenberger et al., *Nat Biotechnol* (2009) 27(1):1186-90), an unstructured hydrophilic polypeptide consisting only of residues G, S, P, T, E, and A. The term "XTEN" as used herein refers to a recombinant peptide or polypeptide lacking hydrophobic amino acid residues. XTEN linkers typically are unstructured and comprise a limited set of natural amino acids. Fusion of XTEN to proteins alters its hydrodynamic properties and reduces the rate of clearance and degradation of the fusion protein. These XTEN fusion proteins are produced using recombinant technology, without the need for chemical modifications, and degraded by natural pathways. The XTEN linker may be, for example, 5, 10, 16, 20, 26, or 80 amino acids in length. In some embodiments, the XTEN linker is 16 amino acids in length. In some embodiments, the XTEN linker is 80 amino acids in length. In certain embodiments, the XTEN linker may be XTEN10, XTEN16, XTEN20, or XTEN80. In certain embodiments, the XTEN linker may comprise the amino acid sequence of any one of SEQ ID NOs: 638-643 or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical thereto. In particular embodiments, the XTEN linker comprises the amino acid sequence of SEQ ID NO: 638. In particular embodiments, the XTEN linker comprises the amino acid sequence of SEQ ID NO: 643.

[0168] In some embodiments, one or more linkers utilized in an epigenetic editor provided herein is a non-peptide linker. For example, the linker may be a carbon bond, a disulfide bond, or carbon-heteroatom bond. In certain embodiments, the linker is a carbon-nitrogen bond of an amide linkage. In certain embodiments, the linker is a cyclic or acyclic, substituted or unsubstituted, or branched or unbranched aliphatic or heteroaliphatic linker.

[0169] In some embodiments, one or more linkers utilized in an epigenetic editor provided herein is polymeric (e.g., polyethylene, polyethylene glycol, polyamide, polyester, etc.). The linker may comprise, for example, a monomer, dimer, or polymer of aminoalkanoic acid; an aminoalkanoic acid (e.g., glycine, ethanoic acid, alanine, beta-

alanine, 3-aminopropanoic acid, 4-aminobutanoic acid, 5-pentanoic acid, etc.); a monomer, dimer, or polymer of aminohexanoic acid (Ahx); or a polyethylene glycol moiety (PEG); or an aryl or heteroaryl moiety. In certain embodiments, the linker may be based on a carbocyclic moiety (e.g., cyclopentane or cyclohexane) or a phenyl ring. The linker may include functionalized moieties to facilitate attachment of a nucleophile (e.g., thiol, amino) from the peptide to the linker. Any electrophile may be used as part of the linker. Exemplary electrophiles include, but are not limited to, activated esters, activated amides, alkyl halides, aryl halides, acyl halides, and isothiocyanates.

[0170] Various linker lengths and flexibilities can be employed between any two components of an epigenetic editor (e.g., between an effector domain (e.g., a repressor domain) and a DNA-binding domain (e.g., a Cas9 domain), between a first effector domain and a second effector domain, etc.). The linkers may range from very flexible linkers, such as glycine/serine-rich linkers, to more rigid linkers, in order to achieve the optimal length for effector domain activity for the specific application. In some embodiments, the more flexible linkers are glycine/serine-rich linkers (GS-rich linkers), where more than 45% (e.g., more than 48, 50, 55, 60, 70, 80, or 90%) of the residues are glycine or serine residues. Non-limiting examples of the GS-rich linkers are (GGGGS)n (SEQ ID NO: 664), (G)n, and W linker (SEQ ID NO: 637). In some embodiments, the more rigid linkers are in the form of the form (EAAAK)n (SEQ ID NO: 665), (SGGS)n (SEQ ID NO: 631, and (XP)n). In the aforementioned formulae of flexible and rigid linkers, n may be any integer between 1 and 30. In some embodiments, n is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15. In some embodiments, the linker comprises a (GGS)n motif, wherein n is 1, 3, or 7. In some embodiments, the linker comprises a (GGGGS)n motif, wherein n is 4 (SEQ ID NO: 636).

[0171] In some embodiments, a linker in an epigenetic editor described herein comprises a nuclear localization signal, for example, with the amino acid sequence of any one of SEQ ID NOs: 644-649. In some embodiments, a linker in an epigenetic editor described herein comprises an expression tag, e.g., a detectable tag such as a green fluorescent protein.

B. Nuclear Localization Signals

[0172] A fusion protein described herein may comprise one or more nuclear localization signals, and in certain embodiments, may comprise two or more nuclear localization signals. For example, the fusion protein may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more nuclear localization signals. As used herein, a "nuclear localization signal" (NLS) is an amino acid sequence that directs proteins to the nucleus. In certain embodiments, the NLS may be an SV40 NLS (e.g., with the amino acid sequence of SEQ ID NO: 644). The fusion protein may

comprise an NLS at its N-terminus, C-terminus, or both, and/or an NLS may be embedded in the middle of the fusion protein (e.g., at the N- or C- terminus of a DNA-binding domain or an effector domain).

[0173] In some embodiments, the fusion protein may comprise two NLSs. The fusion protein may comprise two NLSs at its N-terminus or C-terminus. The fusion protein may comprise one NLS located at its N-terminus and one NLS embedded in the middle of the fusion protein, or one NLS located at its C-terminus and one NLS embedded in the middle of the fusion protein. The fusion protein may comprise two NLSs embedded in the middle of the fusion protein.

[0174] In some embodiments, the fusion protein may comprise four NLSs. The fusion protein may comprise at least two (e.g., two, three, or four) NLSs at its N-terminus or C-terminus. The fusion protein may comprise at least one (e.g., one, two, three, or four) NLSs embedded in the middle of the fusion protein. In particular embodiments, the fusion protein may comprise two NLSs at its N-terminus and two NLSs at its C-terminus.

[0175] An NLS described herein may be an endogenous NLS sequence. In certain embodiments, an NLS described herein comprises the amino acid sequence of any one of SEQ ID NOs: 644-649, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the selected sequence. In particular embodiments, the NLS comprises the amino acid sequence of SEQ ID NO: 644. Additional NLSs are known in the art.

[0176] In some embodiments, an epigenetic editor comprising a fusion protein that comprises at least one NLS at the N-terminus and at least one NLS at the C-terminus may increase the efficiency of the epigenetic editor by at least 5%, at least 10%, at least 15%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 100%, at least 200%, at least 300%, at least 400%, at least 500%, at least 600%, at least 700%, at least 800%, at least 900%, at least 1,000%, at least 5,000%, at least 10,000%, at least 50,000%, at least 100,000%, or more as compared to an epigenetic editor with a corresponding fusion protein that does not have at least one NLS at the N-terminus and at least one NLS at the C-terminus.

[0177] In some embodiments, an epigenetic editor comprising a fusion protein that comprises two NLSs at the N-terminus and two NLSs at the C-terminus may increase the efficiency of the epigenetic editor by at least 5%, at least 10%, at least 15%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 100%, at least 200%, at least 300%, at least 400%, at least 500%, at least 600%, at least

700%, at least 800%, at least 900%, at least 1,000%, at least 5,000%, at least 10,000%, at least 50,000%, at least 100,000%, or more as compared to an epigenetic editor with a corresponding fusion protein that does not have two NLSs at the N-terminus and two NLSs at the C-terminus.

C. Tags

[0178] Epigenetic editors provided herein may comprise one or more additional sequences ("tags") for tracking, detection, and localization of the editors. In some embodiments, the epigenetic editor comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more detectable tags. Each of the detectable tags may be the same or different.

[0179] For example, an epigenetic editor fusion protein may comprise cytoplasmic localization sequences, export sequences, such as nuclear export sequences, or other localization sequences, as well as sequence tags that are useful for solubilization, purification, or detection of the fusion proteins. Suitable protein tags provided herein include, but are not limited to, biotin carboxylase carrier protein (BCCP) tags, myc-tags, calmodulin-tags, FLAG-tags, hemagglutinin (HA)-tags, poly-histidine tags (also referred to as histidine tags or His-tags), maltose binding protein (MBP)-tags, nus-tags, glutathione-S-transferase (GST)-tags, green fluorescent protein (GFP)-tags, thioredoxin-tags, S-tags, Softags (e.g., Softag 1 or Softag 3), strep-tags, biotin ligase tags, FlAsH tags, V5 tags, and SBP-tags. Additional suitable sequences will be apparent to those of skill in the art.

D. Fusion Protein Configurations

[0180] A fusion protein of an epigenetic editor described herein may have its components structured in different configurations. For example, the DNA-binding domain may be at the C-terminus, the N-terminus, or in between two or more epigenetic effector domains or additional domains. In some embodiments, the DNA-binding domain is at the C-terminus of the epigenetic editor. In some embodiments, the DNA-binding domain is at the N-terminus of the epigenetic editor. In some embodiments, the DNA-binding domain is linked to one or more nuclear localization signals. In some embodiments, the DNA-binding domain is flanked by an epigenetic effector domain and/or an additional domain on both sides. In some embodiments, where "DBD" indicates DNA-binding domain and "ED" indicates effector domain, the epigenetic editor comprises the configuration of:

- N']-[ED1]-[DBD]-[ED2]-[C']
- N']-[ED1]-[DBD]-[ED2]-[ED3]-[C']
- N']-[ED1]-[ED2]-[DBD]-[ED3]-[C']

or

- N']-[ED1]-[ED2]-DBD]-[ED3]-[ED4]-[C'.

[0181] In some embodiments, an epigenetic editor comprises a DNA-binding domain (DBD), a DNA methyltransferase (DNMT) domain, and a transcriptional repressor ("repressor") domain that represses or silences expression of a target gene. The DBD, DNMT, and transcriptional repressor domains may be any as described herein, in any combination. The DBD, DNMT domain, and repressor domain may be in any configuration, e.g., with any of said domains at the N-terminus, at the C-terminus, or in the middle of the fusion protein. In some embodiments, the epigenetic editor comprises a fusion protein with the configuration of:

N']-[DNMT domain]-[DBD]-[repressor domain]-[C'

N']-[repressor domain]-[DBD]-[DNMT domain]-[C'

N']-[DNMT domain]-[repressor domain]- [DBD]-[C'

or

N']-[repressor domain]-[DNMT domain]- [DBD]-[C'.

[0182] In some embodiments, a connecting structure "]-["in any one of the epigenetic editor structures is a linker, e.g., a peptide linker; a detectable tag; a peptide bond; a nuclear localization signal; and/or a promoter or regulatory sequence. In an epigenetic editor structure, the multiple connecting structures "]-[" may be the same or may each be a different linker, tag, NLS, or peptide bond. In some embodiments, the DNMT domain may comprise any one of the domains in **Table 6**, or any combinations or homologs thereof. In particular embodiments, the DNMT domain comprises DNMT3A or a truncated version thereof, DNMT3L or a truncated version thereof, or both. In particular embodiments, the DBD is a catalytically inactive polynucleotide guided DNA-binding domain (e.g., a dCas9) or a ZFP domain. In certain embodiments, the repressor domain comprises any one of the domains shown in **Table 4** or **5**, or any combinations or homologs thereof. For example, the repressor domain may be a KRAB domain. In certain embodiments, the repressor domain is a ZFP28, ZN627, KAP1, MeCP2, HP1b, CBX8, CDYL2, TOX, Tox3, Tox4, EED, RBBP4, RCOR1, or SCML2 domain, or a fusion of two of said domains (e.g., a fusion of the N- and C-terminal regions of ZIM3 and KOX1 KRAB). In particular embodiments, the repressor domain is a KRAB domain from ZFP28, ZN627, ZIM3, or KOX1.

[0183] In some embodiments, the epigenetic editor comprises a configuration selected from

N']-[DNMT3A-DNMT3L]-[DBD]-[repressor]-[C'

N']-[repressor]-[DBD]-[DNMT3A-DNMT3L]-[C'

N']-[repressor]-[DBD]-[DNMT3A]-[C']

N']-[DNMT3A]-[DBD]-[repressor]-[C']

N']-[repressor]-[DBD]-[DNMT3A]-[DNMT3L]-[C']

N']-[DNMT3A]- [DNMT3L]-[DBD]-[repressor]-[C']

N']-[DNMT3A]-[DBD]-[C']

N']-[DBD]-[DNMT3A]-[C']

N']-[DNMT3L]-[DBD]-[C']

N']-[DBD]-[DNMT3L]-[C']

wherein [DNMT3A-DNMT3L] indicates that the DNMT3A and DNMT3L domains are directly fused via a peptide bond, and wherein the connecting structure]-[is any one of the linkers as described herein, a detectable tag, an affinity domain, a peptide bond, a nuclear localization signal, a promoter, and/or a regulatory sequence. The DBD, repressor, DNMT3A, and DNMT3L domains may be any as described herein, in any combination. For example, the DNMT3A and DNMT3L domains may be selected from those in **Table 6**. In particular embodiments, the DBD is a CRISPR-associated protein domain (e.g., dCas9) or a ZFP domain; the repressor domain is a KRAB domain derived from KOX1, ZIM3, ZFP28, or ZN627; the DNMT3A domain is a human DNMT3A domain; and the DNMT3L domain is a human or mouse DNMT3L domain; any combination of these components is also contemplated by the present disclosure.

[0184] In some embodiments, the epigenetic editor comprises a configuration selected from

N']-[DNMT3A]-[DBD]-[SETDB1]-[C']

N']-[DNMT3A]- [DNMT3L]- [DBD]-[SETDB1]-[C']

N']- [DNMT3A-DNMT3L]- [DBD]- [SETDB1]- [C']

N']-[SETDB1]-[DBD]-[DNMT3A]-[DNMT3L]-[C']

N']-[SETDB1]-[DBD]-[DNMT3A]-[C']

wherein [DNMT3A-DNMT3L] indicates that the DNMT3A and DNMT3L domains are directly fused via a peptide bond, and wherein the connecting structure]-[is any one of the linkers as described herein, a detectable tag, an affinity domain, a peptide bond, a nuclear localization signal, a promoter, and/or a regulatory sequence. The DBD, SETDB1, DNMT3A, and DNMT3L domains may be any as described herein, in any combination. In particular embodiments, the DBD is a CRISPR-associated protein domain (e.g., dCas9) or a ZFP domain; the SETDB1 domain is derived from human SETDB1, ZIM3, ZFP28, or ZN627; the DNMT3A domain is a human DNMT3A domain; and the DNMT3L domain is a

human or mouse DNMT3L domain; any combination of these components is also contemplated by the present disclosure.

[0185] Particular constructs contemplated herein include:

DNMT3A-DNMT3L-XTEN80-NLS-**dCas9**-NLS-XTEN16-KOX1 KRAB (Configuration 1),

DNMT3A-DNMT3L-XTEN80-NLS-**ZFP domain**-NLS-XTEN16-KOX1 KRAB (Configuration 2),

NLS-DNMT3A-DNMT3L-XTEN80-**dCas9**-XTEN16-KOX1 KRAB-NLS (Configuration 3),

NLS-DNMT3A-DNMT3L-XTEN80-**ZFP domain**-XTEN16-KOX1 KRAB-NLS (Configuration 4),

NLS-NLS-DNMT3A-DNMT3L-XTEN80-**dCas9**-XTEN16-KOX1 KRAB-NLS-NLS (Configuration 5), and

NLS-NLS-DNMT3A-DNMT3L-XTEN80-**ZFP domain**-XTEN16-KOX1 KRAB-NLS-NLS (Configuration 6).

The DNMT3L and DNMT3A may be derived from human parental proteins, mouse parental proteins, or any combination thereof. In certain embodiments, the DNMT3L and DNMT3A are derived from mouse and human parental proteins, respectively (mDNMT3L and hDNMT3A). In certain embodiments, the DNMT3L and DNMT3A are both derived from human parental proteins (hDNMT3L and hDNMT3A). In some embodiments, the dCas9 is dSpCas9. In some embodiments, the KOX1 is human KOX1. Also contemplated is any of Configurations 1-6 wherein the KOX1 KRAB domain is replaced by a ZFP28, ZN627, or ZIM3 KRAB domain. In some embodiments, the ZFP28, ZN627, and ZIM3 are human ZFP28, ZN627, and ZIM3, respectively. In particular embodiments, the fusion construct may have the configuration:

NLS-NLS-hDNMT3A-hDNMT3L-XTEN80-**dCas9**-XTEN16-KOX1 KRAB-NLS-NLS (Configuration 7),

NLS-NLS-DNMT3A-DNMT3L-XTEN80-**ZFP domain**-XTEN16-KOX1 KRAB-NLS-NLS (Configuration 8),

NLS-NLS-hDNMT3A-hDNMT3L-XTEN80-**dCas9**-XTEN16-ZFP28 KRAB-NLS-NLS (Configuration 9),

NLS-NLS-DNMT3A-DNMT3L-XTEN80-**ZFP domain**-XTEN16-ZFP28 KRAB-NLS-NLS (Configuration 10),

NLS-NLS-hDNMT3A-hDNMT3L-XTEN80-**dCas9**-XTEN16-ZN627 KRAB-NLS-NLS (Configuration 11),

NLS-NLS-DNMT3A-DNMT3L-XTEN80-**ZFP domain**-XTEN16-ZN627 KRAB-NLS-NLS (Configuration 12),

NLS-NLS-hDNMT3A-hDNMT3L-XTEN80-**dCas9**-XTEN16-ZIM3 KRAB-NLS-NLS (Configuration 13), or

NLS-NLS-DNMT3A-DNMT3L-XTEN80-**ZFP domain**-XTEN16-ZIM3 KRAB-NLS-NLS (Configuration 14).

[0186] In particular embodiments, a fusion construct described herein may have Configuration 1 and comprise SEQ ID NO: 658, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical thereto. In SEQ ID NO: 658 below, the XTEN linkers are underlined, the W linker is ***bolded, underlined, and italicized***, the NLS sequences are **bolded**, the DNMT3A sequence is *italicized*, the DNMT3L sequence is *underlined and italicized*, the dCas9 domain is ***bolded and italicized***, and the KOX1 KRAB domain is **underlined and bolded**:

MNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE
DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLY
EGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMID
AKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSN
SIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSV
*PVIRHLFAPLKEYFACV**SSGNSNANSRGPSFSSGLVPLSLRGSH**MGPMEIYKTVSAW*
KRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDL
VYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQESQRPFFWIFMDNLLLTEDDQET
TTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEEYLQAQVRSRSK
LDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGPSSGAPPPSGGSPAGSPTSTEEGT
SESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSE**PKK**
KRKV**Y**MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALL
FDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEE
DKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRG
HFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLE
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLA
QIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKA
LVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN
REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYY
VGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVL
PKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLK
EDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLT
LTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTI
LDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKG
ILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGS
QILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVS

DFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKM
IAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR
DFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD
SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSP
EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQA
ENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLS
*QLGGD*PKKKRKV<u>SGSETPGTSESATPESTG</u>**<u>RTLVTFKDVFVDFTREEWKLLDTAQQI</u>**
<u>VYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEP</u> (SEQ ID NO: 658)

[0187] In particular embodiments, a fusion construct described herein may comprise the sequence provided below (SEQ ID NO: 1495), or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical thereto. In SEQ ID NO: 1495 below, the XTEN linkers are <u>underlined</u>, the W linker is ***<u>bolded, underlined, and italicized</u>***, the NLS sequences are **bolded**, the DNMT3A sequence is *italicized*, the DNMT3L sequence is *<u>underlined and italicized</u>*, the dCas9 domain is ***bolded and italicized***, and the KOX1 KRAB domain is **<u>underlined and bolded</u>**:

MNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE
DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLY
EGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMID
AKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSN
SIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSV
*PVIRHLFAPLKEYFACV**<u>SSGNSNANSRGPSFSSGLVPLSLRGSH</u>**MNPLEMFETVPVW*
RRQPVRVLSLFEDIKKELTSLGFLESGSDPGQLKHVVDVTDTVRKDVEEWGPFDLVY
GATPPLGHTCDRPPSWYLFQFHRLLQYARPKPGSPRPFFWMFVDNLVLNKEDLDVAS
RFLEMEPVTIPDVHGGSLQNAVRVWSNIPAIRSRHWALVSEEELSLLAQNKQSSKLA
*AKWPTKLVKNCFLPLREYFKYFSTELTSSL*GGPSSGAPPPSGGSPAGSPTSTEEGTS
ESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSE**PKKK**
RKVY***MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF***
DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGH
FLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLEN
LIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQ
IGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKAL
VRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNR
EDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYV
GPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLP
KHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE
DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTL
TLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTIL
DFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI
LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ
ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDS
IDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL
SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSD
FRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMI

AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRD
FATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDS
PTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKD
LIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPE
DNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAE
NIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQ
*LGGD*PKKKRKVSGSETPGTSESATPESTG**RTLVTFKDVFVDFTREEWKLLDTAQQIV**
YRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEP (SEQ ID NO: 1495)

[0188] In particular embodiments, a fusion construct described herein may have Configuration 2 and comprise SEQ ID NO: 659, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical thereto. In SEQ ID NO: 659 below, the XTEN linkers are underlined, the W linker is ***bolded, underlined, and italicized***, the NLS sequences are **bolded and underlined**, the DNMT3A sequence is *italicized*, the DNMT3L sequence is *underlined and italicized*, the ZFP domain is **bolded**, and the KOX1 KRAB domain is **underlined and bolded**. Variable amino acids represented by Xs are the amino acids of the DNA-recognition helix of the zinc finger and ***XX*** in italics may be either TR, LR or LK.

MNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE
DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLY
EGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMID
AKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSN
SIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSV
*PVIRHLFAPLKEYFACV**SSGNSNANSRGPSFSSGLVPLSLRGSH**MGPMEIYKTVSAW*
KRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDL
VYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQESQRPFFWIFMDNLLLTEDDQET
TTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEEYLQAQVRSRSK
*LDAPKVDLLVKNCLLPLREYFKYFSQNSLPL*GGPSSGAPPPSGGSPAGSPTSTEEGT
SESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSE**PKK**
KRKVYSRPGERPFQCRICMRNFSXXXXXXXHXXTHTGEKPFQCRICMRNFSXXXXXX
XHXXTH[linker]PFQCRICMRNFSXXXXXXXHXXTHTGEKPFQCRICMRNFSXXX
XXXXHXXTH[linker]PFQCRICMRNFSXXXXXXXHXXTHTGEKPFQCRICMRNFS
XXXXXXXHXXTHLRGS**PKKKRKV**SGSETPGTSESATPESTG**RTLVTFKDVFVDFTRE**
EWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEP (SEQ ID NO: 659)

In certain embodiments, the six "XXXXXXX" regions in SEQ ID NO: 659 comprise, in order, the F1-F6 amino acid sequences shown in **Table 1** for any one of ZF001-ZF048. [linker] represents a linker sequence. In some embodiments, one or both linker sequences may be TGSQKP (SEQ ID NO: 651). In some embodiments, one or both linker sequences may be TGGGGSQKP (SEQ ID NO: 652). In some embodiments, one linker sequence may

have the amino acid sequence of SEQ ID NO: 651 and the other linker sequence may have the amino acid sequence of SEQ ID NO: 652.

[0189] In particular embodiments, a fusion construct described herein may comprise the sequence provided below (SEQ ID NO: 1496), or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical thereto. In SEQ ID NO: 1496 below, the XTEN linkers are underlined, the W linker is ***bolded, underlined, and italicized***, the NLS sequences are **bolded and underlined**, the DNMT3A sequence is *italicized*, the DNMT3L sequence is *underlined and italicized*, the ZFP domain is **bolded**, and the KOX1 KRAB domain is **underlined and bolded**. Variable amino acids represented by Xs are the amino acids of the DNA-recognition helix of the zinc finger and ***XX*** in italics may be either TR, LR or LK.

*MNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLY EGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMID AKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSN SIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSV PVIRHLFAPLKEYFACV**SSGNSNANSRGPSFSSGLVPLSLRGSH**MNPLEMFETVPVW RRQPVRVLSLFEDIKKELTSLGFLESGSDPGQLKHVVDVTDTVRKDVEEWGPFDLVY GATPPLGHTCDRPPSWYLFQFHRLLQYARPKPGSPRPFFWMFVDNLVLNKEDLDVAS RFLEMEPVTIPDVHGGSLQNAVRVWSNIPAIRSRHWALVSEEELSLLAQNKQSSKLA AKWPTKLVKNCFLPLREYFKYFSTELTSSLGGPSSGAPPPSGGSPAGSPTSTEEGTS ESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSE**PKKK RKVYSRPGERPFQCRICMRNFSXXXXXXXHXXTHTGEKPFQCRICMRNFSXXXXXXX HXXTH[linker]PFQCRICMRNFSXXXXXXXHXXTHTGEKPFQCRICMRNFSXXXX XXXHXXTH[linker]PFQCRICMRNFSXXXXXXXHXXTHTGEKPFQCRICMRNFSX XXXXXXHXXTHLRGSPKKKRKV**SGSETPGTSESATPESTG**RTLVTFKDVFVDFTREE WKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEP** (SEQ ID NO: 1496)

[0190] In certain embodiments, the six "XXXXXXX" regions in SEQ ID NO: 1496 comprise, in order, the F1-F6 amino acid sequences shown in **Table 1** for any one of ZF001-ZF048. [linker] represents a linker sequence. In some embodiments, one or both linker sequences may be TGSQKP (SEQ ID NO: 651). In some embodiments, one or both linker sequences may be TGGGGSQKP (SEQ ID NO: 652). In some embodiments, one linker sequence may have the amino acid sequence of SEQ ID NO: 651 and the other linker sequence may have the amino acid sequence of SEQ ID NO: 652.

[0191] In some embodiments, the fusion protein may further comprise a Dnmt3A ADD domain, e.g., downstream of the Dnmt3A domain sequence disclosed in SEQ ID Nos 658, 659, 1495, or 1496 disclosed above. In some embodiments, the ADD sequence is situated

between the Dnmt3A and the Dnmt3L sequence of the fusion protein. In some embodiments, the ADD sequence is at the C-terminal end of the Dnmt3A domain. In some embodiments, the Dnmt3A sequence and the ADD sequence are separated by a linker, e.g., a linker disclosed herein. In some embodiments, the ADD sequence and the Dnmt3L sequence are separated by a linker, e.g., a linker disclosed herein. In some embodiments, the ADD domain comprises the sequence:

MAAIPALDPEAEPSMDVILVGSSELSSSVSPGTGRDLIAYEVKANQRNIEDICICCG
SLQVHTQHPLFEGGICAPCKDKFLDALFLYDDDGYQSYCSICCSGETLLICGNPDCT
RCYCFECVDSLVGPGTSGKVHAMSNWVCYLCLPSSRSGLLQRRRKWRSQLKAFYDRE
SENPLEMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSDPGQLKHVVDVTDTV
RKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLLQYARPKPGSPRPFFWMFV
DNLVLNKEDLDVASRFLEMEPVTIPDVHGGSLQNAVRVWSNIPAIRSRHWALVSEEE
LSLLAQNKQSSKLAAKWPTKLVKNCFLPLREYFKYFSTELTSSL (SEQ ID NO:1497).

[0192] In particular embodiments, a fusion construct described herein may have Configuration 7 and comprise SEQ ID NO: 660, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical thereto.

[0193] In particular embodiments, a fusion construct described herein may have Configuration 9 and comprise SEQ ID NO: 661, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical thereto.

[0194] In particular embodiments, a fusion construct described herein may have Configuration 11 and comprise SEQ ID NO: 662, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical thereto.

[0195] In particular embodiments, a fusion construct described herein may have Configuration 13 and comprise SEQ ID NO: 663, or a sequence at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical thereto.

[0196] In some embodiments, a fusion construct described herein (e.g., the fusion construct of any one of Configurations 1-14) is within an expression construct that comprises a WPRE sequence, a polyadenylation site, or both. In certain embodiments, the WPRE sequence is in a 3' noncoding region. In certain embodiments, the WPRE sequence is upstream from a poly-adenylation site. In particular embodiments, the expression construct comprises the fusion construct (e.g., of any one of Configurations 1-14) and a WPRE sequence in a 3' noncoding region upstream from a polyadenylation site.

[0197] In some embodiments, a fusion construct described herein may have the sequence of any one of Fusion Proteins 1-12 as shown in Example 12.

[0198] Multiple fusion proteins may be used to effect activation or repression of a target gene or multiple target genes. For example, an epigenetic editor fusion protein comprising a DNA-binding domain (e.g., a dCas9 domain) and an effector domain may be co-delivered with two or more guide polynucleotides (e.g., gRNAs), each targeting a different target DNA sequence. The target sites for two of the DNA-binding domains may be the same or in the vicinity of each other, or separated by, for example, about 100 base pairs, about 200 base pairs, about 300 base pairs, about 400 base pairs, about 500 base pairs, or about 600 or more base pairs. In addition, when targeting double-strand DNA, such as an endogenous gene locus, the guide polynucleotides may target the same or different strands (one or more to the positive strand and/or one or more to the negative strand).

V. Target Sequences

[0199] An epigenetic editor herein may be directed to a target sequence in *PCSK9* to effect epigenetic modification of the *PCSK9* gene. As used herein, a "target sequence," a "target site," or a "target region" is a nucleic acid sequence present in a gene of interest; in some instances, the target sequence may be outside but in the vicinity of the gene of interest wherein methylation or binding by a repressor of the target sequence represses expression of the gene. In some embodiments, the target sequence may be a hypomethylated or hypermethylated nucleic acid sequence.

[0200] The target sequence may be in any part of a target gene. In some embodiments, the target sequence is part of or near a noncoding sequence of the gene. In some embodiments, the target sequence is part of an exon of the gene. In some embodiments, the target sequence is part of or near a transcriptional regulatory sequence of the gene, such as a promoter or an enhancer. In some embodiments, the target sequence is adjacent to, overlaps with, or encompasses a CpG island. In certain embodiments, the target sequence is within about 3000, 2900, 2800, 2700, 2600, 2500, 2400, 2300, 2200, 2100, 2000, 1900, 1800, 1700, 1600, 1500, 1400, 1300, 1200, 1100, 1000, 900, 800, 700, 600, 500, 400, 300, 200, or 100 base pairs (bp) flanking a *PCSK9* TSS. In certain embodiments, the target sequence is within 500 bp flanking the *PCSK9* TSS. In certain embodiments, the target sequence is within 1000 bp flanking the *PCSK9* TSS.

[0201] In some embodiments, the target sequence may hybridize to a guide polynucleotide sequence (e.g., gRNA) complexed with a fusion protein comprising a polynucleotide guided DNA-binding domain (e.g., a CRISPR protein such as dCas9) and effector domain(s). The guide polynucleotide sequence may be designed to have

complementarity to the target sequence, or identity to the opposing strand of the target sequence. In some embodiments, the guide polynucleotide comprises a spacer sequence that is about 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to a protospacer sequence in the target sequence. In particular embodiments, the guide polynucleotide comprises a spacer sequence that is 100% identical to a protospacer sequence in the target sequence.

[0202] In some embodiments, where the DNA-binding domain of an epigenetic editor described herein is a zinc finger array, the target sequence may be recognized by said zinc finger array.

[0203] In some embodiments, where the DNA-binding domain of an epigenetic editor described herein is a TALE, the target sequence may be recognized by said TALE.

[0204] A target sequence described herein may be specific to one copy of a target gene, or may be specific to one allele of a target gene. Accordingly, the epigenetic modification and modulation of expression thereof may be specific to one copy or one allele of the target gene. For example, an epigenetic editor may repress expression of a specific copy harboring a target sequence recognized by the DNA-binding domain (e.g., a copy associated with a disease or condition, or that harbors a mutation associated with a disease or condition).

[0205] In some embodiments, the target *PCSK9* genomic region may fall within the sequence shown below (chr1:55038548-55040548), with or without the terminal A:

TACCTCATGGAGTCACTGTCAACCCACTGGTTGCACTGTCTTTGTGCACTGGCTCTC
TGGAGTGAGGTCTTTGCAAACAAAGTGGAAAGAGCATCAACTTTGGACTCCAGCACC
TAGATTCAGAGCAGGCCATTTCACTCGGAATCTGCTGTGCATCTGCAAGGGAGGATC
ATAAATTCGCCTTTGTTTCTTCCCAGTATCGACAGCCCTTCCAGAAAGAGCAAGCCT
CATGTCATGCCACATGTACAATCTGAGGCCAGGAGCTCTCTTTCCCCTTTTCATCCT
CCTGCCTGGTACACAATAGGTGTTTACTGGATGCTTGTCCAGTTGATTTCTTGAACA
TGGTGTGTAAAAGGAATCTTTGCAAATTGAATCTTCTGGAAAGCTGAGCTTGTGCCT
ACCATAGAATTCTGAATGTACCTATATGACGTCTTTGCAAACTTAAAACCTGAATCT
TTGTAGTATAAATCCCTTGAAATGCATGTAGGCTGGACATCAAAAGCAAGCAATCTC
TTCAAGGAGCAGCTAGTTGGTAAGGTCAGTGTGCAGGGTGCATAAAGGGCAGAGGCC
GGAGGGGGTCCAGGCTAAGTTTAGAAGGCTGCCAGGTTAAGGCCAGTGGAAAGAATT
CGGTGGGCAGCGAGGAGTCCACAGTAGGATTGATTCAGAAGTCTCACTGGTCAGCAG
GAGACAAGGTGGACCCAGGAAACACTGAAAAGGTGGGCCCGGCAGAACTTGGAGTCT
GGCATCCCACGCAGGGTGAGAGGCGGGAGAGGAGGAGCCCCTAGGGCGCCGGCCTGC
CTTCCAGCCCAGTTAGGATTTGGGAGTTTTTTCTTCCCTCTGCGCGTAATCTGACGC
TGTTTGGGGAGGGCGAGGCCGAAACCTGATCCTCCAGTCCGGGGGTTCCGTTAATGT
TTAATCAGATAGGATCGTCCGATGGGGCTCTGGTGGCGTGATCTGCGCGCCCCAGGC
GTCAAGCACCCACACCCTAGAAGGTTTCCGCAGCGACGTCGAGGCGCTCATGGTTGC
AGGCGGGCGCCGCCGTTCAGTTCAGGGTCTGAGCCTGGAGGAGTGAGCCAGGCAGTG
AGACTGGCTCGGGCGGGCCGGGACGCGTCGTTGCAGCAGCGGCTCCCAGCTCCCAGC
CAGGATTCCGCGCGCCCCTTCACGCGCCCTGCTCCTGAACTTCAGCTCCTGCACAGT
CCTCCCCACCGCAAGGCTCAAGGCGCCGCCGGCGTGGACCGCGCACGGCCTCTAGGT
CTCCTCGCCAGGACAGCAACCTCTCCCCTGGCCCTCATGGGCACCGTCAGCTCCAGG

CGGTCCTGGTGGCCGCTGCCACTGCTGCTGCTGCTGCTGCTCCTGGGTCCCGCG
GGCGCCCGTGCGCAGGAGGACGAGGACGGCGACTACGAGGAGCTGGTGCTAGCCTTG
CGTTCCGAGGAGGACGGCCTGGCCGAAGCACCCGAGCACGGAACCACAGCCACCTTC
CACCGCTGCGCCAAGGTGCGGGTGTAGGGATGGGAGGCCGGGGCGAACCCGCAGCCG
GGACGGTGCGGTGCTGTTTCCTCTCGGGCCTCAGTTTCCCCCCATGTAAGAGAGGAA
GTGGAGTGCAGGTCGCCGAGGGCTCTTCGCTTGGCACGATCTTGGGGACTGCAGGCA
AGGCGGCGGGGGAGGACGGGTAGTGGGGAGCACGGTGGAGAGCGGGGACGGCCGGCT
CTTTGGGGACTTGCTGGGGCGTGCGGCTGCGCTATTCAGTGGGAAGGTTCGCGGGGT
TGGGAGACCCGGAGGCCGAGGAAGGGCGAGCAGAGCACTGCCAGGATATCCTGCCCA
GATTTCCCAGTTTCTGCCTCGCCGCGGCACAGGTGGGTGAAGGAGTGAATGCCTGGA
ACGTACTGGGAACTGCACCAGGCACAGAGAAAGCGGGCTTGCCATTATAGTGGGTTC
CGATTTGGTTTGGAAAACATGGGCAGCGGAGGGTGGAGGGCCTGGAGAGAAGGCCCT
ACCCGA (SEQ ID NO: 1488)

[0206] In some embodiments, the target sequence may be GRCh38 Chr1:55039228-55040296, as shown below:

GCAGGAGACAAGGTGGACCCAGGAAACACTGAAAAGGTGGGCCCGGCAGAACTTGGA
GTCTGGCATCCCACGCAGGGTGAGAGGCGGGAGAGGAGGAGCCCCTAGGGCGCCGGC
CTGCCTTCCAGCCCAGTTAGGATTTGGGAGTTTTTTCTTCCCTCTGCGCGTAATCTG
ACGCTGTTTGGGGAGGGCGAGGCCGAAACCTGATCCTCCAGTCCGGGGGTTCCGTTA
ATGTTTAATCAGATAGGATCGTCCGATGGGGCTCTGGTGGCGTGATCTGCGCGCCCC
AGGCGTCAAGCACCCACACCCTAGAAGGTTTCCGCAGCGACGTCGAGGCGCTCATGG
TTGCAGGCGGGCGCCGCCGTTCAGTTCAGGGTCTGAGCCTGGAGGAGTGAGCCAGGC
AGTGAGACTGGCTCGGGCGGGCCGGGACGCGTCGTTGCAGCAGCGGCTCCCAGCTCC
CAGCCAGGATTCCGCGCGCCCCTTCACGCGCCCTGCTCCTGAACTTCAGCTCCTGCA
CAGTCCTCCCCACCGCAAGGCTCAAGGCGCCGCCGGCGTGGACCGCGCACGGCCTCT
AGGTCTCCTCGCCAGGACAGCAACCTCTCCCCTGGCCCTCATGGGCACCGTCAGCTC
CAGGCGGTCCTGGTGGCCGCTGCCACTGCTGCTGCTGCTGCTGCTGCTCCTGGGTCC
CGCGGGCGCCCGTGCGCAGGAGGACGAGGACGGCGACTACGAGGAGCTGGTGCTAGC
CTTGCGTTCCGAGGAGGACGGCCTGGCCGAAGCACCCGAGCACGGAACCACAGCCAC
CTTCCACCGCTGCGCCAAGGTGCGGGTGTAGGGATGGGAGGCCGGGGCGAACCCGCA
GCCGGGACGGTGCGGTGCTGTTTCCTCTCGGGCCTCAGTTTCCCCCCATGTAAGAGA
GGAAGTGGAGTGCAGGTCGCCGAGGGCTCTTCGCTTGGCACGATCTTGGGGACTGCA
GGCAAGGCGGCGGGGGAGGACGGGTAGTGGGGAGCACGGTGGAGAGCGGGGACGGCC
GGCTCTTTGGGGACTTGCTGGGGCGTGCGGCTGCGCTATTCAG (SEQ ID NO: 1489)

VI. Epigenetic Modifications

[0207] An epigenetic editor described herein may perform sequence-specific epigenetic modification(s) (e.g., alteration of chemical modification(s)) of a target gene that harbors the target sequence. Such epigenetic modulation may be safer and more easily reversible than modulation due to gene editing, e.g., with generation of DNA double-strand breaks. In some embodiments, the epigenetic modulation may reduce or silence the target gene. In some embodiments, the modification is at a specific site of the target sequence. In some embodiments, the modification is at a specific allele of the target gene. Accordingly, the epigenetic modification may result in modulated (e.g., reduced) expression of one copy of a

target gene harboring a specific allele, and not the other copy of the target gene. In some embodiments, the specific allele is associated with a disease, condition, or disorder.

[0208] In some embodiments, the epigenetic modification reduces or abolishes transcription of the target gene harboring the target sequence. In some embodiments, the epigenetic modification reduces or abolishes transcription of a copy of the target gene harboring a specific allele recognized by the epigenetic editor. In some embodiments, the epigenetic editor reduces the level of or eliminates expression of a protein encoded by the target gene. In some embodiments, the epigenetic editor reduces the level of or eliminates expression of a protein encoded by a copy of the target gene harboring a specific allele recognized by the epigenetic editor. The target *PCSK9* gene may be epigenetically modified *in vitro*, *ex vivo*, or *in vivo*.

[0209] The effector domain of an epigenetic editor described herein may alter (e.g., deposit or remove) a chemical modification at a nucleotide of the target gene or at a histone associated with the target gene. The chemical modification may be altered at a single nucleotide or a single histone, or may be altered at 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000 or more nucleotides.

[0210] In some embodiments, an effector domain of an epigenetic editor described herein may alter a CpG dinucleotide within the target gene. In some embodiments, all CpG dinucleotides within 2000, 1500, 1000, 500, or 200 bps flanking a target sequence (e.g., in an alteration site as described herein) are altered according to a modification type described herein, as compared to the original state of the gene or the gene in a comparable cell not contacted with the epigenetic editor. In some embodiments, at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700 or more of the CpG dinucleotides are altered as compared to the original state of the gene or the gene in a comparable cell not contacted with the epigenetic editor. In some embodiments, at least 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% of the CpG dinucleotides are altered as compared to the original state of the gene or the gene in a comparable cell not contacted with the epigenetic editor. In some embodiments, one single CpG dinucleotide is altered, as compared to the original state of the gene or the gene in a comparable cell not contacted with the epigenetic editor.

[0211] An effector domain of an epigenetic editor described herein may alter a histone modification state of a histone associated with or bound to the target gene. For example, an

effector domain may deposit a modification on one or more lysine residues of histone tails of histones associated with the target gene. In some embodiments, the effector domain may result in deacetylation of one or more histone tails of histones associated with the target gene, thereby reducing or silencing expression of the target gene. In some embodiments, the histone modification state is a methylation state. For example, the effector domain may result in a H3K9, H3K27 or H4K20 methylation (e.g. one or more of a H3K9me2, H3K9me3, H3K27me2, H3K27me3, and H4K20me3 methylation) at one or more histone tails associated with the target gene, thereby reducing or silencing expression of the target gene.

[0212] In some embodiments, all histone tails of histones bound to DNA nucleotides within 2000, 1500, 1000, 500, or 200 bps flanking the target sequence are altered according to a modification type as described herein, as compared to the original state of the chromosome or the chromosome in a comparable cell not contacted with the epigenetic editor. In some embodiments, at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120 or more histone tails of the bound histones are altered as compared to the original state of the chromosome or the chromosome in a comparable cell not contacted with the epigenetic editor. In some embodiments, at least 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% of histone tails of the bound histones are altered as compared to the original state of the chromosome or the chromosome in a comparable cell not contacted with the epigenetic editor. For example, one single histone tail of the bound histones may be altered as compared to the original state of the chromosome or the chromosome in a comparable cell not contacted with the epigenetic editor. As another example, one single bound histone octamer may be altered as compared to the original state of the chromosome or the chromosome in a comparable cell not contacted with the epigenetic editor.

[0213] The chemical modification deposited at target gene DNA nucleotides or histone residues may be at or in close proximity to a target sequence in the target gene. In some embodiments, an effector domain of an epigenetic editor described herein alters a chemical modification state of a nucleotide or histone tail bound to a nucleotide 100-200, 200-300, 300-400, 400-55, 500-600, 600-700, or 700-800 nucleotides 5' or 3' to the target sequence in the target gene. In some embodiments, an effector domain alters a chemical modification state of a nucleotide or histone tail bound to a nucleotide within 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, or 2000 nucleotides flanking the target sequence. As used herein,

"flanking" refers to nucleotide positions 5' to the 5' end of and 3' to the 3' end of a particular sequence, e.g. a target sequence.

[0214] In some embodiments, an effector domain mediates or induces a chemical modification change of a nucleotide or a histone tail bound to a nucleotide distant from a target sequence. Such modification may be initiated near the target sequence, and may subsequently spread to one or more nucleotides in the target gene distant from the target sequence. For example, an effector domain may initiate alteration of a chemical modification state of one or more nucleotides or one or more histone residues bound to one or more nucleotides within 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500 nucleotides flanking the target sequence, and the chemical modification state alteration may spread to one or more nucleotides at least 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2500, 3000, or more nucleotides from the target sequence in the target gene, either upstream or downstream of the target sequence. In certain embodiments, the chemical modification may be initiated at less than 2, 3, 5, 10, 20, 30, 40, 50, or 100 nucleotides in the target gene and spread to at least 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, or more nucleotides in the target gene. In some embodiments, the chemical modification spreads to nucleotides in the entire target gene. Additional proteins or transcription factors, for example, transcription repressors, methyltransferases, or transcription regulation scaffold proteins, may be involved in the spreading of the chemical modification. Alternatively, the epigenetic editor alone may be involved.

[0215] In some embodiments, an epigenetic editor described herein reduces expression of a target gene by 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%, at least about 90%, at least about 95%, at least about 99%, or more, as measured by transcription of the target gene in a cell, a tissue, or a subject as compared to a control cell, control tissue, or a control subject (e.g., in the absence of the epigenetic editor). In some embodiments, the epigenetic editors described herein reduces expression of a copy of target gene by 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%, at least about 90%, at least about 95%, at least about 99%, or more, as measured by transcription of the copy of the target gene in a cell, a tissue, or a subject as compared to a control cell, control tissue, or a control subject. In certain embodiments, the copy of the target gene harbors a specific sequence or allele recognized by the epigenetic editor. In particular embodiments, the epigenetically modified copy encodes a functional

protein, and accordingly an epigenetic editor disclosed herein may reduce or abolish expression and/or function of the protein. For example, an epigenetic editor described herein may reduce expression and/or function of a protein encoded by the target gene by at least 3-fold, at least 4-fold, at least 5-fold, at least 6-fold, at least 7-fold, at least 8-fold, at least 9-fold, at least 10-fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 20-fold, at least 25-fold, at least 30-fold, at least 35-fold, at least 40-fold, at least 45-fold, at least 50-fold, at least 60-fold, at least 70-fold, at least 80-fold, at least 90-fold, or at least 100 fold in a cell, a tissue, or a subject as compared to a control cell, control tissue, or a control subject.

[0216] Modulation of target gene expression can be assayed by determining any parameter that is indirectly or directly affected by the expression of the target gene. Such parameters include, e.g., changes in RNA or protein levels; changes in protein activity; changes in product levels; changes in downstream gene expression; changes in transcription or activity of reporter genes such as, for example, luciferase, CAT, beta-galactosidase, or GFP; changes in signal transduction; changes in phosphorylation and dephosphorylation; changes in receptor-ligand interactions; changes in concentrations of second messengers such as, for example, cGMP, cAMP, IP3, and $Ca2^{+}$; changes in cell growth; changes in neovascularization; and/or changes in any functional effect of gene expression. Measurements can be made in vitro, in vivo, and/or ex vivo, and can be made by conventional methods, e.g., measurement of RNA or protein levels, measurement of RNA stability, and/or identification of downstream or reporter gene expression. Readout can be by way of, for example, chemiluminescence, fluorescence, colorimetric reactions, antibody binding, inducible markers, ligand binding assays, changes in intracellular second messengers such as cGMP and inositol triphosphate (IP3), changes in intracellular calcium levels; cytokine release, and the like.

[0217] Methods for determining the expression level of a gene, for example the target of an epigenetic editor, may include, e.g., determining the transcript level of a gene by reverse transcription PCR, quantitative RT-PCR, droplet digital PCR (ddPCR), Northern blot, RNA sequencing, DNA sequencing (e.g., sequencing of complementary deoxyribonucleic acid (cDNA) obtained from RNA); next generation (Next-Gen) sequencing, nanopore sequencing, pyrosequencing, or Nanostring sequencing. Levels of protein expressed from a gene may be determined, e.g., by Western blotting, enzyme linked immuno-absorbance assays, mass-spectrometry, immunohistochemistry, or flow cytometry analysis. Gene expression product

levels may be normalized to an internal standard such as total messenger ribonucleic acid (mRNA) or the expression level of a particular gene, e.g., a housekeeping gene.

[0218] In some embodiments, the effect of an epigenetic editor in modulating target gene expression may be examined using a reporter system. For example, an epigenetic editor may be designed to target a reporter gene encoding a reporter protein, such as a fluorescent protein. Expression of the reporter gene in such a model system may be monitored by, e.g., flow cytometry, fluorescence-activated cell sorting (FACS), or fluorescence microscopy. In some embodiments, a population of cells may be transfected with a vector that harbors a reporter gene. The vector may be constructed such that the reporter gene is expressed when the vector transfects a cell. Suitable reporter genes include genes encoding fluorescent proteins, for example green, yellow, cherry, cyan or orange fluorescent proteins. The population of cells carrying the reporter system may be transfected with DNA, mRNA, or vectors encoding the epigenetic editor targeting the reporter gene.

VII. Pharmaceutical Compositions

[0219] In one aspect, the present disclosure provides a pharmaceutical composition comprising as an active ingredient (or as the sole active ingredient) one or more epigenetic editors described herein or component(s) (e.g., fusion proteins and/or guide polynucleotides) thereof, or nucleic acid molecule(s) encoding said epigenetic editors or component(s) thereof. For example, a pharmaceutical composition may comprise nucleic acid molecule(s) encoding the fusion protein(s) (and guide polynucleotides, where applicable) of an epigenetic editor described herein. In some embodiments, separate pharmaceutical compositions comprise the fusion protein(s) and the guide polynucleotide(s). A pharmaceutical composition may also comprise cells that have undergone epigenetic modification(s) mediated or induced by an epigenetic editor provided herein.

[0220] Generally, the epigenetic editors described herein or component(s) thereof, or nucleic acid molecule(s) encoding said epigenetic editors or component(s) thereof, of the present disclosure are suitable to be administered as a formulation in association with one or more pharmaceutically acceptable excipient(s), e.g., as described below.

[0221] The term “excipient” is used herein to describe any ingredient other than the compound(s) of the present disclosure. The choice of excipient(s) will to a large extent depend on factors such as the particular mode of administration, the effect of the excipient on solubility and stability, and the nature of the dosage form. As used herein, “pharmaceutically acceptable excipient” includes any and all solvents, dispersion media, coatings, antibacterial

and antifungal agents, isotonic and absorption delaying agents, and the like that are physiologically compatible. Some examples of pharmaceutically acceptable excipients are water, saline, phosphate buffered saline, dextrose, glycerol, ethanol and the like, as well as combinations thereof. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as mannitol, sorbitol, or sodium chloride in the composition. Additional examples of pharmaceutically acceptable substances are wetting agents or minor amounts of auxiliary substances such as wetting or emulsifying agents, preservatives, or buffers, which enhance the shelf life or effectiveness of the antibody.

[0222] Formulations of a pharmaceutical composition suitable for parenteral administration typically comprise the active ingredient combined with a pharmaceutically acceptable carrier, such as sterile water or sterile isotonic saline. Such formulations may be prepared, packaged, or sold in a form suitable for bolus administration or for continuous administration.

VIII. Delivery Methods

[0223] In some embodiments, the epigenetic editor or its component(s) are introduced to target cells in the form of nucleic acid molecule(s) encoding the epigenetic editor or its component(s); accordingly, the pharmaceutical compositions herein comprise the nucleic acid molecule(s). Such nucleic acid molecule(s) may be, for example, DNA, RNA or mRNA, and/or modified nucleic acid sequence(s) (e.g., with chemical modifications, a 5' cap, or one or more 3' modifications). In some embodiments, the nucleic acid molecule(s) may be delivered as naked DNA or RNA, for instance by means of transfection or electroporation, or can be conjugated to molecules (e.g., N-acetylgalactosamine) promoting uptake by target cells. In some embodiments, the nucleic acid molecule(s) may be in nucleic acid expression vector(s), which may include expression control sequences such as promoters, enhancers, transcription signal sequences, transcription termination sequences, introns, polyadenylation signals, Kozak consensus sequences, internal ribosome entry sites (IRES), etc. Such expression control sequences are well known in the art. A vector may also comprise a sequence encoding a signal peptide (e.g., for nuclear localization, nucleolar localization, or mitochondrial localization), associated with (e.g., inserted into or fused to) a sequence coding for a protein.

[0224] Examples of vectors include, but are not limited to, plasmid vectors; viral vectors based on vaccinia virus, poliovirus, adenovirus, adeno-associated virus, SV40, herpes simplex virus, human immunodeficiency virus, retrovirus (e.g., Murine Leukemia Virus, or

spleen necrosis virus, vectors derived from retroviruses such as Rous Sarcoma Virus, Harvey Sarcoma Virus, avian leukosis virus, a lentivirus, human immunodeficiency virus, myeloproliferative sarcoma virus, and mammary tumor virus); and other recombinant vectors. In certain embodiments, the vector is a plasmid or a viral vector. Viral particles or virus-like particles (VLPs) may also be used to deliver nucleic acid molecule(s) encoding epigenetic editors or component(s) thereof as described herein. For example, "empty" viral particles can be assembled to contain any suitable cargo. Viral vectors and viral particles may also be engineered to incorporate targeting ligands to alter target tissue specificity.

[0225] In certain embodiments, an epigenetic editor as described herein or component(s) thereof are encoded by nucleic acid sequence(s) present in one or more viral vectors, or a suitable capsid protein of any viral vector. Examples of viral vectors include adeno-associated viral vectors (e.g., derived from AAV3, AAV3b, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAVrh8, AAV10, and/or variants thereof); retroviral vectors (e.g., Maloney murine leukemia virus, MML-V), adenoviral vectors (e.g., AD100), lentiviral vectors (e.g., HIV and FIV-based vectors), and herpesvirus vectors (e.g., HSV-2).

[0226] In some embodiments, delivery involves an adeno-associated virus (AAV) vector. AAV vector delivery may be particularly useful where the DNA-binding domain of an epigenetic editor fusion protein is a zinc finger array. Without wishing to be bound by any theory, the smaller size of zinc finger arrays compared to larger DNA-binding domains such as Cas protein domains may allow such a fusion protein to be conveniently packed in viral vectors such as an AAV vector.

[0227] Any AAV serotype, e.g., human AAV serotype, can be used for an AAV vector as described herein, including, but not limited to, AAV serotype 1 (AAV1), AAV serotype 2 (AAV2), AAV serotype 3 (AAV3), AAV serotype 4 (AAV4), AAV serotype 5 (AAV5), AAV serotype 6 (AAV6), AAV serotype 7 (AAV7), AAV serotype 8 (AAV8), AAV serotype 9 (AAV9), AAV serotype 10 (AAV10), and AAV serotype 11 (AAV11), as well as variants thereof. In some embodiments, an AAV variant has at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% amino acid sequence identity to a wildtype AAV. In certain embodiments, the AAV variant may be engineered such that its capsid proteins have reduced immunogenicity or enhanced transduction ability in humans. In some instances, one or more regions of at least two different AAV serotype viruses are shuffled and reassembled to generate a chimeric variant. For example, a chimeric AAV may comprise inverted terminal repeats (ITRs) that are of a heterologous serotype compared to the serotype of the capsid. The resulting chimeric AAV can have a different antigenic reactivity or recognition

compared to its parental serotypes. In some embodiments, a chimeric variant of an AAV includes amino acid sequences from 2, 3, 4, 5, or more different AAV serotypes.

[0228] Non-viral systems are also contemplated for delivery as described herein. Non-viral systems include, but are not limited to, nucleic acid transfection methods including electroporation, sonoporation, calcium phosphate transfection, microinjection, DNA biolistics, lipid-mediated transfection, transfection through heat shock, compacted DNA-mediated transfection, lipofection, cationic agent-mediated transfection, and transfection with liposomes, immunoliposomes, exosomes, or cationic facial amphiphiles (CFAs). In certain embodiments, one or more mRNAs encoding epigenetic editor fusion proteins as described herein may be co-electroporated with one or more guide polynucleotides (e.g., gRNAs) as described herein. One important category of non-viral nucleic acid vectors is nanoparticles, which can be organic (e.g., lipid) or inorganic (e.g., gold). For instance, organic (e.g. lipid and/or polymer) nanoparticles can be suitable for use as delivery vehicles in certain embodiments of this disclosure.

[0229] In some embodiments, delivery is accomplished using a lipid nanoparticle (LNP). LNP compositions are typically sized on the order of micrometers or smaller and may include a lipid bilayer. In some embodiments, a LNP refers to any particle that has a diameter of less than 1000 nm, 500 nm, 250 nm, 200 nm, 150 nm, 100 nm, 75 nm, 50 nm, or 25 nm. In some embodiments, a nanoparticle may range in size from 1-1000 nm, 1-500 nm, 1-250 nm, 25-200 nm, 25-100 nm, 35-75 nm, or 25-60 nm. Nanoparticle compositions encompass lipid nanoparticles (LNPs), liposomes (e.g., lipid vesicles), and lipoplexes.

[0230] An LNP as described herein may be made from cationic, anionic, or neutral lipids. In some embodiments, an LNP may comprise neutral lipids, such as the fusogenic phospholipid 1,2-Dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE) or the membrane component cholesterol, as helper lipids to enhance transfection activity and nanoparticle stability. In some embodiments, an LNP may comprise hydrophobic lipids, hydrophilic lipids, or both hydrophobic and hydrophilic lipids. Any lipid or combination of lipids that are known in the art can be used to produce an LNP. The lipids may be combined in any molar ratios to produce the LNP. In some embodiments, the LNP is a liver-targeting (e.g., preferentially or specifically targeting the liver) LNP.

[0231] Any type of cell may be targeted for delivery of an epigenetic editor or component(s) thereof as described herein. For example, the cells may be eukaryotic or prokaryotic. In some embodiments, the cells are mammalian (e.g., human) cells. Human

cells may include, for example, hepatocytes, biliary epithelial cells (cholangiocytes), stellate cells, Kupffer cells, and liver sinusoidal endothelial cells.

[0232] In some embodiments, an epigenetic editor described herein, or component(s) thereof, are delivered to a host cell for transient expression, e.g., via a transient expression vector. Transient expression of the epigenetic editor or its component(s) may result in prolonged or permanent epigenetic modification of the target gene. For example, the epigenetic modification may be stable for at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. 11, or 12 weeks or more; or 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 months or more, after introduction of the epigenetic editor into the host cell. The epigenetic modification may be maintained after one or more mitotic and/or meiotic events of the host cell. In particular embodiments, the epigenetic modification is maintained across generations in offspring generated or derived from the host cell.

IX. Therapeutic Uses of Epigenetic Editors

[0233] The present disclosure also provides methods for treating or preventing a condition in a subject, comprising administering to the subject an epigenetic editor or pharmaceutical composition as described herein. The epigenetic editor may effect an epigenetic modification of a target polynucleotide sequence in a target gene associated with a disease, condition, or disorder in the subject, thereby modulating expression of the target gene to treat or prevent the disease, condition, or disorder. In some embodiments, the epigenetic editor reduces the expression of the target gene to an extent sufficient to achieve a desired effect, e.g., a therapeutically relevant effect such as the prevention or treatment of the disease, condition, or disorder.

[0234] In some embodiments, a subject is administered a system for modulating (e.g., repressing) expression of *PCSK9*, wherein the system comprises (1) the fusion protein(s) and, where relevant, guide polynucleotide(s) of an epigenetic editor as described herein, or (2) nucleic acid molecules encoding said fusion protein(s) and, where relevant, guide polynucleotide(s).

[0235] "Treat", "treating" and "treatment" refer to a method of alleviating or abrogating a biological disorder and/or at least one of its attendant symptoms. As used herein, to "alleviate" a disease, disorder or condition means reducing the severity and/or occurrence frequency of the symptoms of the disease, disorder, or condition. Further, references herein to "treatment" include references to curative, palliative and prophylactic treatment. In some embodiments, as compared with an equivalent untreated control, alleviating a symptom may

involve reduction of the symptom by at least 3%, 5%, 10%, 20%, 40%, 50%, 60%, 80%, 90%, 95%, 98%, 99%, 99.5%, 99.9%, or 100% as measured by any standard technique.

[0236] In some embodiments, the subject may be a mammal, e.g., a human. In some embodiments, the subject is selected from a non-human primate such as chimpanzee, cynomolgus monkey, or macaque, and other ape and monkey species.

[0237] In some embodiments, the human patient has a condition selected from hypercholesterolemia (e.g., familial hypercholesterolemia such as heterozygous familial hypercholesterolemia (HeFH) or homozygous familial hypercholesterolemia (HoFH), or established atherosclerotic cardiovascular disease (ASCVD)) or renal insufficiency (RI).

[0238] In some embodiments, a patient to be treated with an epigenetic editor of the present disclosure has received prior treatment for the condition to be treated (e.g., hypercholesterolemia (such as HeFH, HoFH, HF, or established ASCVD) or RI). In other embodiments, the patient has not received such prior treatment. In some embodiments, the patient has failed on a prior treatment for the condition (e.g., a prior hypercholesterolemia treatment).

[0239] An epigenetic editor of the present disclosure may be administered in a therapeutically effective amount to a patient with a condition described herein. "Therapeutically effective amount," as used herein, refers to an amount of the therapeutic agent being administered that will relieve to some extent one or more of the symptoms of the disorder being treated, and/or result in clinical endpoint(s) desired by healthcare professionals. An effective amount for therapy may be measured by its ability to stabilize disease progression and/or ameliorate symptoms in a patient, and preferably to reverse disease progression. The ability of an epigenetic editor of the present disclosure to reduce or silence *PCSK9* expression may be evaluated by *in vitro* assays, e.g., as described herein, as well as in suitable animal models that are predictive of the efficacy in humans. Suitable dosage regimens will be selected in order to provide an optimum therapeutic response in each particular situation, for example, administered as a single bolus or as a continuous infusion, and with possible adjustment of the dosage as indicated by the exigencies of each case.

[0240] An epigenetic editor of the present disclosure may be administered without additional therapeutic treatments, i.e., as a stand-alone therapy (monotherapy). Alternatively, treatment with an epigenetic editor of the present disclosure may include at least one additional therapeutic treatment (combination therapy). In some embodiments, the additional therapeutic agent is any known in the art to treat hypercholesterolemia or RI. Therapeutic agents include, but are not limited to, statins, fibrates, HMG-CoA reductase inhibitors, niacin,

bile acid modulators or sequestrants, cholesterol absorption inhibitors or modulators, CETP inhibitors, MTTP inhibitors, and PPAR agonists.

[0241] The epigenetic editors or components thereof (or nucleic acid molecules encoding the epigenetic editors or components thereof) of the present disclosure may be administered by any method accepted in the art, e.g., subcutaneously, intradermally, intratumorally, intranodally, intramuscularly, intravenously, intralymphatically, or intraperitoneally. In particular embodiments, a pharmaceutical composition of the present disclosure is administered intravenously to the subject.

X. Definitions

[0242] The term "nucleic acid" as used herein refers to any oligonucleotide or polynucleotide containing nucleotides (e.g., deoxyribonucleotides or ribonucleotides) in either single- or double-strand form, and includes DNA and RNA. "Nucleotides" contain a sugar deoxyribose (DNA) or ribose (RNA), a base, and a phosphate group, and are linked together through the phosphate groups. "Bases" include purines and pyrimidines, which include natural compounds such as adenine, thymine, guanine, cytosine, uracil, inosine, and natural analogs; as well as synthetic derivatives of purines and pyrimidines, which include, but are not limited to, modified versions which place new reactive groups such as amines, alcohols, thiols, carboxylates, alkylhalides, etc. Nucleic acids may contain known nucleotide analogs and/or modified backbone residues or linkages, which may be synthetic, naturally occurring, and non-naturally occurring. Such nucleotide analogs, modified residues, and modified linkages are well known in the art, and may provide a nucleic acid molecule with enhanced cellular uptake, reduced immunogenicity, and/or increased stability in the presence of nucleases.

[0243] As used herein, an "isolated" or "purified" nucleic acid molecule is a nucleic acid molecule that exists apart from its native environment. For example, an "isolated" or "purified" nucleic acid molecule (1) has been separated away from the nucleic acids of the genomic DNA or cellular RNA of its source of origin; and/or (2) does not occur in nature. In some embodiments, an "isolated" or "purified" nucleic acid molecule is a recombinant nucleic acid molecule.

[0244] It will be understood that in addition to the specific proteins and nucleic acid molecules mentioned herein, the present disclosure also contemplates the use of variants, derivatives, homologs, and fragments thereof. A variant of any given sequence may have the specific sequence of residues (whether amino acid or nucleic acid residues) modified in such

a manner that the polypeptide or polynucleotide in question substantially retains at least one of its endogenous functions. A variant sequence can be obtained by addition, deletion, substitution, modification, replacement and/or variation of at least one residue present in the naturally-occurring sequence (in some embodiments, no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, or 20 residues). For specific proteins described herein (e.g., KRAB, dCas9, DNMT3A, and DNMT3L proteins described herein), the present disclosure also contemplates any of the protein's naturally occurring forms, or variants or homologs that retain at least one of its endogenous functions (e.g., at least 50%, 60%, 70%, 80%, 90%, 85%, 96%, 97%, 98%, or 99% of its function as compared to the specific protein described).

[0245] Some exemplary fusion proteins embraced by the present disclosure are provided herein. It will be appreciated by the skilled artisan, that these exemplary proteins are non-limiting examples and that additional proteins are within the scope of the present disclosure. For example, where fusion exemplary proteins comprising a specific domain, e.g., a mammalian DNMT3A, DNMT3L and/or KRAB domain, such as a human or mouse DNMT3A, DNMT3L and/or KRAB domain, are provided, the skilled artisan will be able to ascertain that, in some embodiments, fusion proteins with the same configuration, but with one or more of the mammalian domains substituted for a homologous domain from another mammal, e.g., one or more mouse domains substituted for one or more human domains, are also embraced by the present disclosure. For example, where an exemplary fusion protein is provided that comprises a mouse DNMT3L domain, a fusion protein of the same architecture but with the mouse DNMT3L substituted for a human DNMT3L domain is also embraced.

[0246] As used herein, a homologue of any polypeptide or nucleic acid sequence contemplated herein includes sequences having a certain homology with the wildtype amino acid and nucleic sequence. A homologous sequence may include a sequence, e.g. an amino acid sequence which may be at least 50%, 55%, 65%, 75%, 85%, 90%, 91%, 92%< 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the subject sequence. The term "percent identical" in the context of amino acid or nucleotide sequences refers to the percent of residues in two sequences that are the same when aligned for maximum correspondence. In some embodiments, the length of a reference sequence aligned for comparison purposes is at least 30%, (e.g., at least 40, 50, 60, 70, 80, or 90%, or 100%) of the reference sequence. Sequence identity may be measured using sequence analysis software (for example, Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, Wis. 53705, BLAST, BESTFIT, GAP, or PILEUP/PRETTYBOX programs). Such software matches identical or similar

sequences by assigning degrees of homology to various substitutions, deletions, and/or other modifications. In an exemplary approach to determining the degree of identity, a BLAST program may be used, with a probability score between e-3 and e-100 indicating a closely related sequence.

[0247] The percent identity of two nucleotide or polypeptide sequences is determined by, e.g., BLAST® using default parameters (available at the U.S. National Library of Medicine's National Center for Biotechnology Information website). In some embodiments, the length of a reference sequence aligned for comparison purposes is at least 30%, (e.g., at least 40, 50, 60, 70, 80, or 90%) of the reference sequence.

[0248] It will be understood that the numbering of the specific positions or residues in polypeptide sequences depends on the particular protein and numbering scheme used. Numbering might be different, e.g., in precursors of a mature protein and the mature protein itself, and differences in sequences from species to species may affect numbering. One of skill in the art will be able to identify the respective residue in any homologous protein and in the respective encoding nucleic acid by methods well known in the art, e.g., by sequence alignment and determination of homologous residues.

[0249] The term "modulate" or "alter" refers to a change in the quantity, degree, or extent of a function. For example, an epigenetic editor as described herein may modulate the activity of a promoter sequence by binding to a motif within the promoter, thereby inducing, enhancing, or suppressing transcription of a gene operatively linked to the promoter sequence. As other examples, an epigenetic editor as described herein may block RNA polymerase from transcribing a gene, or may inhibit translation of an mRNA transcript. The terms "inhibit," "repress," "suppress," "silence" and the like, when used in reference to an epigenetic editor or a component thereof as described herein, refers to decreasing or preventing the activity (e.g., transcription) of a nucleic acid sequence (e.g., a target gene) or protein relative to the activity of the nucleic acid sequence or protein in the absence of the epigenetic editor or component thereof. The term may include partially or totally blocking activity, or preventing or delaying activity. The inhibited activity may be, e.g., 10%, 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% less than that of a control, or may be, e.g., at least 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold, or 10-fold less than that of a control.

[0250] The term "about" or "approximately" means within an acceptable error range for the particular value as determined by one of ordinary skill in the art, which will depend in part on how the value is measured or determined, e.g., the limitations of the measurement

system. For example, "about" can mean within one or more than one standard deviation, per the practice in the given value. Where particular values are described in the application and claims, unless otherwise stated, the term "about" should be assumed to mean an acceptable error range for the particular value.

Ranges provided herein are understood to be shorthand for all of the values within the range. For example, a range of 1 to 50 is understood to include any number, combination of numbers, or sub-range from the group consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50, as well as all intervening decimal values between the aforementioned integers such as, for example, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, and 1.9. With respect to sub-ranges, "nested sub-ranges" that extend from either end point of the range are specifically contemplated. For example, a nested sub-range of an exemplary range of 1 to 50 may comprise 1 to 10, 1 to 20, 1 to 30, and 1 to 40 in one direction, or 50 to 40, 50 to 30, 50 to 20, and 50 to 10 in the other direction.

[0251] Unless otherwise defined herein, scientific and technical terms used in connection with the present disclosure shall have the meanings that are commonly understood by those of ordinary skill in the art. Exemplary methods and materials are described below, although methods and materials similar or equivalent to those described herein can also be used in the practice or testing of the present disclosure. In case of conflict, the present specification, including definitions, will control. Further, unless otherwise required by context, singular terms shall include pluralities and plural terms shall include the singular. Throughout this specification and embodiments, the words "have" and "comprise," or variations such as "has," "having," "comprises," or "comprising," will be understood to imply the inclusion of a stated integer or group of integers but not the exclusion of any other integer or group of integers. Unless otherwise indicated, the recitation of a listing of elements herein includes any of the elements singly or in any combination. The recitation of an embodiment herein includes that embodiment as a single embodiment, or in combination with any other embodiment(s) herein. All publications, patents, patent applications, and other references mentioned herein are incorporated by reference in their entirety. To the extent that references incorporated by reference contradict the disclosure contained in the specification, the specification is intended to supersede and/or take precedence over any such contradictory material. Although a number of documents are cited herein, this citation does not constitute an admission that any of these documents forms part of the common general knowledge in the art.

[0252] According to the present disclosure, back-references in the dependent claims are meant as short-hand writing for a direct and unambiguous disclosure of each and every combination of claims that is indicated by the back-reference. Further, headers herein are created for ease of organization and are not intended to limit the scope of the claimed compositions and methods in any manner.

[0253] Some protein sequences, e.g., some fusion protein sequences, provided herein include a peptide tag, e.g., a His6 tag, or a DYKDDDDK (SEQ ID NO: 1528) tag, which are useful for detection and/or purification of tagged proteins, but do not affect protein function. These peptide tags, and additional suitable peptide tags, are well known to those of skill in the art. It will be apparent to the person of skill in the art that the disclosed tags can be substituted for other suitable peptide tags, and that fusion proteins of the same or highly similar sequence, but not including such peptide tags, e.g., from which the peptide tag has been cleaved or which are created without a peptide tag, are suitable for carrying out embodiments of the present disclosure as well.

[0254] In order that the present disclosure may be better understood, the following examples are set forth. These examples are for purposes of illustration only and are not to be construed as limiting the scope of the present disclosure in any manner.

EXAMPLES

Example 1: Fusion Protein Design and Synthesis

[0255] Fusion proteins comprising dCas9, DNMT3A, DNMT3L, and KOX1 KRAB ("CRISPR-off") were designed and constructed. From N terminus to C terminus, the proteins have the domains DNMT3A-linker-DNMT3L-XTEN80-NLS-dSpCas9- NLS-XTEN16-KOX1 KRAB (SEQ ID NOs: 658 and 1495). The CRISPR-off plasmid construct has been described in Nuñez (Nuñez et al., *Cell* (2021) 184(9):2503-19) and was ordered from Twist Biosciences.

[0256] ZF fusion proteins ("ZF-off") comprising DNMT3A, 3L, and KOX1 KRAB were also constructed. The constructs have the general structure DNMT3A-linker-DNMT3L-XTEN80-NLS-ZFP domain-NLS-XTEN16-KOX1Krab (SEQ ID NOs: 659 and 1496).

Example 2: Selection of Target *PCSK9* Sequences for gRNA Epigenetic Silencing

[0257] gRNAs targeting +/- 1 kb from the *PCSK9* TSS were computationally designed using the Benchling gRNA platform (Benchling (2021), retrieved from benchling.com) for human (GRCh38), mouse (mm10) and *Macaca fascicularis* (5.0) *PCSK9*. gRNAs containing

poly-TTTT sequences were first discarded. We performed gRNA off-target analysis using CasOFFinder (Bae et al., *Bioinformatics* (2014) 30(10):1473-5). gRNAs were discarded if they matched to multiple locations across the respective genome build for each independent species.

[0258] A cross-reactivity sequence analysis was performed on human *PCSK9* gRNAs in order to annotate sequence mismatches with *Macaca* or mouse gRNA sequences. In particular, gRNA sequence alignments were performed to identify the degree of DNA similarity at each nucleotide, including the annotation of guides that contain up to zero, one, or two nucleotide mismatches. A final set of 226 gRNA sequences was selected for the *PCSK9* primary screen in HeLa cells.

Example 3: Selection of ZF Target Sites and Design of ZF Proteins for Epigenetic Silencing

[0259] A library of two-finger ZFPs (2F units), each recognizing 6 bp DNA sites, was used to design larger six-finger ZFP arrays targeting 18 bp DNA binding sites. The source of the 2F units was a set of three-finger zinc finger proteins that had been selected to bind specific target sites using a bacterial-2-hybrid (B2H) selection system (Hurt et al., *PNAS* (2003) 100:12271-6; Maeder et al., *Mol Cell* (2008) 31(2):294-301). A list of targetable DNA sites was created by generating all possible triplet combinations of 6 bp binding sites represented in the library and allowing either 0 or 1 bp between the 6 bp target sites. To identify zinc finger target sites within *PCSK9*, the sequence +/- 1kb from TSS (human (GRCh38)) was interrogated against this list. For each identified ZF target site, multiple ZF proteins could be designed. Design of the six recognition helices used to generate the full proteins was performed by selecting two-finger units and taking into account a number of factors such as known binding preferences of zinc finger proteins, the frequency with which amino acids in positions –1, 2, 3 and 6 had been selected in the B2H selection system to bind the desired target base, avoidance of amino acids in positions –1, 2, 3 and 6 that had been selected to bind multiple different bases in the B2H, and maintaining context dependencies by matching flanking bases where possible. The full ZF sequence is derived from the naturally occurring Zif268 protein and selected recognition helices were maintained in the sequence context in which they were selected in the B2H (either fingers 1-2 or fingers 2-3 from Zif268). Two-finger units were joined by the linker TGSQKP (SEQ ID NO: 651) where 6 bp binding sites were contiguous and by the linker TGGGGSQKP (SEQ ID NO:

652) where 1 bp separated the 6 bp binding sites. A final set of 209 ZFPs targeted to 49 distinct binding sites were selected for the *PCSK9* primary screen in HeLa cells.

[0260] **FIG. 1** shows the overlap of the gRNAs and zinc finger proteins mapped to the *PCSK9* target region.

Example 4: Guide RNA Screening in HeLa Cells

[0261] A primary screen of gRNAs targeting *PCSK9* was performed in HeLa cells. gRNA sequences were ordered from Twist Biosciences as DNA fragments with a u6 promoter sequence preceding the gRNA coding sequence.

[0262] HeLa cells were transfected with gRNA and CRISPR-off in DNA format. Six 96-well plates (Sigma-Aldrich Catalog No. M2936) were seeded with 12,000 HeLa cells per well (ATCC Catalog No. CCL-2) in standard culture media containing DMEM (Thermo Fisher Catalog No. 11-965-092) supplemented with 10% Fetal Bovine Serum (Thermo Fisher Catalog No. A4766) v/v, 1x GlutaMAX™ (Thermo Fisher Catalog No. 35050061) and 1x Penicillin-Streptomycin (Thermo Fisher Catalog No. 15140122). Following plating, cells were allowed to grow for 24 hours in an incubator at 37° C with 5% CO_2. 25 ng of each gRNA fragment and 50 ng of the CRISPR-off plasmid (SEQ ID NO: 658) were resuspended in DPBS buffer (Thermo Fisher Catalog No. 14190144) to a concentration of 7.5 ng/μL. Additionally, 10 ng of EF1a:PuromycinResistance plasmid was also added to the transfection mix to achieve a total payload of 85 ng of DNA. Transfection mixtures were created by adding the resuspended DNA components to Mirus® TransIT®-LT1 transfection reagent (Mirus Catalog No. MIR2300) following the manufacturer's instructions. 10 μL of each of the transfection mixtures was added in duplicate across a total of six screening plates. The positive controls used were CRISPRi (dCas9-KRAB) with two gRNAs targeting sites proximal to the TSS. The two CRISPRi positive control gRNAs used were gRNA004 and gRNA005 as annotated in the table of gRNA sequences. These control conditions are referred to as "CRISPRi-1" and "CRISPRi-2" respectively in the primary screen data tables. The negative controls were CRISPR-off without gRNA, CRISPR-off with a non-*PCSK9* locus (*CD151*)-targeting gRNA, and empty vector (pUC19; NEB Catalog No. N3041S).

[0263] 24 hours following transfection, a puromycin resistance selection was performed. The cell media was aspirated completely and all wells were washed 3x with DPBS buffer (Thermo Fisher Catalog No. 14190144) and 200 μL of 1 μg/μL puromycin was added to all screening plate wells.

[0264] 48 hours following transfection, the cells were passaged. The cell media was completely aspirated and all wells were washed 3x with DPBS buffer (Thermo Fisher Catalog No. 14190144). Cells were enzymatically lifted by adding 25 µL of Trypsin-EDTA (0.25%) (Thermo Fisher Catalog No. 25200056) for five minutes in a 37° C incubator. Trypsinized cells were resuspended 1:8 in fresh standard culture media and re-plated at a ratio of 1:4 72 hours after the media was changed.

[0265] In order to measure the level of secreted PCSK9 protein, media was harvested 24 hours after the media change and the cell plates were assayed for relative cell counts using the Promega Cell Titer Glo™ protocol (Catalog No. G7570) according to manufacturer's recommendations. PCSK9 protein levels were assessed using the LEGEND MAX™ Human PCSK9 ELISA Kit from BioLegend (Catalog No. 443107). Harvested media was plated and all subsequent steps were performed exactly according to manufacturer's recommendations. Final plate reads at 450 nm were performed on the Perkin Elmer® VICTOR® Nivo™ F instrument. GraphPad Prism software was used to fit a function to the standard curves and interpolate unknowns. PCSK9 ELISA results were normalized by Cell Titer Glo® assay (Promega Catalog No. G7571) results in order to correct for any well-to-well cell number variability.

[0266] Over 200 gRNAs were tested, of which 40 were identified as being top sequences (**FIG. 2**; top sequences designated as darker circles). The sequences and efficacies of the tested gRNAs are shown in **Table 7** (SEQ: SEQ ID NO). The relative PCSK9 secretion ("% Control PCSK9") represents the averaged PCSK9 protein levels of the treated samples expressed as a percent of the average across all non-targeting gRNA (CD151) negative control conditions. Robust silencing of *PCSK9* (30-40% of negative control levels) was observed in cells treated with a number of gRNA candidate treatments. The top 40 gRNAs with the best PCSK9 protein knockdown were selected to be ordered as sgRNAs for further follow-up studies.

Table 7. Targeting Sequences of Tested gRNAs

gRNA No. or Condition	SEQ	Targeting Sequence	bp to TSS	GRCh38 sequence start coordinate	% Control PCSK9
CD151 gRNA	N/A	N/A	N/A	N/A	100
no gRNA	N/A	N/A	N/A	N/A	72.51
pUC	N/A	N/A	N/A	N/A	119.55
CRISPRi-1	N/A	N/A	N/A	N/A	86.89
CRISPRi-2	N/A	N/A	N/A	N/A	64.09

gRNA No. or Condition	SEQ	Targeting Sequence	bp to TSS	GRCh38 sequence start coordinate	% Control PCSK9
gRNA001	1262	GGUGCUAGCCUUGCGUUCCG	431	55039960	43.85
gRNA002	1263	CUGGCCGAAGCACCCGAGCA	462	55039991	74.87
gRNA003	1264	UGCGGAAACCUUCUAGGGUG	0	55039529	27.33
gRNA004	1265	GCGGAAACCUUCUAGGGUGU	-1	55039528	34.93
gRNA005	1266	UCAAGCACCCACACCCUAGA	-11	55039518	41.44
gRNA006	1267	GGGUGUGGGUGCUUGACGCC	-15	55039514	42.41
gRNA007	1268	GGUGUGGGUGCUUGACGCCU	-16	55039513	31.11
gRNA008	1269	GUGUGGGUGCUUGACGCCUG	-17	55039512	36.13
gRNA009	1270	ACUGCCUGGCUCACUCCUCC	81	55039610	21.34
gRNA010	1271	UCACGCCACCAGAGCCCCAU	-48	55039481	52.20
gRNA011	1272	AUCGUCCGAUGGGGCUCUGG	-56	55039473	31.00
gRNA012	1273	AGGAUCGUCCGAUGGGGCUC	-59	55039470	38.05
gRNA013	1274	UCAGAUAGGAUCGUCCGAUG	-65	55039464	41.70
gRNA014	1275	GCGGCUCCCAGCUCCCAGCC	140	55039669	85.56
gRNA015	1276	CGGAAUCCUGGCUGGGAGCU	149	55039678	43.48
gRNA016	1277	GCGGAAUCCUGGCUGGGAGC	150	55039679	75.20
gRNA017	1278	GGGCGCGCGGAAUCCUGGCU	156	55039685	53.76
gRNA018	1279	GGGGCGCGCGGAAUCCUGGC	157	55039686	56.46
gRNA019	1280	UGAGGUCUUUGCAAACAAAG	-919	55038610	76.59
gRNA020	1281	CCAGCACCUAGAUUCAGAGC	-875	55038654	60.05
gRNA021	1282	CCUGCUCUGAAUCUAGGUGC	-872	55038657	56.48
gRNA022	1283	UCGAUACUGGGAAGAAACAA	-798	55038731	67.82
gRNA023	1284	CUGGAAGGGCUGUCGAUACU	-786	55038743	59.17
gRNA024	1285	UCUGGAAGGGCUGUCGAUAC	-785	55038744	63.62
gRNA025	1286	GAGGCUUGCUCUUUCUGGAA	-772	55038757	58.91
gRNA026	1287	UGAGGCUUGCUCUUUCUGGA	-771	55038758	74.33
gRNA027	1288	GACAUGAGGCUUGCUCUUUC	-767	55038762	87.93
gRNA028	1289	UGUACAUGUGGCAUGACAUG	-753	55038776	68.47
gRNA029	1290	CAUGCCACAUGUACAAUCUG	-748	55038781	88.26
gRNA030	1291	CACAUGUACAAUCUGAGGCC	-743	55038786	79.41
gRNA031	1292	CUGGCCUCAGAUUGUACAUG	-741	55038788	106.36
gRNA032	1293	AAAAGGGGAAAGAGAGCUCC	-722	55038807	91.18
gRNA033	1294	CCAGGCAGGAGGAUGAAAAG	-707	55038822	89.34
gRNA034	1295	ACCAGGCAGGAGGAUGAAAA	-706	55038823	79.11
gRNA035	1296	UACCAGGCAGGAGGAUGAAA	-705	55038824	82.01
gRNA036	1297	CCUCCUGCCUGGUACACAAU	-699	55038830	69.28
gRNA037	1298	CCUAUUGUGUACCAGGCAGG	-696	55038833	82.57
gRNA038	1299	ACACCUAUUGUGUACCAGGC	-693	55038836	58.95
gRNA039	1300	GGUACACAAUAGGUGUUUAC	-689	55038840	81.55
gRNA040	1301	GUAAACACCUAUUGUGUACC	-689	55038840	89.49
gRNA041	1302	UCCAGUUGAUUUCUUGAACA	-659	55038870	69.65
gRNA042	1303	ACCAUGUUCAAGAAAUCAAC	-655	55038874	73.51
gRNA043	1304	UCUUGAACAUGGUGUGUAAA	-648	55038881	86.96
gRNA044	1305	UCUUUGCAAAUUGAAUCUUC	-623	55038906	80.70
gRNA045	1306	AGUUUGCAAAGACGUCAUAU	-559	55038970	85.13

gRNA No. or Condition	SEQ	Targeting Sequence	bp to TSS	GRCh38 sequence start coordinate	% Control PCSK9
gRNA046	1307	GAUUUAUACUACAAAGAUUC	-531	55038998	74.88
gRNA047	1308	AGUUGGUAAGGUCAGUGUGC	-454	55039075	66.05
gRNA048	1309	GUUGGUAAGGUCAGUGUGCA	-453	55039076	60.64
gRNA049	1310	GCAGGGUGCAUAAAGGGCAG	-436	55039093	60.72
gRNA050	1311	GGUGCAUAAAGGGCAGAGGC	-432	55039097	92.12
gRNA051	1312	GCAUAAAGGGCAGAGGCCGG	-429	55039100	53.71
gRNA052	1313	UUUAGAAGGCUGCCAGGUUA	-392	55039137	56.64
gRNA053	1314	GCCCACCGAAUUCUUUCCAC	-366	55039163	62.84
gRNA054	1315	AAAGAAUUCGGUGGGCAGCG	-362	55039167	78.78
gRNA055	1316	CUUCUGAAUCAAUCCUACUG	-333	55039196	52.04
gRNA056	1317	CUGGUCAGCAGGAGACAAGG	-328	55039200	90.01
gRNA057	1318	GAUUGAUUCAGAAGUCUCAC	-327	55039202	64.68
gRNA058	1319	CAGAAGUCUCACUGGUCAGC	-319	55039210	101.68
gRNA059	1320	UCACUGGUCAGCAGGAGACA	-311	55039218	75.12
gRNA060	1321	GCAGGAGACAAGGUGGACCC	-301	55039228	43.48
gRNA061	1322	GGACCCAGGAAACACUGAAA	-287	55039242	50.45
gRNA062	1323	CCCAGGAAACACUGAAAAGG	-284	55039245	104.49
gRNA063	1324	CCAGGAAACACUGAAAAGGU	-283	55039246	74.99
gRNA064	1325	AAACACUGAAAAGGUGGGCC	-278	55039251	78.05
gRNA065	1326	UGGAGUCUGGCAUCCCACGC	-248	55039281	86.50
gRNA066	1327	GGAGUCUGGCAUCCCACGCA	-247	55039282	64.96
gRNA067	1328	CGGGAGAGGAGGAGCCCCUA	-217	55039312	51.16
gRNA068	1329	AGGAGGAGCCCCUAGGGCGC	-211	55039318	49.85
gRNA069	1330	AAGGCAGGCCGGCGCCCUAG	-200	55039329	57.25
gRNA070	1331	GAAGGCAGGCCGGCGCCCUA	-199	55039330	71.39
gRNA071	1332	GGAAGGCAGGCCGGCGCCCU	-198	55039331	55.58
gRNA072	1333	AACUGGGCUGGAAGGCAGGC	-189	55039340	81.36
gRNA073	1334	GCCUGCCUUCCAGCCCAGUU	-189	55039340	68.77
gRNA074	1335	UCCUAACUGGGCUGGAAGGC	-185	55039344	67.81
gRNA075	1336	CUUCCAGCCCAGUUAGGAUU	-183	55039346	58.27
gRNA076	1337	UUCCAGCCCAGUUAGGAUUU	-182	55039347	53.06
gRNA077	1338	CAAAUCCUAACUGGGCUGGA	-181	55039348	31.55
gRNA078	1339	CUCCCAAAUCCUAACUGGGC	-177	55039352	52.30
gRNA079	1340	AAAACUCCCAAAUCCUAACU	-173	55039356	52.38
gRNA080	1341	AAAAACUCCCAAAUCCUAAC	-172	55039357	67.62
gRNA081	1342	AGCGUCAGAUUACGCGCAGA	-145	55039384	58.52
gRNA082	1343	CAGCGUCAGAUUACGCGCAG	-144	55039385	76.23
gRNA083	1344	GCGCGUAAUCUGACGCUGUU	-142	55039387	73.94
gRNA084	1345	CGCGUAAUCUGACGCUGUUU	-141	55039388	71.95
gRNA085	1346	GCGUAAUCUGACGCUGUUUG	-140	55039389	77.63
gRNA086	1347	UAAUCUGACGCUGUUUGGGG	-137	55039392	70.70
gRNA087	1348	AAUCUGACGCUGUUUGGGGA	-136	55039393	48.40
gRNA088	1349	GACGCUGUUUGGGGAGGGCG	-131	55039398	64.83
gRNA089	1350	CGAAACCUGAUCCUCCAGUC	-107	55039422	56.51
gRNA090	1351	GAAACCUGAUCCUCCAGUCC	-106	55039423	45.68

gRNA No. or Condition	SEQ	Targeting Sequence	bp to TSS	GRCh38 sequence start coordinate	% Control PCSK9
gRNA091	1352	AAACCUGAUCCUCCAGUCCG	-105	55039424	50.44
gRNA092	1353	CGGACUGGAGGAUCAGGUUU	-105	55039424	30.88
gRNA093	1354	AACCUGAUCCUCCAGUCCGG	-104	55039425	29.56
gRNA094	1355	AACCCCCGGACUGGAGGAUC	-99	55039430	33.66
gRNA095	1356	UAACGGAACCCCCGGACUGG	-93	55039436	35.57
gRNA096	1357	CAUUAACGGAACCCCCGGAC	-90	55039439	30.87
gRNA097	1358	UUAAACAUUAACGGAACCCC	-85	55039444	58.22
gRNA098	1359	CCGUUAAUGUUUAAUCAGAU	-79	55039450	58.99
gRNA099	1360	CCUAUCUGAUUAAACAUUAA	-76	55039453	45.83
gRNA100	1361	AAUCAGAUAGGAUCGUCCGA	-67	55039462	47.05
gRNA101	1362	AUCAGAUAGGAUCGUCCGAU	-66	55039463	57.76
gRNA102	1363	UGGCGUGAUCUGCGCGCCCC	-36	55039493	41.66
gRNA103	1364	GUCGCUGCGGAAACCUUCUA	5	55039534	52.14
gRNA104	1365	CGUCGCUGCGGAAACCUUCU	6	55039535	54.97
gRNA105	1366	AGGUUUCCGCAGCGACGUCG	9	55039538	61.69
gRNA106	1367	CAGCGACGUCGAGGCGCUCA	18	55039547	54.52
gRNA107	1368	UGAGCGCCUCGACGUCGCUG	18	55039547	35.11
gRNA108	1369	GUCGAGGCGCUCAUGGUUGC	25	55039554	25.10
gRNA109	1370	GAGGCGCUCAUGGUUGCAGG	28	55039557	43.19
gRNA110	1371	AGGCGCUCAUGGUUGCAGGC	29	55039558	29.00
gRNA111	1372	AGUUCAGGGUCUGAGCCUGG	43	55039571	39.04
gRNA112	1373	CGGGCGCCGCCGUUCAGUUC	48	55039577	45.13
gRNA113	1374	GGGCGCCGCCGUUCAGUUCA	49	55039578	32.30
gRNA114	1375	UCAGACCCUGAACUGAACGG	57	55039586	55.95
gRNA115	1376	GGCUCAGACCCUGAACUGAA	60	55039589	48.59
gRNA116	1377	UUCAGUUCAGGGUCUGAGCC	60	55039589	57.38
gRNA117	1378	UGAGCCUGGAGGAGUGAGCC	74	55039603	45.58
gRNA118	1379	AGUGAGCCAGGCAGUGAGAC	86	55039615	46.98
gRNA119	1380	GCCAGGCAGUGAGACUGGCU	91	55039620	62.91
gRNA120	1381	CCAGGCAGUGAGACUGGCUC	92	55039621	54.17
gRNA121	1382	CCCGAGCCAGUCUCACUGCC	95	55039624	48.29
gRNA122	1383	GGCAGUGAGACUGGCUCGGG	95	55039624	34.80
gRNA123	1384	GCAGUGAGACUGGCUCGGGC	96	55039625	53.96
gRNA124	1385	UGAGACUGGCUCGGGCGGGC	100	55039629	59.99
gRNA125	1386	GAGACUGGCUCGGGCGGGCC	101	55039630	34.59
gRNA126	1387	GGGACGCGUCGUUGCAGCAG	121	55039650	38.46
gRNA127	1388	GCUGCUGCAACGACGCGUCC	122	55039651	44.41
gRNA128	1389	UGAAGGGGCGCGCGGAAUCC	161	55039690	28.00
gRNA129	1390	AGGGCGCGUGAAGGGGCGCG	169	55039698	64.66
gRNA130	1391	CAGGAGCAGGGCGCGUGAAG	176	55039705	52.46
gRNA131	1392	UCAGGAGCAGGGCGCGUGAA	177	55039706	61.68
gRNA132	1393	UUCAGGAGCAGGGCGCGUGA	178	55039707	48.06
gRNA133	1394	GGAGCUGAAGUUCAGGAGCA	188	55039717	62.68
gRNA134	1395	AGGAGCUGAAGUUCAGGAGC	189	55039718	66.77
gRNA135	1396	CUGUGCAGGAGCUGAAGUUC	195	55039724	47.35

gRNA No. or Condition	SEQ	Targeting Sequence	bp to TSS	GRCh38 sequence start coordinate	% Control PCSK9
gRNA136	1397	GCACAGUCCUCCCCACCGCA	209	55039738	40.95
gRNA137	1398	UGCGGUGGGGAGGACUGUGC	209	55039738	51.32
gRNA138	1399	CCUCCCCACCGCAAGGCUCA	216	55039745	76.44
gRNA139	1400	CCUUGAGCCUUGCGGUGGGG	219	55039748	71.95
gRNA140	1401	GCGCCUUGAGCCUUGCGGUG	222	55039751	63.12
gRNA141	1402	GGCGCCUUGAGCCUUGCGGU	223	55039752	64.08
gRNA142	1403	CGGCGCCUUGAGCCUUGCGG	224	55039753	68.16
gRNA143	1404	CGCAAGGCUCAAGGCGCCGC	225	55039754	58.50
gRNA144	1405	GGCUCAAGGCGCCGCCGGCG	230	55039759	41.31
gRNA145	1406	AGGCCGUGCGCGGUCCACGC	247	55039776	46.31
gRNA146	1407	GUGGACCGCGCACGGCCUCU	249	55039778	52.25
gRNA147	1408	GGAGACCUAGAGGCCGUGCG	257	55039786	62.38
gRNA148	1409	CAGGACAGCAACCUCUCCCC	281	55039810	59.44
gRNA149	1410	UGGGCACCGUCAGCUCCAGG	310	55039839	54.56
gRNA150	1411	CCGUCAGCUCCAGGCGGUCC	316	55039845	47.26
gRNA151	1412	CCAGGACCGCCUGGAGCUGA	319	55039848	71.17
gRNA152	1413	UCAGCUCCAGGCGGUCCUGG	319	55039848	46.68
gRNA153	1414	CAGCGGCCACCAGGACCGCC	328	55039857	67.26
gRNA154	1415	CUGCUGCUCCUGGGUCCCGC	366	55039895	54.47
gRNA155	1416	CACGGGCGCCCGCGGGACCC	377	55039906	48.18
gRNA156	1417	UCCCGCGGGCGCCCGUGCGC	380	55039909	53.18
gRNA157	1418	CGCGGGCGCCCGUGCGCAGG	383	55039912	61.58
gRNA158	1419	UCCUGCGCACGGGCGCCCGC	384	55039913	67.30
gRNA159	1420	CUCCUGCGCACGGGCGCCCG	385	55039914	48.64
gRNA160	1421	CGCCCGUGCGCAGGAGGACG	389	55039918	84.11
gRNA161	1422	CGUGCGCAGGAGGACGAGGA	393	55039922	86.55
gRNA162	1423	GUCCUCGUCCUCCUGCGCAC	394	55039923	54.45
gRNA163	1424	CGUCCUCGUCCUCCUGCGCA	395	55039924	80.30
gRNA164	1425	GGACGAGGACGGCGACUACG	404	55039933	69.15
gRNA165	1426	GGACGGCGACUACGAGGAGC	410	55039939	62.54
gRNA166	1427	GCUAGCCUUGCGUUCCGAGG	434	55039963	60.30
gRNA167	1428	GCCUUGCGUUCCGAGGAGGA	438	55039967	57.71
gRNA168	1429	GCCGUCCUCCUCGGAACGCA	442	55039971	62.04
gRNA169	1430	GCGUUCCGAGGAGGACGGCC	443	55039972	79.68
gRNA170	1431	UUCGGCCAGGCCGUCCUCCU	451	55039980	40.06
gRNA171	1432	CGUGCUCGGGUGCUUCGGCC	464	55039993	59.35
gRNA172	1433	GUGGCUGUGGUUCCGUGCUC	477	55040006	45.91
gRNA173	1434	GGUGGCUGUGGUUCCGUGCU	478	55040007	70.81
gRNA174	1435	GCAGCGGUGGAAGGUGGCUG	490	55040019	72.70
gRNA175	1436	CACCUUCCACCGCUGCGCCA	494	55040023	86.51
gRNA176	1437	CUUGGCGCAGCGGUGGAAGG	496	55040025	76.28
gRNA177	1438	CACCUUGGCGCAGCGGUGGA	499	55040028	66.94
gRNA178	1439	UCCACCGCUGCGCCAAGGUG	499	55040028	48.71
gRNA179	1440	CCACCGCUGCGCCAAGGUGC	500	55040029	60.24
gRNA180	1441	CCCGCACCUUGGCGCAGCGG	503	55040032	44.64

gRNA No. or Condition	SEQ	Targeting Sequence	bp to TSS	GRCh38 sequence start coordinate	% Control PCSK9
gRNA181	1442	GGGCGAACCCGCAGCCGGGA	541	55040070	46.89
gRNA182	1443	UCCCGGCUGCGGGUUCGCCC	541	55040070	57.20
gRNA183	1444	CACCGCACCGUCCCGGCUGC	551	55040080	76.92
gRNA184	1445	GCACCGCACCGUCCCGGCUG	552	55040081	85.00
gRNA185	1446	GAAACAGCACCGCACCGUCC	558	55040087	75.31
gRNA186	1447	GGUGCGGUGCUGUUUCCUCU	562	55040091	51.49
gRNA187	1448	GUGCGGUGCUGUUUCCUCUC	563	55040092	57.43
gRNA188	1449	GGGGGAAACUGAGGCCCGAG	577	55040105	55.16
gRNA189	1450	AGUUUCCCCCCAUGUAAGAG	590	55040119	45.91
gRNA190	1451	CCCCCAUGUAAGAGAGGAAG	596	55040125	63.89
gRNA191	1452	CACUUCCUCUCUUACAUGGG	598	55040127	63.57
gRNA192	1453	CCACUUCCUCUCUUACAUGG	599	55040128	66.47
gRNA193	1454	UAAGAGAGGAAGUGGAGUGC	604	55040133	63.94
gRNA194	1455	GUCGCCGAGGGCUCUUCGCU	626	55040155	55.19
gRNA195	1456	CGUGCCAAGCGAAGAGCCCU	633	55040162	66.66
gRNA196	1457	CUCUUCGCUUGGCACGAUCU	637	55040166	57.10
gRNA197	1458	UCUUCGCUUGGCACGAUCUU	638	55040167	44.24
gRNA198	1459	CUUCGCUUGGCACGAUCUUG	639	55040168	50.40
gRNA199	1460	GGCACGAUCUUGGGGACUGC	647	55040176	73.29
gRNA200	1461	GAUCUUGGGGACUGCAGGCA	652	55040181	56.39
gRNA201	1462	CUUGGGGACUGCAGGCAAGG	655	55040184	66.89
gRNA202	1463	GGACUGCAGGCAAGGCGGCG	660	55040189	61.44
gRNA203	1464	GACUGCAGGCAAGGCGGCGG	661	55040190	49.55
gRNA204	1465	UGCAGGCAAGGCGGCGGGGG	664	55040193	41.85
gRNA205	1466	GGCAAGGCGGCGGGGGAGGA	668	55040197	95.50
gRNA206	1467	GCAAGGCGGCGGGGGAGGAC	669	55040198	120.71
gRNA207	1468	GUGGGGAGCACGGUGGAGAG	694	55040223	73.17
gRNA208	1469	UGGGGAGCACGGUGGAGAGC	695	55040224	61.04
gRNA209	1470	GGGGAGCACGGUGGAGAGCG	696	55040225	79.49
gRNA210	1471	AGCACGGUGGAGAGCGGGGA	700	55040229	42.81
gRNA211	1472	CGGUGGAGAGCGGGGACGGC	704	55040233	79.46
gRNA212	1473	CGUGCGGCUGCGCUAUUCAG	748	55040277	39.48
gRNA213	1474	GUGCGGCUGCGCUAUUCAGU	749	55040278	56.25
gRNA214	1475	GGCUGCGCUAUUCAGUGGGA	753	55040282	82.51
gRNA215	1476	CUAUUCAGUGGGAAGGUUCG	760	55040289	77.53
gRNA216	1477	UAUUCAGUGGGAAGGUUCGC	761	55040290	71.30
gRNA217	1478	AUUCAGUGGGAAGGUUCGCG	762	55040291	68.82
gRNA218	1479	AGUGGGAAGGUUCGCGGGGU	766	55040295	73.55
gRNA219	1480	GUGGGAAGGUUCGCGGGGUU	767	55040296	66.13
gRNA220	1481	AGGGCGAGCAGAGCACUGCC	808	55040337	68.85
gRNA221	1482	UUUCUGCCUCGCCGCGGCAC	853	55040382	65.59
gRNA222	1483	CUGCCUCGCCGCGGCACAGG	856	55040385	76.58
gRNA223	1484	UGCCUCGCCGCGGCACAGGU	857	55040386	75.75
gRNA224	1485	CACCCACCUGUGCCGCGGCG	862	55040391	81.79
gRNA225	1486	UCCUUCACCCACCUGUGCCG	867	55040396	70.86

gRNA No. or Condition	SEQ	Targeting Sequence	bp to TSS	GRCh38 sequence start coordinate	% Control PCSK9
gRNA226	1487	UGGGUGAAGGAGUGAAUGCC	876	55040405	57.69

[0267] Best-performing gRNAs were found to closely align to the *PCSK9* gene transcription start site (**FIG. 3**).

[0268] Following this primary screen, a secondary screen was performed with the top 40 gRNAs in RNA form. The top 40 guides were chemically synthesized and co-transfected with *in vitro* transcribed mRNA encoding the CRISPR-off, CRISPRi or WT Cas9 constructs. Secreted PCSK9 levels were measured 7 and 28 days after transfection.

[0269] To generate *in vitro* transcribed CRISPR-off, CRISPRi and WT Cas9 effector mRNA, plasmid constructs encoding these proteins were linearized using MfeI restriction enzyme from NEB® (Catalog No. R3589S). 1 μg of linearized template was used to set up *in vitro* transcription reactions using T7 mScript™ Standard mRNA Production System from CellScript (Catalog No. C-MSC100625) according to manufacturer's instructions. The resulting RNA had a Cap 1 structure on the 5' end and was 3' polyadenlylated. The transcribed RNA was purified using the RNeasy® Mini Kit from Qiagen (Catalog No. 74104).

[0270] End-modified sgRNAs purified using standard desalting were obtained from Integrated DNA Technologies. The three nucleotides at the 5' end and the three nucleotides at the 3' end of each guide were 2'-O-methyl modified. The three internucleoside linkages at the 3' end and the three internucleoside linkages at the 5' end were phosphorothioate internucleoside linkages (**Table 8**; SEQ: SEQ ID NO). In the Table, mX (i.e., mA, mC, mG, or mU) represents a 2'-O-methyl modified ribonucleoside, rX (i.e., rA, rC, rG, or rU) indicates a natural ribonucleoside, and * indicates a phosphorothioate linkage. All internucleoside linkages that are not phosphorothioate linkages are phosphate linkages.

Table 8. Targeting Sequences of Top 40 gRNAs from Secondary HeLa Cell Screen

gRNA No.	SEQ	Targeting Sequence	Modified gRNA No.	Full Modified RNA Targeting Sequence
gRNA 009	1270	ACUGCCUGGCU CACUCCUCC	modRNA009	mA*mC*mU*rGrCrCrUrGrGrCrUrCr ArCrUrCrCrUrCrCrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU

gRNA No.	SEQ	Targeting Sequence	Modified gRNA No.	Full Modified RNA Targeting Sequence
gRNA 108	1369	GUCGAGGCGCU CAUGGUUGC	modRNA108	mG*mU*mC*rGrArGrGrCrGrCrUrCr ArUrGrGrUrUrGrCrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 003	1264	UGCGGAAACCU UCUAGGGUG	modRNA003	mU*mG*mC*rGrGrArArArCrCrUrUr CrUrArGrGrGrUrGrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 128	1389	UGAAGGGGCGC GCGGAAUCC	modRNA128	mU*mG*mA*rArGrGrGrGrCrGrCrGr CrGrGrArArUrCrCrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 110	1371	AGGCGCUCAUG GUUGCAGGC	modRNA110	mA*mG*mG*rCrGrCrUrCrArUrGrGr UrUrGrCrArGrGrCrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 093	1354	AACCUGAUCCU CCAGUCCGG	modRNA093	mA*mA*mC*rCrUrGrArUrCrCrUrCr CrArGrUrCrCrGrGrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 096	1357	CAUUAACGGAA CCCCCGGAC	modRNA096	mC*mA*mU*rUrArArCrGrGrArArCr CrCrCrCrGrGrArCrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU

gRNA No.	SEQ	Targeting Sequence	Modified gRNA No.	Full Modified RNA Targeting Sequence
gRNA 092	1353	CGGACUGGAGG AUCAGGUUU	modRNA092	mC*mG*mG*rArCrUrGrGrArGrGrAr UrCrArGrGrUrUrUrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 011	1272	AUCGUCCGAUG GGGCUCUGG	modRNA011	mA*mU*mC*rGrUrCrCrGrArUrGrGr GrGrCrUrCrUrGrGrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 007	1268	GGUGUGGGUGC UUGACGCCU	modRNA007	mG*mG*mU*rGrUrGrGrGrUrGrCrUr UrGrArCrGrCrCrUrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 077	1338	CAAAUCCUAAC UGGGCUGGA	modRNA077	mC*mA*mA*rArUrCrCrUrArArCrUr GrGrGrCrUrGrGrArGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 113	1374	GGGCGCCGCCG UUCAGUUCA	modRNA113	mG*mG*mG*rCrGrCrCrGrCrCrGrUr UrCrArGrUrUrCrArGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 094	1355	AACCCCCGGAC UGGAGGAUC	modRNA094	mA*mA*mC*rCrCrCrCrGrGrArCrUr GrGrArGrGrArUrCrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU

gRNA No.	SEQ	Targeting Sequence	Modified gRNA No.	Full Modified RNA Targeting Sequence
gRNA 125	1386	GAGACUGGCUC GGGCGGGCC	modRNA125	mG*mA*mG*rArCrUrGrGrCrUrCrGr GrGrCrGrGrGrCrCrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 122	1383	GGCAGUGAGAC UGGCUCGGG	modRNA122	mG*mG*mC*rArGrUrGrArGrArCrUr GrGrCrUrCrGrGrGrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 004	1265	GCGGAAACCUU CUAGGGUGU	modRNA004	mG*mC*mG*rGrArArArCrCrUrUrCr UrArGrGrGrUrGrUrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 107	1368	UGAGCGCCUCG ACGUCGCUG	modRNA107	mU*mG*mA*rGrCrGrCrCrUrCrGrAr CrGrUrCrGrCrUrGrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 095	1356	UAACGGAACCC CCGGACUGG	modRNA095	mU*mA*mA*rCrGrGrArArCrCrCrCr CrGrGrArCrUrGrGrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 008	1269	GUGUGGGUGCU UGACGCCUG	modRNA008	mG*mU*mG*rUrGrGrGrUrGrCrUrUr GrArCrGrCrCrUrGrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU

gRNA No.	SEQ	Targeting Sequence	Modified gRNA No.	Full Modified RNA Targeting Sequence
gRNA 012	1273	AGGAUCGUCCG AUGGGGCUC	modRNA012	mA*mG*mG*rArUrCrGrUrCrCrGrAr UrGrGrGrGrCrUrCrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 126	1387	GGGACGCGUCG UUGCAGCAG	modRNA126	mG*mG*mG*rArCrGrCrGrUrCrGrUr UrGrCrArGrCrArGrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 111	1372	AGUUCAGGGUC UGAGCCUGG	modRNA111	mA*mG*mU*rUrCrArGrGrGrUrCrUr GrArGrCrCrUrGrGrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 212	1473	CGUGCGGCUGC GCUAUUCAG	modRNA212	mC*mG*mU*rGrCrGrGrCrUrGrCrGr CrUrArUrUrCrArGrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 170	1431	UUCGGCCAGGC CGUCCUCCU	modRNA170	mU*mU*mC*rGrGrCrCrArGrGrCrCr GrUrCrCrUrCrCrUrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 136	1397	GCACAGUCCUC CCCACCGCA	modRNA136	mG*mC*mA*rCrArGrUrCrCrUrCrCr CrCrArCrCrGrCrArGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU

gRNA No.	SEQ	Targeting Sequence	Modified gRNA No.	Full Modified RNA Targeting Sequence
gRNA 144	1405	GGCUCAAGGCG CCGCCGGCG	modRNA144	mG*mG*mC*rUrCrArArGrGrCrGrCr CrGrCrCrGrGrCrGrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 005	1266	UCAAGCACCCA CACCCUAGA	modRNA005	mU*mC*mA*rArGrCrArCrCrCrArCr ArCrCrCrUrArGrArGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 102	1363	UGGCGUGAUCU GCGCGCCCC	modRNA102	mU*mG*mG*rCrGrUrGrArUrCrUrGr CrGrCrGrCrCrCrCrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 013	1274	UCAGAUAGGAU CGUCCGAUG	modRNA013	mU*mC*mA*rGrArUrArGrGrArUrCr GrUrCrCrGrArUrGrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 204	1465	UGCAGGCAAGG CGGCGGGGG	modRNA204	mU*mG*mC*rArGrGrCrArArGrGrCr GrGrCrGrGrGrGrGrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 006	1267	GGGUGUGGGUG CUUGACGCC	modRNA006	mG*mG*mG*rUrGrUrGrGrGrUrGrCr UrUrGrArCrGrCrCrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU

gRNA No.	SEQ	Targeting Sequence	Modified gRNA No.	Full Modified RNA Targeting Sequence
gRNA 210	1471	AGCACGGUGGA GAGCGGGGA	modRNA210	mA*mG*mC*rArCrGrGrUrGrGrArGr ArGrCrGrGrGrGrArGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 109	1370	GAGGCGCUCAU GGUUGCAGG	modRNA109	mG*mA*mG*rGrCrGrCrUrCrArUrGr GrUrUrGrCrArGrGrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 060	1321	GCAGGAGACAA GGUGGACCC	modRNA060	mG*mC*mA*rGrGrArGrArCrArArGr GrUrGrGrArCrCrCrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 015	1276	CGGAAUCCUGG CUGGGAGCU	modRNA015	mC*mG*mG*rArArUrCrCrUrGrGrCr UrGrGrGrArGrCrUrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 001	1262	GGUGCUAGCCU UGCGUUCCG	modRNA001	mG*mG*mU*rGrCrUrArGrCrCrUrUr GrCrGrUrUrCrCrGrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 197	1458	UCUUCGCUUGG CACGAUCUU	modRNA197	mU*mC*mU*rUrCrGrCrUrUrGrGrCr ArCrGrArUrCrUrUrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArGrCrArArGrUrUrUrAr ArArUrArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU

gRNA No.	SEQ	Targeting Sequence	Modified gRNA No.	Full Modified RNA Targeting Sequence
gRNA 127	1388	GCUGCUGCAAC GACGCGUCC	modRNA127	mG*mC*mU*rGrCrUrGrCrArArCrGr ArCrGrCrGrUrCrCrGrUrUrUrArArGr ArGrCrUrArArGrCrUrGrGrArArArCr ArGrCrArUrArArGrCrArArGrUrUrUrAr ArArUrArArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 180	1441	CCCGCACCUUG GCGCAGCGG	modRNA180	mC*mC*mC*rGrCrArCrCrUrUrGrGr CrGrCrArGrCrGrGrGrUrUrUrArArGr ArGrCrUrArArArGrCrUrGrGrArArArCr ArGrCrArUrArArGrCrArArGrUrUrUrAr ArArUrArArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU
gRNA 112	1373	CGGGCGCCGCC GUUCAGUUC	modRNA112	mC*mG*mG*rGrCrGrCrCrGrCrCrGr UrUrCrArGrUrUrCrGrUrUrUrArArGr ArGrCrUrArArArGrCrUrGrGrArArArCr ArGrCrArUrArArGrCrArArGrUrUrUrAr ArArUrArArArGrGrCrUrArGrUrCrCrGr UrUrArUrCrArArCrUrUrGrArArArArAr ArGrUrGrGrCrArCrCrGrArGrUrCrGr GrUrGrCrUrUrU*mU*mU*mU

[0271] HeLa cells were reverse transfected with 25 ng effector and 12.5 ng sgRNA in a 96 well plate format using TransIT®-X2 transfection reagent from Mirus (Catalog No. MIR6003). Conditioned media was harvested every week for up to four weeks and used to measure secreted PCSK9 levels using LEGEND MAX™ Human PCSK9 ELISA Kit from BioLegend (Catalog No. 443107). ELISA data were normalized for cell numbers using the CellTiter-Glo® kit from Promega (Catalog No. G7571). While *PCSK9* silencing was transient with CRISPRi (dCas9-KRAB), and returned to baseline by day 28, several sgRNAs co-transfected with CRISPR-off (DNMT3A-3L-dCas9-KRAB) construct showed robust and durable silencing (**FIG. 4A**). 16 out of the 40 guides tested with CRISPR-off showed silencing efficiency greater than WT Cas9 (**Table 9**).

Table 9. Relative *PCSK9* Expression in HeLa Cells Treated with Modified gRNAs and CRISPR-off

Condition	D28, % of control (Average)
Negative Controls	
Transfection reagent only	100
CRISPRi IVT + modRNA009	108.1
CRISPRi IVT + modRNA004	88.1
Positive Control	
WT Cas9 IVT + modRNA180	10.5
Experimental Conditions (CRISPR-off and indicated modified gRNA)	
modRNA009	5.7
modRNA108	12.9
modRNA003	7.5
modRNA128	16.2
modRNA110	5.7
modRNA093	11.7
modRNA096	20.7
modRNA092	12.7
modRNA011	9.6
modRNA007	11.6
modRNA077	4.2
modRNA113	4.7
modRNA094	12.0
modRNA125	67.4
modRNA122	9.5
modRNA004	1.0
modRNA107	7.7
modRNA095	30.5
modRNA008	4.7
modRNA012	6.0
modRNA126	6.1
modRNA111	4.0
modRNA212	25.7
modRNA170	14.8
modRNA136	17.2
modRNA144	35.0
modRNA005	7.9
modRNA102	25.7
modRNA013	5.2
modRNA204	64.4

modRNA006	36.6
modRNA210	30.0
modRNA109	25.3
modRNA060	33.3
modRNA015	19.4
modRNA001	18.2
modRNA197	45.1
modRNA127	13.4
modRNA180	30.2
modRNA112	6.1

[0272] At the two-week time point, RNA was extracted using Quick-RNA 96 Kit from Zymo Research (Cat# R1053). qPCR was performed using qScript XLT One-Step RT-qPCR ToughMix from Quantabio (Cat# 95134-500) and TaqMan assays (PCSK9: Hs00545399_m1, PPIA: Hs99999904_m1). PCSK9 levels were normalized with PPIA. Relative quantification was done using delta-delta Ct method.

[0273] Suppression of PCSK9 secretion was found to correlate with mRNA silencing at day 14 (**FIG. 4B**).

[0274] modRNA004 and modRNA111 were tested in HeLa cells for suppressing PCSK9 secretion over 60 days (**FIG. 5**). WT Cas9 was co-transfected with modRNA180 as a positive control. Cells were treated with 25 ng of the effector and 12.5 ng of the gRNA. modRNA004 and modRNA111 were shown to mediate durable silencing of *PCSK9*, comparable to what was achieved by WT Cas9 in HeLa cells via gene editing.

[0275] Furthermore, results suggest that epigenetic silencing is maintained in HeLa cells treated with simvastatin. Statin treatment is known to increase PCSK9 secretion via a transcriptional mechanism. In epigenetically silenced HeLa cells, statin treatment was shown to not increase PCSK9 secretion (**FIG. 6**).

Example 5: Guide RNA Assays in Huh7 Hepatoma Cell Line

[0276] The Huh7 hepatoma cell line is amenable to high-throughput screening and transfection. The top 13 guides from the HeLa screen that had either perfect homology or a single mismatch with the cynomolgus *PCSK9* gene were tested in Huh7 hepatoma cell line. Epigenetic silencing with the CRISPR-off construct was shown to be stable over seven days **(FIG. 7)**.

Example 6: Guide RNA Assays in Primary Human and Cynomolgus Hepatocytes

[0277] Primary human and cynomolgus HepatoPac® cultures from BioIVT are used to test the efficacy of the gRNAs in primary hepatocytes. HepatoPac® cultures are maintained according to manufacturer recommendations. Briefly, HepatoPac® maintenance media is thawed and made up within 30 minutes of cells arrival. Upon media change, cells are allowed to acclimate for two days in 37°C, 10% CO_2 incubator. On the second day after receipt, the manufacturer's instructions are followed to formulate LNP's with CRISPR-off + sgRNA, GFP-mRNA and WT CRISPR Cas9 in various concentrations using the SPARK™ (Precision Nanosystems) and the hepato9 mRNA LNP formulation kit (CAT. Number NWS0016). Subsequently, the LNP's are characterized for encapsulation efficiency and total mRNA payload delivery via the Quant-it™ RiboGreen RNA assay kit. LNPs are then added to the media in specified quantities of total mRNA to achieve clinically relevant levels of silencing of PCSK9. Media is changed every other day for a duration of up to four weeks to assess durability and/or inheritability of silencing. *PCSK9* silencing will be assessed every seven days by ELISA to measure secreted PCSK9 levels in the media. PCSK9 concentration will be controlled to total hepatocytes using a human albumin ELISA (Thermo Fisher®). Data is then presented as total PCSK9 secretion as percent of GFP-mRNA negative control. Specificity may be assessed by isolating the primary human and cynomolgus hepatocytes from the mouse fibroblast feeder layer using a magnetic bead based antibody approach (Miltenyi Biotec). Following separation of primary hepatocytes from the feeder layer, cells are processed for RNAseq evaluation and genome-wide bisulfite sequencing.

[0278] The top 13 gRNAs in RNA format are selected to be tested in primary human hepatocytes (PHH). The top gRNAs are selected based on (i) *PCSK9* silencing efficiency and durability in HeLa cells (ii) whether they have a perfect alignment with the human *PCSK9* gene and up to one mismatch with the non-human primate *PCSK9* gene. Combinations of gRNAs are also tested to determine their efficacy and durability. The negative controls are CRISPR-off only, gRNA Fragment Only (modRNA003), and CRISPRi only. The positive control is CRISPRi co-transfected with modRNA004 (**Table 10**; SEQ: SEQ ID NO; NHP: non-human primate). All tested gRNAs are predicted to bind to both human and non-human primate *PCSK9*.

Table 10. gRNAs Being Screened in Primary Human Hepatocytes

gRNA No.	SEQ	gRNA Targeting Sequence	TSS distance	Modified RNA No.	Match to NHP
gRNA009	1270	ACUGCCUGGCUCACUCCUCC	81	modRNA009	Exact
gRNA003	1264	UGCGGAAACCUUCUAGGGUG	0	modRNA003	Exact
gRNA093	1354	AACCUGAUCCUCCAGUCCGG	-104	modRNA093	Exact
gRNA011	1272	AUCGUCCGAUGGGGCUCUGG	-56	modRNA011	Exact
gRNA007	1268	GGUGUGGGUGCUUGACGCCU	-16	modRNA007	Exact
gRNA077	1338	CAAAUCCUAACUGGGCUGGA	-181	modRNA077	1 mismatch
gRNA113	1374	GGGCGCCGCCGUUCAGUUCA	49	modRNA113	1 mismatch
gRNA004	1265	GCGGAAACCUUCUAGGGUGU	-1	modRNA004	Exact
gRNA008	1269	GUGUGGGUGCUUGACGCCUG	-17	modRNA008	Exact
gRNA012	1273	AGGAUCGUCCGAUGGGGCUC	-59	modRNA012	Exact
gRNA111	1372	AGUUCAGGGUCUGAGCCUGG	43	modRNA111	1 mismatch
gRNA005	1266	UCAAGCACCCACACCCUAGA	-11	modRNA005	Exact
gRNA013	1274	UCAGAUAGGAUCGUCCGAUG	-65	modRNA013	Exact

[0279] Robust *PCSK9* silencing is observed. For some gRNAs, >70% reduction in secreted PCSK9 at day 7 is observed (depending on transfection efficiency).

Example 7: ZF Assays in HeLa Cells

[0280] A total of 209 zinc finger proteins (architecture as shown in SEQ ID NO: 659) were designed from the ZF library to 49 *PCSK9* target sites (selected from GRCh38 chromosome 1 between 55038548 to 55040548). The target sites had no other exact matches in the human genome (GRCh38).

[0281] HeLa cells were transfected with ZF-off constructs in DNA format. Six 96-well plates (Sigma-Aldrich Catalog No. M2936) were seeded with 12,000 HeLa cells per well (ATCC Catalog No. CCL-2) in standard culture media containing DMEM (Thermo Fisher Catalog No. 11-965-092) supplemented with 10% Fetal Bovine Serum (Thermo Fisher Catalog No. A4766) v/v, 1x GlutaMAX™ (Thermo Fisher Catalog No. 35050061) and 1x Penicillin-Streptomycin (Thermo Fisher Catalog No. 15140122). Following plating, cells were allowed to grow for 24 hours in a 37° C incubator at 5% CO_2. 10 ng of the ZF-off plasmid was resuspended in DPBS buffer (Thermo Fisher Catalog No. 14190144) to a concentration of 7.5 ng/μL. In addition, 10 ng of EF1a:PuromycinResistance plasmid and 65 ng of empty vector (pUC19) were also added to the transfection mix to achieve a total payload of 85 ng of DNA. Transfection mixtures were created by adding resuspended DNA in serum-free OPTI-MEM media (Thermo Fisher® Catalog No. 31985062) and adding Mirus® TransIT®-LT1 transfection reagent (MIR2300) following the manufacturer's

instructions. 10 μL of transfection mixtures were added in duplicate across a total of six screening plates. The positive control was CRISPR-off (SEQ ID NO: 658) with a high performing gRNA (gRNA009). Negative controls were ZF-off with a non-*PCSK9* locus target (CLTA) and empty vector (pUC19; NEB Catalog No. N3041S).

[0282] The ZF screen yielded hits with activity comparable to CRISPR (**FIG. 8**). The candidates with high silencing efficiency are taken forward to follow-up experiments. **FIG. 9** shows the ZF screening results by distance to TSS. In total, 209 ZFs were screened, with their PCSK9 knockdown activity relative to the negative control shown in **Table 11** below.

Table 11. ZF-off Construct Activity

ZF No.	% Control	ZF No.	% Control	ZF No.	% Control
ZF028	37.43	ZF071	66.51	ZF141	104.50
ZF001	35.90	ZF072	62.59	ZF142	71.70
ZF047	32.08	ZF073	69.50	ZF143	49.70
ZF048	15.19	ZF074	55.74	ZF144	44.01
ZF029	40.42	ZF075	51.58	ZF145	60.80
ZF030	7.74	ZF076	58.82	ZF146	42.85
ZF031	29.80	ZF077	49.52	ZF147	58.42
ZF032	40.65	ZF078	52.50	ZF148	70.50
ZF033	30.71	ZF079	66.30	ZF149	95.42
ZF034	31.95	ZF080	51.52	ZF150	74.00
ZF035	38.29	ZF081	140.85	ZF151	61.01
ZF002	33.02	ZF082	105.84	ZF152	52.34
ZF003	29.17	ZF083	72.30	ZF153	80.11
ZF004	35.39	ZF084	95.05	ZF154	79.46
ZF005	37.90	ZF085	44.18	ZF155	50.09
ZF006	37.47	ZF086	88.95	ZF156	46.99
ZF036	35.19	ZF087	44.87	ZF157	89.03
ZF037	28.76	ZF088	56.44	ZF158	99.89
ZF038	21.11	ZF089	116.41	ZF159	101.66
ZF039	9.35	ZF090	51.30	ZF160	119.21
ZF040	5.20	ZF091	57.46	ZF161	49.64
ZF041	30.50	ZF092	46.82	ZF162	68.98
ZF007	36.13	ZF093	44.03	ZF163	89.71
ZF008	34.72	ZF094	65.55	ZF164	103.15
ZF009	28.91	ZF095	56.92	ZF165	48.37
ZF010	33.58	ZF096	93.13	ZF166	81.13
ZF011	41.01	ZF097	57.08	ZF167	67.29
ZF012	19.39	ZF098	57.03	ZF168	50.88
ZF013	40.46	ZF099	59.36	ZF169	77.25
ZF014	33.06	ZF100	95.34	ZF170	108.08
ZF015	22.48	ZF101	57.13	ZF171	57.10
ZF016	34.09	ZF102	145.81	ZF172	53.73
ZF017	24.71	ZF103	63.08	ZF173	62.61
ZF018	24.84	ZF104	50.06	ZF174	71.84

ZF No.	% Control
ZF019	14.70
ZF020	33.15
ZF021	33.87
ZF022	29.02
ZF023	30.45
ZF024	12.55
ZF042	41.13
ZF025	39.37
ZF026	28.27
ZF043	40.92
ZF044	37.27
ZF045	39.72
ZF027	38.67
ZF046	36.12
ZF049	43.53
ZF050	44.02
ZF051	46.90
ZF052	45.89
ZF053	42.71
ZF054	44.31
ZF055	44.57
ZF056	44.82
ZF057	46.52
ZF058	46.61
ZF059	47.62
ZF060	44.75
ZF061	59.99
ZF062	48.69
ZF063	50.63
ZF064	54.60
ZF065	54.25
ZF066	54.95
ZF067	69.99
ZF068	54.64
ZF069	55.99
ZF070	61.26

ZF No.	% Control
ZF105	72.80
ZF106	75.56
ZF107	56.29
ZF108	103.79
ZF109	60.41
ZF110	53.71
ZF111	69.30
ZF112	94.36
ZF113	57.16
ZF114	51.72
ZF115	78.55
ZF116	76.55
ZF117	66.49
ZF118	60.58
ZF119	78.70
ZF120	63.82
ZF121	117.94
ZF122	82.54
ZF123	84.94
ZF124	71.86
ZF125	76.73
ZF126	119.64
ZF127	144.15
ZF128	66.74
ZF129	86.03
ZF130	103.82
ZF131	116.46
ZF132	130.63
ZF133	100.63
ZF134	42.35
ZF135	61.14
ZF136	47.14
ZF137	101.88
ZF138	99.22
ZF139	69.52
ZF140	48.41

ZF No.	% Control
ZF175	68.12
ZF176	51.52
ZF177	48.59
ZF178	71.10
ZF179	53.08
ZF180	54.56
ZF181	78.86
ZF182	70.20
ZF183	69.07
ZF184	49.73
ZF185	69.04
ZF186	57.09
ZF187	64.90
ZF188	59.91
ZF189	59.92
ZF190	79.85
ZF191	71.95
ZF192	78.47
ZF193	74.79
ZF194	67.87
ZF195	83.86
ZF196	67.71
ZF197	64.36
ZF198	62.06
ZF199	59.08
ZF200	53.09
ZF201	57.39
ZF202	63.53
ZF203	69.94
ZF204	67.61
ZF205	58.51
ZF206	77.67
ZF207	72.94
ZF208	55.34
ZF209	64.42

[0283] Several ZF-off constructs were shown to be more effective at silencing *PCSK9* than WTCas9 and CRISPR-off in combination with gRNA003. The target sites and ZF sequences (F1 through F6) of the ZFP domains in these ZF-off constructs are as shown in **Table 1**.

Example 8: Full Specificity Screen of Constructs in Primary Human Hepatocytes

[0284] The specificity of CRISPR-off and ZF-off constructs for silencing *PCSK9* is tested in primary human hepatocytes. The readouts to assess specificity are RNAseq,

methylation array and whole genome bisulfite sequencing assays. Genome-wide expression and methylation changes after epigenetic editing compared to negative controls will be profiled.

Example 9: CpG Methylation Patterns

[0285] The CpG methylation patterns in human hepatocytes (e.g., primary cells or cell lines) treated with CRISPR-off or ZF-off are investigated. Hybrid capture assay is performed on bisulfite treated DNA to investigate methylation patterns at CpG sites that are induced by CRISPR-Off or ZF-Off at the 1kb region around the *PCSK9* TSS.

Example 10: Stable *PCSK9* Silencing via Epigenetic Editing in Mice with Wildtype *PCSK9*

[0286] The ability of CRISPR-off and ZF-off constructs to mediate epigenetic silencing of endogenous *PCSK9 in vivo* is tested. Constructs are delivered using a single IV administration of mRNA (and, for CRISPR-off silencing, gRNA) formulated into an LNP. Silencing is tested in wildtype mice over a period of two to six months. The readout is serum PCSK9 levels and serum cholesterol levels. A subset of each cohort is selected for liver hematoxylin and eosin (H&E) stain RNAseq and analysis. For several constructs, robust, stable, and inheritable *PCSK9* silencing is observed.

Example 11: Stable *PCSK9* Silencing via Epigenetic Editing in Mice Expressing Transgenic Human *PCSK9*

[0287] Three different mouse strains expressing transgenic human *PCSK9* are used: hPCSK9-Tg (mPCSK9+/-) heterozygous mouse, hPCSK9-Tg (mPCSK9+/+) homozygous mouse, and hPCSK9-Tg (mPCSK9-/-) mouse. The hPCSK9-Tg (mPCSK9-/-) mouse line used is C57BL/6J-Pcsk9-/- Tg(RP11-55M23-AbsI), which expresses human *PCSK9* under the control of its own promoter (**FIG. 10**). *See, e.g.*, Weider et al., *J Biol Chem* (2016) 291(32):16659-71.

[0288] The CRISPR-off and ZF-off constructs are tested. Constructs are delivered via single IV administration of mRNA/gRNA formulated into LNP. The readouts are liver H&E stain, RNAseq to measure *PCSK9* mRNA levels, and AST/ASL measurements. Efficacy is also tested, including durability of *PCSK9* silencing over three to four months as measured by the level of serum PCSK9 protein. A durable and significant reduction in the levels of serum PCSK9 is observed for some constructs.

[0289] Durability is tested over six to twelve months. Readouts are serum PCSK9 levels and serum cholesterol levels. A subset of the cohort is selected for liver H&E and RNAseq analysis.

Example 12: Fusion proteins with Variant NLS Configurations

[0290] Several improved fusion protein constructs were developed using variant nuclear localization sequence (NLS) configurations to have significantly higher epi-silencing activity.

[0291] Several constructs with variant configurations of NLS domains (**FIGs. 11A and 11B**) were constructed and tested in *PCSK9* loci in HeLa cells (**FIGs. 12A-12B**). The constructs were additionally tested in *PCSK9* loci in Hepa1-6 (**FIG. 13**) and in HuH7 (**FIGs. 14A-14C and FIG. 15**). Exemplary fusion protein construct amino acid and DNA sequences are shown below:

Fusion Protein 1 Amino Acid Sequence NLS-3A-3L-dCas9-KRAB-NLS (SEQ ID NO: 1498)	MPKKKRKVPKKKRKVNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVD RYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPAR KGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDA KEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGK DQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLK EYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHNPLEMFETVPVWRRQPVRVLSLFEDIKKE LTSLGFLESGSDPGQLKHVVDVTDTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFH RLLQYARPKPGSPRPFFWMFVDNLVLNKEDLDVASRFLEMEPVTIPDVHGGSLQNAVRVWSN IPAIRSRHWALVSEEELSLLAQNKQSSKLAAKWPTKLVKNCFLPLREYFKYFSTELTSSLGG PSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTST EPSEGSAPGTSTEPSEMDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKN LIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVE EDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAK NLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSK NGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGE LHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEE VVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSG EQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDK DFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKL INGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKE LGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSI DNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKA GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVR EINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFF YSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEV QTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKE LLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGN ELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADAN LDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLI HQSITGLYETRIDLSQLGGDSGSETPGTSESATPESTGRTLVTFKDVFVDFTREEWKLLDTA QQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPPKKKRKVPKKKRKV
Fusion Protein 1 DNA Sequence (SEQ ID NO: 1499)	ATGCCAAAAAAGAAGAGAAAGGTACCGAAGAAAAAAAGAAAGGTCAACCATGATCAAGAATT CGACCCACCTAAAGTCTACCCACCTGTGCCCGCCGAAAAAAGGAAACCCATAAGGGTGCTGT CACTCTTTGATGGCATCGCCACTGGTCTCCTGGTTCTTAAGGATCTGGGAATTCAGGTCGAT CGGTACATTGCTAGCGAGGTTTGTGAGGATAGTATTACAGTGGGTATGGTGCGCCACCAGGG AAAGATCATGTATGTTGGTGACGTTAGGAGCGTCACCCAGAAACATATCCAGGAGTGGGGAC CCTTTGATTTGGTGATCGGAGGTAGTCCCTGCAATGACCTTTCCATCGTGAATCCAGCCAGG

	AAAGGGCTGTATGAAGGGACTGGTAGGCTCTTTTTCGAGTTTTATCGCCTGCTTCACGACGC
	TAGACCTAAGGAAGGTGACGATAGGCCTTTCTTTTGGCTTTTTGAGAACGTCGTGGCAATGG
	GAGTCTCCGACAAAAGGGACATTTCTCGCTTTCTGGAATCTAACCCCGTTATGATCGATGCC
	AAGGAAGTTTCTGCCGCTCACAGGGCAAGGTACTTCTGGGGCAATCTGCCCGGAATGAATCG
	CCCACTGGCCAGTACCGTGAATGACAAACTGGAGCTGCAGGAGTGCCTGGAGCACGGAAGAA
	TCGCAAAGTTTTCTAAAGTCAGGACCATTACCACTCGCAGTAACTCCATAAAACAGGGTAAG
	GACCAGCATTTTCCCGTCTTCATGAATGAAAAGGAAGATATTCTGTGGTGCACTGAAATGGA
	GAGAGTTTTCGGGTTTCCCGTGCACTATACCGATGTTTCCAACATGTCCCGCCTTGCAAGAC
	AAAGGCTTTTGGGCCGCTCTTGGTCTGTGCCAGTGATCCGGCACTTGTTTGCTCCCCTCAAA
	GAGTACTTCGCTTGCGTCAGTTCCGGAAATTCAAACGCTAACTCTCGGGGTCCATCTTTCTC
	CAGTGGTCTCGTGCCACTGTCTCTCCGGGGCTCTCACAATCCCCTGGAGATGTTTGAGACAG
	TGCCAGTCTGGCGGAGGCAGCCCGTTCGCGTTCTCTCTCTGTTCGAAGATATTAAAAAGGAA
	CTCACCTCCCTTGGGTTCCTGGAGAGCGGGAGCGACCCCGGACAGCTTAAGCACGTGGTCGA
	CGTGACTGACACCGTCCGCAAAGACGTGGAGGAATGGGGCCCCTTCGATCTGGTCTATGGGG
	CAACCCCTCCCCTTGGGCATACATGTGATCGGCCTCCATCCTGGTACCTGTTCCAGTTTCAC
	AGACTCCTGCAGTATGCCAGGCCAAAGCCAGGGAGCCCAAGGCCCTTTTTCTGGATGTTCGT
	CGACAACCTGGTCCTGAACAAAGAAGATCTCGACGTTGCTAGTCGCTTTCTCGAAATGGAGC
	CCGTGACCATTCCCGACGTGCATGGCGGTTCCCTCCAGAATGCAGTCAGGGTTTGGAGCAAT
	ATCCCTGCCATCAGGTCAAGGCACTGGGCACTGGTTTCAGAGGAAGAGCTGTCCCTCCTTGC
	CCAGAACAAGCAGTCATCCAAACTGGCAGCCAAGTGGCCAACTAAGCTGGTCAAGAACTGCT
	TTCTTCCCCTCAGAGAATATTTTAAGTATTTCAGTACTGAACTGACTAGCAGTCTGGGAGGG
	CCGAGCTCTGGCGCACCCCCACCAAGTGGAGGGTCTCCTGCCGGGTCCCCAACATCTACTGA
	AGAAGGCACCAGCGAATCCGCAACGCCCGAGTCAGGCCCTGGTACCTCCACAGAACCATCTG
	AAGGTAGTGCGCCTGGTTCCCCAGCTGGAAGCCCTACTTCCACCGAAGAAGGCACGTCAACC
	GAACCAAGTGAAGGATCTGCCCCTGGGACCAGCACTGAACCATCTGAGATGGACAAGAAGTA
	CAGCATCGGCCTGGCCATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACA
	AGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAAC
	CTGATCGGCGCCCTGCTGTTCGACAGCGGAGAAACAGCCGAGGCCACCCGGCTGAAGAGAAC
	CGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCA
	ACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAA
	GAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCA
	CGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCACCGACAAGGCCG
	ACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATC
	GAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAGAC
	CTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCC
	TGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAG
	AAGAAGAATGGCCTGTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAA
	GAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACG
	ACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAG
	AACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGC
	CCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGA
	AAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAG
	AACGGCTACGCCGGCTACATCGATGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAA
	GCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACC
	TGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAG
	CTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAA
	GATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACA
	GCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAA
	GTGGTGGACAAGGGCGCCAGCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAA
	CCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTACA
	ACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGC
	GAGCAGAAAAAAGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCA
	GCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGG
	AAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGGACAAG
	GACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCCTGACACT
	GTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCTGTTCGACGACA
	AAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTG
	ATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGG
	CTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACA
	TCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCC
	GGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAA
	AGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACCAGACCA

	CCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAG CTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCT GTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGGACATCAACC GGCTGTCCGACTACGATGTGGACGCTATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATC GATAACAAAGTGCTGACTCGGAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGA AGAGGTCGTGAAGAAGATGAAGAACTACTGGCGCCAGCTGCTGAATGCCAAGCTGATTACCC AGAGGAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCC GGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCT GGACTCCCGGATGAACACTAAGTACGACGAGAACGACAAACTGATCCGGGAAGTGAAAGTGA TCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGC GAGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCT GATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACG TGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTC TACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAA GCGGCCTCTGATCGAGACAAACGGCGAAACAGGCGAGATCGTGTGGGATAAGGGCCGGGACT TTGCCACCGTGCGGAAAGTGCTGTCTATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTG CAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGACAAGCTGATCGC CAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATT CTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAAGAG CTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGA AGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGT TCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAAC GAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCT GAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAACACTACC TGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAAT CTGGACAAGGTGCTGAGCGCCTACAACAAGCACAGAGACAAGCCTATCAGAGAGCAGGCCGA GAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTG ACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGATC CACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACAG CGGAAGTGAGACCCCAGGTACATCCGAATCAGCAACGCCTGAAAGCACCGGTCGGACACTGG TGACCTTCAAGGATGTATTTGTGGACTTCACCAGGGAGGAGTGGAAGCTGCTGGACACTGCT CAGCAGATCGTGTACAGAAATGTGATGCTGGAGAACTATAAGAACCTGGTTTCCTTGGGTTA TCAGCTTACTAAGCCAGATGTGATCCTCCGGTTGGAGAAGGGAGAAGAGCCCCCAAAAAAGA AGAGAAAGGTACCGAAGAAAAAAAGAAAGGTC
Fusion Protein 2 Amino Acid Sequence 3A-3L-NLS-dCas9-NLS-KRAB (SEQ ID NO: 1500)	MNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITV GMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEF YRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRARYFWG NLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDI LWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNAN SRGPSFSSGLVPLSLRGSHNPLEMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSDPG QLKHVVDVTDTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLLQYARPKPGSPR PFFWMFVDNLVLNKEDLDVASRFLEMEPVTIPDVHGGSLQNAVRVWSNIPAIRSRHWALVSE EELSLLAQNKQSSKLAAKWPTKLVKNCFLPLREYFKYFSTELTSSLGGPSSGAPPPSGGSPA GSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEP SEPKKKRKVMDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHER HPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPD NSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFG NLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAIL LSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYI DGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR QEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEE NEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDK QSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKK GILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILK EHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTR SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQL VETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHH AHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNF FKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSK

	ESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIM ERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSA YNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGL YETRIDLSQLGGDPKKKRKVSGSETPGTSESATPESTGRTLVTFKDVFVDFTREEWKLLDTA QQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEP
Fusion Protein 2 DNA Sequence (SEQ ID NO: 1501)	ATGAACCACGACCAGGAATTTGACCCTCCAAAGGTTTACCCACCTGTCCCAGCTGAGAAGAG GAAGCCCATCCGGGTGCTGTCTCTCTTTGATGGAATCGCTACAGGGCTCCTGGTGCTGAAGG ACTTGGGCATTCAGGTGGACCGCTACATTGCCTCGGAGGTGTGTGAGGACTCCATCACGGTG GGCATGGTGCGGCACCAGGGGAAGATCATGTACGTCGGGGACGTCCGCAGCGTCACACAGAA GCATATCCAGGAGTGGGGCCCATTCGATCTGGTGATTGGGGGCAGTCCCTGCAATGACCTCT CCATCGTCAACCCTGCTCGCAAGGGCCTCTACGAGGGCACTGGCCGGCTCTTCTTTGAGTTC TACCGCCTCCTGCATGATGCGCGGCCCAAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTT TGAGAATGTGGTGGCCATGGGCGTTAGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCA ACCCTGTGATGATTGATGCCAAAGAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGT AACCTTCCCGGTATGAACAGGCCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGA GTGTCTGGAGCATGGCAGGATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAA ACTCCATAAAGCAGGGCAAAGACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATC TTATGGTGCACTGAAATGGAAAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAA CATGAGCCGCTTGGCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCC ACCTCTTCGCTCCGCTGAAGGAGTATTTTGCGTGTGTGTCTAGCGGCAATAGTAACGCTAAC AGCCGCGGGCCGAGCTTCAGCAGCGGCCTGGTGCCGTTAAGCTTGCGCGGCAGCCATAATCC CCTTGAGATGTTCGAAACCGTGCCTGTGTGGAGGAGACAGCCAGTCCGGGTGCTGTCCCTTT TTGAAGACATCAAGAAAGAGCTGACGAGTTTGGGCTTTTTGGAAAGTGGTTCTGACCCGGGA CAACTGAAGCATGTGGTTGATGTCACAGACACAGTGAGGAAGGATGTGGAGGAGTGGGGACC CTTCGATCTTGTGTACGGCGCCACACCTCCCCTGGGCCACACCTGTGACCGTCCTCCCAGCT GGTACCTGTTCCAGTTCCACCGGCTCCTGCAGTACGCACGGCCCAAGCCAGGCAGCCCCAGG CCCTTCTTCTGGATGTTCGTGGACAATCTGGTGCTGAACAAGGAAGACCTGGACGTCGCATC TCGCTTCCTGGAGATGGAGCCAGTCACCATCCCAGATGTCCACGGCGGATCCTTGCAGAATG CTGTCCGCGTGTGGAGCAACATCCCAGCCATAAGGAGCAGGCACTGGGCTCTGGTTTCGGAA GAAGAATTGTCCCTGCTGGCCCAGAACAAGCAGAGCTCGAAGCTCGCGGCCAAGTGGCCCAC CAAGCTGGTGAAGAACTGCTTTCTCCCCCTAAGAGAATATTTCAAGTATTTTTCAACAGAAC TCACTTCCTCTTTAGGAGGGCCGAGCTCTGGCGCACCCCCACCAAGTGGAGGGTCTCCTGCC GGGTCCCCAACATCTACTGAAGAAGGCACCAGCGAATCCGCAACGCCCGAGTCAGGCCCTGG TACCTCCACAGAACCATCTGAAGGTAGTGCGCCTGGTTCCCCAGCTGGAAGCCCTACTTCCA CCGAAGAAGGCACGTCAACCGAACCAAGTGAAGGATCTGCCCCTGGGACCAGCACTGAACCA TCTGAGCCAAAAAAGAAGAGAAAGGTAATGGACAAGAAGTACAGCATCGGCCTGGCCATCGG CACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCA AGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGCGCCCTGCTGTTC GACAGCGGAGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAG ACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACG ACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGG CACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCACCATCTA CCACCTGAGAAAGAAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGG CCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGAC AACAGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGA AAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGA GCAGACGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGC AACCTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGA GGATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCC AGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATCCTG CTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGAT CAAGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGC TGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATC GATGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGA CGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCT TCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGG CAGGAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTT CCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGATTCGCCTGGATGACCA GAAAGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCCAGC GCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCT GCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTACAACGAGCTGACCAAAGTGAAAT

	ACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAAGCCATCGTG GACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAA GAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCC TGGGCACATACCACGATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAA AACGAGGACATTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGAT CGAGGAACGGCTGAAAACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGC GGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAG CAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCAT GCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCG GCCAGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAG GGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCC CGAGAACATCGTGATCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACA GCCGCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAA GAACACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAA TGGGCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGG ACGCTATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGATAACAAAGTGCTGACTCGG AGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAA GAACTACTGGCGCCAGCTGCTGAATGCCAAGCTGATTACCCAGAGGAAGTTCGACAATCTGA CCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTG GTGGAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAA GTACGACGAGAACGACAAACTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGG TGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCACCAC GCCCACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCT GGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGA GCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATGAACTTT TTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAA CGGCGAAACAGGCGAGATCGTGTGGGATAAGGGCCGGGACTTTGCCACCGTGCGGAAAGTGC TGTCTATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAA GAGTCTATCCTGCCCAAGAGGAACAGCGACAAGCTGATCGCCAGAAAGAAGGACTGGGACCC TAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAG TGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATG GAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGT GAAAAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGA AGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAA TATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAA TGAGCAGAAACAGCTGTTTGTGGAACAGCACAAACACTACCTGGACGAGATCATCGAGCAGA TCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAGGTGCTGAGCGCC TACAACAAGCACAGAGACAAGCCTATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTAC CCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGA GGTACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTG TACGAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACCCAAAAAAGAAGAGAAAGGTAAG CGGAAGTGAGACCCCAGGTACATCCGAATCAGCAACGCCTGAAAGCACCGGTCGGACACTGG TGACCTTCAAGGATGTATTTGTGGACTTCACCAGGGAGGAGTGGAAGCTGCTGGACACTGCT CAGCAGATCGTGTACAGAAATGTGATGCTGGAGAACTATAAGAACCTGGTTTCCTTGGGTTA TCAGCTTACTAAGCCAGATGTGATCCTCCGGTTGGAGAAGGGAGAAGAGCCC
Fusion Protein 3 Amino Acid Sequence 3A-ADD-3L-NLS-dCas9-NLS-KRAB (SEQ ID NO: 1502)	MNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITV GMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEF YRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRARYFWG NLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDI LWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNAN SRGPSFSSGLVPLSLRGSHMEVKVNRRSIEDICLCCGTLQVYTRHPLFEGGLCAPCKDKFLE SLFLYDDDGHQSYCTICCSGGTLFICESPDCTRCYCFECVDILVGPGTSERINAMACWVCFL CLPFSRSGLLQRRKRWRHQLKAFHDQEGAGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSL GFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRI LQYALPRQESQRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIP GLKSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGPSS GAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPS EGSAPGTSTEPSEPKKKRKVMDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHS IKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEES FLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRG HFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQ LPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLF

	LAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFF DQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQI HLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPW NFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPA FLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRL SRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHI ANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSE LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQF YKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVK KTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLK SVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGEL QKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVIL ADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD ATLIHQSITGLYETRIDLSQLGGDPKKKRKVSGSETPGTSESATPESTGRTLVTFKDVFVDF TREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEP
Fusion Protein 3 DNA Sequence (SEQ ID NO: 1503)	ATGAACCATGACCAGGAATTTGACCCCCCAAAGGTTTACCCACCTGTGCCAGCTGAGAAGAG GAAGCCCATCCGCGTGCTGTCTCTCTTTGATGGGATTGCTACAGGGCTCCTGGTGCTGAAGG ACCTGGGCATCCAAGTGGACCGCTACATTGCCTCCGAGGTGTGTGAGGACTCCATCACGGTG GGCATGGTGCGGCACCAGGGAAAGATCATGTACGTCGGGGACGTCCGCAGCGTCACACAGAA GCATATCCAGGAGTGGGGCCCATTCGACCTGGTGATTGGAGGCAGTCCCTGCAATGACCTCT CCATTGTCAACCCTGCCCGCAAGGGACTTTATGAGGGTACTGGCCGCCTCTTCTTTGAGTTC TACCGCCTCCTGCATGATGCGCGGCCCAAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTT TGAGAATGTGGTGGCCATGGGCGTTAGTGACAAGAGGGACATCTCGCGATTTCTTGAGTCTA ACCCCGTGATGATTGACGCCAAAGAAGTGTCTGCTGCACACAGGGCCCGTTACTTCTGGGGT AACCTTCCTGGCATGAACAGGCCTTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAAGA GTGTCTGGAGCACGGCAGAATAGCCAAGTTCAGCAAAGTGAGGACCATTACCACCAGGTCAA ACTCTATAAAGCAGGGCAAAGACCAGCATTTCCCCGTCTTCATGAACGAGAAGGAGGACATC CTGTGGTGCACTGAAATGGAAAGGGTGTTTGGCTTCCCCGTCCACTACACAGACGTCTCCAA CATGAGCCGCTTGGCGAGGCAGAGACTGCTGGGCCGATCGTGGAGCGTGCCGGTCATCCGCC ACCTCTTCGCTCCGCTGAAGGAATATTTTGCTTGTGTGTCTAGCGGCAATAGTAACGCTAAC AGCCGCGGGCCGAGCTTCAGCAGCGGCCTGGTGCCGTTAAGCTTGCGCGGCAGCCATATGGA AGTCAAAGTGAACCGACGGAGCATTGAAGACATCTGCCTCTGCTGTGGAACTCTCCAGGTGT ACACTCGGCACCCCTTGTTTGAGGGAGGGTTATGTGCCCCATGTAAGGATAAGTTCCTGGAG TCCCTCTTCCTGTATGATGATGATGGACACCAGAGTTACTGCACCATCTGCTGTTCCGGGGG TACCCTGTTCATCTGTGAGAGCCCCGACTGTACCAGATGCTACTGTTTCGAGTGTGTGGACA TCCTGGTGGGCCCCGGGACCTCAGAGAGGATCAATGCCATGGCCTGCTGGGTTTGCTTCCTG TGCCTGCCCTTCTCACGGAGTGGACTGCTGCAGAGGCGCAAGAGGTGGCGGCACCAGCTGAA GGCCTTCCATGATCAAGAGGGAGCGGGCCCTATGGAGATATACAAGACAGTGTCTGCATGGA AGAGACAGCCAGTGCGGGTACTGAGCCTCTTCAGAAACATCGACAAGGTACTAAAGAGTTTG GGCTTCTTGGAAAGCGGTTCTGGTTCTGGGGGAGGAACGCTGAAGTACGTGGAAGATGTCAC AAATGTCGTGAGGAGAGACGTGGAGAAATGGGGCCCCTTTGACCTGGTGTACGGCTCGACGC AGCCCCTAGGCAGCTCTTGTGATCGCTGTCCCGGCTGGTACATGTTCCAGTTCCACCGGATC CTGCAGTATGCGCTGCCTCGCCAGGAGAGTCAGCGGCCCTTCTTCTGGATATTCATGGACAA TCTGCTGCTGACTGAGGATGACCAAGAGACAACTACCCGCTTCCTTCAGACAGAGGCTGTGA CCCTCCAGGATGTCCGTGGCAGAGACTACCAGAATGCTATGCGGGTGTGGAGCAACATTCCA GGGCTGAAGAGCAAGCATGCGCCCCTGACCCCAAAGGAAGAAGAGTATCTGCAAGCCCAAGT CAGAAGCAGGAGCAAGCTGGACGCCCCGAAAGTTGACCTCCTGGTGAAGAACTGCCTTCTCC CGCTGAGAGAGTACTTCAAGTATTTTTCTCAAAACTCACTTCCTCTTGGAGGGCCGAGCTCT GGCGCACCCCCACCAAGTGGAGGGTCTCCTGCCGGGTCCCCAACATCTACTGAAGAAGGCAC CAGCGAATCCGCAACGCCCGAGTCAGGCCCTGGTACCTCCACAGAACCATCTGAAGGTAGTG CGCCTGGTTCCCCAGCTGGAAGCCCTACTTCCACCGAAGAAGGCACGTCAACCGAACCAAGT GAAGGATCTGCCCCTGGGACCAGCACTGAACCATCTGAGCCAAAAAAGAAGAGAAAGGTAAT GGACAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACTCTGTGGGCTGGGCCGTGATCA CCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGC ATCAAGAAGAACCTGATCGGCGCCCTGCTGTTCGACAGCGGAGAAACAGCCGAGGCCACCCG GCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAG AGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCC TTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGA

	GGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCA
	CCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGC
	CACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCA
	GCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACG
	CCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAG
	CTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGAC
	CCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACA
	CCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTT
	CTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGA
	GATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACC
	TGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTC
	GACCAGAGCAAGAACGGCTACGCCGGCTACATCGATGGCGGAGCCAGCCAGGAAGAGTTCTA
	CAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGA
	ACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATC
	CACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGA
	CAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGG
	CCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGG
	AACTTCGAGGAAGTGGTGGACAAGGGCGCCAGCGCCCAGAGCTTCATCGAGCGGATGACCAA
	CTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACT
	TCACCGTGTACAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCC
	TTCCTGAGCGGCGAGCAGAAAAAAGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGT
	GACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAA
	TCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATT
	ATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCT
	GACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACC
	TGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTG
	AGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCT
	GAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCT
	TTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATT
	GCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGA
	CGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAG
	AGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAG
	GGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCA
	GAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAAC
	TGGACATCAACCGGCTGTCCGACTACGATGTGGACGCTATCGTGCCTCAGAGCTTTCTGAAG
	GACGACTCCATCGATAACAAAGTGCTGACTCGGAGCGACAAGAACCGGGGCAAGAGCGACAA
	CGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGCCAGCTGCTGAATGCCA
	AGCTGATTACCCAGAGGAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAA
	CTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGT
	GGCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAACGACAAACTGATCCGGG
	AAGTGAAAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTT
	TACAAAGTGCGCGAGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGT
	GGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACA
	AGGTGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCC
	AAGTACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGG
	CGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACAGGCGAGATCGTGTGGGATA
	AGGGCCGGGACTTTGCCACCGTGCGGAAAGTGCTGTCTATGCCCCAAGTGAATATCGTGAAA
	AAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGA
	CAAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCA
	CCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAG
	AGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCAT
	CGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTA
	AGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTG
	CAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCA
	CTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGC
	ACAAACACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTG
	GCCGACGCTAATCTGGACAAGGTGCTGAGCGCCTACAACAAGCACAGAGACAAGCCTATCAG
	AGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCT
	TCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGAC
	GCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCT
	GGGAGGCGACCCAAAAAAGAAGAGAAAGGTAAGCGGAAGTGAGACCCCAGGTACATCCGAAT
	CAGCAACGCCTGAAAGCACCGGTCGGACACTGGTGACCTTCAAGGATGTATTTGTGGACTTC

	ACCAGGGAGGAGTGGAAGCTGCTGGACACTGCTCAGCAGATCGTGTACAGAAATGTGATGCT GGAGAACTATAAGAACCTGGTTTCCTTGGGTTATCAGCTTACTAAGCCAGATGTGATCCTCC GGTTGGAGAAGGGAGAAGAGCCC
Fusion Protein 4 Amino Acid Sequence 3A-ADD-h3L-NLS-dCas9-NLS-KRAB (SEQ ID NO: 1504)	MNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITV GMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEF YRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRARYFWG NLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDI LWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNAN SRGPSFSSGLVPLSLRGSHMEVKANQRNIEDICICCGSLQVHTQHPLFEGGICAPCKDKFLD ALFLYDDDGYQSYCSICCSGETLLICGNPDCTRCYCFECVDSLVGPGTSGKVHAMSNWVCYL CLPSSRSGLLQRRRKWRSQLKAFYDRESENPLEMFETVPVWRRQPVRVLSLFEDIKKELTSL GFLESGSDPGQLKHVVDVTDTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLLQ YARPKPGSPRPFFWMFVDNLVLNKEDLDVASRFLEMEPVTIPDVHGGSLQNAVRVWSNIPAI RSRHWALVSEEELSLLAQNKQSSKLAAKWPTKLVKNCFLPLREYFKYFSTELTSSLGGPSSG APPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSE GSAPGTSTEPSEPKKKRKVMDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESF LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGH FLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQL PGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFL AAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFD QSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIH LGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWN FEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAF LSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKII KDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLS RKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIA NLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEG IKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKD DSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFY KVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKK TEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKS VKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILA DANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDA TLIHQSITGLYETRIDLSQLGGDPKKKRKVSGSETPGTSESATPESTGRTLVTFKDVFVDFT REEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEP
Fusion Protein 4 DNA Sequence (SEQ ID NO: 1505)	ATGAACCATGATCAAGAATTCGACCCACCTAAAGTCTACCCACCTGTGCCCGCCGAAAAAAG GAAACCCATAAGGGTGCTGTCACTCTTTGATGGCATCGCCACTGGTCTCCTGGTTCTTAAGG ATCTGGGAATTCAGGTCGATCGGTACATTGCTAGCGAGGTTTGTGAGGATAGTATTACAGTG GGTATGGTGCGCCACCAGGGAAAGATCATGTATGTTGGTGACGTTAGGAGCGTCACCCAGAA ACATATCCAGGAGTGGGGACCCTTTGATTTGGTGATCGGAGGTAGTCCCTGCAATGACCTTT CCATCGTGAATCCAGCCAGGAAAGGGCTGTATGAAGGGACTGGTAGGCTCTTTTTCGAGTTT TATCGCCTGCTTCACGACGCTAGACCTAAGGAAGGTGACGATAGGCCTTTCTTTTGGCTTTT TGAGAACGTCGTGGCAATGGGAGTCTCCGACAAAAGGGACATTTCTCGCTTTCTGGAATCTA ACCCCGTTATGATCGATGCCAAGGAAGTTTCTGCCGCTCACAGGGCAAGGTACTTCTGGGGC AATCTGCCCGGAATGAATCGCCCACTGGCCAGTACCGTGAATGACAAACTGGAGCTGCAGGA GTGCCTGGAGCACGGAAGAATCGCAAAGTTTTCTAAAGTCAGGACCATTACCACTCGCAGTA ACTCCATAAAACAGGGTAAGGACCAGCATTTTCCCGTCTTCATGAATGAAAAGGAAGATATT CTGTGGTGCACTGAAATGGAGAGAGTTTTCGGGTTTCCCGTGCACTATACCGATGTTTCCAA CATGTCCCGCCTTGCAAGACAAAGGCTTTTGGGCCGCTCTTGGTCTGTGCCAGTGATCCGGC ACTTGTTTGCTCCCCTCAAAGAGTACTTCGCTTGCGTCAGTTCCGGAAATTCAAACGCTAAC TCTCGGGGTCCATCTTTCTCCAGTGGTCTCGTGCCACTGTCTCTCCGGGGCTCTCACATGGA AGTCAAGGCTAACCAGCGAAATATAGAAGACATCTGCATCTGCTGCGGAAGTCTCCAGGTTC ACACACAGCACCCTCTGTTTGAGGGAGGGATCTGCGCCCCATGTAAGGACAAGTTCCTGGAT GCCCTCTTCCTGTACGACGATGACGGGTACCAATCCTACTGCTCCATCTGCTGCTCCGGAGA GACGCTGCTCATCTGCGGAAACCCTGATTGCACCCGATGCTACTGCTTCGAGTGTGTGGATA GCCTGGTCGGCCCCGGGACCTCGGGGAAGGTGCACGCCATGAGCAACTGGGTGTGCTACCTG TGCCTGCCGTCCTCCCGAAGCGGGCTGCTGCAGCGTCGGAGGAAGTGGCGCAGCCAGCTCAA GGCCTTCTACGACCGAGAGTCGGAGAATCCCCTGGAGATGTTTGAGACAGTGCCAGTCTGGC

	GGAGGCAGCCCGTTCGCGTTCTCTCTCTGTTCGAAGATATTAAAAAGGAACTCACCTCCCTT
	GGGTTCCTGGAGAGCGGGAGCGACCCCGGACAGCTTAAGCACGTGGTCGACGTGACTGACAC
	CGTCCGCAAAGACGTGGAGGAATGGGGCCCCTTCGATCTGGTCTATGGGGCAACCCCTCCCC
	TTGGGCATACATGTGATCGGCCTCCATCCTGGTACCTGTTCCAGTTTCACAGACTCCTGCAG
	TATGCCAGGCCAAAGCCAGGGAGCCCAAGGCCCTTTTTCTGGATGTTCGTCGACAACCTGGT
	CCTGAACAAAGAAGATCTCGACGTTGCTAGTCGCTTTCTCGAAATGGAGCCCGTGACCATTC
	CCGACGTGCATGGCGGTTCCCTCCAGAATGCAGTCAGGGTTTGGAGCAATATCCCTGCCATC
	AGGTCAAGGCACTGGGCACTGGTTTCAGAGGAAGAGCTGTCCCTCCTTGCCCAGAACAAGCA
	GTCATCCAAACTGGCAGCCAAGTGGCCAACTAAGCTGGTCAAGAACTGCTTTCTTCCCCTCA
	GAGAATATTTTAAGTATTTCAGTACTGAACTGACTAGCAGTCTGGGAGGGCCGAGCTCTGGC
	GCACCCCCACCAAGTGGAGGGTCTCCTGCCGGGTCCCCAACATCTACTGAAGAAGGCACCAG
	CGAATCCGCAACGCCCGAGTCAGGCCCTGGTACCTCCACAGAACCATCTGAAGGTAGTGCGC
	CTGGTTCCCCAGCTGGAAGCCCTACTTCCACCGAAGAAGGCACGTCAACCGAACCAAGTGAA
	GGATCTGCCCCTGGGACCAGCACTGAACCATCTGAGCCAAAAAAGAAGAGAAAGGTAATGGA
	CAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCG
	ACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATC
	AAGAAGAACCTGATCGGCGCCCTGCTGTTCGACAGCGGAGAAACAGCCGAGGCCACCCGGCT
	GAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGA
	TCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTC
	CTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGT
	GGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCACCG
	ACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCCAC
	TTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCAGCT
	GGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACGCCA
	AGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAGCTG
	CCCGGCGAGAAGAAGAATGGCCTGTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGACCCC
	CAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCT
	ACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTG
	GCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGAT
	CACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGA
	CCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGAC
	CAGAGCAAGAACGGCTACGCCGGCTACATCGATGGCGGAGCCAGCCAGGAAGAGTTCTACAA
	GTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACA
	GAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCAC
	CTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAA
	CCGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCA
	GGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAAC
	TTCGAGGAAGTGGTGGACAAGGGCGCCAGCGCCCAGAGCTTCATCGAGCGGATGACCAACTT
	CGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCA
	CCGTGTACAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTC
	CTGAGCGGCGAGCAGAAAAAAGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGAC
	CGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCT
	CCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATC
	AAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGAC
	CCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCTGT
	TCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGC
	CGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAA
	GTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTA
	AAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCC
	AATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGA
	GCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGA
	ACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGC
	ATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAA
	CGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGG
	ACATCAACCGGCTGTCCGACTACGATGTGGACGCTATCGTGCCTCAGAGCTTTCTGAAGGAC
	GACTCCATCGATAACAAAGTGCTGACTCGGAGCGACAAGAACCGGGGCAAGAGCGACAACGT
	GCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGCCAGCTGCTGAATGCCAAGC
	TGATTACCCAGAGGAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTG
	GATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGC
	ACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAACGACAAACTGATCCGGGAAG
	TGAAAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTAC
	AAAGTGCGCGAGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGG

	AACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGG TGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAG TACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGA GATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACAGGCGAGATCGTGTGGGATAAGG GCCGGGACTTTGCCACCGTGCGGAAAGTGCTGTCTATGCCCCAAGTGAATATCGTGAAAAAG ACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGACAA GCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCG TGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGT GTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGA CTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGT ACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAG AAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTA TGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACA AACACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCC GACGCTAATCTGGACAAGGTGCTGAGCGCCTACAACAAGCACAGAGACAAGCCTATCAGAGA GCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCA AGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCC ACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGG AGGCGACCCAAAAAAGAAGAGAAAGGTAAGCGGAAGTGAGACCCCAGGTACATCCGAATCAG CAACGCCTGAAAGCACCGGTCGGACACTGGTGACCTTCAAGGATGTATTTGTGGACTTCACC AGGGAGGAGTGGAAGCTGCTGGACACTGCTCAGCAGATCGTGTACAGAAATGTGATGCTGGA GAACTATAAGAACCTGGTTTCCTTGGGTTATCAGCTTACTAAGCCAGATGTGATCCTCCGGT TGGAGAAGGGAGAAGAGCCC
Fusion Protein 5 Amino Acid Sequence NLS-3A-3L-dCas9-KRAB-NLS (SEQ ID NO: 1506)	MPKKKRKVPKKKRKVNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVD RYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPAR KGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDA KEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGK DQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLK EYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHNPLEMFETVPVWRRQPVRVLSLFEDIKKE LTSLGFLESGSDPGQLKHVVDVTDTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFH RLLQYARPKPGSPRPFFWMFVDNLVLNKEDLDVASRFLEMEPVTIPDVHGGSLQNAVRVWSN IPAIRSRHWALVSEEELSLLAQNKQSSKLAAKWPTKLVKNCFLPLREYFKYFSTELTSSLGG PSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTST EPSEGSAPGTSTEPSEMDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKN LIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVE EDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI EGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAK NLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSK NGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGE LHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEE VVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSG EQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDK DFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKL INGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKE LGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSI DNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKA GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVR EINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFF YSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEV QTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKE LLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGN ELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADAN LDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLI HQSITGLYETRIDLSQLGGDSGSETPGTSESATPESTGRTLVTFKDVFVDFTREEWKLLDTA QQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPSADYKDDDDKAPKKKRKVPKKKRK V
Fusion Protein 5 DNA Sequence	ATGCCAAAAAAGAAGAGAAAGGTACCGAAGAAAAAAAGAAAGGTCAACCACGACCAGGAATT CGACCCTCCAAAGGTTTACCCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGT CTCTCTTTGATGGAATCGCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATTCAGGTGGAC CGCTACATTGCCTCGGAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCAGGG

(SEQ ID NO: 1507)	GAAGATCATGTACGTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCC CATTCGATCTGGTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCGC AAGGGCCTCTACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGC GCGGCCCAAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCATGG GCGTTAGTGACAAGAGGGACATCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCC AAAGAAGTGTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAG GCCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGGA TAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACTCCATAAAGCAGGGCAAA GACCAGCATTTTCCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAAATGGA AAGGGTATTTGGTTTCCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTGGCGAGGC AGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCTCCGCTGAAG GAGTATTTTGCGTGTGTGTCTAGCGGCAATAGTAACGCTAACAGCCGCGGGCCGAGCTTCAG CAGCGGCCTGGTGCCGTTAAGCTTGCGCGGCAGCCATAATCCCCTTGAGATGTTCGAAACCG TGCCTGTGTGGAGGAGACAGCCAGTCCGGGTGCTGTCCCTTTTTGAAGACATCAAGAAAGAG CTGACGAGTTTGGGCTTTTTGGAAAGTGGTTCTGACCCGGGACAACTGAAGCATGTGGTTGA TGTCACAGACACAGTGAGGAAGGATGTGGAGGAGTGGGGACCCTTCGATCTTGTGTACGGCG CCACACCTCCCCTGGGCCACACCTGTGACCGTCCTCCCAGCTGGTACCTGTTCCAGTTCCAC CGGCTCCTGCAGTACGCACGGCCCAAGCCAGGCAGCCCCAGGCCCTTCTTCTGGATGTTCGT GGACAATCTGGTGCTGAACAAGGAAGACCTGGACGTCGCATCTCGCTTCCTGGAGATGGAGC CAGTCACCATCCCAGATGTCCACGGCGGATCCTTGCAGAATGCTGTCCGCGTGTGGAGCAAC ATCCCAGCCATAAGGAGCAGGCACTGGGCTCTGGTTTCGGAAGAAGAATTGTCCCTGCTGGC CCAGAACAAGCAGAGCTCGAAGCTCGCGGCCAAGTGGCCCACCAAGCTGGTGAAGAACTGCT TTCTCCCCCTAAGAGAATATTTCAAGTATTTTTCAACAGAACTCACTTCCTCTTTAGGAGGG CCGAGCTCTGGCGCACCCCCACCAAGTGGAGGGTCTCCTGCCGGGTCCCCAACATCTACTGA AGAAGGCACCAGCGAATCCGCAACGCCCGAGTCAGGCCCTGGTACCTCCACAGAACCATCTG AAGGTAGTGCGCCTGGTTCCCCAGCTGGAAGCCCTACTTCCACCGAAGAAGGCACGTCAACC GAACCAAGTGAAGGATCTGCCCCTGGGACCAGCACTGAACCATCTGAGATGGACAAGAAGTA CAGCATCGGCCTGGCCATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACA AGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAAC CTGATCGGCGCCCTGCTGTTCGACAGCGGAGAAACAGCCGAGGCCACCCGGCTGAAGAGAAC CGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCA ACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAA GAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCA CGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCACCGACAAGGCCG ACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATC GAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAGAC CTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCC TGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAG AAGAAGAATGGCCTGTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAA GAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACG ACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAG AACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGC CCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGA AAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAG AACGGCTACGCCGGCTACATCGATGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAA GCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACC TGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAG CTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAA GATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACA GCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAA GTGGTGGACAAGGGCGCCAGCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAA CCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTACA ACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGC GAGCAGAAAAAAGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCA GCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGG AAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGGACAAG GACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCCTGACACT GTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCTGTTCGACGACA AAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTG ATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGG CTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACA TCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCC

	GGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAA AGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACCAGACCA CCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAG CTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCT GTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGGACATCAACC GGCTGTCCGACTACGATGTGGACGCTATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATC GATAACAAAGTGCTGACTCGGAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGA AGAGGTCGTGAAGAAGATGAAGAACTACTGGCGCCAGCTGCTGAATGCCAAGCTGATTACCC AGAGGAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCC GGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCT GGACTCCCGGATGAACACTAAGTACGACGAGAACGACAAACTGATCCGGGAAGTGAAAGTGA TCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGC GAGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCT GATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACG TGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTC TACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAA GCGGCCTCTGATCGAGACAAACGGCGAAACAGGCGAGATCGTGTGGGATAAGGGCCGGGACT TTGCCACCGTGCGGAAAGTGCTGTCTATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTG CAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGACAAGCTGATCGC CAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATT CTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAAGAG CTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGA AGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGT TCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAAC GAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCT GAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAACACTACC TGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAAT CTGGACAAGGTGCTGAGCGCCTACAACAAGCACAGAGACAAGCCTATCAGAGAGCAGGCCGA GAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTG ACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGATC CACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACAG CGGAAGTGAGACCCCAGGTACATCCGAATCAGCAACGCCTGAAAGCACCGGTCGGACACTGG TGACCTTCAAGGATGTATTTGTGGACTTCACCAGGGAGGAGTGGAAGCTGCTGGACACTGCT CAGCAGATCGTGTACAGAAATGTGATGCTGGAGAACTATAAGAACCTGGTTTCCTTGGGTTA TCAGCTTACTAAGCCAGATGTGATCCTCCGGTTGGAGAAGGGAGAAGAGCCCAGCGCTGATT ACAAAGATGATGACGATAAAGCCCCAAAAAAGAAGAGAAAGGTACCGAAGAAAAAAAGAAAG GTC
Fusion Protein 6 Amino Acid Sequence NLS-3A-3L-dCas9-KRAB-NLS (SEQ ID NO: 1508)	MPKKKRKVPKKKRKVNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVD RYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPAR KGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDA KEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGK DQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLK EYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRQPVRVLSLFRNIDK VLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMF QFHRILQYALPRQESQRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRV WSNIPGLKSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGT STEPSEGSAPGTSTEPSEMDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIK KNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHF LIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLA AKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQ SKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL GELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFL SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIK DKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIAN LAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGI KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDD SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELD

	KAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKY FFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKT EVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSV KELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQK GNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILAD ANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDAT LIHQSITGLYETRIDLSQLGGDSGSETPGTSESATPESTGRTLVTFKDVFVDFTREEWKLLD TAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPSADYKDDDDKAPKKKRKVPKKK RKV
Fusion Protein 6 DNA Sequence (SEQ ID NO: 1509)	ATGCCAAAAAAGAAGAGAAAGGTACCGAAGAAAAAAAGAAAGGTATACAACCATGACCAGGA ATTCGACCCCCCAAAGGTTTACCCACCTGTGCCAGCTGAGAAGAGGAAGCCCATCCGCGTGC TGTCTCTCTTTGATGGGATTGCTACAGGGCTCCTGGTGCTGAAGGACCTGGGCATCCAAGTG GACCGCTACATTGCCTCCGAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCA GGGAAAGATCATGTACGTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGG GCCCATTCGACCTGGTGATTGGAGGCAGTCCCTGCAATGACCTCTCCATTGTCAACCCTGCC CGCAAGGGACTTTATGAGGGTACTGGCCGCCTCTTCTTTGAGTTCTACCGCCTCCTGCATGA TGCGCGGCCCAAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCA TGGGCGTTAGTGACAAGAGGGACATCTCGCGATTTCTTGAGTCTAACCCCGTGATGATTGAC GCCAAAGAAGTGTCTGCTGCACACAGGGCCCGTTACTTCTGGGGTAACCTTCCTGGCATGAA CAGGCCTTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAAGAGTGTCTGGAGCACGGCA GAATAGCCAAGTTCAGCAAAGTGAGGACCATTACCACCAGGTCAAACTCTATAAAGCAGGGC AAAGACCAGCATTTCCCCGTCTTCATGAACGAGAAGGAGGACATCCTGTGGTGCACTGAAAT GGAAAGGGTGTTTGGCTTCCCCGTCCACTACACAGACGTCTCCAACATGAGCCGCTTGGCGA GGCAGAGACTGCTGGGCCGATCGTGGAGCGTGCCGGTCATCCGCCACCTCTTCGCTCCGCTG AAGGAATATTTTGCTTGTGTGTCTAGCGGCAATAGTAACGCTAACAGCCGCGGGCCGAGCTT CAGCAGCGGCCTGGTGCCGTTAAGCTTGCGCGGCAGCCATATGGGCCCTATGGAGATATACA AGACAGTGTCTGCATGGAAGAGACAGCCAGTGCGGGTACTGAGCCTCTTCAGAAACATCGAC AAGGTACTAAAGAGTTTGGGCTTCTTGGAAAGCGGTTCTGGTTCTGGGGGAGGAACGCTGAA GTACGTGGAAGATGTCACAAATGTCGTGAGGAGAGACGTGGAGAAATGGGGCCCCTTTGACC TGGTGTACGGCTCGACGCAGCCCCTAGGCAGCTCTTGTGATCGCTGTCCCGGCTGGTACATG TTCCAGTTCCACCGGATCCTGCAGTATGCGCTGCCTCGCCAGGAGAGTCAGCGGCCCTTCTT CTGGATATTCATGGACAATCTGCTGCTGACTGAGGATGACCAAGAGACAACTACCCGCTTCC TTCAGACAGAGGCTGTGACCCTCCAGGATGTCCGTGGCAGAGACTACCAGAATGCTATGCGG GTGTGGAGCAACATTCCAGGGCTGAAGAGCAAGCATGCGCCCCTGACCCCAAAGGAAGAAGA GTATCTGCAAGCCCAAGTCAGAAGCAGGAGCAAGCTGGACGCCCCGAAAGTTGACCTCCTGG TGAAGAACTGCCTTCTCCCGCTGAGAGAGTACTTCAAGTATTTTTCTCAAAACTCACTTCCT CTTGGAGGGCCGAGCTCTGGCGCACCCCCACCAAGTGGAGGGTCTCCTGCCGGGTCCCCAAC ATCTACTGAAGAAGGCACCAGCGAATCCGCAACGCCCGAGTCAGGCCCTGGTACCTCCACAG AACCATCTGAAGGTAGTGCGCCTGGTTCCCCAGCTGGAAGCCCTACTTCCACCGAAGAAGGC ACGTCAACCGAACCAAGTGAAGGATCTGCCCCTGGGACCAGCACTGAACCATCTGAGATGGA CAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCG ACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATC AAGAAGAACCTGATCGGCGCCCTGCTGTTCGACAGCGGAGAAACAGCCGAGGCCACCCGGCT GAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGA TCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTC CTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGT GGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCACCG ACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCCAC TTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCAGCT GGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACGCCA AGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAGCTG CCCGGCGAGAAGAAGAATGGCCTGTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGACCCC CAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCT ACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTG GCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGAT CACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGA CCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGAC CAGAGCAAGAACGGCTACGCCGGCTACATCGATGGCGGAGCCAGCCAGGAAGAGTTCTACAA GTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACA GAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCAC CTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAA

	CCGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCA GGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAAC TTCGAGGAAGTGGTGGACAAGGGCGCCAGCGCCCAGAGCTTCATCGAGCGGATGACCAACTT CGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCA CCGTGTACAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTC CTGAGCGGCGAGCAGAAAAAAGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGAC CGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCT CCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATC AAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGAC CCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCTGT TCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGC CGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAA GTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTA AAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCC AATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGA GCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGA ACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGC ATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAA CGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGG ACATCAACCGGCTGTCCGACTACGATGTGGACGCTATCGTGCCTCAGAGCTTTCTGAAGGAC GACTCCATCGATAACAAAGTGCTGACTCGGAGCGACAAGAACCGGGGCAAGAGCGACAACGT GCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGCCAGCTGCTGAATGCCAAGC TGATTACCCAGAGGAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTG GATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGC ACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAACGACAAACTGATCCGGGAAG TGAAAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTAC AAAGTGCGCGAGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGG AACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGG TGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAG TACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGA GATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACAGGCGAGATCGTGTGGGATAAGG GCCGGGACTTTGCCACCGTGCGGAAAGTGCTGTCTATGCCCCAAGTGAATATCGTGAAAAAG ACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGACAA GCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCG TGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGT GTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGA CTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGT ACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAG AAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTA TGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACA AACACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCC GACGCTAATCTGGACAAGGTGCTGAGCGCCTACAACAAGCACAGAGACAAGCCTATCAGAGA GCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCA AGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCC ACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGG AGGCGACAGCGGAAGTGAGACCCCAGGTACATCCGAATCAGCAACGCCTGAAAGCACCGGTC GGACACTGGTGACCTTCAAGGATGTATTTGTGGACTTCACCAGGGAGGAGTGGAAGCTGCTG GACACTGCTCAGCAGATCGTGTACAGAAATGTGATGCTGGAGAACTATAAGAACCTGGTTTC CTTGGGTTATCAGCTTACTAAGCCAGATGTGATCCTCCGGTTGGAGAAGGGAGAAGAGCCCA GCGCTGATTACAAAGATGATGACGATAAAGCCCCAAAAAAGAAGAGAAAGGTACCGAAGAAA AAAAGAAAGGTC
Fusion Protein 7 Amino Acid Sequence NLS-3A-3L-dCas9-ZIM-NLS (SEQ ID NO: 1510)	MPKKKRKVPKKKRKVNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVD RYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPAR KGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDA KEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGK DQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLK EYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRQPVRVLSLFRNIDK VLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMF QFHRILQYALPRQESQRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRV WSNIPGLKSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGT STEPSEGSAPGTSTEPSEMDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIK

	KNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHF LIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLA AKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQ SKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL GELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFL SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIK DKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIAN LAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGI KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDD SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELD KAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKY FFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKT EVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSV KELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQK GNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILAD ANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDAT LIHQSITGLYETRIDLSQLGGDSGSETPGTSESATPESTGMNNSQGRVTFEDVTVNFTQGEW QRLNPEQRNLYRDVMLENYSNLVSVGQGETTKPDVILRLEQGKEPWLEEEEVLGSGRAEKNG DIGGQIWKPKDVKESLSADYKDDDDKAPKKKRKVPKKKRKV
Fusion Protein 7 DNA Sequence (SEQ ID NO: 1511)	ATGCCAAAAAAGAAGAGAAAGGTACCGAAGAAAAAAAGAAAGGTATACAACCATGACCAGGA ATTCGACCCCCCAAAGGTTTACCCACCTGTGCCAGCTGAGAAGAGGAAGCCCATCCGCGTGC TGTCTCTCTTTGATGGGATTGCTACAGGGCTCCTGGTGCTGAAGGACCTGGGCATCCAAGTG GACCGCTACATTGCCTCCGAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCA GGGAAAGATCATGTACGTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGG GCCCATTCGACCTGGTGATTGGAGGCAGTCCCTGCAATGACCTCTCCATTGTCAACCCTGCC CGCAAGGGACTTTATGAGGGTACTGGCCGCCTCTTCTTTGAGTTCTACCGCCTCCTGCATGA TGCGCGGCCCAAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCA TGGGCGTTAGTGACAAGAGGGACATCTCGCGATTTCTTGAGTCTAACCCCGTGATGATTGAC GCCAAAGAAGTGTCTGCTGCACACAGGGCCCGTTACTTCTGGGGTAACCTTCCTGGCATGAA CAGGCCTTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAAGAGTGTCTGGAGCACGGCA GAATAGCCAAGTTCAGCAAAGTGAGGACCATTACCACCAGGTCAAACTCTATAAAGCAGGGC AAAGACCAGCATTTCCCCGTCTTCATGAACGAGAAGGAGGACATCCTGTGGTGCACTGAAAT GGAAAGGGTGTTTGGCTTCCCCGTCCACTACACAGACGTCTCCAACATGAGCCGCTTGGCGA GGCAGAGACTGCTGGGCCGATCGTGGAGCGTGCCGGTCATCCGCCACCTCTTCGCTCCGCTG AAGGAATATTTTGCTTGTGTGTCTAGCGGCAATAGTAACGCTAACAGCCGCGGGCCGAGCTT CAGCAGCGGCCTGGTGCCGTTAAGCTTGCGCGGCAGCCATATGGGCCCTATGGAGATATACA AGACAGTGTCTGCATGGAAGAGACAGCCAGTGCGGGTACTGAGCCTCTTCAGAAACATCGAC AAGGTACTAAAGAGTTTGGGCTTCTTGGAAAGCGGTTCTGGTTCTGGGGGAGGAACGCTGAA GTACGTGGAAGATGTCACAAATGTCGTGAGGAGAGACGTGGAGAAATGGGGCCCCTTTGACC TGGTGTACGGCTCGACGCAGCCCCTAGGCAGCTCTTGTGATCGCTGTCCCGGCTGGTACATG TTCCAGTTCCACCGGATCCTGCAGTATGCGCTGCCTCGCCAGGAGAGTCAGCGGCCCTTCTT CTGGATATTCATGGACAATCTGCTGCTGACTGAGGATGACCAAGAGACAACTACCCGCTTCC TTCAGACAGAGGCTGTGACCCTCCAGGATGTCCGTGGCAGAGACTACCAGAATGCTATGCGG GTGTGGAGCAACATTCCAGGGCTGAAGAGCAAGCATGCGCCCCTGACCCCAAAGGAAGAAGA GTATCTGCAAGCCCAAGTCAGAAGCAGGAGCAAGCTGGACGCCCCGAAAGTTGACCTCCTGG TGAAGAACTGCCTTCTCCCGCTGAGAGAGTACTTCAAGTATTTTTCTCAAAACTCACTTCCT CTTGGAGGGCCGAGCTCTGGCGCACCCCCACCAAGTGGAGGGTCTCCTGCCGGGTCCCCAAC ATCTACTGAAGAAGGCACCAGCGAATCCGCAACGCCCGAGTCAGGCCCTGGTACCTCCACAG AACCATCTGAAGGTAGTGCGCCTGGTTCCCCAGCTGGAAGCCCTACTTCCACCGAAGAAGGC ACGTCAACCGAACCAAGTGAAGGATCTGCCCCTGGGACCAGCACTGAACCATCTGAGATGGA CAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCG ACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATC AAGAAGAACCTGATCGGCGCCCTGCTGTTCGACAGCGGAGAAACAGCCGAGGCCACCCGGCT GAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGA TCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTC CTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGT GGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCACCG

	ACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCCAC
	TTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCAGCT
	GGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACGCCA
	AGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAGCTG
	CCCGGCGAGAAGAAGAATGGCCTGTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGACCCC
	CAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCT
	ACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTG
	GCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGAT
	CACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGA
	CCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGAC
	CAGAGCAAGAACGGCTACGCCGGCTACATCGATGGCGGAGCCAGCCAGGAAGAGTTCTACAA
	GTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACA
	GAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCAC
	CTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAA
	CCGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCA
	GGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAAC
	TTCGAGGAAGTGGTGGACAAGGGCGCCAGCGCCCAGAGCTTCATCGAGCGGATGACCAACTT
	CGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCA
	CCGTGTACAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTC
	CTGAGCGGCGAGCAGAAAAAAGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGAC
	CGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCT
	CCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATC
	AAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGAC
	CCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCTGT
	TCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGC
	CGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAA
	GTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTA
	AAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCC
	AATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGA
	GCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGA
	ACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGC
	ATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAA
	CGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGG
	ACATCAACCGGCTGTCCGACTACGATGTGGACGCTATCGTGCCTCAGAGCTTTCTGAAGGAC
	GACTCCATCGATAACAAAGTGCTGACTCGGAGCGACAAGAACCGGGGCAAGAGCGACAACGT
	GCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGCCAGCTGCTGAATGCCAAGC
	TGATTACCCAGAGGAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTG
	GATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGC
	ACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAACGACAAACTGATCCGGGAAG
	TGAAAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTAC
	AAAGTGCGCGAGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGG
	AACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGG
	TGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAG
	TACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGA
	GATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACAGGCGAGATCGTGTGGGATAAGG
	GCCGGGACTTTGCCACCGTGCGGAAAGTGCTGTCTATGCCCCAAGTGAATATCGTGAAAAAG
	ACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGACAA
	GCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCG
	TGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGT
	GTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGA
	CTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGT
	ACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAG
	AAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTA
	TGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACA
	AACACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCC
	GACGCTAATCTGGACAAGGTGCTGAGCGCCTACAACAAGCACAGAGACAAGCCTATCAGAGA
	GCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCA
	AGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCC
	ACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGG
	AGGCGACAGCGGAAGTGAGACCCCAGGTACATCCGAATCAGCAACGCCTGAAAGCACCGGTA
	TGAACAATTCACAGGGGAGAGTGACATTCGAAGACGTGACCGTGAACTTCACCCAGGGAGAA
	TGGCAGCGCTTGAACCCAGAACAAAGGAACCTCTATCGGGACGTGATGCTGGAAAACTACTC

	AAATTTGGTGAGCGTTGGGCAGGGTGAGACCACTAAGCCTGACGTGATCCTGAGATTGGAAC AGGGCAAGGAGCCTTGGCTCGAGGAAGAGGAAGTCCTGGGCTCAGGGAGGGCCGAGAAAAAC GGTGATATAGGAGGCCAGATATGGAAGCCTAAGGACGTCAAGGAGAGCCTGAGCGCTGATTA CAAAGATGATGACGATAAAGCCCCAAAAAAGAAGAGAAAGGTACCGAAGAAAAAAAGAAAGG TC
Fusion Protein 8 Amino Acid Sequence NLS-3A-3L-dCas9-ZFP-NLS (SEQ ID NO: 1512)	MPKKKRKVPKKKRKVNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVD RYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPAR KGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDA KEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGK DQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLK EYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRQPVRVLSLFRNIDK VLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMF QFHRILQYALPRQESQRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRV WSNIPGLKSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGT STEPSEGSAPGTSTEPSEMDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIK KNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHF LIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLA AKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQ SKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL GELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFL SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIK DKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIAN LAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGI KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDD SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELD KAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKY FFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKT EVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSV KELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQK GNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILAD ANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDAT LIHQSITGLYETRIDLSQLGGDSGSETPGTSESATPESTGNKKLEAVGTGIEPKAMSQGLVT FGDVAVDFSQEEWEWLNPIQRNLYRKVMLENYRNLASLGLCVSKPDVISSLEQGKEPWSADY KDDDDKAPKKKRKVPKKKRKV
Fusion Protein 8 DNA Sequence (SEQ ID NO: 1513)	ATGCCAAAAAAGAAGAGAAAGGTACCGAAGAAAAAAAGAAAGGTATACAACCATGACCAGGA ATTCGACCCCCCAAAGGTTTACCCACCTGTGCCAGCTGAGAAGAGGAAGCCCATCCGCGTGC TGTCTCTCTTTGATGGGATTGCTACAGGGCTCCTGGTGCTGAAGGACCTGGGCATCCAAGTG GACCGCTACATTGCCTCCGAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGGCACCA GGGAAAGATCATGTACGTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGG GCCCATTCGACCTGGTGATTGGAGGCAGTCCCTGCAATGACCTCTCCATTGTCAACCCTGCC CGCAAGGGACTTTATGAGGGTACTGGCCGCCTCTTCTTTGAGTTCTACCGCCTCCTGCATGA TGCGCGGCCCAAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATGTGGTGGCCA TGGGCGTTAGTGACAAGAGGGACATCTCGCGATTTCTTGAGTCTAACCCCGTGATGATTGAC GCCAAAGAAGTGTCTGCTGCACACAGGGCCCGTTACTTCTGGGGTAACCTTCCTGGCATGAA CAGGCCTTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAAGAGTGTCTGGAGCACGGCA GAATAGCCAAGTTCAGCAAAGTGAGGACCATTACCACCAGGTCAAACTCTATAAAGCAGGGC AAAGACCAGCATTTCCCCGTCTTCATGAACGAGAAGGAGGACATCCTGTGGTGCACTGAAAT GGAAAGGGTGTTTGGCTTCCCCGTCCACTACACAGACGTCTCCAACATGAGCCGCTTGGCGA GGCAGAGACTGCTGGGCCGATCGTGGAGCGTGCCGGTCATCCGCCACCTCTTCGCTCCGCTG AAGGAATATTTTGCTTGTGTGTCTAGCGGCAATAGTAACGCTAACAGCCGCGGGCCGAGCTT CAGCAGCGGCCTGGTGCCGTTAAGCTTGCGCGGCAGCCATATGGGCCCTATGGAGATATACA AGACAGTGTCTGCATGGAAGAGACAGCCAGTGCGGGTACTGAGCCTCTTCAGAAACATCGAC AAGGTACTAAAGAGTTTGGGCTTCTTGGAAAGCGGTTCTGGTTCTGGGGGAGGAACGCTGAA GTACGTGGAAGATGTCACAAATGTCGTGAGGAGAGACGTGGAGAAATGGGGCCCCTTTGACC TGGTGTACGGCTCGACGCAGCCCCTAGGCAGCTCTTGTGATCGCTGTCCCGGCTGGTACATG TTCCAGTTCCACCGGATCCTGCAGTATGCGCTGCCTCGCCAGGAGAGTCAGCGGCCCTTCTT

	CTGGATATTCATGGACAATCTGCTGCTGACTGAGGATGACCAAGAGACAACTACCCGCTTCC
	TTCAGACAGAGGCTGTGACCCTCCAGGATGTCCGTGGCAGAGACTACCAGAATGCTATGCGG
	GTGTGGAGCAACATTCCAGGGCTGAAGAGCAAGCATGCGCCCCTGACCCCAAAGGAAGAAGA
	GTATCTGCAAGCCCAAGTCAGAAGCAGGAGCAAGCTGGACGCCCCGAAAGTTGACCTCCTGG
	TGAAGAACTGCCTTCTCCCGCTGAGAGAGTACTTCAAGTATTTTTCTCAAAACTCACTTCCT
	CTTGGAGGGCCGAGCTCTGGCGCACCCCCACCAAGTGGAGGGTCTCCTGCCGGGTCCCCAAC
	ATCTACTGAAGAAGGCACCAGCGAATCCGCAACGCCCGAGTCAGGCCCTGGTACCTCCACAG
	AACCATCTGAAGGTAGTGCGCCTGGTTCCCCAGCTGGAAGCCCTACTTCCACCGAAGAAGGC
	ACGTCAACCGAACCAAGTGAAGGATCTGCCCCTGGGACCAGCACTGAACCATCTGAGATGGA
	CAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCG
	ACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATC
	AAGAAGAACCTGATCGGCGCCCTGCTGTTCGACAGCGGAGAAACAGCCGAGGCCACCCGGCT
	GAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGA
	TCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTC
	CTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGT
	GGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCACCG
	ACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCCAC
	TTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCAGCT
	GGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACGCCA
	AGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAGCTG
	CCCGGCGAGAAGAAGAATGGCCTGTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGACCCC
	CAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCT
	ACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTG
	GCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGAT
	CACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGA
	CCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGAC
	CAGAGCAAGAACGGCTACGCCGGCTACATCGATGGCGGAGCCAGCCAGGAAGAGTTCTACAA
	GTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACA
	GAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCAC
	CTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAA
	CCGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCA
	GGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAAC
	TTCGAGGAAGTGGTGGACAAGGGCGCCAGCGCCCAGAGCTTCATCGAGCGGATGACCAACTT
	CGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCA
	CCGTGTACAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTC
	CTGAGCGGCGAGCAGAAAAAAGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGAC
	CGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCT
	CCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATC
	AAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGAC
	CCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCTGT
	TCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGC
	CGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAA
	GTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTA
	AAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCC
	AATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGA
	GCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGA
	ACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGC
	ATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAA
	CGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGG
	ACATCAACCGGCTGTCCGACTACGATGTGGACGCTATCGTGCCTCAGAGCTTTCTGAAGGAC
	GACTCCATCGATAACAAAGTGCTGACTCGGAGCGACAAGAACCGGGGCAAGAGCGACAACGT
	GCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGCCAGCTGCTGAATGCCAAGC
	TGATTACCCAGAGGAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTG
	GATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGC
	ACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAACGACAAACTGATCCGGGAAG
	TGAAAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTAC
	AAAGTGCGCGAGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGG
	AACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGG
	TGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAG
	TACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGA
	GATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACAGGCGAGATCGTGTGGGATAAGG
	GCCGGGACTTTGCCACCGTGCGGAAAGTGCTGTCTATGCCCCAAGTGAATATCGTGAAAAAG

	ACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGACAA GCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCG TGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGT GTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGA CTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGT ACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAG AAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTA TGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACA AACACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCC GACGCTAATCTGGACAAGGTGCTGAGCGCCTACAACAAGCACAGAGACAAGCCTATCAGAGA GCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCA AGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCC ACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGG AGGCGACAGCGGAAGTGAGACCCCAGGTACATCCGAATCAGCAACGCCTGAAAGCACCGGTA ACAAAAAGCTTGAGGCCGTCGGAACCGGAATCGAACCAAAAGCAATGTCCCAGGGTTTGGTG ACATTTGGCGACGTGGCTGTCGATTTTTCCCAGGAAGAGTGGGAGTGGCTCAATCCTATCCA GAGGAACTTGTACCGGAAGGTGATGCTGGAGAATTATAGAAATTTGGCATCACTGGGGTTGT GCGTTAGCAAACCAGATGTTATATCTTCCCTGGAACAGGGAAAGGAGCCCTGGAGCGCTGAT TACAAAGATGATGACGATAAAGCCCCAAAAAAGAAGAGAAAGGTACCGAAGAAAAAAAGAAA GGTC
Fusion Protein 9 Variant 1 Amino Acid Sequence NLS-NLS-3A-h3L-dCas9-KOX1KRAB-NLS-NLS (SEQ ID NO: 1514)	MGTMPKKKRKVPKKKRKVYNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLG IQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIV NPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPV MIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSI KQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLF APLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMAAIPALDPEAEPSMDVILVGSSEL SSSVSPGTGRDLIAYEVKANQRNIEDICICCGSLQVHTQHPLFEGGICAPCKDKFLDALFLY DDDGYQSYCSICCSGETLLICGNPDCTRCYCFECVDSLVGPGTSGKVHAMSNWVCYLCLPSS RSGLLQRRRKWRSQLKAFYDRESENPLEMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLES GSDPGQLKHVVDVTDTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLLQYARPK PGSPRPFFWMFVDNLVLNKEDLDVASRFLEMEPVTIPDVHGGSLQNAVRVWSNIPAIRSRHW ALVSEEELSLLAQNKQSSKLAAKWPTKLVKNCFLPLREYFKYFSTELTSSLGGPSSGAPPPS GGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPG TSTEPSELEDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFD SGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGN LIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILL SDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYID GGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASA QSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKG ILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKE HPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRS DKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLV ETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHA HDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFF KTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAY NKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLY ETRIDLSQLGGDSPKKKRKVGVDGSSGSETPGTSESATPESRTLVTFKDVFVDFTREEWKLL DTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPSADYKDDDDKAPKKKRKVPKK KRKV
Fusion Protein 9 Variant 1 DNA Sequence	ATGGGTACCATGCCAAAAAAGAAGAGAAAGGTACCGAAGAAAAAAAGAAAGGTATACAATCA CGATCAGGAGTTCGACCCCCCTAAGGTGTACCCACCAGTGCCTGCAGAGAAGAGGAAGCCAA TCCGGGTGCTGAGCCTGTTTGATGGCATCGCCACCGGCCTGCTGGTGCTGAAGGATCTGGGC ATCCAGGTGGACCGGTACATCGCCTCCGAGGTGTGCGAGGATTCTATCACCGTGGGCATGGT

(SEQ ID NO: 1515)	GCGCCACCAGGGCAAGATCATGTATGTGGGCGACGTGCGGTCCGTGACACAGAAGCACATCC AGGAGTGGGGCCCATTCGATCTGGTGATCGGCGGCAGCCCCTGTAATGACCTGTCCATCGTG AACCCTGCAAGGAAGGGACTGTACGAGGGAACCGGCCGGCTGTTCTTTGAGTTTTATAGACT GCTGCACGACGCCAGGCCTAAGGAGGGCGACGATAGACCATTCTTTTGGCTGTTCGAGAATG TGGTGGCTATGGGCGTGAGCGATAAGAGGGACATCTCCAGGTTTCTGGAGTCTAACCCCGTG ATGATCGATGCAAAGGAGGTGTCCGCCGCACACAGAGCCAGGTATTTCTGGGGCAATCTGCC AGGAATGAACAGGCCACTGGCAAGCACCGTGAATGACAAGCTGGAGCTGCAGGAGTGCCTGG AGCACGGAAGGATCGCCAAGTTTTCCAAGGTGCGCACAATCACCACACGGAGCAATTCCATC AAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAACGAGAAGGAGGACATCCTGTGGTG TACCGAGATGGAGAGAGTGTTCGGCTTTCCAGTGCACTACACAGACGTGTCTAACATGAGCA GGCTGGCAAGGCAGCGGCTGCTGGGCAGATCTTGGAGCGTGCCCGTGATCAGGCACCTGTTC GCCCCTCTGAAGGAGTATTTTGCCTGCGTGAGCAGCGGCAACTCCAATGCCAACAGCCGGGG CCCCTCTTTCAGCTCCGGATTGGTGCCTCTGAGCCTGAGGGGCTCCCACATGGCAGCAATCC CCGCCCTGGACCCCGAGGCCGAGCCTAGCATGGACGTGATCCTGGTGGGCTCTAGCGAGCTG TCCTCTAGCGTGTCTCCAGGAACCGGAAGGGATCTGATCGCATACGAGGTGAAGGCCAATCA GCGGAACATCGAGGACATCTGTATCTGCTGTGGCAGCCTGCAGGTGCACACACAGCACCCAC TGTTCGAGGGAGGAATCTGCGCACCCTGTAAGGATAAGTTCCTGGACGCCCTGTTTCTGTAC GACGATGACGGCTACCAGTCCTATTGCTCTATCTGCTGTTCCGGCGAGACCCTGCTGATCTG CGGCAATCCAGATTGTACAAGGTGCTATTGTTTTGAGTGCGTGGACTCTCTGGTGGGACCAG GCACCAGCGGAAAGGTGCACGCCATGTCCAACTGGGTGTGCTACCTGTGCCTGCCATCCTCT CGCAGCGGACTGCTGCAGCGGAGAAGGAAGTGGAGATCCCAGCTGAAGGCCTTCTATGATAG GGAGTCTGAGAACCCCCTGGAGATGTTTGAGACCGTGCCAGTGTGGCGCCGGCAGCCCGTGA GGGTGCTGAGCCTGTTCGAGGATATCAAGAAGGAGCTGACATCCCTGGGCTTTCTGGAGTCC GGCTCTGACCCCGGACAGCTGAAGCACGTGGTGGATGTGACCGACACAGTGCGGAAGGATGT GGAGGAGTGGGGCCCTTTCGACCTGGTGTACGGAGCAACCCCTCCACTGGGACACACATGCG ACAGACCCCCTTCTTGGTACCTGTTCCAGTTTCACCGCCTGCTGCAGTATGCAAGGCCAAAG CCAGGCAGCCCTAGACCATTCTTTTGGATGTTCGTGGATAATCTGGTGCTGAACAAGGAGGA TCTGGACGTGGCCAGCAGGTTTCTGGAGATGGAGCCAGTGACCATCCCAGACGTGCACGGCG GCTCCCTGCAGAATGCCGTGCGCGTGTGGTCTAACATCCCTGCCATCAGAAGCAGGCACTGG GCACTGGTGAGCGAGGAGGAGCTGTCCCTGCTGGCCCAGAATAAGCAGAGCAGCAAGCTGGC CGCCAAGTGGCCTACAAAGCTGGTGAAGAACTGCTTCCTGCCACTGCGGGAGTACTTCAAGT ATTTTTCCACCGAGCTGACATCTAGCCTGGGAGGACCCTCCTCTGGCGCCCCACCACCTAGC GGCGGCTCCCCTGCCGGCTCTCCAACCAGCACAGAGGAGGGCACCAGCGAGTCCGCCACACC AGAGTCTGGACCTGGCACCAGCACAGAGCCATCCGAGGGCTCTGCCCCAGGCTCTCCTGCAG GCAGCCCTACCTCCACCGAAGAGGGCACCAGCACAGAGCCTTCTGAGGGCAGCGCCCCAGGC ACCTCTACAGAGCCAAGCGAGCTCGAGGACAAGAAGTACAGCATCGGCCTGGCCATCGGCAC CAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGG TGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGAC AGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACG GAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACA GCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCAC CCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCA CCTGAGAAAGAAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCC TGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAAC AGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAA CCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCA GACGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAAC CTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGA TGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGA TCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTG AGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAA GAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGC CTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGAC GGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGG CACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCG ACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAG GAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCG CATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAA AGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCC CAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCC CAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACG TGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAGGCCATCGTGGAC

	CTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAA AATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGG GCACATACCACGATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAAC GAGGACATTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGA GGAACGGCTGAAAACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGC GGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAG TCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCA GCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCC AGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGC ATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGA GAACATCGTGATCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCC GCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAA CACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGG GCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACG CCATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGC GACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAA CTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAGAAAGTTCGACAATCTGACCA AGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTG GAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTA CGACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGGTGT CCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCACCACGCC CACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGA AAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGAGCG AGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATGAACTTTTTC AAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGG CGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGA GCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAG TCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAA GAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGG AAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAA AGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAA AAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGA GAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATAT GTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGA GCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCA GCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTAC AACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCT GACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGT ACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTAC GAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACAGCCCCAAGAAGAAGAGAAAGGTGGG AGTCGACGGATCCAGCGGCTCCGAGACCCCAGGCACATCTGAGAGCGCCACCCCTGAGTCCC GGACCCTGGTGACCTTCAAGGATGTATTTGTGGACTTCACCAGGGAGGAGTGGAAGCTGCTG GACACTGCTCAGCAGATCGTGTACAGAAATGTGATGCTGGAGAACTATAAGAACCTGGTTTC CTTGGGTTATCAGCTTACTAAGCCAGATGTGATCCTCCGGTTGGAGAAGGGAGAAGAGCCCA GCGCTGATTACAAAGATGATGACGATAAAGCCCCAAAAAAGAAGAGAAAGGTACCGAAGAAA AAAAGAAAGGTCTGA
Fusion Protein 9 Variant 2 Amino Acid Sequence NLS-NLS-3A-h3L-dCas9-KOX1KRAB-NLS-NLS (SEQ ID NO: 1523)	MPKKKRKVPKKKRKVYNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQV DRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPA RKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMID AKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQG KDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPL KEYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMAAIPALDPEAEPSMDVILVGSSELSSS VSPGTGRDLIAYEVKANQRNIEDICICCGSLQVHTQHPLFEGGICAPCKDKFLDALFLYDDD GYQSYCSICCSGETLLICGNPDCTRCYCFECVDSLVGPGTSGKVHAMSNWVCYLCLPSSRSG LLQRRRKWRSQLKAFYDRESENPLEMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSD PGQLKHVVDVTDTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLLQYARPKPGS PRPFFWMFVDNLVLNKEDLDVASRFLEMEPVTIPDVHGGSLQNAVRVWSNIPAIRSRHWALV SEEELSLLAQNKQSSKLAAKWPTKLVKNCFLPLREYFKYFSTELTSSPGGPSSGAPPPSGGS PAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTST EPSELEDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDV

	DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIA LSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGA SQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDF YPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSF IERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDI LEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGK TILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQ TVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPV ENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKN RGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR QITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDA YLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTE ITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESIL PKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNF LYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETR IDLSQLGGDSPKKKRKVGVDGSSGSETPGTSESATPESRTLVTFKDVFVDFTREEWKLLDTA QQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPSADYKDDDDKAPKKKRKVPKKKRK V
Fusion Protein 9 Variant 2 DNA Sequence (SEQ ID NO: 1524)	ATGGGTACCATGCCAAAAAAGAAGAGAAAGGTACCGAAGAAAAAAAGAAAGGTATACAATCA CGATCAGGAGTTCGACCCCCCTAAGGTGTACCCACCAGTGCCTGCAGAGAAGAGGAAGCCAA TCCGGGTGCTGAGCCTGTTTGATGGCATCGCCACCGGCCTGCTGGTGCTGAAGGATCTGGGC ATCCAGGTGGACCGGTACATCGCCTCCGAGGTGTGCGAGGATTCTATCACCGTGGGCATGGT GCGCCACCAGGGCAAGATCATGTATGTGGGCGACGTGCGGTCCGTGACACAGAAGCACATCC AGGAGTGGGGCCCATTCGATCTGGTGATCGGCGGCAGCCCCTGTAATGACCTGTCCATCGTG AACCCTGCAAGGAAGGGACTGTACGAGGGAACCGGCCGGCTGTTCTTTGAGTTTTATAGACT GCTGCACGACGCCAGGCCTAAGGAGGGCGACGATAGACCATTCTTTTGGCTGTTCGAGAATG TGGTGGCTATGGGCGTGAGCGATAAGAGGGACATCTCCAGGTTTCTGGAGTCTAACCCCGTG ATGATCGATGCAAAGGAGGTGTCCGCCGCACACAGAGCCAGGTATTTCTGGGGCAATCTGCC AGGAATGAACAGGCCACTGGCAAGCACCGTGAATGACAAGCTGGAGCTGCAGGAGTGCCTGG AGCACGGAAGGATCGCCAAGTTTTCCAAGGTGCGCACAATCACCACACGGAGCAATTCCATC AAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAACGAGAAGGAGGACATCCTGTGGTG TACCGAGATGGAGAGAGTGTTCGGCTTTCCAGTGCACTACACAGACGTGTCTAACATGAGCA GGCTGGCAAGGCAGCGGCTGCTGGGCAGATCTTGGAGCGTGCCCGTGATCAGGCACCTGTTC GCCCCTCTGAAGGAGTATTTTGCCTGCGTGAGCAGCGGCAACTCCAATGCCAACAGCCGGGG CCCCTCTTTCAGCTCCGGATTGGTGCCTCTGAGCCTGAGGGGCTCCCACATGGCAGCAATCC CCGCCCTGGACCCCGAGGCCGAGCCTAGCATGGACGTGATCCTGGTGGGCTCTAGCGAGCTG TCCTCTAGCGTGTCTCCAGGAACCGGAAGGGATCTGATCGCATACGAGGTGAAGGCCAATCA GCGGAACATCGAGGACATCTGTATCTGCTGTGGCAGCCTGCAGGTGCACACACAGCACCCAC TGTTCGAGGGAGGAATCTGCGCACCCTGTAAGGATAAGTTCCTGGACGCCCTGTTTCTGTAC GACGATGACGGCTACCAGTCCTATTGCTCTATCTGCTGTTCCGGCGAGACCCTGCTGATCTG CGGCAATCCAGATTGTACAAGGTGCTATTGTTTTGAGTGCGTGGACTCTCTGGTGGGACCAG GCACCAGCGGAAAGGTGCACGCCATGTCCAACTGGGTGTGCTACCTGTGCCTGCCATCCTCT CGCAGCGGACTGCTGCAGCGGAGAAGGAAGTGGAGATCCCAGCTGAAGGCCTTCTATGATAG GGAGTCTGAGAACCCCCTGGAGATGTTTGAGACCGTGCCAGTGTGGCGCCGGCAGCCCGTGA GGGTGCTGAGCCTGTTCGAGGATATCAAGAAGGAGCTGACATCCCTGGGCTTTCTGGAGTCC GGCTCTGACCCCGGACAGCTGAAGCACGTGGTGGATGTGACCGACACAGTGCGGAAGGATGT GGAGGAGTGGGGCCCTTTCGACCTGGTGTACGGAGCAACCCCTCCACTGGGACACACATGCG ACAGACCCCCTTCTTGGTACCTGTTCCAGTTTCACCGCCTGCTGCAGTATGCAAGGCCAAAG CCAGGCAGCCCTAGACCATTCTTTTGGATGTTCGTGGATAATCTGGTGCTGAACAAGGAGGA TCTGGACGTGGCCAGCAGGTTTCTGGAGATGGAGCCAGTGACCATCCCAGACGTGCACGGCG GCTCCCTGCAGAATGCCGTGCGCGTGTGGTCTAACATCCCTGCCATCAGAAGCAGGCACTGG GCACTGGTGAGCGAGGAGGAGCTGTCCCTGCTGGCCCAGAATAAGCAGAGCAGCAAGCTGGC CGCCAAGTGGCCTACAAAGCTGGTGAAGAACTGCTTCCTGCCACTGCGGGAGTACTTCAAGT ATTTTTCCACCGAGCTGACATCTAGCCCCGGAGGACCCTCCTCTGGCGCCCCACCACCTAGC GGCGGCTCCCCTGCCGGCTCTCCAACCAGCACAGAGGAGGGCACCAGCGAGTCCGCCACACC AGAGTCTGGACCTGGCACCAGCACAGAGCCATCCGAGGGCTCTGCCCCAGGCTCTCCTGCAG GCAGCCCTACCTCCACCGAAGAGGGCACCAGCACAGAGCCTTCTGAGGGCAGCGCCCCAGGC ACCTCTACAGAGCCAAGCGAGCTCGAGGACAAGAAGTACAGCATCGGCCTGGCCATCGGCAC

	CAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGG
	TGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGAC
	AGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACG
	GAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACA
	GCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCAC
	CCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCA
	CCTGAGAAAGAAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCC
	TGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAAC
	AGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAA
	CCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCA
	GACGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAAC
	CTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGA
	TGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGA
	TCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTG
	AGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAA
	GAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGC
	CTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGAC
	GGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGG
	CACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCG
	ACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAG
	GAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCG
	CATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAA
	AGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCC
	CAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCC
	CAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACG
	TGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAGGCCATCGTGGAC
	CTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAA
	AATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGG
	GCACATACCACGATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAAC
	GAGGACATTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGA
	GGAACGGCTGAAAACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGC
	GGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAG
	TCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCA
	GCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCC
	AGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGC
	ATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGA
	GAACATCGTGATCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCC
	GCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAA
	CACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGG
	GCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACG
	CCATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGC
	GACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAA
	CTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAGAAAGTTCGACAATCTGACCA
	AGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTG
	GAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTA
	CGACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGGTGT
	CCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCACCACGCC
	CACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGA
	AAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGAGCG
	AGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATGAACTTTTTC
	AAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGG
	CGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGA
	GCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAG
	TCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAA
	GAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGG
	AAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAA
	AGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAA
	AAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGA
	GAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATAT
	GTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGA
	GCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCA
	GCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTAC

	AACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCT GACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGT ACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTAC GAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACAGCCCCAAGAAGAAGAGAAAGGTGGG AGTCGACGGATCCAGCGGCTCCGAGACCCCAGGCACATCTGAGAGCGCCACCCCTGAGTCCC GGACCCTGGTGACCTTCAAGGATGTATTTGTGGACTTCACCAGGGAGGAGTGGAAGCTGCTG GACACTGCTCAGCAGATCGTGTACAGAAATGTGATGCTGGAGAACTATAAGAACCTGGTTTC CTTGGGTTATCAGCTTACTAAGCCAGATGTGATCCTCCGGTTGGAGAAGGGAGAAGAGCCCA GCGCTGATTACAAAGATGATGACGATAAAGCCCCAAAAAAGAAGAGAAAGGTACCGAAGAAA AAAAGAAAGGTCTGA
Fusion Protein 10 Amino Acid Sequence NLS-NLS-3A-h3L-dCas9-ZFP-28-NLS-NLS (SEQ ID NO: 1516)	MGTMPKKKRKVPKKKRKVYNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLG IQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIV NPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPV MIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSI KQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLF APLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMAAIPALDPEAEPSMDVILVGSSEL SSSVSPGTGRDLIAYEVKANQRNIEDICICCGSLQVHTQHPLFEGGICAPCKDKFLDALFLY DDDGYQSYCSICCSGETLLICGNPDCTRCYCFECVDSLVGPGTSGKVHAMSNWVCYLCLPSS RSGLLQRRRKWRSQLKAFYDRESENPLEMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLES GSDPGQLKHVVDVTDTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLLQYARPK PGSPRPFFWMFVDNLVLNKEDLDVASRFLEMEPVTIPDVHGGSLQNAVRVWSNIPAIRSRHW ALVSEEELSLLAQNKQSSKLAAKWPTKLVKNCFLPLREYFKYFSTELTSSLGGPSSGAPPPS GGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPG TSTEPSELEDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFD SGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGN LIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILL SDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYID GGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASA QSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKG ILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKE HPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRS DKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLV ETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHA HDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFF KTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAY NKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLY ETRIDLSQLGGDSPKKKRKVGVDGSSGSETPGTSESATPESTGNKKLEAVGTGIEPKAMSQG LVTFGDVAVDFSQEEWEWLNPIQRNLYRKVMLENYRNLASLGLCVSKPDVISSLEQGKEPWS ADYKDDDDKAPKKKRKVPKKKRKV
Fusion Protein 10 DNA Sequence (SEQ ID NO: 1517)	ATGGGTACCATGCCAAAAAAGAAGAGAAAGGTACCGAAGAAAAAAAGAAAGGTATACAATCA CGATCAGGAGTTCGACCCCCCTAAGGTGTACCCACCAGTGCCTGCAGAGAAGAGGAAGCCAA TCCGGGTGCTGAGCCTGTTTGATGGCATCGCCACCGGCCTGCTGGTGCTGAAGGATCTGGGC ATCCAGGTGGACCGGTACATCGCCTCCGAGGTGTGCGAGGATTCTATCACCGTGGGCATGGT GCGCCACCAGGGCAAGATCATGTATGTGGGCGACGTGCGGTCCGTGACACAGAAGCACATCC AGGAGTGGGGCCCATTCGATCTGGTGATCGGCGGCAGCCCCTGTAATGACCTGTCCATCGTG AACCCTGCAAGGAAGGGACTGTACGAGGGAACCGGCCGGCTGTTCTTTGAGTTTTATAGACT GCTGCACGACGCCAGGCCTAAGGAGGGCGACGATAGACCATTCTTTTGGCTGTTCGAGAATG TGGTGGCTATGGGCGTGAGCGATAAGAGGGACATCTCCAGGTTTCTGGAGTCTAACCCCGTG ATGATCGATGCAAAGGAGGTGTCCGCCGCACACAGAGCCAGGTATTTCTGGGGCAATCTGCC AGGAATGAACAGGCCACTGGCAAGCACCGTGAATGACAAGCTGGAGCTGCAGGAGTGCCTGG AGCACGGAAGGATCGCCAAGTTTTCCAAGGTGCGCACAATCACCACACGGAGCAATTCCATC AAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAACGAGAAGGAGGACATCCTGTGGTG TACCGAGATGGAGAGAGTGTTCGGCTTTCCAGTGCACTACACAGACGTGTCTAACATGAGCA

	GGCTGGCAAGGCAGCGGCTGCTGGGCAGATCTTGGAGCGTGCCCGTGATCAGGCACCTGTTC
	GCCCCTCTGAAGGAGTATTTTGCCTGCGTGAGCAGCGGCAACTCCAATGCCAACAGCCGGGG
	CCCCTCTTTCAGCTCCGGATTGGTGCCTCTGAGCCTGAGGGGCTCCCACATGGCAGCAATCC
	CCGCCCTGGACCCCGAGGCCGAGCCTAGCATGGACGTGATCCTGGTGGGCTCTAGCGAGCTG
	TCCTCTAGCGTGTCTCCAGGAACCGGAAGGGATCTGATCGCATACGAGGTGAAGGCCAATCA
	GCGGAACATCGAGGACATCTGTATCTGCTGTGGCAGCCTGCAGGTGCACACACAGCACCCAC
	TGTTCGAGGGAGGAATCTGCGCACCCTGTAAGGATAAGTTCCTGGACGCCCTGTTTCTGTAC
	GACGATGACGGCTACCAGTCCTATTGCTCTATCTGCTGTTCCGGCGAGACCCTGCTGATCTG
	CGGCAATCCAGATTGTACAAGGTGCTATTGTTTTGAGTGCGTGGACTCTCTGGTGGGACCAG
	GCACCAGCGGAAAGGTGCACGCCATGTCCAACTGGGTGTGCTACCTGTGCCTGCCATCCTCT
	CGCAGCGGACTGCTGCAGCGGAGAAGGAAGTGGAGATCCCAGCTGAAGGCCTTCTATGATAG
	GGAGTCTGAGAACCCCCTGGAGATGTTTGAGACCGTGCCAGTGTGGCGCCGGCAGCCCGTGA
	GGGTGCTGAGCCTGTTCGAGGATATCAAGAAGGAGCTGACATCCCTGGGCTTTCTGGAGTCC
	GGCTCTGACCCCGGACAGCTGAAGCACGTGGTGGATGTGACCGACACAGTGCGGAAGGATGT
	GGAGGAGTGGGGCCCTTTCGACCTGGTGTACGGAGCAACCCCTCCACTGGGACACACATGCG
	ACAGACCCCCTTCTTGGTACCTGTTCCAGTTTCACCGCCTGCTGCAGTATGCAAGGCCAAAG
	CCAGGCAGCCCTAGACCATTCTTTTGGATGTTCGTGGATAATCTGGTGCTGAACAAGGAGGA
	TCTGGACGTGGCCAGCAGGTTTCTGGAGATGGAGCCAGTGACCATCCCAGACGTGCACGGCG
	GCTCCCTGCAGAATGCCGTGCGCGTGTGGTCTAACATCCCTGCCATCAGAAGCAGGCACTGG
	GCACTGGTGAGCGAGGAGGAGCTGTCCCTGCTGGCCCAGAATAAGCAGAGCAGCAAGCTGGC
	CGCCAAGTGGCCTACAAAGCTGGTGAAGAACTGCTTCCTGCCACTGCGGGAGTACTTCAAGT
	ATTTTTCCACCGAGCTGACATCTAGCCTGGGAGGACCCTCCTCTGGCGCCCCACCACCTAGC
	GGCGGCTCCCCTGCCGGCTCTCCAACCAGCACAGAGGAGGGCACCAGCGAGTCCGCCACACC
	AGAGTCTGGACCTGGCACCAGCACAGAGCCATCCGAGGGCTCTGCCCCGGGCTCTCCTGCAG
	GCAGCCCTACCTCCACCGAAGAGGGCACCAGCACAGAGCCTTCTGAGGGCAGCGCCCCAGGC
	ACCTCTACAGAGCCAAGCGAGCTCGAGGACAAGAAGTACAGCATCGGCCTGGCCATCGGCAC
	CAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGG
	TGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGAC
	AGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACG
	GAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACA
	GCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCAC
	CCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCA
	CCTGAGAAAGAAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCC
	TGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAAC
	AGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAA
	CCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCA
	GACGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAAC
	CTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGA
	TGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGA
	TCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTG
	AGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAA
	GAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGC
	CTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGAC
	GGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGG
	CACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCG
	ACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAG
	GAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCG
	CATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAA
	AGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCC
	CAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCC
	CAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACG
	TGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAGGCCATCGTGGAC
	CTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAA
	AATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGG
	GCACATACCACGATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAAC
	GAGGACATTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGA
	GGAACGGCTGAAAACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGC
	GGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAG
	TCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCA
	GCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCC
	AGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGC
	ATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGA

	GAACATCGTGATCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCC GCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAA CACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGG GCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACG CCATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGC GACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAA CTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAGAAAGTTCGACAATCTGACCA AGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTG GAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTA CGACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGGTGT CCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCACCACGCC CACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGA AAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGAGCG AGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATGAACTTTTTC AAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGG CGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGA GCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAG TCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAA GAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGG AAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAA AGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAA AAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGA GAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATAT GTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGA GCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCA GCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTAC AACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCT GACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGT ACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTAC GAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACAGCCCCAAGAAGAAGAGAAAGGTGGG AGTCGACGGATCCAGCGGCTCCGAGACCCCAGGCACATCTGAGAGCGCCACCCCTGAGTCCA CCGGTAACAAAAAGCTTGAGGCCGTCGGAACCGGAATCGAACCAAAAGCAATGTCCCAGGGT TTGGTGACATTTGGCGACGTGGCTGTCGATTTTTCCCAGGAAGAGTGGGAGTGGCTCAATCC TATCCAGAGGAACTTGTACCGGAAGGTGATGCTGGAGAATTATAGAAATTTGGCATCACTGG GGTTGTGCGTTAGCAAACCAGATGTTATATCTTCCCTGGAACAGGGAAAGGAGCCCTGGAGC GCTGATTACAAAGATGATGACGATAAAGCCCCCAAGAAGAAAAGGAAGGTCCCAAAGAAAAA AAGAAAGGTGTGA
Fusion Protein 11 Amino Acid Sequence NLS-NLS-3A-h3L-dCas9-ZIM3-NLS-NLS (SEQ ID NO: 1518)	MGTMPKKKRKVPKKKRKVYNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLG IQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIV NPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPV MIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSI KQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLF APLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMAAIPALDPEAEPSMDVILVGSSEL SSSVSPGTGRDLIAYEVKANQRNIEDICICCGSLQVHTQHPLFEGGICAPCKDKFLDALFLY DDDGYQSYCSICCSGETLLICGNPDCTRCYCFECVDSLVGPGTSGKVHAMSNWVCYLCLPSS RSGLLQRRRKWRSQLKAFYDRESENPLEMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLES GSDPGQLKHVVDVTDTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLLQYARPK PGSPRPFFWMFVDNLVLNKEDLDVASRFLEMEPVTIPDVHGGSLQNAVRVWSNIPAIRSRHW ALVSEEELSLLAQNKQSSKLAAKWPTKLVKNCFLPLREYFKYFSTELTSSLGGPSSGAPPPS GGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPG TSTEPSELEDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFD SGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGN LIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILL SDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYID GGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASA QSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKG

	ILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKE HPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRS DKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLV ETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHA HDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFF KTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAY NKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLY ETRIDLSQLGGDSPKKKRKVGVDGSSGSETPGTSESATPESTGMNNSQGRVTFEDVTVNFTQ GEWQRLNPEQRNLYRDVMLENYSNLVSVGQGETTKPDVILRLEQGKEPWLEEEEVLGSGRAE KNGDIGGQIWKPKDVKESLSADYKDDDDKAPKKKRKVPKKKRKV
Fusion Protein 11 DNA Sequence (SEQ ID NO: 1519)	ATGGGTACCATGCCAAAAAAGAAGAGAAAGGTACCGAAGAAAAAAAGAAAGGTATACAATCA CGATCAGGAGTTCGACCCCCCTAAGGTGTACCCACCAGTGCCTGCAGAGAAGAGGAAGCCAA TCCGGGTGCTGAGCCTGTTTGATGGCATCGCCACCGGCCTGCTGGTGCTGAAGGATCTGGGC ATCCAGGTGGACCGGTACATCGCCTCCGAGGTGTGCGAGGATTCTATCACCGTGGGCATGGT GCGCCACCAGGGCAAGATCATGTATGTGGGCGACGTGCGGTCCGTGACACAGAAGCACATCC AGGAGTGGGGCCCATTCGATCTGGTGATCGGCGGCAGCCCCTGTAATGACCTGTCCATCGTG AACCCTGCAAGGAAGGGACTGTACGAGGGAACCGGCCGGCTGTTCTTTGAGTTTTATAGACT GCTGCACGACGCCAGGCCTAAGGAGGGCGACGATAGACCATTCTTTTGGCTGTTCGAGAATG TGGTGGCTATGGGCGTGAGCGATAAGAGGGACATCTCCAGGTTTCTGGAGTCTAACCCCGTG ATGATCGATGCAAAGGAGGTGTCCGCCGCACACAGAGCCAGGTATTTCTGGGGCAATCTGCC AGGAATGAACAGGCCACTGGCAAGCACCGTGAATGACAAGCTGGAGCTGCAGGAGTGCCTGG AGCACGGAAGGATCGCCAAGTTTTCCAAGGTGCGCACAATCACCACACGGAGCAATTCCATC AAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAACGAGAAGGAGGACATCCTGTGGTG TACCGAGATGGAGAGAGTGTTCGGCTTTCCAGTGCACTACACAGACGTGTCTAACATGAGCA GGCTGGCAAGGCAGCGGCTGCTGGGCAGATCTTGGAGCGTGCCCGTGATCAGGCACCTGTTC GCCCCTCTGAAGGAGTATTTTGCCTGCGTGAGCAGCGGCAACTCCAATGCCAACAGCCGGGG CCCCTCTTTCAGCTCCGGATTGGTGCCTCTGAGCCTGAGGGGCTCCCACATGGCAGCAATCC CCGCCCTGGACCCCGAGGCCGAGCCTAGCATGGACGTGATCCTGGTGGGCTCTAGCGAGCTG TCCTCTAGCGTGTCTCCAGGAACCGGAAGGGATCTGATCGCATACGAGGTGAAGGCCAATCA GCGGAACATCGAGGACATCTGTATCTGCTGTGGCAGCCTGCAGGTGCACACACAGCACCCAC TGTTCGAGGGAGGAATCTGCGCACCCTGTAAGGATAAGTTCCTGGACGCCCTGTTTCTGTAC GACGATGACGGCTACCAGTCCTATTGCTCTATCTGCTGTTCCGGCGAGACCCTGCTGATCTG CGGCAATCCAGATTGTACAAGGTGCTATTGTTTTGAGTGCGTGGACTCTCTGGTGGGACCAG GCACCAGCGGAAAGGTGCACGCCATGTCCAACTGGGTGTGCTACCTGTGCCTGCCATCCTCT CGCAGCGGACTGCTGCAGCGGAGAAGGAAGTGGAGATCCCAGCTGAAGGCCTTCTATGATAG GGAGTCTGAGAACCCCCTGGAGATGTTTGAGACCGTGCCAGTGTGGCGCCGGCAGCCCGTGA GGGTGCTGAGCCTGTTCGAGGATATCAAGAAGGAGCTGACATCCCTGGGCTTTCTGGAGTCC GGCTCTGACCCCGGACAGCTGAAGCACGTGGTGGATGTGACCGACACAGTGCGGAAGGATGT GGAGGAGTGGGGCCCTTTCGACCTGGTGTACGGAGCAACCCCTCCACTGGGACACACATGCG ACAGACCCCCTTCTTGGTACCTGTTCCAGTTTCACCGCCTGCTGCAGTATGCAAGGCCAAAG CCAGGCAGCCCTAGACCATTCTTTTGGATGTTCGTGGATAATCTGGTGCTGAACAAGGAGGA TCTGGACGTGGCCAGCAGGTTTCTGGAGATGGAGCCAGTGACCATCCCAGACGTGCACGGCG GCTCCCTGCAGAATGCCGTGCGCGTGTGGTCTAACATCCCTGCCATCAGAAGCAGGCACTGG GCACTGGTGAGCGAGGAGGAGCTGTCCCTGCTGGCCCAGAATAAGCAGAGCAGCAAGCTGGC CGCCAAGTGGCCTACAAAGCTGGTGAAGAACTGCTTCCTGCCACTGCGGGAGTACTTCAAGT ATTTTTCCACCGAGCTGACATCTAGCCTGGGAGGACCCTCCTCTGGCGCCCCACCACCTAGC GGCGGCTCCCCTGCCGGCTCTCCAACCAGCACAGAGGAGGGCACCAGCGAGTCCGCCACACC AGAGTCTGGACCTGGCACCAGCACAGAGCCATCCGAGGGCTCTGCCCCAGGCTCTCCTGCAG GCAGCCCTACCTCCACCGAAGAGGGCACCAGCACAGAGCCTTCTGAGGGCAGCGCCCCAGGC ACCTCTACAGAGCCAAGCGAGCTCGAGGACAAGAAGTACAGCATCGGCCTGGCCATCGGCAC CAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGG TGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGAC AGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACG GAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACA GCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCAC CCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCA CCTGAGAAAGAAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCC TGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAAC AGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAA

	CCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCA
	GACGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAAC
	CTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGA
	TGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGA
	TCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTG
	AGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAA
	GAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGC
	CTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGAC
	GGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGG
	CACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCG
	ACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAG
	GAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCG
	CATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAA
	AGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCC
	CAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCC
	CAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACG
	TGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAGGCCATCGTGGAC
	CTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAA
	AATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGG
	GCACATACCACGATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAAC
	GAGGACATTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGA
	GGAACGGCTGAAAACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGC
	GGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAG
	TCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCA
	GCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCC
	AGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGC
	ATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGA
	GAACATCGTGATCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCC
	GCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAA
	CACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGG
	GCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACG
	CCATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGC
	GACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAA
	CTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAGAAAGTTCGACAATCTGACCA
	AGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTG
	GAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTA
	CGACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGGTGT
	CCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCACCACGCC
	CACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGA
	AAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGAGCG
	AGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATGAACTTTTTC
	AAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGG
	CGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGA
	GCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAG
	TCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAA
	GAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGG
	AAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAA
	AGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAA
	AAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGA
	GAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATAT
	GTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGA
	GCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCA
	GCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTAC
	AACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCT
	GACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGT
	ACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTAC
	GAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACAGCCCCAAGAAGAAGAGAAAGGTGGG
	AGTCGACGGATCCAGCGGCTCCGAGACCCCAGGCACATCTGAGAGCGCCACCCCTGAGTCCA
	CCGGTATGAACAATTCACAGGGGAGAGTGACATTCGAAGACGTGACCGTGAACTTCACCCAG
	GGAGAATGGCAGCGCTTGAACCCAGAACAAAGGAACCTCTATCGGGACGTGATGCTGGAAAA
	CTACTCAAATTTGGTGAGCGTTGGGCAGGGTGAGACCACTAAGCCTGACGTGATCCTGAGAT
	TGGAACAGGGCAAGGAGCCTTGGCTCGAGGAAGAGGAAGTCCTGGGCTCAGGGAGGGCCGAG

	AAAAACGGTGATATAGGAGGCCAGATATGGAAGCCTAAGGACGTCAAGGAGAGCCTGAGCGC TGATTACAAAGATGATGACGATAAAGCCCCCAAGAAGAAAAGGAAGGTCCCAAAGAAAAAAA GAAAGGTGTGA
Fusion Protein 12 Amino Acid Sequence NLS-NLS-3A-h3L-dCas9-ZN627-NLS-NLS (SEQ ID NO: 1520)	MGTMPKKKRKVPKKKRKVYNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLG IQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIV NPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPV MIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSI KQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLF APLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMAAIPALDPEAEPSMDVILVGSSEL SSSVSPGTGRDLIAYEVKANQRNIEDICICCGSLQVHTQHPLFEGGICAPCKDKFLDALFLY DDDGYQSYCSICCSGETLLICGNPDCTRCYCFECVDSLVGPGTSGKVHAMSNWVCYLCLPSS RSGLLQRRRKWRSQLKAFYDRESENPLEMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLES GSDPGQLKHVVDVTDTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLLQYARPK PGSPRPFFWMFVDNLVLNKEDLDVASRFLEMEPVTIPDVHGGSLQNAVRVWSNIPAIRSRHW ALVSEEELSLLAQNKQSSKLAAKWPTKLVKNCFLPLREYFKYFSTELTSSLGGPSSGAPPPS GGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPG TSTEPSELEDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFD SGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGN LIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILL SDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYID GGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASA QSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKG ILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKE HPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRS DKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLV ETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHA HDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFF KTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAY NKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLY ETRIDLSQLGGDSPKKKRKVGVDGSSGSETPGTSESATPESTGDSVAFEDVAVNFTLEEWAL LDPSQKNLYRDVMRETFRNLASVGKQWEDQNIEDPFKIPRRNISHIPERLCESKEGGQGEES ADYKDDDDKAPKKKRKVPKKKRKV
Fusion Protein 12 DNA Sequence (SEQ ID NO: 1521)	ATGGGTACCATGCCAAAAAAGAAGAGAAAGGTACCGAAGAAAAAAAGAAAGGTATACAATCA CGATCAGGAGTTCGACCCCCCTAAGGTGTACCCACCAGTGCCTGCAGAGAAGAGGAAGCCAA TCCGGGTGCTGAGCCTGTTTGATGGCATCGCCACCGGCCTGCTGGTGCTGAAGGATCTGGGC ATCCAGGTGGACCGGTACATCGCCTCCGAGGTGTGCGAGGATTCTATCACCGTGGGCATGGT GCGCCACCAGGGCAAGATCATGTATGTGGGCGACGTGCGGTCCGTGACACAGAAGCACATCC AGGAGTGGGGCCCATTCGATCTGGTGATCGGCGGCAGCCCCTGTAATGACCTGTCCATCGTG AACCCTGCAAGGAAGGGACTGTACGAGGGAACCGGCCGGCTGTTCTTTGAGTTTTATAGACT GCTGCACGACGCCAGGCCTAAGGAGGGCGACGATAGACCATTCTTTTGGCTGTTCGAGAATG TGGTGGCTATGGGCGTGAGCGATAAGAGGGACATCTCCAGGTTTCTGGAGTCTAACCCCGTG ATGATCGATGCAAAGGAGGTGTCCGCCGCACACAGAGCCAGGTATTTCTGGGGCAATCTGCC AGGAATGAACAGGCCACTGGCAAGCACCGTGAATGACAAGCTGGAGCTGCAGGAGTGCCTGG AGCACGGAAGGATCGCCAAGTTTTCCAAGGTGCGCACAATCACCACACGGAGCAATTCCATC AAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAACGAGAAGGAGGACATCCTGTGGTG TACCGAGATGGAGAGAGTGTTCGGCTTTCCAGTGCACTACACAGACGTGTCTAACATGAGCA GGCTGGCAAGGCAGCGGCTGCTGGGCAGATCTTGGAGCGTGCCCGTGATCAGGCACCTGTTC GCCCCTCTGAAGGAGTATTTTGCCTGCGTGAGCAGCGGCAACTCCAATGCCAACAGCCGGGG CCCCTCTTTCAGCTCCGGATTGGTGCCTCTGAGCCTGAGGGGCTCCCACATGGCAGCAATCC CCGCCCTGGACCCCGAGGCCGAGCCTAGCATGGACGTGATCCTGGTGGGCTCTAGCGAGCTG TCCTCTAGCGTGTCTCCAGGAACCGGAAGGGATCTGATCGCATACGAGGTGAAGGCCAATCA GCGGAACATCGAGGACATCTGTATCTGCTGTGGCAGCCTGCAGGTGCACACACAGCACCCAC TGTTCGAGGGAGGAATCTGCGCACCCTGTAAGGATAAGTTCCTGGACGCCCTGTTTCTGTAC

	GACGATGACGGCTACCAGTCCTATTGCTCTATCTGCTGTTCCGGCGAGACCCTGCTGATCTG
	CGGCAATCCAGATTGTACAAGGTGCTATTGTTTTGAGTGCGTGGACTCTCTGGTGGGACCAG
	GCACCAGCGGAAAGGTGCACGCCATGTCCAACTGGGTGTGCTACCTGTGCCTGCCATCCTCT
	CGCAGCGGACTGCTGCAGCGGAGAAGGAAGTGGAGATCCCAGCTGAAGGCCTTCTATGATAG
	GGAGTCTGAGAACCCCCTGGAGATGTTTGAGACCGTGCCAGTGTGGCGCCGGCAGCCCGTGA
	GGGTGCTGAGCCTGTTCGAGGATATCAAGAAGGAGCTGACATCCCTGGGCTTTCTGGAGTCC
	GGCTCTGACCCCGGACAGCTGAAGCACGTGGTGGATGTGACCGACACAGTGCGGAAGGATGT
	GGAGGAGTGGGGCCCTTTCGACCTGGTGTACGGAGCAACCCCTCCACTGGGACACACATGCG
	ACAGACCCCCTTCTTGGTACCTGTTCCAGTTTCACCGCCTGCTGCAGTATGCAAGGCCAAAG
	CCAGGCAGCCCTAGACCATTCTTTTGGATGTTCGTGGATAATCTGGTGCTGAACAAGGAGGA
	TCTGGACGTGGCCAGCAGGTTTCTGGAGATGGAGCCAGTGACCATCCCAGACGTGCACGGCG
	GCTCCCTGCAGAATGCCGTGCGCGTGTGGTCTAACATCCCTGCCATCAGAAGCAGGCACTGG
	GCACTGGTGAGCGAGGAGGAGCTGTCCCTGCTGGCCCAGAATAAGCAGAGCAGCAAGCTGGC
	CGCCAAGTGGCCTACAAAGCTGGTGAAGAACTGCTTCCTGCCACTGCGGGAGTACTTCAAGT
	ATTTTTCCACCGAGCTGACATCTAGCCTGGGAGGACCCTCCTCTGGCGCCCCACCACCTAGC
	GGCGGCTCCCCTGCCGGCTCTCCAACCAGCACAGAGGAGGGCACCAGCGAGTCCGCCACACC
	AGAGTCTGGACCTGGCACCAGCACAGAGCCATCCGAGGGCTCTGCCCCAGGCTCTCCTGCAG
	GCAGCCCTACCTCCACCGAAGAGGGCACCAGCACAGAGCCTTCTGAGGGCAGCGCCCCAGGC
	ACCTCTACAGAGCCAAGCGAGCTCGAGGACAAGAAGTACAGCATCGGCCTGGCCATCGGCAC
	CAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGG
	TGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGAC
	AGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACG
	GAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACA
	GCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCAC
	CCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCA
	CCTGAGAAAGAAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCC
	TGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAAC
	AGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAA
	CCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCA
	GACGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAAC
	CTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGA
	TGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGA
	TCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTG
	AGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAA
	GAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGC
	CTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGAC
	GGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGG
	CACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCG
	ACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAG
	GAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCG
	CATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAA
	AGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCC
	CAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCC
	CAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACG
	TGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAGGCCATCGTGGAC
	CTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAA
	AATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGG
	GCACATACCACGATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAAC
	GAGGACATTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGA
	GGAACGGCTGAAAACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGC
	GGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAG
	TCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCA
	GCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCC
	AGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGC
	ATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGA
	GAACATCGTGATCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCC
	GCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAA
	CACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGG
	GCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACG
	CCATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGC
	GACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAA
	CTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAGAAAGTTCGACAATCTGACCA

	AGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTG GAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTA CGACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGGTGT CCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCACCACGCC CACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGA AAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGAGCG AGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATGAACTTTTTC AAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGG CGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGA GCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAG TCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAA GAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGG AAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAA AGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAA AAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGA GAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATAT GTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGA GCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCA GCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTAC AACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCT GACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGT ACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTAC GAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACAGCCCCAAGAAGAAGAGAAAGGTGGG AGTCGACGGATCCAGCGGCTCCGAGACCCCAGGCACATCTGAGAGCGCCACCCCTGAGTCCA CCGGTGACTCCGTTGCTTTCGAGGACGTGGCCGTGAACTTCACACTTGAGGAATGGGCCTTG CTCGACCCAAGTCAGAAGAATCTGTACAGAGACGTGATGCGGGAGACATTCAGGAATCTCGC CAGTGTCGGAAAGCAGTGGGAAGACCAGAACATCGAAGATCCTTTCAAGATACCACGGCGCA ATATCTCCCACATTCCTGAGAGGCTGTGTGAATCTAAGGAAGGCGGACAAGGTGAGGAAAGC GCTGATTACAAAGATGATGACGATAAAGCCCCCAAGAAGAAAAGGAAGGTCCCAAAGAAAAA AAGAAAGGTGTGA
Fusion Protein 13 Amino Acid Sequence 3A-3L-NLS-dCas9-NLS-KOX1KRAB (SEQ ID NO: 1525)	MGTMNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDS ITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLF FEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRARY FWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEK EDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNS NANSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESG SGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALP RQESQRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKH APLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGPSSGAPPPS GGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPG TSTEPSEPKKKRKVYMDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNL IGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEE DKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEK KNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKN LSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGEL HAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEV VDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGE QKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKD FLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLI NGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG SPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKEL GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSID NKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAG FIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVRE INNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQ TGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKEL LGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNE LALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANL DKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH

	QSITGLYETRIDLSQLGGDPKKKRKVSGSETPGTSESATPESTGRTLVTFKDVFVDFTREEW KLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEP
Fusion Protein 13 DNA Sequence (SEQ ID NO: 1526)	ATGGGTACCATGAACCATGACCAGGAATTTGACCCCCCAAAGGTTTACCCACCTGTGCCAGC TGAGAAGAGGAAGCCCATCCGCGTGCTGTCTCTCTTTGATGGGATTGCTACAGGGCTCCTGG TGCTGAAGGACCTGGGCATCCAAGTGGACCGCTACATTGCCTCCGAGGTGTGTGAGGACTCC ATCACGGTGGGCATGGTGCGGCACCAGGGAAAGATCATGTACGTCGGGGACGTCCGCAGCGT CACACAGAAGCATATCCAGGAGTGGGGCCCATTCGACCTGGTGATTGGAGGCAGTCCCTGCA ATGACCTCTCCATTGTCAACCCTGCCCGCAAGGGACTTTATGAGGGTACTGGCCGCCTCTTC TTTGAGTTCTACCGCCTCCTGCATGATGCGCGGCCCAAGGAGGGAGATGATCGCCCCTTCTT CTGGCTCTTTGAGAATGTGGTGGCCATGGGCGTTAGTGACAAGAGGGACATCTCGCGATTTC TTGAGTCTAACCCCGTGATGATTGACGCCAAAGAAGTGTCTGCTGCACACAGGGCCCGTTAC TTCTGGGGTAACCTTCCTGGCATGAACAGGCCTTTGGCATCCACTGTGAATGATAAGCTGGA GCTGCAAGAGTGTCTGGAGCACGGCAGAATAGCCAAGTTCAGCAAAGTGAGGACCATTACCA CCAGGTCAAACTCTATAAAGCAGGGCAAAGACCAGCATTTCCCCGTCTTCATGAACGAGAAG GAGGACATCCTGTGGTGCACTGAAATGGAAAGGGTGTTTGGCTTCCCCGTCCACTACACAGA CGTCTCCAACATGAGCCGCTTGGCGAGGCAGAGACTGCTGGGCCGATCGTGGAGCGTGCCGG TCATCCGCCACCTCTTCGCTCCGCTGAAGGAATATTTTGCTTGTGTGTCTAGCGGCAATAGT AACGCTAACAGCCGCGGGCCGAGCTTCAGCAGCGGCCTGGTGCCGTTAAGCTTGCGCGGCAG CCATATGGGCCCTATGGAGATATACAAGACAGTGTCTGCATGGAAGAGACAGCCAGTGCGGG TACTGAGCCTCTTCAGAAACATCGACAAGGTACTAAAGAGTTTGGGCTTCTTGGAAAGCGGT TCTGGTTCTGGGGGAGGAACGCTGAAGTACGTGGAAGATGTCACAAATGTCGTGAGGAGAGA CGTGGAGAAATGGGGCCCCTTTGACCTGGTGTACGGCTCGACGCAGCCCCTAGGCAGCTCTT GTGATCGCTGTCCCGGCTGGTACATGTTCCAGTTCCACCGGATCCTGCAGTATGCGCTGCCT CGCCAGGAGAGTCAGCGGCCCTTCTTCTGGATATTCATGGACAATCTGCTGCTGACTGAGGA TGACCAAGAGACAACTACCCGCTTCCTTCAGACAGAGGCTGTGACCCTCCAGGATGTCCGTG GCAGAGACTACCAGAATGCTATGCGGGTGTGGAGCAACATTCCAGGGCTGAAGAGCAAGCAT GCGCCCCTGACCCCAAAGGAAGAAGAGTATCTGCAAGCCCAAGTCAGAAGCAGGAGCAAGCT GGACGCCCCGAAAGTTGACCTCCTGGTGAAGAACTGCCTTCTCCCGCTGAGAGAGTACTTCA AGTATTTTTCTCAAAACTCACTTCCTCTTGGAGGGCCGAGCTCTGGCGCACCCCCACCAAGT GGAGGGTCTCCTGCCGGGTCCCCAACATCTACTGAAGAAGGCACCAGCGAATCCGCAACGCC CGAGTCAGGCCCTGGTACCTCCACAGAACCATCTGAAGGTAGTGCGCCTGGTTCCCCAGCTG GAAGCCCTACTTCCACCGAAGAAGGCACGTCAACCGAACCAAGTGAAGGATCTGCCCCTGGG ACCAGCACTGAACCATCTGAGCCAAAAAAGAAGAGAAAGGTATACATGGACAAGAAGTACAG CATCGGCCTGGCCATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAAGG TGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTG ATCGGCGCCCTGCTGTTCGACAGCGGAGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGC CAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACG AGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAG GATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGA GAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCACCGACAAGGCCGACC TGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAG GGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTA CAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGT CTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAG AAGAATGGCCTGTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAG CAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACC TGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAAC CTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCC CCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAG CTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAAC GGCTACGCCGGCTACATCGATGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCC CATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGC TGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTG CACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGAT CGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGCA GATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAAGTG GTGGACAAGGGCGCCAGCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACCT GCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTACAACG AGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAG CAGAAAAAAGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCT GAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAG ATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGGACAAGGAC

	TTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCCTGACACTGTT TGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCTGTTCGACGACAAAG TGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATC AACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTT CGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCC AGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCCGGC AGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAAGT GATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACCAGACCACCC AGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAGCTG GGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTA CCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGC TGTCCGACTACGATGTGGACGCTATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGAT AACAAAGTGCTGACTCGGAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGA GGTCGTGAAGAAGATGAAGAACTACTGGCGCCAGCTGCTGAATGCCAAGCTGATTACCCAGA GGAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGC TTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGA CTCCCGGATGAACACTAAGTACGACGAGAACGACAAACTGATCCGGGAAGTGAAAGTGATCA CCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAG ATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGAT CAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGC GGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTAC AGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCG GCCTCTGATCGAGACAAACGGCGAAACAGGCGAGATCGTGTGGGATAAGGGCCGGGACTTTG CCACCGTGCGGAAAGTGCTGTCTATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAG ACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGACAAGCTGATCGCCAG AAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTG TGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAAGAGCTG CTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGC CAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCG AGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAA CTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAA GGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAACACTACCTGG ACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTG GACAAGGTGCTGAGCGCCTACAACAAGCACAGAGACAAGCCTATCAGAGAGCAGGCCGAGAA TATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACA CCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGATCCAC CAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACCCAAA AAAGAAGAGAAAGGTAAGCGGAAGTGAGACCCCAGGTACATCCGAATCAGCAACGCCTGAAA GCACCGGTCGGACACTGGTGACCTTCAAGGATGTATTTGTGGACTTCACCAGGGAGGAGTGG AAGCTGCTGGACACTGCTCAGCAGATCGTGTACAGAAATGTGATGCTGGAGAACTATAAGAA CCTGGTTTCCTTGGGTTATCAGCTTACTAAGCCAGATGTGATCCTCCGGTTGGAGAAGGGAG AAGAGCCCTGA
Fusion Protein 14 Amino Acid Sequence NLS-NLS-3A-3L-dCas9-ZIM3-NLS-NLS (SEQ ID NO: 1527)	MGTMPKKKRKVPKKKRKVYNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLG IQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIV NPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPV MIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSI KQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLF APLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRQPVRVLSLFR NIDKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPG WYMFQFHRILQYALPRQESQRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQN AMRVWSNIPGLKSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQN SLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEMDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDR HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLE ESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKF RGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLI AQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRK PAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLL

	KIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWG RLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRI EEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSF LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDF QFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKA TAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNI VKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKK LKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAG ELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRV ILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEV LDATLIHQSITGLYETRIDLSQLGGDSGSETPGTSESATPESTGMNNSQGRVTFEDVTVNFT QGEWQRLNPEQRNLYRDVMLENYSNLVSVGQGETTKPDVILRLEQGKEPWLEEEEVLGSGRA EKNGDIGGQIWKPKDVKESLSADYKDDDDKAPKKKRKVPKKKRKV
Fusion Protein 14 DNA Sequence (SEQ ID NO: 1528)	ATGGGTACCATGCCAAAAAAGAAGAGAAAGGTACCGAAGAAAAAAAGAAAGGTATACAACCA TGACCAGGAATTCGACCCCCCAAAGGTTTACCCACCTGTGCCAGCTGAGAAGAGGAAGCCCA TCCGCGTGCTGTCTCTCTTTGATGGGATTGCTACAGGGCTCCTGGTGCTGAAGGACCTGGGC ATCCAAGTGGACCGCTACATTGCCTCCGAGGTGTGTGAGGACTCCATCACGGTGGGCATGGT GCGGCACCAGGGAAAGATCATGTACGTCGGGGACGTCCGCAGCGTCACACAGAAGCATATCC AGGAGTGGGGCCCATTCGACCTGGTGATTGGAGGCAGTCCCTGCAATGACCTCTCCATTGTC AACCCTGCCCGCAAGGGACTTTATGAGGGTACTGGCCGCCTCTTCTTTGAGTTCTACCGCCT CCTGCATGATGCGCGGCCCAAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATG TGGTGGCCATGGGCGTTAGTGACAAGAGGGACATCTCGCGATTTCTTGAGTCTAACCCCGTG ATGATTGACGCCAAAGAAGTGTCTGCTGCACACAGGGCCCGTTACTTCTGGGGTAACCTTCC TGGCATGAACAGGCCTTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAAGAGTGTCTGG AGCACGGCAGAATAGCCAAGTTCAGCAAAGTGAGGACCATTACCACCAGGTCAAACTCTATA AAGCAGGGCAAAGACCAGCATTTCCCCGTCTTCATGAACGAGAAGGAGGACATCCTGTGGTG CACTGAAATGGAAAGGGTGTTTGGCTTCCCCGTCCACTACACAGACGTCTCCAACATGAGCC GCTTGGCGAGGCAGAGACTGCTGGGCCGATCGTGGAGCGTGCCGGTCATCCGCCACCTCTTC GCTCCGCTGAAGGAATATTTTGCTTGTGTGTCTAGCGGCAATAGTAACGCTAACAGCCGCGG GCCGAGCTTCAGCAGCGGCCTGGTGCCGTTAAGCTTGCGCGGCAGCCATATGGGCCCTATGG AGATATACAAGACAGTGTCTGCATGGAAGAGACAGCCAGTGCGGGTACTGAGCCTCTTCAGA AACATCGACAAGGTACTAAAGAGTTTGGGCTTCTTGGAAAGCGGTTCTGGTTCTGGGGGAGG AACGCTGAAGTACGTGGAAGATGTCACAAATGTCGTGAGGAGAGACGTGGAGAAATGGGGCC CCTTTGACCTGGTGTACGGCTCGACGCAGCCCCTAGGCAGCTCTTGTGATCGCTGTCCCGGC TGGTACATGTTCCAGTTCCACCGGATCCTGCAGTATGCGCTGCCTCGCCAGGAGAGTCAGCG GCCCTTCTTCTGGATATTCATGGACAATCTGCTGCTGACTGAGGATGACCAAGAGACAACTA CCCGCTTCCTTCAGACAGAGGCTGTGACCCTCCAGGATGTCCGTGGCAGAGACTACCAGAAT GCTATGCGGGTGTGGAGCAACATTCCAGGGCTGAAGAGCAAGCATGCGCCCCTGACCCCAAA GGAAGAAGAGTATCTGCAAGCCCAAGTCAGAAGCAGGAGCAAGCTGGACGCCCCGAAAGTTG ACCTCCTGGTGAAGAACTGCCTTCTCCCGCTGAGAGAGTACTTCAAGTATTTTTCTCAAAAC TCACTTCCTCTTGGAGGGCCGAGCTCTGGCGCACCCCCACCAAGTGGAGGGTCTCCTGCCGG GTCCCCAACATCTACTGAAGAAGGCACCAGCGAATCCGCAACGCCCGAGTCAGGCCCTGGTA CCTCCACAGAACCATCTGAAGGTAGTGCGCCTGGTTCCCCAGCTGGAAGCCCTACTTCCACC GAAGAAGGCACGTCAACCGAACCAAGTGAAGGATCTGCCCCTGGGACCAGCACTGAACCATC TGAGATGGACAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACTCTGTGGGCTGGGCCG TGATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGG CACAGCATCAAGAAGAACCTGATCGGCGCCCTGCTGTTCGACAGCGGAGAAACAGCCGAGGC CACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATC TGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAA GAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGT GGACGAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGG ACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTC CGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTT CATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCG TGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATC GCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAACCTGATTGCCCTGAGCCTGGG CCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCA AGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGAC CTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAA CACCGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACC

	AGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATT TTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATCGATGGCGGAGCCAGCCAGGAAGA GTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGA AGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCAC CAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCT GAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCC CTCTGGCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACC CCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCCAGCGCCCAGAGCTTCATCGAGCGGAT GACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACG AGTACTTCACCGTGTACAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAG CCCGCCTTCCTGAGCGGCGAGCAGAAAAAAGCCATCGTGGACCTGCTGTTCAAGACCAACCG GAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCG TGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTG AAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATAT CGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATG CCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGC AGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGA TTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGCC TGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAG CACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGT GGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGG CCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATC GAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCA GCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACC AGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACGCTATCGTGCCTCAGAGCTTT CTGAAGGACGACTCCATCGATAACAAAGTGCTGACTCGGAGCGACAAGAACCGGGGCAAGAG CGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGCCAGCTGCTGA ATGCCAAGCTGATTACCCAGAGGAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTG AGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAA GCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAACGACAAACTGA TCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTC CAGTTTTACAAAGTGCGCGAGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGC CGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCG ACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCT ACCGCCAAGTACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGC CAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACAGGCGAGATCGTGT GGGATAAGGGCCGGGACTTTGCCACCGTGCGGAAAGTGCTGTCTATGCCCCAAGTGAATATC GTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAA CAGCGACAAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACA GCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAA CTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAA TCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGC TGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGC GAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGC CAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGG AACAGCACAAACACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTG ATCCTGGCCGACGCTAATCTGGACAAGGTGCTGAGCGCCTACAACAAGCACAGAGACAAGCC TATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTG CCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTG CTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTC TCAGCTGGGAGGCGACAGCGGAAGTGAGACCCCAGGTACATCCGAATCAGCAACGCCTGAAA GCACCGGTATGAACAATTCACAGGGGAGAGTGACATTCGAAGACGTGACCGTGAACTTCACC CAGGGAGAATGGCAGCGCTTGAACCCAGAACAAAGGAACCTCTATCGGGACGTGATGCTGGA AAACTACTCAAATTTGGTGAGCGTTGGGCAGGGTGAGACCACTAAGCCTGACGTGATCCTGA GATTGGAACAGGGCAAGGAGCCTTGGCTCGAGGAAGAGGAAGTCCTGGGCTCAGGGAGGGCC GAGAAAAACGGTGATATAGGAGGCCAGATATGGAAGCCTAAGGACGTCAAGGAGAGCCTGAG CGCTGATTACAAAGATGATGACGATAAAGCCCCAAAAAAGAAGAGAAAGGTACCGAAGAAAA AAAGAAAGGTCTGA

[0292] Sequences from **FIG. 14A** can be found below:

Description	Sequence	SEQ ID NO:
gRNA009	ACUGCCUGGCUCACUCCUCC	1270
gRNA003	UGCGGAAACCUUCUAGGGUG	1264
gRNA093	AACCUGAUCCUCCAGUCCGG	1354
gRNA011	AUCGUCCGAUGGGGCUCUGG	1272
gRNA007	GGUGUGGGUGCUUGACGCCU	1268
gRNA077	CAAAUCCUAACUGGGCUGGA	1338
gRNA113	GGGCGCCGCCGUUCAGUUCA	1374
gRNA004	GCGGAAACCUUCUAGGGUGU	1265
gRNA008	GUGUGGGUGCUUGACGCCUG	1269
gRNA012	AGGAUCGUCCGAUGGGGCUC	1273
gRNA111	AGUUCAGGGUCUGAGCCUGG	1372
gRNA005	UCAAGCACCCACACCCUAGA	1266
gRNA013	UCAGAUAGGAUCGUCCGAUG	1274

[0293] Sequences from **FIG. 15** can be found below:

ZFoff-15 Target Sequence	GATGGGGCTCTGGTGGCG (SEQ ID NO: 676)
ZFoff-15 ZFP Amino Acid Sequence	SRPGERPFQCRICMRNFSRTDTLARHLRTHTGEKPFQCRICMRNFSRKTALNRHLKTHTGSQ KPFQCRICMRNFSRNESLKVHLRTHTGEKPFQCRICMRNFSVKNTLTRHLKTHTGSQKPFQC RICMRNFSRREHLVRHLRTHTGEKPFQCRICMRNFSLTHNLRRHLKTHLRGS (SEQ ID NO: 677)
ZFoff-15 Full Amino Acid Sequence	MGTMYNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCED SITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRL FFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRAR YFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNE KEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGN SNANSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLES GSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYAL PRQESQRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSK HAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGPSSGAPPP SGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAP GTSTEPSEPKKKRKVSRPGERPFQCRICMRNFSRTDTLARHLRTHTGEKPFQCRICMRNFSR KTALNRHLKTHTGSQKPFQCRICMRNFSRNESLKVHLRTHTGEKPFQCRICMRNFSVKNTLT RHLKTHTGSQKPFQCRICMRNFSRREHLVRHLRTHTGEKPFQCRICMRNFSLTHNLRRHLKT HLRGSPKKKRKVSGSETPGTSESATPESTGRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNV MLENYKNLVSLGYQLTKPDVILRLEKGEEPSADYKDDDDKA (SEQ ID NO: 678)
ZFoff-15 with 2xNLS	MGTMPKKKRKVPKKKRKVYNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLG IQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIV NPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPV MIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSI KQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLF APLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRQPVRVLSLFR NIDKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPG WYMFQFHRILQYALPRQESQRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQN AMRVWSNIPGLKSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQN SLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSESRPGERPFQCRICMRNFSRTDTLARHLRTHTGEKPFQCRI CMRNFSRKTALNRHLKTHTGSQKPFQCRICMRNFSRNESLKVHLRTHTGEKPFQCRICMRNF SVKNTLTRHLKTHTGSQKPFQCRICMRNFSRREHLVRHLRTHTGEKPFQCRICMRNFSLTHN LRRHLKTHLRGSSGSETPGTSESATPESTGRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNV MLENYKNLVSLGYQLTKPDVILRLEKGEEPSADYKDDDDKAPKKKRKVPKKKRKV (SEQ ID NO: 679)
ZFoff-60 Target Sequence	GCAGAGGCCGGAGGGGGTC (SEQ ID NO: 740)
ZFoff-60 ZFP Amino Acid Sequence	SRPGERPFQCRICMRNFSDPSVLKRHLRTHTGEKPFQCRICMRNFSRTEHLARHLKTHTGGG GSQKPFQCRICMRNFSQSPHLKRHLRTHTGEKPFQCRICMRNFSDQTTLRRHLKTHTGSQKP

	FQCRICMRNFSKHSNLTRHTRTHTGEKPFQCRICMRNFSQMETLKRHLRTHLRGS (SEQ ID NO: 680)
ZFoff-60 Full Amino Acid Sequence	MGTMYNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCED SITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRL FFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRAR YFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNE KEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGN SNANSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLES GSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYAL PRQESQRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSK HAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGPSSGAPPP SGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAP GTSTEPSEPKKKRKVSRPGERPFQCRICMRNFSDPSVLKRHLRTHTGEKPFQCRICMRNFSR TEHLARHLKTHTGGGGSQKPFQCRICMRNFSQSPHLKRHLRTHTGEKPFQCRICMRNFSDQT TLRRHLKTHTGSQKPFQCRICMRNFSKHSNLTRHTRTHTGEKPFQCRICMRNFSQMETLKRH LRTHLRGSPKKKRKVSGSETPGTSESATPESTGRTLVTFKDVFVDFTREEWKLLDTAQQIVY RNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPSADYKDDDDKA (SEQ ID NO: 681)
ZFoff-60 with 2xNLS	MGTMPKKKRKVPKKKRKVYNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLG IQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIV NPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPV MIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSI KQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLF APLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRQPVRVLSLFR NIDKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPG WYMFQFHRILQYALPRQESQRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQN AMRVWSNIPGLKSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQN SLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSESRPGERPFQCRICMRNFSDPSVLKRHLRTHTGEKPFQCRI CMRNFSRTEHLARHLKTHTGGGGSQKPFQCRICMRNFSQSPHLKRHLRTHTGEKPFQCRICM RNFSDQTTLRRHLKTHTGSQKPFQCRICMRNFSKHSNLTRHTRTHTGEKPFQCRICMRNFSQ METLKRHLRTHLRGSSGSETPGTSESATPESTGRTLVTFKDVFVDFTREEWKLLDTAQQIVY RNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPSADYKDDDDKAPKKKRKVPKKKRKV (SEQ ID NO: 682)
ZFoff-137 Target Sequence	GGTGGGGAGGACTGTGCA (SEQ ID NO: 683)
ZFoff-137 ZFP Amino Acid Sequence	SRPGERPFQCRICMRNFSQSTTLKRHLRTHTGEKPFQCRICMRNFSGAHGLAGHLKTHTGSQ KPFQCRICMRNFSDQTNLRRHLRTHTGEKPFQCRICMRNFSRQDNLQRHLKTHTGSQKPFQC RICMRNFSKGDHLRRHTRTHTGEKPFQCRICMRNFSEAHHLSRHLRTHLRGS (SEQ ID NO: 684)
ZFoff-137 Full Amino Acid Sequence	MGTMYNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCED SITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRL FFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRAR YFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNE KEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGN SNANSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLES GSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYAL PRQESQRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSK HAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGPSSGAPPP SGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAP GTSTEPSEPKKKRKVSRPGERPFQCRICMRNFSQSTTLKRHLRTHTGEKPFQCRICMRNFSG AHGLAGHLKTHTGSQKPFQCRICMRNFSDQTNLRRHLRTHTGEKPFQCRICMRNFSRQDNLQ RHLKTHTGSQKPFQCRICMRNFSKGDHLRRHTRTHTGEKPFQCRICMRNFSEAHHLSRHLRT HLRGSPKKKRKVSGSETPGTSESATPESTGRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNV MLENYKNLVSLGYQLTKPDVILRLEKGEEPSADYKDDDDKA (SEQ ID NO: 685)
ZFoff-137 with 2xNLS	MGTMPKKKRKVPKKKRKVYNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLG IQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIV NPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPV MIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSI KQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLF APLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRQPVRVLSLFR NIDKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPG WYMFQFHRILQYALPRQESQRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQN

	AMRVWSNIPGLKSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQN SLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSESRPGERPFQCRICMRNFSQSTTLKRHLRTHTGEKPFQCRI CMRNFSGAHGLAGHLKTHTGSQKPFQCRICMRNFSDQTNLRRHLRTHTGEKPFQCRICMRNF SRQDNLQRHLKTHTGSQKPFQCRICMRNFSKGDHLRRHTRTHTGEKPFQCRICMRNFSEAHH LSRHLRTHLRGSSGSETPGTSESATPESTGRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNV MLENYKNLVSLGYQLTKPDVILRLEKGEEPSADYKDDDDKAPKKKRKVPKKKRKV (SEQ ID NO: 686)
ZFoff-152 Target Sequence	GCAGGAGGACGAGGACGGC (SEQ ID NO: 687)
ZFoff-152 ZFP Amino Acid Sequence	SRPGERPFQCRICMRNFSTPSKLDRHTRTHTGEKPFQCRICMRNFSLAENLRRHLRTHTGSQ KPFQCRICMRNFSRQDNLGRHLRTHTGEKPFQCRICMRNFSDGGNLGRHLKTHTGGGGSQKP FQCRICMRNFSQSPHLKRHLRTHTGEKPFQCRICMRNFSQSTSLQRHLKTHLRGS (SEQ ID NO: 688)
ZFoff-152 Full Amino Acid Sequence	MGTMYNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCED SITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRL FFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRAR YFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNE KEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGN SNANSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLES GSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYAL PRQESQRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSK HAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGPSSGAPPP SGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAP GTSTEPSEPKKKRKVSRPGERPFQCRICMRNFSTPSKLDRHTRTHTGEKPFQCRICMRNFSL AENLRRHLRTHTGSQKPFQCRICMRNFSRQDNLGRHLRTHTGEKPFQCRICMRNFSDGGNLG RHLKTHTGGGGSQKPFQCRICMRNFSQSPHLKRHLRTHTGEKPFQCRICMRNFSQSTSLQRH LKTHLRGSPKKKRKVSGSETPGTSESATPESTGRTLVTFKDVFVDFTREEWKLLDTAQQIVY RNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPSADYKDDDDKA (SEQ ID NO: 689)
ZFoff-152 with 2xNLS	MGTMPKKKRKVPKKKRKVYNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLG IQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIV NPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPV MIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSI KQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLF APLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRQPVRVLSLFR NIDKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPG WYMFQFHRILQYALPRQESQRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQN AMRVWSNIPGLKSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQN SLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSESRPGERPFQCRICMRNFSTPSKLDRHTRTHTGEKPFQCRI CMRNFSLAENLRRHLRTHTGSQKPFQCRICMRNFSRQDNLGRHLRTHTGEKPFQCRICMRNF SDGGNLGRHLKTHTGGGGSQKPFQCRICMRNFSQSPHLKRHLRTHTGEKPFQCRICMRNFSQ STSLQRHLKTHLRGSSGSETPGTSESATPESTGRTLVTFKDVFVDFTREEWKLLDTAQQIVY RNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPSADYKDDDDKAPKKKRKVPKKKRKV (SEQ ID NO: 690)
ZFoff-153 Target Sequence	GCAGGAGGACGAGGACGGC (SEQ ID NO: 691)
ZFoff-153 ZFP Amino Acid Sequence	SRPGERPFQCRICMRNFSERAKLIRHLRTHTGEKPFQCRICMRNFSDPSNLRRHLKTHTGSQ KPFQCRICMRNFSRQDNLGRHLRTHTGEKPFQCRICMRNFSDQGNLGRHLKTHTGGGGSQKP FQCRICMRNFSQSAHLKRHLRTHTGEKPFQCRICMRNFSQDVSLVRHLKTHLRGS (SEQ ID NO: 692)
ZFoff-153 Full Amino Acid Sequence	MGTMYNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCED SITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRL FFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRAR YFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNE KEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGN SNANSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLES GSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYAL PRQESQRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSK HAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGPSSGAPPP SGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAP GTSTEPSEPKKKRKVSRPGERPFQCRICMRNFSERAKLIRHLRTHTGEKPFQCRICMRNFSD

	PSNLRRHLKTHTGSQKPFQCRICMRNFSRQDNLGRHLRTHTGEKPFQCRICMRNFSDQGNLG RHLKTHTGGGGSQKPFQCRICMRNFSQSAHLKRHLRTHTGEKPFQCRICMRNFSQDVSLVRH LKTHLRGSPKKKRKVSGSETPGTSESATPESTGRTLVTFKDVFVDFTREEWKLLDTAQQIVY RNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPSADYKDDDDKA (SEQ ID NO:693)
ZFoff-153 with 2xNLS	MGTMPKKKRKVPKKKRKVYNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLG IQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIV NPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPV MIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSI KQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLF APLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRQPVRVLSLFR NIDKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPG WYMFQFHRILQYALPRQESQRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQN AMRVWSNIPGLKSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQN SLPLGGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSESRPGERPFQCRICMRNFSERAKLIRHLRTHTGEKPFQCRI CMRNFSDPSNLRRHLKTHTGSQKPFQCRICMRNFSRQDNLGRHLRTHTGEKPFQCRICMRNF SDQGNLGRHLKTHTGGGGSQKPFQCRICMRNFSQSAHLKRHLRTHTGEKPFQCRICMRNFSQ DVSLVRHLKTHLRGSSGSETPGTSESATPESTGRTLVTFKDVFVDFTREEWKLLDTAQQIVY RNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPSADYKDDDDKAPKKKRKVPKKKRKV (SEQ ID NO: 694)

Cell Culture and Transfection

[0294] HeLa (ATCC-CRM-CCL-2), Hepa1-6 (PCSK9-IRES-TdTomato), Huh7 (Sekisui XenoTech, LLC) and HEK293T Griptite (CLTA-GFP) cells were cultured in DMEM with 10% FBS. All experiments in HeLa and Huh7 cells were done using chemically synthesized guide RNA and *in vitro* transcribed effector construct. HeLa cells were reverse transfected using TransIT-X2 transfection reagent from Mirus (Cat# MIR6003). Huh7 cells were reverse transfected using MessengerMAX reagent from Invitrogen (Cat# LMRNA003). Secreted PCSK9 levels were measured at the indicated time points using LEGEND MAX™ Human PCSK9 ELISA Kit from Biolegend (Cat# 443107). All ELISA data was normalized for cell numbers using CellTiter-Glo kit from Promega (Cat# G7571).

[0295] HEK293T Griptite cells with GFP knocked into the CLTA locus as an in-frame CLTA fusion were co-transfected with plasmids encoding effector construct and human CLTA guide RNA using TransIT-X2 transfection reagent from Mirus (Cat# MIR6003). GFP was measured by FACS for GFP expression as a surrogate for CLTA expression.

[0296] Hepa1-6 cells were co-transfected with plasmids encoding effector construct and mouse *PCSK9* guide RNA using SF Cell Line 96-well Nucleofector Kit (Cat # V4SC-2096, program code: CM-138) in Amaxa 4D nucleofector device from Lonza. At the indicated timepoint, cells were FACS analyzed for TdTomato expression as a surrogate for PCSK9 levels.

In Vitro *Transcription of Effector Constructs and Synthetic gRNA*

[0297] 1 μg of linearized effector template was used to set up in-vitro transcription reactions using T7 mScript™ Standard mRNA Production System from CellScript (Cat# C-MSC100625) according to manufacturer's instructions to obtain RNA that had a Cap 1 structure on the 5' end and was 3'polyadenylated. End-modified sgRNA that had three 2'O-methyl modified nucleotides with phosphorothioate linkages on both 5' and 3' ends were obtained from Integrated DNA Technologies.

Methylation Profiling

[0298] Genomic DNA was extracted from each well of a 96-well culture plate using a DNAdvance DNA Extraction from Tissue Kit (Beckman Coulter). After quantification of genomic DNA via High-Sensitivity DNA 1X kit (Quant-IT), each genomic DNA sample was bisulfite converted using an EZ-96 DNA Methylation-Gold MagPrep kit (Zymo Research) according to manufacturer's instructions. For hybridization capture experiments, DNA libraries were prepared using the xGen™ Methyl-Seq DNA Library Prep Kit (IDT) and hybrid capture was conducted using the xGen™ Hybridization Capture of DNA libraries kit (IDT). For amplicon sequencing experiments, DNA libraries were prepared using the xGen™ Methyl-Seq DNA Library Prep Kit (IDT) and hybrid capture was conducted using the xGen™ Hybridization Capture of DNA libraries (IDT). Bisulfite-converted DNA from each sample was used to seed PCR corresponding to each of the two VIM amplicons using a Platinum Taq kit (Invitrogen). Pooled products were cleaned using the AMPure XP kit (Beckman Coulter) and fragment size assessment via D1000 screentape on a Tapestation 4200 (Agilent) prior to sequencing by commercial service (Azenta).

Example 13: Bacterial DNA Methyltransferases

[0299] In this experiment, a panel of bacterial proteins were screened for DNA methyltransferase activity in mammalian cells. These bacterial DNA methyltransferases (**Table 12**) were tested for epigenetic silencing activity by fusing them N-terminally to a dCas9 domain using the experimental procedure of Example 1. These constructs were then transfected in a reporter cell line that expresses GFP under the control of the mammalian promoter of CTLA4.

Table 12: Bacterial DNA methyltransferases

Description	Sequence
M.AluI methyltransferase	MSKANAKYSFVDLFAGIGGFHAALAATGGVCEYAVEIDREAAAVYERNWNKPALGDITDDAN DEGVTLRGYDGPIDVLTGGFPCQPFSKSGAQHGMAETRGTLFWNIARIIEEREPTVLILENV RNLVGPRHRHEWLTIIETLRFFGYEVSGAPAIFSPHLLPAWMGGTPQVRERVFITATLVPER MRDERSTIRRATGRPLEGFPYWADSWTDFRELSRLVVIRGFQAPEREVVGDRKRYVARTDMP EGFVPASVTRPAIDETLPAWKQSHLRRNYDFFERHFAEVVAWAYRWGVYTDLFPASRRKLEW QAQDAPRLWDTVMHFRPSGIRAKRPTYLPALVAITQTSIVGPLERRLSPRETARLQGLPEWF DFGEQRAAATYKQMGNGVNVGVVRHILREHVRRDRALLKLTPAGQRIINAVLADEPDATVGA LGAAE (SEQ ID NO: 675)
M.MspI methyltransferase	MKPEILKLIRSKLDLTQKQASEIIEVSDKTWQQWESGKTEMHPAYYSFLQEKLKDKINFEEL SAQKTLQKKIFDKYNQNQITKNAEELAEITHIEERKDAYSSDFKFIDLFSGIGGIRQSFEVN GGKCVFSSEIDPFAKFTYYTNFGVVPFGDITKVEATTIPQHDILCAGFPCQPFSHIGKREGF EHPTQGTMFHEIVRIIETKKTPVLFLENVPGLINHDDGNTLKVIIETLEDMGYKVHHTVLDA SHFGIPQKRKRFYLVAFLNQNIHFEFPKPPMISKDIGEVLESDVTGYSISEHLQKSYLFKKD DGKPSLIDKNTTGAVKTLVSTYHKIQRLTGTFVKDGETGIRLLTTNECKAIMGFPKDFVIPV SRTQMYRQMGNSVVVPVVTKIAEQISLALKTVNQQSPQENFELELV (SEQ ID NO: 610)
M.HaeIII methyltransferase	MNLISLFSGAGGLDLGFQKAGFRIICANEYDKSIWKTYESNHSAKLIKGDISKISSDEFPKC DGIIGGPPCQSWSEGGSLRGIDDPRGKLFYEYIRILKQKKPIFFLAENVKGMMAQRHNKAVQ EFIQEFDNAGYDVHIILLNANDYGVAQDRKRVFYIGFRKELNINYLPPIPHLIKPTFKDVIW DLKDNPIPALDKNKTNGNKCIYPNHEYFIGSYSTIFMSRNRVRQWNEPAFTVQASGRQCQLH PQAPVMLKVSKNLNKFVEGKEHLYRRLTVRECARVQGFPDDFIFHYESLNDGYKMIGNAVPV NLAYEIAKTIKSALEICKGN (SEQ ID NO: 608)
M. HhaI methyltransferase	MIEIKDKQLTGLRFIDLFAGLGGFRLALESCGAECVYSNEWDKYAQEVYEMNFGEKPEGDIT QVNEKTIPDHDILCAGFPCQAFSISGKQKGFEDSRGTLFFDIARIVREKKPKVVFMENVKNF ASHDNGNTLEVVKNTMNELDYSFHAKVLNALDYGIPQKRERIYMICFRNDLNIQNFQFPKPF ELNTFVKDLLLPDSEVEHLVIDRKDLVMTNQEIEQTTPKTVRLGIVGKGGQGERIYSTRGIA ITLSAYGGGIFAKTGGYLVNGKTRKLHPRECARVMGYPDSYKVHPSTSQAYKQFGNSVVINV LQYIAYNIGSSLNFKPY(SEQ ID NO: 609)
M. SssI methyltransferase	MSKVENKTKKLRVFEAFAGIGAQRKALEKVRKDEYEIVGLAEWYVPAIVMYQAIHNNFHTKL EYKSVSREEMIDYLENKTLSWNSKNPVSNGYWKRKKDDELKIIYNAIKLSEKEGNIFDIRDL YKRTLKNIDLLTYSFPCQDLSQQGIQKGMKRGSGTRSGLLWEIERALDSTEKNDLPKYLLME NVGALLHKKNEEELNQWKQKLESLGYQNSIEVLNAADFGSSQARRRVFMISTLNEFVELPKG DKKPKSIKKVLNKIVSEKDILNNLLKYNLTEFKKTKSNINKASLIGYSKFNSEGYVYDPEFT GPTLTASGANSRIKIKDGSNIRKMNSDETFLYIGFDSQDGKRVNEIEFLTENQKIFVCGNSI SVEVLEAIIDKIGG (SEQ ID NO: 605)

[0300] *M.SssI* DNA methyltransferase was able to efficiently methylate DNA in mammalian cells (**FIG. 16**) with stable silencing up to 30 days. Sequences from **FIG. 16** can be found below:

Description	Sequence	SEQ ID NO:
TAR087 Target DNA Sequence	GGCGGAAGTGCTGCCTGATA	667
TAR089 Target DNA Sequence	GCAACACCGCCTAGACCGAC	668
TAR090 Target DNA Sequence	GGAGGTTCAGAAAGCCTAAG	669
TAR091 Target DNA Sequence	GCAGCACTAGAGTCCCCTCA	670
TAR092 Target DNA Sequence	GAACAACTCGTGACTGGGGT	671
TAR093 Target DNA Sequence	GGGAGAGGAGGACAAGCGCC	672
TAR094 Target DNA Sequence	GGACCCACCGACACCACGCC	673
TAR114 Target DNA Sequence	CGCGGGAAAATGCAAGACGA	674

[0301] The methylation profile of these cells was also analyzed at day 29 confirming a 20% methylation of the target gene (**FIGs. 17-18**).

[0302] Another three orthologous DNA methyltransferases, predicted to be closely related to *M. SssI*, are identified and tested for epigenetic silencing activity using the experimental procedures of Example 1 (**Table 13**).

Table 13: Bacterial methyltransferases

Description	SEQ ID NO
DNA cytosine methyltransferase *Mycoplasmatales bacterium*	601
DNA cytosine methyltransferase *Mycoplasma marinum*	602
DNA (cytosine-5-)-methyltransferase *Spiroplasma chinense*	603

[0303] The DNA methyltransferases of **Table 13** are predicted to have similar or improved function to *M. SssI*. Sequences are tested in the context of CRISPR-off, in place of murine DNMT3A/DNMT3L, and their function is compared with the function of *M. SssI* DNA methyltransferase in silencing the *PCSK9* locus in a HeLa TdTomato system, to identify novel characteristics and improved function.

Example 14: Alternative KRAB Domains

[0304] In this example, fusion proteins were constructed with alternative KRAB domains (**Table 14**) and showed improved activity as compared to CRISPR-off when tested using the experimental procedures of Example 1 (**FIGs. 19A-19D**).

Table 14: Alternative KRAB Domains

Description	Sequence
ZFP28	MKKLEAVGTGIEPKAMSQGLVTFGDVAVDFSQEEWEWLNPIQRNLYRKVMLENYRNLASLGL CVSKPDVISSLEQGKEPWTVKRKMTRAWCPDLKAVWKIKELPLKKDFCEG (SEQ ID NO: 74)
ZN627	MDSVAFEDVAVNFTLEEWALLDPSQKNLYRDVMRETFRNLASVGKQWEDQNIEDPFKIPRRN ISHIPERLCESKEGGQGEETFSQIPDGILNKKTPGVKPCESSVCGEVGMGPSSLNRHIRDHT GREPNEYQEYGKKSYTRNQCGRALSYHRSFPVRERTHPGGKPYDCKECGETFISLVSIRRHM LTHRGGVPYKCKVCGKAFDYPSLFRIHERSHTGEKPYECKQCGKAFSCSSYIRIHERTHTGD KPYECKQCGKAFSCSKYIRIHERTHTGEKPYECKQCGKAFRCASSVRSHERTHTGEKLFECK ECGKALTCLASVRRHMIKHTGNGPYKCKVCGKAFDFPSSFRIHERTHTGEKPYDCKQCGKAF SCSSSFRKHERIHTGEKPYKCTKCGKAFSRSSYFRIHERTHTGEKPYECKQCGKAFSRSTYF RVHEKIHTGEKPYENPNPNASVVPVLS (SEQ ID NO: 666)
KAP1	MAASAAAASAAAASAASGSPGPGEGSAGGEKRSTAPSAAASASASASAAASSPAGGGAEALELL EHCGVCRERLRPEREPRLLPCLHSACSACLGPAAPAAANSSGDGGAAGDGTVVDCPVCKQQC FSKDIVENYFMRDSGSKAATDAQDANQCCTSCEDNAPATSYCVECSEPLCETCVEAHQRVKY TKDHTVRSTGPAKSRDGERTVYCNVHKHEPLVLFCESCDTLTCRDCQLNAHKDHQYQFLEDA VRNQRKLLASLVKRLGDKHATLQKSTKEVRSSIRQVSDVQKRVQVDVKMAILQIMKELNKRG RVLVNDAQKVTEGQQERLERQHWTMTKIQKHQEHILRFASWALESDNNTALLLSKKLIYFQL HRALKMIVDPVEPHGEMKFQWDLNAWTKSAEAFGKIVAERPGTNSTGPAPMAPPRAPGPLSK QGSGSSQPMEVQEGYGFGSGDDPYSSAEPHVSGVKRSRSGEGEVSGLMRKVPRVSLERLDLD LTADSQPPVFKVFPGSTTEDYNLIVIERGAAAAATGQPGTAPAGTPGAPPLAGMAIVKEEET EAAIGAPPTATEGPETKPVLMALAEGPGAEGPRLASPSGSTSSGLEVVAPEGTSAPGGGPGT LDDSATICRVCQKPGDLVMCNQCEFCFHLDCHLPALQDVPGEEWSCSLCHVLPDLKEEDGSL SLDGADSTGVVAKLSPANQRKCERVLLALFCHEPCRPLHQLATDSTFSLDQPGGTLDLTLIR

	ARLQEKLSPPYSSPQEFAQDVGRMFKQFNKLTEDKADVQSIIGLQRFFETRMNEAFGDTKFS AVLVEPPPMSLPGAGLSSQELSGGPGDGP (SEQ ID NO: 629)
MeCP2	MVAGMLGLREEKSEDQDLQGLKDKPLKFKKVKKDKKEEKEGKHEPVQPSAHHSAEPAEAGKA ETSEGSGSAPAVPEASASPKQRRSIIRDRGPMYDDPTLPEGWTRKLKQRKSGRSAGKYDVYL INPQGKAFRSKVELIAYFEKVGDTSLDPNDFDFTVTGRGSPSRREQKPPKKPKSPKAPGTGR GRGRPKGSGTTRPKAATSEGVQVKRVLEKSPGKLLVKMPFQTSPGGKAEGGGATTSTQVMVI KRPGRKRKAEADPQAIPKKRGRKPGSVVAAAAAEAKKKAVKESSIRSVQETVLPIKKRKTRE TVSIEVKEVVKPLLVSTLGEKSGKGLKTCKSPGRKSKESSPKGRSSSASSPPKKEHHHHHHH SESPKAPVPLLPPLPPPPPEPESSEDPTSPPEPQDLSSSVCKEEKMPRGGSLESDGCPKEPA KTQPAVATAATAAEKYKHRGEGERKDIVSSSMPRPNREEPVDSRTPVTERVS (SEQ ID NO: 630)
HP1b	MGKKQNKKKVEEVLEEEEEEYVVEKVLDRRVVKGKVEYLLKWKGFSDEDNTWEPEENLDCPD LIAEFLQSQKTAHETDKSEGGKRKADSDSEDKGEESKPKKKKEESEKPRGFARGLEPERIIG ATDSSGELMFLMKWKNSDEADLVPAKEANVKCPQVVISFYEERLTWHSYPSEDDDKKDDKN (SEQ ID NO: 94)
CBX8	GSGPPSSGGGLYRDMGAQGGRPSLIARIPVARILGDPEEESWSPSLTNLEKVVVTDVTSNFL TVTIKESNTDQGFFKEKR (SEQ ID NO: 387)
CDYL2	ASGDLYEVERIVDKRKNKKGKWEYLIRWKGYGSTEDTWEPEHHLLHCEEFIDEFNGLHMSKD KRIKSGKQSSTSKLLRDS (SEQ ID NO: 365)
TOX	KDPNEPQKPVSAYALFFRDTQAAIKGQNPNATFGEVSKIVASMWDGLGEEQKQVYKKKTEAA KKEYLKQLAAYRASLVSK (SEQ ID NO: 372)
TOX3	KDPNEPQKPVSAYALFFRDTQAAIKGQNPNATFGEVSKIVASMWDSLGEEQKQVYKRKTEAA KKEYLKALAAYRASLVSK (SEQ ID NO: 394)
TOX4	KDPNEPQKPVSAYALFFRDTQAAIKGQNPNATFGEVSKIVASMWDSLGEEQKQVYKRKTEAA KKEYLKALAAYKDNQECQ (SEQ ID NO: 1529)
EED	MSEREVSTAPAGTDMPAAKKQKLSSDENSNPDLSGDENDDAVSIESGTNTERPDTPTNTPNA PGRKSWGKGKWKSKKCKYSFKCVNSLKEDHNQPLFGVQFNWHSKEGDPLVFATVGSNRVTLY ECHSQGEIRLLQSYVDADADENFYTCAWTYDSNTSHPLLAVAGSRGIIRIINPITMQCIKHY VGHGNAINELKFHPRDPNLLLSVSKDHALRLWNIQTDTLVAIFGGVEGHRDEVLSADYDLLG EKIMSCGMDHSLKLWRINSKRMMNAIKESYDYNPNKTNRPFISQKIHFPDFSTRDIHRNYVD CVRWLGDLILSKSCENAIVCWKPGKMEDDIDKIKPSESNVTILGRFDYSQCDIWYMRFSMDF WQKMLALGNQVGKLYVWDLEVEDPHKAKCTTLTHHKCGAAIRQTSFSRDSSILIAVCDDASI WRWDRLR (SEQ ID NO: 563)
RBBP4	VWDLSKIGEEQSPEDAEDGPPELLFIHGGHTAKISDFSWNPNEPWVICSVSEDNIMQVWQMA ENIYNDEDPEGSVDPEGQ (SEQ ID NO: 1530)
RCOR1	MPAMVEKGPEVSGKRRGRNNAAASASAAAASAAASAACASPAATAASGAAASSASAAAASAA AAPNNGQNKSLAAAAPNGNSSSNSWEEGSSGSSSDEEHGGGGMRVGPQYQAVVPDFDPAKLA RRSQERDNLGMLVWSPNQNLSEAKLDEYIAIAKEKHGYNMEQALGMLFWHKHNIEKSLADLP NFTPFPDEWTVEDKVLFEQAFSFHGKTFHRIQQMLPDKSIASLVKFYYSWKKTRTKTSVMDR HARKQKREREESEDELEEANGNNPIDIEVDQNKESKKEVPPTETVPQVKKEKHSTQAKNRAK RKPPKGMFLSQEDVEAVSANATAATTVLRQLDMELVSVKRQIQNIKQTNSALKEKLDGGIEP YRLPEVIQKCNARWTTEEQLLAVQAIRKYGRDFQAISDVIGNKSVVQVKNFFVNYRRRFNID EVLQEWEAEHGKEETNGPSNQKPVKSPDNSIKMPEEEDEAPVLDVRYASAS (SEQ ID NO: 559)
SCML2	KQGFSKDPSTWSVDEVIQFMKHTDPQISGPLADLFRQHEIDGKALFLLKSDVMMKYMGLKLG PALKLCYYIEKLKEGKYS (SEQ ID NO: 382)

Example 15: ZIM 3 Fusion Constructs

[0305] Novel fusions of ZIM3 and KOX1KRAB are generated. Both ZIM 3 and KOX1KRAB are KRAB family proteins with extensive homology. Thus, sequences are designed which represent halfway points between ZIM3 and KOX1KRAB. These KOX1KRAB and ZIM3 constructs encode a small region of KOX1KRAB and ZIM3 focused around the zinc finger domain of the protein. While the regions used of KOX1KRAB and ZIM3 are very similar within the first ~75bp of their sequence, ZIM3 also possesses a small alpha-helical region at the C-terminus, not present in KOX1KRAB. The KOX1KRAB-FL

sequence includes the KOX1KRAB sequence equivalent of this extra piece, while the ZIM3 truncation has this extra piece removed from the ZIM3 sequence. The ZIM3/KOX1KRAB chimeras are fusions of the N- and C-terminal pieces of the two proteins. The ZIM3-like KOX1KRAB variants were both assembled by first, BLAST of ZIM3 or KOX1KRAB proteins from nonhuman species to assemble the closest 100 homologs ('families') of each gene; second, identifying the 3 members of the KOX1KRAB family that most closely resemble ZIM3 and the 3 members of the ZIM3 family that most closely resemble KOX1KRAB; and third, rationally modifying the KOX1KRAB-FL sequence to resemble each set of three (**Table 15**).

Table 15: ZIM-KOX1KRAB Chimera Proteins

Description	Sequence
KOX1KRAB Residues 11-72	58
KOX1KRAB-FL Residues 11-108	59
ZIM3 Residues 1-100	60
KOX1KRAB-ZIM3 chimera 1 (Residues 11-72 of KOX1KRAB, Residues 68-100 of ZIM3)	61
ZIM3-KOX1KRAB chimera 2 (Residues 1-67 of ZIM3, Residues 73-108 of KOX1KRAB)	62
Truncated ZIM3 (Residues 1-75 of ZIM3)	63
KOX1KRAB-FL Zim3-like (Modified Residues 11-108)	64
KOX1KRAB-like ZIM3 1	65
KOX1KRAB-like ZIM3 2	66
KOX1KRAB-like ZIM3 3	67
KOX1KRAB-like ZIM3 4	68
ZIM-like KOX1KRAB 1	69
ZIM-like KOX1KRAB 2	70
ZIM-like KOX1KRAB 3	71

SEQUENCES

[0306] The SEQ ID NOs (SEQ) of nucleotide (nt) and amino acid (aa) sequences described in the present disclosure are listed below.

SEQ	Description	Sequence
1	*S. pyogenes* WT Cas9 Sequence (nt)	ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGC GTCGGATGGGCGGTGATCACTGATGAATATAAGGTTCCGTCTAAA AAGTTCAAGGTTCTGGGAAATACAGACCGCCACAGTATCAAAAAA AATCTTATAGGGGCTCTTTTATTTGACAGTGGAGAGACAGCGGAA GCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCGG AAGAATCGTATTTGTTATCTACAGGAGATTTTTTCAAATGAGATG GCGAAAGTAGATGATAGTTTCTTTCATCGACTTGAAGAGTCTTTT TTGGTGGAAGAAGACAAGAAGCATGAACGTCATCCTATTTTTGGA AATATAGTAGATGAAGTTGCTTATCATGAGAAATATCCAACTATC TATCATCTGCGAAAAAAATTGGTAGATTCTACTGATAAAGCGGAT TTGCGCTTAATCTATTTGGCCTTAGCGCATATGATTAAGTTTCGT GGTCATTTTTTGATTGAGGGAGATTTAAATCCTGATAATAGTGAT GTGGACAAACTATTTATCCAGTTGGTACAAACCTACAATCAATTA TTTGAAGAAAACCCTATTAACGCAAGTGGAGTAGATGCTAAAGCG ATTCTTTCTGCACGATTGAGTAAATCAAGACGATTAGAAAATCTC ATTGCTCAGCTCCCCGGTGAGAAGAAAAATGGCTTATTTGGGAAT CTCATTGCTTTGTCATTGGGTTTGACCCCTAATTTTAAATCAAAT TTTGATTTGGCAGAAGATGCTAAATTACAGCTTTCAAAAGATACT TACGATGATGATTTAGATAATTTATTGGCGCAAATTGGAGATCAA TATGCTGATTTGTTTTTGGCAGCTAAGAATTTATCAGATGCTATT TTACTTTCAGATATCCTAAGAGTAAATACTGAAATAACTAAGGCT CCCCTATCAGCTTCAATGATTAAACGCTACGATGAACATCATCAA GACTTGACTCTTTTAAAAGCTTTAGTTCGACAACAACTTCCAGAA AAGTATAAAGAAATCTTTTTTGATCAATCAAAAAACGGATATGCA GGTTATATTGATGGGGGAGCTAGCCAAGAAGAATTTTATAAATTT ATCAAACCAATTTTAGAAAAAATGGATGGTACTGAGGAATTATTG GTGAAACTAAATCGTGAAGATTTGCTGCGCAAGCAACGGACCTTT GACAACGGCTCTATTCCCCATCAAATTCACTTGGGTGAGCTGCAT GCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTAAAAGAC AATCGTGAGAAGATTGAAAAAATCTTGACTTTTCGAATTCCTTAT TATGTTGGTCCATTGGCGCGTGGCAATAGTCGTTTTGCATGGATG ACTCGGAAGTCTGAAGAAACAATTACCCCATGGAATTTTGAAGAA GTTGTCGATAAAGGTGCTTCAGCTCAATCATTTATTGAACGCATG ACAAACTTTGATAAAAATCTTCCAAATGAAAAAGTACTACCAAAA CATAGTTTGCTTTATGAGTATTTTACGGTTTATAACGAATTGACA AAGGTCAAATATGTTACTGAAGGAATGCGAAAACCAGCATTTCTT TCAGGTGAACAGAAGAAAGCCATTGTTGATTTACTCTTCAAAACA AATCGAAAAGTAACCGTTAAGCAATTAAAAGAAGATTATTTCAAA AAAATAGAATGTTTTGATAGTGTTGAAATTTCAGGAGTTGAAGAT AGATTTAATGCTTCATTAGGTACCTACCATGATTTGCTAAAAATT ATTAAAGATAAAGATTTTTTGGATAATGAAGAAAATGAAGATATC TTAGAGGATATTGTTTTAACATTGACCTTATTTGAAGATAGGGAG ATGATTGAGGAAAGACTTAAAACATATGCTCACCTCTTTGATGAT AAGGTGATGAAACAGCTTAAACGTCGCCGTTATACTGGTTGGGGA CGTTTGTCTCGAAAATTGATTAATGGTATTAGGGATAAGCAATCT GGCAAAACAATATTAGATTTTTTGAAATCAGATGGTTTTGCCAAT CGCAATTTTATGCAGCTGATCCATGATGATAGTTTGACATTTAAA GAAGACATTCAAAAAGCACAAGTGTCTGGACAAGGCGATAGTTTA CATGAACATATTGCAAATTTAGCTGGTAGCCCTGCTATTAAAAAA GGTATTTTACAGACTGTAAAAGTTGTTGATGAATTGGTCAAAGTA

SEQ	Description	Sequence
		ATGGGGCGGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGT GAAAATCAGACAACTCAAAAGGGCCAGAAAAATTCGCGAGAGCGT ATGAAACGAATCGAAGAAGGTATCAAAGAATTAGGAAGTCAGATT CTTAAAGAGCATCCTGTTGAAAATACTCAATTGCAAAATGAAAAG CTCTATCTCTATTATCTCCAAAATGGAAGAGACATGTATGTGGAC CAAGAATTAGATATTAATCGTTTAAGTGATTATGATGTCGATCAC ATTGTTCCACAAAGTTTCCTTAAAGACGATTCAATAGACAATAAG GTCTTAACGCGTTCTGATAAAAATCGTGGTAAATCGGATAACGTT CCAAGTGAAGAAGTAGTCAAAAAGATGAAAAACTATTGGAGACAA CTTCTAAACGCCAAGTTAATCACTCAACGTAAGTTTGATAATTTA ACGAAAGCTGAACGTGGAGGTTTGAGTGAACTTGATAAAGCTGGT TTTATCAAACGCCAATTGGTTGAAACTCGCCAAATCACTAAGCAT GTGGCACAAATTTTGGATAGTCGCATGAATACTAAATACGATGAA AATGATAAACTTATTCGAGAGGTTAAAGTGATTACCTTAAAATCT AAATTAGTTTCTGACTTCCGAAAAGATTTCCAATTCTATAAAGTA CGTGAGATTAACAATTACCATCATGCCCATGATGCGTATCTAAAT GCCGTCGTTGGAACTGCTTTGATTAAGAAATATCCAAAACTTGAA TCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTCGTAAA ATGATTGCTAAGTCTGAGCAAGAAATAGGCAAAGCAACCGCAAAA TATTTCTTTTACTCTAATATCATGAACTTCTTCAAAACAGAAATT ACACTTGCAAATGGAGAGATTCGCAAACGCCCTCTAATCGAAACT AATGGGGAAACTGGAGAAATTGTCTGGGATAAAGGGCGAGATTTT GCCACAGTGCGCAAAGTATTGTCCATGCCCCAAGTCAATATTGTC AAGAAAACAGAAGTACAGACAGGCGGATTCTCCAAGGAGTCAATT TTACCAAAAAGAAATTCGGACAAGCTTATTGCTCGTAAAAAAGAC TGGGATCCAAAAAAATATGGTGGTTTTGATAGTCCAACGGTAGCT TATTCAGTCCTAGTGGTTGCTAAGGTGGAAAAAGGGAAATCGAAG AAGTTAAAATCCGTTAAAGAGTTACTAGGGATCACAATTATGGAA AGAAGTTCCTTTGAAAAAAATCCGATTGACTTTTTAGAAGCTAAA GGATATAAGGAAGTTAAAAAAGACTTAATCATTAAACTACCTAAA TATAGTCTTTTTGAGTTAGAAAACGGTCGTAAACGGATGCTGGCT AGTGCCGGAGAATTACAAAAAGGAAATGAGCTGGCTCTGCCAAGC AAATATGTGAATTTTTTATATTTAGCTAGTCATTATGAAAAGTTG AAGGGTAGTCCAGAAGATAACGAACAAAAACAATTGTTTGTGGAG CAGCATAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGAA TTTTCTAAGCGTGTTATTTTAGCAGATGCCAATTTAGATAAAGTT CTTAGTGCATATAACAAACATAGAGACAAACCAATACGTGAACAA GCAGAAAATATTATTCATTTATTTACGTTGACGAATCTTGGAGCT CCCGCTGCTTTTAAATATTTTGATACAACAATTGATCGTAAACGA TATACGTCTACAAAAGAAGTTTTAGATGCCACTCTTATCCATCAA TCCATCACTGGTCTTTATGAAACACGCATTGATTTGAGTCAGCTA GGAGGTGACTGA
2	*S. pyogenes* WT Cas9 Sequence (aa)	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKK NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEM AKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTI YHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENL IAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDT YDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKA PLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPY YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERM TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFL SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED

SEQ	Description	Sequence
		RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRE MIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMAR ENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEK LYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNK VLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDE NDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLN AVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLA SAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQ AENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD
3	SaCas9	MKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNE GRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYE ARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTK EQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEA KQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKD IKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRD ENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRV TSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQS SEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILD ELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVV KRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQ KRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEA IPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGN RTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEER DINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKS INGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKEWKK LDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKD FKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYD KDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLY KYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNS RNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSK CYEEAKKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLN RIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDIL GNLYEVKSKKHPQIIKKG
4	*F. novicida* WT Cpf1	MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAK DYKKAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSD DDNLQKDFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQE SDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFK GFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKA PEAINYEQIKKDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFN NYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYSQQINDKT LKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTMQSFYEQIA AFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQ VFDDYSVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKY LSLETIKLALEEFNKHRDIDKQCRFEEILANFAAIPMIFDEIAQN KDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKL KIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNY

SEQ	Description	Sequence
		ITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYL GVMNKKNNKIFDDKAIKENKGEGYKKIVYKLLPGANKMLPKVFFS AKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKF IDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKLT FENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKA LFDERNLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKD NPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEI NLLLKEKANDVHILSIDRGERHLAYYTLVDGKGNIIKQDTFNIIG NDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVH EIAKLVIEYNAIVVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKL NYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAG FTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFE FSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKNHNWDTREVYP TKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTIL QMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAY HIGLKGLMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRNN
5	CasX	MEKRINKIRKKLSADNATKPVSRSGPMKTLLVRVMTDDLKKRLEK RRKKPEVMPQVISNNAANNLRMLLDDYTKMKEAILQVYWQEFKDD HVGLMCKFAQPASKKIDQNKLKPEMDEKGNLTTAGFACSQCGQPL FVYKLEQVSEKGKAYTNYFGRCNVAEHEKLILLAQLKPEKDSDEA VTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGPVG KALSDACMGTIASFLSKYQDIIIEHQKVVKGNQKRLESLRELAGK ENLEYPSVTLPPQPHTKEGVDAYNEVIARVRMWVNLNLWQKLKLS RDDAKPLLRLKGFPSFPVVERRENEVDWWNTINEVKKLIDAKRDM GRVFWSGVTAEKRNTILEGYNYLPNENDHKKREGSLENPKKPAKR QFGDLLLYLEKKYAGDWGKVFDEAWERIDKKIAGLTSHIEREEAR NAEDAQSKAVLTDWLRAKASFVLERLKEMDEKEFYACEIQLQKWY GDLRGNPFAVEAENRVVDISGFSIGSDGHSIQYRNLLAWKYLENG KREFYLLMNYGKKGRIRFTDGTDIKKSGKWQGLLYGGGKAKVIDL TFDPDDEQLIILPLAFGTRQGREFIWNDLLSLETGLIKLANGRVI EKTIYNKKIGRDEPALFVALTFERREVVDPSNIKPVNLIGVDRGE NIPAVIALTDPEGCPLPEFKDSSGGPTDILRIGEGYKEKQRAIQA AKEVEQRRAGGYSRKFASKSRNLADDMVRNSARDLFYHAVTHDAV LVFENLSRGFGRQGKRTFMTERQYTKMEDWLTAKLAYEGLTSKTY LSKTLAQYTSKTCSNCGFTITTADYDGMLVRLKKTSDGWATTLNN KELKAEGQITYYNRYKRQTVEKELSAELDRLSEESGNNDISKWTK GRRDEALFLLKKRFSHRPVQEQFVCLDCGHEVHADEQAALNIARS WLFLNSNSTEFKSYKSGKQPFVGAWQAFYKRRLKEVWKPNA
6	CasY	MRKKLFKGYILHNKRLVYTGKAAIRSIKYPLVAPNKTALNNLSEK IIYDYEHLFGPLNVASYARNSNRYSLVDFWIDSLRAGVIWQSKST SLIDLISKLEGSKSPSEKIFEQIDFELKNKLDKEQFKDIILLNTG IRSSSNVRSLRGRFLKCFKEEFRDTEEVIACVDKWSKDLIVEGKS ILVSKQFLYWEEEFGIKIFPHFKDNHDLPKLTFFVEPSLEFSPHL PLANCLERLKKFDISRESLLGLDNNFSAFSNYFNELFNLLSRGEI KKIVTAVLAVSKSWENEPELEKRLHFLSEKAKLLGYPKLTSSWAD YRMIIGGKIKSWHSNYTEQLIKVREDLKKHQIALDKLQEDLKKVV DSSLREQIEAQREALLPLLDTMLKEKDFSDDLELYRFILSDFKSL LNGSYQRYIQTEEERKEDRDVTKKYKDLYSNLRNIPRFFGESKKE QFNKFINKSLPTIDVGLKILEDIRNALETVSVRKPPSITEEYVTK QLEKLSRKYKINAFNSNRFKQITEQVLRKYNNGELPKISEVFYRY PRESHVAIRILPVKISNPRKDISYLLDKYQISPDWKNSNPGEVVD LIEIYKLTLGWLLSCNKDFSMDFSSYDLKLFPEAASLIKNFGSCL SGYYLSKMIFNCITSEIKGMITLYTRDKFVVRYVTQMIGSNQKFP LLCLVGEKQTKNFSRNWGVLIEEKGDLGEEKNQEKCLIFKDKTDF AKAKEVEIFKNNIWRIRTSKYQIQFLNRLFKKTKEWDLMNLVLSE

SEQ	Description	Sequence
		PSLVLEEEWGVSWDKDKLLPLLKKEKSCEERLYYSLPLNLVPATD YKEQSAEIEQRNTYLGLDVGEFGVAYAVVRIVRDRIELLSWGFLK DPALRKIRERVQDMKKKQVMAVFSSSSTAVARVREMAIHSLRNQI HSIALAYKAKIIYEISISNFETGGNRMAKIYRSIKVSDVYRESGA DTLVSEMIWGKKNKQMGNHISSYATSYTCCNCARTPFELVIDNDK EYEKGGDEFIFNVGDEKKVRGFLQKSLLGKTIKGKEVLKSIKEYA RPPIREVLLEGEDVEQLLKRRGNSYIYRCPFCGYKTDADIQAALN IACRGYISDNAKDAVKEGERKLDYILEVRKLWEKNGAVLRSAKFL
7	CasPhi	MADTPTLFTQFLRHHLPGQRFRKDILKQAGRILANKGEDATIAFL RGKSEESPPDFQPPVKCPIIACSRPLTEWPIYQASVAIQGYVYGQ SLAEFEASDPGCSKDGLLGWFDKTGVCTDYFSVQGLNLIFQNARK RYIGVQTKVTNRNEKRHKKLKRINAKRIAEGLPELTSDEPESALD ETGHLIDPPGLNTNIYCYQQVSPKPLALSEVNQLPTAYAGYSTSG DDPIQPMVTKDRLSISKGQPGYIPEHQRALLSQKKHRRMRGYGLK ARALLVIVRIQDDWAVIDLRSLLRNAYWRRIVQTKEPSTITKLLK LVTGDPVLDATRMVATFTYKPGIVQVRSAKCLKNKQGSKLFSERY LNETVSVTSIDLGSNNLVAVATYRLVNGNTPELLQRFTLPSHLVK DFERYKQAHDTLEDSIQKTAVASLPQGQQTEIRMWSMYGFREAQE RVCQELGLADGSIPWNVMTATSTILTDLFLARGGDPKKCMFTSEP KKKKNSKQVLYKIRDRAWAKMYRTLLSKETREAWNKALWGLKRGS PDYARLSKRKEELARRCVNYTISTAEKRAQCGRTIVALEDLNIGF FHGRGKQEPGWVGLFTRKKENRWLMQALHKAFLELAHHRGYHVIE VNPAYTSQTCPVCRHCDPDNRDQHNREAFHCIGCGFRGNADLDVA THNIAMVAITGESLKRARGSVASKTPQPLAAE
8	Cas12f1 (Cas14a)	MIKVYRYEIVKPLDLDWKEFGTILRQLQQETRFALNKATQLAWEW MGFSSDYKDNHGEYPKSKDILGYTNVHGYAYHTIKTKAYRLNSGN LSQTIKRATDRFKAYQKEILRGDMSIPSYKRDIPLDLIKENISVN RMNHGDYIASLSLLSNPAKQEMNVKRKISVIIIVRGAGKTIMDRI LSGEYQVSASQIIHDDRKNKWYLNISYDFEPQTRVLDLNKIMGID LGVAVAVYMAFQHTPARYKLEGGEIENFRRQVESRRISMLRQGKY AGGARGGHGRDKRIKPIEQLRDKIANFRDTTNHRYSRYIVDMAIK EGCGTIQMEDLTNIRDIGSRFLQNWTYYDLQQKIIYKAEEAGIKV IKIDPQYTSQRCSECGNIDSGNRIGQAIFKCRACGYEANADYNAA RNIAIPNIDKIIAESIKSGGS
9	Cas12f2 (Cas14b)	NAMIAQKTIKIKLNPTKEQIIKLNSIIEEYIKVSNFTAKKIAEIQ ESFTDSGLTQGTCSECGKEKTYRKYHLLKKDNKLFCITCYKRKYS QFTLQKVEFQNKTGLRNVAKLPKTYYTNAIRFASDTFSGFDEIIK KKQNRLNSIQNRLNFWKELLYNPSNRNEIKIKVVKYAPKTDTREH PHYYSEAEIKGRIKRLEKQLKKFKMPKYPEFTSETISLQRELYSW KNPDELKISSITDKNESMNYYGKEYLKRYIDLINSQTPQILLEKE NNSFYLCFPITKNIEMPKIDDTFEPVGIDWGITRNIAVVSILDSK TKKPKFVKFYSAGYILGKRKHYKSLRKHFGQKKRQDKINKLGTKE DRFIDSNIHKLAFLIVKEIRNHSNKPIILMENITDNREEAEKSMR QNILLHSVKSRLQNYIAYKALWNNIPTNLVKPEHTSQICNRCGHQ DRENRPKGSKLFKCVKCNYMSNADFNASINIARKFYIGEYEPFYK DNEKMKSGVNSISM
10	Cas12f3 (Cas14c)	MEVQKTVMKTLSLRILRPLYSQEIEKEIKEEEKERRKQAGGTGEL DGGFYKKLEKKHSEMFSFDRLNLLLNQLQREIAKVYNHAISELYI ATIAQGNKSNKHYISSIVYNRAYGYFYNAYIALGICSKVEANFRS NELLTQQSALPTAKSDNFPIVLHKQKGAEGEDGGFRISTEGSDLI FEIPIPFYEYNGENRKEPYKWVKKGGQKPVLKLILSTFRRQRNKG WAKDEGTDAEIRKVTEGKYQVSQIEINRGKKLGEHQKWFANFSIE QPIYERKPNRSIVGGLDVGIRSPLVCAINNSFSRYSVDSNDVFKF SKQVFAFRRRLLSKNSLKRKHGHAAHKLEPITEMTEKNDKFRKKI IERWAKEVTNFFVKNQVGIVQIEDLSTMKDREDHFFNQYLRGFWP

SEQ	Description	Sequence
		YYQMQTLIENKLKEYGIEVKRVQAKYTSQLCSNPNCRYWNNYFNF EYRKVNKFPKFKCEKCNLEISADYNAARNLSTPDIEKFVAKATKG INLPEK
11	C2c8	MKVLEFKIHPTEEQVSKIDQSLAACKLLWNLSIALKEESKQRYYR KKHKFDEFSPEIWGLSYSGHYDEKEFKTLKDKEKKLLIGNPCCKI AYFKKTSNGKEYTPLNSIPIRRFMNAENIDKDAVNYLNRKKLAFY FRENTAKFIGEIETEFKKGFFKSVIKPAYDAAKKGIRGIPRFKGR RDKVETLVNGQPETIKIKSNGVIVSSKIGLLKIRGLDRLQGKAPR MAKITRKATGYYLQLTIETDDTIYKESDKCVGLDMGAVAIFTDDL GRQSEAKRYAKIQKKRLNRLQRQASRQKDNSNNQRKTYAKLARVH EKIARQRKGRNAQLAHKITSEYQSVILEDLNLKNMTAAAKPKERE DGDGYKQNGKKRKSGLNKALLDNAIGQLRTFIENKANERGRKIIR VNPKHTSQTCPNCGNIDKANRVSQSKFKCVSCGYEAHADQNAAAN ILIRGLRDEFLRAIGSLYKFPVSMIGKYPGLAGEFTPDLDANQES IGDAPIENAEHSISKQMKQEGNRTPTQPENGSQSLIFLSAPPQPC GDSHGTNNPKALPNKASKRSSKKPRGAIPENPDQLTIWDLLD
12	dSpCas9	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKK NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEM AKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTI YHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENL IAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDT YDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKA PLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPY YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERM TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFL SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRE MIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMAR ENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEK LYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNK VLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDE NDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLN AVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLA SAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQ AENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD
13	dSaCas9	MKRNYILGLAIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNE GRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYE ARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTK EQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEA KQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKD IKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRD ENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRV TSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQS

SEQ	Description	Sequence
		SEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILD ELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVV KRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQ KRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEA IPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEEASKKGN RTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEER DINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKS INGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKEWKK LDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKD FKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYD KDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLY KYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNS RNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSK CYEEAKKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLN RIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDIL GNLYEVKSKKHPQIIKKG
14	inactive FnCpf1	MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAK DYKKAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSD DDNLQKDFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQE SDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFK GFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKA PEAINYEQIKKDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFN NYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYSQQINDKT LKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTMQSFYEQIA AFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQ VFDDYSVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKY LSLETIKLALEEFNKHRDIDKQCRFEEILANFAAIPMIFDEIAQN KDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKL KIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNY ITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYL GVMNKKNNKIFDDKAIKENKGEGYKKIVYKLLPGANKMLPKVFFS AKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKF IDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKLT FENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKA LFDERNLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKD NPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEI NLLLKEKANDVHILSIARGERHLAYYTLVDGKGNIIKQDTFNIIG NDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVH EIAKLVIEYNAIVVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKL NYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAG FTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFE FSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKNHNWDTREVYP TKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTIL QMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAY HIGLKGLMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRNN
15	dNmeCas9	MAAFKPNSINYILGLAIGIASVGWAMVEIDEEENPIRLIDLGVRV FERAEVPKTGDSLAMARRLARSVRRLTRRRAHRLLRTRRLLKREG VLQAANFDENGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLI KHRGYLSQRKNEGETADKELGALLKGVAGNAHALQTGDFRTPAEL ALNKFEKESGHIRNQRSDYSHTFSRKDLQAELILLFEKQKEFGNP HVSGGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKN TYTAERFIWLTKLNNLRILEQGSERPLTDTERATLMDEPYRKSKL TYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRAL EKEGLKDKKSPLNLSPELQDEIGTAFSLFKTDEDITGRLKDRIQP EILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEIYG

SEQ	Description	Sequence
		DHYGKKNTEEKIYLPPIPADEIRNPVVLRALSQARKVINGVVRRY GSPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREY FPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEINLGRLNEKGYVE IDAALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSR EWQEFKARVETSRFPRSKKQRILLQKFDEDGFKERNLNDTRYVNR FLCQFVADRMRLTGKGKKRVFASNGQITNLLRGFWGLRKVRAEND RHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGEV LHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTLEKLRTLL AEKLSSRPEAVHEYVTPLFVSRAPNRKMSGQGHMETVKSAKRLDE GVSVLRVPLTQLKLKDLEKMVNREREPKLYEALKARLEAHKDDPA KAFAEPFYKYDKAGNRTQQVKAVRVEQVQKTGVWVRNHNGIADNA TMVRVDVFEKGDKYYLVPIYSWQVAKGILPDRAVVQGKDEEDWQL IDDSFNFKFSLHPNDLVEVITKKARMFGYFASCHRGTGNINIRIH DLDHKIGKNGILEGIGVKTALSFQKYQIDELGKEIRPCRLKKRPP VR
16	dCjCas9	MARILAFAIGISSIGWAFSENDELKDCGVRIFTKVENPKTGESLA LPRRLARSARKRLARRKARLNHLKHLIANEFKLNYEDYQSFDESL AKAYKGSLISPYELRFRALNELLSKQDFARVILHIAKRRGYDDIK NSDDKEKGAILKAIKQNEEKLANYQSVGEYLYKEYFQKFKENSKE FTNVRNKKESYERCIAQSFLKDELKLIFKKQREFGFSFSKKFEEE VLSVAFYKRALKDFSHLVGNCSFFTDEKRAPKNSPLAFMFVALTR IINLLNNLKNTEGILYTKDDLNALLNEVLKNGTLTYKQTKKLLGL SDDYEFKGEKGTYFIEFKKYKEFIKALGEHNLSQDDLNEIAKDIT LIKDEIKLKKALAKYDLNQNQIDSLSKLEFKDHLNISFKALKLVT PLMLEGKKYDEACNELNLKVAINEDKKDFLPAFNETYYKDEVTNP VVLRAIKEYRKVLNALLKKYGKVHKINIELAREVGKNHSQRAKIE KEQNENYKAKKDAELECEKLGLKINSKNILKLRLFKEQKEFCAYS GEKIKISDLQDEKMLEIDAIYPYSRSFDDSYMNKVLVFTKQNQEK LNQTPFEAFGNDSAKWQKIEVLAKNLPTKKQKRILDKNYKDKEQK NFKDRNLNDTRYIARLVLNYTKDYLDFLPLSDDENTKLNDTQKGS KVHVEAKSGMLTSALRHTWGFSAKDRNNHLHHAIDAVIIAYANNS IVKAFSDFKKEQESNSAELYAKKISELDYKNKRKFFEPFSGFRQK VLDKIDEIFVSKPERKKPSGALHEETFRKEEEFYQSYGGKEGVLK ALELGKIRKVNGKIVKNGDMFRVDIFKHKKTNKFYAVPIYTMDFA LKVLPNKAVARSKKGEIKDWILMDENYEFCFSLYKDSLILIQTKD MQEPEFVYYNAFTSSTVSLIVSKHDNKFETLSKNQKILFKNANEK EVIAKSIGIQNLKVFEKYIVSALGEVTKAEFRQREDFKK
17	dSt1Cas9	MGSDLVLGLAIGIGSVGVGILNKVTGEIIHKNSRIFPAAQAENNL VRRTNRQGRRLARRKKHRRVRLNRLFEESGLITDFTKISINLNPY QLRVKGLTDELSNEELFIALKNMVKHRGISYLDDASDDGNSSVGD YAQIVKENSKQLETKTPGQIQLERYQTYGQLRGDFTVEKDGKKHR LINVFPTSAYRSEALRILQTQQEFNPQITDEFINRYLEILTGKRK YYHGPGNEKSRTDYGRYRTSGETLDNIFGILIGKCTFYPDEFRAA KASYTAQEFNLLNDLNNLTVPTETKKLSKEQKNQIINYVKNEKAM GPAKLFKYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTL ETLDIEQMDRETLDKLAYVLTLNTEREGIQEALEHEFADGSFSQK QVDELVQFRKANSSIFGKGWHNFSVKLMMELIPELYETSEEQMTI LTRLGKQKTTSSSNKTKYIDEKLLTEEIYNPVVAKSVRQAIKIVN AAIKEYGDFDNIVIEMARETNEDDEKKAIQKIQKANKDEKDAAML KAANQYNGKAELPHSVFHGHKQLATKIRLWHQQGERCLYTGKTIS IHDLINNSNQFEVDAILPLSITFDDSLANKVLVYATANQEKGQRT PYQALDSMDDAWSFRELKAFVRESKTLSNKKKEYLLTEEDISKFD VRKKFIERNLVDTRYASRVVLNALQEHFRAHKIDTKVSVVRGQFT SQLRRHWGIEKTRDTYHHHAVDALIIAASSQLNLWKKQKNTLVSY SEDQLLDIETGELISDDEYKESVFKAPYQHFVDTLKSKEFEDSIL

SEQ	Description	Sequence
		FSYQVDSKFNRKISDATIYATRQAKVGKDKADETYVLGKIKDIYT QDGYDAFMKIYKKDKSKFLMYRHDPQTFEKVIEPILENYPNKQIN EKGKEVPCNPFLKYKEEHGYIRKYSKKGNGPEIKSLKYYDSKLGN HIDITPKDSNNKVVLQSVSPWRADVYFNKTTGKYEILGLKYADLQ FEKGTGTYKISQEKYNDIKKKEGVDSDSEFKFTLYKNDLLLVKDT ETKEQQLFRFLSRTMPKQKHYVELKPYDKQKFEGGEALIKVLGNV ANSGQCKKGLGKSNISIYKVRTDVLGNQHIIKNEGDKPKLDF
18	dSt3Cas9	MTKPYSIGLAIGTNSVGWAVITDNYKVPSKKMKVLGNTSKKYIKK NLLGVLLFDSGITAEGRRLKRTARRRYTRRRNRILYLQEIFSTEM ATLDDAFFQRLDDSFLVPDDKRDSKYPIFGNLVEEKVYHDEFPTI YHLRKYLADSTKKADLRLVYLALAHMIKYRGHFLIEGEFNSKNND IQKNFQDFLDTYNAIFESDLSLENSKQLEEIVKDKISKLEKKDRI LKLFPGEKNSGIFSEFLKLIVGNQADFRKCFNLDEKASLHFSKES YDEDLETLLGYIGDDYSDVFLKAKKLYDAILLSGFLTVTDNETEA PLSSAMIKRYNEHKEDLALLKEYIRNISLKTYNEVFKDDTKNGYA GYIDGKTNQEDFYVYLKNLLAEFEGADYFLEKIDREDFLRKQRTF DNGSIPYQIHLQEMRAILDKQAKFYPFLAKNKERIEKILTFRIPY YVGPLARGNSDFAWSIRKRNEKITPWNFEDVIDKESSAEAFINRM TSFDLYLPEEKVLPKHSLLYETFNVYNELTKVRFIAESMRDYQFL DSKQKKDIVRLYFKDKRKVTDKDIIEYLHAIYGYDGIELKGIEKQ FNSSLSTYHDLLNIINDKEFLDDSSNEAIIEEIIHTLTIFEDREM IKQRLSKFENIFDKSVLKKLSRRHYTGWGKLSAKLINGIRDEKSG NTILDYLIDDGISNRNFMQLIHDDALSFKKKIQKAQIIGDEDKGN IKEVVKSLPGSPAIKKGILQSIKIVDELVKVMGGRKPESIVVEMA RENQYTNQGKSNSQQRLKRLEKSLKELGSKILKENIPAKLSKIDN NALQNDRLYLYYLQNGKDMYTGDDLDIDRLSNYDIDHIIPQAFLK DNSIDNKVLVSSASARGKSDDFPSLEVVKKRKTFWYQLLKSKLIS QRKFDNLTKAERGGLLPEDKAGFIQRQLVETRQITKHVARLLDEK FNNKKDENNRAVRTVKIITLKSTLVSQFRKDFELYKVREINDFHH AHDAYLNAVIASALLKKYPKLEPEFVYGDYPKYNSFRERKSATEK VYFYSNIMNIFKKSISLADGRVIERPLIEVNEETGESVWNKESDL ATVRRVLSYPQVNVVKKVEEQNHGLDRGKPKGLFNANLSSKPKPN SNENLVGAKEYLDPKKYGGYAGISNSFAVLVKGTIEKGAKKKITN VLEFQGISILDRINYRKDKLNFLLEKGYKDIELIIELPKYSLFEL SDGSRRMLASILSTNNKRGEIHKGNQIFLSQKFVKLLYHAKRISN TINENHRKYVENHKKEFEELFYYILEFNENYVGAKKNGKLLNSAF QSWQNHSIDELCSSFIGPTGSERKGLFELTSRGSAADFEFLGVKI PRYRDYTPSSLLKDATLIHQSVTGLYETRIDLAKLGEG
19	dLbCpf1	MSKLEKFTNCYSLSKTLRFKAIPVGKTQENIDNKRLLVEDEKRAE DYKGVKKLLDRYYLSFINDVLHSIKLKNLNNYISLFRKKTRTEKE NKELENLEINLRKEIAKAFKGNEGYKSLFKKDIIETILPEFLDDK DEIALVNSFNGFTTAFTGFFDNRENMFSEEAKSTSIAFRCINENL TRYISNMDIFEKVDAIFDKHEVQEIKEKILNSDYDVEDFFEGEFF NFVLTQEGIDVYNAIIGGFVTESGEKIKGLNEYINLYNQKTKQKL PKFKPLYKQVLSDRESLSFYGEGYTSDEEVLEVFRNTLNKNSEIF SSIKKLEKLFKNFDEYSSAGIFVKNGPAISTISKDIFGEWNVIRD KWNAEYDDIHLKKKAVVTEKYEDDRRKSFKKIGSFSLEQLQEYAD ADLSVVEKLKEIIIQKVDEIYKVYGSSEKLFDADFVLEKSLKKND AVVAIMKDLLDSVKSFENYIKAFFGEGKETNRDESFYGDFVLAYD ILLKVDHIYDAIRNYVTQKPYSKDKFKLYFQNPQFMGGWDKDKET DYRATILRYGSKYYLAIMDKKYAKCLQKIDKDDVNGNYEKINYKL LPGPNKMLPKVFFSKKWMAYYNPSEDIQKIYKNGTFKKGDMFNLN DCHKLIDFFKDSISRYPKWSNAYDFNFSETEKYKDIAGFYREVEE QGYKVSFESASKKEVDKLVEEGKLYMFQIYNKDFSDKSHGTPNLH TMYFKLLFDENNHGQIRLSGGAELFMRRASLKKEELVVHPANSPI

SEQ	Description	Sequence
		ANKNPDNPKKTTTLSYDVYKDKRFSEDQYELHIPIAINKCPKNIF KINTEVRVLLKHDDNPYVIGIARGERNLLYIVVVDGKGNIVEQYS LNEIINNFNGIRIKTDYHSLLDKKEKERFEARQNWTSIENIKELK AGYISQVVHKICELVEKYDAVIALEDLNSGFKNSRVKVEKQVYQK FEKMLIDKLNYMVDKKSNPCATGGALKGYQITNKFESFKSMSTQN GFIFYIPAWLTSKIDPSTGFVNLLKTKYTSIADSKKFISSFDRIM YVPEEDLFEFALDYKNFSRTDADYIKKWKLYSYGNRIRIFRNPKK NNVFDWEEVCLTSAYKELFNKYGINYQQGDIRALLCEQSDKAFYS SFMALMSLMLQMRNSITGRTDVDFLISPVKNSDGIFYDSRNYEAQ ENAILPKNADANGAYNIARKVLWAIGQFKKAEDEKLDKVKIAISN KEWLEYAQTSVKH
20	inactive AsCpf1	MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARND HYKELKPIIDRIYKTYADQCLQLVQLDWENLSAAIDSYRKEKTEE TRNALIEEQATYRNAIHDYFIGRTDNLTDAINKRHAEIYKGLFKA ELFNGKVLKQLGTVTTTEHENALLRSFDKFTTYFSGFYENRKNVF SAEDISTAIPHRIVQDNFPKFKENCHIFTRLITAVPSLREHFENV KKAIGIFVSTSIEEVFSFPFYNQLLTQTQIDLYNQLLGGISREAG TEKIKGLNEVLNLAIQKNDETAHIIASLPHRFIPLFKQILSDRNT LSFILEEFKSDEEVIQSFCKYKTLLRNENVLETAEALFNELNSID LTHIFISHKKLETISSALCDHWDTLRNALYERRISELTGKITKSA KEKVQRSLKHEDINLQEIISAAGKELSEAFKQKTSEILSHAHAAL DQPLPTTLKKQEEKEILKSQLDSLLGLYHLLDWFAVDESNEVDPE FSARLTGIKLEMEPSLSFYNKARNYATKKPYSVEKFKLNFQMPTL ASGWDVNKEKNNGAILFVKNGLYYLGIMPKQKGRYKALSFEPTEK TSEGFDKMYYDYFPDAAKMIPKCSTQLKAVTAHFQTHTTPILLSN NFIEPLEITKEIYDLNNPEKEPKKFQTAYAKKTGDQKGYREALCK WIDFTRDFLSKYTKTTSIDLSSLRPSSQYKDLGEYYAELNPLLYH ISFQRIAEKEIMDAVETGKLYLFQIYNKDFAKGHHGKPNLHTLYW TGLFSPENLAKTSIKLNGQAELFYRPKSRMKRMAHRLGEKMLNKK LKDQKTPIPDTLYQELYDYVNHRLSHDLSDEARALLPNVITKEVS HEIIKDRRFTSDKFFFHVPITLNYQAANSPSKFNQRVNAYLKEHP ETPIIGIARGERNLIYITVIDSTGKILEQRSLNTIQQFDYQKKLD NREKERVAARQAWSVVGTIKDLKQGYLSQVIHEIVDLMIHYQAVV VLENLNFGFKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYPAEK VGGVLNPYQLTDQFTSFAKMGTQSGFLFYVPAPYTSKIDPLTGFV DPFVWKTIKNHESRKHFLEGFDFLHYDVKTGDFILHFKMNRNLSF QRGLPGFMPAWDIVFEKNETQFDAKGTPFIAGKRIVPVIENHRFT GRYRDLYPANELIALLEEKGIVFRDGSNILPKLLENDDSHAIDTM VALIRSVLQMRNSNAATGEDYINSPVRDLNGVCFDSRFQNPEWPM DADANGAYHIALKGQLLLNHLKESKDLKLQNGISNQDWLAYIQEL RN
21	inactive enAsCpf1	MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARND HYKELKPIIDRIYKTYADQCLQLVQLDWENLSAAIDSYRKEKTEE TRNALIEEQATYRNAIHDYFIGRTDNLTDAINKRHAEIYKGLFKA ELFNGKVLKQLGTVTTTEHENALLRSFDKFTTYFSGFYRNRKNVF SAEDISTAIPHRIVQDNFPKFKENCHIFTRLITAVPSLREHFENV KKAIGIFVSTSIEEVFSFPFYNQLLTQTQIDLYNQLLGGISREAG TEKIKGLNEVLNLAIQKNDETAHIIASLPHRFIPLFKQILSDRNT LSFILEEFKSDEEVIQSFCKYKTLLRNENVLETAEALFNELNSID LTHIFISHKKLETISSALCDHWDTLRNALYERRISELTGKITKSA KEKVQRSLKHEDINLQEIISAAGKELSEAFKQKTSEILSHAHAAL DQPLPTTLKKQEEKEILKSQLDSLLGLYHLLDWFAVDESNEVDPE FSARLTGIKLEMEPSLSFYNKARNYATKKPYSVEKFKLNFQMPTL ARGWDVNREKNNGAILFVKNGLYYLGIMPKQKGRYKALSFEPTEK TSEGFDKMYYDYFPDAAKMIPKCSTQLKAVTAHFQTHTTPILLSN

SEQ	Description	Sequence
		NFIEPLEITKEIYDLNNPEKEPKKFQTAYAKKTGDQKGYREALCK WIDFTRDFLSKYTKTTSIDLSSLRPSSQYKDLGEYYAELNPLLYH ISFQRIAEKEIMDAVETGKLYLFQIYNKDFAKGHHGKPNLHTLYW TGLFSPENLAKTSIKLNGQAELFYRPKSRMKRMAHRLGEKMLNKK LKDQKTPIPDTLYQELYDYVNHRLSHDLSDEARALLPNVITKEVS HEIIKDRRFTSDKFFFHVPITLNYQAANSPSKFNQRVNAYLKEHP ETPIIGIARGERNLIYITVIDSTGKILEQRSLNTIQQFDYQKKLD NREKERVAARQAWSVVGTIKDLKQGYLSQVIHEIVDLMIHYQAVV VLENLNFGFKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYPAEK VGGVLNPYQLTDQFTSFAKMGTQSGFLFYVPAPYTSKIDPLTGFV DPFVWKTIKNHESRKHFLEGFDFLHYDVKTGDFILHFKMNRNLSF QRGLPGFMPAWDIVFEKNETQFDAKGTPFIAGKRIVPVIENHRFT GRYRDLYPANELIALLEEKGIVFRDGSNILPKLLENDDSHAIDTM VALIRSVLQMRNSNAATGEDYINSPVRDLNGVCFDSRFQNPEWPM DADANGAYHIALKGQLLLNHLKESKDLKLQNGISNQDWLAYIQEL RN
22	inactive HFAsCpf1	MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARND HYKELKPIIDRIYKTYADQCLQLVQLDWENLSAAIDSYRKEKTEE TRNALIEEQATYRNAIHDYFIGRTDNLTDAINKRHAEIYKGLFKA ELFNGKVLKQLGTVTTTEHENALLRSFDKFTTYFSGFYRNRKNVF SAEDISTAIPHRIVQDNFPKFKENCHIFTRLITAVPSLREHFENV KKAIGIFVSTSIEEVFSFPFYNQLLTQTQIDLYNQLLGGISREAG TEKIKGLNEVLALAIQKNDETAHIIASLPHRFIPLFKQILSDRNT LSFILEEFKSDEEVIQSFCKYKTLLRNENVLETAEALFNELNSID LTHIFISHKKLETISSALCDHWDTLRNALYERRISELTGKITKSA KEKVQRSLKHEDINLQEIISAAGKELSEAFKQKTSEILSHAHAAL DQPLPTTLKKQEEKEILKSQLDSLLGLYHLLDWFAVDESNEVDPE FSARLTGIKLEMEPSLSFYNKARNYATKKPYSVEKFKLNFQMPTL ARGWDVNREKNNGAILFVKNGLYYLGIMPKQKGRYKALSFEPTEK TSEGFDKMYYDYFPDAAKMIPKCSTQLKAVTAHFQTHTTPILLSN NFIEPLEITKEIYDLNNPEKEPKKFQTAYAKKTGDQKGYREALCK WIDFTRDFLSKYTKTTSIDLSSLRPSSQYKDLGEYYAELNPLLYH ISFQRIAEKEIMDAVETGKLYLFQIYNKDFAKGHHGKPNLHTLYW TGLFSPENLAKTSIKLNGQAELFYRPKSRMKRMAHRLGEKMLNKK LKDQKTPIPDTLYQELYDYVNHRLSHDLSDEARALLPNVITKEVS HEIIKDRRFTSDKFFFHVPITLNYQAANSPSKFNQRVNAYLKEHP ETPIIGIARGERNLIYITVIDSTGKILEQRSLNTIQQFDYQKKLD NREKERVAARQAWSVVGTIKDLKQGYLSQVIHEIVDLMIHYQAVV VLENLNFGFKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYPAEK VGGVLNPYQLTDQFTSFAKMGTQSGFLFYVPAPYTSKIDPLTGFV DPFVWKTIKNHESRKHFLEGFDFLHYDVKTGDFILHFKMNRNLSF QRGLPGFMPAWDIVFEKNETQFDAKGTPFIAGKRIVPVIENHRFT GRYRDLYPANELIALLEEKGIVFRDGSNILPKLLENDDSHAIDTM VALIRSVLQMRNSNAATGEDYINSPVRDLNGVCFDSRFQNPEWPM DADANGAYHIALKGQLLLNHLKESKDLKLQNGISNQDWLAYIQEL RN
23	inactive RVRAsCpf1	MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARND HYKELKPIIDRIYKTYADQCLQLVQLDWENLSAAIDSYRKEKTEE TRNALIEEQATYRNAIHDYFIGRTDNLTDAINKRHAEIYKGLFKA ELFNGKVLKQLGTVTTTEHENALLRSFDKFTTYFSGFYENRKNVF SAEDISTAIPHRIVQDNFPKFKENCHIFTRLITAVPSLREHFENV KKAIGIFVSTSIEEVFSFPFYNQLLTQTQIDLYNQLLGGISREAG TEKIKGLNEVLNLAIQKNDETAHIIASLPHRFIPLFKQILSDRNT LSFILEEFKSDEEVIQSFCKYKTLLRNENVLETAEALFNELNSID LTHIFISHKKLETISSALCDHWDTLRNALYERRISELTGKITKSA

SEQ	Description	Sequence
		KEKVQRSLKHEDINLQEIISAAGKELSEAFKQKTSEILSHAHAAL DQPLPTTLKKQEEKEILKSQLDSLLGLYHLLDWFAVDESNEVDPE FSARLTGIKLEMEPSLSFYNKARNYATKKPYSVEKFKLNFQMPTL ARGWDVNVEKNRGAILFVKNGLYYLGIMPKQKGRYKALSFEPTEK TSEGFDKMYYDYFPDAAKMIPKCSTQLKAVTAHFQTHTTPILLSN NFIEPLEITKEIYDLNNPEKEPKKFQTAYAKKTGDQKGYREALCK WIDFTRDFLSKYTKTTSIDLSSLRPSSQYKDLGEYYAELNPLLYH ISFQRIAEKEIMDAVETGKLYLFQIYNKDFAKGHHGKPNLHTLYW TGLFSPENLAKTSIKLNGQAELFYRPKSRMKRMAHRLGEKMLNKK LKDQKTPIPDTLYQELYDYVNHRLSHDLSDEARALLPNVITKEVS HEIIKDRRFTSDKFFFHVPITLNYQAANSPSKFNQRVNAYLKEHP ETPIIGIARGERNLIYITVIDSTGKILEQRSLNTIQQFDYQKKLD NREKERVAARQAWSVVGTIKDLKQGYLSQVIHEIVDLMIHYQAVV VLENLNFGFKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYPAEK VGGVLNPYQLTDQFTSFAKMGTQSGFLFYVPAPYTSKIDPLTGFV DPFVWKTIKNHESRKHFLEGFDFLHYDVKTGDFILHFKMNRNLSF QRGLPGFMPAWDIVFEKNETQFDAKGTPFIAGKRIVPVIENHRFT GRYRDLYPANELIALLEEKGIVFRDGSNILPKLLENDDSHAIDTM VALIRSVLQMRNSNAATGEDYINSPVRDLNGVCFDSRFQNPEWPM DADANGAYHIALKGQLLLNHLKESKDLKLQNGISNQDWLAYIQEL RN
24	inactive RRAsCpf1	MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARND HYKELKPIIDRIYKTYADQCLQLVQLDWENLSAAIDSYRKEKTEE TRNALIEEQATYRNAIHDYFIGRTDNLTDAINKRHAEIYKGLFKA ELFNGKVLKQLGTVTTTEHENALLRSFDKFTTYFSGFYENRKNVF SAEDISTAIPHRIVQDNFPKFKENCHIFTRLITAVPSLREHFENV KKAIGIFVSTSIEEVFSFPFYNQLLTQTQIDLYNQLLGGISREAG TEKIKGLNEVLNLAIQKNDETAHIIASLPHRFIPLFKQILSDRNT LSFILEEFKSDEEVIQSFCKYKTLLRNENVLETAEALFNELNSID LTHIFISHKKLETISSALCDHWDTLRNALYERRISELTGKITKSA KEKVQRSLKHEDINLQEIISAAGKELSEAFKQKTSEILSHAHAAL DQPLPTTLKKQEEKEILKSQLDSLLGLYHLLDWFAVDESNEVDPE FSARLTGIKLEMEPSLSFYNKARNYATKKPYSVEKFKLNFQMPTL ARGWDVNKEKNNGAILFVKNGLYYLGIMPKQKGRYKALSFEPTEK TSEGFDKMYYDYFPDAAKMIPRCSTQLKAVTAHFQTHTTPILLSN NFIEPLEITKEIYDLNNPEKEPKKFQTAYAKKTGDQKGYREALCK WIDFTRDFLSKYTKTTSIDLSSLRPSSQYKDLGEYYAELNPLLYH ISFQRIAEKEIMDAVETGKLYLFQIYNKDFAKGHHGKPNLHTLYW TGLFSPENLAKTSIKLNGQAELFYRPKSRMKRMAHRLGEKMLNKK LKDQKTPIPDTLYQELYDYVNHRLSHDLSDEARALLPNVITKEVS HEIIKDRRFTSDKFFFHVPITLNYQAANSPSKFNQRVNAYLKEHP ETPIIGIARGERNLIYITVIDSTGKILEQRSLNTIQQFDYQKKLD NREKERVAARQAWSVVGTIKDLKQGYLSQVIHEIVDLMIHYQAVV VLENLNFGFKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYPAEK VGGVLNPYQLTDQFTSFAKMGTQSGFLFYVPAPYTSKIDPLTGFV DPFVWKTIKNHESRKHFLEGFDFLHYDVKTGDFILHFKMNRNLSF QRGLPGFMPAWDIVFEKNETQFDAKGTPFIAGKRIVPVIENHRFT GRYRDLYPANELIALLEEKGIVFRDGSNILPKLLENDDSHAIDTM VALIRSVLQMRNSNAATGEDYINSPVRDLNGVCFDSRFQNPEWPM DADANGAYHIALKGQLLLNHLKESKDLKLQNGISNQDWLAYIQEL RN
25	dCasX	MEKRINKIRKKLSADNATKPVSRSGPMKTLLVRVMTDDLKKRLEK RRKKPEVMPQVISNNAANNLRMLLDDYTKMKEAILQVYWQEFKDD HVGLMCKFAQPASKKIDQNKLKPEMDEKGNLTTAGFACSQCGQPL FVYKLEQVSEKGKAYTNYFGRCNVAEHEKLILLAQLKPEKDSDEA

SEQ	Description	Sequence
		VTYSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGPVG KALSDACMGTIASFLSKYQDIIIEHQKVVKGNQKRLESLRELAGK ENLEYPSVTLPPQPHTKEGVDAYNEVIARVRMWVNLNLWQKLKLS RDDAKPLLRLKGFPSFPVVERRENEVDWWNTINEVKKLIDAKRDM GRVFWSGVTAEKRNTILEGYNYLPNENDHKKREGSLENPKKPAKR QFGDLLLYLEKKYAGDWGKVFDEAWERIDKKIAGLTSHIEREEAR NAEDAQSKAVLTDWLRAKASFVLERLKEMDEKEFYACEIQLQKWY GDLRGNPFAVEAENRVVDISGFSIGSDGHSIQYRNLLAWKYLENG KREFYLLMNYGKKGRIRFTDGTDIKKSGKWQGLLYGGGKAKVIDL TFDPDDEQLIILPLAFGTRQGREFIWNDLLSLETGLIKLANGRVI EKTIYNKKIGRDEPALFVALTFERREVVDPSNIKPVNLIGVARGE NIPAVIALTDPEGCPLPEFKDSSGGPTDILRIGEGYKEKQRAIQA AKEVEQRRAGGYSRKFASKSRNLADDMVRNSARDLFYHAVTHDAV LVFANLSRGFGRQGKRTFMTERQYTKMEDWLTAKLAYEGLTSKTY LSKTLAQYTSKTCSNCGFTITTADYDGMLVRLKKTSDGWATTLNN KELKAEGQITYYNRYKRQTVEKELSAELDRLSEESGNNDISKWTK GRRDEALFLLKKRFSHRPVQEQFVCLDCGHEVHAAEQAALNIARS WLFLNSNSTEFKSYKSGKQPFVGAWQAFYKRRLKEVWKPNA
26	dCasPhi	MPKPAVESEFSKVLKKHFPGERFRSSYMKRGGKILAAQGEEAVVA YLQGKSEEEPPNFQPPAKCHVVTKSRDFAEWPIMKASEAIQRYIY ALSTTERAACKPGKSSESHAAWFAATGVSNHGYSHVQGLNLIFDH TLGRYDGVLKKVQLRNEKARARLESINASRADEGLPEIKAEEEEV ATNETGHLLQPPGINPSFYVYQTISPQAYRPRDEIVLPPEYAGYV RDPNAPIPLGVVRNRCDIQKGCPGYIPEWQREAGTAISPKTGKAV TVPGLSPKKNKRMRRYWRSEKEKAQDALLVTVRIGTDWVVIDVRG LLRNARWRTIAPKDISLNALLDLFTGDPVIDVRRNIVTFTYTLDA CGTYARKWTLKGKQTKATLDKLTATQTVALVAIALGQTNPISAGI SRVTQENGALQCEPLDRFTLPDDLLKDISAYRIAWDRNEEELRAR SVEALPEAQQAEVRALDGVSKETARTQLCADFGLDPKRLPWDKMS SNTTFISEALLSNSVSRDQVFFTPAPKKGAKKAPVEVMRKDRTW ARAYKPRLSVEAQKLKNEALWALKRTSPEYLKLSRRKEELCRRSI NYVIEKTRRRTQCQIVIPVIEDLNVRFFHGSGKRLPGWDNFFTAK KENRWFIQGLHKAFSDLRTHRSFYVFEVRPERTSITCPKCGHCEV GNRDGEAFQCLSCGKTCNADLDVATHNLTQVALTGKTMPKREEPR DAQGTAPARKTKKASKSKAPPAEREDQTPAQEPSQTS
27	inactive VRER SpCas9	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKK NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEM AKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTI YHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENL IAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDT YDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKA PLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPY YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERM TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFL SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRE MIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMAR ENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEK LYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNK VLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL

SEQ	Description	Sequence
		TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDE NDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLN AVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLA SARELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQ AENIIHLFTLTNLGAPAAFKYFDTTIDRKEYRSTKEVLDATLIHQ SITGLYETRIDLSQLGGD
28	inactive EQR SpCas9	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKK NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEM AKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTI YHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENL IAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDT YDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKA PLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPY YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERM TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFL SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRE MIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMAR ENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEK LYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNK VLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDE NDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLN AVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFESPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLA SAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQ AENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIHQ SITGLYETRIDLSQLGGD
29	inactive VQR SpCas9	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKK NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEM AKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTI YHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENL IAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDT YDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKA PLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPY YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERM TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFL SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED

SEQ	Description	Sequence
		RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRE MIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMAR ENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEK LYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNK VLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDE NDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLN AVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLA SAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQ AENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIHQ SITGLYETRIDLSQLGGD
30	inactive SPG SpCas9	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKK NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEM AKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTI YHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENL IAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDT YDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKA PLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPY YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERM TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFL SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRE MIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMAR ENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEK LYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNK VLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDE NDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLN AVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFLWPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLA SAKQLQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQ AENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIHQ SITGLYETRIDLSQLGGD
31	inactive SpRY Cas9	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKK NLIGALLFDSGETAERTRLKRTARRRYTRRKNRICYLQEIFSNEM AKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTI YHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENL IAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDT

SEQ	Description	Sequence
		YDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKA PLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPY YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERM TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFL SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRE MIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMAR ENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEK LYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNK VLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDE NDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLN AVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF ATVRKVLSMPQVNIVKKTEVQTGGFSKESIRPKRNSDKLIARKKD WDPKKYGGFLWPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLA SAKQLQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQ AENIIHLFTLTRLGAPRAFKYFDTTIDPKQYRSTKEVLDATLIHQ SITGLYETRIDLSQLGGD
32	inactive KKH dSaCas9	MKRNYILGLAIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNE GRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYE ARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTK EQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEA KQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKD IKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRD ENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRV TSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQS SEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILD ELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVV KRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQ KRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEA IPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEEASKKGN RTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEER DINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKS INGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKEWKK LDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKD FKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYD KDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLY KYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNS RNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSK CYEEAKKLKKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLN RIEVNMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDIL GNLYEVKSKKHPQIIKKG
33	ZIM3	MNNSQGRVTFEDVTVNFTQGEWQRLNPEQRNLYRDVMLENYSNLV SVGQGETTKPDVILRLEQGKEPWLEEEEVLGSGRAEKNGDIGGQI WKPKDVKESL
34	ZNF436	MAATLLMAGSQAPVTFEDMAMYLTREEWRPLDAAQRDLYRDVMQE NYGNVVSLDFEIRSENEVNPKQEISEDVQFGTTSERPAENAEENP ESEEGFESGDRSERQW

SEQ	Description	Sequence
35	ZNF257	MLENYRNLVFLGIAVSKPDLITCLEQGKEPCNMKRHEMVAKPPVM CSHIAEDLCPERDIKYFFQKVILRRYDKCEHENLQLRKGCKSVDE CKVCK
36	ZNF675	MGLLTFRDVAIEFSLEEWQCLDTAQRNLYKNVILENYRNLVFLGI AVSKQDLITCLEQEKEPLTVKRHEMVNEPPVMCSHFAQEFWPEQN IKDSF
37	ZNF490	MLQMQNSEHHGQSIKTQTDSISLEDVAVNFTLEEWALLDPGQRNI YRDVMRATFKNLACIGEKWKDQDIEDEHKNQGRNLRSPMVEALCE NKEDCPCGKSTSQIPDLNTNLETPTG
38	ZNF320	MALSQGLLTFRDVAIEFSQEEWKCLDPAQRTLYRDVMLENYRNLV SLDISSKCMMNTLSSTGQGNTEVIHTGTLQRQASYHIGAFCSQEI EKDIHDFVFQ
39	ZNF331	MAQGLVTFADVAIDFSQEEWACLNSAQRDLYWDVMLENYSNLVSL DLESAYENKSLPTKKNIHEIRASKRNSDRRSKSLGRNWICEGTLE RPQRSRGR
40	ZNF816	MLREEATKKSKEKEPGMALPQGRLTFRDVAIEFSLEEWKCLNPAQ RALYRAVMLENYRNLEFVDSSLKSMMEFSSTRHSITGEVIHTGTL QRHKSHHIGDFCFPEMKKDIHHFEFQWQ
41	ZNF680	MPGPPGSLEMGPLTFRDVAIEFSLEEWQCLDTAQRNLYRKVMFEN YRNLVFLGIAVSKPHLITCLEQGKEPWNRKRQEMVAKPPVIYSHF TEDLWPEHSIKDSF
42	ZNF41	MSPPWSPALAAEGRGSSCEASVSFEDVTVDFSKEEWQHLDPAQRR LYWDVTLENYSHLLSVGYQIPKSEAAFKLEQGEGPWMLEGEAPHQ SCSGEAIGKMQQQGIPGGIFFHC
43	ZNF189	MASPSPPPESKEEWDYLDPAQRSLYKDVMMENYGNLVSLDVLNRD KDEEPTVKQEIEEIEEEVEPQGVIVTRIKSEIDQDPMGRETFELV GRLDKQRGIFLWEIPRESL
44	ZNF528	MALTQGPLKFMDVAIEFSQEEWKCLDPAQRTLYRDVMLENYRNLV SLGICLPDLSVTSMLEQKRDPWTLQSEEKIANDPDGRECIKGVNT ERSSKLGSN
45	ZNF543	MAASAQVSVTFEDVAVTFTQEEWGQLDAAQRTLYQEVMLETCGLL MSLGCPLFKPELIYQLDHRQELWMATKDLSQSSYPGDNTKPKTTE PTFSHLALPE
46	ZNF554	MFSQEERMAAGYLPRWSQELVTFEDVSMDFSQEEWELLEPAQKNL YREVMLENYRNVVSLEALKNQCTDVGIKEGPLSPAQTSQVTSLSS WTGYLLFQPVASSHLEQREALWIEEKGTPQASCSDWMTVLRNQDS TYKKVALQE
47	ZNF140	MSQGSVTFRDVAIDFSQEEWKWLQPAQRDLYRCVMLENYGHLVSL GLSISKPDVVSLLEQGKEPWLGKREVKRDLFSVSESSGEIKDFSP KNVIYDD
48	ZNF610	MEEAQKRKAKESGMALPQGRLTFMDVAIEFSQEEWKSLDPGQRAL YRDVMLENYRNLVFLGRSCVLGSNAENKPIKNQLGLTLESHLSEL QLFQAGRKIYRSNQVEKFTNHR
49	ZNF264	MAAAVLTDRAQVSVTFDDVAVTFTKEEWGQLDLAQRTLYQEVMLE NCGLLVSLGCPVPKAELICHLEHGQEPWTRKEDLSQDTCPGDKGK PKTTEPTTCEPALSE
50	ZNF350	MIQAQESITLEDVAVDFTWEEWQLLGAAQKDLYRDVMLENYSNLV AVGYQASKPDALFKLEQGEQLWTIEDGIHSGACSDIWKVDHVLER LQSESLVNR
51	ZNF8	MEGVAGVMSVGPPAARLQEPVTFRDVAVDFTQEEWGQLDPTQRIL YRDVMLETFGHLLSIGPELPKPEVISQLEQGTELWVAERGTTQGC HPAWEPRSESQASRKEEGLPEE
52	ZNF582	MSLGSELFRDVAIVFSQEEWQWLAPAQRDLYRDVMLETYSNLVSL GLAVSKPDVISFLEQGKEPWMVERVVSGGLCPVLESRYDTKELFP KQHVYEV

SEQ	Description	Sequence
53	ZNF30	MAHKYVGLQYHGSVTFEDVAIAFSQQEWESLDSSQRGLYRDVMLE NYRNLVSMAGHSRSKPHVIALLEQWKEPEVTVRKDGRRWCTDLQL EDDTIGCKEMPTSEN
54	ZNF324	MAFEDVAVYFSQEEWGLLDTAQRALYRRVMLDNFALVASLGLSTS RPRVVIQLERGEEPWVPSGTDTTLSRTTYRRRNPGSWSLTEDRDV SG
55	ZNF98	MLENYRNLVFVGIAASKPDLITCLEQGKEPWNVKRHEMVTEPPVV YSYFAQDLWPKQGKKNYFQKVILRTYKKCGRENLQLRKYCKSMDE CKVHKECYNGLNQC
56	ZNF669	MHFRRPDPCREPLASPIQDSVAFEDVAVNFTQEEWALLDSSQKNL YREVMQETCRNLASVGSQWKDQNIEDHFEKPGKDIRNHIVQRLCE SKEDGQYGEVVSQIPNLDLNENISTGLKPCECSICGK
57	ZNF677	MALSQGLFTFKDVAIEFSQEEWECLDPAQRALYRDVMLENYRNLL SLDEDNIPPEDDISVGFTSKGLSPKENNKEELYHLVILERKESHG INNFDLKEVWENMPKFDSLW
58	ZNF596	MTFEDIIVDFTQEEWALLDTSQRKLFQDVMLENISHLVSIGKQLC KSVVLSQLEQVEKLSTQRISLLQGREVGIKHQEIPFIHHIYQKGT STISTMRS
59	ZNF214	MAVTFEDVTIIFTWEEWKFLDSSQKRLYREVMWENYTNVMSVENW NESYKSQEEKFRYLEYENFSYWQGWWNAGAQMYENQNYGETVQGT DSKDLTQQDRSQC
60	ZNF37A	MITSQGSVSFRDVTVGFTQEEWQHLDPAQRTLYRDVMLENYSHLV SVGYCIPKPEVILKLEKGEEPWILEEKFPSQSHLELINTSRNYSI MKFNEFNKG
61	ZNF34	MFEDVAVYLSREEWGRLGPAQRGLYRDVMLETYGNLVSLGVGPAG PKPGVISQLERGDEPWVLDVQGTSGKEHLRVNSPALGTRTEYKEL TSQETFGEEDPQGSEPVEACDHIS
62	ZNF250	METYGNVVSLGLPGSKPDIISQLERGEDPWVLDRKGAKKSQGLWS DYSDNLKYDHTTACTQQDSLSCPWECETKGESQNTDLSPKPLISE QTVILGKTPLGRIDQENNETKQ
63	ZNF547	MAEMNPAQGHVVFEDVAIYFSQEEWGHLDEAQRLLYRDVMLENLA LLSSLGCCHGAEDEEAPLEPGVSVGVSQVMAPKPCLSTQNTQPCE TCSSLLKDILRL
64	ZNF273	MLDNYRNLVFLGIAVSKPDLITCLEQGKEPCNMKRHAMVAKPPVV CSHFAQDLWPKQGLKDS
65	ZNF354A	MAAGQREARPQVSLTFEDVAVLFTRDEWRKLAPSQRNLYRDVMLE NYRNLVSLGLPFTKPKVISLLQQGEDPWEVEKDGSGVSSLGSKSS HKTTKSTQTQDSSFQ
66	ZFP82	MALRSVMFSDVSIDFSPEEWEYLDLEQKDLYRDVMLENYSNLVSL GCFISKPDVISSLEQGKEPWKVVRKGRRQYPDLETKYETKKLSLE NDIYEIN
67	ZNF224	MTTFKEAMTFKDVAVVFTEEELGLLDLAQRKLYRDVMLENFRNLL SVGHQAFHRDTFHFLREEKIWMMKTAIQREGNSGDKIQTEMETVS EAGTHQEW
68	ZNF33A	MFQVEQKSQESVSFKDVTVGFTQEEWQHLDPSQRALYRDVMLENY SNLVSVGYCVHKPEVIFRLQQGEEPWKQEEEFPSQSFPEVWTADH LKERSQENQSKHL
69	ZNF45	MTKSKEAVTFKDVAVVFSEEELQLLDLAQRKLYRDVMLENFRNVV SVGHQSTPDGLPQLEREEKLWMMKMATQRDNSSGAKNLKEMETLQ EVGLRYLP
70	ZNF175	MSQKPQVLGPEKQDGSCEASVSFEDVTVDFSREEWQQLDPAQRCL YRDVMLELYSHLFAVGYHIPNPEVIFRMLKEKEPRVEEAEVSHQR CQEREFGLEIPQKEISKKASFQ

SEQ	Description	Sequence
71	ZNF595	MELVTFRDVAIEFSPEEWKCLDPAQQNLYRDVMLENYRNLVSLGF VISNPDLVTCLEQIKEPCNLKIHETAAKPPAICSPFSQDLSPVQG IEDSF
72	ZNF184	MSTLLQGGHNLLSSASFQESVTFKDVIVDFTQEEWKQLDPGQRDL FRDVTLENYTHLVSIGLQVSKPDVISQLEQGTEPWIMEPSIPVGT CADWETRLENSVSAPEPDISEE
73	ZNF419	MDPAQVPVAADLLTDHEEGYVTFEDVAVYFSQEEWRLLDDAQRLL YRNVMLENFTLLASLGLASSKTHEITQLESWEEPFMPAWEVVTSA IPRGCWHGAEAEEAPEQIASVG
74	ZFP28-1	MKKLEAVGTGIEPKAMSQGLVTFGDVAVDFSQEEWEWLNPIQRNL YRKVMLENYRNLASLGLCVSKPDVISSLEQGKEPWTVKRKMTRAW CPDLKAVWKIKELPLKKDFCEG
75	ZFP28-2	MSLLGEHWDYDALFETQPGLVTIKNLAVDFRQQLHPAQKNFCKNG IWENNSDLGSAGHCVAKPDLVSLLEQEKEPWMVKRELTGSLFSGQ RSVHETQELFPKQDSYAE
76	ZNF18	MLALAASQPARLEERLIRDRDLGASLLPAAPQEQWRQLDSTQKEQ YWDLILETYGKMVSGAGISHPKSDLTNSIEFGEELAGIYLHVNEK IPRPTCIGDRQENDKENLNLENH
77	ZNF213	MEGRPGETTDTCFVSGVHGPVALGDIPFYFSREEWGTLDPAQRDL FWDIKRENSRNTTLGFGLKGQSEKSLLQEMVPVVPGQTGSDVTVS WSPEEAEAWESENRPRAALGPVVGARRGRPPTRRRQFRDLA
78	ZNF394	MVAVVRALQRALDGTSSQGMVTFEDTAVSLTWEEWERLDPARRDF CRESAQKDSGSTVPPSLESRVENKELIPMQQILEEAEPQGQLQEA FQGKRPLFSKCGSTHEDRVEKQSGDP
79	ZFP1	MNKSQGSVSFTDVTVDFTQEEWEQLDPSQRILYMDVMLENYSNLL SVEVWKADDQMERDHRNPDEQARQFLILKNQTPIEERGDLFGKAL NLNTDFVSLRQVPYKYDLYEKTL
80	ZFP14	MAHGSVTFRDVAIDFSQEEWEFLDPAQRDLYRDVMWENYSNFISL GPSISKPDVITLLDEERKEPGMVVREGTRRYCPDLESRYRTNTLS PEKDIYEIYSFQWDIMER
81	ZNF416	MAAAVLRDSTSVPVTAEAKLMGFTQGCVTFEDVAIYFSQEEWGLL DEAQRLLYRDVMLENFALITALVCWHGMEDEETPEQSVSVEGVPQ VRTPEASPSTQKIQSCDMCVPFLTDILHLTDLPGQELYLTGACAV FHQDQK
82	ZNF557	MLPPTAASQREGHTEGGELVNELLKSWLKGLVTFEDVAVEFTQEE WALLDPAQRTLYRDVMLENCRNLASLGNQVDKPRLISQLEQEDKV MTEERGILSGTCPDVENPFKAKGLTPKLHVFRKEQSRNMKMER
83	ZNF566	MAQESVMFSDVSVDFSQEEWECLNDDQRDLYRDVMLENYSNLVSM GHSISKPNVISYLEQGKEPWLADRELTRGQWPVLESRCETKKLFL KKEIYEIESTQWEIMEK
84	ZNF729	MPGAPGSLEMGPLTFRDVTIEFSLEEWQCLDTVQQNLYRDVMLEN YRNLVFLGMAVFKPDLITCLKQGKEPWNMKRHEMVTKPPVMRSHF TQDLWPDQSTKDSFQEVILRTYAR
85	ZIM2	MAGSQFPDFKHLGTFLVFEELVTFEDVLVDFSPEELSSLSAAQRN LYREVMLENYRNLVSLGHQFSKPDIISRLEEEESYAMETDSRHTV ICQGE
86	ZNF254	MPGPPRSLEMGLLTFRDVAIEFSLEEWQHLDIAQQNLYRNVMLEN YRNLAFLGIAVSKPDLITCLEQGKEPWNMKRHE
87	ZNF764	MAPPLAPLPPRDPNGAGPEWREPGAVSFADVAVYFCREEWGCLRP AQRALYRDVMRETYGHLSALGIGGNKPALISWVEEEAELWGPAAQ DPE
88	ZNF785	MGPPLAPRPAHVPGEAGPRRTRESRPGAVSFADVAVYFSPEEWEC LRPAQRALYRDVMRETFGHLGALGFSVPKPAFISWVEGEVEAWSP EAQDPDGESS

SEQ	Description	Sequence
89	ZNF10 (KOX1)	MDAKSLTAWSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLE NYKNLVSLGYQLTKPDVILRLEKGEEPWLVEREIHQETHPDSETA FEIKSSVSSRSIFKDKQSCDIKMEGMARNDLWYLSLEEVWKCRDQ LDKYQENPERHLRQVAFTQKKVLTQERVSESGKYGGNCLLPAQLV LREYFHKRDSHTKSLKHDLVLNGHQDSCASNSNECGQTFCQNIHL IQFARTHTGDKSYKCPDNDNSLTHGSSLGISKGIHREKPYECKEC GKFFSWRSNLTRHQLIHTGEKPYECKECGKSFSRSSHLIGHQKTH TGEEPYECKECGKSFSWFSHLVTHQRTHTGDKLYTCNQCGKSFVH SSRLIRHQRTHTGEKPYECPECGKSFRQSTHLILHQRTHVRVRPY ECNECGKSYSQRSHLVVHHRIHTGLKPFECKDCGKCFSRSSHLYS HQRTHTGEKPYECHDCGKSFSQSSALIVHQRIHTGEKPYECCQCG KAFIRKNDLIKHQRIHVGEETYKCNQCGIIFSQNSPFIVHQIAHT GEQFLTCNQCGTALVNTSNLIGYQTNHIRENAY
90	CBX5 (chromoshadow domain)	MGKKTKRTADSSSSEDEEEYVVEKVLDRRVVKGQVEYLLKWKGFS EEHNTWEPEKNLDCPELISEFMKKYKKMKEGENNKPREKSESNKR KSNFSNSADDIKSKKKREQSNDIARGFERGLEPEKIIGATDSCGD LMFLMKWKDTDEADLVLAKEANVKCPQIVIAFYEERLTWHAYPED AENKEKETAKS
91	RYBP (YAF2_RYBP component of PRC1)	MTMGDKKSPTRPKRQAKPAADEGFWDCSVCTFRNSAEAFKCSICD VRKGTSTRKPRINSQLVAQQVAQQYATPPPPKKEKKEKVEKQDKE KPEKDKEISPSVTKKNTNKKTKPKSDILKDPPSEANSIQSANATT KTSETNHTSRPRLKNVDRSTAQQLAVTVGNVTVIITDFKEKTRSS STSSSTVTSSAGSEQQNQSSSGSESTDKGSSRSSTPKGDMSAVND ESF
92	YAF2 (YAF2_RYBP component of PRC1)	MGDKKSPTRPKRQPKPSSDEGYWDCSVCTFRNSAEAFKCMMCDVR KGTSTRKPRPVSQLVAQQVTQQFVPPTQSKKEKKDKVEKEKSEKE TTSKKNSHKKTRPRLKNVDRSSAQHLEVTVGDLTVIITDFKEKTK SPPASSAASADQHSQSGSSSDNTERGMSRSSSPRGEASSLNGESH
93	MGA (component of PRC1.6)	MEEKQQIILANQDGGTVAGAAPTFFVILKQPGNGKTDQGILVTNQ DACALASSVSSPVKSKGKICLPADCTVGGITVTLDNNSMWNEFYH RSTEMILTKQGRRMFPYCRYWITGLDSNLKYILVMDISPVDNHRY KWNGRWWEPSGKAEPHVLGRVFIHPESPSTGHYWMHQPVSFYKLK LTNNTLDQEGHIILHSMHRYLPRLHLVPAEKAVEVIQLNGPGVHT FTFPQTEFFAVTAYQNIQITQLKIDYNPFAKGFRDDGLNNKPQRD GKQKNSSDQEGNNISSSSGHRVRLTEGQGSEIQPGDLDPLSRGHE TSGKGLEKTSLNIKRDFLGFMDTDSALSEVPQLKQEISECLIASS FEDDSRVASPLDQNGSFNVVIKEEPLDDYDYELGECPEGVTVKQE ETDEETDVYSNSDDDPILEKQLKRHNKVDNPEADHLSSKWLPSSP SGVAKAKMFKLDTGKMPVVYLEPCAVTRSTVKISELPDNMLSTSR KDKSSMLAELEYLPTYIENSNETAFCLGKESENGLRKHSPDLRVV QKYPLLKEPQWKYPDISDSISTERILDDSKDSVGDSLSGKEDLGR KRTTMLKIATAAKVVNANQNASPNVPGKRGRPRKLKLCKAGRPPK NTGKSLISTKNTPVSPGSTFPDVKPDLEDVDGVLFVSFESKEALD IHAVDGTTEESSSLQASTTNDSGYRARISQLEKELIEDLKTLRHK QVIHPGLQEVGLKLNSVDPTMSIDLKYLGVQLPLAPATSFPFWNL TGTNPASPDAGFPFVSRTGKTNDFTKIKGWRGKFHSASASRNEGG NSESSLKNRSAFCSDKLDEYLENEGKLMETSMGFSSNAPTSPVVY QLPTKSTSYVRTLDSVLKKQSTISPSTSYSLKPHSVPPVSRKAKS QNRQATFSGRTKSSYKSILPYPVSPKQKYSHVILGDKVTKNSSGI ISENQANNFVVPTLDENIFPKQISLRQAQQQQQQQQGSRPPGLSK SQVKLMDLEDCALWEGKPRTYITEERADVSLTTLLTAQASLKTKP IHTIIRKRAPPCNNDFCRLGCVCSSLALEKRQPAHCRRPDCMFGC TCLKRKVVLVKGGSKTKHFQRKAAHRDPVFYDTLGEEAREEEEGI REEEEQLKEKKKRKKLEYTICETEPEQPVRHYPLWVKVEGEVDPE PVYIPTPSVIEPMKPLLLPQPEVLSPTVKGKLLTGIKSPRSYTPK

SEQ	Description	Sequence
		PNPVIREEDKDPVYLYFESMMTCARVRVYERKKEDQRQPSSSSSP SPSFQQQTSCHSSPENHNNAKEPDSEQQPLKQLTCDLEDDSDKLQ EKSWKSSCNEGESSSTSYMHQRSPGGPTKLIEIISDCNWEEDRNK ILSILSQHINSNMPQSLKVGSFIIELASQRKSRGEKNPPVYSSRV KISMPSCQDQDDMAEKSGSETPDGPLSPGKMEDISPVQTDALDSV RERLHGGKGLPFYAGLSPAGKLVAYKRKPSSSTSGLIQVASNAKV AASRKPRTLLPSTSNSKMASSSGTATNRPGKNLKAFVPAKRPIAA RPSPGGVFTQFVMSKVGALQQKIPGVSTPQTLAGTQKFSIRPSPV MVVTPVVSSEPVQVCSPVTAAVTTTTPQVFLENTTAVTPMTAISD VETKETTYSSGATTTGVVEVSETNTSTSVTSTQSTATVNLTKTTG ITTPVASVAFPKSLVASPSTITLPVASTASTSLVVVTAAASSSMV TTPTSSLGSVPIILSGINGSPPVSQRPENAAQIPVATPQVSPNTV KRAGPRLLLIPVQQGSPTLRPVSNTQLQGHRMVLQPVRSPSGMNL FRHPNGQIVQLLPLHQLRGSNTQPNLQPVMFRNPGSVMGIRLPAP SKPSETPPSSTSSSAFSVMNPVIQAVGSSSAVNVITQAPSLLSSG ASFVSQAGTLTLRISPPEPQSFASKTGSETKITYSSGGQPVGTAS LIPLQSGSFALLQLPGQKPVPSSILQHVASLQMKRESQNPDQKDE TNSIKREQETKKVLQSEGEAVDPEANVIKQNSGAATSEETLNDSL EDRGDHLDEECLPEEGCATVKPSEHSCITGSHTDQDYKDVNEEYG ARNRKSSKEKVAVLEVRTISEKASNKTVQNLSKVQHQKLGDVKVE QQKGFDNPEENSSEFPVTFKEESKFELSGSKVMEQQSNLQPEAKE KECGDSLEKDRERWRKHLKGPLTRKCVGASQECKKEADEQLIKET KTCQENSDVFQQEQGISDLLGKSGITEDARVLKTECDSWSRISNP SAFSIVPRRAAKSSRGNGHFQGHLLLPGEQIQPKQEKKGGRSSAD FTVLDLEEDDEDDNEKTDDSIDEIVDVVSDYQSEEVDDVEKNNCV EYIEDDEEHVDIETVEELSEEINVAHLKTTAAHTQSFKQPSCTHI SADEKAAERSRKAPPIPLKLKPDYWSDKLQKEAEAFAYYRRTHTA NERRRRGEMRDLFEKLKITLGLLHSSKVSKSLILTRAFSEIQGLT DQADKLIGQKNLLTRKRNILIRKVSSLSGKTEEVVLKKLEYIYAK QQALEAQKRKKKMGSDEFDISPRISKQQEGSSASSVDLGQMFINN RRGKPLILSRKKDQATENTSPLNTPHTSANLVMTPQGQLLTLKGP LFSGPVVAVSPDLLESDLKPQVAGSAVALPENDDLFMMPRIVNVT SLATEGGLVDMGGSKYPHEVPDSKPSDHLKDTVRNEDNSLEDKGR ISSRGNRDGRVTLGPTQVFLANKDSGYPQIVDVSNMQKAQEFLPK KISGDMRGIQYKWKESESRGERVKSKDSSFHKLKMKDLKDSSIEM ELRKVTSAIEEAALDSSELLTNMEDEDDTDETLTSLLNEIAFLNQ QLNDDSVGLAELPSSMDTEFPGDARRAFISKVPPGSRATFQVEHL GTGLKELPDVQGESDSISPLLLHLEDDDFSENEKQLAEPASEPDV LKIVIDSEIKDSLLSNKKAIDGGKNTSGLPAEPESVSSPPTLHMK TGLENSNSTDTLWRPMPKLAPLGLKVANPSSDADGQSLKVMPCLA PIAAKVGSVGHKMNLTGNDQEGRESKVMPTLAPVVAKLGNSGASP SSAGK
94	CBX1 (chromoshadow)	MGKKQNKKKVEEVLEEEEEEYVVEKVLDRRVVKGKVEYLLKWKGF SDEDNTWEPEENLDCPDLIAEFLQSQKTAHETDKSEGGKRKADSD SEDKGEESKPKKKKEESEKPRGFARGLEPERIIGATDSSGELMFL MKWKNSDEADLVPAKEANVKCPQVVISFYEERLTWHSYPSEDDDK KDDKN
95	SCMH1 (SAM_1/SPM)	MLVCYSVLACEILWDLPCSIMGSPLGHFTWDKYLKETCSVPAPVH CFKQSYTPPSNEFKISMKLEAQDPRNTTSTCIATVVGLTGARLRL RLDGSDNKNDFWRLVDSAEIQPIGNCEKNGGMLQPPLGFRLNASS WPMFLLKTLNGAEMAPIRIFHKEPPSPSHNFFKMGMKLEAVDRKN PHFICPATIGEVRGSEVLVTFDGWRGAFDYWCRFDSRDIFPVGWC SLTGDNLQPPGTKVVIPKNPYPASDVNTEKPSIHSSTKTVLEHQP GQRGRKPGKKRGRTPKTLISHPISAPSKTAEPLKFPKKRGPKPGS KRKPRTLLNPPPASPTTSTPEPDTSTVPQDAATIPSSAMQAPTVC

SEQ	Description	Sequence
		IYLNKNGSTGPHLDKKKVQQLPDHFGPARASVVLQQAVQACIDCA YHQKTVFSFLKQGHGGEVISAVFDREQHTLNLPAVNSITYVLRFL EKLCHNLRSDNLFGNQPFTQTHLSLTAIEYSHSHDRYLPGETFVL GNSLARSLEPHSDSMDSASNPTNLVSTSQRHRPLLSSCGLPPSTA SAVRRLCSRGVLKGSNERRDMESFWKLNRSPGSDRYLESRDASRL SGRDPSSWTVEDVMQFVREADPQLGPHADLFRKHEIDGKALLLLR SDMMMKYMGLKLGPALKLSYHIDRLKQGKF
96	MPP8 (Chromodomain)	MEQVAEGARVTAVPVSAADSTEELAEVEEGVGVVGEDNDAAARGA EAFGDSEEDGEDVFEVEKILDMKTEGGKVLYKVRWKGYTSDDDTW EPEIHLEDCKEVLLEFRKKIAENKAKAVRKDIQRLSLNNDIFEAN SDSDQQSETKEDTSPKKKKKKLRQREEKSPDDLKKKKAKAGKLKD KSKPDLESSLESLVFDLRTKKRISEAKEELKESKKPKKDEVKETK ELKKVKKGEIRDLKTKTREDPKENRKTKKEKFVESQVESESSVLN DSPFPEDDSEGLHSDSREEKQNTKSARERAGQDMGLEHGFEKPLD SAMSAEEDTDVRGRRKKKTPRKAEDTRENRKLENKNAFLEKKTVP KKQRNQDRSKSAAELEKLMPVSAQTPKGRRLSGEERGLWSTDSAE EDKETKRNESKEKYQKRHDSDKEEKGRKEPKGLKTLKEIRNAFDL FKLTPEEKNDVSENNRKREEIPLDFKTIDDHKTKENKQSLKERRN TRDETDTWAYIAAEGDQEVLDSVCQADENSDGRQQILSLGMDLQL EWMKLEDFQKHLDGKDENFAATDAIPSNVLRDAVKNGDYITVKVA LNSNEEYNLDQEDSSGMTLVMLAAAGGQDDLLRLLITKGAKVNGR QKNGTTALIHAAEKNFLTTVAILLEAGAFVNVQQSNGETALMKAC KRGNSDIVRLVIECGADCNILSKHQNSALHFAKQSNNVLVYDLLK NHLETLSRVAEETIKDYFEARLALLEPVFPIACHRLCEGPDFSTD FNYKPPQNIPEGSGILLFIFHANFLGKEVIARLCGPCSVQAVVLN DKFQLPVFLDSHFVYSFSPVAGPNKLFIRLTEAPSAKVKLLIGAY RVQLQ
97	SUMO3 (Rad60-SLD)	MSEEKPKEGVKTENDHINLKVAGQDGSVVQFKIKRHTPLSKLMKA YCERQGLSMRQIRFRFDGQPINETDTPAQLEMEDEDTIDVFQQQT GGVPESSLAGHSF
98	HERC2 (Cyt-b5)	MPSESFCLAAQARLDSKWLKTDIQLAFTRDGLCGLWNEMVKDGEI VYTGTESTQNGELPPRKDDSVEPSGTKKEDLNDKEKKDEEETPAP IYRAKSILDSWVWGKQPDVNELKECLSVLVKEQQALAVQSATTTL SALRLKQRLVILERYFIALNRTVFQENVKVKWKSSGISLPPVDKK SSRPAGKGVEGLARVGSRAALSFAFAFLRRAWRSGEDADLCSELL QESLDALRALPEASLFDESTVSSVWLEVVERATRFLRSVVTGDVH GTPATKGPGSIPLQDQHLALAILLELAVQRGTLSQMLSAILLLLQ LWDSGAQETDNERSAQGTSAPLLPLLQRFQSIICRKDAPHSEGDM HLLSGPLSPNESFLRYLTLPQDNELAIDLRQTAVVVMAHLDRLAT PCMPPLCSSPTSHKGSLQEVIGWGLIGWKYYANVIGPIQCEGLAN LGVTQIACAEKRFLILSRNGRVYTQAYNSDTLAPQLVQGLASRNI VKIAAHSDGHHYLALAATGEVYSWGCGDGGRLGHGDTVPLEEPKV ISAFSGKQAGKHVVHIACGSTYSAAITAEGELYTWGRGNYGRLGH GSSEDEAIPMLVAGLKGLKVIDVACGSGDAQTLAVTENGQVWSWG DGDYGKLGRGGSDGCKTPKLIEKLQDLDVVKVRCGSQFSIALTKD GQVYSWGKGDNQRLGHGTEEHVRYPKLLEGLQGKKVIDVAAGSTH CLALTEDSEVHSWGSNDQCQHFDTLRVTKPEPAALPGLDTKHIVG IACGPAQSFAWSSCSEWSIGLRVPFVVDICSMTFEQLDLLLRQVS EGMDGSADWPPPQEKECVAVATLNLLRLQLHAAISHQVDPEFLGL GLGSILLNSLKQTVVTLASSAGVLSTVQSAAQAVLQSGWSVLLPT AEERARALSALLPCAVSGNEVNISPGRRFMIDLLVGSLMADGGLE SALHAAITAEIQDIEAKKEAQKEKEIDEQEANASTFHRSRTPLDK DLINTGICESSGKQCLPLVQLIQQLLRNIASQTVARLKDVARRIS SCLDFEQHSRERSASLDLLLRFQRLLISKLYPGESIGQTSDISSP ELMGVGSLLKKYTALLCTHIGDILPVAASIASTSWRHFAEVAYIV

SEQ	Description	Sequence
		EGDFTGVLLPELVVSIVLLLSKNAGLMQEAGAVPLLGGLLEHLDR FNHLAPGKERDDHEELAWPGIMESFFTGQNCRNNEEVTLIRKADL ENHNKDGGFWTVIDGKVYDIKDFQTQSLTGNSILAQFAGEDPVVA LEAALQFEDTRESMHAFCVGQYLEPDQEIVTIPDLGSLSSPLIDT ERNLGLLLGLHASYLAMSTPLSPVEIECAKWLQSSIFSGGLQTSQ IHYSYNEEKDEDHCSSPGGTPASKSRLCSHRRALGDHSQAFLQAI ADNNIQDHNVKDFLCQIERYCRQCHLTTPIMFPPEHPVEEVGRLL LCCLLKHEDLGHVALSLVHAGALGIEQVKHRTLPKSVVDVCRVVY QAKCSLIKTHQEQGRSYKEVCAPVIERLRFLFNELRPAVCNDLSI MSKFKLLSSLPRWRRIAQKIIRERRKKRVPKKPESTDDEEKIGNE ESDLEEACILPHSPINVDKRPIAIKSPKDKWQPLLSTVTGVHKYK WLKQNVQGLYPQSPLLSTIAEFALKEEPVDVEKMRKCLLKQLERA EVRLEGIDTILKLASKNFLLPSVQYAMFCGWQRLIPEGIDIGEPL TDCLKDVDLIPPFNRMLLEVTFGKLYAWAVQNIRNVLMDASAKFK ELGIQPVPLQTITNENPSGPSLGTIPQARFLLVMLSMLTLQHGAN NLDLLLNSGMLALTQTALRLIGPSCDNVEEDMNASAQGASATVLE ETRKETAPVQLPVSGPELAAMMKIGTRVMRGVDWKWGDQDGPPPG LGRVIGELGEDGWIRVQWDTGSTNSYRMGKEGKYDLKLAELPAAA QPSAEDSDTEDDSEAEQTERNIHPTAMMFTSTINLLQTLCLSAGV HAEIMQSEATKTLCGLLRMLVESGTTDKTSSPNRLVYREQHRSWC TLGFVRSIALTPQVCGALSSPQWITLLMKVVEGHAPFTATSLQRQ ILAVHLLQAVLPSWDKTERARDMKCLVEKLFDFLGSLLTTCSSDV PLLRESTLRRRRVRPQASLTATHSSTLAEEVVALLRTLHSLTQWN GLINKYINSQLRSITHSFVGRPSEGAQLEDYFPDSENPEVGGLMA VLAVIGGIDGRLRLGGQVMHDEFGEGTVTRITPKGKITVQFSDMR TCRVCPLNQLKPLPAVAFNVNNLPFTEPMLSVWAQLVNLAGSKLE KHKIKKSTKQAFAGQVDLDLLRCQQLKLYILKAGRALLSHQDKLR QILSQPAVQETGTVHTDDGAVVSPDLGDMSPEGPQPPMILLQQLL ASATQPSPVKAIFDKQELEAAALAVCQCLAVESTHPSSPGFEDCS SSEATTPVAVQHIRPARVKRRKQSPVPALPIVVQLMEMGFSRRNI EFALKSLTGASGNASSLPGVEALVGWLLDHSDIQVTELSDADTVS DEYSDEEVVEDVDDAAYSMSTGAVVTESQTYKKRADFLSNDDYAV YVRENIQVGMMVRCCRAYEEVCEGDVGKVIKLDRDGLHDLNVQCD WQQKGGTYWVRYIHVELIGYPPPSSSSHIKIGDKVRVKASVTTPK YKWGSVTHQSVGVVKAFSANGKDIIVDFPQQSHWTGLLSEMELVP SIHPGVTCDGCQMFPINGSRFKCRNCDDFDFCETCFKTKKHNTRH TFGRINEPGQSAVFCGRSGKQLKRCHSSQPGMLLDSWSRMVKSLN VSSSVNQASRLIDGSEPCWQSSGSQGKHWIRLEIFPDVLVHRLKM IVDPADSSYMPSLVVVSGGNSLNNLIELKTININPSDTTVPLLND CTEYHRYIEIAIKQCRSSGIDCKIHGLILLGRIRAEEEDLAAVPF LASDNEEEEDEKGNSGSLIRKKAAGLESAATIRTKVFVWGLNDKD QLGGLKGSKIKVPSFSETLSALNVVQVAGGSKSLFAVTVEGKVYA CGEATNGRLGLGISSGTVPIPRQITALSSYVVKKVAVHSGGRHAT ALTVDGKVFSWGEGDDGKLGHFSRMNCDKPRLIEALKTKRIRDIA CGSSHSAALTSSGELYTWGLGEYGRLGHGDNTTQLKPKMVKVLLG HRVIQVACGSRDAQTLALTDEGLVFSWGDGDFGKLGRGGSEGCNI PQNIERLNGQGVCQIECGAQFSLALTKSGVVWTWGKGDYFRLGHG SDVHVRKPQVVEGLRGKKIVHVAVGALHCLAVTDSGQVYAWGDND HGQQGNGTTTVNRKPTLVQGLEGQKITRVACGSSHSVAWTTVDVA TPSVHEPVLFQTARDPLGASYLGVPSDADSSAASNKISGASNSKP NRPSLAKILLSLDGNLAKQQALSHILTALQIMYARDAVVGALMPA AMIAPVECPSFSSAAPSDASAMASPMNGEECMLAVDIEDRLSPNP WQEKREIVSSEDAVTPSAVTPSAPSASARPFIPVTDDLGAASIIA ETMTKTKEDVESQNKAAGPEPQALDEFTSLLIADDTRVVVDLLKL SVCSRAGDRGRDVLSAVLSGMGTAYPQVADMLLELCVTELEDVAT DSQSGRLSSQPVVVESSHPYTDDTSTSGTVKIPGAEGLRVEFDRQ

SEQ	Description	Sequence
		CSTERRHDPLTVMDGVNRIVSVRSGREWSDWSSELRIPGDELKWK FISDGSVNGWGWRFTVYPIMPAAGPKELLSDRCVLSCPSMDLVTC LLDFRLNLASNRSIVPRLAASLAACAQLSALAASHRMWALQRLRK LLTTEFGQSININRLLGENDGETRALSFTGSALAALVKGLPEALQ RQFEYEDPIVRGGKQLLHSPFFKVLVALACDLELDTLPCCAETHK WAWFRRYCMASRVAVALDKRTPLPRLFLDEVAKKIRELMADSENM DVLHESHDIFKREQDEQLVQWMNRRPDDWTLSAGGSGTIYGWGHN HRGQLGGIEGAKVKVPTPCEALATLRPVQLIGGEQTLFAVTADGK LYATGYGAGGRLGIGGTESVSTPTLLESIQHVFIKKVAVNSGGKH CLALSSEGEVYSWGEAEDGKLGHGNRSPCDRPRVIESLRGIEVVD VAAGGAHSACVTAAGDLYTWGKGRYGRLGHSDSEDQLKPKLVEAL QGHRVVDIACGSGDAQTLCLTDDDTVWSWGDGDYGKLGRGGSDGC KVPMKIDSLTGLGVVKVECGSQFSVALTKSGAVYTWGKGDYHRLG HGSDDHVRRPRQVQGLQGKKVIAIATGSLHCVCCTEDGEVYTWGD NDEGQLGDGTTNAIQRPRLVAALQGKKVNRVACGSAHTLAWSTSK PASAGKLPAQVPMEYNHLQEIPIIALRNRLLLLHHLSELFCPCIP MFDLEGSLDETGLGPSVGFDTLRGILISQGKEAAFRKVVQATMVR DRQHGPVVELNRIQVKRSRSKGGLAGPDGTKSVFGQMCAKMSSFG PDSLLLPHRVWKVKFVGESVDDCGGGYSESIAEICEELQNGLTPL LIVTPNGRDESGANRDCYLLSPAARAPVHSSMFRFLGVLLGIAIR TGSPLSLNLAEPVWKQLAGMSLTIADLSEVDKDFIPGLMYIRDNE ATSEEFEAMSLPFTVPSASGQDIQLSSKHTHITLDNRAEYVRLAI NYRLHEFDEQVAAVREGMARVVPVPLLSLFTGYELETMVCGSPDI PLHLLKSVATYKGIEPSASLIQWFWEVMESFSNTERSLFLRFVWG RTRLPRTIADFRGRDFVIQVLDKYNPPDHFLPESYTCFFLLKLPR YSCKQVLEEKLKYAIHFCKSIDTDDYARIALTGEPAADDSSDDSD NEDVDSFASDSTQDYLTGH
99	BIN1 (SH3_9)	MAEMGSKGVTAGKIASNVQKKLTRAQEKVLQKLGKADETKDEQFE QCVQNFNKQLTEGTRLQKDLRTYLASVKAMHEASKKLNECLQEVY EPDWPGRDEANKIAENNDLLWMDYHQKLVDQALLTMDTYLGQFPD IKSRIAKRGRKLVDYDSARHHYESLQTAKKKDEAKIAKPVSLLEK AAPQWCQGKLQAHLVAQTNLLRNQAEEELIKAQKVFEEMNVDLQE ELPSLWNSRVGFYVNTFQSIAGLEENFHKEMSKLNQNLNDVLVGL EKQHGSNTFTVKAQPSDNAPAKGNKSPSPPDGSPAATPEIRVNHE PEPAGGATPGATLPKSPSQLRKGPPVPPPPKHTPSKEVKQEQILS LFEDTFVPEISVTTPSQFEAPGPFSEQASLLDLDFDPLPPVTSPV KAPTPSGQSIPWDLWEPTESPAGSLPSGEPSAAEGTFAVSWPSQT AEPGPAQPAEASEVAGGTQPAAGAQEPGETAASEAASSSLPAVVV ETFPATVNGTVEGGSGAGRLDLPPGFMFKVQAQHDYTATDTDELQ LKAGDVVLVIPFQNPEEQDEGWLMGVKESDWNQHKELEKCRGVFP ENFTERVP
100	PCGF2 (RING finger protein domain)	MHRTTRIKITELNPHLMCALCGGYFIDATTIVECLHSFCKTCIVR YLETNKYCPMCDVQVHKTRPLLSIRSDKTLQDIVYKLVPGLFKDE MKRRRDFYAAYPLTEVPNGSNEDRGEVLEQEKGALSDDEIVSLSI EFYEGARDRDEKKGPLENGDGDKEKTGVRFLRCPAAMTVMHLAKF LRNKMDVPSKYKVEVLYEDEPLKEYYTLMDIAYIYPWRRNGPLPL KYRVQPACKRLTLATVPTPSEGTNTSGASECESVSDKAPSPATLP ATSSSLPSPATPSHGSPSSHGPPATHPTSPTPPSTASGATTAANG GSLNCLQTPSSTSRGRKMTVNGAPVPPLT
101	TOX (HMG box)	MDVRFYPPPAQPAAAPDAPCLGPSPCLDPYYCNKFDGENMYMSMT EPSQDYVPASQSYPGPSLESEDFNIPPITPPSLPDHSLVHLNEVE SGYHSLCHPMNHNGLLPFHPQNMDLPEITVSNMLGQDGTLLSNSI SVMPDIRNPEGTQYSSHPQMAAMRPRGQPADIRQQPGMMPHGQLT TINQSQLSAQLGLNMGGSNVPHNSPSPPGSKSATPSPSSSVHEDE GDDTSKINGGEKRPASDMGKKPKTPKKKKKKDPNEPQKPVSAYAL

SEQ	Description	Sequence
		FFRDTQAAIKGQNPNATFGEVSKIVASMWDGLGEEQKQVYKKKTE AAKKEYLKQLAAYRASLVSKSYSEPVDVKTSQPPQLINSKPSVFH GPSQAHSALYLSSHYHQQPGMNPHLTAMHPSLPRNIAPKPNNQMP VTVSIANMAVSPPPPLQISPPLHQHLNMQQHQPLTMQQPLGNQLP MQVQSALHSPTMQQGFTLQPDYQTIINPTSTAAQVVTQAMEYVRS GCRNPPPQPVDWNNDYCSSGGMQRDKALYLT
102	FOXA1 (HNF3A C-terminal domain)	MLGTVKMEGHETSDWNSYYADTQEAYSSVPVSNMNSGLGSMNSMN TYMTMNTMTTSGNMTPASFNMSYANPGLGAGLSPGAVAGMPGGSA GAMNSMTAAGVTAMGTALSPSGMGAMGAQQAASMNGLGPYAAAMN PCMSPMAYAPSNLGRSRAGGGGDAKTFKRSYPHAKPPYSYISLIT MAIQQAPSKMLTLSEIYQWIMDLFPYYRQNQQRWQNSIRHSLSFN DCFVKVARSPDKPGKGSYWTLHPDSGNMFENGCYLRRQKRFKCEK QPGAGGGGGSGSGGSGAKGGPESRKDPSGASNPSADSPLHRGVHG KTGQLEGAPAPGPAASPQTLDHSGATATGGASELKTPASSTAPPI SSGPGALASVPASHPAHGLAPHESQLHLKGDPHYSFNHPFSINNL MSSSEQQHKLDFKAYEQALQYSPYGSTLPASLPLGSASVTTRSPI EPSALEPAYYQGVYSRPVLNTS
103	FOXA2 (HNF3B C-terminal domain)	MLGAVKMEGHEPSDWSSYYAEPEGYSSVSNMNAGLGMNGMNTYMS MSAAAMGSGSGNMSAGSMNMSSYVGAGMSPSLAGMSPGAGAMAGM GGSAGAAGVAGMGPHLSPSLSPLGGQAAGAMGGLAPYANMNSMSP MYGQAGLSRARDPKTYRRSYTHAKPPYSYISLITMAIQQSPNKML TLSEIYQWIMDLFPFYRQNQQRWQNSIRHSLSFNDCFLKVPRSPD KPGKGSFWTLHPDSGNMFENGCYLRRQKRFKCEKQLALKEAAGAA GSGKKAAAGAQASQAQLGEAAGPASETPAGTESPHSSASPCQEHK RGGLGELKGTPAAALSPPEPAPSPGQQQQAAAHLLGPPHHPGLPP EAHLKPEHHYAFNHPFSINNLMSSEQQHHHSHHHHQPHKMDLKAY EQVMHYPGYGSPMPGSLAMGPVTNKTGLDASPLAADTSYYQGVYS RPIMNSS
104	IRF2BP1 (IRF-2BP1_2 N-terminal domain)	MASVQASRRQWCYLCDLPKMPWAMVWDFSEAVCRGCVNFEGADRI ELLIDAARQLKRSHVLPEGRSPGPPALKHPATKDLAAAAAQGPQL PPPQAQPQPSGTGGGVSGQDRYDRATSSGRLPLPSPALEYTLGSR LANGLGREEAVAEGARRALLGSMPGLMPPGLLAAAVSGLGSRGLT LAPGLSPARPLFGSDFEKEKQQRNADCLAELNEAMRGRAEEWHGR PKAVREQLLALSACAPFNVRFKKDHGLVGRVFAFDATARPPGYEF ELKLFTEYPCGSGNVYAGVLAVARQMFHDALREPGKALASSGFKY LEYERRHGSGEWRQLGELLTDGVRSFREPAPAEALPQQYPEPAPA ALCGPPPRAPSRNLAPTPRRRKASPEPEGEAAGKMTTEEQQQRHW VAPGGPYSAETPGVPSPIAALKNVAEALGHSPKDPGGGGGPVRAG GASPAASSTAQPPTQHRLVARNGEAEVSPTAGAEAVSGGGSGTGA TPGAPLCCTLCRERLEDTHFVQCPSVPGHKFCFPCSREFIKAQGP AGEVYCPSGDKCPLVGSSVPWAFMQGEIATILAGDIKVKKERDP
105	IRF2BP2 (IRF-2BP1_2 N-terminal domain)	MAAAVAVAAASRRQSCYLCDLPRMPWAMIWDFTEPVCRGCVNYEG ADRVEFVIETARQLKRAHGCFPEGRSPPGAAASAAAKPPPLSAKD ILLQQQQQLGHGGPEAAPRAPQALERYPLAAAAERPPRLGSDFGS SRPAASLAQPPTPQPPPVNGILVPNGFSKLEEPPELNRQSPNPRR GHAVPPTLVPLMNGSATPLPTALGLGGRAAASLAAVSGTAAASLG SAQPTDLGAHKRPASVSSSAAVEHEQREAAAKEKQPPPPAHRGPA DSLSTAAGAAELSAEGAGKSRGSGEQDWVNRPKTVRDTLLALHQH GHSGPFESKFKKEPALTAGRLLGFEANGANGSKAVARTARKRKPS PEPEGEVGPPKINGEAQPWLSTSTEGLKIPMTPTSSFVSPPPPTA SPHSNRTTPPEAAQNGQSPMAALILVADNAGGSHASKDANQVHST TRRNSNSPPSPSSMNQRRLGPREVGGQGAGNTGGLEPVHPASLPD SSLATSAPLCCTLCHERLEDTHFVQCPSVPSHKFCFPCSRQSIKQ QGASGEVYCPSGEKCPLVGSNVPWAFMQGEIATILAGDVKVKKER DS

SEQ	Description	Sequence
106	IRF2BPL IRF-2BP1_2 N-terminal domain	MSAAQVSSSRRQSCYLCDLPRMPWAMIWDFSEPVCRGCVNYEGAD RIEFVIETARQLKRAHGCFQDGRSPGPPPPVGVKTVALSAKEAAA AAAAAAAAAAAAQQQQQQQQQQQQQQQQQQQQQQQQLNHVDGSS KPAVLAAPSGLERYGLSAAAAAAAAAAAAVEQRSRFEYPPPPVSL GSSSHTARLPNGLGGPNGFPKPTPEEGPPELNRQSPNSSSAAASV ASRRGTHGGLVTGLPNPGGGGGPQLTVPPNLLPQTLLNGPASAAV LPPPPPHALGSRGPPTPAPPGAPGGPACLGGTPGVSATSSSASSS TSSSVAEVGVGAGGKRPGSVSSTDQERELKEKQRNAEALAELSES LRNRAEEWASKPKMVRDTLLTLAGCTPYEVRFKKDHSLLGRVFAF DAVSKPGMDYELKLFIEYPTGSGNVYSSASGVAKQMYQDCMKDFG RGLSSGFKYLEYEKKHGSGDWRLLGDLLPEAVRFFKEGVPGADML PQPYLDASCPMLPTALVSLSRAPSAPPGTGALPPAAPSGRGAAAS LRKRKASPEPPDSAEGALKLGEEQQRQQWMANQSEALKLTMSAGG FAAPGHAAGGPPPPPPPLGPHSNRTTPPESAPQNGPSPMAALMSV ADTLGTAHSPKDGSSVHSTTASARRNSSSPVSPASVPGQRRLASR NGDLNLQVAPPPPSAHPGMDQVHPQNIPDSPMANSGPLCCTICHE RLEDTHFVQCPSVPSHKFCFPCSRESIKAQGATGEVYCPSGEKCP LVGSNVPWAFMQGEIATILAGDVKVKKERDP
107	HOXA13 (homeodomain)	MTASVLLHPRWIEPTVMFLYDNGGGLVADELNKNMEGAAAAAAAA AAAAAAGAGGGGFPHPAAAAAGGNFSVAAAAAAAAAAAANQCRNL MAHPAPLAPGAASAYSSAPGEAPPSAAAAAAAAAAAAAAAAAASS SGGPGPAGPAGAEAAKQCSPCSAAAQSSSGPAALPYGYFGSGYYP CARMGPHPNAIKSCAQPASAAAAAAFADKYMDTAGPAAEEFSSRA KEFAFYHQGYAAGPYHHHQPMPGYLDMPVVPGLGGPGESRHEPLG LPMESYQPWALPNGWNGQMYCPKEQAQPPHLWKSTLPDVVSHPSD ASSYRRGRKKRVPYTKVQLKELEREYATNKFITKDKRRRISATTN LSERQVTIWFQNRRVKEKKVINKLKTTS
108	HOXB13 (homeodomain)	MEPGNYATLDGAKDIEGLLGAGGGRNLVAHSPLTSHPAAPTLMPA VNYAPLDLPGSAEPPKQCHPCPGVPQGTSPAPVPYGYFGGGYYSC RVSRSSLKPCAQAATLAAYPAETPTAGEEYPSRPTEFAFYPGYPG TYQPMASYLDVSVVQTLGAPGEPRHDSLLPVDSYQSWALAGGWNS QMCCQGEQNPPGPFWKAAFADSSGQHPPDACAFRRGRKKRIPYSK GQLRELEREYAANKFITKDKRRKISAATSLSERQITIWFQNRRVK EKKVLAKVKNSATP
109	HOXC13 (homeodomain)	MTTSLLLHPRWPESLMYVYEDSAAESGIGGGGGGGGGGTGGAGGG CSGASPGKAPSMDGLGSSCPASHCRDLLPHPVLGRPPAPLGAPQG AVYTDIPAPEAARQCAPPPAPPTSSSATLGYGYPFGGSYYGCRLS HNVNLQQKPCAYHPGDKYPEPSGALPGDDLSSRAKEFAFYPSFAS SYQAMPGYLDVSVVPGISGHPEPRHDALIPVEGYQHWALSNGWDS QVYCSKEQSQSAHLWKSPFPDVVPLQPEVSSYRRGRKKRVPYTKV QLKELEKEYAASKFITKEKRRRISATTNLSERQVTIWFQNRRVKE KKVVSKSKAPHLHST
110	HOXA11 (homeodomain)	MDFDERGPCSSNMYLPSCTYYVSGPDFSSLPSFLPQTPSSRPMTY SYSSNLPQVQPVREVTFREYAIEPATKWHPRGNLAHCYSAEELVH RDCLQAPSAAGVPGDVLAKSSANVYHHPTPAVSSNFYSTVGRNGV LPQAFDQFFETAYGTPENLASSDYPGDKSAEKGPPAATATSAAAA AAATGAPATSSSDSGGGGGCRETAAAAEEKERRRRPESSSSPESS SGHTEDKAGGSSGQRTRKKRCPYTKYQIRELEREFFFSVYINKEK RLQLSRMLNLTDRQVKIWFQNRRMKEKKINRDRLQYYSANPLL
111	HOXC11 (homeodomain)	MFNSVNLGNFCSPSRKERGADFGERGSCASNLYLPSCTYYMPEFS TVSSFLPQAPSRQISYPYSAQVPPVREVSYGLEPSGKWHHRNSYS SCYAAADELMHRECLPPSTVTEILMKNEGSYGGHHHPSAPHATPA GFYSSVNKNSVLPQAFDRFFDNAYCGGGDPPAEPPCSGKGEAKGE PEAPPASGLASRAEAGAEAEAEEENTNPSSSGSAHSVAKEPAKGA

SEQ	Description	Sequence
		APNAPRTRKKRCPYSKFQIRELEREFFFNVYINKEKRLQLSRMLN LTDRQVKIWFQNRRMKEKKLSRDRLQYFSGNPLL
112	HOXC10 (homeodomain)	MTCPRNVTPNSYAEPLAAPGGGERYSRSAGMYMQSGSDFNCGVMR GCGLAPSLSKRDEGSSPSLALNTYPSYLSQLDSWGDPKAAYRLEQ PVGRPLSSCSYPPSVKEENVCCMYSAEKRAKSGPEAALYSHPLPE SCLGEHEVPVPSYYRASPSYSALDKTPHCSGANDFEAPFEQRASL NPRAEHLESPQLGGKVSFPETPKSDSQTPSPNEIKTEQSLAGPKG SPSESEKERAKAADSSPDTSDNEAKEEIKAENTTGNWLTAKSGRK KRCPYTKHQTLELEKEFLFNMYLTRERRLEISKTINLTDRQVKIW FQNRRMKLKKMNRENRIRELTSNFNFT
113	HOXA10 (homeodomain)	MSARKGYLLPSPNYPTTMSCSESPAANSFLVDSLISSGRGEAGGG GGGAGGGGGGGYYAHGGVYLPPAADLPYGLQSCGLFPTLGGKRNE AASPGSGGGGGGLGPGAHGYGPSPIDLWLDAPRSCRMEPPDGPPP PPQQQPPPPPQPPQPAPQATSCSFAQNIKEESSYCLYDSADKCPK VSATAAELAPFPRGPPPDGCALGTSSGVPVPGYFRLSQAYGTAKG YGSGGGGAQQLGAGPFPAQPPGRGFDLPPALASGSADAARKERAL DSPPPPTLACGSGGGSQGDEEAHASSSAAEELSPAPSESSKASPE KDSLGNSKGENAANWLTAKSGRKKRCPYTKHQTLELEKEFLFNMY LTRERRLEISRSVHLTDRQVKIWFQNRRMKLKKMNRENRIRELTA NFNFS
114	HOXB9 (homeodomain)	MSISGTLSSYYVDSIISHESEDAPPAKFPSGQYASSRQPGHAEHL EFPSCSFQPKAPVFGASWAPLSPHASGSLPSVYHPYIQPQGVPPA ESRYLRTWLEPAPRGEAAPGQGQAAVKAEPLLGAPGELLKQGTPE YSLETSAGREAVLSNQRPGYGDNKICEGSEDKERPDQTNPSANWL HARSSRKKRCPYTKYQTLELEKEFLFNMYLTRDRRHEVARLLNLS ERQVKIWFQNRRMKMKKMNKEQGKE
115	HOXA9 (homeodomain)	MATTGALGNYYVDSFLLGADAADELSVGRYAPGTLGQPPRQAATL AEHPDFSPCSFQSKATVFGASWNPVHAAGANAVPAAVYHHHHHHP YVHPQAPVAAAAPDGRYMRSWLEPTPGALSFAGLPSSRPYGIKPE PLSARRGDCPTLDTHTLSLTDYACGSPPVDREKQPSEGAFSENNA ENESGGDKPPIDPNNPAANWLHARSTRKKRCPYTKHQTLELEKEF LFNMYLTRDRRYEVARLLNLTERQVKIWFQNRRMKMKKINKDRAK DE
116	ZFP28_HUMAN	NKKLEAVGTGIEPKAMSQGLVTFGDVAVDFSQEEWEWLNPIQRNL YRKVMLENYRNLASLGLCVSKPDVISSLEQGKEPW
117	ZN334_HUMAN	KMKKFQIPVSFQDLTVNFTQEEWQQLDPAQRLLYRDVMLENYSNL VSVGYHVSKPDVIFKLEQGEEPWIVEEFSNQNYPD
118	ZN568_HUMAN	CSQESALSEEEEDTTRPLETVTFKDVAVDLTQEEWEQMKPAQRNL YRDVMLENYSNLVTVGCQVTKPDVIFKLEQEEEPW
119	ZN37A_HUMAN	ITSQGSVSFRDVTVGFTQEEWQHLDPAQRTLYRDVMLENYSHLVS VGYCIPKPEVILKLEKGEEPWILEEKFPSQSHLEL
120	ZN181_HUMAN	PQVTFNDVAIDFTHEEWGWLSSAQRDLYKDVMVQNYENLVSVAGL SVTKPYVITLLEDGKEPWMMEKKLSKGMIPDWESR
121	ZN510_HUMAN	PLRFSTLFQEQQKMNISQASVSFKDVTIEFTQEEWQQMAPVQKNL YRDVMLENYSNLVSVGYCCFKPEVIFKLEQGEEPW
122	ZN862_HUMAN	QDPSAEGLSEEVPVVFEELPVVFEDVAVYFTREEWGMLDKRQKEL YRDVMRMNYELLASLGPAAAKPDLISKLERRAAPW
123	ZN140_HUMAN	SQGSVTFRDVAIDFSQEEWKWLQPAQRDLYRCVMLENYGHLVSLG LSISKPDVVSLLEQGKEPWLGKREVKRDLFSVSES
124	ZN208_HUMAN	GSLTFRDVAIEFSLEEWQCLDTAQQNLYRNVMLENYRNLVFLGIA AFKPDLIIFLEEGKESWNMKRHEMVEESPVICSHF
125	ZN248_HUMAN	NKSQEQVSFKDVCVDFTQEEWYLLDPAQKILYRDVILENYSNLVS VGYCITKPEVIFKIEQGEEPWILEKGFPSQCHPER
126	ZN571_HUMAN	PHLLVTFRDVAIDFSQEEWECLDPAQRDLYRDVMLENYSNLISLD LESSCVTKKLSPEKEIYEMESLQWENMGKRINHHL

SEQ	Description	Sequence
127	ZN699_HUMAN	EEERKTAELQKNRIQDSVVFEDVAVDFTQEEWALLDLAQRNLYRD VMLENFQNLASLGYPLHTPHLISQWEQEEDLQTVK
128	ZN726_HUMAN	GLLTFRDVAIEFSLEEWQCLDTAQKNLYRNVMLENYRNLAFLGIA VSKPDLIICLEKEKEPWNMKRDEMVDEPPGICPHF
129	ZIK1_HUMAN	RAPTQVTVSPETHMDLTKGCVTFEDIAIYFSQDEWGLLDEAQRLL YLEVMLENFALVASLGCGHGTEDEETPSDQNVSVG
130	ZNF2_HUMAN	AAVSPTTRCQESVTFEDVAVVFTDEEWSRLVPIQRDLYKEVMLEN YNSIVSLGLPVPQPDVIFQLKRGDKPWMVDLHGSE
131	Z705F_HUMAN	HSLEKVTFEDVAIDFTQEEWDMMDTSKRKLYRDVMLENISHLVSL GYQISKSYIILQLEQGKELWREGRVFLQDQNPDRE
132	ZNF14_HUMAN	DSVSFEDVAVNFTLEEWALLDSSQKKLYEDVMQETFKNLVCLGKK WEDQDIEDDHRNQGKNRRCHMVERLCESRRGSKCG
133	ZN471_HUMAN	NVEVVKVMPQDLVTFKDVAIDFSQEEWQWMNPAQKRLYRSMMLEN YQSLVSLGLCISKPYVISLLEQGREPWEMTSEMTR
134	ZN624_HUMAN	TQPDEDLHLQAEETQLVKESVTFKDVAIDFTLEEWRLMDPTQRNL HKDVMLENYRNLVSLGLAVSKPDMISHLENGKGPW
135	ZNF84_HUMAN	TMLQESFSFDDLSVDFTQKEWQLLDPSQKNLYKDVMLENYSSLVS LGYEVMKPDVIFKLEQGEEPWVGDGEIPSSDSPEV
136	ZNF7_HUMAN	EVVTFGDVAVHFSREEWQCLDPGQRALYREVMLENHSSVAGLAGF LVFKPELISRLEQGEEPWVLDLQGAEGTEAPRTSK
137	ZN891_HUMAN	RNAEEERMIAVFLTTWLQEPMTFKDVAVEFTQEEWMMLDSAQRSL YRDVMLENYRNLTSVEYQLYRLTVISPLDQEEIRN
138	ZN337_HUMAN	GPQGARRQAFLAFGDVTVDFTQKEWRLLSPAQRALYREVTLENYS HLVSLGILHSKPELIRRLEQGEVPWGEERRRRPGP
139	Z705G_HUMAN	HSLKKLTFEDVAIDFTQEEWAMMDTSKRKLYRDVMLENISHLVSL GYQISKSYIILQLEQGKELWREGRVFLQDQNPNRE
140	ZN529_HUMAN	MPEVEFPDQFFTVLTMDHELVTLRDVVINFSQEEWEYLDSAQRNL YWDVMMENYSNLLSLDLESRNETKHLSVGKDIIQN
141	ZN729_HUMAN	PGAPGSLEMGPLTFRDVTIEFSLEEWQCLDTVQQNLYRDVMLENY RNLVFLGMAVFKPDLITCLKQGKEPWNMKRHEMVT
142	ZN419_HUMAN	RDPAQVPVAADLLTDHEEGYVTFEDVAVYFSQEEWRLLDDAQRLL YRNVMLENFTLLASLGLASSKTHEITQLESWEEPF
143	Z705A_HUMAN	HSLKKVTFEDVAIDFTQEEWAMMDTSKRKLYRDVMLENISHLVSL GYQISKSYIILQLEQGKELWREGREFLQDQNPDRE
144	ZNF45_HUMAN	TKSKEAVTFKDVAVVFSEEELQLLDLAQRKLYRDVMLENFRNVVS VGHQSTPDGLPQLEREEKLWMMKMATQRDNSSGAK
145	ZN302_HUMAN	SQVTFSDVAIDFSHEEWACLDSAQRDLYKDVMVQNYENLVSVGLS VTKPYVIMLLEDGKEPWMMEKKLSKAYPFPLSHSV
146	ZN486_HUMAN	PGPLRSLEMESLQFRDVAVEFSLEEWHCLDTAQQNLYRDVMLENY RHLVFLGIIVSKPDLITCLEQGIKPLTMKRHEMIA
147	ZN621_HUMAN	LQTTWPQESVTFEDVAVYFTQNQWASLDPAQRALYGEVMLENYAN VASLVAFPFPKPALISHLERGEAPWGPDPWDTEIL
148	ZN688_HUMAN	APLLAPRPGETRPGCRKPGTVSFADVAVYFSPEEWGCLRPAQRAL YRDVMQETYGHLGALGFPGPKPALISWMEQESEAW
149	ZN33A_HUMAN	NKVEQKSQESVSFKDVTVGFTQEEWQHLDPSQRALYRDVMLENYS NLVSVGYCVHKPEVIFRLQQGEEPWKQEEEFPSQS
150	ZN554_HUMAN	CFSQEERMAAGYLPRWSQELVTFEDVSMDFSQEEWELLEPAQKNL YREVMLENYRNVVSLEALKNQCTDVGIKEGPLSPA
151	ZN878_HUMAN	DSVAFEDVAVNFTQEEWALLDPSQKNLYREVMQETLRNLTSIGKK WNNQYIEDEHQNPRRNLRRLIGERLSESKESHQHG
152	ZN772_HUMAN	MGPAQVPMNSEVIVDPIQGQVNFEDVFVYFSQEEWVLLDEAQRLL YRDVMLENFALMASLGHTSFMSHIVASLVMGSEPW
153	ZN224_HUMAN	TTFKEAMTFKDVAVVFTEEELGLLDLAQRKLYRDVMLENFRNLLS VGHQAFHRDTFHFLREEKIWMMKTAIQREGNSGDK

SEQ	Description	Sequence
154	ZN184_HUMAN	DSTLLQGGHNLLSSASFQEAVTFKDVIVDFTQEEWKQLDPGQRDL FRDVTLENYTHLVSIGLQVSKPDVISQLEQGTEPW
155	ZN544_HUMAN	EARSMLVPPQASVCFEDVAMAFTQEEWEQLDLAQRTLYREVTLET WEHIVSLGLFLSKSDVISQLEQEEDLCRAEQEAPR
156	ZNF57_HUMAN	DSVVFEDVAVDFTLEEWALLDSAQRDLYRDVMLETFRNLASVDDG TQFKANGSVSLQDMYGQEKSKEQTIPNFTGNNSCA
157	ZN283_HUMAN	EESHGALISSCNSRTMTDGLVTFRDVAIDFSQEEWECLDPAQRDL YVDVMLENYSNLVSLDLESKTYETKKIFSENDIFE
158	ZN549_HUMAN	VITPQIPMVTEEFVKPSQGHVTFEDIAVYFSQEEWGLLDEAQRCL YHDVMLENFSLMASVGCLHGIEAEEAPSEQTLSAQ
159	ZN211_HUMAN	VQLRPQTRMATALRDPASGSVTFEDVAVYFSWEEWDLLDEAQKHL YFDVMLENFALTSSLGCWCGVEHEETPSEQRISGE
160	ZN615_HUMAN	MQAQESLTLEDVAVDFTWEEWQFLSPAQKDLYRDVMLENYSNLVA VGYQASKPDALSKLERGEETCTTEDEIYSRICSEI
161	ZN253_HUMAN	GPLQFRDVAIEFSLEEWHCLDTAQRNLYRDVMLENYRNLVFLGIV VSKPDLVTCLEQGKKPLTMERHEMIAKPPVMSSHF
162	ZN226_HUMAN	NMFKEAVTFKDVAVAFTEEELGLLGPAQRKLYRDVMVENFRNLLS VGHPPFKQDVSPIERNEQLWIMTTATRRQGNLGEK
163	ZN730_HUMAN	GALTFRDVAIEFSLEEWQCLDTEQQNLYRNVMLDNYRNLVFLGIA VSKPDLITCLEQEKEPWNLKTHDMVAKPPVICSHI
164	Z585A_HUMAN	SPQKSSALAPEDHGSSYEGSVSFRDVAIDFSREEWRHLDPSQRNL YRDVMLETYSHLLSVGYQVPEAEVVMLEQGKEPWA
165	ZN732_HUMAN	ELLTFRDVAIEFSPEEWKCLDPAQQNLYRDVMLENYRNLISLGVA ISNPDLVIYLEQRKEPYKVKIHETVAKHPAVCSHF
166	ZN681_HUMAN	EPLKFRDVAIEFSLEEWQCLDTIQQNLYRNVMLENYRNLVFLGIV VSKPDLITCLEQEKEPWTRKRHRMVAEPPVICSHF
167	ZN667_HUMAN	PSARGKSKSKAPITFGDLAIYFSQEEWEWLSPIQKDLYEDVMLEN YRNLVSLGLSFRRPNVITLLEKGKAPWMVEPVRRR
168	ZN649_HUMAN	TKAQESLTLEDVAVDFTWEEWQFLSPAQKDLYRDVMLENYSNLVS VGYQAGKPDALTKLEQGEPLWTLEDEIHSPAHPEI
169	ZN470_HUMAN	SQEEVEVAGIKLCKAMSLGSVTFTDVAIDFSQDEWEWLNLAQRSL YKKVMLENYRNLVSVGLCISKPDVISLLEQEKDPW
170	ZN484_HUMAN	TKSLESVSFKDVTVDFSRDEWQQLDLAQKSLYREVMLENYFNLIS VGCQVPKPEVIFSLEQEEPCMLDGEIPSQSRPDGD
171	ZN431_HUMAN	SGCPGAERNLLVYSYFEKETLTFRDVAIEFSLEEWECLNPAQQNL YMNVMLENYKNLVFLGVAVSKQDPVTCLEQEKEPW
172	ZN382_HUMAN	PLQGSVSFKDVTVDFTQEEWQQLDPAQKALYRDVMLENYCHFVSV GFHMAKPDMIRKLEQGEELWTQRIFPSYSYLEEDG
173	ZN254_HUMAN	PGPPRSLEMGLLTFRDVAIEFSLEEWQHLDIAQQNLYRNVMLENY RNLAFLGIAVSKPDLITCLEQGKEPWNMKRHEMVD
174	ZN124_HUMAN	SGHPGSWEMNSVAFEDVAVNFTQEEWALLDPSQKNLYRDVMQETF RNLASIGNKGEDQSIEDQYKNSSRNLRHIISHSGN
175	ZN607_HUMAN	SYGSITFGDVAIDFSHQEWEYLSLVQKTLYQEVMMENYDNLVSLA GHSVSKPDLITLLEQGKEPWMIVREETRGECTDLD
176	ZN317_HUMAN	DLFVCSGLEPHTPSVGSQESVTFQDVAVDFTEKEWPLLDSSQRKL YKDVMLENYSNLTSLGYQVGKPSLISHLEQEEEPR
177	ZN620_HUMAN	FQTAWRQEPVTFEDVAVYFTQNEWASLDSVQRALYREVMLENYAN VASLAFPFTTPVLVSQLEQGELPWGLDPWEPMGRE
178	ZN141_HUMAN	ELLTFRDVAIEFSPEEWKCLDPDQQNLYRDVMLENYRNLVSLGVA ISNPDLVTCLEQRKEPYNVKIHKIVARPPAMCSHF
179	ZN584_HUMAN	AGEAEAQLDPSLQGLVMFEDVTVYFSREEWGLLNVTQKGLYRDVM LENFALVSSLGLAPSRSPVFTQLEDDEQSWVPSWV
180	ZN540_HUMAN	AHALVTFRDVAIDFSQKEWECLDTTQRKLYRDVMLENYNNLVSLG YSGSKPDVITLLEQGKEPCVVARDVTGRQCPGLLS

SEQ	Description	Sequence
181	ZN75D_HUMAN	KRIKHWKMASKLILPESLSLLTFEDVAVYFSEEEWQLLNPLEKTL YNDVMQDIYETVISLGLKLKNDTGNDHPISVSTSE
182	ZN555_HUMAN	DSVVFEDVAVDFTLEEWALLDSAQRDLYRDVMLETFQNLASVDDE TQFKASGSVSQQDIYGEKIPKESKIATFTRNVSWA
183	ZN658_HUMAN	NMSQASVSFQDVTVEFTREEWQHLGPVERTLYRDVMLENYSHLIS VGYCITKPKVISKLEKGEEPWSLEDEFLNQRYPGY
184	ZN684_HUMAN	ISFQESVTFQDVAVDFTAEEWQLLDCAERTLYWDVMLENYRNLIS VGCPITKTKVILKVEQGQEPWMVEGANPHESSPES
185	RBAK_HUMAN	NTLQGPVSFKDVAVDFTQEEWQQLDPDEKITYRDVMLENYSHLVS VGYDTTKPNVIIKLEQGEEPWIMGGEFPCQHSPEA
186	ZN829_HUMAN	HPEEEERMHDELLQAVSKGPVMFRDVSIDFSQEEWECLDADQMNL YKEVMLENFSNLVSVGLSNSKPAVISLLEQGKEPW
187	ZN582_HUMAN	SLGSELFRDVAIVFSQEEWQWLAPAQRDLYRDVMLETYSNLVSLG LAVSKPDVISFLEQGKEPWMVERVVSGGLCPVLES
188	ZN112_HUMAN	TKFQEMVTFKDVAVVFTEEELGLLDSVQRKLYRDVMLENFRNLLL VAHQPFKPDLISQLEREEKLLMVETETPRDGCSGR
189	ZN716_HUMAN	AKRPGPPGSREMGLLTFRDIAIEFSLAEWQCLDHAQQNLYRDVML ENYRNLVSLGIAVSKPDLITCLEQNKEPQNIKRNE
190	HKR1_HUMAN	TCMVHRQTMSCSGAGGITAFVAFRDVAVYFTQEEWRLLSPAQRTL HREVMLETYNHLVSLEIPSSKPKLIAQLERGEAPW
191	ZN350_HUMAN	IQAQESITLEDVAVDFTWEEWQLLGAAQKDLYRDVMLENYSNLVA VGYQASKPDALFKLEQGEQLWTIEDGIHSGACSDI
192	ZN480_HUMAN	AQKRRKRKAKESGMALPQGHLTFRDVAIEFSQAEWKCLDPAQRAL YKDVMLENYRNLVSLGISLPDLNINSMLEQRREPW
193	ZN416_HUMAN	DSTSVPVTAEAKLMGFTQGCVTFEDVAIYFSQEEWGLLDEAQRLL YRDVMLENFALITALVCWHGMEDEETPEQSVSVEG
194	ZNF92_HUMAN	GPLTFRDVKIEFSLEEWQCLDTAQRNLYRDVMLENYRNLVFLGIA VSKPDLITWLEQGKEPWNLKRHEMVDKTPVMCSHF
195	ZN100_HUMAN	SGCPGAERSLLVQSYFEKGPLTFRDVAIEFSLEEWQCLDSAQQGL YRKVMLENYRNLVFLAGIALTKPDLITCLEQGKEP
196	ZN736_HUMAN	GVLTFRDVAVEFSPEEWECLDSAQQRLYRDVMLENYGNLVSLGLA IFKPDLMTCLEQRKEPWKVKRQEAVAKHPAGSFHF
197	ZNF74_HUMAN	KENLEDISGWGLPEARSKESVSFKDVAVDFTQEEWGQLDSPQRAL YRDVMLENYQNLLALGPPLHKPDVISHLERGEEPW
198	CBX1_HUMAN	EESEKPRGFARGLEPERIIGATDSSGELMFLMKWKNSDEADLVPA KEANVKCPQVVISFYEERLTWHSYPSEDDDKKDDK
199	ZN443_HUMAN	ASVALEDVAVNFTREEWALLGPCQKNLYKDVMQETIRNLDCVVMK WKDQNIEDQYRYPRKNLRCRMLERFVESKDGTQCG
200	ZN195_HUMAN	TLLTFRDVAIEFSLEEWKCLDLAQQNLYRDVMLENYRNLFSVGLT VCKPGLITCLEQRKEPWNVKRQEAADGHPEMGFHH
201	ZN530_HUMAN	AAALRAPTQQVFVAFEDVAIYFSQEEWELLDEMQRLLYRDVMLEN FAVMASLGCWCGAVDEGTPSAESVSVEELSQGRTP
202	ZN782_HUMAN	NTFQASVSFQDVTVEFSQEEWQHMGPVERTLYRDVMLENYSHLVS VGYCFTKPELIFTLEQGEDPWLLEKEKGFLSRNSP
203	ZN791_HUMAN	DSVAFEDVSVSFSQEEWALLAPSQKKLYRDVMQETFKNLASIGEK WEDPNVEDQHKNQGRNLRSHTGERLCEGKEGSQCA
204	ZN331_HUMAN	AQGLVTFADVAIDFSQEEWACLNSAQRDLYWDVMLENYSNLVSLD LESAYENKSLPTEKNIHEIRASKRNSDRRSKSLGR
205	Z354C_HUMAN	AVDLLSAQEPVTFRDVAVFFSQDEWLHLDSAQRALYREVMLENYS SLVSLGIPFSMPKLIHQLQQGEDPCMVEREVPSDT
206	ZN157_HUMAN	SPQRFPALIPGEPGRSFEGSVSFEDVAVDFTRQEWHRLDPAQRTM HKDVMLETYSNLASVGLCVAKPEMIFKLERGEELW
207	ZN727_HUMAN	RVLTFRDVAVEFSPEEWECLDSAQQRLYRDVMLENYGNLFSLGLA IFKPDLITYLEQRKEPWNARRQKTVAKHPAGSLHF

SEQ	Description	Sequence
208	ZN550_HUMAN	AETKDAAQMLVTFKDVAVTFTREEWRQLDLAQRTLYREVMLETCG LLVSLGHRVPKPELVHLLEHGQELWIVKRGLSHAT
209	ZN793_HUMAN	IEYQIPVSFKDVVVGFTQEEWHRLSPAQRALYRDVMLETYSNLVS VGYEGTKPDVILRLEQEEAPWIGEAACPGCHCWED
210	ZN235_HUMAN	TKFQEAVTFKDVAVAFTEEELGLLDSAQRKLYRDVMLENFRNLVS VGHQSFKPDMISQLEREEKLWMKELQTQRGKHSGD
211	ZNF8_HUMAN	DEGVAGVMSVGPPAARLQEPVTFRDVAVDFTQEEWGQLDPTQRIL YRDVMLETFGHLLSIGPELPKPEVISQLEQGTELW
212	ZN724_HUMAN	GPLTFMDVAIEFSVEEWQCLDTAQQNLYRNVMLENYRNLVFLGIA VSKPDLITCLEQGKEPWNMERHEMVAKPPGMCCYF
213	ZN573_HUMAN	HQVGLIRSYNSKTMTCFQELVTFRDVAIDFSRQEWEYLDPNQRDL YRDVMLENYRNLVSLGGHSISKPVVVDLLERGKEP
214	ZN577_HUMAN	NATIVMSVRREQGSSSGEGSLSFEDVAVGFTREEWQFLDQSQKVL YKEVMLENYINLVSIGYRGTKPDSLFKLEQGEPPG
215	ZN789_HUMAN	FPPARGKELLSFEDVAMYFTREEWGHLNWGQKDLYRDVMLENYRN MVLLGFQFPKPEMICQLENWDEQWILDLPRTGNRK
216	ZN718_HUMAN	ELLTFKDVAIEFSPEEWKCLDTSQQNLYRDVMLENYRNLVSLGVS ISNPDLVTSLEQRKEPYNLKIHETAARPPAVCSHF
217	ZN300_HUMAN	MKSQGLVSFKDVAVDFTQEEWQQLDPSQRTLYRDVMLENYSHLVS MGYPVSKPDVISKLEQGEEPWIIKGDISNWIYPDE
218	ZN383_HUMAN	AEGSVMFSDVSIDFSQEEWDCLDPVQRDLYRDVMLENYGNLVSMG LYTPKPQVISLLEQGKEPWMVGRELTRGLCSDLES
219	ZN429_HUMAN	GPLTFTDVAIEFSLEEWQCLDTAQQNLYRNVMLENYRNLVFLGIA VSKPDLITCLEKEKEPCKMKRHEMVDEPPVVCSHF
220	ZN677_HUMAN	ALSQGLFTFKDVAIEFSQEEWECLDPAQRALYRDVMLENYRNLLS LDEDNIPPEDDISVGFTSKGLSPKENNKEELYHLV
221	ZN850_HUMAN	NMEGLVMFQDLSIDFSQEEWECLDAAQKDLYRDVMMENYSSLVSL GLSIPKPDVISLLEQGKEPWMVSRDVLGGWCRDSE
222	ZN454_HUMAN	AVSHLPTMVQESVTFKDVAILFTQEEWGQLSPAQRALYRDVMLEN YSNLVSLGLLGPKPDTFSQLEKREVWMPEDTPGGF
223	ZN257_HUMAN	GPLTIRDVTVEFSLEEWHCLDTAQQNLYRDVMLENYRNLVFLGIA VSKPDLITCLEQGKEPCNMKRHEMVAKPPVMCSHI
224	ZN264_HUMAN	AAAVLTDRAQVSVTFDDVAVTFTKEEWGQLDLAQRTLYQEVMLEN CGLLVSLGCPVPKAELICHLEHGQEPWTRKEDLSQ
225	ZFP82_HUMAN	ALRSVMFSDVSIDFSPEEWEYLDLEQKDLYRDVMLENYSNLVSLG CFISKPDVISSLEQGKEPWKVVRKGRRQYPDLETK
226	ZFP14_HUMAN	AHGSVTFRDVAIDFSQEEWEFLDPAQRDLYRDVMWENYSNFISLG PSISKPDVITLLDEERKEPGMVVREGTRRYCPDLE
227	ZN485_HUMAN	APRAQIQGPLTFGDVAVAFTRIEWRHLDAAQRALYRDVMLENYGN LVSVGLLSSKPKLITQLEQGAEPWTEVREAPSGTH
228	ZN737_HUMAN	GPLQFRDVAIEFSLEEWHCLDTAQRNLYRNVMLENYRNLVFLGIV VSKPDLITCLEQGKKPLTMKKHEMVANPSVTCSHF
229	ZNF44_HUMAN	TLPRGQPEVLEWGLPKDQDSVAFEDVAVNFTHEEWALLGPSQKNL YRDVMRETIRNLNCIGMKWENQNIDDQHQNLRRNP
230	ZN596_HUMAN	PSPDSMTFEDIIVDFTQEEWALLDTSQRKLFQDVMLENISHLVSI GKQLCKSVVLSQLEQVEKLSTQRISLLQGREVGIK
231	ZN565_HUMAN	EESREIRAGQIVLKAMAQGLVTFRDVAIEFSLEEWKCLEPAQRDL YREVTLENFGHLASLGLSISKPDVVSLLEQGKEPW
232	ZN543_HUMAN	AASAQVSVTFEDVAVTFTQEEWGQLDAAQRTLYQEVMLETCGLLM SLGCPLFKPELIYQLDHRQELWMATKDLSQSSYPG
233	ZFP69_HUMAN	RESLEDEVTPGLPTAESQELLTFKDISIDFTQEEWGQLAPAHQNL YREVMLENYSNLVSVGYQLSKPSVISQLEKGEEPW
234	SUMO1_HUMAN	EGEYIKLKVIGQDSSEIHFKVKMTTHLKKLKESYCQRQGVPMNSL RFLFEGQRIADNHTPKELGMEEEDVIEVYQEQTGG

SEQ	Description	Sequence
235	ZNF12_HUMAN	NKSLGPVSFKDVAVDFTQEEWQQLDPEQKITYRDVMLENYSNLVS VGYHIIKPDVISKLEQGEEPWIVEGEFLLQSYPDE
236	ZN169_HUMAN	SPGLLTTRKEALMAFRDVAVAFTQKEWKLLSSAQRTLYREVMLEN YSHLVSLGIAFSKPKLIEQLEQGDEPWREENEHLL
237	ZN433_HUMAN	MFQDSVAFEDVAVTFTQEEWALLDPSQKNLCRDVMQETFRNLASI GKKWKPQNIYVEYENLRRNLRIVGERLFESKEGHQ
238	SUMO3_HUMAN	ENDHINLKVAGQDGSVVQFKIKRHTPLSKLMKAYCERQGLSMRQI RFRFDGQPINETDTPAQLEMEDEDTIDVFQQQTGG
239	ZNF98_HUMAN	PGPLGSLEMGVLTFRDVALEFSLEEWQCLDTAQQNLYRNVMLENY RNLVFVGIAASKPDLITCLEQGKEPWNVKRHEMVT
240	ZN175_HUMAN	LSQKPQVLGPEKQDGSCEASVSFEDVTVDFSREEWQQLDPAQRCL YRDVMLELYSHLFAVGYHIPNPEVIFRMLKEKEPR
241	ZN347_HUMAN	ALTQGQVTFRDVAIEFSQEEWTCLDPAQRTLYRDVMLENYRNLAS LGISCFDLSIISMLEQGKEPFTLESQVQIAGNPDG
242	ZNF25_HUMAN	NKFQGPVTLKDVIVEFTKEEWKLLTPAQRTLYKDVMLENYSHLVS VGYHVNKPNAVFKLKQGKEPWILEVEFPHRGFPED
243	ZN519_HUMAN	ELLTFRDVAIEFSPEEWKCLDPAQQNLYRDVMLENYRNLVSLAVY SYYNQGILPEQGIQDSFKKATLGRYGSCGLENICL
244	Z585B_HUMAN	SPQKSSALAPEDHGSSYEGSVSFRDVAIDFSREEWRHLDLSQRNL YRDVMLETYSHLLSVGYQVPKPEVVMLEQGKEPWA
245	ZIM3_HUMAN	NNSQGRVTFEDVTVNFTQGEWQRLNPEQRNLYRDVMLENYSNLVS VGQGETTKPDVILRLEQGKEPWLEEEEVLGSGRAE
246	ZN517_HUMAN	AMALPMPGPQEAVVFEDVAVYFTRIEWSCLAPDQQALYRDVMLEN YGNLASLGFLVAKPALISLLEQGEEPGALILQVAE
247	ZN846_HUMAN	DSSQHLVTFEDVAVDFTQEEWTLLDQAQRDLYRDVMLENYKNLII LAGSELFKRSLMSGLEQMEELRTGVTGVLQELDLQ
248	ZN230_HUMAN	TTFKEAVTFKDVAVFFTEEELGLLDPAQRKLYQDVMLENFTNLLS VGHQPFHPFHFLREEKFWMMETATQREGNSGGKTI
249	ZNF66_HUMAN	GPLQFRDVAIEFSLEEWHCLDMAQRNLYRDVMLENYRNLVFLGIV VSKPDLITHLEQGKKPSTMQRHEMVANPSVLCSHF
250	ZFP1_HUMAN	NKSQGSVSFTDVTVDFTQEEWEQLDPSQRILYMDVMLENYSNLLS VEVWKADDQMERDHRNPDEQARQFLILKNQTPIEE
251	ZN713_HUMAN	EEEEMNDGSQMVRSQESLTFQDVAVDFTREEWDQLYPAQKNLYRD VMLENYRNLVALGYQLCKPEVIAQLELEEEWVIER
252	ZN816_HUMAN	EEATKKSKEKEPGMALPQGRLTFRDVAIEFSLEEWKCLNPAQRAL YRAVMLENYRNLEFVDSSLKSMMEFSSTRHSITGE
253	ZN426_HUMAN	EKTPAGRIVADCLTDCYQDSVTFDDVAVDFTQEEWTLLDSTQRSL YSDVMLENYKNLATVGGQIIKPSLISWLEQEESRT
254	ZN674_HUMAN	AMSQESLTFKDVFVDFTLEEWQQLDSAQKNLYRDVMLENYSHLVS VGHLVGKPDVIFRLGPGDESWMADGGTPVRTCAGE
255	ZN627_HUMAN	DSVAFEDVAVNFTLEEWALLDPSQKNLYRDVMRETFRNLASVGKQ WEDQNIEDPFKIPRRNISHIPERLCESKEGGQGEE
256	ZNF20_HUMAN	MFQDSVAFEDVAVSFTQEEWALLDPSQKNLYRDVMQETFKNLTSV GKTWKVQNIEDEYKNPRRNLSLMREKLCESKESHH
257	Z587B_HUMAN	AVVATLRLSAQGTVTFEDVAVKFTQEEWNLLSEAQRCLYRDVTLE NLALMSSLGCWCGVEDEAAPSKQSIYIQRETQVRT
258	ZN316_HUMAN	EEEEEDEDEDDLLTAGCQELVTFEDVAVYFSLEEWERLEADQRGL YQEVMQENYGILVSLGYPIPKPDLIFRLEQGEEPW
259	ZN233_HUMAN	TKFQEMVTFKDVAVVFTREELGLLDLAQRKLYQDVMLENFRNLLS VGYQPFKLDVILQLGKEDKLRMMETEIQGDGCSGH
260	ZN611_HUMAN	EEAAQKRKGKEPGMALPQGRLTFRDVAIEFSLAEWKCLNPSQRAL YREVMLENYRNLEAVDISSKCMMKEVLSTGQGNTE
261	ZN556_HUMAN	DTVVFEDVVVDFTLEEWALLNPAQRKLYRDVMLETFKHLASVDNE AQLKASGSISQQDTSGEKLSLKQKIEKFTRKNIWA

SEQ	Description	Sequence
262	ZN234_HUMAN	TTFKEGLTFKDVAVVFTEEELGLLDPVQRNLYQDVMLENFRNLLS VGHHPFKHDVFLLEKEKKLDIMKTATQRKGKSADK
263	ZN560_HUMAN	SALQQEFWKIQTSNGIQMDLVTFDSVAVEFTQEEWTLLDPAQRNL YSDVMLENYKNLSSVGYQLFKPSLISWLEEEEELS
264	ZNF77_HUMAN	DCVIFEEVAVNFTPEEWALLDHAQRSLYRDVMLETCRNLASLDCY IYVRTSGSSSQRDVFGNGISNDEEIVKFTGSDSWS
265	ZN682_HUMAN	ELLTFRDVTIEFSLEEWEFLNPAQQSLYRKVMLENYRNLVSLGLT VSKPELISRLEQRQEPWNVKRHETIAKPPAMSSHY
266	ZN614_HUMAN	IKTQESLTLEDVAVEFSWEEWQLLDTAQKNLYRDVMVENYNHLVS LGYQTSKPDVLSKLAHGQEPWTTDAKIQNKNCPGI
267	ZN785_HUMAN	PAHVPGEAGPRRTRESRPGAVSFADVAVYFSPEEWECLRPAQRAL YRDVMRETFGHLGALGFSVPKPAFISWVEGEVEAW
268	ZN445_HUMAN	GCPGDQVTPTRSLTAQLQETMTFKDVEVTFSQDEWGWLDSAQRNL YRDVMLENYRNMASLVGPFTKPALISWLEAREPWG
269	ZFP30_HUMAN	ARDLVMFRDVAVDFSQEEWECLNSYQRNLYRDVILENYSNLVSLA GCSISKPDVITLLEQGKEPWMVVRDEKRRWTLDLE
270	ZN225_HUMAN	TTLKEAVTFKDVAVVFTEEELRLLDLAQRKLYREVMLENFRNLLS VGHQSLHRDTFHFLKEEKFWMMETATQREGNLGGK
271	ZN551_HUMAN	SPPSPRSSMAAVALRDSAQGMTFEDVAIYFSQEEWELLDESQRFL YCDVMLENFAHVTSLGYCHGMENEAIASEQSVSIQ
272	ZN610_HUMAN	DEEAQKRKAKESGMALPQGRLTFMDVAIEFSQEEWKSLDPGQRAL YRDVMLENYRNLVFLGICLPDLSIISMLKQRREPL
273	ZN528_HUMAN	ALTQGPLKFMDVAIEFSQEEWKCLDPAQRTLYRDVMLENYRNLVS LGICLPDLSVTSMLEQKRDPWTLQSEEKIANDPDG
274	ZN284_HUMAN	TMFKEAVTFKDVAVVFTEEELGLLDVSQRKLYRDVMLENFRNLLS VGHQLSHRDTFHFQREEKFWIMETATQREGNSGGK
275	ZN418_HUMAN	QGTVAFEDVAVNFSQEEWSLLSEVQRCLYHDVMLENWVLISSLGC WCGSEDEEAPSKKSISIQRVSQVSTPGAGVSPKKA
276	MPP8_HUMAN	AEAFGDSEEDGEDVFEVEKILDMKTEGGKVLYKVRWKGYTSDDDT WEPEIHLEDCKEVLLEFRKKIAENKAKAVRKDIQR
277	ZN490_HUMAN	VLQMQNSEHHGQSIKTQTDSISLEDVAVNFTLEEWALLDPGQRNI YRDVMRATFKNLACIGEKWKDQDIEDEHKNQGRNL
278	ZN805_HUMAN	AMALTDPAQVSVTFDDVAVTFTQEEWGQLDLAQRTLYQEVMLENC GLLVSLGCPVPRPELIYHLEHGQEPWTRKEDLSQG
279	Z780B_HUMAN	VHGSVTFRDVAIDFSQEEWECLQPDQRTLYRDVMLENYSHLISLG SSISKPDVITLLEQEKEPWIVVSKETSRWYPDLES
280	ZN763_HUMAN	DPVACEDVAVNFTQEEWALLDISQRKLYREVMLETFRNLTSIGKK WKDQNIEYEYQNPRRNFRSLIEGNVNEIKEDSHCG
281	ZN285_HUMAN	IKFQERVTFKDVAVVFTKEELALLDKAQINLYQDVMLENFRNLML VRDGIKNNILNLQAKGLSYLSQEVLHCWQIWKQRI
282	ZNF85_HUMAN	GPLTFRDVAIEFSLKEWQCLDTAQRNLYRNVMLENYRNLVFLGIT VSKPDLITCLEQGKEAWSMKRHEIMVAKPTVMCSH
283	ZN223_HUMAN	TMSKEAVTFKDVAVVFTEEELGLLDLAQRKLYRDVMLENFRNLLS VGHQPFHRDTFHFLREEKFWMMDIATQREGNSGGK
284	ZNF90_HUMAN	GPLEFRDVAIEFSLEEWHCLDTAQQNLYRDVMLENYRHLVFLGIV VTKPDLITCLEQGKKPFTVKRHEMIAKSPVMCFHF
285	ZN557_HUMAN	GHTEGGELVNELLKSWLKGLVTFEDVAVEFTQEEWALLDPAQRTL YRDVMLENCRNLASLGNQVDKPRLISQLEQEDKVM
286	ZN425_HUMAN	AEPASVTVTFDDVALYFSEQEWEILEKWQKQMYKQEMKTNYETLD SLGYAFSKPDLITWMEQGRMLLISEQGCLDKTRRT
287	ZN229_HUMAN	HSQASAISQDREEKIMSQEPLSFKDVAVVFTEEELELLDSTQRQL YQDVMQENFRNLLSVGERNPLGDKNGKDTEYIQDE
288	ZN606_HUMAN	GSLEEGRRATGLPAAQVQEPVTFKDVAVDFTQEEWGQLDLVQRTL YRDVMLETYGHLLSVGNQIAKPEVISLLEQGEEPW

SEQ	Description	Sequence
289	ZN155_HUMAN	TTFKEAVTFKDVAVVFTEEELGLLDPAQRKLYRDVMLENFRNLLS VGHQPFHQDTCHFLREEKFWMMGTATQREGNSGGK
290	ZN222_HUMAN	AKLYEAVTFKDVAVIFTEEELGLLDPAQRKLYRDVMLENFRNLLS VGGKIQTEMETVPEAGTHEEFSCKQIWEQIASDLT
291	ZN442_HUMAN	RSDLFLPDSQTNEERKQYDSVAFEDVAVNFTQEEWALLGPSQKSL YRDVMWETIRNLDCIGMKWEDTNIEDQHRNPRRSL
292	ZNF91_HUMAN	PGTPGSLEMGLLTFRDVAIEFSPEEWQCLDTAQQNLYRNVMLENY RNLAFLGIALSKPDLITYLEQGKEPWNMKQHEMVD
293	ZN135_HUMAN	TPGVRVSTDPEQVTFEDVVVGFSQEEWGQLKPAQRTLYRDVMLDT FRLLVSVGHWLPKPNVISLLEQEAELWAVESRLPQ
294	ZN778_HUMAN	EQTQAAGMVAGWLINCYQDAVTFDDVAVDFTQEEWTLLDPSQRDL YRDVMLENYENLASVEWRLKTKGPALRQDRSWFRA
295	RYBP_HUMAN	PSEANSIQSANATTKTSETNHTSRPRLKNVDRSTAQQLAVTVGNV TVIITDFKEKTRSSSTSSSTVTSSAGSEQQNQSSS
296	ZN534_HUMAN	ALTQGQLSFSDVAIEFSQEEWKCLDPGQKALYRDVMLENYRNLVS LGEDNVRPEACICSGICLPDLSVTSMLEQKRDPWT
297	ZN586_HUMAN	AAAAALRAPAQSSVTFEDVAVNFSLEEWSLLNEAQRCLYRDVMLE TLTLISSLGCWHGGEDEAAPSKQSTCIHIYKDQGG
298	ZN567_HUMAN	AQGSVSFNDVTVDFTQEEWQHLDHAQKTLYMDVMLENYCHLISVG CHMTKPDVILKLERGEEPWTSFAGHTCLEENWKAE
299	ZN440_HUMAN	DPVAFKDVAVNFTQEEWALLDISQRKLYREVMLETFRNLTSLGKR WKDQNIEYEHQNPRRNFRSLIEEKVNEIKDDSHCG
300	ZN583_HUMAN	SKDLVTFGDVAVNFSQEEWEWLNPAQRNLYRKVMLENYRSLVSLG VSVSKPDVISLLEQGKEPWMVKKEGTRGPCPDWEY
301	ZN441_HUMAN	DSVAFEDVAINFTCEEWALLGPSQKSLYRDVMQETIRNLDCIGMI WQNHDIEEDQYKDLRRNLRCHMVERACEIKDNSQC
302	ZNF43_HUMAN	GPLTFMDVAIEFCLEEWQCLDIAQQNLYRNVMLENYRNLVFLGIA VSKPDLITCLEQEKEPWEPMRRHEMVAKPPVMCSH
303	CBX5_HUMAN	QSNDIARGFERGLEPEKIIGATDSCGDLMFLMKWKDTDEADLVLA KEANVKCPQIVIAFYEERLTWHAYPEDAENKEKET
304	ZN589_HUMAN	ALPAKDSAWPWEEKPRYLGPVTFEDVAVLFTEAEWKRLSLEQRNL YKEVMLENLRNLVSLAESKPEVHTCPSCPLAFGSQ
305	ZNF10_HUMAN	DAKSLTAWSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLEN YKNLVSLGYQLTKPDVILRLEKGEEPWLVEREIHQ
306	ZN563_HUMAN	DAVAFEDVAVNFTQEEWALLGPSQKNLYRYVMQETIRNLDCIRMI WEEQNTEDQYKNPRRNLRCHMVERFSESKDSSQCG
307	ZN561_HUMAN	EKTKVERMVEDYLASGYQDSVTFDDVAVDFTPEEWALLDTTEKYL YRDVMLENYMNLASVEWEIQPRTKRSSLQQGFLKN
308	ZN136_HUMAN	DSVAFEDVDVNFTQEEWALLDPSQKNLYRDVMWETMRNLASIGKK WKDQNIKDHYKHRGRNLRSHMLERLYQTKDGSQRG
309	ZN630_HUMAN	IESQEPVTFEDVAVDFTQEEWQQLNPAQKTLHRDVMLETYNHLVS VGCSGIKPDVIFKLEHGKDPWIIESELSRWIYPDR
310	ZN527_HUMAN	AVGLCKAMSQGLVTFRDVALDFSQEEWEWLKPSQKDLYRDVMLEN YRNLVWLGLSISKPNMISLLEQGKEPWMVERKMSQ
311	ZN333_HUMAN	DKVEEEAMAPGLPTACSQEPVTFADVAVVFTPEEWVFLDSTQRSL YRDVMLENYRNLASVADQLCKPNALSYLEERGEQW
312	Z324B_HUMAN	TFEDVAVYFSQEEWGLLDTAQRALYRHVMLENFTLVTSLGLSTSR PRVVIQLERGEEPWVPSGKDMTLARNTYGRLNSGS
313	ZN786_HUMAN	AEPPRLPLTFEDVAIYFSEQEWQDLEAWQKELYKHVMRSNYETLV SLDDGLPKPELISWIEHGGEPFRKWRESQKSGNII
314	ZN709_HUMAN	DSVVFEDVAVNFTQEEWALLGPSQKKLYRDVMQETFVNLASIGEN WEEKNIEDHKNQGRKLRSHMVERLCERKEGSQFGE
315	ZN792_HUMAN	AAAALRDPAQGCVTFEDVTIYFSQEEWVLLDEAQRLLYCDVMLEN FALIASLGLISFRSHIVSQLEMGKEPWVPDSVDMT

SEQ	Description	Sequence
316	ZN599_HUMAN	AAPALALVSFEDVVVTFTGEEWGHLDLAQRTLYQEVMLETCRLLV SLGHPVPKPELIYLLEHGQELWTVKRGLSQSTCAG
317	ZN613_HUMAN	IKSQESLTLEDVAVEFTWEEWQLLGPAQKDLYRDVMLENYSNLVS VGYQASKPDALFKLEQGEPWTVENEIHSQICPEIK
318	ZF69B_HUMAN	GESLESRVTLGSLTAESQELLTFKDVSVDFTQEEWGQLAPAHRNL YREVMLENYGNLVSVGCQLSKPGVISQLEKGEEPW
319	ZN799_HUMAN	ASVALEDVAVNFTREEWALLGPCQKNLYKDVMQETIRNLDCVGMK WKDQNIEDQYRYPRKNLRCRMLERFVESKDGTQCG
320	ZN569_HUMAN	TESQGTVTFKDVAIDFTQEEWKRLDPAQRKLYRNVMLENYNNLIT VGYPFTKPDVIFKLEQEEEPWVMEEEVLRRHWQGE
321	ZN564_HUMAN	DSVASEDVAVNFTLEEWALLDPSQKKLYRDVMRETFRNLACVGKK WEDQSIEDWYKNQGRILRNHMEEGLSESKEYDQCG
322	ZN546_HUMAN	EETQGELTSSCGSKTMANVSLAFRDVSIDLSQEEWECLDAVQRDL YKDVMLENYSNLVSLGYTIPKPDVITLLEQEKEPW
323	ZFP92_HUMAN	AAILLTTRPKVPVSFEDVSVYFTKTEWKLLDLRQKVLYKRVMLEN YSHLVSLGFSFSKPHLISQLERGEGPWVADIPRTW
324	YAF2_HUMAN	KDKVEKEKSEKETTSKKNSHKKTRPRLKNVDRSSAQHLEVTVGDL TVIITDFKEKTKSPPASSAASADQHSQSGSSSDNT
325	ZN723_HUMAN	GPLTFTDVAIKFSLEEWQFLDTAQQNLYRDVMLENYRNLVFLGVG VSKPDLITCLEQGKEPWNMKRHKMVAKPPVVCSHF
326	ZNF34_HUMAN	RKPNPQAMAALFLSAPPQAEVTFEDVAVYLSREEWGRLGPAQRGL YRDVMLETYGNLVSLGVGPAGPKPGVISQLERGDE
327	ZN439_HUMAN	LSLSPILLYTCEMFQDPVAFKDVAVNFTQEEWALLDISQKNLYRE VMLETFWNLTSIGKKWKDQNIEYEYQNPRRNFRSV
328	ZFP57_HUMAN	AAGEPRSLLFFQKPVTFEDVAVNFTQEEWDCLDASQRVLYQDVMS ETFKNLTSVARIFLHKPELITKLEQEEEQWRETRV
329	ZNF19_HUMAN	AAMPLKAQYQEMVTFEDVAVHFTKTEWTGLSPAQRALYRSVMLEN FGNLTALGYPVPKPALISLLERGDMAWGLEAQDDP
330	ZN404_HUMAN	ARVPLTFSDVAIDFSQEEWEYLNSDQRDLYRDVMLENYTNLVSLD FNFTTESNKLSSEKRNYEVNAYHQETWKRNKTFNL
331	ZN274_HUMAN	ASRLPTAWSCEPVTFEDVTLGFTPEEWGLLDLKQKSLYREVMLEN YRNLVSVEHQLSKPDVVSQLEEAEDFWPVERGIPQ
332	CBX3_HUMAN	SKKKRDAADKPRGFARGLDPERIIGATDSSGELMFLMKWKDSDEA DLVLAKEANMKCPQIVIAFYEERLTWHSCPEDEAQ
333	ZNF30_HUMAN	AHKYVGLQYHGSVTFEDVAIAFSQQEWESLDSSQRGLYRDVMLEN YRNLVSMGHSRSKPHVIALLEQWKEPEVTVRKDGR
334	ZN250_HUMAN	AAARLLPVPAGPQPLSFQAKLTFEDVAVLLSQDEWDRLCPAQRGL YRNVMMETYGNVVSLGLPGSKPDIISQLERGEDPW
335	ZN570_HUMAN	AVGLLKAMYQELVTFRDVAVDFSQEEWDCLDSSQRHLYSNVMLEN YRILVSLGLCFSKPSVILLLEQGKAPWMVKRELTK
336	ZN675_HUMAN	GLLTFRDVAIEFSLEEWQCLDTAQRNLYKNVILENYRNLVFLGIA VSKQDLITCLEQEKEPLTVKRHEMVNEPPVMCSHF
337	ZN695_HUMAN	GLLAFRDVALEFSPEEWECLDPAQRSLYRDVMLENYRNLISLGED SFNMQFLFHSLAMSKPELIICLEARKEPWNVNTEK
338	ZN548_HUMAN	NLTEGRVVFEDVAIYFSQEEWGHLDEAQRLLYRDVMLENLALLSS LGSWHGAEDEEAPSQQGFSVGVSEVTASKPCLSSQ
339	ZN132_HUMAN	GPAQHTSWPCGSAVPTLKSMVTFEDVAVYFSQEEWELLDAAQRHL YHSVMLENLELVTSLGSWHGVEGEGAHPKQNVSVE
340	ZN738_HUMAN	SGYPGAERNLLEYSYFEKGPLTFRDVVIEFSQEEWQCLDTAQQDL YRKVMLENFRNLVFLGIDVSKPDLITCLEQGKDPW
341	ZN420_HUMAN	ARKLVMFRDVAIDFSQEEWECLDSAQRDLYRDVMLENYSNLVSLD LPSRCASKDLSPEKNTYETELSQWEMSDRLENCDL
342	ZN626_HUMAN	GPLQFRDVAIEFSLEEWHCLDTAQRNLYRNVMLENYSNLVFLGIT VSKPDLITCLEQGRKPLTMKRNEMIAKPSVMCSHF

SEQ	Description	Sequence
343	ZN559_HUMAN	VAGWLTNYSQDSVTFEDVAVDFTQEEWTLLDQTQRNLYRDVMLEN YKNLVAVDWESHINTKWSAPQQNFLQGKTSSVVEM
344	ZN460_HUMAN	AAAWMAPAQESVTFEDVAVTFTQEEWGQLDVTQRALYVEVMLETC GLLVALGDSTKPETVEPIPSHLALPEEVSLQEQLA
345	ZN268_HUMAN	VLEWLFISQEQPKITKSWGPLSFMDVFVDFTWEEWQLLDPAQKCL YRSVMLENYSNLVSLGYQHTKPDIIFKLEQGEELC
346	ZN304_HUMAN	AAAVLMDRVQSCVTFEDVFVYFSREEWELLEEAQRFLYRDVMLEN FALVATLGFWCEAEHEAPSEQSVSVEGVSQVRTAE
347	ZIM2_HUMAN	AGSQFPDFKHLGTFLVFEELVTFEDVLVDFSPEELSSLSAAQRNL YREVMLENYRNLVSLGHQFSKPDIISRLEEEESYA
348	ZN605_HUMAN	IQSQISFEDVAVDFTLEEWQLLNPTQKNLYRDVMLENYSNLVFLE VWLDNPKMWLRDNQDNLKSMERGHKYDVFGKIFNS
349	ZN844_HUMAN	DLVAFEDVAVNFTQEEWSLLDPSQKNLYREVMQETLRNLASIGEK WKDQNIEDQYKNPRNNLRSLLGERVDENTEENHCG
350	SUMO5_HUMAN	KDEDIKLRVIGQDSSEIHFKVKMTTPLKKLKKSYCQRQGVPVNSL RFLFEGQRIADNHTPEELGMEEEDVIEVYQEQIGG
351	ZN101_HUMAN	DSVAFEDVAVNFTQEEWALLSPSQKNLYRDVTLETFRNLASVGIQ WKDQDIENLYQNLGIKLRSLVERLCGRKEGNEHRE
352	ZN783_HUMAN	RNFWILRLPPGSKGEAPKVPVTFDDVAVYFSELEWGKLEDWQKEL YKHVMRGNYETLVSLDYAISKPDILTRIERGEEPC
353	ZN417_HUMAN	AAAAPRRPTQQGTVTFEDVAVNFSQEEWCLLSEAQRCLYRDVMLE NLALISSLGCWCGSKDEEAPCKQRISVQRESQSRT
354	ZN182_HUMAN	SGEDSGSFYSWQKAKREQGLVTFEDVAVDFTQEEWQYLNPPQRTL YRDVMLETYSNLVFVGQQVTKPNLILKLEVEECPA
355	ZN823_HUMAN	DSVAFEDVAVNFTQEEWALLGPSQKSLYRNVMQETIRNLDCIEMK WEDQNIGDQCQNAKRNLRSHTCEIKDDSQCGETFG
356	ZN177_HUMAN	AAGWLTTWSQNSVTFQEVAVDFSQEEWALLDPAQKNLYKDVMLEN FRNLASVGYQLCRHSLISKVDQEQLKTDERGILQG
357	ZN197_HUMAN	ENPRNQLMALMLLTAQPQELVMFEEVSVCFTSEEWACLGPIQRAL YWDVMLENYGNVTSLEWETMTENEEVTSKPSSSQR
358	ZN717_HUMAN	LETYNSLVSLQELVSFEEVAVHFTWEEWQDLDDAQRTLYRDVMLE TYSSLVSLGHCITKPEMIFKLEQGAEPWIVEETPN
359	ZN669_HUMAN	RHFRRPEPCREPLASPIQDSVAFEDVAVNFTQEEWALLDSSQKNL YREVMQETCRNLASVGSQWKDQNIEDHFEKPGKDI
360	ZN256_HUMAN	AAAELTAPAQGIVTFEDVAVYFSWKEWGLLDEAQKCLYHDVMLEN LTLTTSLGGSGAGDEEAPYQQSTSPQRVSQVRIPK
361	ZN251_HUMAN	AATFQLPGHQEMPLTFQDVAVYFSQAEGRQLGPQQRALYRDVMLE NYGNVASLGFPVPKPELISQLEQGKELWVLNLLGA
362	CBX4_HUMAN	RSEAGEPPSSLQVKPETPASAAVAVAAAAAPTTTAEKPPAEAQDE PAESLSEFKPFFGNIIITDVTANCLTVTFKEYVTV
363	PCGF2_HUMAN	HRTTRIKITELNPHLMCALCGGYFIDATTIVECLHSFCKTCIVRY LETNKYCPMCDVQVHKTRPLLSIRSDKTLQDIVYK
364	CDY2_HUMAN	ASQEFEVEAIVDKRQDKNGNTQYLVRWKGYDKQDDTWEPEQHLMN CEKCVHDFNRRQTEKQKKLTWTTTSRIFSNNARRR
365	CDYL2_HUMAN	ASGDLYEVERIVDKRKNKKGKWEYLIRWKGYGSTEDTWEPEHHLL HCEEFIDEFNGLHMSKDKRIKSGKQSSTSKLLRDS
366	HERC2_HUMAN	TLIRKADLENHNKDGGFWTVIDGKVYDIKDFQTQSLTGNSILAQF AGEDPVVALEAALQFEDTRESMHAFCVGQYLEPDQ
367	ZN562_HUMAN	EKTKIGTMVEDHRSNSYQDSVTFDDVAVEFTPEEWALLDTTQKYL YRDVMLENYMNLASVDFFFCLTSEWEIQPRTKRSS
368	ZN461_HUMAN	AHELVMFRDVAIDVSQEEWECLNPAQRNLYKEVMLENYSNLVSLG LSVSKPAVISSLEQGKEPWMVVREETGRWCPGTWK
369	Z324A_HUMAN	AFEDVAVYFSQEEWGLLDTAQRALYRRVMLDNFALVASLGLSTSR PRVVIQLERGEEPWVPSGTDTTLSRTTYRRRNPGS

SEQ	Description	Sequence
370	ZN766_HUMAN	AQLRRGHLTFRDVAIEFSQEEWKCLDPVQKALYRDVMLENYRNLV SLGICLPDLSIISMMKQRTEPWTVENEMKVAKNPD
371	ID2_HUMAN	SDHSLGISRSKTPVDDPMSLLYNMNDCYSKLKELVPSIPQNKKVS KMEILQHVIDYILDLQIALDSHPTIVSLHHQRPGQ
372	TOX_HUMAN	KDPNEPQKPVSAYALFFRDTQAAIKGQNPNATFGEVSKIVASMWD GLGEEQKQVYKKKTEAAKKEYLKQLAAYRASLVSK
373	ZN274_HUMAN	QEEKQEDAAICPVTVLPEEPVTFQDVAVDFSREEWGLLGPTQRTE YRDVMLETFGHLVSVGWETTLENKELAPNSDIPEE
374	SCMH1_HUMAN	DASRLSGRDPSSWTVEDVMQFVREADPQLGPHADLFRKHEIDGKA LLLLRSDMMMKYMGLKLGPALKLSYHIDRLKQGKF
375	ZN214_HUMAN	AVTFEDVTIIFTWEEWKFLDSSQKRLYREVMWENYTNVMSVENWN ESYKSQEEKFRYLEYENFSYWQGWWNAGAQMYENQ
376	CBX7_HUMAN	ELSAIGEQVFAVESIRKKRVRKGKVEYLVKWKGWPPKYSTWEPEE HILDPRLVMAYEEKEERDRASGYRKRGPKPKRLLL
377	ID1_HUMAN	GGAGARLPALLDEQQVNVLLYDMNGCYSRLKELVPTLPQNRKVSK VEILQHVIDYIRDLQLELNSESEVGTPGGRGLPVR
378	CREM_HUMAN	VVMAASPGSLHSPQQLAEEATRKRELRLMKNREAAKECRRRKKEY VKCLESRVAVLEVQNKKLIEELETLKDICSPKTDY
379	SCX_HUMAN	GGGPGGRPGREPRQRHTANARERDRTNSVNTAFTALRTLIPTEPA DRKLSKIETLRLASSYISHLGNVLLAGEACGDGQP
380	ASCL1_HUMAN	SGFGYSLPQQQPAAVARRNERERNRVKLVNLGFATLREHVPNGAA NKKMSKVETLRSAVEYIRALQQLLDEHDAVSAAFQ
381	ZN764_HUMAN	APLPPRDPNGAGPEWREPGAVSFADVAVYFCREEWGCLRPAQRAL YRDVMRETYGHLSALGIGGNKPALISWVEEEAELW
382	SCML2_HUMAN	KQGFSKDPSTWSVDEVIQFMKHTDPQISGPLADLFRQHEIDGKAL FLLKSDVMMKYMGLKLGPALKLCYYIEKLKEGKYS
383	TWST1_HUMAN	SGGGSPQSYEELQTQRVMANVRERQRTQSLNEAFAALRKIIPTLP SDKLSKIQTLKLAARYIDFLYQVLQSDELDSKMAS
384	CREB1_HUMAN	IAPGVVMASSPALPTQPAEEAARKREVRLMKNREAARECRRKKKE YVKCLENRVAVLENQNKTLIEELKALKDLYCHKSD
385	TERF1_HUMAN	SRIPVSKSQPVTPEKHRARKRQAWLWEEDKNLRSGVRKYGEGNWS KILLHYKFNNRTSVMLKDRWRTMKKLKLISSDSED
386	ID3_HUMAN	SLAIARGRGKGPAAEEPLSLLDDMNHCYSRLRELVPGVPRGTQLS QVEILQRVIDYILDLQVVLAEPAPGPPDGPHLPIQ
387	CBX8_HUMAN	GSGPPSSGGGLYRDMGAQGGRPSLIARIPVARILGDPEEESWSPS LTNLEKVVVTDVTSNFLTVTIKESNTDQGFFKEKR
388	CBX4_HUMAN	ELPAVGEHVFAVESIEKKRIRKGRVEYLVKWRGWSPKYNTWEPEE NILDPRLLIAFQNRERQEQLMGYRKRGPKPKPLVV
389	GSX1_HUMAN	VDSSSNQLPSSKRMRTAFTSTQLLELEREFASNMYLSRLRRIEIA TYLNLSEKQVKIWFQNRRVKHKKEGKGSNHRGGGG
390	NKX22_HUMAN	TPGGGGDAGKKRKRRVLFSKAQTYELERRFRQQRYLSAPEREHLA SLIRLTPTQVKIWFQNHRYKMKRARAEKGMEVTPL
391	ATF1_HUMAN	QTVVMTSPVTLTSQTTKTDDPQLKREIRLMKNREAARECRRKKKE YVKCLENRVAVLENQNKTLIEELKTLKDLYSNKSV
392	TWST2_HUMAN	KGSPSAQSFEELQSQRILANVRERQRTQSLNEAFAALRKIIPTLP SDKLSKIQTLKLAARYIDFLYQVLQSDEMDNKMTS
393	ZNF17_HUMAN	NLTEDYMVFEDVAIHFSQEEWGILNDVQRHLHSDVMLENFALLSS VGCWHGAKDEEAPSKQCVSVGVSQVTTLKPALSTQ
394	TOX3_HUMAN	KDPNEPQKPVSAYALFFRDTQAAIKGQNPNATFGEVSKIVASMWD SLGEEQKQVYKRKTEAAKKEYLKALAAYRASLVSK
395	TOX4_HUMAN	KDPNEPQKPVSAYALFFRDTQAAIKGQNPNATFGEVSKIVASMWD SLGEEQKQVYKRKTEAAKKEYLKALAAYKDNQECQ
396	ZMYM3_HUMAN	LDGSTWDFCSEDCKSKYLLWYCKAARCHACKRQGKLLETIHWRGQ IRHFCNQQCLLRFYSQQNQPNLDTQSGPESLLNSQ

SEQ	Description	Sequence
397	I2BP1_HUMAN	ASVQASRRQWCYLCDLPKMPWAMVWDFSEAVCRGCVNFEGADRIE LLIDAARQLKRSHVLPEGRSPGPPALKHPATKDLA
398	RHXF1_HUMAN	MEGPQPENMQPRTRRTKFTLLQVEELESVFRHTQYPDVPTRRELA ENLGVTEDKVRVWFKNKRARCRRHQRELMLANELR
399	SSX2_HUMAN	PKIMPKKPAEEGNDSEEVPEASGPQNDGKELCPPGKPTTSEKIHE RSGPKRGEHAWTHRLRERKQLVIYEEISDPEEDDE
400	I2BPL_HUMAN	SAAQVSSSRRQSCYLCDLPRMPWAMIWDFSEPVCRGCVNYEGADR IEFVIETARQLKRAHGCFQDGRSPGPPPPVGVKTV
401	ZN680_HUMAN	PGPPGSLEMGPLTFRDVAIEFSLEEWQCLDTAQRNLYRKVMFENY RNLVFLGIAVSKPHLITCLEQGKEPWNRKRQEMVA
402	CBX1_HUMAN	NKKKVEEVLEEEEEEYVVEKVLDRRVVKGKVEYLLKWKGFSDEDN TWEPEENLDCPDLIAEFLQSQKTAHETDKSEGGKR
403	TRI68_HUMAN	LANVVEKVRLLRLHPGMGLKGDLCERHGEKLKMFCKEDVLIMCEA CSQSPEHEAHSVVPMEDVAWEYKWELHEALEHLKK
404	HXA13_HUMAN	VVSHPSDASSYRRGRKKRVPYTKVQLKELEREYATNKFITKDKRR RISATTNLSERQVTIWFQNRRVKEKKVINKLKTTS
405	PHC3_HUMAN	ENSDLLPVAQTEPSIWTVDDVWAFIHSLPGCQDIADEFRAQEIDG QALLLLKEDHLMSAMNIKLGPALKICARINSLKES
406	TCF24_HUMAN	AGPGGGSRSGSGRPAAANAARERSRVQTLRHAFLELQRTLPSVPP DTKLSKLDVLLLATTYIAHLTRSLQDDAEAPADAG
407	CBX3_HUMAN	QNGKSKKVEEAEPEEFVVEKVLDRRVVNGKVEYFLKWKGFTDADN TWEPEENLDCPELIEAFLNSQKAGKEKDGTKRKSL
408	HXB13_HUMAN	QHPPDACAFRRGRKKRIPYSKGQLRELEREYAANKFITKDKRRKI SAATSLSERQITIWFQNRRVKEKKVLAKVKNSATP
409	HEY1_HUMAN	SMSPTTSSQILARKRRRGIIEKRRRDRINNSLSELRRLVPSAFEK QGSAKLEKAEILQMTVDHLKMLHTAGGKGYFDAHA
410	PHC2_HUMAN	LVGMGHHFLPSEPTKWNVEDVYEFIRSLPGCQEIAEEFRAQEIDG QALLLLKEDHLMSAMNIKLGPALKIYARISMLKDS
411	ZNF81_HUMAN	PANEDAPQPGEHGSACEVSVSFEDVTVDFSREEWQQLDSTQRRLY QDVMLENYSHLLSVGFEVPKPEVIFKLEQGEGPWT
412	FIGLA_HUMAN	GYSSTENLQLVLERRRVANAKERERIKNLNRGFARLKALVPFLPQ SRKPSKVDILKGATEYIQVLSDLLEGAKDSKKQDP
413	SAM11_HUMAN	EEAPAPEDVTKWTVDDVCSFVGGLSGCGEYTRVFREQGIDGETLP LLTEEHLLTNMGLKLGPALKIRAQVARRLGRVFYV
414	KMT2B_HUMAN	GGTLAHTPRRSLPSHHGKKMRMARCGHCRGCLRVQDCGSCVNCLD KPKFGGPNTKKQCCVYRKCDKIEARKMERLAKKGR
415	HEY2_HUMAN	LNSPTTTSQIMARKKRRGIIEKRRRDRINNSLSELRRLVPTAFEK QGSAKLEKAEILQMTVDHLKMLQATGGKGYFDAHA
416	JDP2_HUMAN	QPVKSELDEEEERRKRRREKNKVAAARCRNKKKERTEFLQRESER LELMNAELKTQIEELKQERQQLILMLNRHRPTCIV
417	HXC13_HUMAN	LQPEVSSYRRGRKKRVPYTKVQLKELEKEYAASKFITKEKRRRIS ATTNLSERQVTIWFQNRRVKEKKVVSKSKAPHLHS
418	ASCL4_HUMAN	LPVPLDSAFEPAFLRKRNERERQRVRCVNEGYARLRDHLPRELAD KRLSKVETLRAAIDYIKHLQELLERQAWGLEGAAG
419	HHEX_HUMAN	SPFLQRPLHKRKGGQVRFSNDQTIELEKKFETQKYLSPPERKRLA KMLQLSERQVKTWFQNRRAKWRRLKQENPQSNKKE
420	HERC2_HUMAN	IAIATGSLHCVCCTEDGEVYTWGDNDEGQLGDGTTNAIQRPRLVA ALQGKKVNRVACGSAHTLAWSTSKPASAGKLPAQV
421	GSX2_HUMAN	GGSDASQVPNGKRMRTAFTSTQLLELEREFSSNMYLSRLRRIEIA TYLNLSEKQVKIWFQNRRVKHKKEGKGTQRNSHAG
422	BIN1_HUMAN	RLDLPPGFMFKVQAQHDYTATDTDELQLKAGDVVLVIPFQNPEEQ DEGWLMGVKESDWNQHKELEKCRGVFPENFTERVP
423	ETV7_HUMAN	GICKLPGRLRIQPALWSREDVLHWLRWAEQEYSLPCTAEHGFEMN GRALCILTKDDFRHRAPSSGDVLYELLQYIKTQRR

SEQ	Description	Sequence
424	ASCL3_HUMAN	PNYRGCEYSYGPAFTRKRNERERQRVKCVNEGYAQLRHHLPEEYL EKRLSKVETLRAAIKYINYLQSLLYPDKAETKNNP
425	PHC1_HUMAN	LHGINPVFLSSNPSRWSVEEVYEFIASLQGCQEIAEEFRSQEIDG QALLLLKEEHLMSAMNIKLGPALKICAKINVLKET
426	OTP_HUMAN	QAGQQQGQQKQKRHRTRFTPAQLNELERSFAKTHYPDIFMREELA LRIGLTESRVQVWFQNRRAKWKKRKKTTNVFRAPG
427	I2BP2_HUMAN	AAAVAVAAASRRQSCYLCDLPRMPWAMIWDFTEPVCRGCVNYEGA DRVEFVIETARQLKRAHGCFPEGRSPPGAAASAAA
428	VGLL2_HUMAN	FSSQTPASIKEEEGSPEKERPPEAEYINSRCVLFTYFQGDISSVV DEHFSRALSQPSSYSPSCTSSKAPRSSGPWRDCSF
429	HXA11_HUMAN	DKAGGSSGQRTRKKRCPYTKYQIRELEREFFFSVYINKEKRLQLS RMLNLTDRQVKIWFQNRRMKEKKINRDRLQYYSAN
430	PDLI4_HUMAN	GAPLSGLQGLPECTRCGHGIVGTIVKARDKLYHPECFMCSDCGLN LKQRGYFFLDERLYCESHAKARVKPPEGYDVVAVY
431	ASCL2_HUMAN	RRPATAETGGGAAAVARRNERERNRVKLVNLGFQALRQHVPHGGA SKKLSKVETLRSAVEYIRALQRLLAEHDAVRNALA
432	CDX4_HUMAN	TVQVTGKTRTKEKYRVVYTDHQRLELEKEFHCNRYITIQRKSELA VNLGLSERQVKIWFQNRRAKERKMIKKKISQFENS
433	ZN860_HUMAN	EEAAQKRKEKEPGMALPQGHLTFRDVAIEFSLEEWKCLDPTQRAL YRAMMLENYRNLHSVDISSKCMMKKFSSTAQGNTE
434	LMBL4_HUMAN	DIRASQVARWTVDEVAEFVQSLLGCEEHAKCFKKEQIDGKAFLLL TQTDIVKVMKIKLGPALKIYNSILMFRHSQELPEE
435	PDIP3_HUMAN	LSPLEGTKMTVNNLHPRVTEEDIVELFCVCGALKRARLVHPGVAE VVFVKKDDAITAYKKYNNRCLDGQPMKCNLHMNGN
436	NKX25_HUMAN	DNAERPRARRRRKPRVLFSQAQVYELERRFKQQRYLSAPERDQLA SVLKLTSTQVKIWFQNRRYKCKRQRQDQTLELVGL
437	CEBPB_HUMAN	SQVKSKAKKTVDKHSDEYKIRRERNNIAVRKSRDKAKMRNLETQH KVLELTAENERLQKKVEQLSRELSTLRNLFKQLPE
438	ISL1_HUMAN	KRDYIRLYGIKCAKCSIGFSKNDFVMRARSKVYHIECFRCVACSR QLIPGDEFALREDGLFCRADHDVVERASLGAGDPL
439	CDX2_HUMAN	SLGSQVKTRTKDKYRVVYTDHQRLELEKEFHYSRYITIRRKAELA ATLGLSERQVKIWFQNRRAKERKINKKKLQQQQQQ
440	PROP1_HUMAN	QGGQRGRPHSRRRHRTTFSPVQLEQLESAFGRNQYPDIWARESLA RDTGLSEARIQVWFQNRRAKQRKQERSLLQPLAHL
441	SIN3B_HUMAN	DALTYLDQVKIRFGSDPATYNGFLEIMKEFKSQSIDTPGVIRRVS QLFHEHPDLIVGFNAFLPLGYRIDIPKNGKLNIQS
442	SMBT1_HUMAN	RLHLDSNPLKWSVADVVRFIRSTDCAPLARIFLDQEIDGQALLLL TLPTVQECMDLKLGPAIKLCHHIERIKFAFYEQFA
443	HXC11_HUMAN	AKGAAPNAPRTRKKRCPYSKFQIRELEREFFFNVYINKEKRLQLS RMLNLTDRQVKIWFQNRRMKEKKLSRDRLQYFSGN
444	HXC10_HUMAN	TTGNWLTAKSGRKKRCPYTKHQTLELEKEFLFNMYLTRERRLEIS KTINLTDRQVKIWFQNRRMKLKKMNRENRIRELTS
445	PRS6A_HUMAN	YLVSNVIELLDVDPNDQEEDGANIDLDSQRKGKCAVIKTSTRQTY FLPVIGLVDAEKLKPGDLVGVNKDSYLILETLPTE
446	VSX1_HUMAN	KASPTLGKRKKRRHRTVFTAHQLEELEKAFSEAHYPDVYAREMLA VKTELPEDRIQVWFQNRRAKWRKREKRWGGSSVMA
447	NKX23_HUMAN	EESERPKPRSRRKPRVLFSQAQVFELERRFKQQRYLSAPEREHLA SSLKLTSTQVKIWFQNRRYKCKRQRQDKSLELGAH
448	MTG16_HUMAN	VVPGSRQEEVIDHKLTEREWAEEWKHLNNLLNCIMDMVEKTRRSL TVLRRCQEADREELNHWARRYSDAEDTKKGPAPAA
449	HMX3_HUMAN	ESPEKKPACRKKKTRTVFSRSQVFQLESTFDMKRYLSSSERAGLA ASLHLTETQVKIWFQNRRNKWKRQLAAELEAANLS
450	HMX1_HUMAN	RGGVGVGGGRKKKTRTVFSRSQVFQLESTFDLKRYLSSAERAGLA ASLQLTETQVKIWFQNRRNKWKRQLAAELEAASLS

SEQ	Description	Sequence
451	KIF22_HUMAN	ELLAHGRQKILDLLNEGSARDLRSLQRIGPKKAQLIVGWRELHGP FSQVEDLERVEGITGKQMESFLKANILGLAAGQRC
452	CSTF2_HUMAN	ESPYGETISPEDAPESISKAVASLPPEQMFELMKQMKLCVQNSPQ EARNMLLQNPQLAYALLQAQVVMRIVDPEIALKIL
453	CEBPE_HUMAN	AGPLHKGKKAVNKDSLEYRLRRERNNIAVRKSRDKAKRRILETQQ KVLEYMAENERLRSRVEQLTQELDTLRNLFRQIPE
454	DLX2_HUMAN	IRIVNGKPKKVRKPRTIYSSFQLAALQRRFQKTQYLALPERAELA ASLGLTQTQVKIWFQNRRSKFKKMWKSGEIPSEQH
455	ZMYM3_HUMAN	TVYQFCSPSCWTKFQRTSPEGGIHLSCHYCHSLFSGKPEVLDWQD QVFQFCCRDCCEDFKRLRGVVSQCEHCRQEKLLHE
456	PPARG_HUMAN	TMVDTEMPFWPTNFGISSVDLSVMEDHSHSFDIKPFTTVDFSSIS TPHYEDIPFTRTDPVVADYKYDLKLQEYQSAIKVE
457	PRIC1_HUMAN	GRHHAELLKPRCSACDEIIFADECTEAEGRHWHMKHFCCLECETV LGGQRYIMKDGRPFCCGCFESLYAEYCETCGEHIG
458	UNC4_HUMAN	DPDKESPGCKRRRTRTNFTGWQLEELEKAFNESHYPDVFMREALA LRLDLVESRVQVWFQNRRAKWRKKENTKKGPGRPA
459	BARX2_HUMAN	TEQPTPRQKKPRRSRTIFTELQLMGLEKKFQKQKYLSTPDRLDLA QSLGLTQLQVKTWYQNRRMKWKKMVLKGGQEAPTK
460	ALX3_HUMAN	SMELAKNKSKKRRNRTTFSTFQLEELEKVFQKTHYPDVYAREQLA LRTDLTEARVQVWFQNRRAKWRKRERYGKIQEGRN
461	TCF15_HUMAN	GGGGGAGPVVVVRQRQAANARERDRTQSVNTAFTALRTLIPTEPV DRKLSKIETVRLASSYIAHLANVLLLGDSADDGQP
462	TERA_HUMAN	IDDTVEGITGNLFEVYLKPYFLEAYRPIRKGDIFLVRGGMRAVEF KVVETDPSPYCIVAPDTVIHCEGEPIKREDEEESL
463	VSX2_HUMAN	SALNQTKKRKKRRHRTIFTSYQLEELEKAFNEAHYPDVYAREMLA MKTELPEDRIQVWFQNRRAKWRKREKCWGRSSVMA
464	HXD12_HUMAN	DGLPWGAAPGRARKKRKPYTKQQIAELENEFLVNEFINRQKRKEL SNRLNLSDQQVKIWFQNRRMKKKRVVLREQALALY
465	CDX1_HUMAN	GGGGSGKTRTKDKYRVVYTDHQRLELEKEFHYSRYITIRRKSELA ANLGLTERQVKIWFQNRRAKERKVNKKKQQQQQPP
466	TCF23_HUMAN	TRAGGLALGRSEASPENAARERSRVRTLRQAFLALQAALPAVPPD TKLSKLDVLVLAASYIAHLTRTLGHELPGPAWPPF
467	ALX1_HUMAN	KCDSNVSSSKKRRHRTTFTSLQLEELEKVFQKTHYPDVYVREQLA LRTELTEARVQVWFQNRRAKWRKRERYGQIQQAKS
468	HXA10_HUMAN	NAANWLTAKSGRKKRCPYTKHQTLELEKEFLFNMYLTRERRLEIS RSVHLTDRQVKIWFQNRRMKLKKMNRENRIRELTA
469	RX_HUMAN	LSEEEQPKKKHRRNRTTFTTYQLHELERAFEKSHYPDVYSREELA GKVNLPEVRVQVWFQNRRAKWRRQEKLEVSSMKLQ
470	CXXC5_HUMAN	HMAGLAEYPMQGELASAISSGKKKRKRCGMCAPCRRRINCEQCSS CRNRKTGHQICKFRKCEELKKKPSAALEKVMLPTG
471	SCML1_HUMAN	SITKHPSTWSVEAVVLFLKQTDPLALCPLVDLFRSHEIDGKALLL LTSDVLLKHLGVKLGTAVKLCYYIDRLKQGKCFEN
472	NFIL3_HUMAN	ACRRKREFIPDEKKDAMYWEKRRKNNEAAKRSREKRRLNDLVLEN KLIALGEENATLKAELLSLKLKFGLISSTAYAQEI
473	DLX6_HUMAN	EIRFNGKGKKIRKPRTIYSSLQLQALNHRFQQTQYLALPERAELA ASLGLTQTQVKIWFQNKRSKFKKLLKQGSNPHESD
474	MTG8_HUMAN	GLHGTRQEEMIDHRLTDREWAEEWKHLDHLLNCIMDMVEKTRRSL TVLRRCQEADREELNYWIRRYSDAEDLKKGGGSSS
475	CBX8_HUMAN	ELSAVGERVFAAEALLKRRIRKGRMEYLVKWKGWSQKYSTWEPEE NILDARLLAAFEEREREMELYGPKKRGPKPKTFLL
476	CEBPD_HUMAN	AREKSAGKRGPDRGSPEYRQRRERNNIAVRKSRDKAKRRNQEMQQ KLVELSAENEKLHQRVEQLTRDLAGLRQFFKQLPS
477	SEC13_HUMAN	SGGCDNLIKLWKEEEDGQWKEEQKLEAHSDWVRDVAWAPSIGLPT STIASCSQDGRVFIWTCDDASSNTWSPKLLHKFND

SEQ	Description	Sequence
478	FIP1_HUMAN	VKGVDLDAPGSINGVPLLEVDLDSFEDKPWRKPGADLSDYFNYGF NEDTWKAYCEKQKRIRMGLEVIPVTSTTNKITAED
479	ALX4_HUMAN	KADSESNKGKKRRNRTTFTSYQLEELEKVFQKTHYPDVYAREQLA MRTDLTEARVQVWFQNRRAKWRKRERFGQMQQVRT
480	LHX3_HUMAN	TAKQREAEATAKRPRTTITAKQLETLKSAYNTSPKPARHVREQLS SETGLDMRVVQVWFQNRRAKEKRLKKDAGRQRWGQ
481	PRIC2_HUMAN	GRHHAECLKPRCAACDEIIFADECTEAEGRHWHMKHFCCFECETV LGGQRYIMKEGRPYCCHCFESLYAEYCDTCAQHIG
482	MAGI3_HUMAN	IIGGDRPDEFLQVKNVLKDGPAAQDGKIAPGDVIVDINGNCVLGH THADVVQMFQLVPVNQYVNLTLCRGYPLPDDSEDP
483	NELL1_HUMAN	CCPECDTRVTSQCLDQNGHKLYRSGDNWTHSCQQCRCLEGEVDCW PLTCPNLSCEYTAILEGECCPRCVSDPCLADNITY
484	PRRX1_HUMAN	LNSEEKKKRKQRRNRTTFNSSQLQALERVFERTHYPDAFVREDLA RRVNLTEARVQVWFQNRRAKFRRNERAMLANKNAS
485	MTG8R_HUMAN	GLNGGYQDELVDHRLTEREWADEWKHLDHALNCIMEMVEKTRRSM AVLRRCQESDREELNYWKRRYNENTELRKTGTELV
486	RAX2_HUMAN	GPGEEAPKKKHRRNRTTFTTYQLHQLERAFEASHYPDVYSREELA AKVHLPEVRVQVWFQNRRAKWRRQERLESGSGAVA
487	DLX3_HUMAN	VRMVNGKPKKVRKPRTIYSSYQLAALQRRFQKAQYLALPERAELA AQLGLTQTQVKIWFQNRRSKFKKLYKNGEVPLEHS
488	DLX1_HUMAN	EVRFNGKGKKIRKPRTIYSSLQLQALNRRFQQTQYLALPERAELA ASLGLTQTQVKIWFQNKRSKFKKLMKQGGAALEGS
489	NKX26_HUMAN	GRSEQPKARQRRKPRVLFSQAQVLALERRFKQQRYLSAPEREHLA SALQLTSTQVKIWFQNRRYKCKRQRQDKSLELAGH
490	NAB1_HUMAN	LPRTLGELQLYRILQKANLLSYFDAFIQQGGDDVQQLCEAGEEEF LEIMALVGMASKPLHVRRLQKALRDWVTNPGLFNQ
491	SAMD7_HUMAN	NLSLDEDIQKWTVDDVHSFIRSLPGCSDYAQVFKDHAIDGETLPL LTEEHLRGTMGLKLGPALKIQSQVSQHVGSMFYKK
492	PITX3_HUMAN	SPEDGSLKKKQRRQRTHFTSQQLQELEATFQRNRYPDMSTREEIA VWTNLTEARVRVWFKNRRAKWRKRERSQQAELCKG
493	WDR5_HUMAN	SNLLVSASDDKTLKIWDVSSGKCLKTLKGHSNYVFCCNFNPQSNL IVSGSFDESVRIWDVKTGKCLKTLPAHSDPVSAVH
494	MEOX2_HUMAN	GNYKSEVNSKPRKERTAFTKEQIRELEAEFAHHNYLTRLRRYEIA VNLDLTERQVKVWFQNRRMKWKRVKGGQQGAAARE
495	NAB2_HUMAN	LPRTLGELQLYRVLQRANLLSYYETFIQQGGDDVQQLCEAGEEEF LEIMALVGMATKPLHVRRLQKALREWATNPGLFSQ
496	DHX8_HUMAN	PEEPTIGDIYNGKVTSIMQFGCFVQLEGLRKRWEGLVHISELRRE GRVANVADVVSKGQRVKVKVLSFTGTKTSLSMKDV
497	FOXA2_HUMAN	YAFNHPFSINNLMSSEQQHHHSHHHHQPHKMDLKAYEQVMHYPGY GSPMPGSLAMGPVTNKTGLDASPLAADTSYYQGVY
498	CBX6_HUMAN	TAAAGPAPPTAPEPAGASSEPEAGDWRPEMSPCSNVVVTDVTSNL LTVTIKEFCNPEDFEKVAAGVAGAAGGGGSIGASK
499	EMX2_HUMAN	FLLHNALARKPKRIRTAFSPSQLLRLEHAFEKNHYVVGAERKQLA HSLSLTETQVKVWFQNRRTKFKRQKLEEEGSDSQQ
500	CPSF6_HUMAN	KRIALYIGNLTWWTTDEDLTEAVHSLGVNDILEIKFFENRANGQS KGFALVGVGSEASSKKLMDLLPKRELHGQNPVVTP
501	HXC12_HUMAN	SGAPWYPINSRSRKKRKPYSKLQLAELEGEFLVNEFITRQRRREL SDRLNLSDQQVKIWFQNRRMKKKRLLLREQALSFF
502	KDM4B_HUMAN	SDNLYPESITSRDCVQLGPPSEGELVELRWTDGNLYKAKFISSVT SHIYQVEFEDGSQLTVKRGDIFTLEEELPKRVRSR
503	LMBL3_HUMAN	GIPASKVSKWSTDEVSEFIQSLPGCEEHGKVFKDEQIDGEAFLLM TQTDIVKIMSIKLGPALKIFNSILMFKAAEKNSHN
504	PHX2A_HUMAN	EPSGLHEKRKQRRIRTTFTSAQLKELERVFAETHYPDIYTREELA LKIDLTEARVQVWFQNRRAKFRKQERAASAKGAAG

SEQ	Description	Sequence
505	EMX1_HUMAN	LLLHGPFARKPKRIRTAFSPSQLLRLERAFEKNHYVVGAERKQLA GSLSLSETQVKVWFQNRRTKYKRQKLEEEGPESEQ
506	NC2B_HUMAN	SSGNDDDLTIPRAAINKMIKETLPNVRVANDARELVVNCCTEFIH LISSEANEICNKSEKKTISPEHVIQALESLGFGSY
507	DLX4_HUMAN	ERRPQAPAKKLRKPRTIYSSLQLQHLNQRFQHTQYLALPERAQLA AQLGLTQTQVKIWFQNKRSKYKKLLKQNSGGQEGD
508	SRY_HUMAN	NVQDRVKRPMNAFIVWSRDQRRKMALENPRMRNSEISKQLGYQWK MLTEAEKWPFFQEAQKLQAMHREKYPNYKYRPRRK
509	ZN777_HUMAN	EITRLAVWAAVQAVERKLEAQAMRLLTLEGRTGTNEKKIADCEKT AVEFANHLESKWVVLGTLLQEYGLLQRRLENMENL
510	NELL1_HUMAN	CEKDIDECSEGIIECHNHSRCVNLPGWYHCECRSGFHDDGTYSLS GESCIDIDECALRTHTCWNDSACINLAGGFDCLCP
511	ZN398_HUMAN	AAISLWTVVAAVQAIERKVEIHSRRLLHLEGRTGTAEKKLASCEK TVTELGNQLEGKWAVLGTLLQEYGLLQRRLENLEN
512	GATA3_HUMAN	GQNRPLIKPKRRLSAARRAGTSCANCQTTTTTLWRRNANGDPVCN ACGLYYKLHNINRPLTMKKEGIQTRNRKMSSKSKK
513	BSH_HUMAN	HAELPGKHCRRRKARTVFSDSQLSGLEKRFEIQRYLSTPERVELA TALSLSETQVKTWFQNRRMKHKKQLRKSQDEPKAP
514	SF3B4_HUMAN	QDATVYVGGLDEKVSEPLLWELFLQAGPVVNTHMPKDRVTGQHQG YGFVEFLSEEDADYAIKIMNMIKLYGKPIRVNKAS
515	TEAD1_HUMAN	PIDNDAEGVWSPDIEQSFQEALAIYPPCGRRKIILSDEGKMYGRN ELIARYIKLRTGKTRTRKQVSSHIQVLARRKSRDF
516	TEAD3_HUMAN	GLDNDAEGVWSPDIEQSFQEALAIYPPCGRRKIILSDEGKMYGRN ELIARYIKLRTGKTRTRKQVSSHIQVLARKKVREY
517	RGAP1_HUMAN	DSVGTPQSNGGMRLHDFVSKTVIKPESCVPCGKRIKFGKLSLKCR DCRVVSHPECRDRCPLPCIPTLIGTPVKIGEGMLA
518	PHF1_HUMAN	SAPHSMTASSSSVSSPSPGLPRRSAPPSPLCRSLSPGTGGGVRGG VGYLSRGDPVRVLARRVRPDGSVQYLVEWGGGGIF
519	FOXA1_HUMAN	GDPHYSFNHPFSINNLMSSSEQQHKLDFKAYEQALQYSPYGSTLP ASLPLGSASVTTRSPIEPSALEPAYYQGVYSRPVL
520	GATA2_HUMAN	GQNRPLIKPKRRLSAARRAGTCCANCQTTTTTLWRRNANGDPVCN ACGLYYKLHNVNRPLTMKKEGIQTRNRKMSNKSKK
521	FOXO3_HUMAN	DSLSGSSLYSTSANLPVMGHEKFPSDLDLDMFNGSLECDMESIIR SELMDADGLDFNFDSLISTQNVVGLNVGNFTGAKQ
522	ZN212_HUMAN	TEISLWTVVAAIQAVEKKMESQAARLQSLEGRTGTAEKKLADCEK MAVEFGNQLEGKWAVLGTLLQEYGLLQRRLENVEN
523	IRX4_HUMAN	MDSGTRRKNATRETTSTLKAWLQEHRKNPYPTKGEKIMLAIITKM TLTQVSTWFANARRRLKKENKMTWPPRNKCADEKR
524	ZBED6_HUMAN	NIEKQIYLPSTRAKTSIVWHFFHVDPQYTWRAICNLCEKSVSRGK PGSHLGTSTLQRHLQARHSPHWTRANKFGVASGEE
525	LHX4_HUMAN	AKQNDDSEAGAKRPRTTITAKQLETLKNAYKNSPKPARHVREQLS SETGLDMRVVQVWFQNRRAKEKRLKKDAGRHRWGQ
526	SIN3A_HUMAN	DALSYLDQVKLQFGSQPQVYNDFLDIMKEFKSQSIDTPGVISRVS QLFKGHPDLIMGFNTFLPPGYKIEVQTNDMVNVTT
527	RBBP7_HUMAN	DDHTVCLWDINAGPKEGKIVDAKAIFTGHSAVVEDVAWHLLHESL FGSVADDQKLMIWDTRSNTTSKPSHLVDAHTAEVN
528	NKX61_HUMAN	GSILLDKDGKRKHTRPTFSGQQIFALEKTFEQTKYLAGPERARLA YSLGMTESQVKVWFQNRRTKWRKKHAAEMATAKKK
529	TRI68_HUMAN	DPTALVEAIVEEVACPICMTFLREPMSIDCGHSFCHSCLSGLWEI PGESQNWGYTCPLCRAPVQPRNLRPNWQLANVVEK
530	R51A1_HUMAN	QSLPKKVSLSSDTTRKPLEIRSPSAESKKPKWVPPAASGGSRSSS SPLVVVSVKSPNQSLRLGLSRLARVKPLHPNATST
531	MB3L1_HUMAN	AKSSQRKQRDCVNQCKSKPGLSTSIPLRMSSYTFKRPVTRITPHP GNEVRYHQWEESLEKPQQVCWQRRLQGLQAYSSAG

SEQ	Description	Sequence
532	DLX5_HUMAN	VRMVNGKPKKVRKPRTIYSSFQLAALQRRFQKTQYLALPERAELA ASLGLTQTQVKIWFQNKRSKIKKIMKNGEMPPEHS
533	NOTC1_HUMAN	LQCNNHACGWDGGDCSLNFNDPWKNCTQSLQCWKYFSDGHCDSQC NSAGCLFDGFDCQRAEGQCNPLYDQYCKDHFSDGH
534	TERF2_HUMAN	ETWVEEDELFQVQAAPDEDSTTNITKKQKWTVEESEWVKAGVQKY GEGNWAAISKNYPFVNRTAVMIKDRWRTMKRLGMN
535	ZN282_HUMAN	AEISLWTVVAAIQAVERKVDAQASQLLNLEGRTGTAEKKLADCEK TAVEFGNHMESKWAVLGTLLQEYGLLQRRLENLEN
536	RGS12_HUMAN	LEKRTLFRLDLVPINRSVGLKAKPTKPVTEVLRPVVARYGLDLSG LLVRLSGEKEPLDLGAPISSLDGQRVVLEEKDPSR
537	ZN840_HUMAN	PNCLSSSMQLPHGGGRHQELVRFRDVAVVFSPEEWDHLTPEQRNL YKDVMLDNCKYLASLGNWTYKAHVMSSLKQGKEPW
538	SPI2B_HUMAN	DDYKEGDLRIMPESSESPPTEREPGGVVDGLIGKHVEYTKEDGSK RIGMVIHQVEAKPSVYFIKFDDDFHIYVYDLVKKS
539	PAX7_HUMAN	SEPDLPLKRKQRRSRTTFTAEQLEELEKAFERTHYPDIYTREELA QRTKLTEARVQVWFSNRRARWRKQAGANQLAAFNH
540	NKX62_HUMAN	AGGVLDKDGKKKHSRPTFSGQQIFALEKTFEQTKYLAGPERARLA YSLGMTESQVKVWFQNRRTKWRKRHAVEMASAKKK
541	ASXL2_HUMAN	DVMSFSVTVTTTIPASQAMNPSSHGQTIPVQAFSEENSIEGTPSKC YCRLKAMIMCKGCGAFCHDDCIGPSKLCVSCLVVR
542	FOXO1_HUMAN	GGYSSVSSCNGYGRMGLLHQEKLPSDLDGMFIERLDCDMESIIRN DLMDGDTLDFNFDNVLPNQSFPHSVKTTTHSWVSG
543	GATA3_HUMAN	GGSPTGFGCKSRPKARSSTGRECVNCGATSTPLWRRDGTGHYLCN ACGLYHKMNGQNRPLIKPKRRLSAARRAGTSCANC
544	GATA1_HUMAN	GQNRPLIRPKKRLIVSKRAGTQCTNCQTTTTTLWRRNASGDPVCN ACGLYYKLHQVNRPLTMRKDGIQTRNRKASGKGKK
545	ZMYM5_HUMAN	PVALLRKQNFQPTAQQQLTKPAKITCANCKKPLQKGQTAYQRKGS AHLFCSTTCLSSFSHKRTQNTRSIICKKDASTKKA
546	ZN783_HUMAN	TEITLWTVVAAIQALEKKVDSCLTRLLTLEGRTGTAEKKLADCEK TAVEFGNQLEGKWAVLGTLLQEYGLLQRRLENVEN
547	SPI2B_HUMAN	KKQRGRPSSQPRRNIVGCRISHGWKEGDEPITQWKGTVLDQVPIN PSLYLVKYDGIDCVYGLELHRDERVLSLKILSDRV
548	LRP1_HUMAN	WTCDLDDDCGDRSDESASCAYPTCFPLTQFTCNNGRCININWRCD NDNDCGDNSDEAGCSHSCSSTQFKCNSGRCIPEHW
549	MIXL1_HUMAN	PKGAAAPSASQRRKRTSFSAEQLQLLELVFRRTRYPDIHLRERLA ALTLLPESRIQVWFQNRRAKSRRQSGKSFQPLARP
550	SGT1_HUMAN	KIKYDWYQTESQVVITLMIKNVQKNDVNVEFSEKELSALVKLPSG EDYNLKLELLHPIIPEQSTFKVLSTKIEIKLKKPE
551	LMCD1_HUMAN	DPSKEVEYVCELCKGAAPPDSPVVYSDRAGYNKQWHPTCFVCAKC SEPLVDLIYFWKDGAPWCGRHYCESLRPRCSGCDE
552	CEBPA_HUMAN	GSGAGKAKKSVDKNSNEYRVRRERNNIAVRKSRDKAKQRNVETQQ KVLELTSDNDRLRKRVEQLSRELDTLRGIFRQLPE
553	GATA2_HUMAN	GPASSFTPKQRSKARSCSEGRECVNCGATATPLWRRDGTGHYLCN ACGLYHKMNGQNRPLIKPKRRLSAARRAGTCCANC
554	SOX14_HUMAN	KPSDHIKRPMNAFMVWSRGQRRKMAQENPKMHNSEISKRLGAEWK LLSEAEKRPYIDEAKRLRAQHMKEHPDYKYRPRRK
555	WTIP_HUMAN	LYSGFQQTADKCSVCGHLIMEMILQALGKSYHPGCFRCSVCNECL DGVPFTVDVENNIYCVRDYHTVFAPKCASCARPIL
556	PRP19_HUMAN	HPSQDLVFSASPDATIRIWSVPNASCVQVVRAHESAVTGLSLHAT GDYLLSSSDDQYWAFSDIQTGRVLTKVTDETSGCS
557	CBX6_HUMAN	ELSAVGERVFAAESIIKRRIRKGRIEYLVKWKGWAIKYSTWEPEE NILDSRLIAAFEQKERERELYGPKKRGPKPKTFLL
558	NKX11_HUMAN	RTGSDSKSGKPRRARTAFTYEQLVALENKFKATRYLSVCERLNLA LSLSLTETQVKIWFQNRRTKWKKQNPGADTSAPTG

SEQ	Description	Sequence
559	RBBP4_HUMAN	VWDLSKIGEEQSPEDAEDGPPELLFIHGGHTAKISDFSWNPNEPW VICSVSEDNIMQVWQMAENIYNDEDPEGSVDPEGQ
560	DMRT2_HUMAN	ERCTPAGGGAEPRKLSRTPKCARCRNHGVVSCLKGHKRFCRWRDC QCANCLLVVERQRVMAAQVALRRQQATEDKKGLSG
561	SMCA2_HUMAN	SQPGALIPGDPQAMSQPNRGPSPFSPVQLHQLRAQILAYKMLARG QPLPETLQLAVQGKRTLPGLQQQQQQQQQQQQQQQ
562	ZNF10	MDAKSLTAWSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLE NYKNLVSLGYQLTKPDVILRLEKGEEPWLVEREIHQETHPDSETA FEIKSSVSSRSIFKDKQSCDIKMEGMARNDLWYLSLEEVWKCRDQ LDKYQENPERHLRQVAFTQKKVLTQERVSESGKYGGNCLLPAQLV LREYFHKRDSHTKSLKHDLVLNGHQDSCASNSNECGQTFCQNIHL IQFARTHTGDKSYKCPDNDNSLTHGSSLGISKGIHREKPYECKEC GKFFSWRSNLTRHQLIHTGEKPYECKECGKSFSRSSHLIGHQKTH TGEEPYECKECGKSFSWFSHLVTHQRTHTGDKLYTCNQCGKSFVH SSRLIRHQRTHTGEKPYECPECGKSFRQSTHLILHQRTHVRVRPY ECNECGKSYSQRSHLVVHHRIHTGLKPFECKDCGKCFSRSSHLYS HQRTHTGEKPYECHDCGKSFSQSSALIVHQRIHTGEKPYECCQCG KAFIRKNDLIKHQRIHVGEETYKCNQCGIIFSQNSPFIVHQIAHT GEQFLTCNQCGTALVNTSNLIGYQTNHIRENAY
563	EED_HUMAN	MSEREVSTAPAGTDMPAAKKQKLSSDENSNPDLSGDENDDAVSIE SGTNTERPDTPTNTPNAPGRKSWGKGKWKSKKCKYSFKCVNSLKE DHNQPLFGVQFNWHSKEGDPLVFATVGSNRVTLYECHSQGEIRLL QSYVDADADENFYTCAWTYDSNTSHPLLAVAGSRGIIRIINPITM QCIKHYVGHGNAINELKFHPRDPNLLLSVSKDHALRLWNIQTDTL VAIFGGVEGHRDEVLSADYDLLGEKIMSCGMDHSLKLWRINSKRM MNAIKESYDYNPNKTNRPFISQKIHFPDFSTRDIHRNYVDCVRWL GDLILSKSCENAIVCWKPGKMEDDIDKIKPSESNVTILGRFDYSQ CDIWYMRFSMDFWQKMLALGNQVGKLYVWDLEVEDPHKAKCTTLT HHKCGAAIRQTSFSRDSSILIAVCDDASIWRWDRLR
564	RCOR1_HUMAN	MPAMVEKGPEVSGKRRGRNNAAASASAAAASAAASAACASPAATA ASGAAASSASAAAASAAAAPNNGQNKSLAAAAPNGNSSSNSWEEG SSGSSSDEEHGGGGMRVGPQYQAVVPDFDPAKLARRSQERDNLGM LVWSPNQNLSEAKLDEYIAIAKEKHGYNMEQALGMLFWHKHNIEK SLADLPNFTPFPDEWTVEDKVLFEQAFSFHGKTFHRIQQMLPDKS IASLVKFYYSWKKTRTKTSVMDRHARKQKREREESEDELEEANGN NPIDIEVDQNKESKKEVPPTETVPQVKKEKHSTQAKNRAKRKPPK GMFLSQEDVEAVSANATAATTVLRQLDMELVSVKRQIQNIKQTNS ALKEKLDGGIEPYRLPEVIQKCNARWTTEEQLLAVQAIRKYGRDF QAISDVIGNKSVVQVKNFFVNYRRRFNIDEVLQEWEAEHGKEETN GPSNQKPVKSPDNSIKMPEEEDEAPVLDVRYASAS
565	KOX1/ZNF10 KRAB 1	TGRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSL GYQLTKPDVILRLEKGEEPLEINLWITKFVKD
566	KOX1/ZNF10 KRAB 2	MYPYDVPDYASPKKKRKVGGGASMDAKSLTAWS*RTLVTFKDVFVD FTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLE KGEEPWLV*EREIHQETHPDSETAFEIKSSV
567	KOX1/ZNF10 KRAB 3	ALSPQHSAVTQGSIIKNKEGMDAKSLTAWSRTLVTFKDVFVDFTR EEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGE EPWLVEREIHQETHPDSETAFEIKSSV
568	KOX1/ZNF10 (aa 11-72)	RTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGY QLTKPDVILRLEKGEEP
569	KOX1/ZNF10 (aa 11-108)	RTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGY QLTKPDVILRLEKGEEPWLVEREIHQETHPDSETAFEIKSSVSSR SIFKDKQS

SEQ	Description	Sequence
570	KOX1/ZNF10 variant	RTLVTFKDVAVDFTQEEWQQLDPAQKIVYRDVMLENYSNLVSVGY QLTKPDVILRLEQKGEEPWLVEEEIHQETHPDSETAFEIKSSVSS RSIFKDKQS
571	KOX1 KRAB-ZIM3 chimera	RTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGY QLTKPDVILRLEKGEEPWLEEEEVLGSGRAEKNGDIGGQIWKPKD VKESL
572	ZIM3-KOX1 KRAB chimera	MNNSQGRVTFEDVTVNFTQGEWQRLNPEQRNLYRDVMLENYSNLV SVGQGETTKPDVILRLEQGKEPWLVEREIHQETHPDSETAFEIKS SVSSRSIFKDKQS
573	human DNMT1	MPARTAPARVPTLAVPAISLPDDVRRRLKDLERDSLTEKECVKEK LNLLHEFLQTEIKNQLCDLETKLRKEELSEEGYLAKVKSLLNKDL SLENGAHAYNREVNGRLENGNQARSEARRVGMADANSPPKPLSKP RTPRRSKSDGEAKPEPSPSPRITRKSTRQTTITSHFAKGPAKRKP QEESERAKSDESIKEEDKDQDEKRRRVTSRERVARPLPAEEPERA KSGTRTEKEEERDEKEEKRLRSQTKEPTPKQKLKEEPDREARAGV QADEDEDGDEKDEKKHRSQPKDLAAKRRPEEKEPEKVNPQISDEK DEDEKEEKRRKTTPKEPTEKKMARAKTVMNSKTHPPKCIQCGQYL DDPLKYGQHPPDAVDEPQMLTNEKLSIFDANESGFESYEALPQHK LTCFSVYCKHGHLCPIDTGLIEKNIELFFSGSAKPIYDDDPSLEG GVNGKNLGPINEWWITGFDGGEKALIGFSTSFAEYILMDPSPEYA PIFGLMQEKIYISKIVVEFLQSNSDSTYEDLINKIETTVPPSGLN LNRFTEDSLLRHAQFVVEQVESYDEAGDSDEQPIFLTPCMRDLIK LAGVTLGQRRAQARRQTIRHSTREKDRGPTKATTTKLVYQIFDTF FAEQIEKDDREDKENAFKRRRCGVCEVCQQPECGKCKACKDMVKF GGSGRSKQACQERRCPNMAMKEADDDEEVDDNIPEMPSPKKMHQG KKKKQNKNRISWVGEAVKTDGKKSYYKKVCIDAETLEVGDCVSVI PDDSSKPLYLARVTALWEDSSNGQMFHAHWFCAGTDTVLGATSDP LELFLVDECEDMQLSYIHSKVKVIYKAPSENWAMEGGMDPESLLE GDDGKTYFYQLWYDQDYARFESPPKTQPTEDNKFKFCVSCARLAE MRQKEIPRVLEQLEDLDSRVLYYSATKNGILYRVGDGVYLPPEAF TFNIKLSSPVKRPRKEPVDEDLYPEHYRKYSDYIKGSNLDAPEPY RIGRIKEIFCPKKSNGRPNETDIKIRVNKFYRPENTHKSTPASYH ADINLLYWSDEEAVVDFKAVQGRCTVEYGEDLPECVQVYSMGGPN RFYFLEAYNAKSKSFEDPPNHARSPGNKGKGKGKGKGKPKSQACE PSEPEIEIKLPKLRTLDVFSGCGGLSEGFHQAGISDTLWAIEMWD PAAQAFRLNNPGSTVFTEDCNILLKLVMAGETTNSRGQRLPQKGD VEMLCGGPPCQGFSGMNRFNSRTYSKFKNSLVVSFLSYCDYYRPR FFLLENVRNFVSFKRSMVLKLTLRCLVRMGYQCTFGVLQAGQYGV AQTRRRAIILAAAPGEKLPLFPEPLHVFAPRACQLSVVVDDKKFV SNITRLSSGPFRTITVRDTMSDLPEVRNGASALEISYNGEPQSWF QRQLRGAQYQPILRDHICKDMSALVAARMRHIPLAPGSDWRDLPN IEVRLSDGTMARKLRYTHHDRKNGRSSSGALRGVCSCVEAGKACD PAARQFNTLIPWCLPHTGNRHNHWAGLYGRLEWDGFFSTTVTNPE PMGKQGRVLHPEQHRVVSVRECARSQGFPDTYRLFGNILDKHRQV GNAVPPPLAKAIGLEIKLCMLAKARESASAKIKEEEAAKD
574	human DNMT3A	MPAMPSSGPGDTSSSAAEREEDRKDGEEQEEPRGKEERQEPSTTA RKVGRPGRKRKHPPVESGDTPKDPAVISKSPSMAQDSGASELLPN GDLEKRSEPQPEEGSPAGGQKGGAPAEGEGAAETLPEASRAVENG CCTPKEGRGAPAEAGKEQKETNIESMKMEGSRGRLRGGLGWESSL RQRPMPRLTFQAGDPYYISKRKRDEWLARWKREAEKKAKVIAGMN AVEENQGPGESQKVEEASPPAVQQPTDPASPTVATTPEPVGSDAG DKNATKAGDDEPEYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWM TGRSRAAEGTRWVMWFGDGKFSVVCVEKLMPLSSFCSAFHQATYN KQPMYRKAIYEVLQVASSRAGKLFPVCHDSDESDTAKAVEVQNKP MIEWALGGFQPSGPKGLEPPEEEKNPYKEVYTDMWVEPEAAAYAP

SEQ	Description	Sequence
		PPPAKKPRKSTAEKPKVKEIIDERTRERLVYEVRQKCRNIEDICI SCGSLNVTLEHPLFVGGMCQNCKNCFLECAYQYDDDGYQSYCTIC CGGREVLMCGNNNCCRCFCVECVDLLVGPGAAQAAIKEDPWNCYM CGHKGTYGLLRRREDWPSRLQMFFANNHDQEFDPPKVYPPVPAEK RKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVGMV RHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPAR KGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSD KRDISRFLESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVN DKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEK EDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRH LFAPLKEYFACV
575	human DNMT3A catalytic domain	NHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQ VDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPF DLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPKEG DDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRA RYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTR SNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMS RLARQRLLGRSWSVPVIRHLFAPLKEYFACV
576	human DNMT3B	MKGDTRHLNGEEDAGGREDSILVNGACSDQSSDSPPILEAIRTPE IRGRRSSSRLSKREVSSLLSYTQDLTGDGDGEDGDGSDTPVMPKL FRETRTRSESPAVRTRNNNSVSSRERHRPSPRSTRGRQGRNHVDE SPVEFPATRSLRRRATASAGTPWPSPPSSYLTIDLTDDTEDTHGT PQSSSTPYARLAQDSQQGGMESPQVEADSGDGDSSEYQDGKEFGI GDLVWGKIKGFSWWPAMVVSWKATSKRQAMSGMRWVQWFGDGKFS EVSADKLVALGLFSQHFNLATFNKLVSYRKAMYHALEKARVRAGK TFPSSPGDSLEDQLKPMLEWAHGGFKPTGIEGLKPNNTQPVVNKS KVRRAGSRKLESRKYENKTRRRTADDSATSDYCPAPKRLKTNCYN NGKDRGDEDQSREQMASDVANNKSSLEDGCLSCGRKNPVSFHPLF EGGLCQTCRDRFLELFYMYDDDGYQSYCTVCCEGRELLLCSNTSC CRCFCVECLEVLVGTGTAAEAKLQEPWSCYMCLPQRCHGVLRRRK DWNVRLQAFFTSDTGLEYEAPKLYPAIPAARRRPIRVLSLFDGIA TGYLVLKELGIKVGKYVASEVCEESIAVGTVKHEGNIKYVNDVRN ITKKNIEEWGPFDLVIGGSPCNDLSNVNPARKGLYEGTGRLFFEF YHLLNYSRPKEGDDRPFFWMFENVVAMKVGDKRDISRFLECNPVM IDAIKVSAAHRARYFWGNLPGMNRPVIASKNDKLELQDCLEYNRI AKLKKVQTITTKSNSIKQGKNQLFPVVMNGKEDVLWCTELERIFG FPVHYTDVSNMGRGARQKLLGRSWSVPVIRHLFAPLKDYFACE
577	mouse DNMT3C	MRGGSRHLSNEEDVSGCEDCIIISGTCSDQSSDPKTVPLTQVLEA VCTVENRGCRTSSQPSKRKASSLISYVQDLTGDGDEDRDGEVGGS SGSGTPVMPQLFCETRIPSKTPAPLSWQANTSASTPWLSPASPYP IIDLTDEDVIPQSISTPSVDWSQDSHQEGMDTTQVDAESRDGGNI EYQVSADKLLLSQSCILAAFYKLVPYRESIYRTLEKARVRAGKAC PSSPGESLEDQLKPMLEWAHGGFKPTGIEGLKPNKKQPENKSRRR TTNDPAASESSPPKRLKTNSYGGKDRGEDEESREQMASDVTNNKG NLEDHCLSCGRKDPVSFHPLFEGGLCQSCRDRFLELFYMYDEDGY QSYCTVCCEGRELLLCSNTSCCRCFCVECLEVLVGAGTAEDVKLQ EPWSCYMCLPQRCHGVLRRRKDWNMRLQDFFTTDPDLEEFEPPKL YPAIPAAKRRPIRVLSLFDGIATGYLVLKELGIKVEKYIASEVCA ESIAVGTVKHEGQIKYVDDIRNITKEHIDEWGPFDLVIGGSPCND LSCVNPVRKGLFEGTGRLFFEFYRLLNYSCPEEEDDRPFFWMFEN VVAMEVGDKRDISRFLECNPVMIDAIKVSAAHRARYFWGNLPGMN RPVMASKNDKLELQDCLEFSRTAKLKKVQTITTKSNSIRQGKNQL FPVVMNGKDDVLWCTELERIFGFPEHYTDVSNMGRGARQKLLGRS WSVPVIRHLFAPLKDHFACE

SEQ	Description	Sequence
578	human DNMT3L	MAAIPALDPEAEPSMDVILVGSSELSSSVSPGTGRDLIAYEVKAN QRNIEDICICCGSLQVHTQHPLFEGGICAPCKDKFLDALFLYDDD GYQSYCSICCSGETLLICGNPDCTRCYCFECVDSLVGPGTSGKVH AMSNWVCYLCLPSSRSGLLQRRRKWRSQLKAFYDRESENPLEMFE TVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSDPGQLKHVVDVT DTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLLQYA RPKPGSPRPFFWMFVDNLVLNKEDLDVASRFLEMEPVTIPDVHGG SLQNAVRVWSNIPAIRSSRHWALVSEEELSLLAQNKQSSKLAAKW PTKLVKNCFLPLREYFKYFSTELTSSL
579	human DNMT3L catalytic domain	NPLEMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSDPGQL KHVVDVTDTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQF HRLLQYARPKPGSPRPFFWMFVDNLVLNKEDLDVASRFLEMEPVT IPDVHGGSLQNAVRVWSNIPAIRSRHWALVSEEELSLLAQNKQSS KLAAKWPTKLVKNCFLPLREYFKYFSTELTSSL
580	mouse DNMT3L	MGSRETPSSCSKTLETLDLETSDSSSPDADSPLEEQWLKSSPALK EDSVDVVLEDCKEPLSPSSPPTGREMIRYEVKVNRRSIEDICLCC GTLQVYTRHPLFEGGLCAPCKDKFLESLFLYDDDGHQSYCTICCS GGTLFICESPDCTRCYCFECVDILVGPGTSERINAMACWVCFLCL PFSRSGLLQRRKRWRHQLKAFHDQEGAGPMEIYKTVSAWKRQPVR VLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEK WGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQESQRP FFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVW SNIPGLKSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLL PLREYFKYFSQNSLPL
581	mouse DNMT3L catalytic domain	GPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGG TLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMF QFHRILQYALPRQESQRPFFWIFMDNLLLTEDDQETTTRFLQTEA VTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEEYLQAQVRS RSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL
582	*Ailuropoda melanoleuca* DNMT3L	MALSPTGTLSVETLDRSDPDPLDEGPWQATCEILLEPDAEHSTDV ILVGSSELSAPASPGPRRDLLAYEVKVNQRDIEDVCICCGSLRVH TQHPLFEGGMCAPCKDKFLDCLFLYDDDGYQSYCSICCAGETLLI CENPDCTRPSLMMKLRLFRECACLIFPSEGMLLQTVWFWKMTVVW QPGLRHLPQENPLETYKTVPVWKREPVRVLSLFGDIRRELMSLGF LESGSAPGRLKHLDDVTDVVRKDVEGWGPFDLVYGSTPPIGHACD HPPVWYLLQFHRILQYARPRPGSQQPFFWMFVDNLVLSQDDQTAA TRFLEADPVTIQDVCGRAVRNTVHVWSNIPAVRSRHSALALCEEL SLLAQDRQRTKPPAQGPAQLVKNCFLPLREYFKYFSTELTSSL
583	*Ailuropoda melanoleuca* DNMT3L catalytic domain	NPLETYKTVPVWKREPVRVLSLFGDIRRELMSLGFLESGSAPGRL KHLDDVTDVVRKDVEGWGPFDLVYGSTPPIGHACDHPPVWYLLQF HRILQYARPRPGSQQPFFWMFVDNLVLSQDDQTAATRFLEADPVT IQDVCGRAVRNTVHVWSNIPAVRSRHSALALCEELSLLAQDRQRT KPPAQGPAQLVKNCFLPLREYFKYFSTELTSSL
584	*Carlito syrichta* DNMT3L	MALSCRRTLPLESLHSSNSDLASQLDKEQWRPPCETHGIPVAAAP VLDLEAECSLDVILVGSSELSTSSSPRLGRDHIAYEVKVNQRNIE DICLCCGSFLVHTQHPLFEGGMCAPCKDKFLDTLFLYDEDGYQSY CSICCSGETLLICENPDCTRCYCFECLDTLVSPGTSEKVHAMSNW VCFLCLPFTRSGLLQRRRKWRGQLKAFYDRESESSLEMYKTVPVW KREPVRVLSLFGDIKKELMSLGFVETGSDPGRLRHLDDTTNIVRR NVEEWGPFHLLYGATPPLGHTCDRPPGWYLFQFHRLLQYARPQPG SPQPFFWMFVDNVMLTREDRAIASRFLETEPVTIPDIHGRALQNA VCVWSNIPAVRSKHSALVSEEELSLLAQDRQRAKLPTQGPTKLVK NCFLPLREYFKYFSTELTSFL

SEQ	Description	Sequence
585	*Carlito syrichta* DNMT3L catalytic domain	SSLEMYKTVPVWKREPVRVLSLFGDIKKELMSLGFVETGSDPGRL RHLDDTTNIVRRNVEEWGPFHLLYGATPPLGHTCDRPPGWYLFQF HRLLQYARPQPGSPQPFFWMFVDNVMLTREDRAIASRFLETEPVT IPDIHGRALQNAVCVWSNIPAVRSKHSALVSEEELSLLAQDRQRA KLPTQGPTKLVKNCFLPLREYFKYFSTELTSFL
586	*Meriones unguiculatus* DNMT3L	MGSQETPSTRAKTPGTWNLESTDSSSPESLGHLEEQWANSSPDLK DEHSKDVEPEDSKELISSASPPSGREIIRYEISVNQRNIEDICLC CGTLQVYKQHPLFEGGICAPCKDKFLETFFLYDEDGHQSYCSICC SGGTLFICESPDCTRCYCFECVDILVGPGTSERINAMPCWVCFLC LPFTRSGLLQRRRKWRHQLKAFFDEGGASPLEMYKTVSAWKRKPM RVLSLFKNIDKELKNLGFLESGSGSEEERLKYLEDVTNVVRRDVE KWGPFDLVYGSTRPRGSSCDHCPAWYMFQFHRILQYARPPSGSEQ PFFWVFVDNLLMTEDDQITADRFLQMKAVTLQDVRGRVLQNAVRV WSNIPGVKSKHMALTEKEEQSLEAQAGTRTKLSAQKVDPLVKNCL LPLREYFKFFSQNSLPLDK
587	*Meriones unguiculatus* DNMT3L catalytic domain	SPLEMYKTVSAWKRKPMRVLSLFKNIDKELKNLGFLESGSGSEEE RLKYLEDVTNVVRRDVEKWGPFDLVYGSTRPRGSSCDHCPAWYMF QFHRILQYARPPSGSEQPFFWVFVDNLLMTEDDQITADRFLQMKA VTLQDVRGRVLQNAVRVWSNIPGVKSKHMALTEKEEQSLEAQAGT RTKLSAQKVDPLVKNCLLPLREYFKFFSQNSLPLDK
588	*Ochotona princeps* DNMT3L	MALPSPETLDSLDRVPASHPDEQHWTVCDNSDPILEVEAEGSMDV ILVDDSPAPSGRDRIELEVKVNQRSIEDLCLCCGSSQVHRQHPLF QGGLCAPCKDKFLEALFLYDEDGYQSYCSICGLGDTLLVCESPDC TRGYCFACVDGLVGAGSSGHMHTVSPWVCFLCVPGSRHGLLQRRR RWRTQLKVFHEQEAAQPLEIYETVPACRRKPLRVLSLFEHIEKEL ASLGFLETGSSPGRIRHLDDVTDVVRRDVEQWGPFDLVYGSTPPL GHASPRSPGWYLFQFHRMLQYTQPTASTQRPFFWMFVDNLLLTRD DLVTATRFLEVEPATLQDVRGRVLQGAMRVWSNIPAVNSRHTELA PEAETALLAQSCRRAKASGEGLARLLKSCFLPLREYFKYFPQSPL PLRK
589	*Ochotona princeps* DNMT3L catalytic domain	QPLEIYETVPACRRKPLRVLSLFEHIEKELASLGFLETGSSPGRI RHLDDVTDVVRRDVEQWGPFDLVYGSTPPLGHASPRSPGWYLFQF HRMLQYTQPTASTQRPFFWMFVDNLLLTRDDLVTATRFLEVEPAT LQDVRGRVLQGAMRVWSNIPAVNSRHTELAPEAETALLAQSCRRA KASGEGLARLLKSCFLPLREYFKYFPQSPLPLRK
590	*Neosciurus carolinensis* DNMT3L	MGGPRPAAVEESPHEIYKTVPAWKREPMRVLSLFGDIGKELTSLG FLETGSEAGRLKHLEDVTDTVRRDVEEWGPFDLVYGSTPALGHSC DRSPGWYLFQFHRLLQYARPRLGSPKPFFWMFVDNLLLTKDDQAI ASRFLEMEPVTLQDVHGRVLQNAVRVWTNVPAVKSRHSALASEEE LLLVQDGQRGRLPAQGPAALVKHCFLPLREYFKYFSQNTLPLYK
591	*Neosciurus carolinensis* DNMT3L catalytic domain	SPHEIYKTVPAWKREPMRVLSLFGDIGKELTSLGFLETGSEAGRL KHLEDVTDTVRRDVEEWGPFDLVYGSTPALGHSCDRSPGWYLFQF HRLLQYARPRLGSPKPFFWMFVDNLLLTKDDQAIASRFLEMEPVT LQDVHGRVLQNAVRVWTNVPAVKSRHSALASEEELLLVQDGQRGR LPAQGPAALVKHCFLPLREYFKYFSQNTLPLYK
592	*Bison bison* DNMT3L	MARSSPGTLNLEIMDGSDPDPALPPDREQWPPPCEILLDPEPEHS LDIILVGSSELSSPPSPGPRRDFIAYEVKVNQRDIEDVCICCGSL QLHTQHPLFEGGMCAPCKDKFLECLFLYDDDGYQSYCSICCAGET LLICENPDCTRCYCFECVDTLVGPGTSGKVHAMSNWVCFLCLPFP RSGLLQRRRKWRTWLKAFYDREAESPLVMYKTVPVWKREPIRVLS LFGDIKKELTSLGFLEDGSKPGRLKHLDDVTNIVRRDIDEWGPFD LTYGSTPTLGHTCDHPPGWYVYQFHRILQYARPLPGSPQPFFWMF VDNLVLTEEDLDVATRFLETDPVTIQDVRGRTVQNAVHVWSNIPA VKSRHSALVSQEELSLLAQDRQRVKSPVQGPATLVKNCFLPLREY FKYFSTELTSSL

SEQ	Description	Sequence
593	*Bison bison* DNMT3L catalytic domain	SPLVMYKTVPVWKREPIRVLSLFGDIKKELTSLGFLEDGSKPGRL KHLDDVTNIVRRDIDEWGPFDLTYGSTPTLGHTCDHPPGWYVYQF HRILQYARPLPGSPQPFFWMFVDNLVLTEEDLDVATRFLETDPVT IQDVRGRTVQNAVHVWSNIPAVKSRHSALVSQEELSLLAQDRQRV KSPVQGPATLVKNCFLPLREYFKYFSTELTSSL
594	*Equus przewalskii* DNMT3L	MALSSPGTLSLETLDSWDPDVAGQLDEERWQPSSEIVGRPMAAAP VLDLEEEPSMDIILVDSSELSSPPSPGPSRDMCICCGSFQVHTQH PLFEGGMCAACKDKFLSCLFLYDDDGNQSYCSICCSGETLLICEN PDCTRCYCFECVDTLVSPRTSEKVQAMSNWVCFLCLPFPRSGLLQ RRRKWRGWLKAFYDQEAVRSRSAWGRRMRSGPHLVGFLWLLVAKC PSALESPLEMYKTVPVWKREPVRVLSLFGDIKKELTTLGFLENGS DPGRLKHLDDVTNTVRRDVEEWGPFDLVYGSTPPLGHACDHPPGW YLFQFHRVLQYARPRPGSPQAFFWMFVDNLVLTEDDRAVATRFLE TDPVTIQDVCGRAVRNAVHVWSNIPAVKSRHSALFSQEESFLRAQ DRQRAKPPARGPAKLVKNCFLPLREYFKYFSTEFTSSL
595	*Equus przewalskii* DNMT3L catalytic domain	SPLEMYKTVPVWKREPVRVLSLFGDIKKELTTLGFLENGSDPGRL KHLDDVTNTVRRDVEEWGPFDLVYGSTPPLGHACDHPPGWYLFQF HRVLQYARPRPGSPQAFFWMFVDNLVLTEDDRAVATRFLETDPVT IQDVCGRAVRNAVHVWSNIPAVKSRHSALFSQEESFLRAQDRQRA KPPARGPAKLVKNCFLPLREYFKYFSTEFTSSL
596	*Mus caroli* DNMT3L	MGSRETPSSFSKTLETLDLETSDSSSPDADSPLEEQWLKSSPALK EDNVDMVLEDCKEPLSPSSPPTGREMIRYEVKVNRRSIEDICLCC GTLQVYTQHPLFEGGICAPCKDKFLESLFLYDDDGHQSYCTICCS GGTLFICESPDCTRCYCFECVDILVGPGTSERINAMACWVCFLCL PFSRSGLLQRRKRWRHQLKAFHDQEGAGPMEIYKTVSTWKRQPVR VLSLFGNIDKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEK WGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQESQRP FFWIFMDNLLMTEDDQETTARFLQTEAVTLQDVRGRDYQNVMRVW SNIPGLKSKHVPLTPKEEEYLQAQVRTRSKLDAQKVDLLVKNCLL PLREYFKYFS
597	*Mus caroli* DNMT3L catalytic domain	GPMEIYKTVSTWKRQPVRVLSLFGNIDKVLKSLGFLESGSGSGGG TLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMF QFHRILQYALPRQESQRPFFWIFMDNLLMTEDDQETTARFLQTEA VTLQDVRGRDYQNVMRVWSNIPGLKSKHVPLTPKEEEYLQAQVRT RSKLDAQKVDLLVKNCLLPLREYFKYFS
598	*Pan troglodytes* DNMT3L	MAAIPALDPEAEPSMDVILVGSSELSSSISPRTGRDLIAYEVKAN QRNIEDICICCGSLQVHTQHPLFEGGICAPCKDKSLDALFLYDDD GYQSYCSICCSGETLLICGNPDCTRCYCFECVDSLVGPGTSGKVH AMSNWVCYLCLPSSRSGLLQRRRKWRSQLKAFYDRESENPLEMFE TVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSDPGQLKHVVDVT DTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLLQYA RPKPGSPRPFFWMFVDNLVLNKEDLDVASRFLEMEPVTIPDVHGG SLQNAVRVWSNIPAIRSSRHWALVSEEELSLLAQNKQSSKLAAKW PTKLVKNCFLPLREYFKYFSTELTSSL
599	*Pan troglodytes* DNMT3L catalytic domain	NPLEMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSDPGQL KHVVDVTDTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQF HRLLQYARPKPGSPRPFFWMFVDNLVLNKEDLDVASRFLEMEPVT IPDVHGGSLQNAVRVWSNIPAIRSSRHWALVSEEELSLLAQNKQS SKLAAKWPTKLVKNCFLPLREYFKYFSTELTSSL
600	human TRDMT1 (DNMT2)	MEPLRVLELYSGVGGMHHALRESCIPAQVVAAIDVNTVANEVYKY NFPHTQLLAKTIEGITLEEFDRLSFDMILMSPPCQPFTRIGRQGD MTDSRTNSFLHILDILPRLQKLPKYILLENVKGFEVSSTRDLLIQ TIENCGFQYQEFLLSPTSLGIPNSRLRYFLIAKLQSEPLPFQAPG QVLMEFPKIESVHPQKYAMDVENKIQEKNVEPNISFDGSIQCSGK DAILFKLETAEEIHRKNQQDSDLSVKMLKDFLEDDTDVNQYLLPP

SEQ	Description	Sequence
		KSLLRYALLLDIVQPTCRRSVCFTKGYGSYIEGTGSVLQTAEDVQ VENIYKSLTNLSQEEQITKLLILKLRYFTPKEIANLLGFPPEFGF PEKITVKQRYRLLGNSLNVHVVAKLIKILYE
601	*M. bacterium* methyltransferase	MAEWYIPAIVSYQAIHNGFTLNKINHKIELQTMIDYLESKTLSMN SKEPVKRGFWYKKHLDEIRIVYTAVKMSEQEGNIFDVRTLFERGL SDIDLLTYSFPCQDLSQQGKQKGMGRDSQTRSGLLWEIEKALDTS KKEDLPKYLLMENVVALTHKVNAEELDEWMMKLESLGYKNDLRIL NAGDFGSSQARRRTFMISTLNEKVELPVGNKKPKSMNKILNDEPT RKDFLPALDKFDLTEYKWTKSNINKAKLINYSTFNSEAYVYDSNF TGPTLTASGANSRIKFEYNGKIRKIGAEEAYAYMGFKKSDYIKVN KLNYLNETKMIYTCGNSISVEVLRSIMTNINNNFKENK
602	*M. marinum* methyltransferase	MLFLIGTFKYVLIYITKVIRIFEAFAGIGAQRKALRNIKSNYEVS GMAEWYIPAIVSYQAIHNGFTLSRVDKKTKLTEMIKYLESKTLSM DSKEPVRTGYWFKKHKDMVRIVYSAVKLSEAEGNIFDVRTLHERK LEDIDLLTYSFPCQDLSQQGKQRGMKKDSGTRSGLLWEIEKALEA TPKDKLPKYLLMENVVALTHKTNKKDLDNWKRKLRSLGYYNDINV LNAGDFGSSQARRRAFMISTLDSKVTLPLGDKKPQAISKILNKET RSQDFMPALDEYEKTDFKRTLSNIKKCKLIDYTSFNSEAYVYDPK YTGPTLTASGANSRIKFTHQGKMRKINAEEAYRYMGFSTNDYKKV NNLNFLSETKMIYTCGNSISVEVLEEIMLKIIREDNNG
603	*S. chinense* methyltransferase	MKKIRLFEAFAGIGSQRRALKSVVGNNFEIAGLAEWYVPAIVMYQ IINNDFSKKNVLDNVPRDEVIDYLNSKCLSWDSKKPVSKNFWNRK SQDILNVIYSAVKKSEEEGNIFDVRTLHERTLESIDILTYSFPCQ DLSQQGIQKGMKKNSGTRSGLLWEIEKAIDNTPKNNLPKILLMEN VPALLNKTNELELKEWLIKLENMGYKNSIGILNAADFGSPQARRR VFMISSRNKKIELPVGKSKPGKLNDILEKNVEDKFIMTNLEKYDF SEFSLTKSNIKKCSLINYTKFNSEAYVYDPDFTGPTLTASGANSR IKIYDKGFIRRMSPLESFRYMGFDDEDYKKIDEFEFLTDTQKIFV CGNSISIEVLKAIFERIDSNE
604	*M. penetrans M MpeI*	MNSNKDKIKVIKVFEAFAGIGSQFKALKNIARSKNWEIQHSGMVE WFVDAIVSYVAIHSKNFNPKIEQLDKDILSISNDSKMPISEYGIK KINNTIKASYLNYAKKHFNNLFDIKKVNKDNFPKNIDIFTYSFPC QDLSVQGLQKGIDKELNTRSGLLWEIERILEEIKNSFSKEEMPKY LLMENVKNLLSHKNKKNYNTWLKQLEKFGYKSKTYLLNSKNFDNC QNRERVFCLSIRDDYLEKTGFKFKELEKVKNPPKKIKDILVDSSN YKYLNLNKYETTTFRETKSNIISRSLKNYTTFNSENYVYNINGIG PTLTASGANSRIKIETQQGVRYLTPLECFKYMQFDVNDFKKVQST NLISENKMIYIAGNSIPVKILEAIFNTLEFVNNEE
605	*S. monobiae M SssI*	MSKVENKTKKLRVFEAFAGIGAQRKALEKVRKDEYEIVGLAEWYV PAIVMYQAIHNNFHTKLEYKSVSREEMIDYLENKTLSWNSKNPVS NGYWKRKKDDELKIIYNAIKLSEKEGNIFDIRDLYKRTLKNIDLL TYSFPCQDLSQQGIQKGMKRGSGTRSGLLWEIERALDSTEKNDLP KYLLMENVGALLHKKNEEELNQWKQKLESLGYQNSIEVLNAADFG SSQARRRVFMISTLNEFVELPKGDKKPKSIKKVLNKIVSEKDILN NLLKYNLTEFKKTKSNINKASLIGYSKFNSEGYVYDPEFTGPTLT ASGANSRIKIKDGSNIRKMNSDETFLYIGFDSQDGKRVNEIEFLT ENQKIFVCGNSISVEVLEAIIDKIGG
606	*H. parainfluenzae M HpaII*	MKDVLDDNLLEEPAAQYSLFEPESNPNLREKFTFIDLFAGIGGFR IAMQNLGGKCIFSSEWDEQAQKTYEANFGDLPYGDITLEETKAFI PEKFDILCAGFPCQAFSIAGKRGGFEDTRGTLFFDVAEIIRRHQP KAFFLENVKGLKNHDKGRTLKTILNVLREDLGYFVPEPAIVNAKN FGVPQNRERIYIVGFHKSTGVNSFSYPEPLDKIVTFADIREEKTV PTKYYLSTQYIDTLRKHKERHESKGNGFGYEIIPDDGIANAIVVG GMGRERNLVIDHRITDFTPTTNIKGEVNREGIRKMTPREWARLQG FPDSYVIPVSDASAYKQFGNSVAVPAIQATGKKILEKLGNLYD

SEQ	Description	Sequence
607	*A. luteus M AluI*	MSKANAKYSFVDLFAGIGGFHAALAATGGVCEYAVEIDREAAAVY ERNWNKPALGDITDDANDEGVTLRGYDGPIDVLTGGFPCQPFSKS GAQHGMAETRGTLFWNIARIIEEREPTVLILENVRNLVGPRHRHE WLTIIETLRFFGYEVSGAPAIFSPHLLPAWMGGTPQVRERVFITA TLVPERMRDERIPRTETGEIDAEAIGPKPVATMNDRFPIKKGGTE LFHPGDRKSGWNLLTSGIIREGDPEPSNVDLRLTETETLWIDAWD DLESTIRRATGRPLEGFPYWADSWTDFRELSRLVVIRGFQAPERE VVGDRKRYVARTDMPEGFVPASVTRPAIDETLPAWKQSHLRRNYD FFERHFAEVVAWAYRWGVYTDLFPASRRKLEWQAQDAPRLWDTVM HFRPSGIRAKRPTYLPALVAITQTSIVGPLERRLSPRETARLQGL PEWFDFGEQRAAATYKQMGNGVNVGVVRHILREHVRRDRALLKLT PAGQRIINAVLADEPDATVGALGAAE
608	*H. aegyptius M HaeIII*	MNLISLFSGAGGLDLGFQKAGFRIICANEYDKSIWKTYESNHSAK LIKGDISKISSDEFPKCDGIIGGPPCQSWSEGGSLRGIDDPRGKL FYEYIRILKQKKPIFFLAENVKGMMAQRHNKAVQEFIQEFDNAGY DVHIILLNANDYGVAQDRKRVFYIGFRKELNINYLPPIPHLIKPT FKDVIWDLKDNPIPALDKNKTNGNKCIYPNHEYFIGSYSTIFMSR NRVRQWNEPAFTVQASGRQCQLHPQAPVMLKVSKNLNKFVEGKEH LYRRLTVRECARVQGFPDDFIFHYESLNDGYKMIGNAVPVNLAYE IAKTIKSALEICKGN
609	*H. haemolyticus M HhaI*	MIEIKDKQLTGLRFIDLFAGLGGFRLALESCGAECVYSNEWDKYA QEVYEMNFGEKPEGDITQVNEKTIPDHDILCAGFPCQAFSISGKQ KGFEDSRGTLFFDIARIVREKKPKVVFMENVKNFASHDNGNTLEV VKNTMNELDYSFHAKVLNALDYGIPQKRERIYMICFRNDLNIQNF QFPKPFELNTFVKDLLLPDSEVEHLVIDRKDLVMTNQEIEQTTPK TVRLGIVGKGGQGERIYSTRGIAITLSAYGGGIFAKTGGYLVNGK TRKLHPRECARVMGYPDSYKVHPSTSQAYKQFGNSVVINVLQYIA YNIGSSLNFKPY
610	*Moraxella M MspI*	MKPEILKLIRSKLDLTQKQASEIIEVSDKTWQQWESGKTEMHPAY YSFLQEKLKDKINFEELSAQKTLQKKIFDKYNQNQITKNAEELAE ITHIEERKDAYSSDFKFIDLFSGIGGIRQSFEVNGGKCVFSSEID PFAKFTYYTNFGVVPFGDITKVEATTIPQHDILCAGFPCQPFSHI GKREGFEHPTQGTMFHEIVRIIETKKTPVLFLENVPGLINHDDGN TLKVIIETLEDMGYKVHHTVLDASHFGIPQKRKRFYLVAFLNQNI HFEFPKPPMISKDIGEVLESDVTGYSISEHLQKSYLFKKDDGKPS LIDKNTTGAVKTLVSTYHKIQRLTGTFVKDGETGIRLLTTNECKA IMGFPKDFVIPVSRTQMYRQMGNSVVVPVVTKIAEQISLALKTVN QQSPQENFELELV
611	*Ascobolus* Masc1	MSERRYEAGMTVALHEGSFLKIQRVYIRQYHADNRREHMLVGPLF RRTKYLKALSKKVNEVAIVHESIHVPVQDVIGVRELIITNRPFPE CRKGDEHTGRLVCRWVYNLDERAKGREYKKQRYIRRITEAEADPE YRVEDRVLRRRWFQEGYIGDEISYKEHGNGDIVDIRSESPLQVLD GWGGDLVDLENGEETSIPGPCRSASSYGRLMKPPLAQAADSNTSR KYTFGDTFCGGGGVSLGARQAGLEVKWAFDMNPNAGANYRRNFPN TDFFLAEAEQFIQLSVGISQHVDILHLSPPCQTFSRAHTIAGKND ENNEASFFAVVNLIKAVRPRLFTVEETDGIMDRQSRQFIDTALMG ITELGYSFRICVLNAIEYGVCQNRKRLIIIGAAPGEELPPFPLPT HQDFFSKDPRRDLLPAVTLDDALSTITPESTDHHLNHVWQPAEWK TPYDAHRPFKNAIRAGGGEYDIYPDGRRKFTVRELACIQGFPDEY EFVGTLTDKRRIIGNAVPPPLSAAIMSTLRQWMTEKDFERME
612	*Arabidopsis* MET1	MVENGAKAAKRKKRPLPEIQEVEDVPRTRRPRRAAACTSFKEKSI RVCEKSATIEVKKQQIVEEEFLALRLTALETDVEDRPTRRLNDFV LFDSDGVPQPLEMLEIHDIFVSGAILPSDVCTDKEKEKGVRCTSF GRVEHWSISGYEDGSPVIWISTELADYDCRKPAASYRKVYDYFYE KARASVAVYKKLSKSSGGDPDIGLEELLAAVVRSMSSGSKYFSSG

SEQ	Description	Sequence
		AAIIDFVISQGDFIYNQLAGLDETAKKHESSYVEIPVLVALREKS SKIDKPLQRERNPSNGVRIKEVSQVAESEALTSDQLVDGTDDDRR YAILLQDEENRKSMQQPRKNSSSGSASNMFYIKINEDEIANDYPL PSYYKTSEEETDELILYDASYEVQSEHLPHRMLHNWALYNSDLRF ISLELLPMKQCDDIDVNIFGSGVVTDDNGSWISLNDPDSGSQSHD PDGMCIFLSQIKEWMIEFGSDDIISISIRTDVAWYRLGKPSKLYA PWWKPVLKTARVGISILTFLRVESRVARLSFADVTKRLSGLQAND KAYISSDPLAVERYLVVHGQIILQLFAVYPDDNVKRCPFVVGLAS KLEDRHHTKWIIKKKKISLKELNLNPRAGMAPVASKRKAMQATTT RLVNRIWGEFYSNYSPEDPLQATAAENGEDEVEEEGGNGEEEVEE EGENGLTEDTVPEPVEVQKPHTPKKIRGSSGKREIKWDGESLGKT SAGEPLYQQALVGGEMVAVGGAVTLEVDDPDEMPAIYFVEYMFES TDHCKMLHGRFLQRGSMTVLGNAANERELFLTNECMTTQLKDIKG VASFEIRSRPWGHQYRKKNITADKLDWARALERKVKDLPTEYYCK SLYSPERGGFFSLPLSDIGRSSGFCTSCKIREDEEKRSTIKLNVS KTGFFINGIEYSVEDFVYVNPDSIGGLKEGSKTSFKSGRNIGLRA YVVCQLLEIVPKESRKADLGSFDVKVRRFYRPEDVSAEKAYASDI QELYFSQDTVVLPPGALEGKCEVRKKSDMPLSREYPISDHIFFCD LFFDTSKGSLKQLPANMKPKFSTIKDDTLLRKKKGKGVESEIESE IVKPVEPPKEIRLATLDIFAGCGGLSHGLKKAGVSDAKWAIEYEE PAGQAFKQNHPESTVFVDNCNVILRAIMEKGGDQDDCVSTTEANE LAAKLTEEQKSTLPLPGQVDFINGGPPCQGFSGMNRFNQSSWSKV QCEMILAFLSFADYFRPRYFLLENVRTFVSFNKGQTFQLTLASLL EMGYQVRFGILEAGAYGVSQSRKRAFIWAAAPEEVLPEWPEPMHV FGVPKLKISLSQGLHYAAVRSTALGAPFRPITVRDTIGDLPSVEN GDSRTNKEYKEVAVSWFQKEIRGNTIALTDHICKAMNELNLIRCK LIPTRPGADWHDLPKRKVTLSDGRVEEMIPFCLPNTAERHNGWKG LYGRLDWQGNFPTSVTDPQPMGKVGMCFHPEQHRILTVRECARSQ GFPDSYEFAGNINHKHRQIGNAVPPPLAFALGRKLKEALHLKKSP QHQP
613	*Ascobolus* Masc2	MELTPELSGVSTDLGGGGSIFAHWRMKEESPAPTEILDDLNVLEW EKTTRDYSKEDLRIADQLFSIEDEHQSLPFETADAEDGTPTEEEE EKELPMRTLDNFVLYDASDLELAALDLIGTELNIHAVGTVGPIYT EGEEDEQEDEDEDVSPPVRTGTQATSASVTQMTVELYIRNIVQYE FCFNDDGTVETWIQTTNAHYKLLQPAKCYTSLYRPVNDCLNVITA IITLAPESTTMSLKDLLKVMDDKAQAVSYEEVERMSEFIVQHLDQ WMETAPKKKSKLIEKSKVYIDLNNLAGIDMVSGVRPPPVRRVTGR SSAPKKRIVRNMNDAVLLHQNETTVTNWIHQLSAGMFGRALNVLG AETADVENLTCDPASAKFVVPQRRLHKRLKWETRGHIPVSEEEYK HIYQGKKYAKFFEAVRAVDESKLTIKLGDLVYVLDQDPKVTQTQF ATAGREGRKKGAEKEKIQVRFGRVLSIRQPDSNSKDAQNVFIHVQ WLVLGCDTILQEMASRRELFLTDSCDTVFADVIYGVAKLTPLGAK DIPTVEFHESMATMMGENEFFVRFKYNYQDGSFTDLKDVDAEQIG TLQPRVNTHRNPGYCSNCRIKYDNERTGDKWIYENDTEGEPRLFR SSKGWCIYAQEFVYLQPVEKQPGTTFRVGYISEINKSSVIVELLA RVDDDDKSGHISYSDPRHLYFTGTDIKVTFDKIIRKCFVFHDSGD QKAKAPLMYGTLQRDLYYYRYEKRKGKAELVPVREIRSIHEQTLN DWESRTQIERHGAVSGKKLKGLDIFAGCGGLTLGLDLSGAVDTKW DIEFAPSAANTLALNFPDAQVFNQCANVLLSRAIQSEDEGSLDIE YDLQGRVLPDLPKKGEVDFIYGGPPCQGFSGVNRYKKGNDIKNSL VATFLSYVDHYKPRFVLLENVKGLITTKLGNSKNAEGKWEGGISN GVVKFIYRTLISMNYQCRIGLVQSGEYGVPQSRPRVIFLAARMGE RLPDLPEPMHAFEVLDSQYALPHIKRYHTTQNGVAPLPRITIGEA VSDLPKFQYANPGVWPRHDPYSSAKAQPSDKTIEKFSVSKATSFV GYLLQPYHSRPQSEFQRRLRTKLVPSDEPAEKTSLLTTKLVTAHV

SEQ	Description	Sequence
		TRLFNKETTQRIVCVPMWPGADHRSLPKEMRPWCLVDPNSQAEKH RFWPGLFGRLGMEDFFSTALTDVQPCGKQGKVLHPTQRRVYTVRE LARAQGFPDWFAFTDGDADSGLGGVKKWHRNIGNAVPVPLGEQIG RCIGYSVWWKDDMIAQLREDGADEDEEMIDGNDQWVEELNTQMAA DMPGLPLLVTHLLNLCVYRRLYGPNAKEFLPARVYDKKLEGGRRR LVWAML
614	*Neurospora* Dim2	MDSPDRSHGGMFIDVPAETMGFQEDYLDMFASVLSQGLAKEGDYA HHQPLPAGKEECLEPIAVATTITPSPDDPQLQLQLELEQQFQTES GLNGVDPAPAPESEDEADLPDGFSDESPDDDFVVQRSKHITVDLP VSTLINPRSTFQRIDENDNLVPPPQSTPERVAVEDLLKAAKAAGK NKEDYIEFELHDFNFYVNYAYHPQEMRPIQLVATKVLHDKYYFDG VLKYGNTKHYVTGMQVLELPVGNYGASLHSVKGQIWVRSKHNAKK EIYYLLKKPAFEYQRYYQPFLWIADLGKHVVDYCTRMVERKREVT LGCFKSDFIQWASKAHGKSKAFQNWRAQHPSDDFRTSVAANIGYI WKEINGVAGAKRAAGDQLFRELMIVKPGQYFRQEVPPGPVVTEGD RTVAATIVTPYIKECFGHMILGKVLRLAGEDAEKEKEVKLAKRLK IENKNATKADTKDDMKNDTATESLPTPLRSLPVQVLEATPIESDI VSIVSSDLPPSENNPPPLTNGSVKPKAKANPKPKPSTQPLHAAHV KYLSQELVNKIKVGDVISTPRDDSSNTDTKWKPTDTDDHRWFGLV QRVHTAKTKSSGRGLNSKSFDVIWFYRPEDTPCCAMKYKWRNELF LSNHCTCQEGHHARVKGNEVLAVHPVDWFGTPESNKGEFFVRQLY ESEQRRWITLQKDHLTCYHNQPPKPPTAPYKPGDTVLATLSPSDK FSDPYEVVEYFTQGEKETAFVRLRKLLRRRKVDRQDAPANELVYT EDLVDVRAERIVGKCIMRCFRPDERVPSPYDRGGTGNMFFITHRQ DHGRCVPLDTLPPTLRQGFNPLGNLGKPKLRGMDLYCGGGNFGRG LEEGGVVEMRWANDIWDKAIHTYMANTPDPNKTNPFLGSVDDLLR LALEGKFSDNVPRPGEVDFIAAGSPCPGFSLLTQDKKVLNQVKNQ SLVASFASFVDFYRPKYGVLENVSGIVQTFVNRKQDVLSQLFCAL VGMGYQAQLILGDAWAHGAPQSRERVFLYFAAPGLPLPDPPLPSH SHYRVKNRNIGFLCNGESYVQRSFIPTAFKFVSAGEGTADLPKIG DGKPDACVRFPDHRLASGITPYIRAQYACIPTHPYGMNFIKAWNN GNGVMSKSDRDLFPSEGKTRTSDASVGWKRLNPKTLFPTVTTTSN PSDARMGPGLHWDEDRPYTVQEMRRAQGYLDEEVLVGRTTDQWKL VGNSVSRHMALAIGLKFREAWLGTLYDESAVVATATATATTAAAV GVTVPVMEEPGIGTTESSRPSRSPVHTAVDLDDSKSERSRSTTPA TVLSTSSAAGDGSANAAGLEDDDNDDMEMMEVTRKRSSPAVDEEG MRPSKVQKVEVTVASPASRRSSRQASRNPTASPSSKASKATTHEA PAPEELESDAESYSETYDKEGFDGDYHSGHEDQYSEEDEEEEYAE PETMTVNGMTIVKL
615	*Drosophila* dDnmt2	MVFRVLELFSGIGGMHYAFNYAQLDGQIVAALDVNTVANAVYAHN YGSNLVKTRNIQSLSVKEVTKLQANMLLMSPPCQPHTRQGLQRDT EDKRSDALTHLCGLIPECQELEYILMENVKGFESSQARNQFIESL ERSGFHWREFILTPTQFNVPNTRYRYYCIARKGADFPFAGGKIWE EMPGAIAQNQGLSQIAEIVEENVSPDFLVPDDVLTKRVLVMDIIH PAQSRSMCFTKGYTHYTEGTGSAYTPLSEDESHRIFELVKEIDTS NQDASKSEKILQQRLDLLHQVRLRYFTPREVARLMSFPENFEFPP ETTNRQKYRLLGNSINVKVVGELIKLLTIK
616	*S. pombe* Pmt1	MLSTKRLRVLELYSGIGGMHYALNLANIPADIVCAIDINPQANEI YNLNHGKLAKHMDISTLTAKDFDAFDCKLWTMSPSCQPFTRIGNR KDILDPRSQAFLNILNVLPHVNNLPEYILIENVQGFEESKAAEEC RKVLRNCGYNLIEGILSPNQFNIPNSRSRWYGLARLNFKGEWSID DVFQFSEVAQKEGEVKRIRDYLEIERDWSSYMVLESVLNKWGHQF DIVKPDSSSCCCFTRGYTHLVQGAGSILQMSDHENTHEQFERNRM ALQLRYFTAREVARLMGFPESLEWSKSNVTEKCMYRLLGNSINVK VVSYLISLLLEPLNF

SEQ	Description	Sequence
617	*Arabidopsis* DRM1	MVMSHIFLISQIQEVEHGDSDDVNWNTDDDELAIDNFQFSPSPVH ISATSPNSIQNRISDETVASFVEMGFSTQMIARAIEETAGANMEP MMILETLFNYSASTEASSSKSKVINHFIAMGFPEEHVIKAMQEHG DEDVGEITNALLTYAEVDKLRESEDMNININDDDDDNLYSLSSDD EEDELNNSSNEDRILQALIKMGYLREDAAIAIERCGEDASMEEVV DFICAAQMARQFDEIYAEPDKKELMNNNKKRRTYTETPRKPNTDQ LISLPKEMIGFGVPNHPGLMMHRPVPIPDIARGPPFFYYENVAMT PKGVWAKISSHLYDIVPEFVDSKHFCAAARKRGYIHNLPIQNRFQ IQPPQHNTIQEAFPLTKRWWPSWDGRTKLNCLLTCIASSRLTEKI REALERYDGETPLDVQKWVMYECKKWNLVWVGKNKLAPLDADEME KLLGFPRDHTRGGGISTTDRYKSLGNSFQVDTVAYHLSVLKPLFP NGINVLSLFTGIGGGEVALHRLQIKMNVVVSVEISDANRNILRSF WEQTNQKGILREFKDVQKLDDNTIERLMDEYGGFDLVIGGSPCNN LAGGNRHHRVGLGGEHSSLFFDYCRILEAVRRKARHMRR
618	*Arabadopsis* DRM2	MVIWNNDDDDFLEIDNFQSSPRSSPIHAMQCRVENLAGVAVTTSS LSSPTETTDLVQMGFSDEVFATLFDMGFPVEMISRAIKETGPNVE TSVIIDTISKYSSDCEAGSSKSKAIDHFLAMGFDEEKVVKAIQEH GEDNMEAIANALLSCPEAKKLPAAVEEEDGIDWSSSDDDTNYTDM LNSDDEKDPNSNENGSKIRSLVKMGFSELEASLAVERCGENVDIA ELTDFLCAAQMAREFSEFYTEHEEQKPRHNIKKRRFESKGEPRSS VDDEPIRLPNPMIGFGVPNEPGLITHRSLPELARGPPFFYYENVA LTPKGVWETISRHLFEIPPEFVDSKYFCVAARKRGYIHNLPINNR FQIQPPPKYTIHDAFPLSKRWWPEWDKRTKLNCILTCTGSAQLTN RIRVALEPYNEEPEPPKHVQRYVIDQCKKWNLVWVGKNKAAPLEP DEMESILGFPKNHTRGGGMSRTERFKSLGNSFQVDTVAYHLSVLK PIFPHGINVLSLFTGIGGGEVALHRLQIKMKLVVSVEISKVNRNI LKDFWEQTNQTGELIEFSDIQHLTNDTIEGLMEKYGGFDLVIGGS PCNNLAGGNRVSRVGLEGDQSSLFFEYCRILEVVRARMRGS
619	Arabadopsis CMT1	MAARNKQKKRAEPESDLCFAGKPMSVVESTIRWPHRYQSKKTKLQ APTKKPANKGGKKEDEEIIKQAKCHFDKALVDGVLINLNDDVYVT GLPGKLKFIAKVIELFEADDGVPYCRFRWYYRPEDTLIERFSHLV QPKRVFLSNDENDNPLTCIWSKVNIAKVPLPKITSRIEQRVIPPC DYYYDMKYEVPYLNFTSADDGSDASSSLSSDSALNCFENLHKDEK FLLDLYSGCGAMSTGFCMGASISGVKLITKWSVDINKFACDSLKL NHPETEVRNEAAEDFLALLKEWKRLCEKFSLVSSTEPVESISELE DEEVEENDDIDEASTGAELEPGEFEVEKFLGIMFGDPQGTGEKTL QLMVRWKGYNSSYDTWEPYSGLGNCKEKLKEYVIDGFKSHLLPLP GTVYTVCGGPPCQGISGYNRYRNNEAPLEDQKNQQLLVFLDIIDF LKPNYVLMENVVDLLRFSKGFLARHAVASFVAMNYQTRLGMMAAG SYGLPQLRNRVFLWAAQPSEKLPPYPLPTHEVAKKFNTPKEFKDL QVGRIQMEFLKLDNALTLADAISDLPPVTNYVANDVMDYNDAAPK TEFENFISLKRSETLLPAFGGDPTRRLFDHQPLVLGDDDLERVSY IPKQKGANYRDMPGVLVHNNKAEINPRFRAKLKSGKNVVPAYAIS FIKGKSKKPFGRLWGDEIVNTVVTRAEPHNQCVIHPMQNRVLSVR ENARLQGFPDCYKLCGTIKEKYIQVGNAVAVPVGVALGYAFGMAS QGLTDDEPVIKLPFKYPECMQAKDQI
620	*Arabadopsis* CMT2	MLSPAKCESEEAQAPLDLHSSSRSEPECLSLVLWCPNPEEAAPSS TRELIKLPDNGEMSLRRSTTLNCNSPEENGGEGRVSQRKSSRGKS QPLLMLTNGCQLRRSPRFRALHANFDNVCSVPVTKGGVSQRKFSR GKSQPLLTLTNGCQLRRSPRFRAVDGNFDSVCSVPVTGKFGSRKR KSNSALDKKESSDSEGLTFKDIAVIAKSLEMEIISECQYKNNVAE GRSRLQDPAKRKVDSDTLLYSSINSSKQSLGSNKRMRRSQRFMKG TENEGEENLGKSKGKGMSLASCSFRRSTRLSGTVETGNTETLNRR KDCGPALCGAEQVRGTERLVQISKKDHCCEAMKKCEGDGLVSSKQ ELLVFPSGCIKKTVNGCRDRTLGKPRSSGLNTDDIHTSSLKISKN

SEQ	Description	Sequence
		DTSNGLTMTTALVEQDAMESLLQGKTSACGAADKGKTREMHVNST VIYLSDSDEPSSIEYLNGDNLTQVESGSALSSGGNEGIVSLDLNN PTKSTKRKGKRVTRTAVQEQNKRSICFFIGEPLSCEEAQERWRWR YELKERKSKSRGQQSEDDEDKIVANVECHYSQAKVDGHTFSLGDF AYIKGEEEETHVGQIVEFFKTTDGESYFRVQWFYRATDTIMERQA TNHDKRRLFYSTVMNDNPVDCLISKVTVLQVSPRVGLKPNSIKSD YYFDMEYCVEYSTFQTLRNPKTSENKLECCADVVPTESTESILKK KSFSGELPVLDLYSGCGGMSTGLSLGAKISGVDVVTKWAVDQNTA ACKSLKLNHPNTQVRNDAAGDFLQLLKEWDKLCKRYVFNNDQRTD TLRSVNSTKETSGSSSSSDDDSDSEEYEVEKLVDICFGDHDKTGK NGLKFKVHWKGYRSDEDTWELAEELSNCQDAIREFVTSGFKSKIL PLPGRVGVICGGPPCQGISGYNRHRNVDSPLNDERNQQIIVFMDI VEYLKPSYVLMENVVDILRMDKGSLGRYALSRLVNMRYQARLGIM TAGCYGLSQFRSRVFMWGAVPNKNLPPFPLPTHDVIVRYGLPLEF ERNVVAYAEGQPRKLEKALVLKDAISDLPHVSNDEDREKLPYESL PKTDFQRYIRSTKRDLTGSAIDNCNKRTMLLHDHRPFHINEDDYA RVCQIPKRKGANFRDLPGLIVRNNTVCRDPSMEPVILPSGKPLVP GYVFTFQQGKSKRPFARLWWDETVPTVLTVPTCHSQALLHPEQDR VLTIRESARLQGFPDYFQFCGTIKERYCQIGNAVAVSVSRALGYS LGMAFRGLARDEHLIKLPQNFSHSTYPQLQETIPH
621	*Arabadopsis* CMT3	MAPKRKRPATKDDTTKSIPKPKKRAPKRAKTVKEEPVTVVEEGEK HVARFLDEPIPESEAKSTWPDRYKPIEVQPPKASSRKKTKDDEKV EIIRARCHYRRAIVDERQIYELNDDAYVQSGEGKDPFICKIIEMF EGANGKLYFTARWFYRPSDTVMKEFEILIKKKRVFFSEIQDTNEL GLLEKKLNILMIPLNENTKETIPATENCDFFCDMNYFLPYDTFEA IQQETMMAISESSTISSDTDIREGAAAISEIGECSQETEGHKKAT LLDLYSGCGAMSTGLCMGAQLSGLNLVTKWAVDMNAHACKSLQHN HPETNVRNMTAEDFLFLLKEWEKLCIHFSLRNSPNSEEYANLHGL NNVEDNEDVSEESENEDDGEVFTVDKIVGISFGVPKKLLKRGLYL KVRWLNYDDSHDTWEPIEGLSNCRGKIEEFVKLGYKSGILPLPGG VDVVCGGPPCQGISGHNRFRNLLDPLEDQKNKQLLVYMNIVEYLK PKFVLMENVVDMLKMAKGYLARFAVGRLLQMNYQVRNGMMAAGAY GLAQFRLRFFLWGALPSEIIPQFPLPTHDLVHRGNIVKEFQGNIV AYDEGHTVKLADKLLLKDVISDLPAVANSEKRDEITYDKDPTTPF QKFIRLRKDEASGSQSKSKSKKHVLYDHHPLNLNINDYERVCQVP KRKGANFRDFPGVIVGPGNVVKLEEGKERVKLESGKTLVPDYALT YVDGKSCKPFGRLWWDEIVPTVVTRAEPHNQVIIHPEQNRVLSIR ENARLQGFPDDYKLFGPPKQKYIQVGNAVAVPVAKALGYALGTAF QGLAVGKDPLLTLPEGFAFMKPTLPSELA
622	*Neurospora* Rid	MAEQNPFVIDDEDDVIQIHDEEEVEEEVAEVIDITEDDIEPSELD RAFGSRPKEETLPSLLLRDQGFIVRPGMTVELKAPIGRFAISFVR VNSIVKVRQAHVNNVTIRGHGFTRAKEMNGMLPKQLNECCLVASI DTRDPRP
623	*E. coli* strain 12 hsdM	MNNNDLVAKLWKLCDNLRDGGVSYQNYVNELASLLFLKMCKETGQ EAEYLPEGYRWDDLKSRIGQEQLQFYRKMLVHLGEDDKKLVQAVF HNVSTTITEPKQITALVSNMDSLDWYNGAHGKSRDDFGDMYEGLL QKNANETKSGAGQYFTPRPLIKTIIHLLKPQPREVVQDPAAGTAG FLIEADRYVKSQTNDLDDLDGDTQDFQIHRAFIGLELVPGTRRLA LMNCLLHDIEGNLDHGGAIRLGNTLGSDGENLPKAHIVATNPPFG SAAGTNITRTFVHPTSNKQLCFMQHIIETLHPGGRAAVVVPDNVL FEGGKGTDIRRDLMDKCHLHTILRLPTGIFYAQGVKTNVLFFTKG TVANPNQDKNCTDDVWVYDLRTNMPSFGKRTPFTDEHLQPFERVY GEDPHGLSPRTEGEWSFNAEETEVADSEENKNTDQHLATSRWRKF SREWIRTAKSDSLDISWLKDKDSIDADSLPEPDVLAAEAMGELVQ ALSELDALMRELGASDEADLQRQLLEEAFGGVKE

SEQ	Description	Sequence
624	*E. coli* strain 12 hsdS	MSAGKLPEGWVIAPVSTVTTLIRGVTYKKEQAINYLKDDYLPLIR ANNIQNGKFDTTDLVFVPKNLVKESQKISPEDIVIAMSSGSKSVV GKSAHQHLPFECSFGAFCGVLRPEKLIFSGFIAHFTKSSLYRNKI SSLSAGANINNIKPASFDLINIPIPPLAEQKIIAEKLDTLLAQVD STKARFEQIPQILKRFRQAVLGGAVNGKLTEKWRNFEPQHSVFKK LNFESILTELRNGLSSKPNESGVGHPILRISSVRAGHVDQNDIRF LECSESELNRHKLQDGDLLFTRYNGSLEFVGVCGLLKKLQHQNLL YPDKLIRARLTKDALPEYIEIFFSSPSARNAMMNCVKTTSGQKGI SGKDIKSQVVLLPPVKEQAEIVRRVEQLFAYADTIEKQVNNALAR VNNLTQSILAKAFRGELTAQWRAENPDLISGENSAAALLEKIKAE RAASGGKKASRKKS
625	*T. aquaticus* M TaqI	MGLPPLLSLPSNSAPRSLGRVETPPEVVDFMVSLAEAPRGGRVLE PACAHGPFLRAFREAHGTAYRFVGVEIDPKALDLPPWAEGILADF LLWEPGEAFDLILGNPPYGIVGEASKYPIHVFKAVKDLYKKAFST WKGKYNLYGAFLEKAVRLLKPGGVLVFVVPATWLVLEDFALLREF LAREGKTSVYYLGEVFPQKKVSAVVIRFQKSGKGLSLWDTQESES GFTPILWAEYPHWEGEIIRFETEETRKLEISGMPLGDLFHIRFAA RSPEFKKHPAVRKEPGPGLVPVLTGRNLKPGWVDYEKNHSGLWMP KERAKELRDFYATPHLVVAHTKGTRVVAAWDERAYPWREEFHLLP KEGVRLDPSSLVQWLNSEAMQKHVRTLYRDFVPHLTLRMLERLPV RREYGFHTSPESARNF
626	*E. coli* M EcoDam	MKKNRAFLKWAGGKYPLLDDIKRHLPKGECLVEPFVGAGSVFLNT DFSRYILADINSDLISLYNIVKMRTDEYVQAARELFVPETNCAEV YYQFREEFNKSQDPFRRAVLFLYLNRYGYNGLCRYNLRGEFNVPF GRYKKPYFPEAELYHFAEKAQNAFFYCESYADSMARADDASVVYC DPPYAPLSATANFTAYHTNSFTLEQQAHLAEIAEGLVERHIPVLI SNHDTMLTREWYQRAKLHVVKVRRSISSNGGTRKKVDELLALYKP GVVSPAKK
627	*C. crescentus* M CcrMI	MKFGPETIIHGDCIEQMNALPEKSVDLIFADPPYNLQLGGDLLRP DNSKVDAVDDHWDQFESFAAYDKFTREWLKAARRVLKDDGAIWVI GSYHNIFRVGVAVQDLGFWILNDIVWRKSNPMPNFKGTRFANAHE TLIWASKSQNAKRYTFNYDALKMANDEVQMRSDWTIPLCTGEERI KGADGQKAHPTQKPEALLYRVILSTTKPGDVILDPFFGVGTTGAA AKRLGRKFIGIEREAEYLEHAKARIAKVVPIAPEDLDVMGSKRAE PRVPFGTIVEAGLLSPGDTLYCSKGTHVAKVRPDGSITVGDLSGS IHKIGALVQSAPACNGWTYWHFKTDAGLAPIDVLRAQVRAGMN
628	*C. difficile* CamA	MDDISQDNFLLSKEYENSLDVDTKKASGIYYTPKIIVDYIVKKTL KNHDIIKNPYPRILDISCGCGNFLLEVYDILYDLFEENIYELKKK YDENYWTVDNIHRHILNYCIYGADIDEKAISILKDSLTNKKVVND LDESDIKINLFCCDSLKKKWRYKFDYIVGNPPYIGHKKLEKKYKK FLLEKYSEVYKDKADLYFCFYKKIIDILKQGGIGSVITPRYFLES LSGKDLREYIKSNVNVQEIVDFLGANIFKNIGVSSCILTFDKKKT KETYIDVFKIKNEDICINKFETLEELLKSSKFEHFNINQRLLSDE WILVNKDDETFYNKIQEKCKYSLEDIAISFQGIITGCDKAFILSK DDVKLNLVDDKFLKCWIKSKNINKYIVDKSEYRLIYSNDIDNENT NKRILDEIIGLYKTKLENRRECKSGIRKWYELQWGREKLFFERKK IMYPYKSNENRFAIDYDNNFSSADVYSFFIKEEYLDKFSYEYLVG ILNSSVYDKYFKITAKKMSKNIYDYYPNKVMKIRIFRDNNYEEIE NLSKQIISILLNKSIDKGKVEKLQIKMDNLIMDSLGI
629	KAP1	MAASAAAASAAAASAASGSPGPGEGSAGGEKRSTAPSAAASASAS AAASSPAGGGAEALELLEHCGVCRERLRPEREPRLLPCLHSACSA CLGPAAPAAANSSGDGGAAGDGTVVDCPVCKQQCFSKDIVENYFM RDSGSKAATDAQDANQCCTSCEDNAPATSYCVECSEPLCETCVEA HQRVKYTKDHTVRSTGPAKSRDGERTVYCNVHKHEPLVLFCESCD TLTCRDCQLNAHKDHQYQFLEDAVRNQRKLLASLVKRLGDKHATL

SEQ	Description	Sequence
		QKSTKEVRSSIRQVSDVQKRVQVDVKMAILQIMKELNKRGRVLVN DAQKVTEGQQERLERQHWTMTKIQKHQEHILRFASWALESDNNTA LLLSKKLIYFQLHRALKMIVDPVEPHGEMKFQWDLNAWTKSAEAF GKIVAERPGTNSTGPAPMAPPRAPGPLSKQGSGSSQPMEVQEGYG FGSGDDPYSSAEPHVSGVKRSRSGEGEVSGLMRKVPRVSLERLDL DLTADSQPPVFKVFPGSTTEDYNLIVIERGAAAAATGQPGTAPAG TPGAPPLAGMAIVKEEETEAAIGAPPTATEGPETKPVLMALAEGP GAEGPRLASPSGSTSSGLEVVAPEGTSAPGGGPGTLDDSATICRV CQKPGDLVMCNQCEFCFHLDCHLPALQDVPGEEWSCSLCHVLPDL KEEDGSLSLDGADSTGVVAKLSPANQRKCERVLLALFCHEPCRPL HQLATDSTFSLDQPGGTLDLTLIRARLQEKLSPPYSSPQEFAQDV GRMFKQFNKLTEDKADVQSIIGLQRFFETRMNEAFGDTKFSAVLV EPPPMSLPGAGLSSQELSGGPGDGP
630	MECP2	MVAGMLGLREEKSEDQDLQGLKDKPLKFKKVKKDKKEEKEGKHEP VQPSAHHSAEPAEAGKAETSEGSGSAPAVPEASASPKQRRSIIRD RGPMYDDPTLPEGWTRKLKQRKSGRSAGKYDVYLINPQGKAFRSK VELIAYFEKVGDTSLDPNDFDFTVTGRGSPSRREQKPPKKPKSPK APGTGRGRGRPKGSGTTRPKAATSEGVQVKRVLEKSPGKLLVKMP FQTSPGGKAEGGGATTSTQVMVIKRPGRKRKAEADPQAIPKKRGR KPGSVVAAAAAEAKKKAVKESSIRSVQETVLPIKKRKTRETVSIE VKEVVKPLLVSTLGEKSGKGLKTCKSPGRKSKESSPKGRSSSASS PPKKEHHHHHHHSESPKAPVPLLPPLPPPPPEPESSEDPTSPPEP QDLSSSVCKEEKMPRGGSLESDGCPKEPAKTQPAVATAATAAEKY KHRGEGERKDIVSSSMPRPNREEPVDSRTPVTERVS
631	linker	SGGS
632	linker	SGGSSGSETPGTSESATPESSGGS
633	linker	SGGSSGGSSGSETPGTSESATPESSGGSSGGS
634	linker	GGSGGSPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGS PAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESG PGSEPATSGGSGGS
635	G linker	GSGGG
636	GX4 linker	GGGGSGGGGSGGGGSGGGGS
637	W linker	SSGNSNANSRGPSFSSGLVPLSLRGSH
638	XTEN linker (XTEN16)	SGSETPGTSESATPES
639	XTEN linker	SGGSSGGSSGSETPGTSESATPES
640	XTEN linker	SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGS
641	XTEN linker	SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGSSGSET PGTSESATPESSGGSSGGS
642	XTEN linker	PGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPT STEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPA TS
643	XTEN linker (XTEN80)	GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEG SAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSE
644	NLS	PKKKRKV
645	NLS	AVKRPAATKKAGQAKKKKLD
646	NLS	MSRRRKANPTKLSENAKKLAKEVEN
647	NLS	PAAKRVKLD
648	NLS	KLKIKRPVK
649	NLS	MDSLLMNRRKFLYQFKNVRWAKGRRETYLC
660	fusion protein (Configuration 7)	MGTMPKKKRKVPKKKRKVYNHDQEFDPPKVYPPVPAEKRKPIRVL SLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIM YVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGT

SEQ	Description	Sequence
		GRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRF LESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQE CLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCT EMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKE YFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMAAIPALDPEAEP SMDVILVGSSELSSSVSPGTGRDLIAYEVKANQRNIEDICICCGS LQVHTQHPLFEGGICAPCKDKFLDALFLYDDDGYQSYCSICCSGE TLLICGNPDCTRCYCFECVDSLVGPGTSGKVHAMSNWVCYLCLPS SRSGLLQRRRKWRSQLKAFYDRESENPLEMFETVPVWRRQPVRVL SLFEDIKKELTSLGFLESGSDPGQLKHVVDVTDTVRKDVEEWGPF DLVYGATPPLGHTCDRPPSWYLFQFHRLLQYARPKPGSPRPFFWM FVDNLVLNKEDLDVASRFLEMEPVTIPDVHGGSLQNAVRVWSNIP AIRSRHWALVSEEELSLLAQNKQSSKLAAKWPTKLVKNCFLPLRE YFKYFSTELTSSLGGPSSGAPPPSGGSPAGSPTSTEEGTSESATP ESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTE PSELEDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRH SIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIF SNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEK YPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR LENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQL SKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTE ITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSK NGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRK QRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTF RIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSF IERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRK PAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEIS GVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLF EDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR DKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQ GDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVI EMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDS IDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDK GRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIA RKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGI TIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRK RMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKP IREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDAT LIHQSITGLYETRIDLSQLGGDSPKKKRKVGVDGSSGSETPGTSE SATPESRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKN LVSLGYQLTKPDVILRLEKGEEPSADYKDDDDKAPKKKRKVPKKK RKV
661	fusion protein (Configuration 9)	MGTMPKKKRKVPKKKRKVYNHDQEFDPPKVYPPVPAEKRKPIRVL SLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIM YVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGT GRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRF LESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQE CLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCT

SEQ	Description	Sequence
		EMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKE YFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMAAIPALDPEAEP SMDVILVGSSELSSSVSPGTGRDLIAYEVKANQRNIEDICICCGS LQVHTQHPLFEGGICAPCKDKFLDALFLYDDDGYQSYCSICCSGE TLLICGNPDCTRCYCFECVDSLVGPGTSGKVHAMSNWVCYLCLPS SRSGLLQRRRKWRSQLKAFYDRESENPLEMFETVPVWRRQPVRVL SLFEDIKKELTSLGFLESGSDPGQLKHVVDVTDTVRKDVEEWGPF DLVYGATPPLGHTCDRPPSWYLFQFHRLLQYARPKPGSPRPFFWM FVDNLVLNKEDLDVASRFLEMEPVTIPDVHGGSLQNAVRVWSNIP AIRSRHWALVSEEELSLLAQNKQSSKLAAKWPTKLVKNCFLPLRE YFKYFSTELTSSLGGPSSGAPPPSGGSPAGSPTSTEEGTSESATP ESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTE PSELEDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRH SIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIF SNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEK YPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR LENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQL SKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTE ITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSK NGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRK QRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTF RIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSF IERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRK PAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEIS GVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLF EDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR DKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQ GDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVI EMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDS IDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDK GRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIA RKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGI TIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRK RMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKP IREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDAT LIHQSITGLYETRIDLSQLGGDSPKKKRKVGVDGSSGSETPGTSE SATPESTGNKKLEAVGTGIEPKAMSQGLVTFGDVAVDFSQEEWEW LNPIQRNLYRKVMLENYRNLASLGLCVSKPDVISSLEQGKEPWSA DYKDDDDKAPKKKRKVPKKKRKV
662	fusion protein (Configuration 11)	MGTMPKKKRKVPKKKRKVYNHDQEFDPPKVYPPVPAEKRKPIRVL SLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIM YVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGT GRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRF LESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQE CLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCT EMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKE YFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMAAIPALDPEAEP SMDVILVGSSELSSSVSPGTGRDLIAYEVKANQRNIEDICICCGS

SEQ	Description	Sequence
		LQVHTQHPLFEGGICAPCKDKFLDALFLYDDDGYQSYCSICCSGE TLLICGNPDCTRCYCFECVDSLVGPGTSGKVHAMSNWVCYLCLPS SRSGLLQRRRKWRSQLKAFYDRESENPLEMFETVPVWRRQPVRVL SLFEDIKKELTSLGFLESGSDPGQLKHVVDVTDTVRKDVEEWGPF DLVYGATPPLGHTCDRPPSWYLFQFHRLLQYARPKPGSPRPFFWM FVDNLVLNKEDLDVASRFLEMEPVTIPDVHGGSLQNAVRVWSNIP AIRSRHWALVSEEELSLLAQNKQSSKLAAKWPTKLVKNCFLPLRE YFKYFSTELTSSLGGPSSGAPPPSGGSPAGSPTSTEEGTSESATP ESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTE PSELEDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRH SIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIF SNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEK YPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR LENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQL SKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTE ITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSK NGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRK QRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTF RIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSF IERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRK PAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEIS GVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLF EDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR DKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQ GDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVI EMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDS IDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDK GRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIA RKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGI TIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRK RMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKP IREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDAT LIHQSITGLYETRIDLSQLGGDSPKKKRKVGVDGSSGSETPGTSE SATPESTGDSVAFEDVAVNFTLEEWALLDPSQKNLYRDVMRETFR NLASVGKQWEDQNIEDPFKIPRRNISHIPERLCESKEGGQGEESA DYKDDDDKAPKKKRKVPKKKRKV
663	fusion protein (Configuration 13)	MGTMPKKKRKVPKKKRKVYNHDQEFDPPKVYPPVPAEKRKPIRVL SLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIM YVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGT GRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRF LESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQE CLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCT EMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKE YFACVSSGNSNANSRGPSFSSGLVPLSLRGSHMAAIPALDPEAEP SMDVILVGSSELSSSVSPGTGRDLIAYEVKANQRNIEDICICCGS LQVHTQHPLFEGGICAPCKDKFLDALFLYDDDGYQSYCSICCSGE TLLICGNPDCTRCYCFECVDSLVGPGTSGKVHAMSNWVCYLCLPS SRSGLLQRRRKWRSQLKAFYDRESENPLEMFETVPVWRRQPVRVL

SEQ	Description	Sequence
		SLFEDIKKELTSLGFLESGSDPGQLKHVVDVTDTVRKDVEEWGPF DLVYGATPPLGHTCDRPPSWYLFQFHRLLQYARPKPGSPRPFFWM FVDNLVLNKEDLDVASRFLEMEPVTIPDVHGGSLQNAVRVWSNIP AIRSRHWALVSEEELSLLAQNKQSSKLAAKWPTKLVKNCFLPLRE YFKYFSTELTSSLGGPSSGAPPPSGGSPAGSPTSTEEGTSESATP ESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTE PSELEDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRH SIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIF SNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEK YPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR LENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQL SKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTE ITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSK NGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRK QRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTF RIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSF IERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRK PAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEIS GVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLF EDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR DKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQ GDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVI EMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDS IDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDK GRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIA RKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGI TIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRK RMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKP IREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDAT LIHQSITGLYETRIDLSQLGGDSPKKKRKVGVDGSSGSETPGTSE SATPESTGMNNSQGRVTFEDVTVNFTQGEWQRLNPEQRNLYRDVM LENYSNLVSVGQGETTKPDVILRLEQGKEPWLEEEEVLGSGRAEK NGDIGGQIWKPKDVKESLSADYKDDDDKAPKKKRKVPKKKRKV
664	linker	GGGGS
665	linker	EAAAK
631	linker	SGGS

CLAIMS

1. A system for repressing transcription of a human *PCSK9* gene in a human cell, optionally a human hepatocyte, comprising

a) one or more fusion proteins that collectively comprise

a DNA methyltransferase (DNMT) domain and/or a domain that recruits a DNMT, optionally wherein the DNMT domain and/or the recruiter domain comprise a DNMT3A domain and/or a DNMT3L domain, and optionally wherein the recruited DNMT is DNMT3A, and

a transcriptional repressor domain,

each domain being linked to a DNA-binding domain that binds to a target region in the human *PCSK9* gene; or

b) one or more nucleic acid molecules encoding the one or more fusion proteins.

2. The system of claim 1, wherein the DNA-binding domain binds to a target sequence in SEQ ID NO: 1488 or 1489.

3. The system of claim 1 or 2, wherein the DNA-binding domain targets the fusion protein(s) to one or more sequences in the *PCSK9* gene selected from SEQ ID NOs: 700-747 and 1036-1261.

4. The system of any one of claims 1-3, wherein the DNA-binding domain comprises a dead CRISPR Cas (dCas) domain, a ZFP domain, or a TALE domain.

5. The system of claim 4, wherein the DNA-binding domain comprises a dCas9 domain and the system further comprises (i) one or more guide RNAs comprising any one of SEQ ID NOs: 1262-1487, or (ii) nucleic acid molecules coding for the one or more guide RNAs.

6. The system of claim 4 or 5, wherein the dCas domain comprises a dCas9 sequence, optionally a sequence with at least 90% identity to SEQ ID NO: 12 or 13.

7. The system of claim 4, wherein the ZFP domain targets a nucleotide sequence selected from SEQ ID NOs: 700-747.

8. The system of claim 7, wherein the ZFP domain comprises the F1-F6 amino acid sequences of any one of ZF001 through ZF048 as shown in **Table 1**.

9. The system of any one of claims 1-8, wherein the DNMT3A domain comprises a sequence with at least 90% identity to SEQ ID NO: 574 or 575.

10. The system of any one of claims 1-9, wherein the DNMT3L domain comprises a sequence with at least 90% identity to a sequence selected from SEQ ID NOs: 578-581.

11. The system of any one of claims 1-9, wherein the DNMT3L domain comprises a sequence with at least 90% identity to a sequence selected from SEQ ID NOs: 582-603.

12. The system of any one of claims 1-8, wherein the DNMT domain comprises a sequence with at least 90% identity to a sequence selected from SEQ ID NOs: 601-603.

13. The system of any one of claims 1-12, wherein the transcriptional repressor domain comprises a sequence with at least 90% identity to a sequence selected from SEQ ID NOs: 33-570.

14. The system of any one of claims 1-12, wherein the transcriptional repressor domain comprises a KRAB domain derived from KOX1, ZIM3, ZFP28, or ZN627.

15. The system of claim 14, wherein the KRAB domain comprises a sequence with at least 90% identity to a sequence selected from SEQ ID NOs: 89, 116, 245, and 255.

16. The system of any one of claims 1-12, wherein the transcriptional repressor domain comprises a fusion of the N- and C-terminal regions of ZIM3 and KOX1 KRAB, and optionally comprises the amino acid sequence of SEQ ID NO: 571 or 572.

17. The system of any one of claims 1-12, wherein the transcriptional repressor domain is derived from KAP1, MECP2, HP1a/CBX5, HP1b, CBX8, CDYL2, TOX, TOX3, TOX4, EED, EZH2, RBBP4, RCOR1, or SCML2.

18. The system of any one of claims 1-17, wherein the system comprises

a) a fusion protein comprising the DNMT3A domain, the DNMT3L domain, the transcriptional repressor domain, and the DNA-binding domain,

optionally wherein one or both of the DNMT3A domain and the DNMT3L domain are human, and

optionally wherein the DNA-binding domain is a dead CRISPR Cas domain or a ZFP domain; or

b) a nucleic acid molecule encoding the fusion protein.

19. The system of claim 18, wherein the fusion protein comprises, from N-terminus to C-terminus, the DNMT3A domain, a first peptide linker, the DNMT3L domain, a second peptide linker, the DNA-binding domain, a third peptide linker, and the transcriptional repressor domain.

20. The system of claim 19, wherein the fusion protein comprises, from N-terminus to C-terminus, the DNMT3A domain, the first peptide linker, the DNMT3L domain, the second peptide linker, a first nuclear localization signal (NLS), the DNA-binding domain, a second NLS, the third peptide linker, and the transcriptional repressor domain.

21. The system of claim 18, wherein the fusion protein comprises, from N-terminus to C-terminus, a first nuclear localization signal (NLS), the DNMT3A domain, the first peptide linker, the DNMT3L domain, the second peptide linker, the DNA-binding domain, the third peptide linker, the transcriptional repressor domain, and a second NLS.

22. The system of claim 18, wherein the fusion protein comprises, from N-terminus to C-terminus, first and second nuclear localization signals (NLSs), the DNMT3A domain, the first peptide linker, the DNMT3L domain, the second peptide linker, the DNA-

binding domain, the third peptide linker, the transcriptional repressor domain, and third and fourth NLSs.

23. The system of any one of claims 18-22, wherein the transcriptional repressor domain is a KRAB domain, optionally a human KOX1, ZFP28, ZN627, or ZIM3 KRAB domain.

24. The system of any one of claims 19-23, wherein one or both of the second and third peptide linkers are XTEN linkers, optionally selected from XTEN80 and XTEN16, and further optionally wherein the second peptide linker is XTEN80, and the third peptide linker is XTEN16.

25. The system of claim 18, wherein the fusion protein comprises, from N-terminus to C-terminus, a human DNMT3A domain, a first peptide linker, a human DNMT3L domain, an XTEN80 peptide linker, a first NLS, a dSpCas9 domain, a second NLS, an XTEN16 peptide linker, and a human KOX1 KRAB domain.

26. The system of claim 25, wherein the fusion protein comprises SEQ ID NO: 658 or a sequence at least 90% identical thereto, or SEQ ID NO: 1495 or a sequence at least 90% identical thereto.

27. The system of claim 18, wherein the fusion protein comprises, from N-terminus to C-terminus, a human DNMT3A domain, a first peptide linker, a human DNMT3L domain, an XTEN80 peptide linker, a first NLS, a ZFP domain, a second NLS, an XTEN16 linker, and a human KOX1 KRAB domain.

28. The system of claim 27, wherein the fusion protein comprises SEQ ID NO: 659 or a sequence at least 90% identical thereto, or SEQ ID NO: 1496 or a sequence at least 90% identical thereto, optionally wherein the ZFP comprises the F1-F6 amino acid sequences of any one of ZF001 through ZF048 as shown in **Table 1**.

29. The system of claim 18, wherein the fusion protein comprises, from N-terminus to C-terminus, first and second NLSs, a human DNMT3A domain, a first peptide linker, a

human DNMT3L domain, an XTEN80 peptide linker, a dSpCas9 domain, an XTEN16 peptide linker, a human KOX1 KRAB domain, and third and fourth NLSs.

30. The system of claim 29, wherein the fusion protein comprises SEQ ID NO: 660 or a sequence at least 90% identical thereto.

31. The system of claim 18, wherein the fusion protein comprises, from N-terminus to C-terminus, first and second NLSs, a human DNMT3A domain, a first peptide linker, a human DNMT3L domain, an XTEN80 peptide linker, a ZFP domain, an XTEN16 peptide linker, a human KOX1 KRAB domain, and third and fourth NLSs, optionally wherein the fusion protein comprises SEQ ID NO: 1514 or a sequence at least 90% identical thereto, or SEQ ID NO: 1523 or a sequence at least 90% identical thereto.

32. The system of claim 18, wherein the fusion protein comprises, from N-terminus to C-terminus, first and second NLSs, a human DNMT3A domain, a first peptide linker, a human DNMT3L domain, an XTEN80 peptide linker, a dSpCas9 domain, an XTEN16 peptide linker, a human ZFP28 KRAB domain, and third and fourth NLSs.

33. The system of claim 32, wherein the fusion protein comprises SEQ ID NO: 661 or a sequence at least 90% identical thereto, or SEQ ID NO: 1516 or a sequence at least 90% identical thereto.

34. The system of claim 18, wherein the fusion protein comprises, from N-terminus to C-terminus, first and second NLSs, a human DNMT3A domain, a first peptide linker, a human DNMT3L domain, an XTEN80 peptide linker, a ZFP domain, an XTEN16 peptide linker, a human ZFP28 KRAB domain, and third and fourth NLSs.

35. The system of claim 18, wherein the fusion protein comprises, from N-terminus to C-terminus, first and second NLSs, a human DNMT3A domain, a first peptide linker, a human DNMT3L domain, an XTEN80 peptide linker, a dSpCas9 domain, an XTEN16 peptide linker, a human ZN627 KRAB domain, and third and fourth NLSs.

36. The system of claim 35, wherein the fusion protein comprises SEQ ID NO: 662 or a sequence at least 90% identical thereto, or SEQ ID NO: 1520 or a sequence at least 90% identical thereto.

37. The system of claim 18, wherein the fusion protein comprises, from N-terminus to C-terminus, first and second NLSs, a human DNMT3A domain, a first peptide linker, a human DNMT3L domain, an XTEN80 peptide linker, a ZFP domain, an XTEN16 peptide linker, a human ZN627 KRAB domain, and third and fourth NLSs.

38. The system of claim 18, wherein the fusion protein comprises, from N-terminus to C-terminus, first and second NLSs, a human DNMT3A domain, a first peptide linker, a human DNMT3L domain, an XTEN80 peptide linker, a dSpCas9 domain, an XTEN16 peptide linker, a human ZIM3 KRAB domain, and third and fourth NLSs.

39. The system of claim 38, wherein the fusion protein comprises SEQ ID NO: 663 or a sequence at least 90% identical thereto or SEQ ID NO: 1518 or a sequence at least 90% identical thereto.

40. The system of claim 18, wherein the fusion protein comprises, from N-terminus to C-terminus, first and second NLSs, a human DNMT3A domain, a first peptide linker, a human DNMT3L domain, an XTEN80 peptide linker, a ZFP domain, an XTEN16 peptide linker, a human ZIM3 KRAB domain, and third and fourth NLSs.

41. The system of any one of claims 20-40, wherein at least one of the NLSs is an SV40 NLS.

42. The system of any one of claims 1-17, wherein the system comprises:

a) a first fusion protein comprising a first DNA-binding domain and comprising or recruiting the DNMT3A domain,

a second fusion protein comprising a second DNA-binding domain and comprising or recruiting the DNMT3L domain, and

a third fusion protein comprising a third DNA-binding domain and comprising or recruiting the transcriptional repressor domain; or

b) one or more nucleic acid molecules encoding the fusion proteins.

43. A human cell comprising the system of any one of claims 1-42, or progeny of the cell, optionally wherein the cell is a hepatocyte.

44. A pharmaceutical composition comprising the system of any one of claims 1-42 and a pharmaceutically acceptable excipient, optionally wherein

the composition comprises lipid nanoparticles (LNPs) comprising the system, and/or

the DNA-binding domain is a dCas domain and the LNPs further comprise one or more gRNAs.

45. A method of treating a patient in need thereof, comprising administering the system of any one of claims 1-42 or the pharmaceutical composition of claim 44 to the patient, optionally intravenously.

46. The method of claim 45, wherein the patient

has heart disease,

has elevated low-density lipoprotein cholesterol (LDL-C) or hypercholesterolemia,

is at risk of developing myocardial infarction, stroke, or unstable angina, and/or

has primary hyperlipidemia, optionally heterozygous familial hypercholesterolemia (HeFH), or homozygous familial hypercholesterolemia (HoFH).

47. The system of any one of claims 1-42, or the pharmaceutical composition of claim 44, for use in treating a patient in need thereof, optionally in the method of claim 45 or 46.

48. Use of the system of any one of claims 1-42 in the manufacture of a medicament for treating a patient in need thereof, optionally in the method of claim 45 or 46.

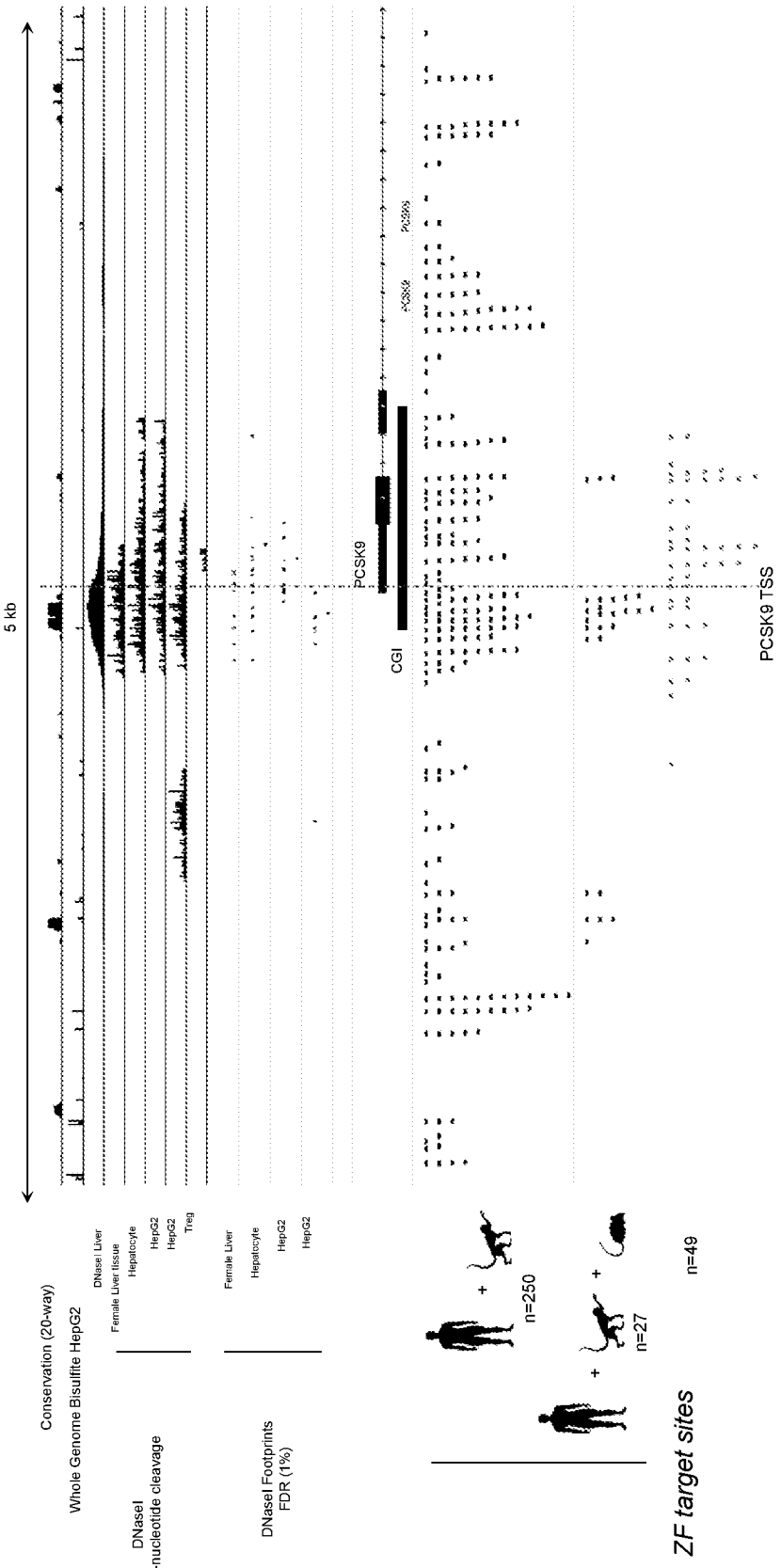

5 kb
Conservation (20-way)
Whole Genome Bisulfite HepG2
DNaseI
Single-nucleotide cleavage
DNaseI Liver
Female Liver tissue
Hepatocyte
HepG2
HepG2
Treg
DNaseI Footprints
FDR (1%)
Female Liver
Hepatocyte
HepG2
HepG2
PCSK9
CGI
PCSK9 TSS
PCSK9 gRNA
n=250
n=27
n=49
ZF target sites

FIG. 1

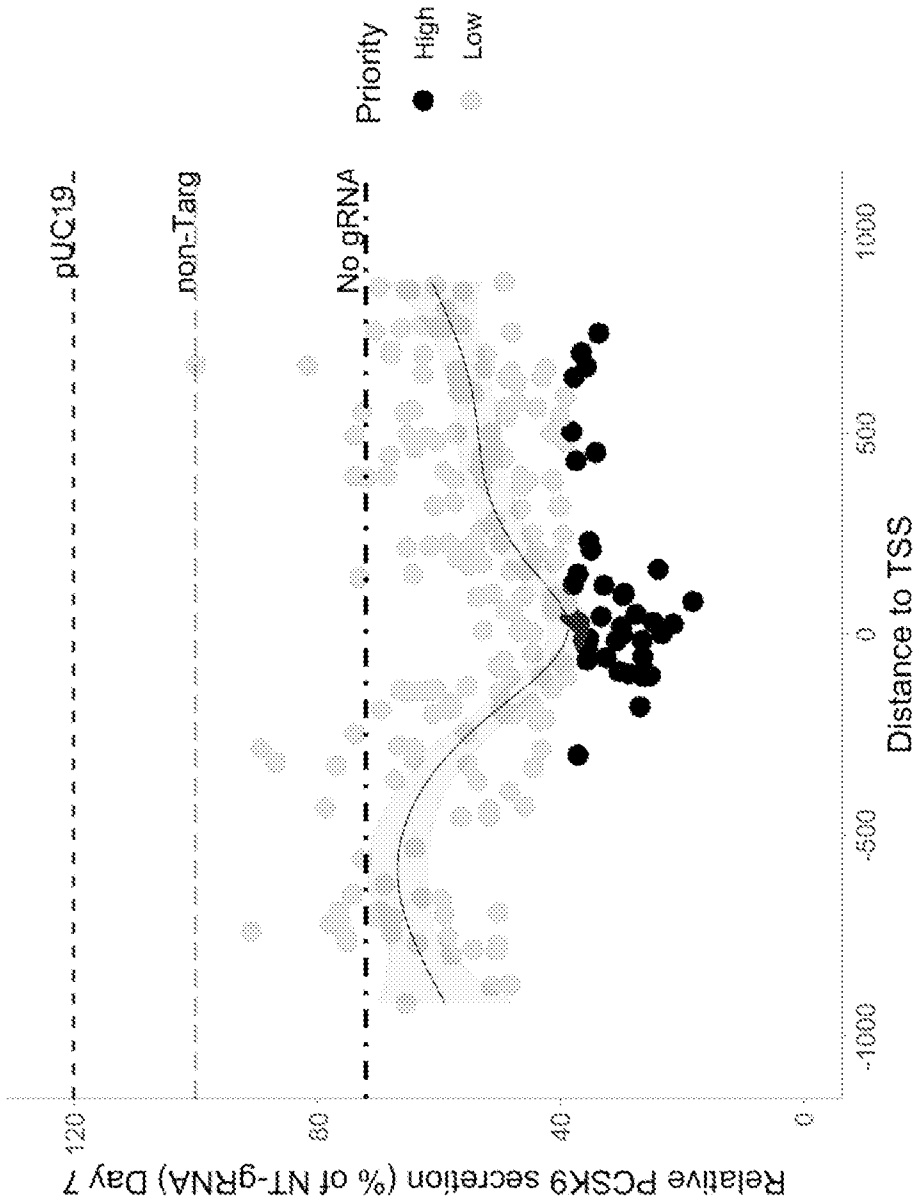

Relative PCSK9 secretion (% of NT-gRNA) Day 7
120
80
40
0
pUC19_
non-Targ
No_gRNA
Priority
High
Low
-1000
-500
0
500
1000
Distance to TSS

FIG. 2

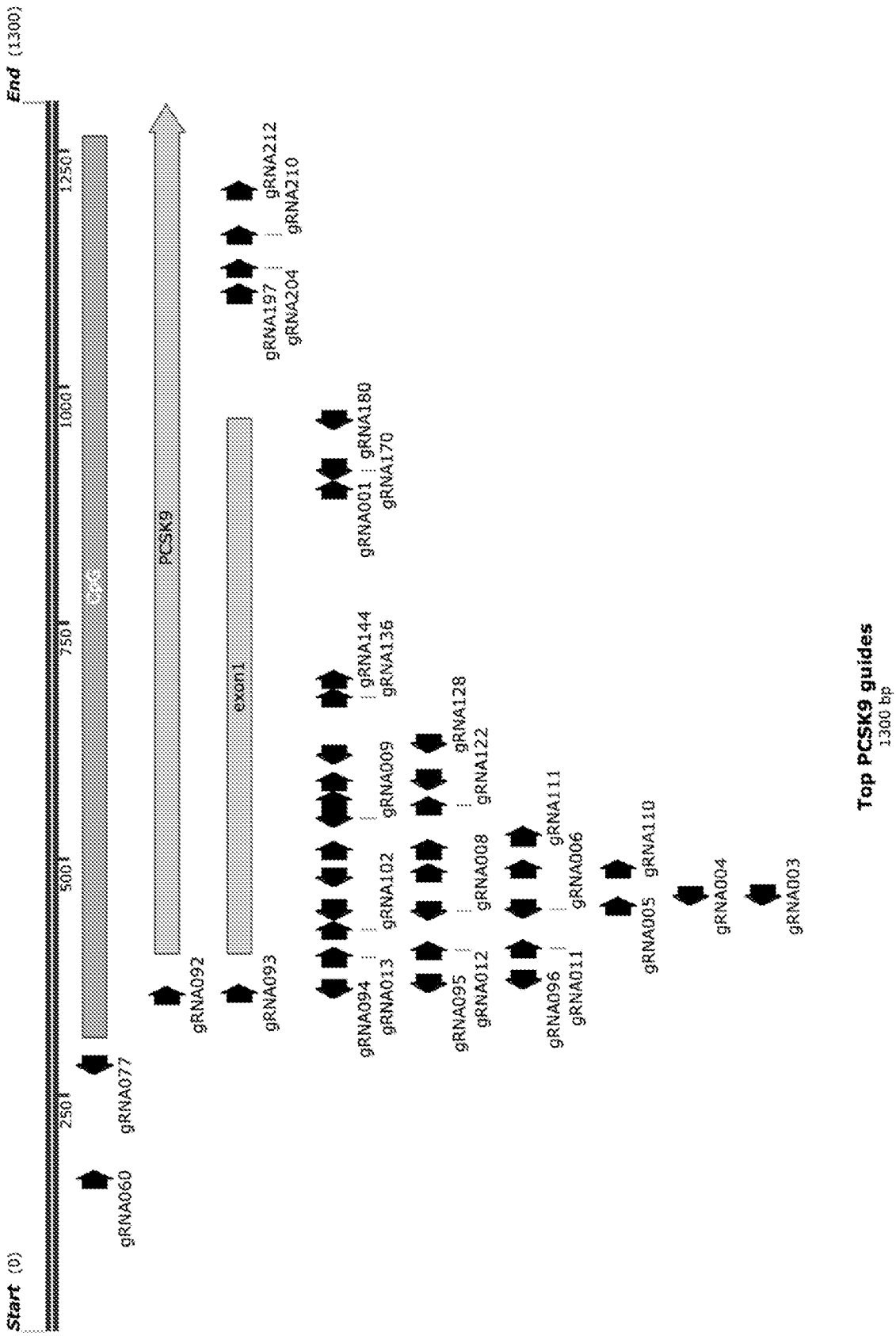

Start (0)
End (1300)
2501
5001
7501
10001
12501
gRNA060
gRNA077
gRNA092
gRNA093
PCSK9
exon1
gRNA094
gRNA013
gRNA102
gRNA009
gRNA144
gRNA136
gRNA095
gRNA012
gRNA008
gRNA122
gRNA128
gRNA096
gRNA011
gRNA006
gRNA111
gRNA005
gRNA110
gRNA004
gRNA003
gRNA001
gRNA180
gRNA170
gRNA197
gRNA204
gRNA212
gRNA210
Top PCSK9 guides
1300 bp

FIG. 3

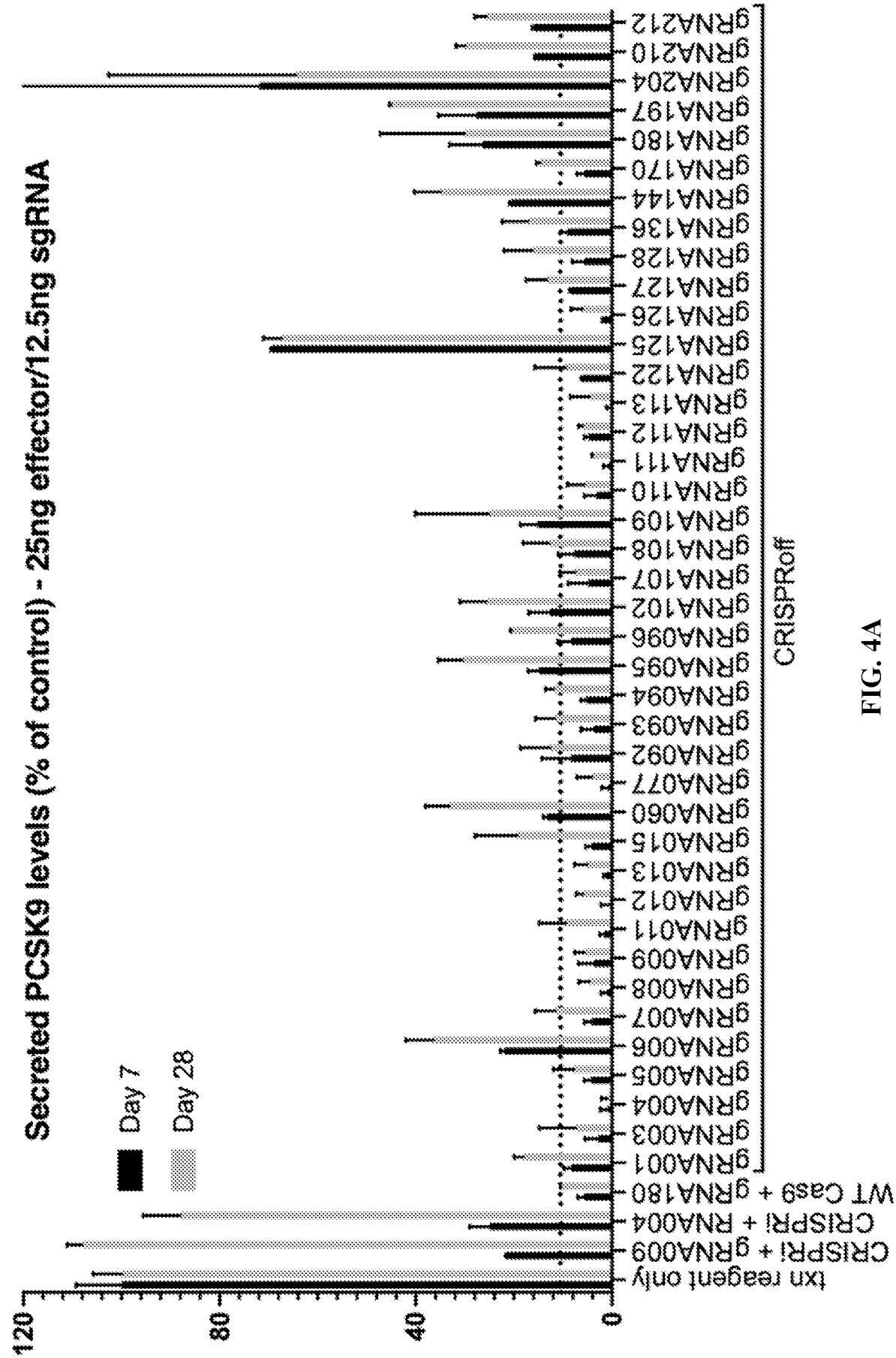

Secreted PCSK9 levels (% of control) - 25ng effector/12.5ng sgRNA
Day 7
Day 28
120
80
40
0
txn reagent only
CRISPRi + gRNA009
CRISPRi + RNA004
WT Cas9 + gRNA180
gRNA001
gRNA003
gRNA004
gRNA005
gRNA006
gRNA007
gRNA008
gRNA009
gRNA011
gRNA012
gRNA013
gRNA015
gRNA060
gRNA077
gRNA092
gRNA093
gRNA094
gRNA095
gRNA096
gRNA102
gRNA107
gRNA108
gRNA109
gRNA110
gRNA111
gRNA112
gRNA113
gRNA122
gRNA125
gRNA126
gRNA127
gRNA128
gRNA136
gRNA144
gRNA170
gRNA180
gRNA197
gRNA204
gRNA210
gRNA212
CRISPRoff

FIG. 4A

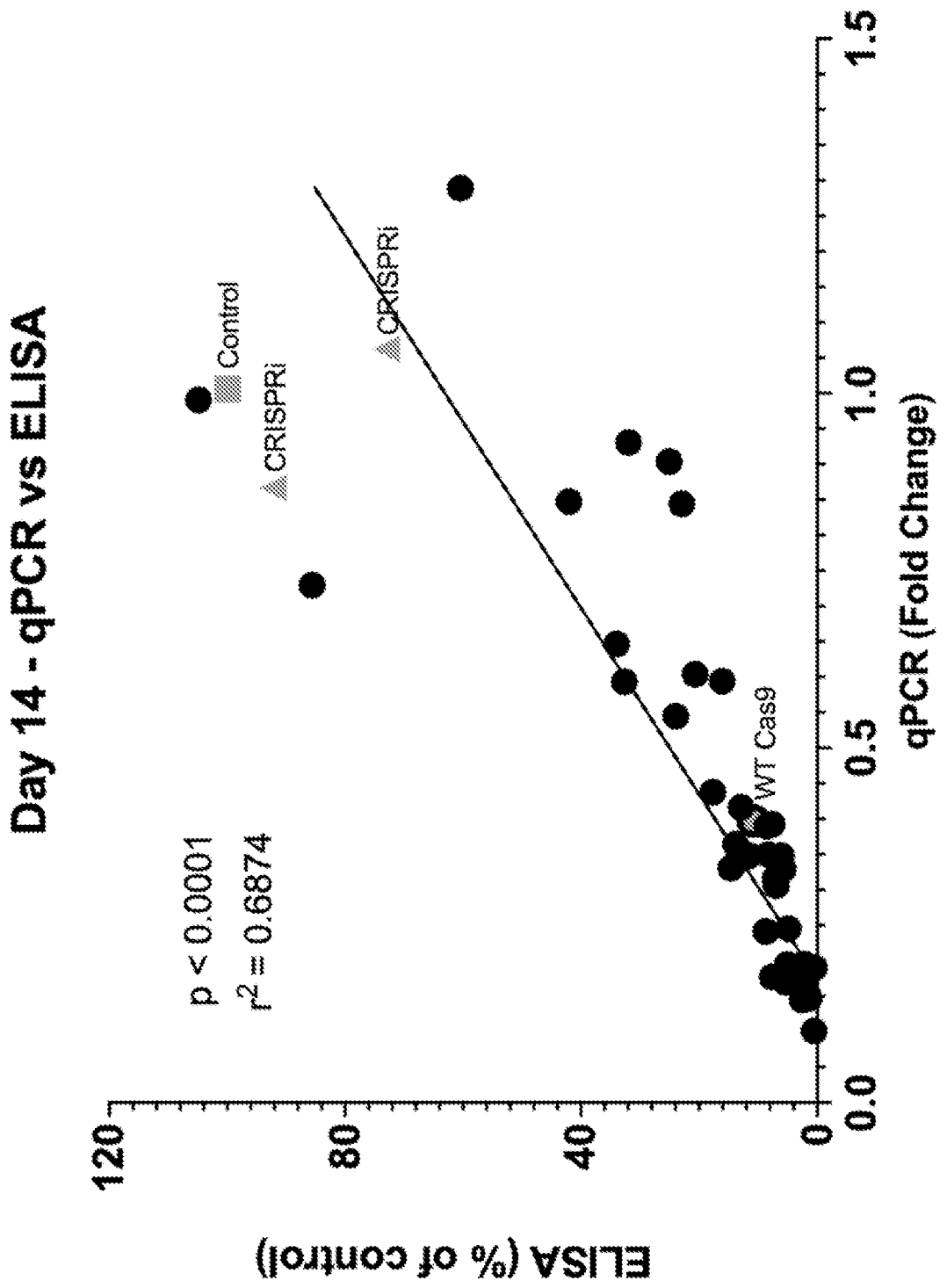

Day 14 - qPCR vs ELISA
p < 0.0001
r2 = 0.6874
Control
CRISPRi
CRISPRi
WT Cas9
ELISA (% of control)
qPCR (Fold Change)
120
80
40
0
0.0
0.5
1.0
1.5

FIG. 4B

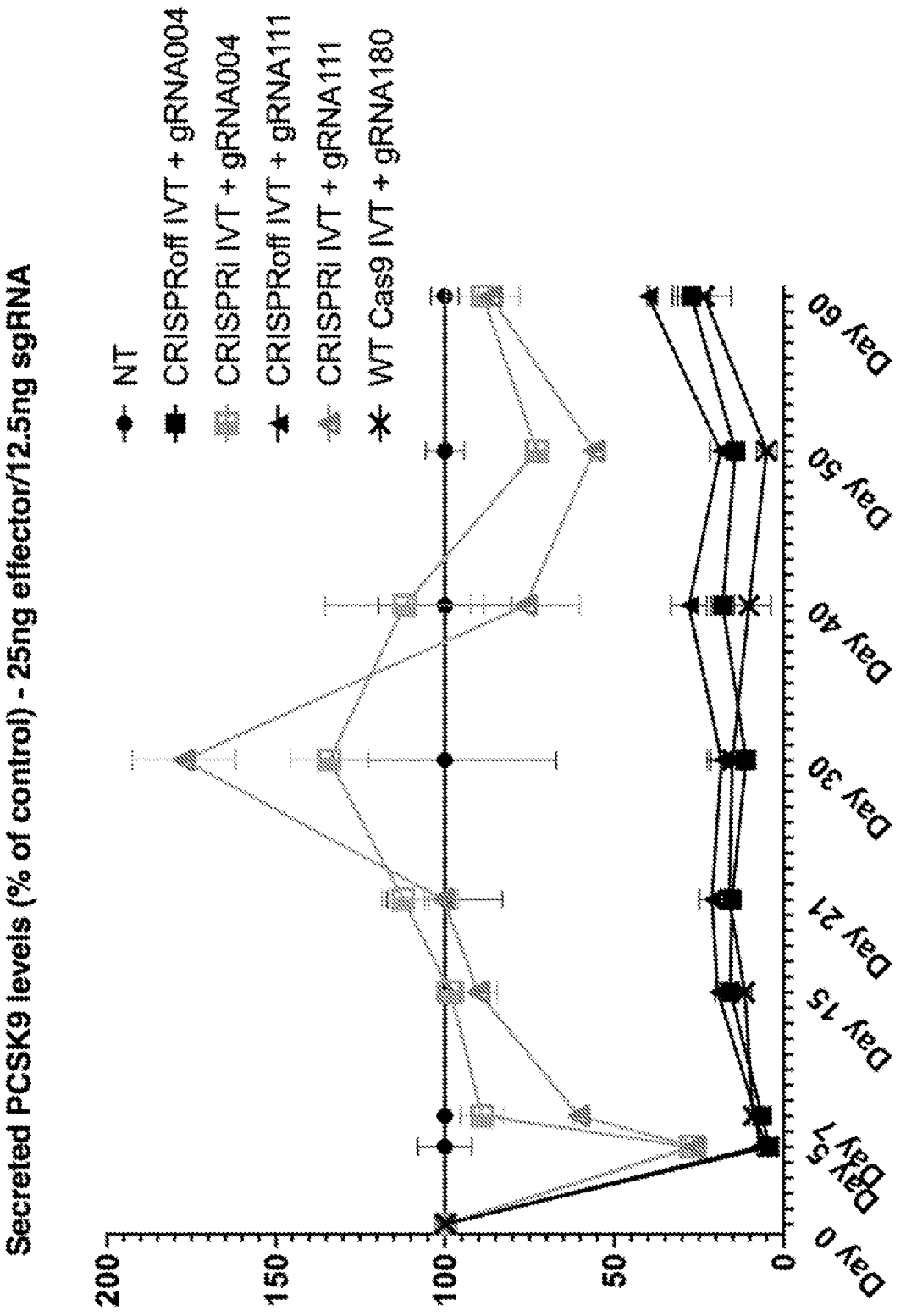

Secreted PCSK9 levels (% of control) - 25ng effector/12.5ng sgRNA
NT
CRISPRoff IVT + gRNA004
CRISPRi IVT + gRNA004
CRISPRoff IVT + gRNA111
CRISPRi IVT + gRNA111
WT Cas9 IVT + gRNA180
200
150
100
50
0
Day 0
Day 5
Day 7
Day 15
Day 21
Day 30
Day 40
Day 50
Day 60

FIG. 5

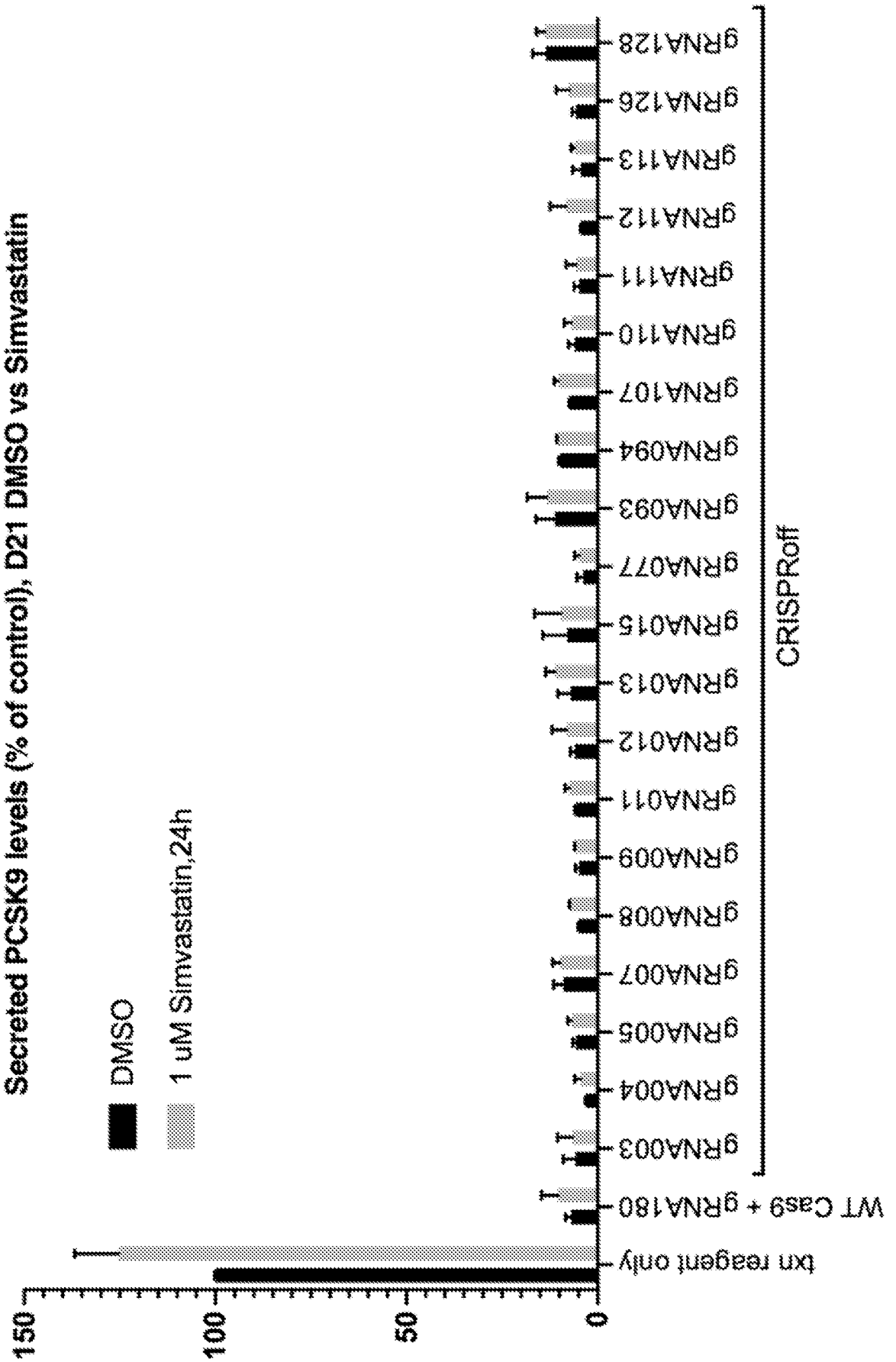

Secreted PCSK9 levels (% of control), D21 DMSO vs Simvastatin
DMSO
1 uM Simvastatin,24h
150
100
50
0
txn reagent only
WT Cas9 + gRNA180
gRNA003
gRNA004
gRNA005
gRNA007
gRNA008
gRNA009
gRNA011
gRNA012
gRNA013
gRNA015
gRNA077
gRNA093
gRNA094
gRNA107
gRNA110
gRNA111
gRNA112
gRNA113
gRNA126
gRNA128
CRISPRoff

FIG. 6

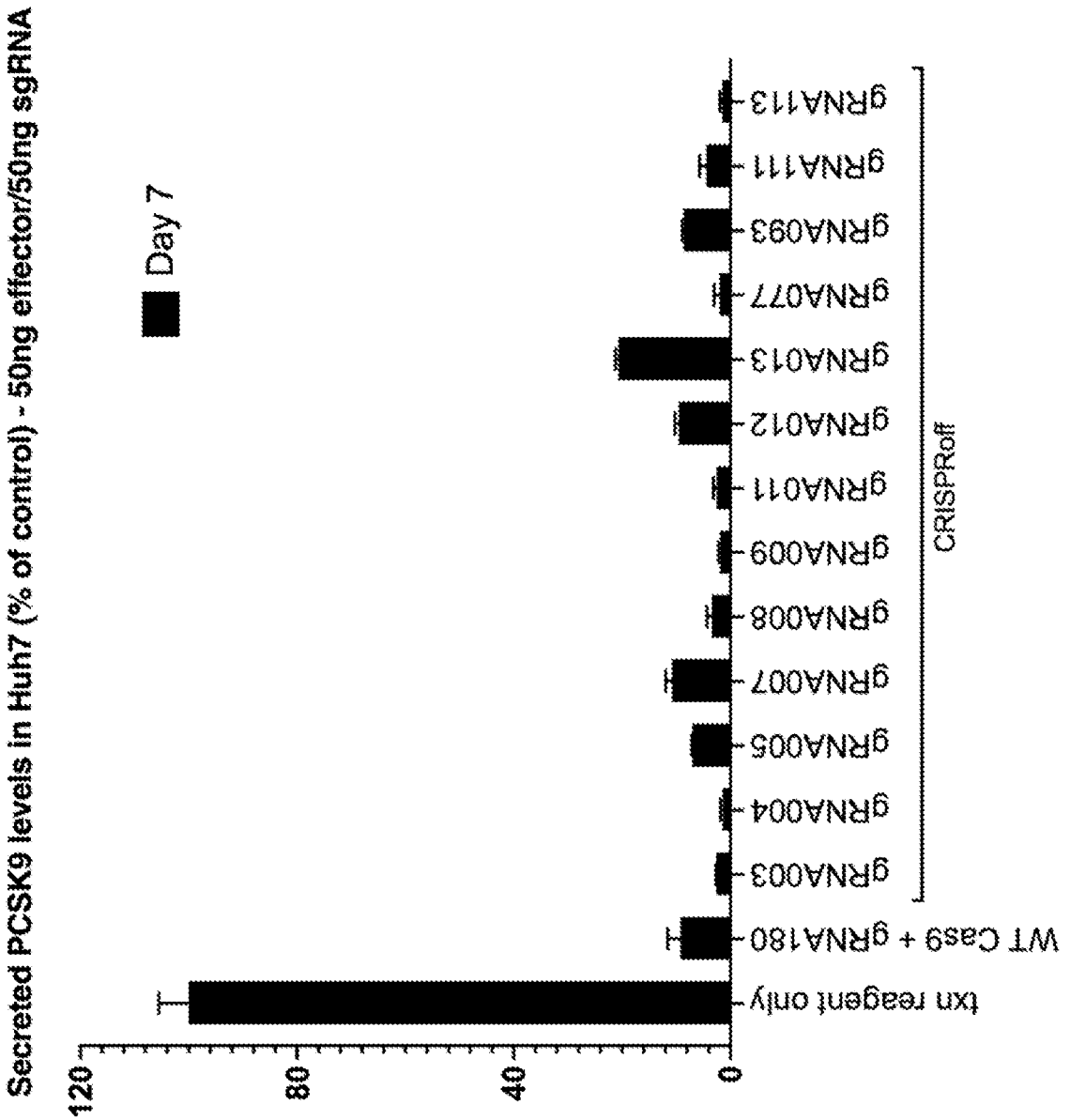

Secreted PCSK9 levels in Huh7 (% of control) - 50ng effector/50ng sgRNA
Day 7
120
80
40
0
txn reagent only
WT Cas9 + gRNA180
gRNA003
gRNA004
gRNA005
gRNA007
gRNA008
gRNA009
gRNA011
gRNA012
gRNA013
gRNA077
gRNA093
gRNA111
gRNA113
CRISPRoff

FIG. 7

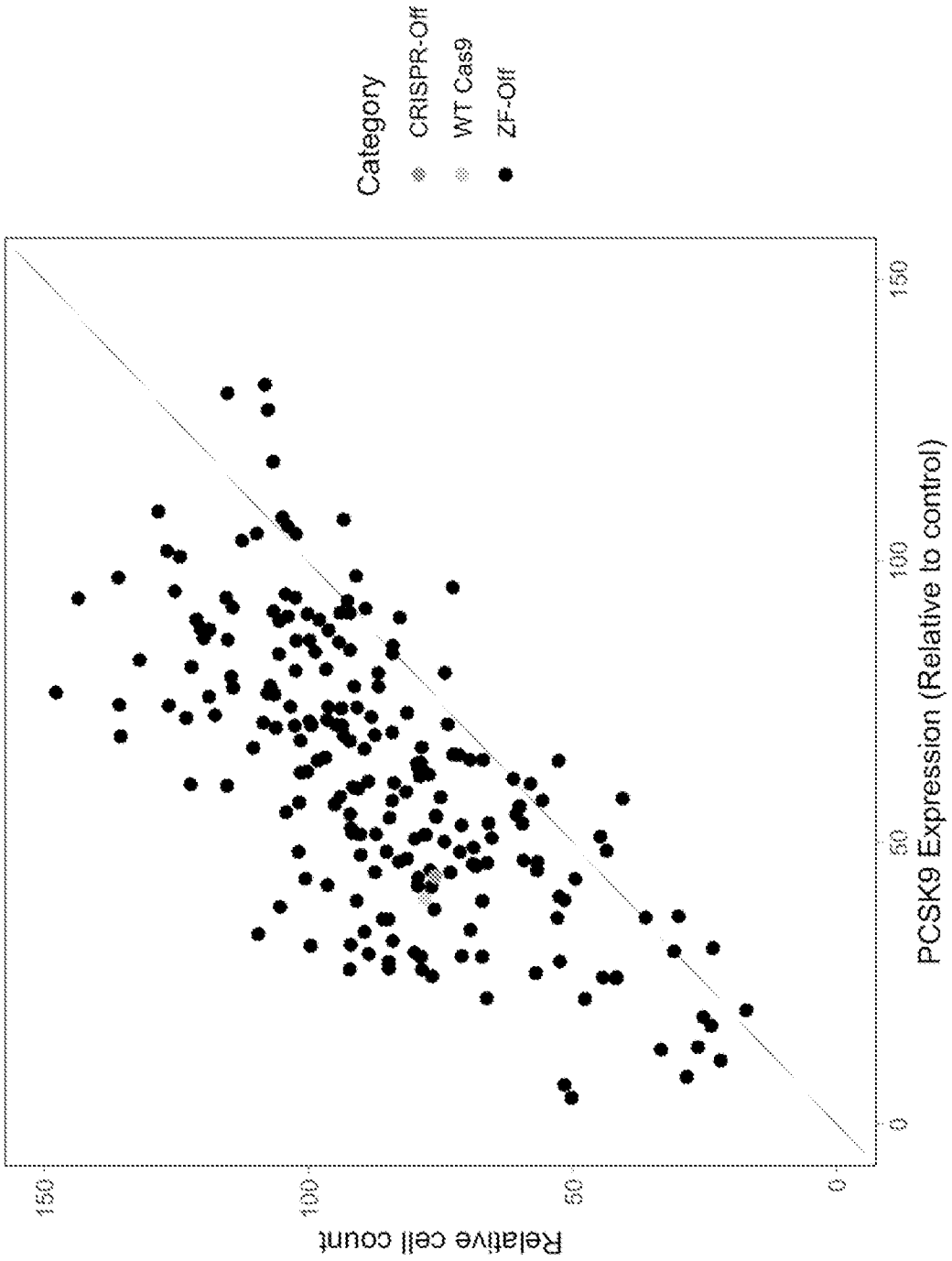

Category
CRISPR-Off
WT Cas9
ZF-Off
Relative cell count
150
100
50
0
0
50
100
150
PCSK9 Expression (Relative to control)

FIG. 8

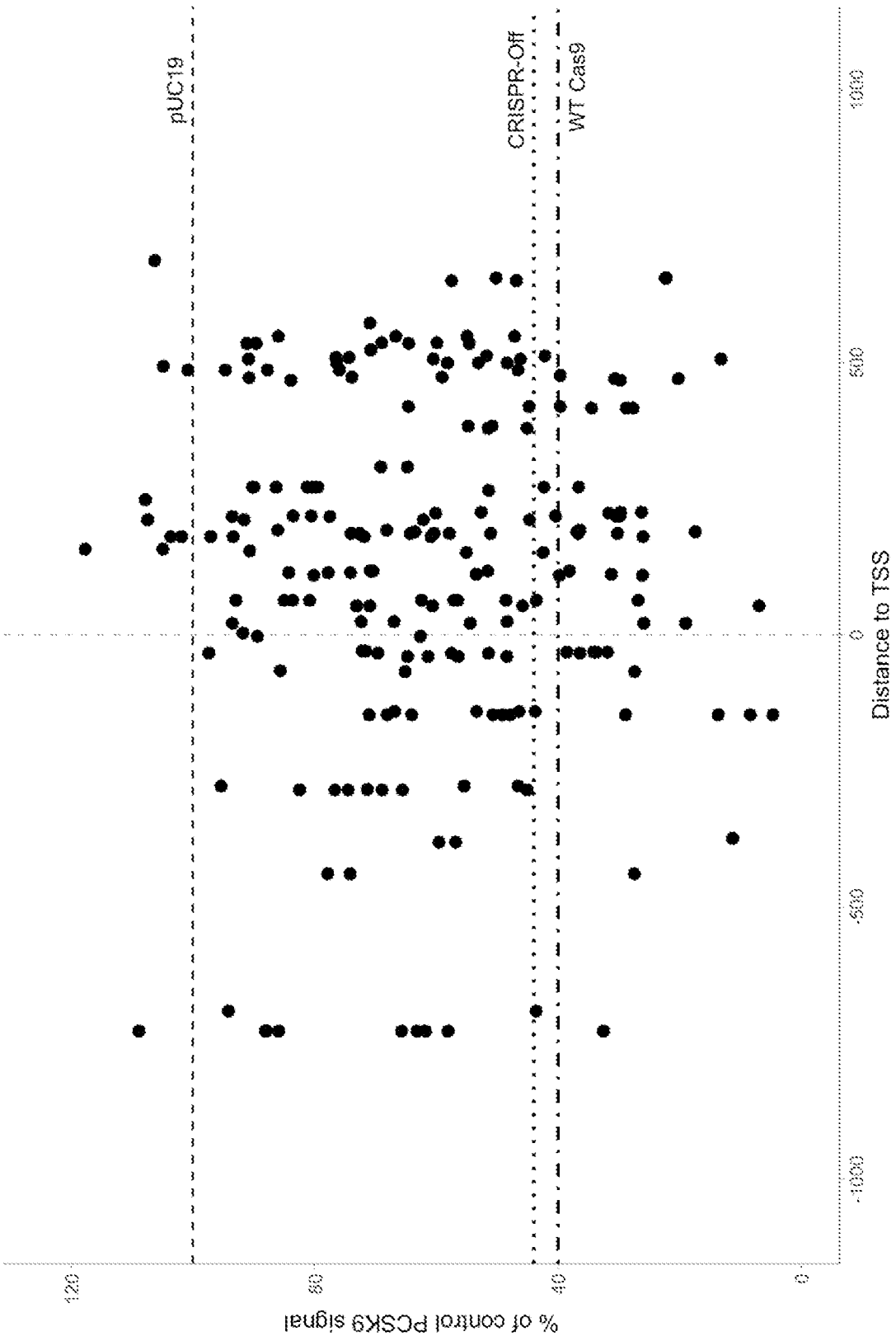

pUC19
CRISPR-Off
WT Cas9
% of control PCSK9 signal
120
80
40
0
-1000
-500
0
500
1000
Distance to TSS

FIG. 9

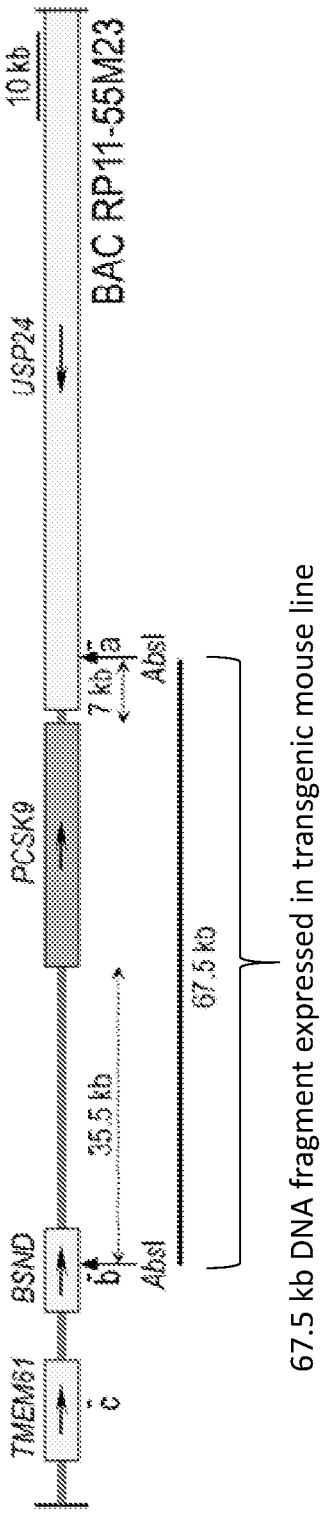

10 kb
BAC RP11-55M23
USP24
PCSK9
BSMD
TMEM61
c
b
a
7 kb
35.5 kb
AbsI
AbsI
67.5 kb
67.5 kb DNA fragment expressed in transgenic mouse line

FIG. 10

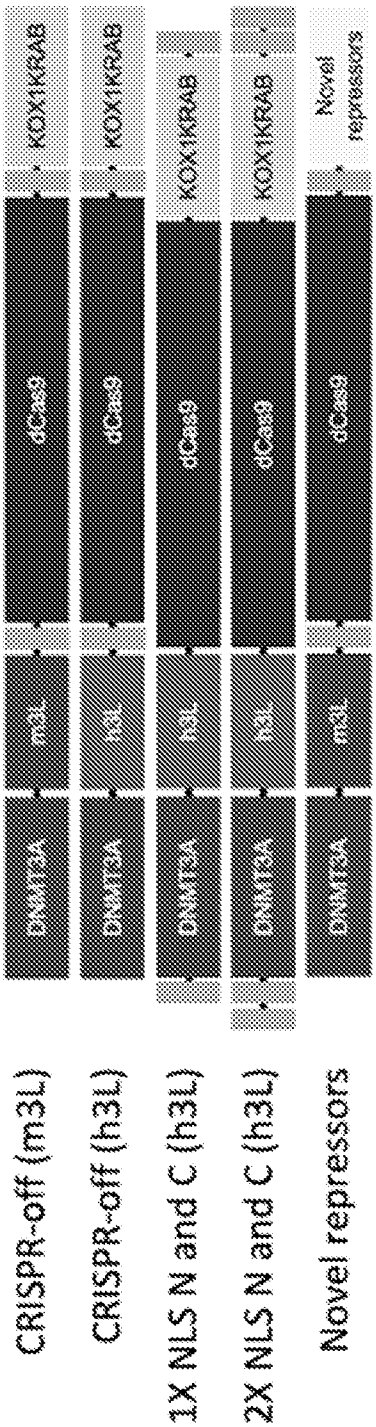


FIG. 11A

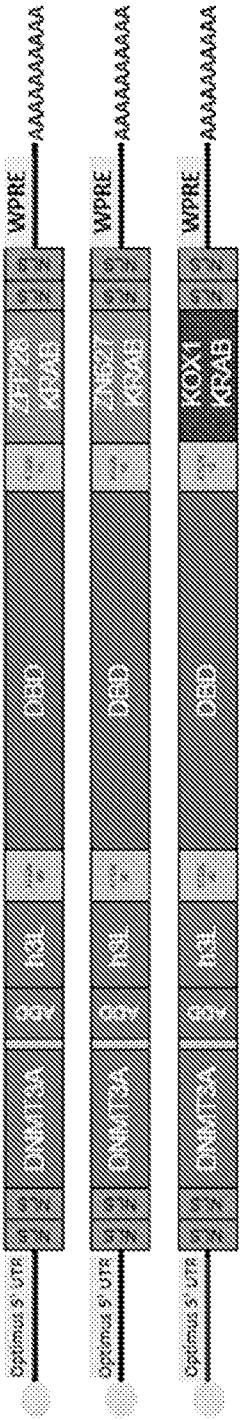


FIG. 11B

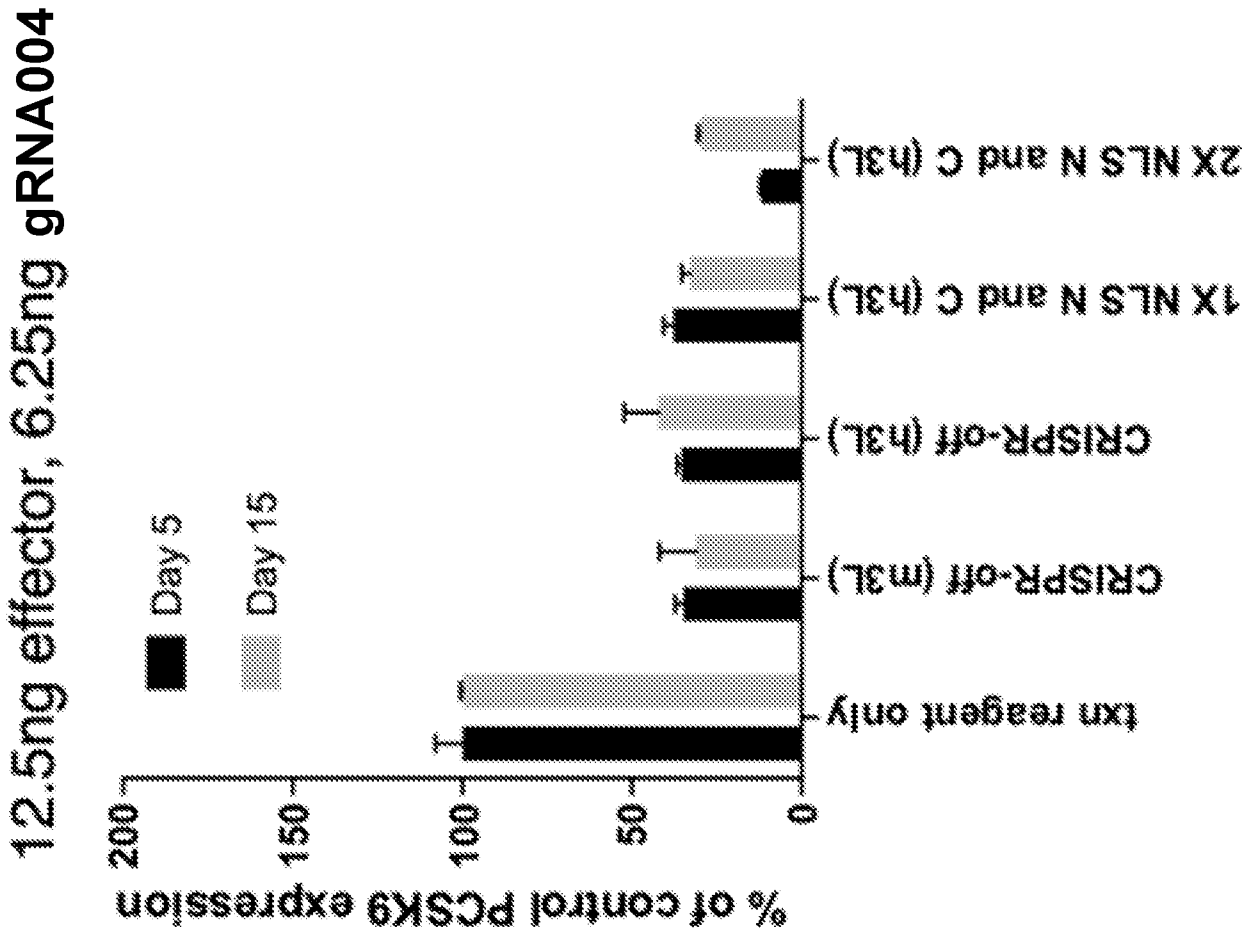

12.5ng effector, 6.25ng gRNA004
% of control PCSK9 expression
200
150
100
50
0
Day 5
Day 15
txn reagent only
CRISPR-off (m3L)
CRISPR-off (h3L)
1X NLS N and C (h3L)
2X NLS N and C (h3L)

FIG. 12A

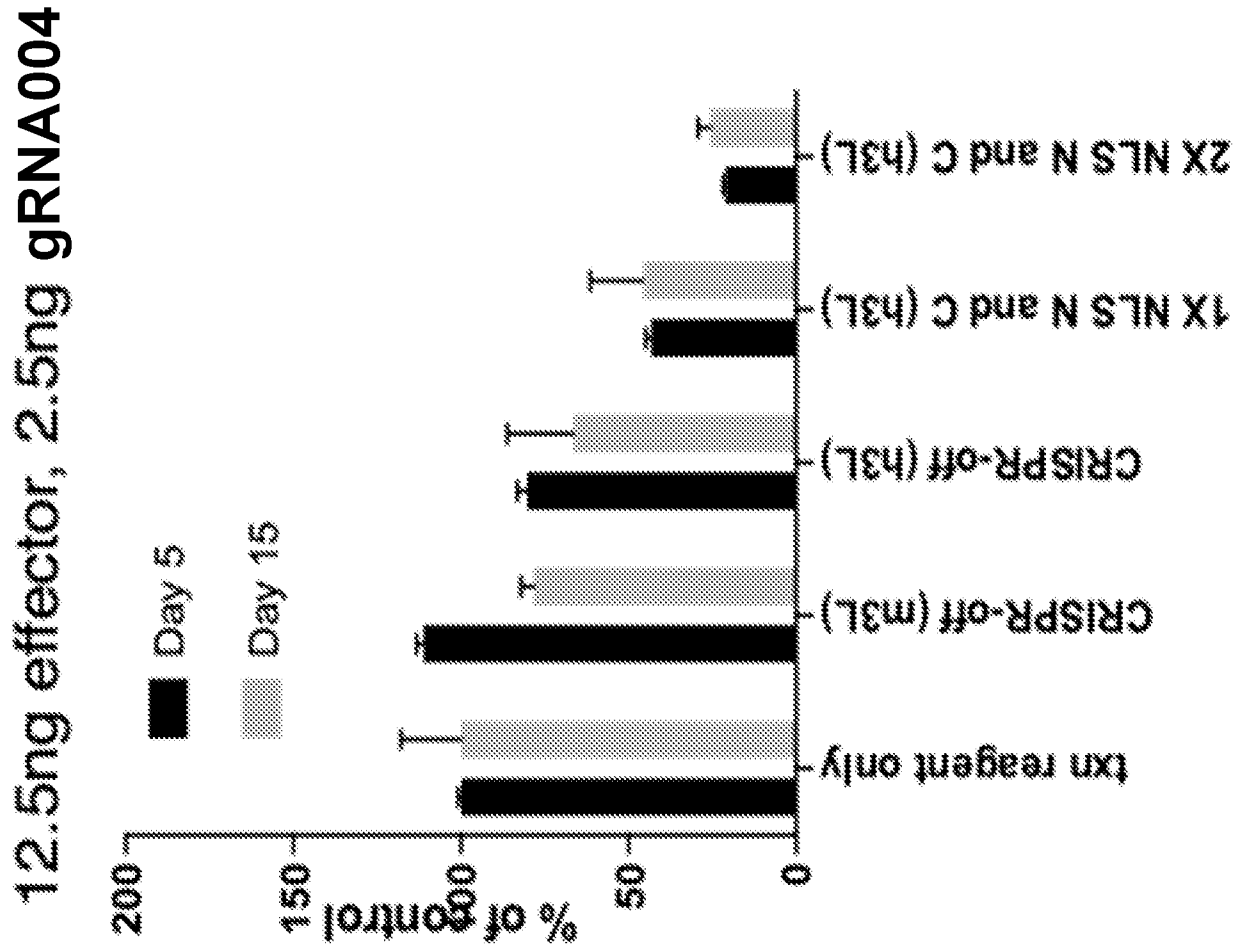

12.5ng effector, 2.5ng gRNA004
Day 5
Day 15
% of control
200
150
100
50
0
txn reagent only
CRISPR-off (m3L)
CRISPR-off (h3L)
1X NLS N and C (h3L)
2X NLS N and C (h3L)

FIG. 12B

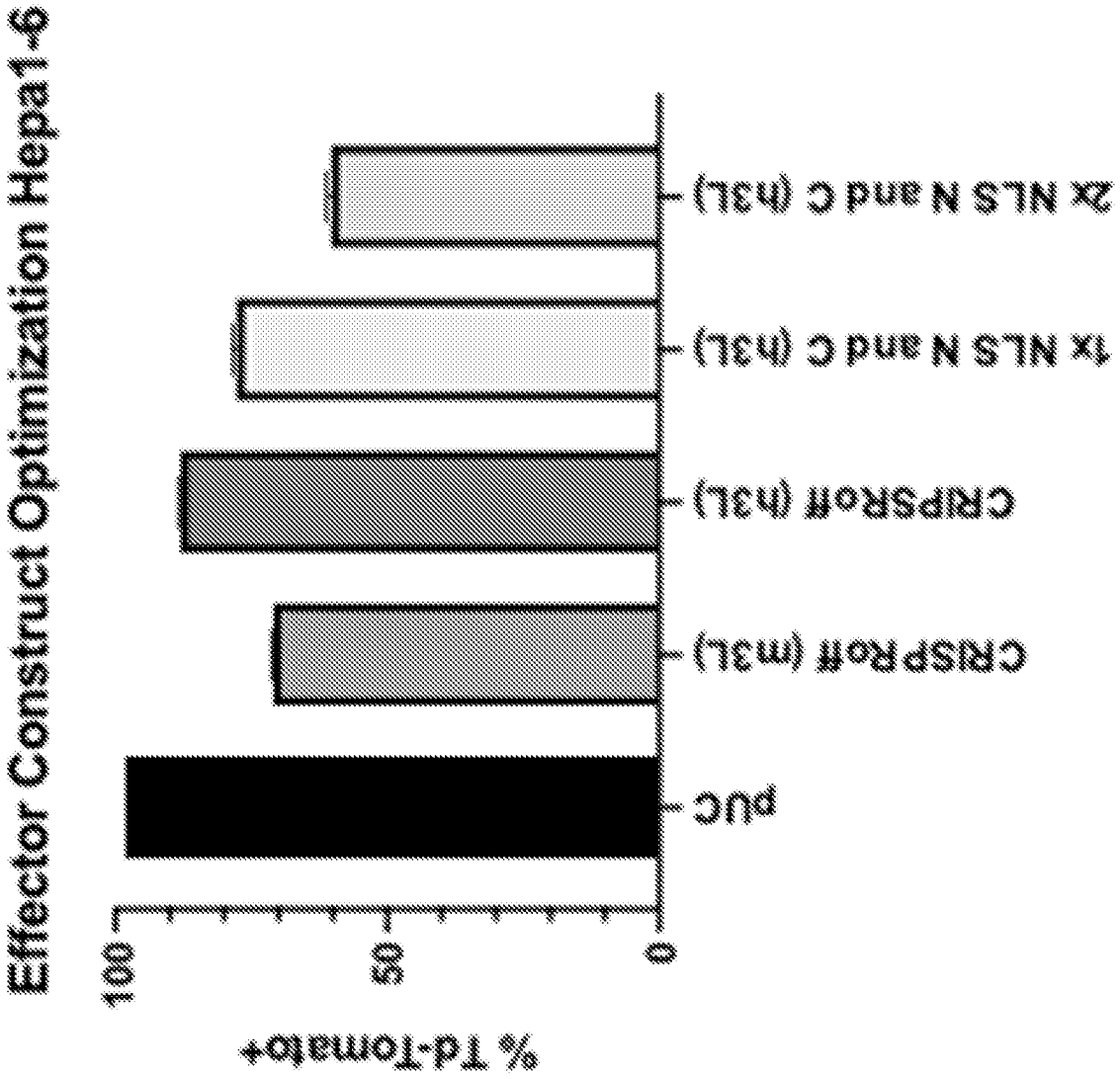

Effector Construct Optimization Hepa1-6
% Td-Tomato+
100
50
0
pUC
CRISPRoff (m3L)
CRIPSRoff (h3L)
1x NLS N and C (h3L)
2x NLS N and C (h3L)

FIG. 13

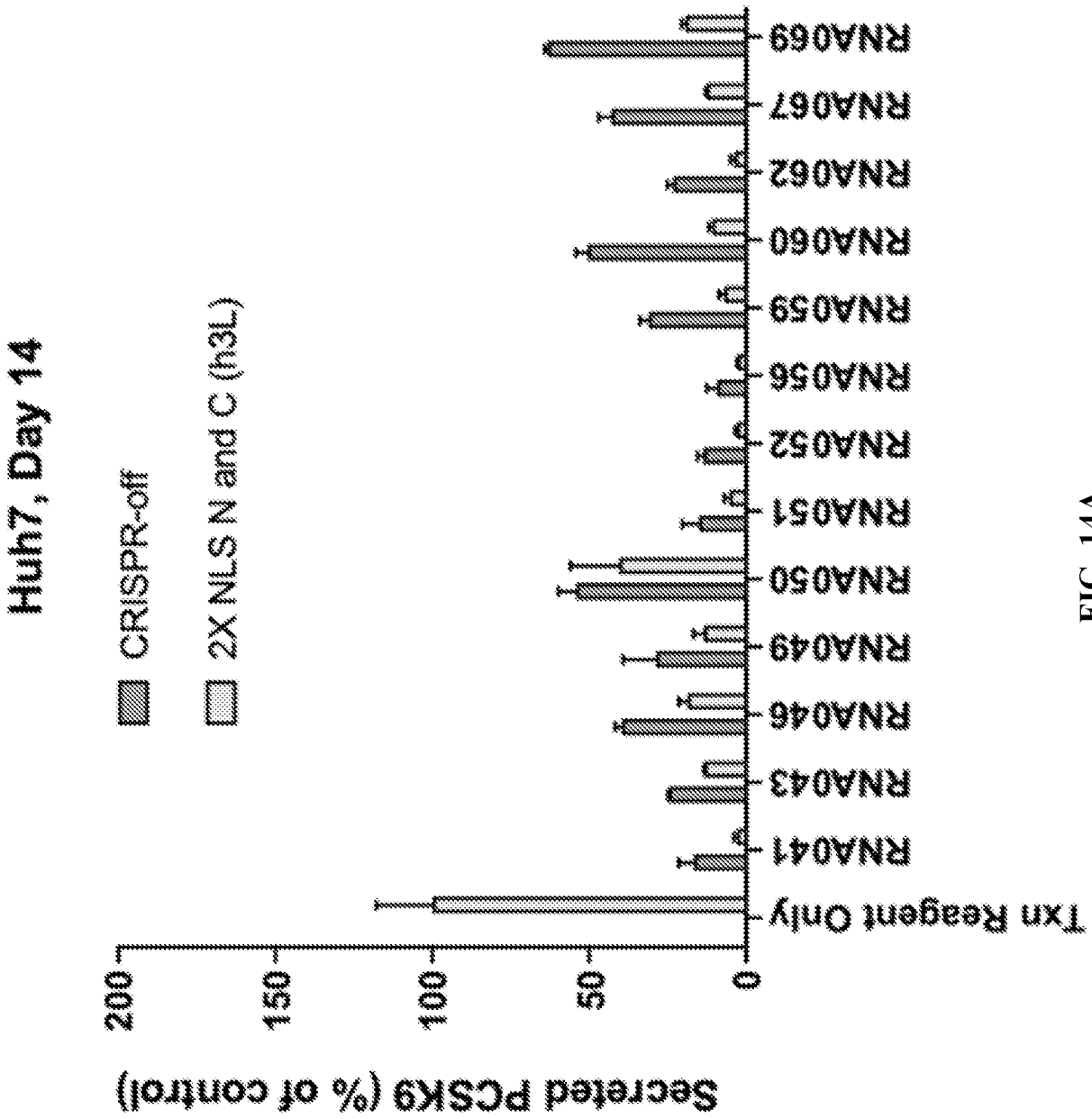

Huh7, Day 14
CRISPR-off
2X NLS N and C (h3L)
Secreted PCSK9 (% of control)
200
150
100
50
0
Txn Reagent Only
RNA041
RNA043
RNA046
RNA049
RNA050
RNA051
RNA052
RNA056
RNA059
RNA060
RNA062
RNA067
RNA069

FIG. 14A

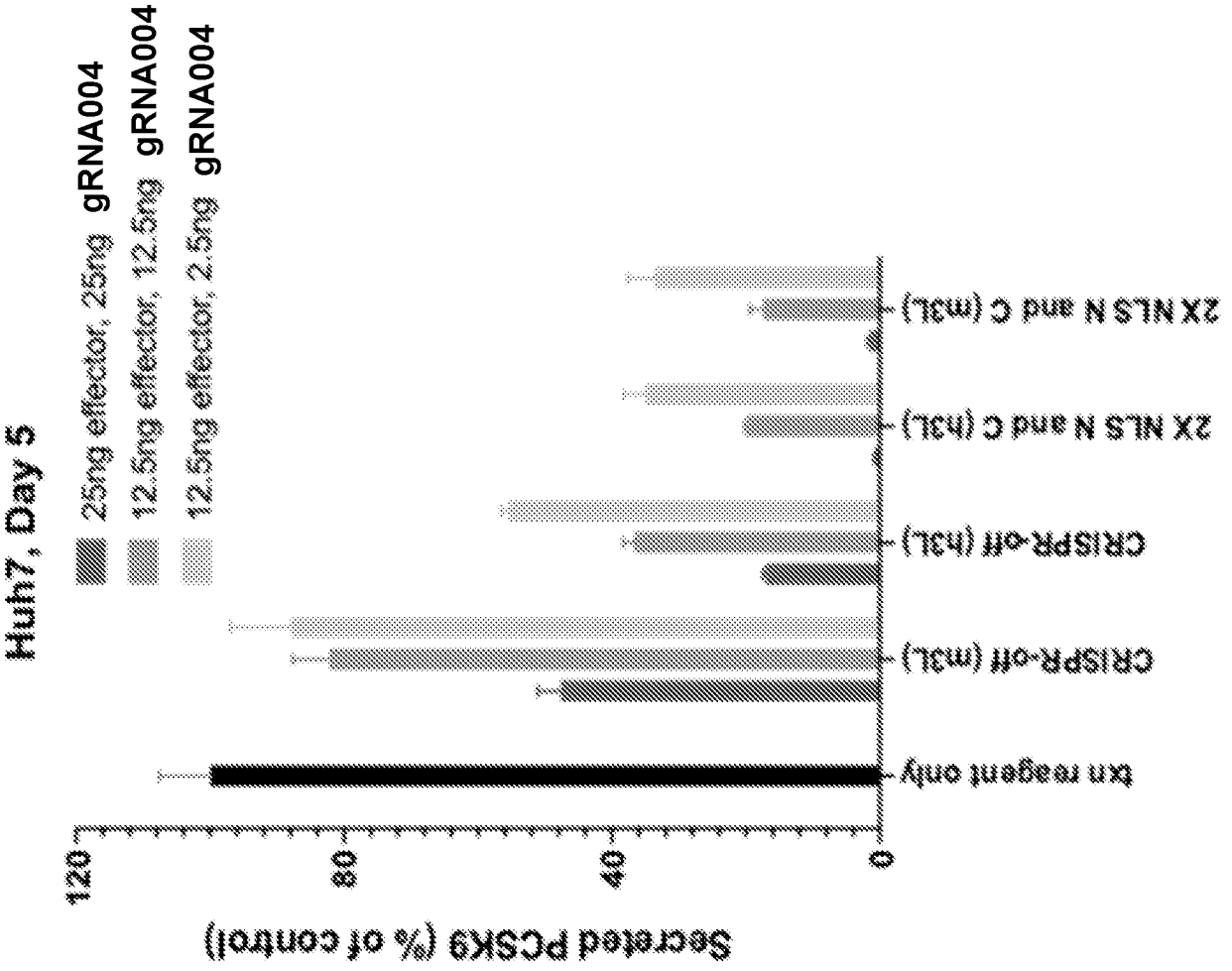

Huh7, Day 5
25ng effector, 25ng gRNA004
12.5ng effector, 12.5ng gRNA004
12.5ng effector, 2.5ng gRNA004
Secreted PCSK9 (% of control)
120
80
40
0
txn reagent only
CRISPR-off (m3L)
CRISPR-off (h3L)
2X NLS N and C (h3L)
2X NLS N and C (m3L)

FIG. 14B

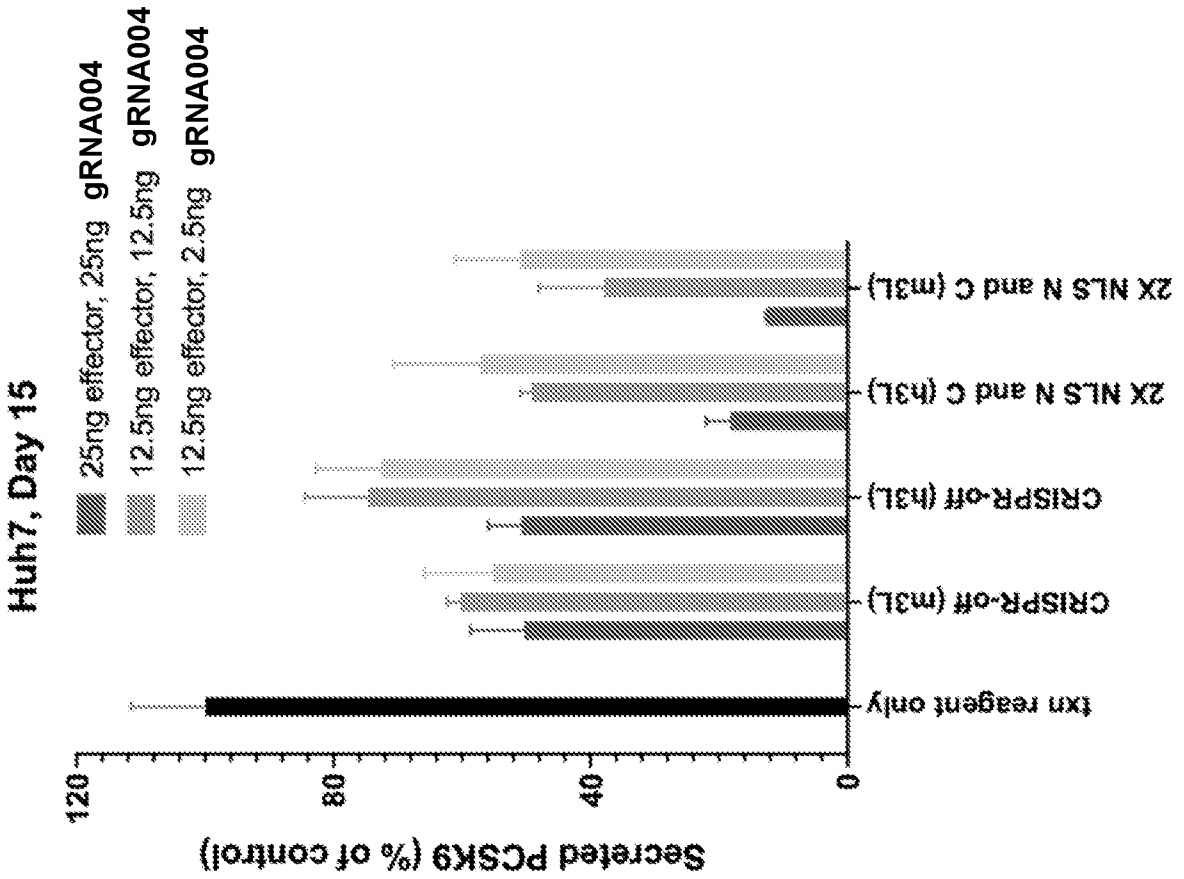

Huh7, Day 15
25ng effector, 25ng gRNA004
12.5ng effector, 12.5ng gRNA004
12.5ng effector, 2.5ng gRNA004
Secreted PCSK9 (% of control)
120
80
40
0
txn reagent only
CRISPR-off (m3L)
CRISPR-off (h3L)
2X NLS N and C (h3L)
2X NLS N and C (m3L)

FIG. 14C

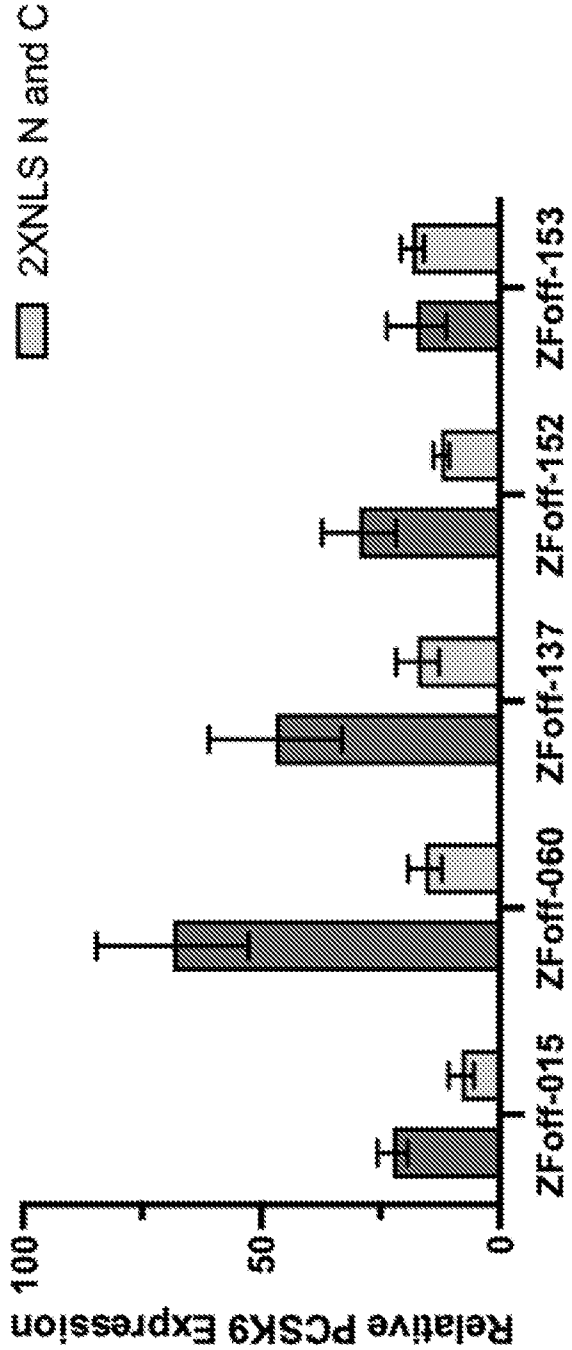

Huh7 Zinc Finger-off, Day 5
ZF-off
2XNLS N and C
Relative PCSK9 Expression
100
50
0
ZFoff-015
ZFoff-060
ZFoff-137
ZFoff-152
ZFoff-153

FIG. 15

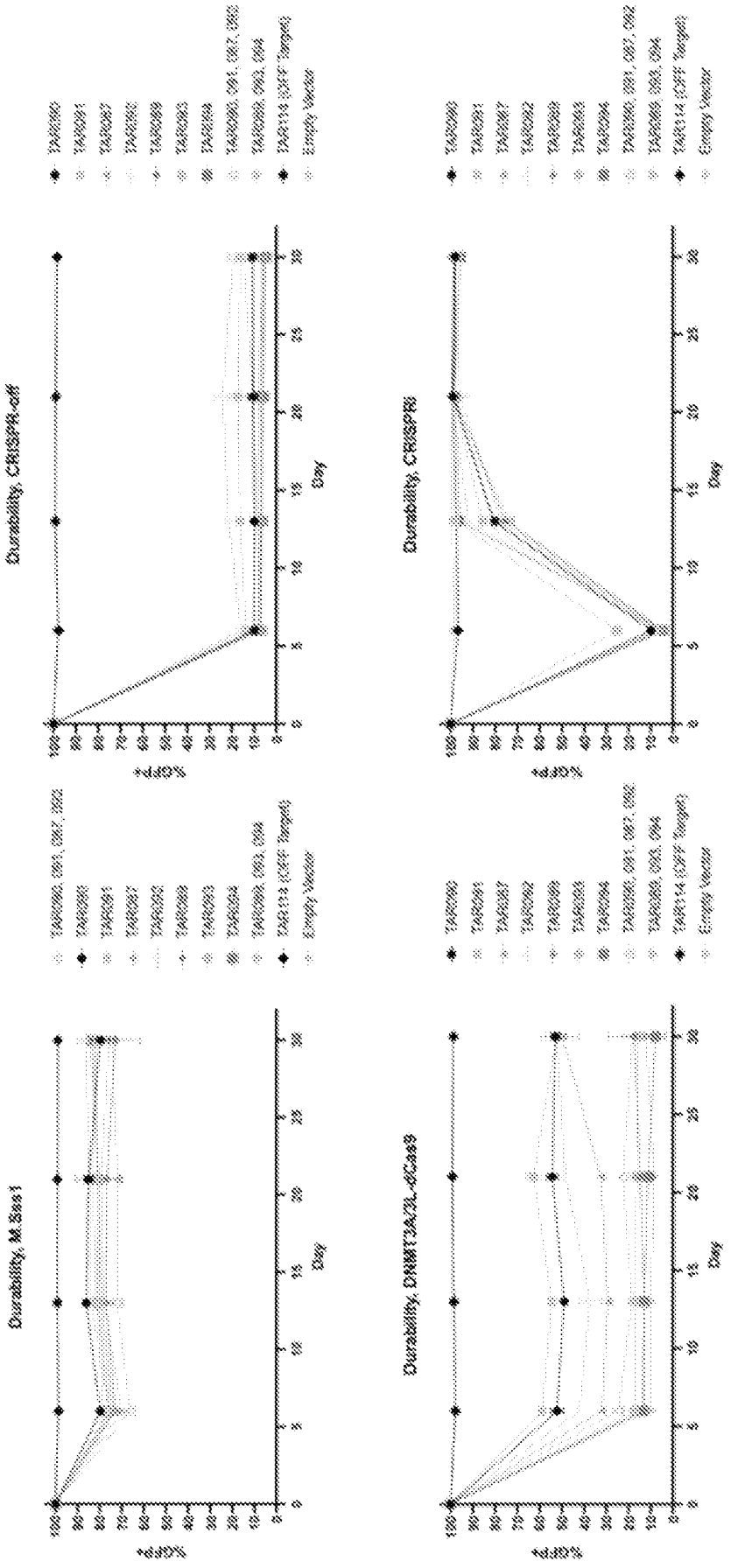

Durability, M.Sss1
Durability, CRISPR-off
Durability, DNMT3A/3L-dCas9
Durability, CRISPR
%GFP+
Day
TAR360
TAR361
TAR367
TAR362
TAR368
TAR363
TAR369
TAR360, 361, 367, 362
TAR368, 363, 369
TAR114 (OFF Target)
Empty Vector

FIG. 16


Name = empty vector
Name = M.MspI
Name = M.SssI
Name = CRISPRoff
Name = hD3A (CD) - mD3L (CD)
% Methylation Day 13
0
25
50
75
100
17229300
17229400
17229500
17229600
17229700
17229800
17229900
17230000
17230100

FIG. 17

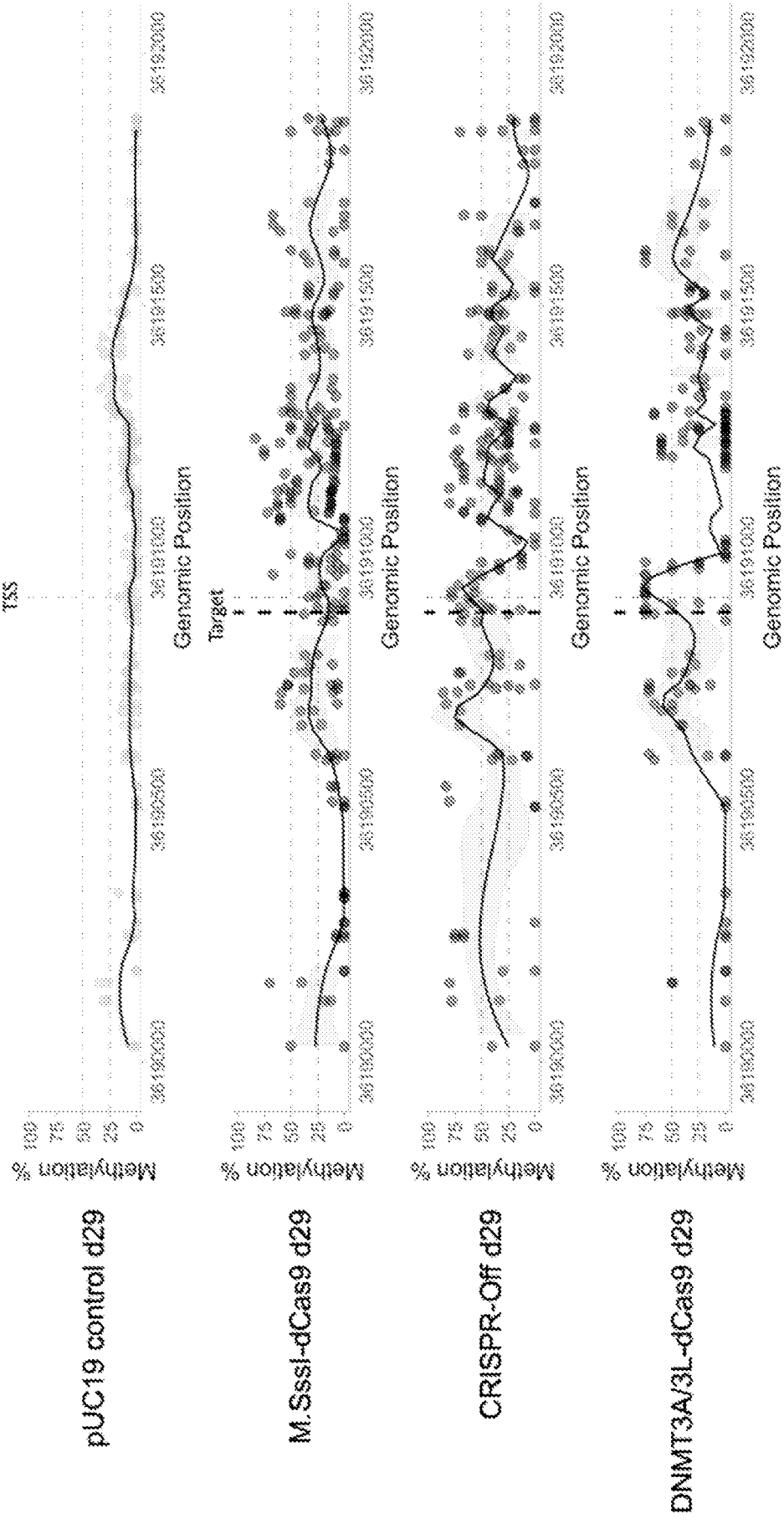

pUC19 control d29
M.SssI-dCas9 d29
CRISPR-Off d29
DNMT3A/3L-dCas9 d29
Methylation %
100
75
50
25
0
TSS
Target
Genomic Position
36190000
36190500
36191000
36191500
36192000

FIG. 18

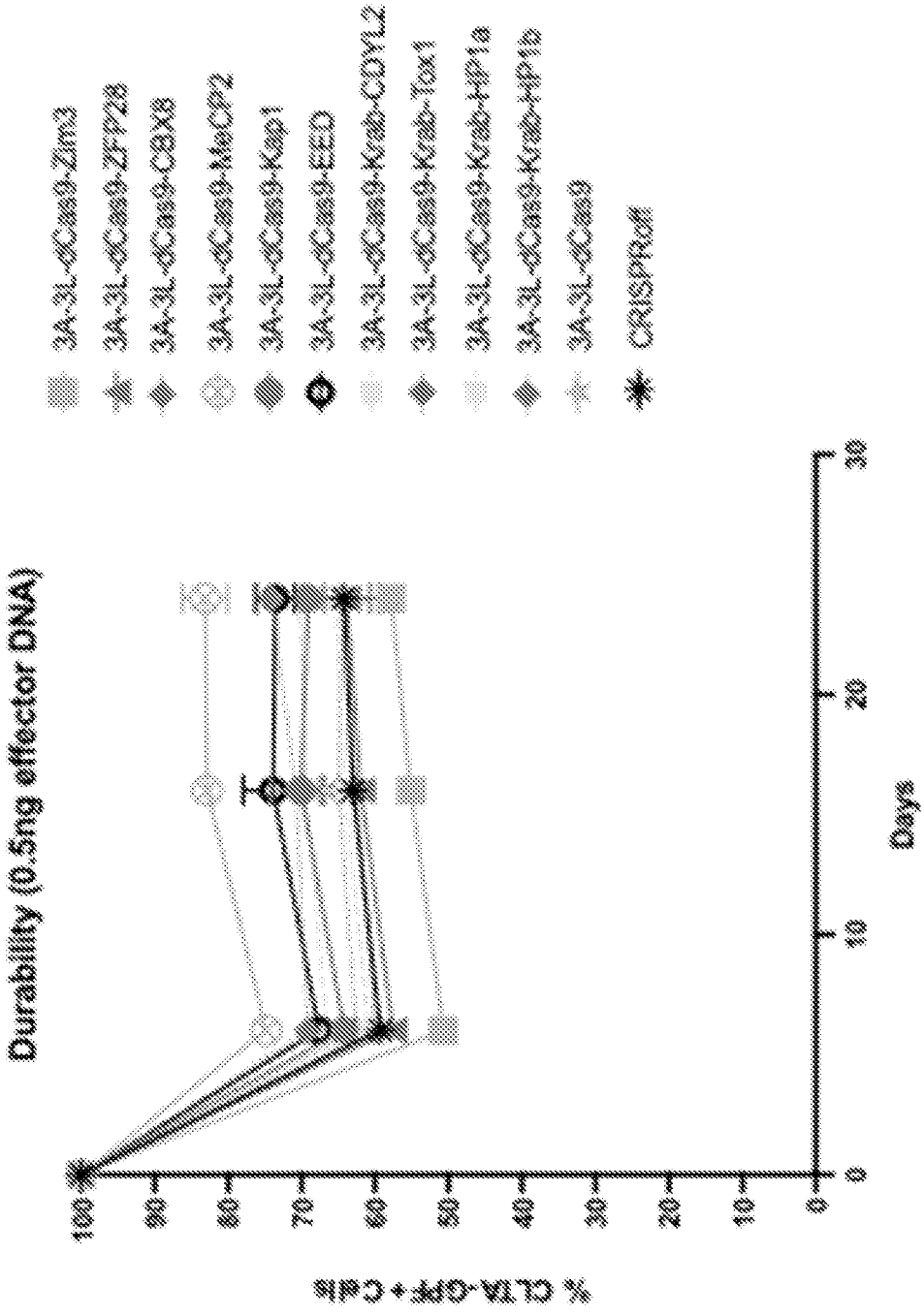

Durability (0.5ng effector DNA)
% CLTA-OFF + Cells
Days
0
10
20
30
40
50
60
70
80
90
100
3A-3L-dCas9-Zim3
3A-3L-dCas9-ZFP28
3A-3L-dCas9-CBX8
3A-3L-dCas9-MeCP2
3A-3L-dCas9-Kap1
3A-3L-dCas9-EED
3A-3L-dCas9-Krab-CDYL2
3A-3L-dCas9-Krab-Tox1
3A-3L-dCas9-Krab-HP1a
3A-3L-dCas9-Krab-HP1b
3A-3L-dCas9
CRISPRoff

FIG. 19A

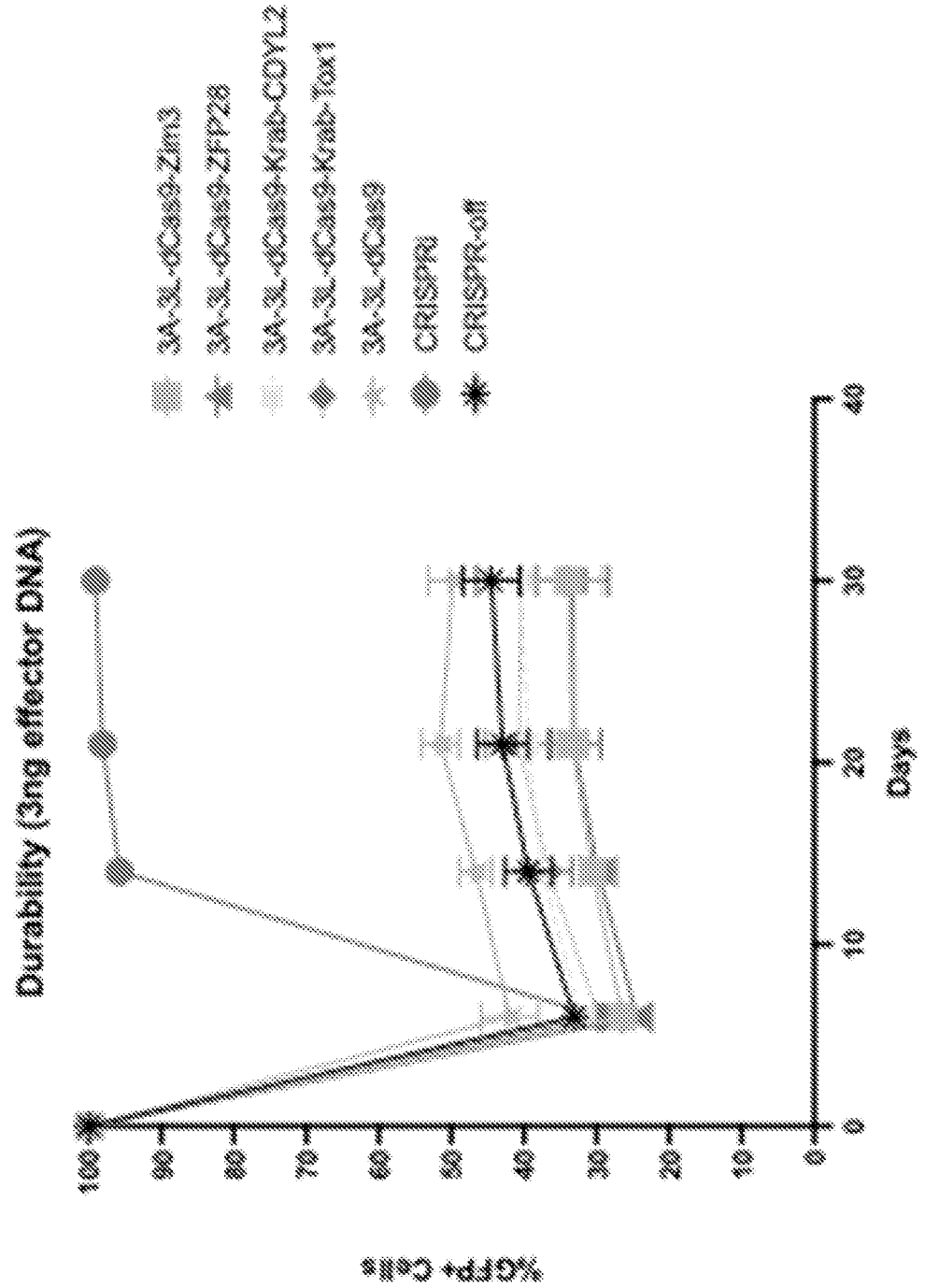

Durability (3ng effector DNA)
3A-3L-dCas9-Zim3
3A-3L-dCas9-ZFP28
3A-3L-dCas9-Krab-COYL2
3A-3L-dCas9-Krab-Tox1
3A-3L-dCas9
CRISPRi
CRISPR-off
%GFP+ Cells
Days
100
90
80
70
60
50
40
30
20
10
0
0
10
20
30
40

FIG. 19B

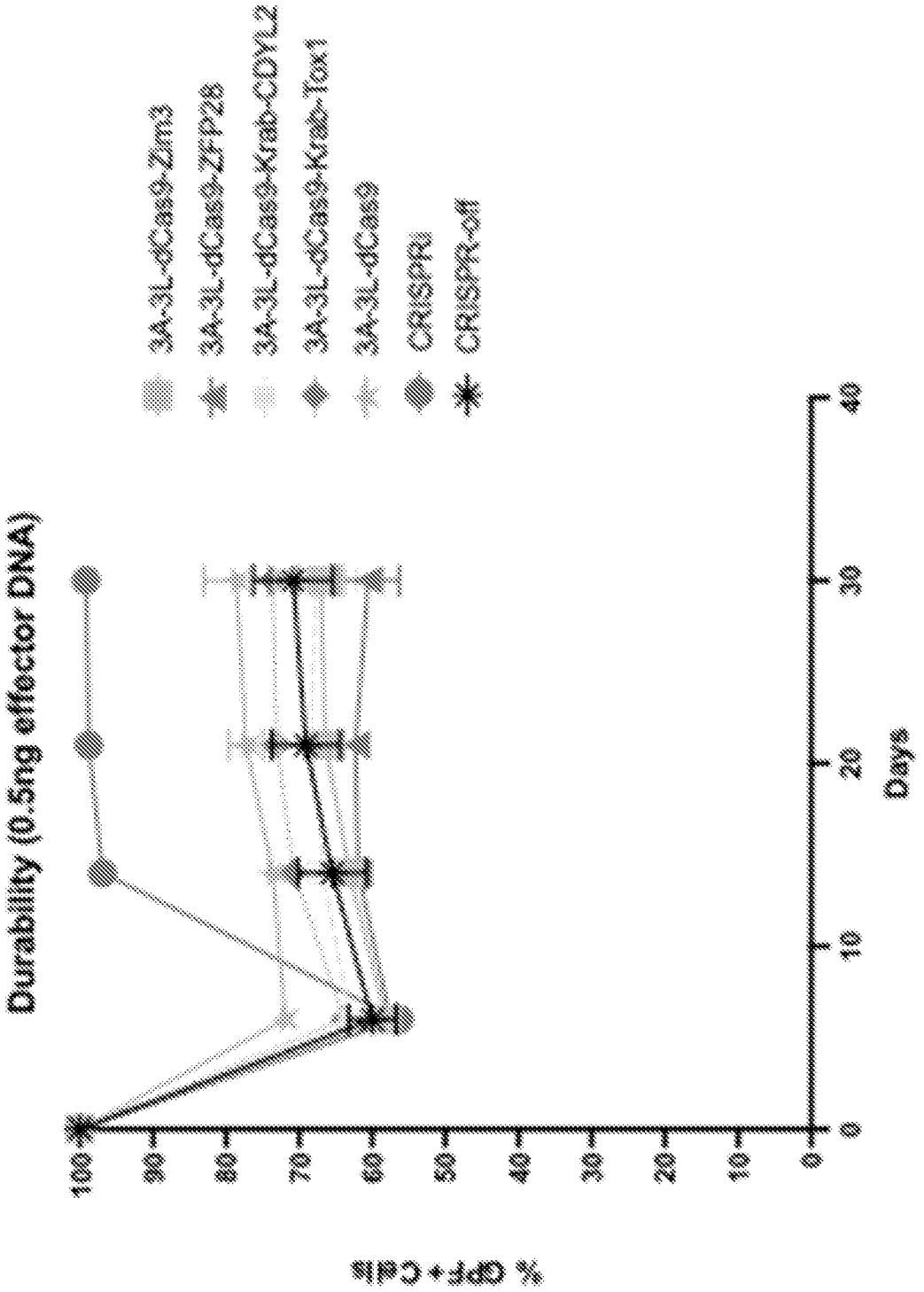

Durability (0.5ng effector DNA)
3A-3L-dCas9-Zim3
3A-3L-dCas9-ZFP28
3A-3L-dCas9-Krab-CDYL2
3A-3L-dCas9-Krab-Tox1
3A-3L-dCas9
CRISPR
CRISPR-off
% GFP+ Cells
Days
0
10
20
30
40
50
60
70
80
90
100

FIG. 19C

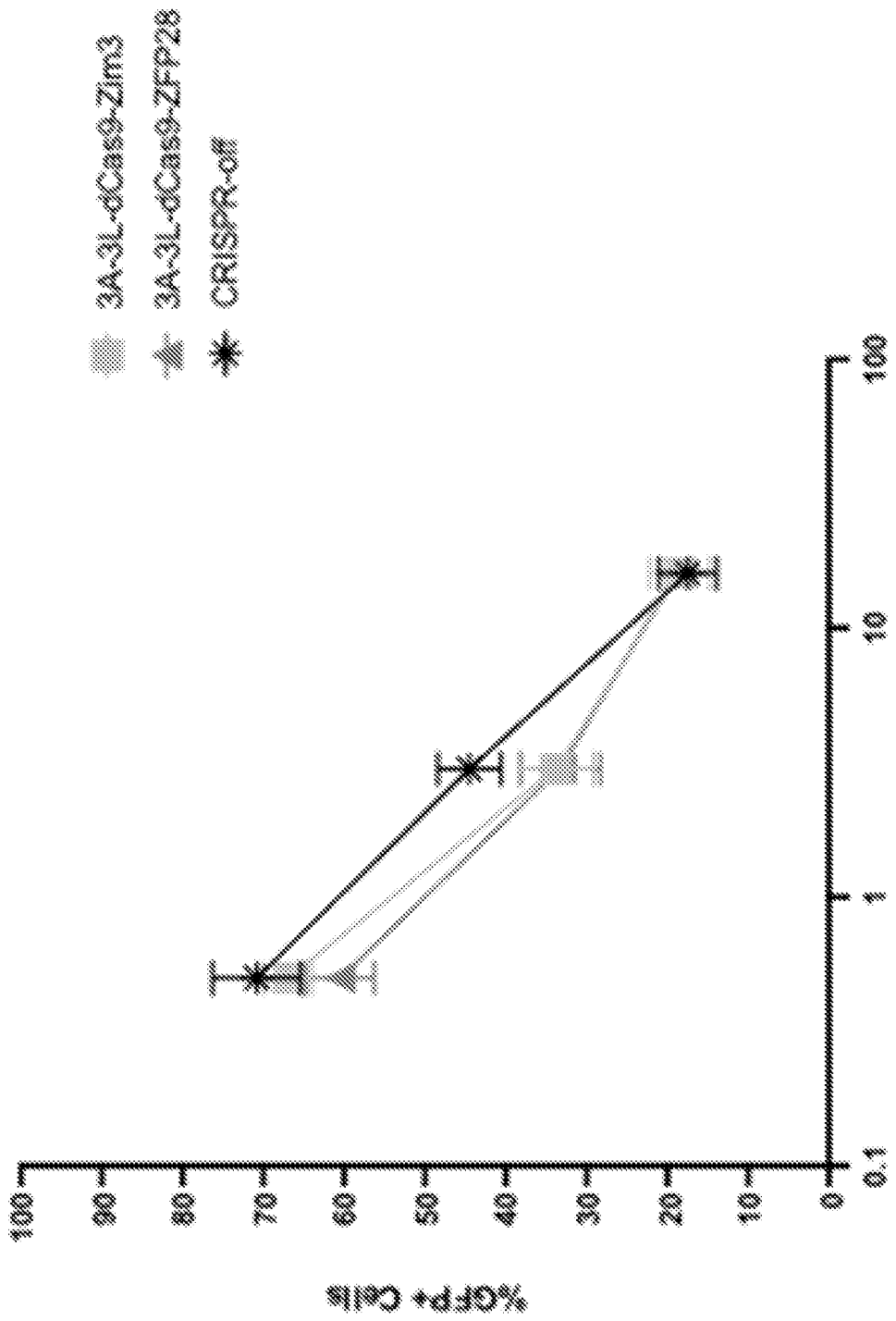

%GFP-CLTA+ cells day 30
3A-3L-dCas9-Zim3
3A-3L-dCas9-ZFP28
CRISPR-off
%GFP+ Cells
100
90
80
70
60
50
40
30
20
10
0
0.1
1
10
100
ng DNA (effector)

FIG. 19D

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2023/066436

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. [X] forming part of the international application as filed.

   b. [ ] furnished subsequent to the international filing date for the purposes of international search (Rule 13*ter*.1(a)).

      [ ] accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.

2. [ ] With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2023/066436

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/47 A61K38/00 A61P3/06 C12N9/10 C12N9/22
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K A61P A61K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2019/204766 A1 (UNIV CALIFORNIA [US]) 24 October 2019 (2019-10-24)	1-6, 9-15,18, 19,23, 24,26, 43,44
Y	[0010]-[0012], [0015], [0071], [0140], [0248], [0272], [0273], [0286], [0297], [0299]; Example 4; Table 3; Fig. 1A-1F, Fig. 4, Fig. 5; SEQ ID NOs: 1, 2, 15, 16, 23, 25-28, 32; claim 28	7,8,16, 17, 20-22, 25, 27-42, 45-48

Y	WO 2021/247570 A2 (THE REGENTS OF THE UNIV OFCALIFORNIA [US]) 9 December 2021 (2021-12-09) [0005]-[0007], [0029]-[0044], [0422]; SEQ ID NO: 99	1-48

	-/--	

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance;; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance;; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 August 2023	17/08/2023

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Benzerari, Maëlis

1

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2023/066436

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PRATIKSHA I THAKORE ET AL: "Editing the epigenome: technologies for programmable transcription and epigenetic modulation", NATURE METHODS, vol. 13, no. 2, 1 February 2016 (2016-02-01), pages 127-137, XP055623879, New York ISSN: 1548-7091, DOI: 10.1038/nmeth.3733 Abstract; page 127, left-hand column, first sentence; page 128, left-hand column; page130; left-hand column; Figure 2	1-48
Y	WO 2017/180915 A2 (UNIV DUKE [US]) 19 October 2017 (2017-10-19) claims 34,63,69,76,143; Examples, Section 3. Hypercholesterolemia; Section 4.2, results; p.59-63, p.94-123; Fig. 10-12	1-48
Y	THAKORE PRATIKSHA I. ET AL: "RNA-guided transcriptional silencing in vivo with S. aureus CRISPR-Cas9 repressors", NATURE COMMUNICATIONS, vol. 9, no. 1, 26 April 2018 (2018-04-26), XP093048361, DOI: 10.1038/s41467-018-04048-4 Retrieved from the Internet: URL:https://www.nature.com/articles/s41467-018-04048-4> the whole document	1-48
X,P	WO 2022/140577 A2 (CHROMA MEDICINE INC [US]) 30 June 2022 (2022-06-30) SEQ ID NO: 1020	1-48
E	WO 2023/093862 A1 (EPIGENIC THERAPEUTICS INC [CN]) 1 June 2023 (2023-06-01) SEQ ID NO: 97	1-48

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2023/066436

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2019204766 A1	24-10-2019	AU 2019255789 A1	22-10-2020
		BR 112020021229 A2	02-02-2021
		CA 3096713 A1	24-10-2019
		CN 112334577 A	05-02-2021
		EP 3781705 A1	24-02-2021
		GB 2587970 A	14-04-2021
		IL 277837 A	30-11-2020
		JP 2021521788 A	30-08-2021
		KR 20210003819 A	12-01-2021
		SG 11202009783W A	27-11-2020
		US 2021139918 A1	13-05-2021
		US 2023103909 A1	06-04-2023
		US 2023124253 A1	20-04-2023
		WO 2019204766 A1	24-10-2019
WO 2021247570 A2	09-12-2021	AU 2021283185 A1	05-01-2023
		CA 3184884 A1	09-12-2021
		EP 4158009 A2	05-04-2023
		IL 298434 A	01-01-2023
		JP 2023529611 A	11-07-2023
		KR 20230031832 A	07-03-2023
		WO 2021247570 A2	09-12-2021
WO 2017180915 A2	19-10-2017	EP 3443081 A2	20-02-2019
		US 2019127713 A1	02-05-2019
		WO 2017180915 A2	19-10-2017
WO 2022140577 A2	30-06-2022	AU 2021409729 A1	13-07-2023
		CA 3202977 A1	30-06-2022
		IL 303923 A	01-08-2023
		WO 2022140577 A2	30-06-2022
WO 2023093862 A1	01-06-2023	NONE	