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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
08 December 2022 (08.12.2022)

WIPO | PCT

(10) International Publication Number
WO 2022/256462 A1

(51) International Patent Classification:
C12N 15/10 (2006.01) *C12N 15/90* (2006.01)
C12N 15/113 (2010.01) *C12N 9/22* (2006.01)

(21) International Application Number:
PCT/US2022/031849

(22) International Filing Date:
01 June 2022 (01.06.2022)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

63/196,127	02 June 2021 (02.06.2021)	US
63/233,653	16 August 2021 (16.08.2021)	US
63/261,436	21 September 2021 (21.09.2021)	US
63/262,169	06 October 2021 (06.10.2021)	US
63/280,026	16 November 2021 (16.11.2021)	US
63/299,664	14 January 2022 (14.01.2022)	US
63/308,766	10 February 2022 (10.02.2022)	US
63/323,014	23 March 2022 (23.03.2022)	US
63/331,076	14 April 2022 (14.04.2022)	US

(71) Applicant: **METAGENOMI, INC.** [US/US]; 1545 Park Avenue, Emeryville, CA 94608 (US).

(72) Inventors: **THOMAS, Brian**; 1545 Park Avenue, Emeryville, CA 94608 (US). **BROWN, Christopher**; 1545 Park Avenue, Emeryville, CA 94608 (US). **DEVOTO, Audra**; 1545 Park Avenue, Emeryville, CA 94608 (US). **BUTTERFIELD, Cristina**; 1545 Park Avenue, Emeryville, CA 94608 (US). **ALEXANDER, Lisa**; 1545 Park Avenue, Emeryville, CA 94608 (US). **GOLTSMAN, Daniela, S., A.**; 1545 Park Avenue, Emeryville, CA 94608 (US).

(74) Agent: **TAVSHANJIAN, Brandon**; Wilson Sonsini Goodrich & Rosati, 650 Page Mill Road, Palo Alto, CA 94304 (US).

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

(54) Title: CLASS II, TYPE V CRISPR SYSTEMS

(57) Abstract: Described herein are methods, compositions, and systems derived from uncultivated microorganisms useful for gene editing.

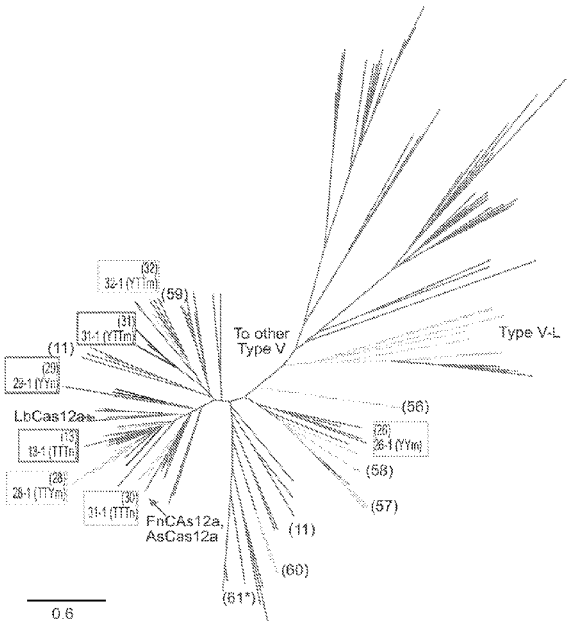


FIG. 4B

WO 2022/256462 A1

[Continued on next page]

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

Published:

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

CLASS II, TYPE V CRISPR SYSTEMS

CROSS-REFERENCE

[0001] This application claims the benefit of U.S. Provisional Application Nos: 63/196,127, filed June 2, 2021; 63/233,653, filed August 16, 2021; 63/261,436, filed September 21, 2021; 63/262,169, filed October 6, 2021; 63/280,026, filed November 16, 2021; 63/299,664, filed January 14, 2022; 63/308,766, filed February 10, 2022; 63/323,014, filed March 23, 2022; and 63/331,076, filed April 14, 2022; each of which is incorporated by reference herein in its entirety. This application is related to PCT Application No. PCT/US21/21259, which is incorporated by reference herein in its entirety.

BACKGROUND

[0002] Cas enzymes along with their associated Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) guide ribonucleic acids (RNAs) appear to be a pervasive (~45% of bacteria, ~84% of archaea) component of prokaryotic immune systems, serving to protect such microorganisms against non-self nucleic acids, such as infectious viruses and plasmids by CRISPR-RNA guided nucleic acid cleavage. While the deoxyribonucleic acid (DNA) elements encoding CRISPR RNA elements may be relatively conserved in structure and length, their CRISPR-associated (Cas) proteins are highly diverse, containing a wide variety of nucleic acid-interacting domains. While CRISPR DNA elements have been observed as early as 1987, the programmable endonuclease cleavage ability of CRISPR/Cas complexes has only been recognized relatively recently, leading to the use of recombinant CRISPR/Cas systems in diverse DNA manipulation and gene editing applications.

SEQUENCE LISTING

[0003] The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on March 5, 2021, is named 55921-728_601_SL.txt and is 23,886,645 bytes in size.

SUMMARY

[0004] In some aspects, the present disclosure provides for an engineered nuclease system comprising: (a) an endonuclease comprising a RuvC domain, wherein the endonuclease is derived from an uncultivated microorganism, and wherein the endonuclease is a Cas12a endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is

configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence. In some embodiments, the Cas12a endonuclease comprises the sequence GWxxxK. In some embodiments, the engineered guide RNA comprises UCUAC[N_{3-5}]GUAGAU (N_4). In some embodiments, the engineered guide RNA comprises CCUGC[N_4]GCAGG (N_{3-4}). In some aspects, the present disclosure provides for an engineered nuclease system comprising: (a) an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence. In some embodiments, the endonuclease comprises a RuvCI, II, or III domain. In some embodiments, the endonuclease has at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% identity to a RuvCI, II, or III domain of any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the RuvCI domain comprises a D catalytic residue. In some embodiments the RuvCII domain comprises an E catalytic residue. In some embodiments the RuvCIII domain comprises a D catalytic residue. In some embodiments, said RuvC domain does not have nuclease activity. In some embodiments, said endonuclease further comprises a WED II domain having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% identity to a WED II domain of any one of SEQ ID NOs: 1-3470 or a variant thereof. In some aspects, the present disclosure provides for an engineered nuclease system comprising: (a) an endonuclease configured to bind to a protospacer adjacent motif (PAM) sequence comprising any one of SEQ ID NOs: 3862-3913, wherein the endonuclease is a class 2, type V Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence. In some embodiments, the endonuclease further comprises a zinc finger-like domain. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence

identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, -3678, 3695-3696, 3729-3730, 3734-3735, and 3851-3857. In some aspects, the present disclosure provides for an engineered nuclease system comprising: (a) an engineered guide RNA comprising a sequence with at least 80% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3677-3678, 3695-3696, 3729-3730, 3734-3735, or 3851-3857, and (b) a class 2, type V Cas endonuclease configured to bind to the engineered guide RNA. In some embodiments, the endonuclease is configured to bind to a protospacer adjacent motif (PAM) sequence comprising any one of SEQ ID NOs: 3863-3913. In some embodiments, the guide RNA comprises a sequence complementary to a eukaryotic, fungal, plant, mammalian, or human genomic polynucleotide sequence. In some embodiments, the guide RNA is 30-250 nucleotides in length. In some embodiments, the endonuclease comprises one or more nuclear localization sequences (NLSs) proximal to an N- or C-terminus of the endonuclease. In some embodiments, the NLS comprises a sequence at least 80% identical to a sequence from the group consisting of SEQ ID NO: 3938-3953. In some embodiments, the endonuclease comprises at least one of the following mutations: S168R, E172R, N577R, or Y170R when a sequence of the endonuclease is optimally aligned to SEQ ID NO: 215. In some embodiments, the endonuclease comprises the mutations S168R and E172R when a sequence of the endonuclease is optimally aligned to SEQ ID NO: 215. In some embodiments, the endonuclease comprises the mutations N577R or Y170R when a sequence of the endonuclease is optimally aligned to SEQ ID NO: 215. In some embodiments, the endonuclease comprises the mutation S168R when a sequence of the endonuclease is optimally aligned to SEQ ID NO: 215. In some embodiments, the endonuclease does not comprise a mutation of E172, N577, or Y170. In some embodiments, the engineered nuclease system further comprises

[0005] a single- or double-stranded DNA repair template comprising from 5′ to 3′: a first homology arm comprising a sequence of at least 20 nucleotides 5′ to the target deoxyribonucleic acid sequence, a synthetic DNA sequence of at least 10 nucleotides, and a second homology arm comprising a sequence of at least 20 nucleotides 3′ to the target sequence. In some embodiments, the first or second homology arm comprises a sequence of at least 40, 80, 120, 150, 200, 300, 500, or 1,000 nucleotides. In some embodiments, the first and second homology arms are homologous to a genomic sequence of a prokaryote, bacteria, fungus, or eukaryote. In some embodiments, the single- or double-stranded DNA repair template comprises a transgene donor.

In some embodiments, the engineered nuclease system further comprises a DNA repair template comprising a double-stranded DNA segment flanked by one or two single-stranded DNA segments. In some embodiments, single-stranded DNA segments are conjugated to the 5′ ends of the double-stranded DNA segment. In some embodiments, the single stranded DNA segments are conjugated to the 3′ ends of the double-stranded DNA segment. In some embodiments, the single-stranded DNA segments have a length from 4 to 10 nucleotide bases. In some embodiments, the single-stranded DNA segments have a nucleotide sequence complementary to a sequence within the spacer sequence. In some embodiments, the double-stranded DNA sequence comprises a barcode, an open reading frame, an enhancer, a promoter, a protein-coding sequence, a miRNA coding sequence, an RNA coding sequence, or a transgene. In some embodiments, the double-stranded DNA sequence is flanked by a nuclease cut site. In some embodiments, the nuclease cut site comprises a spacer and a PAM sequence. In some embodiments, the system further comprises a source of Mg^{2+}. In some embodiments, the guide RNA comprises a hairpin comprising at least 8, at least 10, or at least 12 base-paired ribonucleotides. In some embodiments, the hairpin comprises 10 base-paired ribonucleotides. In some embodiments: (a) the endonuclease comprises a sequence at least 75%, 80%, or 90% identical to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1721 or a variant thereof; and (b) the guide RNA structure comprises a sequence at least 80%, or 90% identical to the non-degenerate nucleotides of any one of SEQ ID NOs: 3608-3609, 3853, or 3851-3857. In some embodiments, the endonuclease is configured to bind to a PAM comprising any one of SEQ ID NOs: 3863-3913. In some embodiments, the endonuclease is configured to bind to a PAM comprising SEQ ID NO: 3871. In some embodiments, the sequence identity is determined by a BLASTP, CLUSTALW, MUSCLE, MAFFT algorithm, or a CLUSTALW algorithm with the Smith-Waterman homology search algorithm parameters. In some embodiments, the sequence identity is determined by the BLASTP homology search algorithm using parameters of a wordlength (W) of 3, an expectation (E) of 10, and a BLOSUM62 scoring matrix setting gap costs at existence of 11, extension of 1, and using a conditional compositional score matrix adjustment.

[0006] In some aspects, the present disclosure provides for an engineered guide RNA comprising: (a) a DNA-targeting segment comprising a nucleotide sequence that is complementary to a target sequence in a target DNA molecule; and (b) a protein-binding segment comprising two complementary stretches of nucleotides that hybridize to form a double-stranded RNA (dsRNA) duplex, wherein the two complementary stretches of nucleotides are covalently linked to one another with intervening nucleotides, and wherein the engineered guide ribonucleic

acid polynucleotide is capable of forming a complex with an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470, and targeting the complex to the target sequence of the target DNA molecule. In some embodiments, the DNA-targeting segment is positioned 3′ of both of the two complementary stretches of nucleotides. In some embodiments, the protein binding segment comprises a sequence having at least 70%, at least 80%, or at least 90% identity to the non-degenerate nucleotides of SEQ ID NO: 3608-3609. In some embodiments, the double-stranded RNA (dsRNA) duplex comprises at least 5, at least 8, at least 10, or at least 12 ribonucleotides.

[0007] In some aspects, the present disclosure provides for a deoxyribonucleic acid polynucleotide encoding the engineered guide ribonucleic acid polynucleotide described herein.

[0008] In some aspects, the present disclosure provides for a nucleic acid comprising an engineered nucleic acid sequence optimized for expression in an organism, wherein the nucleic acid encodes a class 2, type V Cas endonuclease, and wherein the endonuclease is derived from an uncultivated microorganism, wherein the organism is not the uncultivated organism. In some embodiments, the endonuclease comprises a variant having at least 70% or at least 80% sequence identity to any one of SEQ ID NOs: 1-3470. In some embodiments, the endonuclease comprises a sequence encoding one or more nuclear localization sequences (NLSs) proximal to an N- or C-terminus of the endonuclease. In some embodiments, the NLS comprises a sequence selected from SEQ ID NOs: 3938-3953. In some embodiments, the NLS comprises SEQ ID NO: 3939. In some embodiments, the NLS is proximal to the N-terminus of the endonuclease. In some embodiments, the NLS comprises SEQ ID NO: 3938. In some embodiments, the NLS is proximal to the C-terminus of the endonuclease. In some embodiments, the organism is prokaryotic, bacterial, eukaryotic, fungal, plant, mammalian, rodent, or human.

[0009] In some aspects, the present disclosure provides for an engineered vector comprising a nucleic acid sequence encoding a class 2, type V Cas endonuclease or a Cas12a endonuclease, wherein the endonuclease is derived from an uncultivated microorganism.

[0010] In some aspects, the present disclosure provides for an engineered vector comprising a nucleic acid described herein.

[0011] In some aspects, the present disclosure provides for an engineered vector comprising a deoxyribonucleic acid polynucleotide described herein. In some embodiments, the vector is a plasmid, a minicircle, a CELiD, an adeno-associated virus (AAV) derived virion, a lentivirus, or an adenovirus.

[0012] In some aspects, the present disclosure provides for a cell comprising a vector described herein.

[0013] In some aspects, the present disclosure provides for a method of manufacturing an endonuclease, comprising cultivating any of the host cells described herein.

[0014] In some aspects, the present disclosure provides for a method for binding, cleaving, marking, or modifying a double-stranded deoxyribonucleic acid polynucleotide, comprising: (a) contacting the double-stranded deoxyribonucleic acid polynucleotide with a class 2, type V Cas endonuclease in complex with an engineered guide RNA configured to bind to the endonuclease and the double-stranded deoxyribonucleic acid polynucleotide; (b) wherein the double-stranded deoxyribonucleic acid polynucleotide comprises a protospacer adjacent motif (PAM); and (c) wherein the PAM comprises a sequence comprising any one of SEQ ID NOs: 3863-3913. In some embodiments, the double-stranded deoxyribonucleic acid polynucleotide comprises a first strand comprising a sequence complementary to a sequence of the engineered guide RNA and a second strand comprising the PAM. In some embodiments, the PAM is directly adjacent to the 5′ end of the sequence complementary to the sequence of the engineered guide RNA. In some embodiments, the PAM comprises SEQ ID NO: 3871. In some embodiments, the class 2, type V Cas endonuclease is derived from an uncultivated microorganism. In some embodiments, the double-stranded deoxyribonucleic acid polynucleotide is a eukaryotic, plant, fungal, mammalian, rodent, or human double-stranded deoxyribonucleic acid polynucleotide. In some embodiments, the method comprising delivering to the target nucleic acid locus the engineered nuclease system described herein, wherein the endonuclease is configured to form a complex with the engineered guide ribonucleic acid structure, and wherein the complex is configured such that upon binding of the complex to the target nucleic acid locus, the complex modifies the target nucleic acid locus. In some embodiments, modifying the target nucleic acid locus comprises binding, nicking, cleaving, or marking the target nucleic acid locus. In some embodiments, the target nucleic acid locus comprises deoxyribonucleic acid (DNA) or ribonucleic acid (RNA). In some embodiments, the target nucleic acid comprises genomic DNA, viral DNA, viral RNA, or bacterial DNA. In some embodiments, the target nucleic acid locus is *in vitro*. In some embodiments, the target nucleic acid locus is within a cell. In some embodiments, the cell is a prokaryotic cell, a bacterial cell, a eukaryotic cell, a fungal cell, a plant cell, an animal cell, a mammalian cell, a rodent cell, a primate cell, a human cell, or a primary cell. In some embodiments, the cell is a primary cell. In some embodiments, the primary cell is a T cell. In some embodiments, the primary cell is a hematopoietic stem cell (HSC). In some embodiments, delivering the engineered nuclease system to the target nucleic acid locus comprises delivering the nucleic acid described herein or the vector described herein. In some embodiments, delivering the engineered nuclease system to the target nucleic acid locus comprises delivering a nucleic acid comprising an open reading frame

encoding the endonuclease. In some embodiments, the nucleic acid comprises a promoter to which the open reading frame encoding the endonuclease is operably linked. In some embodiments, delivering the engineered nuclease system to the target nucleic acid locus comprises delivering a capped mRNA containing the open reading frame encoding the endonuclease. In some embodiments, delivering the engineered nuclease system to the target nucleic acid locus comprises delivering a translated polypeptide. In some embodiments, delivering the engineered nuclease system to the target nucleic acid locus comprises delivering a deoxyribonucleic acid (DNA) encoding the engineered guide RNA operably linked to a ribonucleic acid (RNA) pol III promoter. In some embodiments, the endonuclease induces a single-stranded break or a double-stranded break at or proximal to the target locus. In some embodiments, the endonuclease induces a staggered single stranded break within or 3′ to the target locus.

[0015] In some aspects, the present disclosure provides for a method of editing a TRAC locus in a cell, comprising contacting to the cell (a) an RNA-guided endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a region of the TRAC locus, wherein the engineered guide RNA comprises a targeting sequence having at least 85% identity at least 18 consecutive nucleotides of any one of SEQ ID NOs: 4316-4369. In some embodiments, the RNA-guided nuclease is a Cas endonuclease. In some embodiments, the Cas endonuclease is a class 2, type V Cas endonuclease. In some embodiments, the class 2, type V Cas endonuclease comprises a RuvC domain comprising a RuvCI subdomain, a RuvCII subdomain, and a RuvCIII subdomain. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the engineered guide RNA further comprises a sequence with at least 80% sequence identity to at least 19 of the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3677-3678, 3695-3696, 3729-3730, 3734-3735, and 3851-3857 . In some embodiments, the endonuclease comprises a sequence at least 75%, 80%, or 90% identical to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1721 or a variant thereof. In some embodiments, the guide RNA structure comprises a sequence at least 80%, or at least 90% identical to at least 19 of the non-degenerate nucleotides of any one of SEQ ID NOs: 3608-3609, 3853, or 3851-3857. In some embodiments, the method further comprises contacting to the cell or introducing to the cell a donor nucleic acid comprising a cargo sequence

flanked on a 3' or 5' end by sequence having at least 80% identity to any one of SEQ ID NOs: 4424 or 4425. In some embodiments, the cell is a peripheral blood mononuclear cell (PBMC). In some embodiments, the cell is a T-cell or a precursor thereof or a hematopoietic stem cell (HSC). In some embodiments, the cargo sequence comprises a sequence encoding a T-cell receptor polypeptide, a CAR-T polypeptide, or a fragment or derivative thereof. In some embodiments, the engineered guide RNA comprises a sequence having at least 80% identity to any one of SEQ ID NOs:4370-4423. In some embodiments, the engineered guide RNA comprises the nucleotide sequence of sgRNAs 1-54 from Table 5A comprising the corresponding chemical modifications listed in Table 5A. In some embodiments, the engineered guide RNA comprises a targeting sequence having at least 80% sequence identity to any one of SEQ ID NOs: 4334, 4350, or 4324. In some embodiments, the engineered guide RNA comprises a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 4388, 4404, or 4378. In some embodiments, the engineered guide RNA comprises the nucleotide sequence of sgRNAs 9, 35, or 19 from Table 5A.

[0016] In some aspects, the present disclosure provides for an engineered nuclease system comprising: (a) an RNA-guided endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence, wherein the engineered guide RNA comprises at least one of the following modifications:(i) a 2'-O methyl or a 2'-fluoro base modification of at least one nucleotide within the first 4 bases of the 5' end of the engineered guide RNA or the last 4 bases of a 3' end of the engineered guide RNA; (ii) a thiophosphate (PS) linkage between at least 2 of the first five bases of a 5' end of the engineered guide RNA, or a thiophosphate linkage between at least two of the last five bases of a 3' end of the engineered guide RNA; (iii) a thiophosphate linkage within a 3' stem or a 5' stem of the engineered guide RNA; (iv) a 2'-O methyl or 2'base modification within a 3' stem or a 5' stem of the engineered guide RNA; (v) a 2'-fluoro base modification of at least 7 bases of a spacer region of the engineered guide RNA; and (vi) a thiophosphate linkage within a loop region of the engineered guide RNA. In some embodiments, the engineered guide RNA comprises a 2'-O methyl or a 2'-fluoro base modification of at least one nucleotide within the first 5 bases of a 5' end of the engineered guide RNA or the last 5 bases of a 3' end of the engineered guide RNA. In some embodiments, the engineered guide RNA comprises a 2'-O methyl or a 2'-fluoro base modification at a 5' end of the engineered guide RNA or a 3' end of the engineered guide RNA. In some embodiments, the engineered guide RNA comprises a thiophosphate (PS) linkage between at least 2 of the first five bases of a 5' end of the engineered

guide RNA, or a thiophosphate linkage between at least two of the last five bases of a 3' end of the engineered guide RNA. In some embodiments, the engineered guide RNA comprises a thiophosphate linkage within a 3' stem or a 5' stem of the engineered guide RNA. In some embodiments, the engineered guide RNA comprises a 2'-O methyl base modification within a 3' stem or a 5' stem of the engineered guide RNA. In some embodiments, the engineered guide RNA comprises a 2'-fluoro base modification of at least 7 bases of a spacer region of the engineered guide RNA. In some embodiments, the engineered guide RNA comprises a thiophosphate linkage within a loop region of the engineered guide RNA. In some embodiments, the engineered guide RNA comprises at least three 2'-O methyl or 2'-fluoro bases at the 5' end of the engineered guide RNA, two thiophosphate linkages between the first 3 bases of the 5' end of the engineered guide RNA, at least 4 2'-O methyl or 2'-fluoro bases at the 4' end of the engineered guide RNA, and three thiophosphate linkages between the last three bases of the 3' end of the engineered guide RNA. In some embodiments, the engineered guide RNA comprises at least two 2'-O-methyl bases and at least two thiophosphate linkages at a 5' end of the engineered guide RNA and at least one 2'-O-methyl bases and at least one thiophosphate linkage at a 3' end of the engineered guide RNA. In some embodiments, the engineered guide RNA comprises at least one 2'-O-methyl base in both the 3' stem or the 5' stem region of the engineered guide RNA. In some embodiments, the engineered guide RNA comprises at least one to at least fourteen 2'-fluoro bases in the spacer region excluding a seed region of the engineered guide RNA. In some embodiments, the engineered guide RNA comprises at least one 2'-O-methyl base in the 5' stem region of the engineered guide RNA and at least one to at least fourteen 2'-fluoro bases in the spacer region excluding a seed region of the guide RNA. In some embodiments, the guide RNA comprises a spacer sequence targeting a VEGF-A gene. In some embodiments, the guide RNA comprises a spacer sequence having at least 80% identity to SEQ ID NO: 3985. In some embodiments, the guide RNA comprises the nucleotides of guide RNAs 1-7 from Table 7 comprising the chemical modifications listed in Table 7. In some embodiments, the RNA-guided nuclease is a Cas endonuclease. In some embodiments, the Cas endonuclease is a class 2, type V Cas endonuclease In some embodiments, the class 2, type V Cas endonuclease comprises a RuvC domain comprising a RuvCI subdomain, a RuvCII subdomain, and a RuvCIII subdomain. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1721 or a variant thereof. In some embodiments, the engineered guide RNA

comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3677-3678, 3695-3696, 3729-3730, 3734-3735, and 3851-3857. In some embodiments, the engineered guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3608 -3609, 3853, or 3851-3857.

[0017] In some aspects, the present disclosure provides for a host cell comprising an open reading frame encoding a heterologous endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the endonuclease has at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1721, or a variant thereof. In some embodiments, the host cell is an *E. coli* cell or a mammalian cell. In some embodiments, the host cell is an *E. coli* cell, wherein the *E. coli* cell is a λDE3 lysogen or the *E. coli* cell is a BL21(DE3) strain. In some embodiments, the *E. coli* cell has an *ompT lon* genotype. In some embodiments, the open reading frame is operably linked to a T7 promoter sequence, a T7-lac promoter sequence, a lac promoter sequence, a tac promoter sequence, a trc promoter sequence, a ParaBAD promoter sequence, a PrhaBAD promoter sequence, a T5 promoter sequence, a *cspA* promoter sequence, an $araP_{BAD}$ promoter, a strong leftward promoter from phage lambda (pL promoter), or any combination thereof. In some embodiments, the open reading frame comprises a sequence encoding an affinity tag linked in-frame to a sequence encoding the endonuclease. In some embodiments, the affinity tag is an immobilized metal affinity chromatography (IMAC) tag. In some embodiments, the IMAC tag is a polyhistidine tag. In some embodiments, the affinity tag is a myc tag, a human influenza hemagglutinin (HA) tag, a maltose binding protein (MBP) tag, a glutathione S-transferase (GST) tag, a streptavidin tag, a FLAG tag, or any combination thereof. In some embodiments, the affinity tag is linked in-frame to the sequence encoding the endonuclease via a linker sequence encoding a protease cleavage site. In some embodiments, the protease cleavage site is a tobacco etch virus (TEV) protease cleavage site, a PreScission® protease cleavage site, a Thrombin cleavage site, a Factor Xa cleavage site, an enterokinase cleavage site, or any combination thereof. In some embodiments, the open reading frame is codon-optimized for expression in the host cell. In some embodiments, the open reading frame is provided on a vector. In some embodiments, the open reading frame is integrated into a genome of the host cell.

[0018] In some aspects, the present disclosure provides for a culture comprising any of the host cells described herein in compatible liquid medium.

[0019] In some aspects, the present disclosure provides for a method of producing an endonuclease, comprising cultivating any of the host cells described herein in compatible growth medium. In some embodiments, the method further comprises inducing expression of the endonuclease. In some embodiments, the inducing expression of the nuclease is by addition of an additional chemical agent or an increased amount of a nutrient, or by temperature increase or decrease. In some embodiments, an additional chemical agent or an increased amount of a nutrient comprises Isopropyl β-D-1-thiogalactopyranoside (IPTG) or additional amounts of lactose. In some embodiments, the method further comprises isolating the host cell after the cultivation and lysing the host cell to produce a protein extract. In some embodiments, the method further comprises isolating the endonuclease. In some embodiments, the isolating comprises subjecting the protein extract to IMAC, ion-exchange chromatography, anion exchange chromatography, or cation exchange chromatography. In some embodiments, the open reading frame comprises a sequence encoding an affinity tag linked in-frame to a sequence encoding the endonuclease. In some embodiments, the affinity tag is linked in-frame to the sequence encoding the endonuclease via a linker sequence encoding protease cleavage site. In some embodiments, the protease cleavage site comprises a tobacco etch virus (TEV) protease cleavage site, a PreScission® protease cleavage site, a Thrombin cleavage site, a Factor Xa cleavage site, an enterokinase cleavage site, or any combination thereof. In some embodiments, the method further comprises cleaving the affinity tag by contacting a protease corresponding to the protease cleavage site to the endonuclease. In some embodiments, the affinity tag is an IMAC affinity tag. In some embodiments, the method further comprises performing subtractive IMAC affinity chromatography to remove the affinity tag from a composition comprising the endonuclease.

[0020] In some aspects, the present disclosure provides for a system comprising (a) a class 2, Type V-A Cas endonuclease configured to bind a 3- or 4-nucleotide PAM sequence, wherein the endonuclease has increased cleavage activity relative to sMbCas12a; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the class 2, Type V-A Cas endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid comprising a target nucleic acid sequence. In some embodiments, the cleavage activity is measured *in vitro* by introducing the endonucleases alongside compatible guide RNAs to cells comprising the target nucleic acid and detecting cleavage of the target nucleic acid sequence in the cells. In some embodiments, the class 2, Type V-A Cas endonuclease comprises a sequence having at least 75% identity to any one of 215-225 or a variant thereof. In some embodiments, the engineered guide RNA comprises a sequence

having at least 80% identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the target nucleic acid further comprises a YYN PAM sequence proximal to the target nucleic acid sequence. In some embodiments, the class 2, Type V-A Cas endonuclease has at least about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or 200%, or more increased activity relative to sMbCas12a.

[0021] In some aspects, the present disclosure provides for a system comprising: (a) a class 2, Type V-A' Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA comprises a sequence having at least 80% identity to about 19 to about 25 or about 19 to about 31 consecutive nucleotides of a natural effector repeat sequence of a class 2, Type V Cas endonuclease. In some embodiments, the natural effector repeat sequence is any one of SEQ ID NOs: 3560-3572. In some embodiments, the class 2, Type V-A' Cas endonuclease has at least 75% identity to SEQ ID NO: 126.

[0022] In some aspects, the present disclosure provides for a system comprising: (a) a class 2, Type V-L endonuclease, and (b) an engineered guide RNA, wherein the engineered guide RNA comprises a sequence having at least 80% identity to about 19 to about 25 or about 19 to about 31 consecutive nucleotides of a natural effector repeat sequence of a class 2, Type V Cas endonuclease. In some embodiments, the class 2, Type V-L endonuclease has at least 75% sequence identity to any one of SEQ ID NOs: 793-1163.

[0023] In some aspects, the present disclosure provides for a method of disrupting the VEGF-A locus in a cell, comprising introducing to the cell: (a) a class 2, type V Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a region of the VEGF-A locus, wherein the engineered guide RNA comprises a targeting sequence having at least 80% identity to SEQ ID NO: 3985; or wherein the engineered guide RNA comprises the nucleotide sequence of any one of guide RNAs 1-7 from Table 7 In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1721 or a variant thereof. In some embodiments, the engineered guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3677-3678, 3695-3696, 3729-3730, 3734-3735, and 3851-3857. In some embodiments, the engineered guide RNA comprises a sequence with at

least 80% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3608 - 3609, 3853, or 3851-3857.

[0024] In some aspects, the present disclosure provides for a method of disrupting a locus in a cell, comprising contacting to the cell a composition comprising: (a) a class 2, type V Cas endonuclease having at least 75% identity to any one of SEQ ID NOs: 215-225 or a variant thereof; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a region of the locus, wherein the class 2, type V Cas endonuclease has at least equivalent cleavage activity to spCas9 in the cell. In some embodiments, the cleavage activity is measured in vitro by introducing the endonucleases alongside compatible guide RNAs to cells comprising the target nucleic acid and detecting cleavage of the target nucleic acid sequence in the cells. In some embodiments, the composition comprises 20 pmoles or less of the class 2, type V Cas endonuclease. In some embodiments, the composition comprises 1 pmol or less of the class 2, type V Cas endonuclease.

In some aspects, the present disclosure provides for a method of disrupting a CD38 locus in a cell, comprising introducing to the cell: (a) a class 2, type V Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a region of the CD38 locus, wherein the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4466-4503 and 5686; orwherein the engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4428-4465 and 5685. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least

about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, and 6033-6036. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 4466, 4467, 4468, 4479, 4484, 4490, 4492, 4493, 4495, 4498. In some embodiments, the engineered guide RNA comprises a nucleotide sequence having at least 80% identity to any one of SEQ ID NOs: 4428, 4429, 4430, 4436, 4441, 4446, 4452, 4454, 4455, 4460, or 4461. In some embodiments, the cell is a eukaryotic cell, T-cell, hematopoietic stem cell, or precursor thereof.

In some aspects, the present disclosure provides for a method of disrupting a TIGIT locus in a cell, comprising introducing to the cell: (a) a class 2, type V Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a region of the TIGIT locus, wherein the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4521-4537; or wherein the engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4504-4520. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735,

3851-3857, and 6033-6036. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 4521, 4527, 4528, 4535, or 4536. In some embodiments, the engineered guide RNA comprises a nucleotide sequence having at least 80% identity to any one of SEQ ID NOs: 4504, 4510, 4511, 4518, or 4519. In some embodiments, the cell is a eukaryotic cell, T-cell, hematopoietic stem cell, or precursor thereof.

[0025] In some aspects, the present disclosure provides for a method of disrupting an AAVS1 locus in a cell, comprising introducing to the cell: (a) a class 2, type V Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a region of the AAVS1 locus, wherein the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4569-4599; or wherein the engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4538-4568. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, 6033-6036. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the engineered guide

RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 4574, 4577, 4578, 4579, 4582, 4584, 4585, 4586, 4587, 4589, 4590, 4591, 4592, 4593, 4595, 4596, or 4598. In some embodiments, the engineered guide RNA comprises a nucleotide sequence having at least 80% identity to any one of SEQ ID NOs: 4543, 4546, 4547, 4548, 4551, 4553, 4554, 4555, 4556, 4558, 4559, 4560, 4561, 4562, 4565, or 4567. In some embodiments, the cell is a eukaryotic cell, T-cell, hematopoietic stem cell, hepatocyte, or precursor thereof.

[0026] In some aspects, the present disclosure provides for a method of disrupting a B2M locus in a cell, comprising introducing to the cell: (a) a class 2, type V Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a region of the B2M locus, wherein the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4676-4751; or wherein the engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4600-4675. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857 and 6033-6036. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides

having at least 80% identity to any one of SEQ ID NOs: 4676, 4678-4687, 4690, 4692, 4698-4707, 4720-4723, 4725-4726, 4732-4733, 4736-4737, 4741, or 4750-4751. In some embodiments, the engineered guide RNA comprises a nucleotide sequence having at least 80% identity to any one of SEQ ID NOs: 4600, 4602-4611, 4614, 4616, 4622-4631, 4644-4647, 4649-4650, 4656-4657, 4660-4661, 4665, or 4674-4675. In some embodiments, the cell is a eukaryotic cell, T-cell, hematopoietic stem cell, or precursor thereof.

[0027] In some aspects, the present disclosure provides for a method of disrupting a CD2 locus in a cell, comprising introducing to the cell: (a) a class 2, type V Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a region of the CD2 locus, wherein the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4837-4921; or wherein the engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4752-4836. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, and 6033-6036. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 4837, 4844, 4845, 4848, 4857-4858,

4883, 4887, 4892-4893, 4904-4909, 4914, 4916, or 4918. In some embodiments, the engineered guide RNA has at least 80% sequence identity to any of the guide RNAs from Table 14E that target any one of SEQ ID NOs: 4837, 4844, 4845, 4848, 4857-4858, 4883, 4887, 4892-4893, 4904-4909, 4914, 4916, or 4918. In some embodiments, the cell is a eukaryotic cell, T-cell, hematopoietic stem cell, or precursor thereof.

[0028] In some aspects, the present disclosure provides for a method of disrupting a CD5 locus in a cell, comprising introducing to the cell: (a) a class 2, type V Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a region of the CD5 locus, wherein the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4946-4969; or wherein the engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4922-4945. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, and 6033-6036. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 4946-4947, 4949, 4951, 4957-4960, 4963, 4967, or 4969. In some embodiments, the engineered guide RNA has at least 80%

sequence identity to any of the guide RNAs from Table 14F that target any one of SEQ ID NOs: 4946-4947, 4949, 4951, 4957-4960, 4963, 4967, or 4969. In some embodiments, the cell is a eukaryotic cell, T-cell, hematopoietic stem cell, or precursor thereof.

[0029] In some aspects, the present disclosure provides for a method of disrupting a mouse TRAC locus in a cell, comprising introducing to the cell: (a) a class 2, type V Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a region of the mouse TRAC locus, wherein the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5126-5195, 5682, or 5684; or wherein the engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5056-5125, 5681, or 5683. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3677-3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 5126-5130, 5133-5143, 5147-5150, 5172-5173, 5184-5189, or 5192-5194. In some embodiments, the engineered guide RNA has at least 80% sequence identity to any of the guide RNAs from Table 14G that target any one of SEQ ID NOs: 5126-5130, 5133-5143, 5147-5150,

5172-5173, 5184-5189, or 5192-5194. In some embodiments, the cell is a eukaryotic cell, T-cell, hematopoietic stem cell, or precursor thereof.

[0030] In some aspects, the present disclosure provides for a method of disrupting a mouse TRBC1 or TRBC2 locus in a cell, comprising introducing to the cell: (a) a class 2, type V Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a region of the mouse TRBC1 or TRBC2 locus, wherein the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5211-5225 or 5247-5267; or wherein the engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5196-5210 or 5226-5246. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3677-3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 5211, 5213-5215, 5217, 5221, 5223, 5247, 5249-5250, 5252-5253, 5258-5259, or 5264. In some embodiments, the engineered guide RNA has at least 80% sequence identity to any of the guide RNAs from Table 14H that target any one of SEQ ID NOs: 5211,

5213-5215, 5217, 5221, 5223, 5247, 5249-5250, 5252-5253, 5258-5259, or 5264. In some embodiments, the cell is a eukaryotic cell, T-cell, hematopoietic stem cell, or precursor thereof.

[0031] In some aspects, the present disclosure provides for a method of disrupting a human TRBC1 or TRBC2 locus in a cell, comprising introducing to the cell: (a) a class 2, type V Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a region of the human TRBC1 or TRBC2 locus, wherein the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at last about 99% sequence identity to any one of SEQ ID NOs: 5661-5679; or wherein the engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5642-5660. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 5661-5663, 5672-5675, or 5678. In some embodiments, the engineered guide RNA has at least 80% sequence identity to any of the guide RNAs from Table 14I that target any one of SEQ ID NOs: 5661-5663, 5672-5675, or 5678. In some embodiments, the cell is a eukaryotic cell, T-cell, hematopoietic stem cell, or precursor thereof.

[0032] In some aspects, the present disclosure provides for a method of disrupting an HPRT locus in a cell, comprising introducing to the cell: (a) a class 2, type V Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a region of the HPRT locus, wherein the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5562-5641; or wherein the engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5482-5561. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 5562-5564 or 5568. In some embodiments, the engineered guide RNA has at least 80% sequence identity to any of the guide RNAs from Table 14J that target any one of SEQ ID NOs: 5562-5564 or 5568. In some embodiments, the cell is a eukaryotic cell, hepatocyte, T-cell, hematopoietic stem cell, or precursor thereof.

[0033] In some aspects, the present disclosure provides for a method of disrupting an APO-A1 locus in a cell, comprising introducing to the cell: (a) a class 2, type V Cas endonuclease; and (b)

an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a region of the APO-A1 locus, wherein the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5861-5874; or wherein the engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5847-5860. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 5861-5866 or 5868-5869. In some embodiments, the engineered guide RNA has at least 80% sequence identity to any of the guide RNAs from Table 43A that target any one of SEQ ID NOs: 5861-5866 or 5868-5869. In some embodiments, the cell is a eukaryotic cell, hepatocyte, T-cell, hematopoietic stem cell, or precursor thereof.

[0034] In some aspects, the present disclosure provides for a method of disrupting an ANGPTL3 locus in a cell, comprising introducing to the cell: (a) a class 2, type V Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to

hybridize to a region of the ANGPTL3 locus, wherein the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5953-6030; or wherein the engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5875-5952. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 5955-5963, 5968-5975, 5979-5987, 5989-5993, 5997, 5999, 6003-6010, 6014-6016, 6024-6025, or 6027-6030. In some embodiments, the engineered guide RNA has at least 80% sequence identity to any of the guide RNAs from Table 43B that target any one of SEQ ID NOs: 5955-5963, 5968-5975, 5979-5987, 5989-5993, 5997, 5999, 6003-6010, 6014-6016, 6024-6025, or 6027-6030. In some embodiments, the cell is a eukaryotic cell, hepatocyte, T-cell, hematopoietic stem cell, or precursor thereof.

[0035] In some aspects, the present disclosure provides for a method of disrupting a human Rosa26 locus in a cell, comprising introducing to the cell: (a) a class 2, type V Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence

configured to hybridize to a region of the human Rosa26 locus, wherein the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5013-5055; or wherein the engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4970-5012. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the cell is a eukaryotic cell, hepatocyte, T-cell, hematopoietic stem cell, or precursor thereof.

[0036] In some aspects, the present disclosure provides for a method of disrupting a FAS locus in a cell, comprising introducing to the cell: (a) a class 2, type V Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a region of the FAS locus, wherein the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5367-5465; or wherein the engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least

about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5268-5366. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the cell is a eukaryotic cell, hepatocyte, T-cell, hematopoietic stem cell, or precursor thereof.

[0037] In some aspects, the present disclosure provides for a method of disrupting a PD-1 locus in a cell, comprising introducing to the cell: (a) a class 2, type V Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a region of the PD-1 locus, wherein the engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5474-5481; or wherein the engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5466-5473. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the engineered guide RNA

comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the cell is a eukaryotic cell, hepatocyte, T-cell, hematopoietic stem cell, or precursor thereof.

[0038] In some aspects, the present disclosure provides for an engineered nuclease system comprising: (a) an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 215 or a variant thereof; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence, wherein the system has reduced immunogenicity when administered to a human subject compared to an equivalent system comprising a Cas9 enzyme. In some embodiments, the Cas9 enzyme is an SpCas9 enzyme. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the immunogenicity is antibody immunogenicity.

[0039] An aspsect of the present disclosure provides for a method of disrupting a mouse HAO-1 locus in a cell, comprising introducing to the cell: (a) a class 2, type V Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a region of the mouse HAO-1 locus, wherein the engineered guide RNA comprises the nucleotides of guide RNAs mH29-1_37, mH29-15_37, mH29-29_37 from Table 25 comprising the nucleotide modifications described in Table 25; or wherein the engineered guide RNA comprises any one of SEQ ID NOs: 4184-4225. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99%

sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the cell is a eukaryotic cell, hepatocyte, T-cell, hematopoietic stem cell, or precursor thereof. In some embodiments, the engineered guide RNA comprises the nucleotides of guide RNAs mH29-15_37 or mH29-29_37 from Table 25 comprising the nucleotide modifications described in Table 25. In some embodiments, the method further comprises disrupting expression of glycolate oxidase from the HAO-1 locus.

[0040] In some aspects, the present disclosure provides for a method of disrupting a human TRAC locus in a cell, comprising introducing to the cell: (a) a class 2, type V Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a region of the human TRAC locus, wherein the engineered guide RNA comprises the nucleotides of MG29-1-TRAC-sgRNA-35 from Table 28B comprising the nucleotide modifications described in Table 28B. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the cell is a eukaryotic cell, hepatocyte, T-cell, hematopoietic stem cell, or precursor thereof.

[0041] An aspec of the present disclosure provides for a method of disrupting an albumin locus in a cell, comprising introducing to the cell: (a) a class 2, type V Cas endonuclease; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to

hybridize to a region of the albumin locus, wherein the engineered guide RNA comprises the nucleotides of mAlb298-37, mAlb2912-37, mAlb2918-37, or mAlb298-34 from Table 29 comprising the nucleotide modifications described in Table 29; or wherein the engineered guide RNA comprises the nucleotides of mAlb29-8-44, mAlb29-8-50, mAlb29-8-50b, mAlb29-8-51b, mAlb29-8-52b, mAlb29-8-53b, or mAlb29-8-54b comprising the nucleotide modifications described in Table 46. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036. In some embodiments, the guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609. In some embodiments, the cell is a eukaryotic cell, hepatocyte, T-cell, hematopoietic stem cell, or precursor thereof. In some embodiments, the engineered guide RNA comprises the nucleotides of mAlb298-37, mAlb2912-37, mAlb2918-37, or mAlb298-34 from Table 29 comprising the nucleotide modifications described in Table 29.

[0042] In some aspects, the present disclosure provides for an engineered guide RNA comprising: (a) a DNA-targeting segment comprising a nucleotide sequence that is complementary to a target sequence in a target DNA molecule; and (b) a protein-binding segment configured to bind to a class 2, type V Cas endonuclease, and wherein the guide RNA comprises a nucleotide modification pattern depicted in any one of SEQ ID NOs: 5695-5701 in Table 34. In some embodiments, the guide RNA comprises mAlb29-8-44, mAlb29-8-50, mAlb29-8-37, or mAlb29-12-44. In some embodiments, the guide RNA comprises hH29-4_50, hH29-21_50, hH29-23_50, hH29-41_50, hH29-4_50b, hH29-21_50b, hH29-23_50b, or hH29-41_50b, mH29-1-50, mH29-15-50, mH29-29-50, mH29-1-50b, mH29-15-50b, or mH29-29-50b. In some embodiments, the DNA-targeting segment is configured to hybridize to an HAO-1 gene or an albumin gene. In some embodiments, the class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the class 2, type V Cas

endonuclease comprises an endonuclease having at least 75% sequence identity to SEQ ID NO: 215.

[0043] In some aspects, the present disclosure provides for an engineered nuclease system comprising: (a) an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof, or a nucleotide sequence encoding the enonuclease; and (b) a polynucleotide sequence encoding a CRISPR array, wherein the CRISPR array is configured to be processed by the endonuclease to an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence, wherein the spacer sequence is configured to hybridize to an albumin gene. In some embodiments, the polynucleotide sequence comprises a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to SEQ ID NO: 5712. In some embodiments, the endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof. In some embodiments, the endonuclease comprises an endonuclease having at least 75% sequence identity to SEQ ID NO: 215.

[0044] In some aspects, the present disclosure provides for an engineered nuclease system comprising: (a) an endonuclease having at least 75% sequence identity to SEQ ID NOs: 470 or a variant thereof; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and the engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence. In some embodiments, the engineered guide RNA comprises a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to SEQ ID NO: 6031. In some embodiments, the endonuclease is configured to be selective for a 5' PAM sequence comprising SEQ ID NO: 6032.

[0045] In some aspects, the present disclosure provides for an engineered nuclease system comprising: (a) an endonuclease having at least at least about 80%, at least about 85%, at least about 86%, at least about 87%, at least about 88%, at least about 89%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 2824, 2841, or 2896, or a variant thereof; and (b) an engineered guide RNA, wherein the engineered guide RNA is configured to form a complex with the endonuclease and

the engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence, wherein the engineered guide RNA comprises a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs:6033, 6034, or 6035. In some embodiments, the endonuclease has at least 80% sequence identity to SEQ ID NO: 2824 and the engineered guide RNA has at least 80% sequence identity to SEQ ID NO: 6033. In some embodiments, the endonuclease has at least 80% sequence identity to SEQ ID NO: 2841 and the engineered guide RNA has at least 80% sequence identity to SEQ ID NO: 6034. In some embodiments, the endonuclease has at least 80% sequence identity to SEQ ID NO: 2896 and the engineered guide RNA has at least 80% sequence identity to SEQ ID NO: 6035. In some embodiments, the endonuclease is configured to be selective for a 5' PAM sequence comprising any one of SEQ ID NOs: 6037-6039.

[0046] In some aspects, the present disclosure provides for a lipid nanoparticle comprising: (a) any of the endonucleases described herein; (b) any of the engineered guide RNAs described herein: (c) a cationic lipid; (d) a sterol; (e) a neutral lipid; and (f) a PEG-modified lipid. In some embodiments, the cationic lipid comprises C12-200, the sterol comprises cholesterol, the neutral lipid comprises DOPE, or the PEG-modified lipid comprises DMG-PEG2000. In some embodiments, the cationic lipid comprises any of the cationic lipids depicted in FIG. 109.

[0047] Additional aspects and advantages of the present disclosure will become readily apparent to those skilled in this art from the following detailed description, wherein only illustrative embodiments of the present disclosure are shown and described. As will be realized, the present disclosure is capable of other and different embodiments, and its several details are capable of modifications in various obvious respects, all without departing from the disclosure. Accordingly, the drawings and description are to be regarded as illustrative in nature, and not as restrictive.

INCORPORATION BY REFERENCE

[0048] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

[0049] The novel features of the invention are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present invention will be obtained

by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention are utilized, and the accompanying drawings of which:

[0050] FIG. 1 depicts example organizations of CRISPR/Cas loci of different classes and types that were previously documented before this disclosure.

[0051] FIG. 2 depicts environmental distribution of MG nucleases described herein. Protein length is shown for representatives of the MG29 protein family. Shades of circle indicates the environment or environment type from which each protein was identified (dark gray circle indicates high temperature environment source; light gray circle indicates non-high temperature environment source). N/A denotes the type of environment the sample was collected from is unknown.

[0052] FIG. 3 depicts the number of predicted catalytic residues present in MG nucleases detected from sample types described herein (e.g. **FIG.**). Protein length is shown for representatives of the MG29 protein family. The number of catalytic residues that were predicted for each protein are indicated in the figure legend (3.0 residues). The first, second and third catalytic residues are located in the RuvCI domain, the RuvCII domain and the RuvCIII domain, respectively.

[0053] FIGs. 4A and 4B show the diversity of CRISPR Type V-A effectors. **FIG. 4A** depicts per family distribution of taxonomic classification of contigs encoding the novel Type V-A effectors. **FIG. 4B** depicts the phylogenetic gene tree inferred from an alignment of 119 novel and 89 reference Type V effector sequences. MG families are denoted in parentheses. PAM requirements for active nucleases are outlined with boxes associated with the family. Non-Type V-A reference sequences were used to root the tree (*MG61 family requires a crRNA with an alternative stem-loop sequence).

[0054] FIGs. 1A, 5B, 5C, and 5D provide various characteristic information about nucleases described herein. **FIG. 5A** depicts the per family distribution of effector protein length and the type of sample; **FIG. 5B** shows the presence of RuvC catalytic residues. **FIG. 5C** shows the number of CRISPR arrays having various repeat motifs. **FIG. 5D** depicts the per family distribution of repeat motifs.

[0055] FIGs. 2A and 6B depict multiple sequence alignment of catalytic and PAM interacting regions in Type V-A sequences. *Francisella novicida* Cas12a (FnCas12) is a reference sequence. Other reference sequences are *Acidaminococcus sp. (*AsCas12a), *Moraxella bovoculi* (MbCas12a), and *Lachnospiraceae bacterium* ND2006 (LbCas12a). **FIG. 6A** shows blocks of conservation around the DED catalytic residues in RuvC-I (left), RuvC-II (middle), and RuvC-III (right) regions. **FIG. 6B** shows WED-II and PAM interacting regions containing residues

involved in PAM recognition and interaction. The grey boxes underneath the FnCas12a sequence identify the domains. Darker boxes in the alignments indicate increased sequence identity. Black boxes over the FnCas12a sequence indicate catalytic residues (and positions) of the reference sequence. Grey boxes indicate domains in the reference sequence at the top of the alignment (FnCas12a). Black boxes indicate catalytic residues (and positions) of the reference sequence.

[0056] FIGs. 3A and 7B depicts Type V-A and associated V-A' effectors. **FIG. 7A** shows Type V-A (MG26-1) and V-A' (MG26-2) indicated by arrows pointing in the direction of transcription. The CRISPR array is indicated by a gray bar. Predicted domains for each protein in the contig are indicated by boxes. **FIG. 7B** shows sequence alignments of Type V-A' MG26-2 and AsCas12a reference sequence. Top: RuvC-I domain. Middle: region containing the RuvC-I and RuvC-II catalytic residues. Bottom: region containing the RuvC-III catalytic residue. Catalytic residues are indicated by squares.

[0057] FIG. 8 depicts a schematic representation of the structure of a sgRNA and a target DNA in a ternary complex with AacC2C1 (see Yang, Hui, Pu Gao, Kanagalaghatta R. Rajashankar, and Dinshaw J. Patel. 2016. "PAM-Dependent Target DNA Recognition and Cleavage by C2c1 CRISPR-Cas Endonuclease." Cell 167 (7): 1814–28.e12 which is incorporated by reference herein in its entirety).

[0058] FIG. 9 depicts the effects of mutations or truncations in the R-AR domains of an sgRNA on AacC2c1-mediated cleavage of linear plasmid DNA; WT, wild-type sgRNA. The mutant nucleotides within sgRNA (lanes 1–5) are highlighted in the left panel. Δ15: 15 nt deleted from the sgRNA R-AR 1 region. Δ12: 12 nt have been removed from the sgRNA J2/4 R-AR 1 region (see Liu, Liang, Peng Chen, Min Wang, Xueyan Li, Jiuyu Wang, Maolu Yin, and Yanli Wang. 2017. "C2c1-sgRNA Complex Structure Reveals RNA-Guided DNA Cleavage Mechanism." Molecular Cell 65 (2): 310–22 which is incorporated by reference herein in its entirety).

[0059] FIGs. 10A, 10B, and 10C demonstrate that the CRISPR RNA (crRNA) structure is conserved among Type V-A systems. **FIG. 10A** shows the fold structure of the reference crRNA sequence in the LbCpf1 system. **FIG. 10B** shows multiple sequence alignment of CRISPR repeats associated with novel Type V-A systems. The LbCpf1 processing site is indicated with a black bar. **FIG. 10C** shows the fold structure of MG61-2 putative crRNA with an alternative stem-loop motif CCUGC[N_{3-4}]GCAGG. **FIG. 10D** shows multiple sequence alignment of CRISPR repeats with the alternative repeat motif sequence. The processing sites and loop are indicated.

[0060] FIG. 11 depicts a predicted structure of a guide RNA utilized herein (SEQ ID NO: 3608).

[0061] FIG. 12 depicts predicted structures of corresponding sgRNAs of MG enzymes described herein (clockwise, SEQ ID NOs: 3636, 3637, 3641, 3640).

[0062] FIG. 13 depicts predicted structures of corresponding sgRNAs of MG enzymes described herein (clockwise, SEQ ID NOs: 3644, 3645, 3649, 3648).

[0063] FIG. 14 depicts predicted structures of corresponding sgRNAs of MG enzymes described herein (clockwise, SEQ ID NOs: 3652, 3653, 3657, 3656).

[0064] FIG. 15 depicts predicted structures of corresponding sgRNAs of MG enzymes described herein (clockwise, SEQ ID NOs: 3660, 3661, 3665, 3664).

[0065] FIG. 16 depicts predicted structures of corresponding sgRNAs of MG enzymes described herein (clockwise, SEQ ID NOs: 3666, 3667, 3672, 3671).

[0066] FIGs. 17A and 17B depict an agarose gel showing the results of PAM vector library cleavage in the presence of TXTL extracts containing various MG family nucleases and their corresponding tracrRNA or sgRNAs (as described in Example 12). **FIG. 17A** shows lane 1: ladder. The bands are, from top to bottom, 766, 500, 350, 300, 350, 200, 150, 100, 75, 50; lane 2: 28-1 +MGcrRNA spacer1 (SEQ ID NOs: 141 + 3860); lane 3: 29-1 +MGcrRNA spacer1 (SEQ ID NOs: 215 + 3860); lane 4: 30-1 +MGcrRNA spacer1 (SEQ ID NOs:226 + 3860); lane 5: 31-1 +MGcrRNA spacer1(SEQ ID NOs: 229 + 3860); lane 6: 32-1 +MGcrRNA spacer1 (SEQ ID NOs: 261 + 3860); lane 7: ladder. **FIG. 17B** shows lane 1: ladder; lane 2: LbaCas12a + LbaCas12a crRNA spacer2; lane 3: LbaCas12a + MGcrRNA spacer2; lane 4: Apo 13-1; lane 5: 28-1 +MGcrRNA spacer2 (SEQ ID NOs: 141 + 3861); lane 6: 29-1 + MGcrRNA spacer2 (SEQ ID NOs: 215 + 3861); lane 7: 30-1 + MGcrRNA spacer2 (SEQ ID NOs: 226 + 3861); lane 8: 31-1 + MGcrRNA spacer2 (SEQ ID NOs: 229 + 3861); lane 9: 32-1 + MGcrRNA spacer2 (SEQ ID NOs: 261 +3861)

[0067] FIGs. 18A, 18B, 18C, 18D, and 18E provide data showing that type V-A effectors described herein are active nucleases. **FIG. 18A** depicts seqLogo representations of PAM sequences determined for three nucleases described herein. **FIG.18B** shows a boxplot of plasmid transfection activity assays inferred from frequency of indel edits for active nucleases. The boundaries of the boxplots indicate first and third quartile values. The mean is indicated with an "x" and the median is represented by the midline within each box. **FIG. 18C** shows plasmid transfection editing frequencies at four target sites for MG29-1 and AsCas12a. One side-by-side experiment with AsCas12a was done. **FIG.18D** shows plasmid and RNP editing activity for nuclease MG29-1 at 14 target loci with either TTN or CCN PAMs. **FIG. 18E** shows the editing profile of nuclease MG29-1 from RNP transfection assays. One side-by-side experiment with AsCas12a was done. Editing frequency and profile experiments for MG29-1 were done in

duplicate. The bar plots **FIG. 18C** and **FIG.18D** show mean editing frequency with one standard deviation error bar.

[0068] FIG. 19 depicts in cell indel formation generated by transfection of HEK cells with MG29-1 constructs described in Example 12 alongside their corresponding sgRNAs containing various different targeting sequences targeting various locations in the human genome.

[0069] FIG. 20 depicts seqLogo representations of PAM sequences of specific MG family enzymes derived via NGS as described herein (as described in Example 13).

[0070] FIG. 21 depict seqLogo representations of PAM sequences of specific MG family enzymes derived via NGS as described herein (top to bottom, SEQ ID NOs: 3865, 3867, 3872).

[0071] FIG. 22 depict seqLogo representations of PAM sequences derived via NGS as described herein (top to bottom, SEQ ID NOs: ~~3878,~~ 3879, 3880, 3881).

[0072] FIG. 23 depict seqLogo representations of PAM sequences derived via NGS as described herein (top to bottom, SEQ ID NOs: 3883, 3884, 3885) .

[0073] FIG. 24 depict seqLogo representations of PAM sequences derived via NGS as described herein (SEQ ID NO: 3882).

[0074] FIG. 25 depicts in cell indel formation generated by transfection of HEK cells with MG31-1 constructs described in Example 14 alongside their corresponding sgRNAs containing various different targeting sequences targeting various locations in the human genome.

[0075] FIGs. 4A, 26B, and 26C shows the biochemical characterization of Type V-A nucleases. **FIG. 26A** shows PCR of cleavage products with adaptors ligated to their ends shows activity of nucleases described herein and Cpf1 (positive control) when bound to a universal crRNA. Expected cleavage product band labeled with an arrow. **FIGs. 4B** shows PCR of cleavage products with adaptors ligated to their ends show activity of nucleases described herein when bound to their native crRNA . Cleavage product band indicated with an arrow. **FIG. 26C** shows analysis of the NGS cut sites shows cleavage on the target strand at position 22, sometimes with less frequent cleavage after 21 or 23 nt.

[0076] FIGs. 27A and 27B depict multiple sequence alignments of Type V-L nucleases described herein, showing (**FIG. 27A**) an example locus organization for a Type V-L nuclease, and (**FIG. 27B**) a multiple sequence alignment. Regions containing putative RuvC-III domains are shown as light grey rectangles. Putative RuvC catalytic residues are shown as small dark grey rectangles above each sequence. Putative single-guide RNA binding sequences are small white rectangles, putative scissile phosphate binding sites are indicated by black rectangles above sequences, and residues predicted to disrupt base stacking near the scissile phosphate in the target sequence are indicated by small medium-grey rectangles above sequences.

[0077] FIG. 28 shows a Type V-L candidate labeled MG60 as an example locus organization alongside an effector repeat structure and a phylogenetic tree showing the location of the enzyme in the Type V families.

[0078] FIG. 29 shows examples of smaller Type V effectors one of which may be labeled as MG70.

[0079] FIG. 30 shows characteristic information of MG70 as described herein. Depicted is an example locus organization alongside a phylogenetic tree illustrating the location of these enzymes in the Type V family.

[0080] FIG. 31 shows another example of a small Type V effector MG81 as described herein. Depicted is an example locus organization alongside a phylogenetic tree illustrating the location of these enzymes in the Type V family.

[0081] FIG. 32 shows that the activity individual enzymes of Type V effector families identified herein (e.g. MG20, MG60, MG70, other) is maintained over a variety of different enzyme lengths (e.g. 400-1200 AA). Light dots (True) indicate active enzymes while dark dots (unknown) indicate untested enzymes.

[0082] FIG. 33 depicts sequence conservation of MG nucleases described herein. The black bars indicate putative RuvC catalytic residues.

[0083] FIG. 34 and **FIG. 35** depict an enlarged version of multiple sequence alignments in **FIG. 33** of regions of the MG nucleases described herein containing putative RuvC catalytic residues (dark-grey rectangles), scissile phosphate-binding residues (black rectangles), and residues predicted to disrupt base stacking adjacent to the scissile phosphate (light-grey rectangles).

[0084] FIG. 36 depicts the regions of the MG nucleases described herein containing putative RuvC-III domain & catalytic residues.

[0085] FIG. 37 depicts regions of the MG nucleases containing putative single-guide RNA-binding residues (white rectangles above sequences).

[0086] FIG. 38 depicts multiple protein sequence alignment of representatives from several MG type V Families. Shown are conserved regions containing portions of the RuvC domain predicted to be involved in nuclease activity. Predicted catalytic residues are highlighted.

[0087] FIG. 39 shows a screen of the TRAC locus for MG29-1 gene editing. A bar graph shows indel creation resulting from transfection of MG29-1 with 54 separate guide RNAs targeting the TRAC locus in primary human T cells. The corresponding guide RNAs depicted in the figure are identified in SEQ ID NOs: 4316-4423.

[0088] FIG. 40 depicts the optimization of MG29-1 editing at TRAC. A bar graph shows indel creation resulting from transfection of MG29-1 (at the indicated concentrations) with the four

best 22 nt guide RNAs from **FIG. 39** (9, 19, 25, and 35). Legend: MG29-1 9 is MG29-1 effector (SEQ ID NO: 215) and Guide 9 (SEQ ID NO: 4378), MG29-1 19 is MG29-1 effector (SEQ ID NO: 215) and Guide 19 (SEQ ID NO: 4388), MG29-1 25 is MG29-1 effector (SEQ ID NO: 215) and Guide 25 (SEQ ID NO: 4394), and MG29-1 35 is MG29-1 effector (SEQ ID NO: 215) and Guide 35 (SEQ ID NO: 4404).

[0089] FIG. 41 depicts the optimization of dose and guide length for MG29-1 editing at TRAC. Line graphs show the indel creation resulting from transfection of MG29-1 and either guide RNA #19 (SEQ ID NO: 4388) or guide RNA #35 (SEQ ID NO: 4404). Three different doses of nuclease/guide RNA were used. For each dose, six different guide lengths were tested, successive one-nucleotide 3' truncations of SEQ ID NOs: 4388 and 4404. The guides used in **FIG. 39** and **FIG. 40** are the 22nt-long spacer-containing guides in this case.

[0090] FIG. 42 shows a correlation of indel generation at TRAC and loss of the T cell receptor expression in the Experiment of Example 22.

[0091] FIG. 43 depicts targeted transgene integration at TRAC stimulated by MG29-1 cleavage. Cells receiving transgene donor alone by AAV infection retain TCR expression and lack CAR expression; cells transfected with MG29-1 RNPs and infected with 100,000 vg (vector genomes) of a CAR transgene donor lose TCR expression and gain CAR expression. Shown are FACS plots of CAR antigen binding vs TCR expression for cells transfected with AAV alone containing the CAR-T-containing donor sequence ("AAV"); AAV containing the CAR-T-containing donor sequence with MG29-1 enzyme and sgRNA 19 (SEQ ID NO: 4388) ("AAV+MG29-1-19-22" comprising a 22 nucleotide spacer), or AAV containing the CAR-T-containing donor sequence with MG29-1 enzyme and sgRNA 35 (SEQ ID NO: 4404) ("AAV+MG29-1-35-22" comprising a 22 nucleotide spacer).

[0092] FIG. 44 shows MG29-1 gene editing at TRAC in hematopoietic stem cells. A bar graph shows the extent of indel creation at TRAC after transfection with MG29-1-9-22 ("MG29-1 9"; MG29-1 plus guide RNA #19) and MG29-1-35-22 ("MG29-1 35"; MG29-1 plus guide RNA #35) compared to mock-transfected cells.

[0093] FIG. 45 shows the refinement of the MG29-1 PAM based on analysis of gene editing outcomes in cells. Guide RNAs were designed using a 5'-NTTN-3' PAM sequence and then sorted according to the gene editing activity observed. The identity of the underlined base (the 5'-proximal N) is shown for each bin. All of the guides with activity greater than 10% had a T at this position in the genomic DNA indicating that the MG29-1 PAM may be best described as 5'-TTTN-3'. The statistical significance of the over-representation of T at this position is shown for each bin.

[0094] FIG. 46 depicts the analysis of gene editing activity versus the base composition of MG29-1 spacer sequences. A bar graph shows experimental data illustrating a relationship between GC content (%) and indel frequency ("high" signifies >50% indels (N = 4); "medium" signifies 10-50% indels (N = 15); ">1%" signifies 1-5% indels (N = 12); "<1%" signifies less than 1% indels (N = 82)).

[0095] FIG. 47 depicts MG29-1 guide RNA chemical modifications. The bar graph shows the consequences of modifications from Table 7 on VEGF-A editing activity relative to an unmodified guide RNA (sample #1).

[0096] FIG. 48 depicts a dose titration of a variously chemically modified MG29-1 RNA. The bar graphs show indel generation after transfection of RNPs with guides using modification patterns 1, 4, 5, 7, and 8. RNPs doses were 126 pmol MG29-1 and 160 pmol guide RNA or as indicated. Full dose (A), 1/4th (B), 1/8th (C), 1/16th (D), and 1/32nd (E).

[0097] FIG. 49 depicts a plasmid map of pMG450 (MG29-1 nuclease protein in lac inducible tac promoter *E. coli* BL21 expression vector.

[0098] FIG. 50 depicts the indel profile of MG29-1with spacer mALb29-1-8 (SEQ ID NO: 3999) compared to spCas9 with a guide targeting mouse albumin intron 1.

[0099] FIG. 51 is a representative indel profile of MG29-1 with a guide targeting mouse albumin intron 1 determined by next generation sequencing (approximately 15,000 total reads analyzed) as in Example 29.

[00100] FIG. 52 shows the editing efficiency of MG29-1 compared to spCas9 in mouse liver cell line Hepa1-6 nucleofected with RNP as in Example 29.

[00101] FIGs. 53A, 53B, 53C, and 53D show the editing efficiencies in mammalian cells of MG29-1 variants with single and double amino acid substitutions compared to wild type MG29-1. **FIG. 53A** depicts editing efficiency in Hepa 1-6 cells transfected with plasmids codifying for MG29-1 WT or mutant versions. **FIG. 53B** depicts Editing efficiency in Hepa 1-6 cells transfected with mRNA encoding WT or S168R at various concentrations. **FIG. 53C** depicts the editing efficiency in Hepa 1-6 cells transfected with mRNA codifying versions of MG29-1 with single or double amino acid substitutions. **FIG. 53D** depicts the editing efficiency in Hepa 1-6 and HEK293T cells transfected with MG29-1 WT vs S168R in combination with 13 guides. 12 guides correspond to guides in Table 7. Guide "35 (TRAC)" is a guide targeting the human locus TRAC.

[00102] FIG. 54 shows the predicted secondary structure of the MG29-1 guide mAlb29-1-8.

[00103] FIG. 55 shows the impact of chemical modifications of the MG29-1 sgRNA sequence upon the stability of the sgRNA in whole cell extracts of mammalian cells.

[00104] FIGs. 56A, 56B, and 56C show the use of sequencing to identify the cut site on the target strand in an *in vitro* reaction performed with MG29-1 protein, a guide RNA, and an appropriate template. **FIG. 56A** shows the distance of the cut position from the PAM in nucleotides as determined by next generation sequencing. **FIG. 56B** shows the use of Sanger Sequencing to define the MG29-1 cut site on the target strand. **FIG. 56C** shows the use of Sanger Sequencing to define the MG29-1 cut site on the non-target strand. Run-off Sanger sequencing was performed on *in vitro* reactions containing MG29-1, a guide, and an appropriate template to evaluate the cleavage of both strands. The cleavage site on the target strand is position 23 which is consistent with the NGS data in FIG. 56A which shows cleavage at 21-23 bases. The "A" peak at the end of the sequence is due to polymerase run off and is expected. The cleavage site on the non-target strand can be seen in the reverse read in which the expected terminating base is "T". The marked spot (line) shows cleavage at position 17 from the PAM and then the terminal T. However, there is a mixed T signal at positions 18, 19, and 20 from the PAM suggesting variable cleavage on this strand at positions 17, 18, and 19.

[00105] FIG. 57 depicts the gene editing outcomes at the DNA level for CD38. S. pyogenes (Spy) Cas9 guides for CD38 and TRAC are shown at right.

[00106] FIG. 58 depicts the gene editing outcomes at the phenotypic level for CD38.

[00107] FIG. 59 depicts the gene editing outcomes at the DNA level for TIGIT.

[00108] FIG. 60 depicts the gene editing outcomes at the DNA level for AAVS1.

[00109] FIG. 61 depicts the gene editing outcomes at the DNA level for B2M.

[00110] FIG. 62 depicts the gene editing outcomes at the DNA level for CD2.

[00111] FIG. 63 depicts the gene editing outcomes at the DNA level for CD5.

[00112] FIG. 64A depicts the gene editing outcomes at the DNA level for mouse TRAC. **FIG. 64B** depicts the flow cytometry results for gene editing of mouse TRAC.

[00113] FIG. 65 depicts the percentage of TRAC knock-out versus the percentage of indels.

[00114] FIG. 66A depicts the gene editing outcomes at the DNA level for mouse TRBC1. **FIG. 66B** depicts the gene editing outcomes at the DNA level for mouse TRBC2. **FIG. 66C** depicts the flow cytometry results for gene editing of human TRBC1/2.

[00115] FIG. 67 depicts the gene editing outcomes at the DNA level for HPRT.

[00116] FIG. 68 depicts the activity of chemically modified guides in Hepa1-6 cells when delivered as mRNA and gRNA using lipofectamine Messenger Max.

[00117] FIG. 69 depicts the stability of guides modified with modification 44 versus end-modified or unmodified guides.

[00118] FIGs. 70A and 70B depict a comparison of guide stability for Type II and Type V systems. **FIG. 70A** depicts stability data for unmodified guides. **FIG. 70B** depicts stability data for guides with 5' and 3' end modifications.

[00119] FIG. 71 depicts the predicted secondary structures of MG29-1 (Type V) and MG3-6/3-4 (Type II) guide RNA. The backbone (tracr) portion is shown.

[00120] FIG. 72 depicts the stability of guide mAlb298-34 compared to mAlb298-37 in cell lysates from Hepa1-6 cells.

[00121] FIG. 73 depicts the editing efficiency of MG29-1 in mouse liver following *in vivo* delivery.

[00122] FIG. 74 depicts analysis of gene-editing outcomes by NGS for mRNA electroporation in T cells.

[00123] FIG. 75 depicts analysis of gene-editing outcomes by NGS for chemically modified guides.

[00124] FIG. 76 depicts ELISA results from a screen performed at a serum dilution of 1:50 to detect antibodies against MG29-1 (n = 50). Tetanus toxoid was used as the positive control due to wide-spread vaccination against this antigen. Serum samples above the dashed line were considered antibody-positive; the line represents the mean absorbance of the negative control (human albumin) plus two standard deviations from the mean. *P<0.05, **P<0.01, ****P<0.0001 as determined by an unpaired Student's *t*-test; ns, not significant.

[00125] FIG. 77 depicts HAO-1 editing efficiency in mouse liver as measured by NGS. Each point represents an individual mouse.

[00126] FIGs. 78A-B depict the effects of HAO-1 editing on glycolate oxidase (GO) protein levels in mouse liver as evaluated by Western Blot. 10 μg of total protein was loaded for each sample.

[00127] FIG. 79 depicts Western Blot analysis of glycolate oxidase (GO) protein levels in an untreated mouse compared to two individual mice treated with lipid nanoparticles (LNPs) encapsulating MG29-1 mRNA and either guide mH29-1_37 or mH29-5_37. Three different amounts of total liver protein (40 μg, 20 μg, and 10 μg) from each mouse were loaded on the gel then processed for detection of the mouse glycolate oxidase protein.

[00128] FIG. 80 depicts an example INDEL profile for MG29-1 and an sgRNA targeting the HAO-1 gene in mouse liver. The sample was taken from mouse #17 (treated with a lipid nanoparticle encapsulating mH29-29_37 and MG29-1 WT mRNA).

[00129] FIG. 81 depicts the gene editing outcomes at the DNA level for TRAC in human peripheral blood B cells.

[00130] FIG. 82 depicts the gene editing outcomes at the DNA level for TRAC in hematopoietic stem cells.

[00131] FIG. 83 depicts the gene editing outcomes at the DNA level for TRAC in induced pluripotent stem cells (iPSCs).

[00132] FIG. 84 depicts the results of *in vivo* genome editing with MG29-1 as quantified by next generation sequencing (NGS).

[00133] FIG. 85 depicts an example INDEL profile generated by the MG29-1 nuclease and guide 298-37 as measured by next generation sequencing (NGS).

[00134] FIG. 86 depicts spacer length optimization for MG29-1 guides targeting two loci.

[00135] FIG. 87 depicts *in vitro* stability of sgRNAs for MG29-1 and MG3-6/3-4.

[00136] FIG. 88 depicts the predicted secondary structures of the backbone parts of the guide RNA for MG29-1 and MG3-6/3-4.

[00137] FIG. 89 depicts the predicted secondary structure of an MG3-6/3-4 guide with a spacer targeting mouse albumin.

[00138] FIG. 90 depicts the predicted secondary structure of an MG29-1 guide with stem-loop 1 from MG3-6/3-4 added to the 5' end.

[00139] FIG. 91 depicts the editing efficiency of MG29-1 with mouse albumin guide 8 with chemistries 44 or 50 in Hepa1-6 cells by mRNA transfection or RNP nucleofection.

[00140] FIG. 92 depicts editing in the liver of mice after dosing with LNP encapsulating MG29-1 mRNA and one of four different guide RNAs.

[00141] FIG. 93 depicts the predicted secondary structure of the RNA molecule mAlb29-g8-37-array.

[00142] FIG. 94 depicts a plot showing editing efficiency in the whole liver of mice at 5 days after intravenous injection of LNP encapsulating one of: (1) MG29-1 mRNA and guide mAlb29-8-50 (mA29-8-50) at three different doses; (2) spCas9 mRNA and guide mAlbR2 at three different doses; or (3) PBS buffer (Control). Each circle represents a single mouse and the bars indicate the mean and standard deviation.

[00143] FIG. 95 depicts editing activity in Hep3B cells transfected with MG29-1 mRNA and 6 sgRNA targeting human HAO-1.

[00144] FIG. 96 depicts editing Activity of 4 MG29-1 sgRNA targeting human HAO-1 in HuH7 and Hep3B cells transfected with Ribonuclear Protein Complexes.

[00145] FIG. 97 depicts editing activity of MG29-1 with sgRNA targeting the human HAO-1 gene in primary human hepatocytes.

[00146] FIG. 98A depicts a representative indel profile for MG29-1 sgRNAs hH29-4-37 and hH29-21-37 in Primary Human Hepatocytes. **FIG. 98B** depicts a representative indel profile for MG29-1 sgRNAs hH29-23-37 and hH29-41-37 in Primary Human Hepatocytes.

[00147] FIG. 99 depicts the activity of MG29-1 guide RNAs with 22 nucleotide or 20 nucleotide spacers targeting mouse HAO-1 in mouse liver.

[00148] FIG. 100 depicts the impact of the mRNA/guide RNA ratio and separate or co-formulation on editing efficiency in mouse liver.

[00149] FIG. 101 depicts the evaluation of MG29-1 guide chemistries on editing activity in the liver of mice after *in vivo* delivery in LNP.

[00150] FIG. 102 depicts the gene-editing outcomes at the DNA level for APO-A1 in Hepa1-6 cells.

[00151] FIG. 103 depicts the gene-editing outcomes at the DNA level for ANGPTL3 in Hepa1-6 cells.

[00152] FIG. 104A-E depicts *in vitro* characterization of MG55-43. **FIG. 104A** shows the genomic region in the vicinity of the MG55-43 nuclease. Genes are represented by orange arrows. The gene encoding the candidate nuclease includes a "putative transposase DNA-binding domain". The CRISPR array is represented by repeats and spacers. The predicted tracrRNA is shown as an arrow between the array and the nuclease and labeled "Predicted-trimmed-TracrRNA-CM2". **FIG. 104B** shows active single guide RNA design (tracrRNA and repeat sequences connected by a tetraloop). The color of the nitrogenated bases corresponds to the probability of base pairing of that base, where red is high probability and blue is low probability. **FIG. 104C** shows an agarose gel showing *in vitro* cleavage of plasmid target DNA library with the sgRNA and two different spacers (U67 and U40). Lanes that are not related to the MG55-43 nuclease are not shown. **FIG. 104D** shows sequence logo showing the MG55-43 PAM sequence. **FIG. 104E** shows a histogram showing the cut site position in the spacer sequence tested with MG55-43.

[00153] FIG. 105A-C depicts examples of genomic regions encoding MG91 nucleases. Genes are represented by arrows with genes encoding candidate nuclease labeled as such. The CRISPR array is represented by repeats and spacers. Intergenic regions potentially encoding active tracrRNAs are highlighted as bars labeled IG # or Intergenic region #.

[00154] FIG. 106 depicts multiple sequence alignments of intergenic region nucleotide sequences potentially containing tracrRNAs. Green bars on top indicate a high degree of similarity among the sequence of the intergenic regions. **FIG. 106A** shows intergenic region 2 in the vicinity of the MG91-15 nuclease and its relatives. **FIG. 106B** shows intergenic region 2 in

the vicinity of the MG91-32 nuclease and its relatives. **FIG. 106C** shows intergenic region 2 in the vicinity of the MG91-87 nuclease and its relatives.

[00155] FIG. 107A-D depicts single guide RNA designs and *in vitro* cleavage assay results. For the single guide RNA designs (tracrRNA and repeat sequences), the color of the bases corresponds to the probability of base pairing of that base, where red is high probability and blue is low probability. **FIG. 107A** depicts MG91-15 sgRNA1, **FIG. 107B** depicts MG91-32 sgRNA1, and **FIG. 107C** depicts MG91-87 sgRNA1. **FIG. 107D** depicts an agarose gel showing *in vitro* cleavage of plasmid target DNA library with different sgRNA designs (sgRNA1 and sgRNA2) and two different spacers (U67 and U40). Lanes that are not related to these nucleases are not shown.

[00156] FIG. 108A-F depicts sequence logos of predicted PAM sequence and histograms showing cut site position. **FIGs. 108A, 108B,** and **108C** depict sequence logos showing the PAM sequences of MG91-15, MG91-32, and MG91-87, respectively. **FIGs. 108D, 108E,** and **108F** depict histograms showing the cut site positions in the spacer sequences tested with MG91-15, MG91-32, and MG91-87, respectively.

[00157] FIG. 109 depicts structures of example cationic lipids that can be used in lipid nanoparticles described herein.

BRIEF DESCRIPTION OF THE SEQUENCE LISTING

[00158] The Sequence Listing filed herewith provides exemplary polynucleotide and polypeptide sequences for use in methods, compositions, and systems according to the disclosure. Below are exemplary descriptions of sequences therein.

MG11

[00159] SEQ ID NOs: 1-37 show the full-length peptide sequences of MG11 nucleases.

[00160] SEQ ID NO: 3471 shows a crRNA 5' direct repeats designed to function with an MG11 nuclease.

[00161] SEQ ID NOs: 3472-3538 show effector repeat motifs of MG11 nucleases.

[00162] SEQ ID NOs: 38-118 show the full-length peptide sequences of MG13 nucleases.

[00163] SEQ ID NO: 3540-3550 show effector repeat motifs of MG13 nucleases.

MG19

[00164] SEQ ID NOs: 119-124 show the full-length peptide sequences of MG19 nucleases.

[00165] SEQ ID NOs: 3551-3558 show the nucleotide sequences of sgRNAs engineered to function with a MG19 nuclease.

[00166] SEQ ID NOs: 3863-3866 show PAM sequences compatible with MG19 nucleases.

MG20

[00167] SEQ ID NO: 125 shows the full-length peptide sequence of a MG20 nuclease.

[00168] SEQ ID NO: 3559 shows the nucleotide sequence of a sgRNA engineered to function with a MG20 nuclease.

[00169] SEQ ID NO: 3867 shows a PAM sequence compatible with an MG20 nuclease.

MG26

[00170] SEQ ID NOs: 126-140 show the full-length peptide sequences of MG26 nucleases.

[00171] SEQ ID NOs: 3560-3572 show effector repeat motifs of MG26 nucleases.

MG28

[00172] SEQ ID NOs: 141-214 show the full-length peptide sequences of MG28 nucleases.

[00173] SEQ ID NOs: 3573-3607 show effector repeat motifs of MG28 nucleases.

[00174] SEQ ID NOs: 3608-3609 show crRNA 5′ direct repeats designed to function with an MG28 nuclease.

[00175] SEQ ID NOs: 3868-3869 shows a PAM sequence compatible with an MG28 nuclease.

MG29

[00176] SEQ ID NOs: 215-225 show the full-length peptide sequences of MG29 nucleases.

[00177] SEQ ID NO: 5680 shows the nucleotide sequence of an MG29-1 nuclease containing 5' UTR, NLS, CDS, NLS, 3' UTR, and polyA tail.

[00178] SEQ ID NOs: 3610-3611 show effector repeat motifs of MG29 nucleases.

[00179] SEQ ID NO: 3612 shows the nucleotide sequence of a sgRNA engineered to function with a MG29 nuclease.

[00180] SEQ ID NOs: 3870-3872 show PAM sequences compatible with an MG29 nuclease.

[00181] SEQ ID NO: 5687 shows an MG29-1 coding sequence used for the generation of mRNA.

[00182] SEQ ID NOs: 5830 and 5846 show DNA sequences encoding MG29-1 mRNAs.

MG30

[00183] SEQ ID NOs: 226-228 show the full-length peptide sequences of MG30 nucleases.

[00184] SEQ ID NOs: 3613-3615 show effector repeat motifs of MG30 nucleases.

[00185] SEQ ID NO: 3873 shows a PAM sequence compatible with an MG30 nuclease.

MG31

[00186] SEQ ID NOs: 229-260 show the full-length peptide sequences of MG31 nucleases.

[00187] SEQ ID NOs: 3616-3632 show effector repeat motifs of MG31 nucleases.

[00188] SEQ ID NOs: 3874-3876 show PAM sequences compatible with a MG31 nuclease.

MG32

[00189] SEQ ID NO: 261 shows the full-length peptide sequence of a MG32 nuclease.

[00190] SEQ ID NO: 3633-3634 show effector repeat motifs of MG32 nucleases.

[00191] SEQ ID NO: 3876 shows a PAM sequence compatible with a MG32 nuclease.

MG37

[00192] SEQ ID NOs: 262-426 show the full-length peptide sequences of MG37 nucleases.

[00193] SEQ ID NO: 3635 shows an effector repeat motif of MG37 nucleases.

[00194] SEQ ID NOs: 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, and 3660-3661 show the nucleotide sequence of sgRNA engineered to function with an MG37 nuclease.

[00195] SEQ ID NOs: 3638, 3642, 3646, 3650, 3654, 3658, and 3662 show the nucleotide sequences of MG37 tracrRNAs derived from the same loci as MG37 nucleases above.

[00196] SEQ ID NO: 3639, 3643, 3647, 3651, 3655, and 3659 show 5′ direct repeat sequences derived from native MG37 loci that serve as crRNAs when placed 5′ to a 3′ targeting or spacer sequence.

MG53

[00197] SEQ ID NOs: 427-428 show the full-length peptide sequences of MG53 nucleases.

[00198] SEQ ID NO: 3663 shows a 5′ direct repeat sequence derived from native MG53 loci that serve as a crRNA when placed 5′ to a 3′ targeting or spacer sequence.

[00199] SEQ ID NOs: 3664-3667 show the nucleotide sequence of sgRNAs engineered to function with an MG53 nuclease.

[00200] SEQ ID NOs: 3668-3669 show the nucleotide sequences of MG53 tracrRNAs derived from the same loci as MG53 nucleases above.

MG54

[00201] SEQ ID NOs: 429-430 show the full-length peptide sequences of MG54 nucleases.

[00202] SEQ ID NO: 3670 shows a 5′ direct repeat sequence derived from native MG54 loci that serve as a crRNA when placed 5′ to a 3′ targeting or spacer sequence.

[00203] SEQ ID NOs: 3671-3672 show the nucleotide sequence of sgRNA engineered to function with an MG54 nuclease.

[00204] SEQ ID NOs: 3673-3676 show the nucleotide sequences of MG54 tracrRNAs derived from the same loci as MG54 nucleases above.

MG55

[00205] SEQ ID NOs: 431-688 show the full-length peptide sequences of MG55 nucleases.

[00206] SEQ ID NO: 6031 shows the nucleotide sequence of an sgRNA engineered to function with an MG55 nuclease.

[00207] SEQ ID NO: 6032 shows a PAM sequence compatible with an MG55 nuclease.

MG56

[00208] SEQ ID NOs: 689-690 show the full-length peptide sequences of MG56 nucleases.

[00209] SEQ ID NO: 3678 shows a crRNA 5' direct repeats designed to function with an MG56 nuclease.

[00210] SEQ ID NOs: 3679-3680 show effector repeat motifs of MG56 nucleases.

MG57

[00211] SEQ ID NOs: 691-721 show the full-length peptide sequences of MG57 nucleases.

[00212] SEQ ID NOs: 3681-3694 show effector repeat motifs of MG57 nucleases.

[00213] SEQ ID NOs: 3695-3696 show the nucleotide sequences of sgRNAs engineered to function with an MG57 nuclease.

[00214] SEQ ID NOs: 3879-3880 shows PAM sequences compatible with MG57 nucleases.

MG58

[00215] SEQ ID NOs: 722-779 show the full-length peptide sequences of MG58 nucleases.

[00216] SEQ ID NOs: 3697-3711 show effector repeat motifs of MG58 nucleases.

MG59

[00217] SEQ ID NOs: 780-792 show the full-length peptide sequences of MG59 nucleases.

[00218] SEQ ID NOs: 3712-3728 show effector repeat motifs of MG59 nucleases.

[00219] SEQ ID NOs: 3729-3730 show the nucleotide sequences of sgRNAs engineered to function with an MG59 nuclease.

[00220] SEQ ID NOs: 3881-3882 shows PAM sequences compatible with MG59 nucleases.

MG60

[00221] SEQ ID NOs: 793-1163 show the full-length peptide sequences of MG60 nucleases.

[00222] SEQ ID NOs: 3731-3733 show effector repeat motifs of MG60 nucleases.

MG61

[00223] SEQ ID NOs: 1164-1469 show the full-length peptide sequences of MG61 nucleases.

[00224] SEQ ID NOs: 3734-3735 show crRNA 5' direct repeats designed to function with MG61 nucleases.

[00225] SEQ ID NOs: 3736-3847 show effector repeat motifs of MG61 nucleases.

MG62

[00226] SEQ ID NOs: 1470-1472 show the full-length peptide sequences of MG62 nucleases.

[00227] SEQ ID NOs: 3848-3850 show effector repeat motifs of MG62 nucleases.

MG70

[00228] SEQ ID NOs: 1473-1514 show the full-length peptide sequences of MG70 nucleases.

MG75

[00229] SEQ ID NOs: 1515-1710 show the full-length peptide sequences of MG75 nucleases.

MG77

[00230] SEQ ID NOs: 1711-1712 show the full-length peptide sequences of MG77 nucleases.

[00231] SEQ ID NOs: 3851-3852 show the nucleotide sequences of sgRNAs engineered to function with an MG77 nuclease.

[00232] SEQ ID NOs: 3883-3884 show PAM sequences compatible with MG77 nucleases.

MG78

[00233] SEQ ID NOs: 1713-1717 show the full-length peptide sequences of MG78 nucleases.

[00234] SEQ ID NO: 3853 shows the nucleotide sequence of a sgRNA engineered to function with an MG78 nuclease.

[00235] SEQ ID NO: 3885 shows a PAM sequence compatible with a MG78 nuclease.

MG79

[00236] SEQ ID NOs: 1718-1722 show the full-length peptide sequences of MG79 nucleases.

[00237] SEQ ID NOs: 3854-3857 shows the nucleotide sequences of sgRNAs engineered to function with an MG79 nuclease.

[00238] SEQ ID NOs: 3886-3889 show the PAM sequences compatible with MG79 nucleases.

MG80

[00239] SEQ ID NO: 1723 shows the full-length peptide sequence of a MG80 nuclease.

MG81

[00240] SEQ ID NOs: 1724-2654 show the full-length peptide sequences of MG81 nucleases.

MG82

[00241] SEQ ID NOs: 2655-2657 show the full-length peptide sequences of MG82 nucleases.

MG83

[00242] SEQ ID NOs: 2658-2659 show the full-length peptide sequences of MG83 nucleases.

MG84

[00243] SEQ ID NOs: 2660-2677 show the full-length peptide sequences of MG84 nucleases.

MG85

[00244] SEQ ID NOs: 2678-2680 show the full-length peptide sequences of MG85 nucleases.

MG90

[00245] SEQ ID NOs: 2681-2809 show the full-length peptide sequences of MG90 nucleases.

MG91

[00246] SEQ ID NOs: 2810-3470 show the full-length peptide sequences of MG91 nucleases.

[00247] SEQ ID NOs: 6033-6036 show nucleotide sequences of sgRNAs engineered to function with MG91 nucleases.

[00248] SEQ ID NOs: 6037-6039 show PAM sequences compatible with MG91 nucleases.

[00249] SEQ ID NOs: 6040-6049 show MG91 intergenic regions potentially encoding tracrRNA.

[00250] SEQ ID NOs: 6050-6059 show MG91 CRISPR repeats.

Spacer segments

[00251] SEQ ID NOs: 3858-3861 show the nucleotide sequences of spacer segments.

NLS

[00252] SEQ ID NOs: 3938-3953 show the sequences of example nuclear localization sequences (NLSs) that can be appended to nucleases according to the disclosure.

CD38 Targeting

[00253] SEQ ID NOs: 4428-4465 and 5685 show the nucleotide sequences of sgRNAs engineered to function with an MG29-1 nuclease in order to target CD38.

[00254] SEQ ID NOs: 4466-4503 and 5686 show the DNA sequences of CD38 target sites.

TIGIT Targeting

[00255] SEQ ID NOs: 4504-4520 show the nucleotide sequences of sgRNAs engineered to function with an MG29-1 nuclease in order to target TIGIT.

[00256] SEQ ID NOs: 4521-4537 show the DNA sequences of TIGIT target sites.

AAVS1 Targeting

[00257] SEQ ID NOs: 4538-4568 show the nucleotide sequences of sgRNAs engineered to function with an MG29-1 nuclease in order to target AAVS1.

[00258] SEQ ID NOs: 4569-4599 show the DNA sequences of AAVS1 target sites.

B2M Targeting

[00259] SEQ ID NOs: 4600-4675 show the nucleotide sequences of sgRNAs engineered to function with an MG29-1 nuclease in order to target B2M.

[00260] SEQ ID NOs: 4676-4751 show the DNA sequences of B2M target sites.

CD2 Targeting

[00261] SEQ ID NOs: 4752-4836 show the nucleotide sequences of sgRNAs engineered to function with an MG29-1 nuclease in order to target CD2.

[00262] SEQ ID NOs: 4837-4921 show the DNA sequences of CD2 target sites.

CD5 Targeting

[00263] SEQ ID NOs: 4922-4945 show the nucleotide sequences of sgRNAs engineered to function with an MG29-1 nuclease in order to target CD5.

[00264] SEQ ID NOs: 4946-4969 show the DNA sequences of CD5 target sites.

hRosa26 Targeting

[00265] SEQ ID NOs: 4970-5012 show the nucleotide sequences of sgRNAs engineered to function with an MG29-1 nuclease in order to target hRosa26.

[00266] SEQ ID NOs: 5013-5055 show the DNA sequences of hRosa26 target sites.

TRAC Targeting

[00267] SEQ ID NOs: 5056-5125, 5681, and 5683 show the nucleotide sequences of sgRNAs engineered to function with an MG29-1 nuclease in order to target TRAC.

[00268] SEQ ID NOs: 5126-5195, 5682, and 5684 show the DNA sequences of TRAC target sites.

TRBC1 Targeting

[00269] SEQ ID NOs: 5196-5210 show the nucleotide sequences of sgRNAs engineered to function with an MG29-1 nuclease in order to target TRBC1.

[00270] SEQ ID NOs: 5211-5225 show the DNA sequences of TRBC1 target sites.

TRBC2 Targeting

[00271] SEQ ID NOs: 5226-5246 show the nucleotide sequences of sgRNAs engineered to function with an MG29-1 nuclease in order to target TRBC2.

[00272] SEQ ID NOs: 5247-5267 show the DNA sequences of TRBC2 target sites.

TRBC1/2 Targeting

[00273] SEQ ID NOs: 5642-5660 show the nucleotide sequences of sgRNAs engineered to function with an MG29-1 nuclease in order to target TRBC.

[00274] SEQ ID NOs: 5661-5679 show the DNA sequences of TRBC target sites.

FAS Targeting

[00275] SEQ ID NOs: 5268-5366 show the nucleotide sequences of sgRNAs engineered to function with an MG29-1 nuclease in order to target FAS.

[00276] SEQ ID NOs: 5367-5465 show the DNA sequences of FAS target sites.

PD-1 Targeting

[00277] SEQ ID NOs: 5466-5473 show the nucleotide sequences of sgRNAs engineered to function with an MG29-1 nuclease in order to target PD-1.

[00278] SEQ ID NOs: 5474-5481 show the DNA sequences of PD-1 target sites.

HPRT Targeting

[00279] SEQ ID NOs: 5482-5561show the nucleotide sequences of sgRNAs engineered to function with an MG29-1 nuclease in order to target HPRT.

[00280] SEQ ID NOs: 5562-5641 show the DNA sequences of HPRT target sites.

HAO-1 Targeting

[00281] SEQ ID NOs: 5788-5829 and 5831-5834 show the nucleotide sequences of sgRNAs engineered to function with an MG29-1 nuclease in order to target human HAO-1.

[00282] SEQ ID NOs: 5836-5845 show the nucleotide sequences of sgRNAs engineered to function with an MG29-1 nuclease in order to target mouse HAO-1.

APO-A1 Targeting

[00283] SEQ ID NOs: 5847-5860 show the nucleotide sequences of sgRNAs engineered to function with an MG29-1 nuclease in order to target mouse APO-A1.

[00284] SEQ ID NOs: 5861-5874 show the DNA sequences of APO-A1 target sites.

ANGPTL3 Targeting

[00285] SEQ ID NOs: 5875-5952 show the nucleotide sequences of sgRNAs engineered to function with an MG29-1 nuclease in order to target mouse ANGPTL3.

[00286] SEQ ID NOs: 5953-6030 show the DNA sequences of ANGPTL3 target sites.

DETAILED DESCRIPTION

[00287] While various embodiments of the invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions may occur to those skilled in the art without departing from the invention. It should be understood that various alternatives to the embodiments of the invention described herein may be employed.

[00288] The practice of some methods disclosed herein employ, unless otherwise indicated, techniques of immunology, biochemistry, chemistry, molecular biology, microbiology, cell biology, genomics, and recombinant DNA. See for example Sambrook and Green, Molecular Cloning: A Laboratory Manual, 4th Edition (2012); the series Current Protocols in Molecular Biology (F. M. Ausubel, et al. eds.); the series Methods In Enzymology (Academic Press, Inc.), PCR 2: A Practical Approach (M.J. MacPherson, B.D. Hames and G.R. Taylor eds. (1995)), Harlow and Lane, eds. (1988) Antibodies, A Laboratory Manual, and Culture of Animal Cells: A Manual of Basic Technique and Specialized Applications, 6th Edition (R.I. Freshney, ed. (2010)) (which is entirely incorporated by reference herein).

[00289] As used herein, the singular forms “a”, “an” and “the” are intended to include the plural forms as well, unless the context clearly indicates otherwise. Furthermore, to the extent that the terms “including”, “includes”, “having”, “has”, “with”, or variants thereof are used in either the detailed description and/or the claims, such terms are intended to be inclusive in a manner similar to the term “comprising”.

[00290] 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, i.e., the limitations of the measurement system. For example, "about" can mean within one or more than one standard deviation, per the practice in the art. Alternatively, "about" can mean a range of up to 20%, up to 15%, up to 10%, up to 5%, or up to 1% of a given value.

[00291] As used herein, a "cell" generally refers to a biological cell. A cell may be the basic structural, functional and/or biological unit of a living organism. A cell may originate from any organism having one or more cells. Some non-limiting examples include: a prokaryotic cell, eukaryotic cell, a bacterial cell, an archaeal cell, a cell of a single-cell eukaryotic organism, a protozoa cell, a cell from a plant (e.g., cells from plant crops, fruits, vegetables, grains, soy bean, corn, maize, wheat, seeds, tomatoes, rice, cassava, sugarcane, pumpkin, hay, potatoes, cotton, cannabis, tobacco, flowering plants, conifers, gymnosperms, ferns, clubmosses, hornworts, liverworts, mosses), an algal cell, (e.g., *Botryococcus braunii*, *Chlamydomonas reinhardtii*, *Nannochloropsis gaditana*, *Chlorella pyrenoidosa*, *Sargassum patens C. Agardh*, and the like), seaweeds (e.g., kelp), a fungal cell (e.g.,, a yeast cell, a cell from a mushroom), an animal cell, a cell from an invertebrate animal (e.g., fruit fly, cnidarian, echinoderm, nematode, etc.), a cell from a vertebrate animal (e.g., fish, amphibian, reptile, bird, mammal), a cell from a mammal (e.g., a pig, a cow, a goat, a sheep, a rodent, a rat, a mouse, a non-human primate, a human, etc.), and etcetera. Sometimes a cell is not originating from a natural organism (e.g., a cell can be a synthetically made, sometimes termed an artificial cell).

[00292] The term "nucleotide," as used herein, generally refers to a base-sugar-phosphate combination. A nucleotide may comprise a synthetic nucleotide. A nucleotide may comprise a synthetic nucleotide analog. Nucleotides may be monomeric units of a nucleic acid sequence (e.g., deoxyribonucleic acid (DNA) and ribonucleic acid (RNA)). The term nucleotide may include ribonucleoside triphosphates adenosine triphosphate (ATP), uridine triphosphate (UTP), cytosine triphosphate (CTP), guanosine triphosphate (GTP) and deoxyribonucleoside triphosphates such as dATP, dCTP, dITP, dUTP, dGTP, dTTP, or derivatives thereof. Such derivatives may include, for example, [αS]dATP, 7-deaza-dGTP and 7-deaza-dATP, and nucleotide derivatives that confer nuclease resistance on the nucleic acid molecule containing them. The term nucleotide as used herein may refer to dideoxyribonucleoside triphosphates (ddNTPs) and their derivatives. Illustrative examples of dideoxyribonucleoside triphosphates may include, but are not limited to, ddATP, ddCTP, ddGTP, ddITP, and ddTTP. A nucleotide may be unlabeled or detectably labeled, such as using moieties comprising optically detectable

moieties (e.g., fluorophores). Labeling may also be carried out with quantum dots. Detectable labels may include, for example, radioactive isotopes, fluorescent labels, chemiluminescent labels, bioluminescent labels, and enzyme labels. Fluorescent labels of nucleotides may include but are not limited fluorescein, 5-carboxyfluorescein (FAM), 2′7′-dimethoxy-4′5-dichloro-6-carboxyfluorescein (JOE), rhodamine, 6-carboxyrhodamine (R6G), N,N,N′,N′-tetramethyl-6-carboxyrhodamine (TAMRA), 6-carboxy-X-rhodamine (ROX), 4-(4′dimethylaminophenylazo) benzoic acid (DABCYL), Cascade Blue, Oregon Green, Texas Red, Cyanine and 5-(2′-aminoethyl)aminonaphthalene-1-sulfonic acid (EDANS). Specific examples of fluorescently labeled nucleotides can include [R6G]dUTP, [TAMRA]dUTP, [R110]dCTP, [R6G]dCTP, [TAMRA]dCTP, [JOE]ddATP, [R6G]ddATP, [FAM]ddCTP, [R110]ddCTP, [TAMRA]ddGTP, [ROX]ddTTP, [dR6G]ddATP, [dR110]ddCTP, [dTAMRA]ddGTP, and [dROX]ddTTP available from Perkin Elmer, Foster City, Calif; FluoroLink DeoxyNucleotides, FluoroLink Cy3-dCTP, FluoroLink Cy5-dCTP, FluoroLink Fluor X-dCTP, FluoroLink Cy3-dUTP, and FluoroLink Cy5-dUTP available from Amersham, Arlington Heights, Il.; Fluorescein-15-dATP, Fluorescein-12-dUTP, Tetramethyl-rodamine-6-dUTP, IR770-9-dATP, Fluorescein-12-ddUTP, Fluorescein-12-UTP, and Fluorescein-15-2′-dATP available from Boehringer Mannheim, Indianapolis, Ind.; and Chromosome Labeled Nucleotides, BODIPY-FL-14-UTP, BODIPY-FL-4-UTP, BODIPY-TMR-14-UTP, BODIPY-TMR-14-dUTP, BODIPY-TR-14-UTP, BODIPY-TR-14-dUTP, Cascade Blue-7-UTP, Cascade Blue-7-dUTP, fluorescein-12-UTP, fluorescein-12-dUTP, Oregon Green 488-5-dUTP, Rhodamine Green-5-UTP, Rhodamine Green-5-dUTP, tetramethylrhodamine-6-UTP, tetramethylrhodamine-6-dUTP, Texas Red-5-UTP, Texas Red-5-dUTP, and Texas Red-12-dUTP available from Molecular Probes, Eugene, Oreg. Nucleotides can also be labeled or marked by chemical modification. A chemically-modified single nucleotide can be biotin-dNTP. Some non-limiting examples of biotinylated dNTPs can include, biotin-dATP (e.g., bio-N6-ddATP, biotin-14-dATP), biotin-dCTP (e.g., biotin-11-dCTP, biotin-14-dCTP), and biotin-dUTP (e.g., biotin-11-dUTP, biotin-16-dUTP, biotin-20-dUTP).

[00293] The terms “polynucleotide,” “oligonucleotide,” and “nucleic acid” are used interchangeably to generally refer to a polymeric form of nucleotides of any length, either deoxyribonucleotides or ribonucleotides, or analogs thereof, either in single-, double-, or multi-stranded form. A polynucleotide may be exogenous or endogenous to a cell. A polynucleotide may exist in a cell-free environment. A polynucleotide may be a gene or fragment thereof. A polynucleotide may be DNA. A polynucleotide may be RNA. A polynucleotide may have any three-dimensional structure and may perform any function. A polynucleotide may comprise one or more analogs (e.g., altered backbone, sugar, or nucleobase). If present, modifications to the

nucleotide structure may be imparted before or after assembly of the polymer. Some non-limiting examples of analogs include: 5-bromouracil, peptide nucleic acid, xeno nucleic acid, morpholinos, locked nucleic acids, glycol nucleic acids, threose nucleic acids, dideoxynucleotides, cordycepin, 7-deaza-GTP, fluorophores (e.g., rhodamine or fluorescein linked to the sugar), thiol-containing nucleotides, biotin-linked nucleotides, fluorescent base analogs, CpG islands, methyl-7-guanosine, methylated nucleotides, inosine, thiouridine, pseudouridine, dihydrouridine, queuosine, and wyosine. Non-limiting examples of polynucleotides include coding or non-coding regions of a gene or gene fragment, loci (locus) defined from linkage analysis, exons, introns, messenger RNA (mRNA), transfer RNA (tRNA), ribosomal RNA (rRNA), short interfering RNA (siRNA), short-hairpin RNA (shRNA), micro-RNA (miRNA), ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, cell-free polynucleotides including cell-free DNA (cfDNA) and cell-free RNA (cfRNA), nucleic acid probes, and primers. The sequence of nucleotides may be interrupted by non-nucleotide components.

[00294] The terms "transfection" or "transfected" generally refer to introduction of a nucleic acid into a cell by non-viral or viral-based methods. The nucleic acid molecules may be gene sequences encoding complete proteins or functional portions thereof. See, e.g., Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 18.1-18.88 (which is entirely incorporated by reference herein).

[00295] The terms "peptide," "polypeptide," and "protein" are used interchangeably herein to generally refer to a polymer of at least two amino acid residues joined by peptide bond(s). This term does not connote a specific length of polymer, nor is it intended to imply or distinguish whether the peptide is produced using recombinant techniques, chemical or enzymatic synthesis, or is naturally occurring. The terms apply to naturally occurring amino acid polymers as well as amino acid polymers comprising at least one modified amino acid. In some cases, the polymer may be interrupted by non-amino acids. The terms include amino acid chains of any length, including full length proteins, and proteins with or without secondary and/or tertiary structure (e.g., domains). The terms also encompass an amino acid polymer that has been modified, for example, by disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, oxidation, and any other manipulation such as conjugation with a labeling component. The terms "amino acid" and "amino acids," as used herein, generally refer to natural and non-natural amino acids, including, but not limited to, modified amino acids and amino acid analogues. Modified amino acids may include natural amino acids and non-natural amino acids, which have been

chemically modified to include a group or a chemical moiety not naturally present on the amino acid. Amino acid analogues may refer to amino acid derivatives. The term "amino acid" includes both D-amino acids and L-amino acids.

[00296] As used herein, the "non-native" can generally refer to a nucleic acid or polypeptide sequence that is not found in a native nucleic acid or protein. Non-native may refer to affinity tags. Non-native may refer to fusions. Non-native may refer to a naturally occurring nucleic acid or polypeptide sequence that comprises mutations, insertions and/or deletions. A non-native sequence may exhibit and/or encode for an activity (e.g., enzymatic activity, methyltransferase activity, acetyltransferase activity, kinase activity, ubiquitinating activity, etc.) that may also be exhibited by the nucleic acid and/or polypeptide sequence to which the non-native sequence is fused. A non-native nucleic acid or polypeptide sequence may be linked to a naturally-occurring nucleic acid or polypeptide sequence (or a variant thereof) by genetic engineering to generate a chimeric nucleic acid and/or polypeptide sequence encoding a chimeric nucleic acid and/or polypeptide.

[00297] The term "promoter", as used herein, generally refers to the regulatory DNA region which controls transcription or expression of a gene and which may be located adjacent to or overlapping a nucleotide or region of nucleotides at which RNA transcription is initiated. A promoter may contain specific DNA sequences which bind protein factors, often referred to as transcription factors, which facilitate binding of RNA polymerase to the DNA leading to gene transcription. A 'basal promoter', also referred to as a 'core promoter', may generally refer to a promoter that contains all the basic elements to promote transcriptional expression of an operably linked polynucleotide. Eukaryotic basal promoters can contain a TATA-box and/or a CAAT box.

[00298] The term "expression", as used herein, generally refers to the process by which a nucleic acid sequence or a polynucleotide is transcribed from a DNA template (such as into mRNA or other RNA transcript) and/or the process by which a transcribed mRNA is subsequently translated into peptides, polypeptides, or proteins. Transcripts and encoded polypeptides may be collectively referred to as "gene product." If the polynucleotide is derived from genomic DNA, expression may include splicing of the mRNA in a eukaryotic cell.

[00299] As used herein, "operably linked", "operable linkage", "operatively linked", or grammatical equivalents thereof generally refer to juxtaposition of genetic elements, e.g., a promoter, an enhancer, a polyadenylation sequence, etc., wherein the elements are in a relationship permitting them to operate in the expected manner. For instance, a regulatory element, which may comprise promoter and/or enhancer sequences, is operatively linked to a coding region if the regulatory element helps initiate transcription of the coding sequence. There

may be intervening residues between the regulatory element and coding region so long as this functional relationship is maintained.

[00300] A "vector" as used herein, generally refers to a macromolecule or association of macromolecules that comprises or associates with a polynucleotide and which may be used to mediate delivery of the polynucleotide to a cell. Examples of vectors include plasmids, viral vectors, liposomes, and other gene delivery vehicles. The vector generally comprises genetic elements, e.g., regulatory elements, operatively linked to a gene to facilitate expression of the gene in a target.

[00301] As used herein, "an expression cassette" and "a nucleic acid cassette" are used interchangeably generally to refer to a combination of nucleic acid sequences or elements that are expressed together or are operably linked for expression. In some cases, an expression cassette refers to the combination of regulatory elements and a gene or genes to which they are operably linked for expression.

[00302] A "functional fragment" of a DNA or protein sequence generally refers to a fragment that retains a biological activity (either functional or structural) that is substantially similar to a biological activity of the full-length DNA or protein sequence. A biological activity of a DNA sequence may be its ability to influence expression in a manner attributed to the full-length sequence.

[00303] As used herein, an "engineered" object generally indicates that the object has been modified by human intervention. According to non-limiting examples: a nucleic acid may be modified by changing its sequence to a sequence that does not occur in nature; a nucleic acid may be modified by ligating it to a nucleic acid that it does not associate with in nature such that the ligated product possesses a function not present in the original nucleic acid; an engineered nucleic acid may synthesized *in vitro* with a sequence that does not exist in nature; a protein may be modified by changing its amino acid sequence to a sequence that does not exist in nature; an engineered protein may acquire a new function or property. An "engineered" system comprises at least one engineered component.

[00304] As used herein, "synthetic" and "artificial" can generally be used interchangeably to refer to a protein or a domain thereof that has low sequence identity (e.g., less than 50% sequence identity, less than 25% sequence identity, less than 10% sequence identity, less than 5% sequence identity, less than 1% sequence identity) to a naturally occurring human protein. For example, VPR and VP64 domains are synthetic transactivation domains.

[00305] As used herein, the term "Cas12a" generally refers to a family of Cas endonucleases that are class 2, Type V-A Cas endonucleases and that (a) use a relatively small guide RNA (about

42-44 nucleotides) that is processed by the nuclease itself following transcription from the CRISPR array, and (b) cleave DNA to leave staggered cut sites. Further features of this family of enzymes can be found, e.g. in Zetsche B, Heidenreich M, Mohanraju P, et al. Nat Biotechnol 2017;35:31–34, and Zetsche B, Gootenberg JS, Abudayyeh OO, et al. Cell 2015;163:759–771, which are incorporated by reference herein.

[00306] As used herein, a "guide nucleic acid" can generally refer to a nucleic acid that may hybridize to another nucleic acid. A guide nucleic acid may be RNA. A guide nucleic acid may be DNA. The guide nucleic acid may be programmed to bind to a sequence of nucleic acid site-specifically. The nucleic acid to be targeted, or the target nucleic acid, may comprise nucleotides. The guide nucleic acid may comprise nucleotides. A portion of the target nucleic acid may be complementary to a portion of the guide nucleic acid. The strand of a double-stranded target polynucleotide that is complementary to and hybridizes with the guide nucleic acid may be called the complementary strand. The strand of the double-stranded target polynucleotide that is complementary to the complementary strand, and therefore may not be complementary to the guide nucleic acid may be called noncomplementary strand. A guide nucleic acid may comprise a polynucleotide chain and can be called a "single guide nucleic acid." A guide nucleic acid may comprise two polynucleotide chains and may be called a "double guide nucleic acid." If not otherwise specified, the term "guide nucleic acid" may be inclusive, referring to both single guide nucleic acids and double guide nucleic acids. A guide nucleic acid may comprise a segment that can be referred to as a "nucleic acid-targeting segment" or a "nucleic acid-targeting sequence" or "spacer sequence." A nucleic acid-targeting segment may comprise a sub-segment that may be referred to as a "protein binding segment" or "protein binding sequence" or "Cas protein binding segment".

[00307] The term "sequence identity" or "percent identity" in the context of two or more nucleic acids or polypeptide sequences, generally refers to two (e.g., in a pairwise alignment) or more (e.g., in a multiple sequence alignment) sequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence over a local or global comparison window, as measured using a sequence comparison algorithm. Suitable sequence comparison algorithms for polypeptide sequences include, e.g., BLASTP using parameters of a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix setting gap costs at existence of 11, extension of 1, and using a conditional compositional score matrix adjustment for polypeptide sequences longer than 30 residues; BLASTP using parameters of a wordlength (W) of 2, an expectation (E) of 1000000, and the PAM30 scoring matrix setting gap costs at 9 to open gaps and 1 to extend gaps

for sequences of less than 30 residues (these are the default parameters for BLASTP in the BLAST suite available at https://blast.ncbi.nlm.nih.gov); CLUSTALW with the Smith-Waterman homology search algorithm parameters with a match of 2, a mismatch of -1, and a gap of -1; MUSCLE with default parameters; MAFFT with parameters of a retree of 2 and max iterations of 1000; Novafold with default parameters; HMMER hmmalign with default parameters.

[00308] The term "optimally aligned" in the context of two or more nucleic acids or polypeptide sequences, generally refers to two (e.g., in a pairwise alignment) or more (e.g., in a multiple sequence alignment) sequences that have been aligned to maximal correspondence of amino acids residues or nucleotides, for example, as determined by the alignment producing a highest or "optimized" percent identity score.

[00309] Included in the current disclosure are variants of any of the enzymes described herein with one or more conservative amino acid substitutions. Such conservative substitutions can be made in the amino acid sequence of a polypeptide without disrupting the three-dimensional structure or function of the polypeptide. Conservative substitutions can be accomplished by substituting amino acids with similar hydrophobicity, polarity, and R chain length for one another. Additionally, or alternatively, by comparing aligned sequences of homologous proteins from different species, conservative substitutions can be identified by locating amino acid residues that have been mutated between species (e.g., non-conserved residues) without altering the basic functions of the encoded proteins. Such conservatively substituted variants may include variants with at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% sequence identity to any one of the endonuclease protein sequences described herein (e.g. MG11, MG13, MG26, MG28, MG29, MG30, MG31, MG32, MG37, MG53, MG54, MG55, MG56, MG57, MG58, MG59, MG60, MG61, MG62, MG70, MG82, MG83, MG84 or MG85 family endonucleases described herein, or any other family nuclease described herein). In some embodiments, such conservatively substituted variants are functional variants. Such functional variants can encompass sequences with substitutions such that the activity of one or more critical active site residues or guide RNA binding residues of the endonuclease are not disrupted. In some embodiments, a functional variant of any of the proteins described herein lacks substitution of at least one of the conserved or functional residues called out in FIGURES 17, 18, 10, 20, or 25 or a residue described in Table 1B. In some embodiments, a functional variant of any of the

proteins described herein lacks substitution of all of the conserved or functional residues called out in FIGURES 17, 18, 10, 20, or 25 or a residue described in Table 1B.

[00310] Also included in the current disclosure are variants of any of the enzymes described herein with substitution of one or more catalytic residues to decrease or eliminate activity of the enzyme (e.g. decreased-activity variants). In some embodiments, a decreased activity variant as a protein described herein comprises a disrupting substitution of at least one, at least two, or all three catalytic residues identified in Table 1B.

[00311] Conservative substitution tables providing functionally similar amino acids are available from a variety of references (see, for e.g., Creighton, Proteins: Structures and Molecular Properties (W H Freeman & Co.; 2nd edition (December 1993)). The following eight groups each contain amino acids that are conservative substitutions for one another:

1) Alanine (A), Glycine (G);
2) Aspartic acid (D), Glutamic acid (E);
3) Asparagine (N), Glutamine (Q);
4) Arginine (R), Lysine (K);
5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V);
6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W);
7) Serine (S), Threonine (T); and
8) Cysteine (C), Methionine (M)

Overview

[00312] The discovery of new Cas enzymes with unique functionality and structure may offer the potential to further disrupt deoxyribonucleic acid (DNA) editing technologies, improving speed, specificity, functionality, and ease of use. Relative to the predicted prevalence of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) systems in microbes and the sheer diversity of microbial species, relatively few functionally characterized CRISPR/Cas enzymes exist in the literature. This is partly because a huge number of microbial species may not be readily cultivated in laboratory conditions. Metagenomic sequencing from natural environmental niches containing large numbers of microbial species may offer the potential to drastically increase the number of new CRISPR/Cas systems characterized and speed the discovery of new oligonucleotide editing functionalities. A recent example of the fruitfulness of such an approach is demonstrated by the 2016 discovery of CasX/CasY CRISPR systems from metagenomic analysis of natural microbial communities.

[00313] CRISPR/Cas systems are RNA-directed nuclease complexes that have been described to function as an adaptive immune system in microbes. In their natural context, CRISPR/Cas systems occur in CRISPR (clustered regularly interspaced short palindromic repeats) operons or loci, which generally comprise two parts: (i) an array of short repetitive sequences (30-40bp) separated by equally short spacer sequences, which encode the RNA-based targeting element; and (ii) ORFs encoding the Cas encoding the nuclease polypeptide directed by the RNA-based targeting element alongside accessory proteins/enzymes. Efficient nuclease targeting of a particular target nucleic acid sequence generally requires both (i) complementary hybridization between the first 6-8 nucleic acids of the target (the target seed) and the crRNA guide; and (ii) the presence of a protospacer-adjacent motif (PAM) sequence within a defined vicinity of the target seed (the PAM usually being a sequence not commonly represented within the host genome). Depending on the exact function and organization of the system, CRISPR-Cas systems are commonly organized into 2 classes, 5 types and 16 subtypes based on shared functional characteristics and evolutionary similarity (see **FIG.**).

[00314] Class I CRISPR-Cas systems have large, multi-subunit effector complexes, and comprise Types I, III, and IV. Class II CRISPR-Cas systems generally have single-polypeptide multidomain nuclease effectors, and comprise Types II, V and VI.

[00315] Type II CRISPR-Cas systems are considered the simplest in terms of components. In Type II CRISPR-Cas systems, the processing of the CRISPR array into mature crRNAs does not require the presence of a special endonuclease subunit, but rather a small trans-encoded crRNA (tracrRNA) with a region complementary to the array repeat sequence; the tracrRNA interacts with both its corresponding effector nuclease (e.g. Cas9) and the repeat sequence to form a precursor dsRNA structure, which is cleaved by endogenous RNAse III to generate a mature effector enzyme loaded with both tracrRNA and crRNA. Cas II nucleases are identified as DNA nucleases. Type 2 effectors generally exhibit a structure comprising a RuvC-like endonuclease domain that adopts the RNase H fold with an unrelated HNH nuclease domain inserted within the folds of the RuvC-like nuclease domain. The RuvC-like domain is responsible for the cleavage of the target (e.g., crRNA complementary) DNA strand, while the HNH domain is responsible for cleavage of the displaced DNA strand.

[00316] Type V CRISPR-Cas systems are characterized by a nuclease effector (e.g. Cas12) structure similar to that of Type II effectors, comprising a RuvC-like domain. Similar to Type II, most (but not all) Type V CRISPR systems use a tracrRNA to process pre-crRNAs into mature crRNAs; however, unlike Type II systems which requires RNAse III to cleave the pre-crRNA into multiple crRNAs, type V systems are capable of using the effector nuclease itself to cleave

pre-crRNAs. Like Type-II CRISPR-Cas systems, Type V CRISPR-Cas systems are again identified as DNA nucleases. Unlike Type II CRISPR-Cas systems, some Type V enzymes (e.g., Cas12a) appear to have a robust single-stranded nonspecific deoxyribonuclease activity that is activated by the first crRNA directed cleavage of a double-stranded target sequence.

[00317] CRISPR-Cas systems have emerged in recent years as the gene editing technology of choice due to their targetability and ease of use. The most commonly used systems are the Class 2 Type II SpCas9 and the Class 2 Type V-A Cas12a. The Type V-A systems in particular are becoming more widely used since their reported specificity in cells is higher than other nucleases, with fewer or no off-target effects. The V-A systems are also advantageous in that the guide RNA is small (42-44 nucleotides compared with approximately 100 nt for SpCas9) and is processed by the nuclease itself following transcription from the CRISPR array, simplifying multiplexed applications with multiple gene edits. Furthermore, the V-A systems have staggered cut sites, which may facilitate directed repair pathways, such as microhomology-dependent targeted integration (MITI).

[00318] The most commonly used Type V-A enzymes require a 5' protospacer adjacent motif (PAM) next to the chosen target site: 5'-TTTV-3' for *Lachnospiraceae* bacterium ND2006 LbCas12a and *Acidaminococcus sp.* AsCas12a; and 5'-TTV-3' for *Francisella novicida* FnCas12a. Recent exploration of orthologs has revealed proteins with less restrictive PAM sequences that are also active in mammalian cell culture, for example YTV, YYN or TTN. However, these enzymes do not fully encompass V-A biodiversity and targetability, and may not represent all possible activities and PAM sequence requirements. Here, thousands of genomic fragments were mined from numerous metagenomes for Type V-A nucleases. The diversity of identified V-A enzymes may have been expanded and novel systems may have been developed into highly targetable, compact, and precise gene editing agents.

MG Enzymes

[00319] Type V-A CRISPR systems are quickly being adopted for use in a variety of genome editing applications. These programmable nucleases are part of adaptive microbial immune systems, the natural diversity of which has been largely unexplored. Novel families of Type V-A CRISPR enzymes were identified through a large-scale analysis of metagenomes collected from a variety of complex environments, and developed representatives of these systems into gene-editing platforms. The nucleases are phylogenetically diverse (see **FIG. 4A**) and recognize a single guide RNA with specific motifs. The majority of these systems come from uncultivated organisms, some of which encode a divergent Type V effector within the same CRISPR operon.

Biochemical analysis uncovered unexpected PAM diversity (see **FIG. 4B**), indicating that these systems will facilitate a variety of genome engineering applications. The simplicity of guide sequences and activity in human cell lines suggest utility in gene and cell therapies.

[00320] In some aspects, the present disclosure provides for novel Type V-L candidates (see **FIGs. 27A-B**). Type V-L may be a novel subtype and some sub-families may have been identified. These nucleases are about 1000 – 1100 amino acids in length. Type V-L may be found in the same CRISPR locus as Type V-A effectors. RuvC catalytic residues may have been identified for Type V-L candidates and these Type V-L candidates may not require tracrRNA. One example of a Type V-L are the MG60 nucleases described herein (see **FIG. 28** and **FIG. 32**).

[00321] In some aspects, the present disclosure provides for smaller Type V effectors (see **FIG. 30**). Such effectors may be small putative effectors. These effectors may simplify delivery and may extend therapeutic applications.

[00322] In some aspects, the present disclosure provides for novel type V effector. Such an effector may be MG70 as described herein (see **FIG. 29**). MG70 may be an ultra-small enzyme of about 373 amino acids in length. MG 70 may have a single transposase domain at the N-terminus and may have a predicted tracrRNA (see **FIG. 30** and **FIG. 32**).

[00323] In some aspects, the present disclosure provides for a smaller Type V effector (see **FIG. 31**). Such an effector may be MG81 described herein. MG81 may be about 500-700 amino acids in length and may contain RuvC, and HTH DNA binding domains.

[00324] In one aspect, the present disclosure provides for an engineered nuclease system discovered through metagenomic sequencing. In some cases, the metagenomic sequencing is conducted on samples. In some cases, the samples may be collected from a variety of environments. Such environments may be a human microbiome, an animal microbiome, environments with high temperatures, environments with low temperatures. Such environments may include sediment. An example of the types of such environments of the engineered nuclease systems described herein may be found in **FIG.** .

[00325] In one aspect, the present disclosure provides for an engineered nuclease system comprising (a) an endonuclease. In some cases, the endonuclease is a Cas endonuclease. In some cases, the endonuclease is a class 2, type V Cas endonuclease. In some cases, the endonuclease is a class 2, type V-A Cas endonuclease. In some cases, the endonuclease is derived from an uncultivated microorganism. The endonuclease may comprise a RuvC domain. In some cases, the engineered nuclease system comprises (b) an engineered guide RNA. In some cases, the engineered guide RNA is configured to form a complex with the endonuclease. In some cases,

the engineered guide RNA comprises a spacer sequence. In some cases, the spacer sequence is configured to hybridize to a target nucleic acid sequence.

[00326] In one aspect, the present disclosure provides for an engineered nuclease system comprising (a) an endonuclease. In some cases, the endonuclease has at least about 70% sequence identity to any one of SEQ ID NOs: 1-3470. In some cases, the endonuclease has at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 1-3470.

[00327] In some cases, the endonuclease comprises a variant having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 1-3470. In some cases, the endonuclease may be substantially identical to any one of SEQ ID NOs: 1-3470.

[00328] In some cases, the engineered nuclease system comprises an engineered guide RNA. In some cases, the engineered guide RNA is configured to form a complex with the endonuclease. In some cases, the engineered guide RNA comprises a spacer sequence. In some cases, the spacer sequence is configured to hybridize to a target nucleic acid sequence.

[00329] In one aspect, the present disclosure provides an engineered nuclease system comprising (a) an endonuclease. In some cases, the endonuclease is configured to bind to a protospacer adjacent motif (PAM) sequence. In some cases, the PAM sequence is substantially identical to any one of SEQ ID NOs: 3863-3913. In some cases, the PAM sequence any one of SEQ ID NOs: 3863-3913. In some cases, the endonuclease is a Cas endonuclease. In some cases, the endonuclease is a class 2 Cas endonuclease. In some cases, the endonuclease is a class 2, type V Cas endonuclease. In some cases, the endonuclease is a class 2, type V-A Cas endonuclease. In some cases, the engineered nuclease system comprises (b) an engineered guide RNA. In some cases, the engineered guide RNA is configured to form a complex with the endonuclease. In some cases, the engineered guide RNA comprises a spacer sequence. In some cases, the spacer sequence is configured to hybridize to a target nucleic acid sequence.

[00330] In some cases, the endonuclease is not a Cpf1 or Cms1 endonuclease. In some cases, the endonuclease further comprises a zinc finger-like domain.

[00331] In some cases, the guide RNA comprises a sequence with at least 80% sequence identity to the first 19 nucleotides or the non-degenerate nucleotides of SEQ ID NO: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3677-3678, 3695-3696, 3729-3730, 3734-3735, or 3851-3857. In some cases, the guide RNA comprises a sequence with at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to the first 19 nucleotides or the non-degenerate nucleotides of SEQ ID NO: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3677-3678, 3695-3696, 3729-3730, 3734-3735, or 3851-3857. In some cases, the guide RNA comprises a variant having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to the first 19 nucleotides or the non-degenerate nucleotides of SEQ ID NO: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3677-3678, 3695-3696, 3729-3730, 3734-3735, or 3851-3857. In some cases, the guide RNA comprises a sequence which is substantially identical to the first 19 nucleotides or the non-degenerate nucleotides of SEQ ID NO: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3677-3678, 3695-3696, 3729-3730, 3734-3735, or 3851-3857.

[00332] In some cases, the guide RNA comprises a sequence with at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to the first 19 nucleotides or the

non-degenerate nucleotides of SEQ ID NO: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3677-3678, 3695-3696, 3729-3730, 3734-3735, or 3851-3857. In some cases, the endonuclease is configured to bind to the engineered guide RNA. In some cases, the Cas endonuclease is configured to bind to the engineered guide RNA. In some cases, the class 2 Cas endonuclease is configured to bind to the engineered guide RNA. In some cases, the class 2, type V Cas endonuclease is configured to bind to the engineered guide RNA. .In some cases, the class 2, type V-A Cas endonuclease is configured to bind to the engineered guide RNA.

[00333] In some cases, the endonuclease is configured to bind to a protospacer adjacent motif (PAM) sequence comprising any one of SEQ ID NOs: 3863-3913.

[00334] In some cases, the guide RNA comprises a sequence complementary to a eukaryotic, fungal, plant, mammalian, or human genomic polynucleotide sequence. In some cases, the guide RNA comprises a sequence complementary to a eukaryotic genomic polynucleotide sequence. In some cases, the guide RNA comprises a sequence complementary to a fungal genomic polynucleotide sequence. In some cases, the guide RNA comprises a sequence complementary to a plant genomic polynucleotide sequence. In some cases, the guide RNA comprises a sequence complementary to a mammalian genomic polynucleotide sequence. In some cases, the guide RNA comprises a sequence complementary to a human genomic polynucleotide sequence.

[00335] In some cases, the guide RNA is 30-250 nucleotides in length. In some cases, the guide RNA is 42-44 nucleotides in length. In some cases, the guide RNA is 42 nucleotides in length. In some cases, the guide RNA is 43 nucleotides in length. In some cases, the guide RNA is 44 nucleotides in length. In some cases, the guide RNA is 85-245 nucleotides in length. In some cases, the guide RNA is more than 90 nucleotides in length. In some cases, the guide RNA is less than 245 nucleotides in length.

[00336] In some cases, the endonuclease may comprise a variant having one or more nuclear localization sequences (NLSs). The NLS may be proximal to the N- or C-terminus of the endonuclease. The NLS may be appended N-terminal or C-terminal to any one of SEQ ID NOs: 3938-3953, or to a variant having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 3938-3953. In some cases, the NLS may comprise a sequence substantially identical to any one of SEQ ID NOs: 3938-3953.

Table 1: Example NLS Sequences that may be used with Cas Effectors according to the disclosure.

Source	NLS amino acid sequence	SEQ ID NO:
SV40	PKKKRKV	3938
nucleoplasmin bipartite NLS	KRPAATKKAGQAKKKK	3939
c-myc NLS	PAAKRVKLD	3940
c-myc NLS	RQRRNELKRSP	3941
hRNPA1 M9 NLS	NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY	3942
Importin-alpha IBB domain	RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV	3943
Myoma T protein	VSRKRPRP	3944
Myoma T protein	PPKKARED	3945
p53	PQPKKKPL	3946
mouse c-abl IV	SALIKKKKKMAP	3947
influenza virus NS1	DRLRR	3948
influenza virus NS1	PKQKKRK	3949
Hepatitis virus delta antigen	RKLKKKIKKL	3950
mouse Mx1 protein	REKKKFLKRR	3951
human poly(ADP-ribose) polymerase	KRKGDEVDGVDEVAKKKSKK	3952
steroid hormone receptors (human) glucocorticoid	RKCLQAGMNLEARKTKK	3953

[00337] In some cases, the engineered nuclease system further comprises a single- or double stranded DNA repair template. In some cases, the engineered nuclease system further comprises a single-stranded DNA repair template. In some cases, the engineered nuclease system further comprises a double-stranded DNA repair template. In some cases, the single- or double-stranded DNA repair template may comprise from 5' to 3': a first homology arm comprising a sequence of at least 20 nucleotides 5′ to said target deoxyribonucleic acid sequence, a synthetic DNA sequence of at least 10 nucleotides, and a second homology arm comprising a sequence of at least 20 nucleotides 3′ to said target sequence.

[00338] In some cases, the first homology arm comprises a sequence of at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 110, at least 120, at least 130, at least 140, at least 150, at least 175, at least 200, at least 250, at least 300, at least 400, at least 500, at least 750, or at least 1000 nucleotides. In some cases, the second homology arm comprises a sequence of at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at

least 100, at least 110, at least 120, at least 130, at least 140, at least 150, at least 175, at least 200, at least 250, at least 300, at least 400, at least 500, at least 750, or at least 1000 nucleotides.

[00339] In some cases, the first and second homology arms are homologous to a genomic sequence of a prokaryote. In some cases, the first and second homology arms are homologous to a genomic sequence of a bacteria. In some cases, the first and second homology arms are homologous to a genomic sequence of a fungus. In some cases, the first and second homology arms are homologous to a genomic sequence of a eukaryote.

[00340] In some cases, the engineered nuclease system further comprises a DNA repair template. The DNA repair template may comprise a double-stranded DNA segment. The double-stranded DNA segment may be flanked by one single-stranded DNA segment. The double-stranded DNA segment may be flanked by two single-stranded DNA segments. In some cases, the single-stranded DNA segments are conjugated to the 5' ends of the double-stranded DNA segment. In some cases, the single stranded DNA segments are conjugated to the 3' ends of the double-stranded DNA segment.

[00341] In some cases, the single-stranded DNA segments have a length from 1 to 15 nucleotide bases. In some cases, the single-stranded DNA segments have a length from 4 to 10 nucleotide bases. In some cases, the single-stranded DNA segments have a length of 4 nucleotide bases. In some cases, the single-stranded DNA segments have a length of 5 nucleotide bases. In some cases, the single-stranded DNA segments have a length of 6 nucleotide bases. In some cases, the single-stranded DNA segments have a length of 7 nucleotide bases. In some cases, the single-stranded DNA segments have a length of 8 nucleotide bases. In some cases, the single-stranded DNA segments have a length of 9 nucleotide bases. In some cases, the single-stranded DNA segments have a length of 10 nucleotide bases.

[00342] In some cases, the single-stranded DNA segments have a nucleotide sequence complementary to a sequence within the spacer sequence. In some cases, the double-stranded DNA sequence comprises a barcode, an open reading frame, an enhancer, a promoter, a protein-coding sequence, a miRNA coding sequence, an RNA coding sequence, or a transgene.

[00343] In some cases, the engineered nuclease system further comprises a source of Mg^{2+}.

[00344] In some cases, the guide RNA comprises a hairpin comprising at least 8 base-paired ribonucleotides. In some cases, the guide RNA comprises a hairpin comprising at least 9 base-paired ribonucleotides. In some cases, the guide RNA comprises a hairpin comprising at least 10 base-paired ribonucleotides. In some cases, the guide RNA comprises a hairpin comprising at least 11 base-paired ribonucleotides. In some cases, the guide RNA comprises a hairpin comprising at least 12 base-paired ribonucleotides.

[00345] In some cases, the endonuclease comprises a sequence at least 70% identical to a variant of any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1721 or a variant thereof 141, 215, 229, 261, or 1711-1721 or a variant thereof . In some cases, the endonuclease comprises a sequence at least 75% identical to a variant of any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1721 or a variant thereof . In some cases, the endonuclease comprises a sequence at least 80% identical to a variant of any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1721 or a variant thereof . In some cases, the endonuclease comprises a sequence at least 85% identical to a variant of any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1721 or a variant thereof . In some cases, the endonuclease comprises a sequence at least 90% identical to a variant of any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1721 or a variant thereof . In some cases, the endonuclease comprises a sequence at least 95% identical to a variant of any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1721 or a variant thereof .

[00346] In some cases, the guide RNA structure comprises a sequence of at least 70% identical to the first 19 nucleotides or the non-degenerate nucleotides of SEQ ID NO: 3608. In some cases, the guide RNA structure comprises a sequence of at least 75% identical to the first 19 nucleotides or the non-degenerate nucleotides of SEQ ID NO: 3608. In some cases, the guide RNA structure comprises a sequence of at least 80% identical to the first 19 nucleotides or the non-degenerate nucleotides of SEQ ID NO: 3608. In some cases, the guide RNA structure comprises a sequence of at least 85% identical to the first 19 nucleotides or the non-degenerate nucleotides of SEQ ID NO: 3608. In some cases, the guide RNA structure comprises a sequence of at least 90% identical to the first 19 nucleotides or the non-degenerate nucleotides of SEQ ID NO: 3608. In some cases, the guide RNA structure comprises a sequence of at least 95% identical to the first 19 nucleotides or the non-degenerate nucleotides of SEQ ID NO: 3608. In some cases, the endonuclease is configured to bind to a PAM comprising any one of SEQ ID NOs: 3863-3913.

[00347] In some cases, sequence may be determined by a BLASTP, CLUSTALW, MUSCLE, or MAFFT algorithm, or a CLUSTALW algorithm with the Smith-Waterman homology search algorithm parameters. The sequence identity may be determined by said BLASTP homology search algorithm using parameters of a wordlength (W) of 3, an expectation (E) of 10, and a BLOSUM62 scoring matrix setting gap costs at existence of 11, extension of 1, and using a conditional compositional score matrix adjustment.

[00348] In one aspect, the present disclosure provides an engineered guide RNA comprising (a) a DNA-targeting segment. In some cases, the DNA-targeting segment comprises a nucleotide sequence that is complementary to a target sequence. In some cases, the target sequence is in a target DNA molecule. In some cases, the engineered guide RNA comprises (b) a protein-binding

segment. In some cases, the protein-binding segment comprises two complementary stretches of nucleotides. In some cases, the two complementary stretches of nucleotides hybridize to form a double-stranded RNA (dsRNA) duplex. In some cases, the two complementary stretches of nucleotides are covalently linked to one another with intervening nucleotides. In some cases, the engineered guide ribonucleic acid polynucleotide is capable of forming a complex with an endonuclease. In some cases, the endonuclease has at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 1-3470. In some cases, the complex targets the target sequence of the target DNA molecule.

[00349] In some cases, the DNA-targeting segment is positioned 3' of both of the two complementary stretches of nucleotides. In some cases, the protein binding segment comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to the first 19 nucleotides or the non-degenerate nucleotides of SEQ ID NO: 3608.

[00350] In some cases, the double-stranded RNA (dsRNA) duplex comprises at least 8 ribonucleotides. In some cases, the double-stranded RNA (dsRNA) duplex comprises at least 9 ribonucleotides. In some cases, the double-stranded RNA (dsRNA) duplex comprises at least 10 ribonucleotides. In some cases, the double-stranded RNA (dsRNA) duplex comprises at least 11 ribonucleotides. In some cases, the double-stranded RNA (dsRNA) duplex comprises at least 12 ribonucleotides.

[00351] In some cases, the deoxyribonucleic acid polynucleotide encodes the engineered guide ribonucleic acid polynucleotide.

[00352] In one aspect, the present disclosure provides a nucleic acid comprising an engineered nucleic acid sequence. In some cases, the engineered nucleic acid sequence is optimized for expression in an organism. In some cases, the nucleic acid encodes an endonuclease. In some cases, the endonuclease is a Cas endonuclease. In some cases, the endonuclease is a class 2 endonuclease. In some cases, the endonuclease is a class2, type V Cas endonuclease. In some cases, the endonuclease is a class2, type V-A Cas endonuclease. In some cases, the endonuclease

is derived from an uncultivated microorganism. In some cases, the organism is not the uncultivated organism.

[00353] In some cases, the endonuclease comprises a variant having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 1-3470.

[00354] In some cases, the endonuclease may comprise a variant having one or more nuclear localization sequences (NLSs). The NLS may be proximal to the N- or C-terminus of the endonuclease. The NLS may be appended N-terminal or C-terminal to any one of SEQ ID NOs: 3938-3953, or to a variant having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 3938-3953.

[00355] In some cases, the organism is prokaryotic. In some cases, the organism is bacterial. In some cases, the organism is eukaryotic. In some cases, the organism is fungal. In some cases, the organism is a plant. In some cases, the organism is mammalian. In some cases, the organism is a rodent. In some cases, the organism is human.

[00356] In one aspect, the present disclosure provides an engineered vector. In some cases, the engineered vector comprises a nucleic acid sequence encoding an endonuclease. In some cases, the endonuclease is a Cas endonuclease. In some cases, the endonuclease is a class 2 Cas endonuclease. In some cases, the endonuclease is a class 2, type V Cas endonuclease. In some cases, the endonuclease is a class2, type V-A Cas endonuclease. In some cases, the endonuclease is derived from an uncultivated microorganism.

[00357] In some cases, the engineered vector comprises a nucleic acid described herein. In some cases, the nucleic acid described herein is a deoxyribonucleic acid polynucleotide described herein. In some cases, the vector is a plasmid, a minicircle, a CELiD, an adeno-associated virus (AAV) derived virion, or a lentivirus.

[00358] In one aspect, the present disclosure provides a cell comprising a vector described herein.

[00359] In one aspect, the present disclosure provides a method of manufacturing an endonuclease. In some cases, the method comprises cultivating the cell.

[00360] In one aspect, the present disclosure provides a method for binding, cleaving, marking, or modifying a double-stranded deoxyribonucleic acid polynucleotide. The method may comprise contacting the double-stranded deoxyribonucleic acid polynucleotide with an endonuclease. In some cases, the endonuclease is a Cas endonuclease. In some cases, the endonuclease is a class 2 Cas endonuclease. In some cases, the endonuclease is a class 2, type V Cas endonuclease. In some cases, the endonuclease is a class2, type V-A Cas endonuclease. In some cases, the endonuclease is in complex with an engineered guide RNA. In some cases, the engineered guide RNA is configured to bind to the endonuclease. In some cases, the engineered guide RNA is configured to bind to the double-stranded deoxyribonucleic acid polynucleotide. In some cases, the engineered guide RNA is configured to bind to the endonuclease and to the double-stranded deoxyribonucleic acid polynucleotide. In some cases, the double-stranded deoxyribonucleic acid polynucleotide comprises a protospacer adjacent motif (PAM). In some cases, the PAM comprises a sequence comprising any one of SEQ ID NOs: 3863-3913.

[00361] In some cases, the double-stranded deoxyribonucleic acid polynucleotide comprises a first strand comprising a sequence complementary to a sequence of the engineered guide RNA and a second strand comprising the PAM. In some cases, the PAM is directly adjacent to the 5′ end of the sequence complementary to the sequence of the engineered guide RNA. In some cases, the endonuclease is not a Cpf1 endonuclease or a Cms1 endonuclease. In some cases, the endonuclease is derived from an uncultivated microorganism. In some cases, the double-stranded deoxyribonucleic acid polynucleotide is a eukaryotic, plant, fungal, mammalian, rodent, or human double-stranded deoxyribonucleic acid polynucleotide. In some cases, the PAM comprises any one of SEQ ID NOs: 3863-3913.

[00362] In one aspect, the present disclosure provides a method of modifying a target nucleic acid locus. The method may comprise delivering to the target nucleic acid locus the engineered nuclease system described herein. In some cases, the endonuclease is configured to form a complex with the engineered guide ribonucleic acid structure. In some cases, the complex is configured such that upon binding of the complex to the target nucleic acid locus, the complex modifies the target nucleic acid locus.

[00363] In some cases, modifying the target nucleic acid locus comprises binding, nicking, cleaving, or marking said target nucleic acid locus. In some cases, the target nucleic acid locus comprises deoxyribonucleic acid (DNA) or ribonucleic acid (RNA). In some cases, the target nucleic acid comprises genomic DNA, viral DNA, viral RNA, or bacterial DNA. In some cases,

the target nucleic acid locus is *in vitro*. In some cases, the target nucleic acid locus is within a cell. In some cases, the cell is a prokaryotic cell, a bacterial cell, a eukaryotic cell, a fungal cell, a plant cell, an animal cell, a mammalian cell, a rodent cell, a primate cell, or a human cell.

[00364] In some cases, delivery of the engineered nuclease system to the target nucleic acid locus comprises delivering the nucleic acid described herein or the vector described herein. In some cases, delivery of engineered nuclease system to the target nucleic acid locus comprises delivering a nucleic acid comprising an open reading frame encoding the endonuclease. In some cases, the nucleic acid comprises a promoter. In some cases, the open reading frame encoding the endonuclease is operably linked to the promoter.

[00365] In some cases, delivery of the engineered nuclease system to the target nucleic acid locus comprises delivering a capped mRNA containing the open reading frame encoding the endonuclease. In some cases, delivery of the engineered nuclease system to the target nucleic acid locus comprises delivering a translated polypeptide. In some cases, delivery of the engineered nuclease system to the target nucleic acid locus comprises delivering a deoxyribonucleic acid (DNA) encoding the engineered guide RNA operably linked to a ribonucleic acid (RNA) pol III promoter.

[00366] In some cases, the endonuclease induces a single-stranded break or a double-stranded break at or proximal to the target locus. In some cases, the endonuclease induces a staggered single stranded break within or 3′ to said target locus.

[00367] In some cases, effector repeat motifs are used to inform guide design of MG nucleases. For example, the processed gRNA in Type V-A systems comprises the last 20-22 nucleotides of a CRISPR repeat. This sequence may be synthesized into a crRNA (along with a spacer) and tested *in vitro*, along with the synthesized nucleases, for cleavage on a library of possible targets. Using this method, the PAM may be determined. In some cases, Type V-A enzymes may use a "universal" gRNA. In some cases, Type V enzymes may utilize a unique gRNA.

[00368] *Lipid nanoparticles*

[00369] Lipid nanoparticles as described herein can be 4-component lipid nanoparticles. Such nanoparticles can be configured for delivery of RNA or other nucleic acids (e.g. synthetic RNA, mRNA, or *in vitro*-synthesized mRNA) and can be generally formulated as described in WO2012135805A2, which is incorporated by reference herein for all purposes. Such nanoparticles can generally comprise: (a) a cationic lipid (e.g. any of the lipids described in **FIG. 109**), (b) a neutral lipid (e.g. DSPC or DOPE), (c) a sterol (e.g. cholesterol or a cholesterol analog), and (d) a PEG-modified lipid (e.g. PEG-DMG).

[00370] The cationic lipid referred to herein as "C12-200" is disclosed by Love *et al.*, Proc Natl Acad Sci USA. 2010 107:1864-1869 and Liu and Huang, Molecular Therapy. 2010 669-670; both of which are herein incorporated by reference in their entirety. Cationic lipid formulations can include particles comprising either 3 or 4 or more components in addition to polynucleotide, primary construct, or RNA (e.g. mRNA). As an example, formulations with certain cationic lipids include, but are not limited to, 98N12-5 (or any of the other structures described in **FIG. 109**) and may contain 42% lipidoid, 48% cholesterol, and 10% PEG (C14 or greater alkyl chain length). As another example, formulations with certain lipidoids include, but are not limited to, C12-200 and may contain 50% cationic lipid, 10% disteroylphosphatidyl choline, 38.5% cholesterol, and 1.5% PEG-DMG.

[00371] In some embodiments, lipid nanoparticles are formulated as described in US10709779B2, which is incorporated in its entirety by reference herein. In some embodiments, the cationic lipid nanoparticle comprises a cationic lipid, a PEG-modified lipid, a sterol, and a non-cationic lipid. In some embodiments, the cationic lipid is selected from the group consisting of any of the cationic lipids depicted in **FIG. 109**. In some embodiments, the cationic lipid nanoparticle has a molar ratio of about 20-60% cationic lipid, about 5-25% non-cationic lipid, about 25-55% sterol, and about 0.5-15% PEG-modified lipid. In some embodiments, the cationic lipid nanoparticle comprises a molar ratio of about 50% cationic lipid, about 1.5% PEG-modified lipid, about 38.5% cholesterol, and about 10% non-cationic lipid. In some embodiments, the cationic lipid nanoparticle comprises a molar ratio of about 55% cationic lipid, about 2.5% PEG-modified lipid, about 32.5% cholesterol, and about 10% non-cationic lipid. In some embodiments, the cationic lipid is an ionizable cationic lipid, the non-cationic lipid is a neutral lipid, and the sterol is a cholesterol. In some embodiments, the cationic lipid nanoparticle has a molar ratio of 50:38.5:10:1.5 of cationic lipid: cholesterol: PEG2000-DMG:DSPC or DMG:DOPE. In some embodiments, lipid nanoparticles as described herein can comprise cholesterol, 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), 1,1'-((2-(4-(2-((2-(bis(2-hydroxydodecyl)amino)ethyl)(2-hydroxydodecyl)amino)ethyl)piperazin-1-yl)ethyl)azanediyl)bis(dodecan-2-ol) (C12-200), and DMG-PEG-2000 at molar ratios of 47.5:16:35:1.5.

[00372] Systems of the present disclosure may be used for various applications, such as, for example, nucleic acid editing (e.g., gene editing), binding to a nucleic acid molecule (e.g., sequence-specific binding). Such systems may be used, for example, for addressing (e.g., removing or replacing) a genetically inherited mutation that may cause a disease in a subject, inactivating a gene in order to ascertain its function in a cell, as a diagnostic tool to detect

disease-causing genetic elements (e.g. via cleavage of reverse-transcribed viral RNA or an amplified DNA sequence encoding a disease-causing mutation), as deactivated enzymes in combination with a probe to target and detect a specific nucleotide sequence (e.g. sequence encoding antibiotic resistance int bacteria), to render viruses inactive or incapable of infecting host cells by targeting viral genomes, to add genes or amend metabolic pathways to engineer organisms to produce valuable small molecules, macromolecules, or secondary metabolites, to establish a gene drive element for evolutionary selection, to detect cell perturbations by foreign small molecules and nucleotides as a biosensor.

EXAMPLES

Example 1 – A method of metagenomic analysis for new proteins

[00373] Metagenomic samples were collected from sediment, soil, and animals. Deoxyribonucleic acid (DNA) was extracted with a Zymobiomics DNA mini-prep kit and sequenced on an Illumina HiSeq® 2500. Samples were collected with consent of property owners. Metagenomic sequence data was searched using Hidden Markov Models generated based on identified Cas protein sequences including class II type V Cas effector proteins to identify new Cas effectors (see **FIG**. , which shows distribution of proteins detected in one family, MG29, identified from sample types such as high-temperature samples). Novel effector proteins identified by the search were aligned to identified proteins to identify potential active sites (see e.g. **FIG**. , which shows that all MG29 family effectors identified from various samples have three catalytic residues from RuvCI, RuvCII, and RuvCIII catalytic domains and are predicted to be active). This metagenomic workflow resulted in the delineation of the MG11, MG13, MG19, MG20, MG26, MG28, MG29, MG30, MG31, MG32, MG37, MG53, MG54, MG55, MG56, MG57, MG58, MG59, MG60, MG61, MG62, MG70, MG75, MG77, MG78, MG79, MG80, MG81, MG82, MG83, MG84, MG85, MG90, and MG91 families described herein. Putative spacer sequences were identified by their location adjacent to the genomic loci encoding the effector proteins.

Example 2 – A method of metagenomic analysis for new proteins

[00374] Thirteen animal microbiome, high temperature biofilm and sediment samples were collected and stored on ice or in Zymo DNA/RNA Shield after collection. DNA was extracted from samples using either the Qiagen DNeasy PowerSoil Kit or the ZymoBIOMICS DNA Miniprep Kit. DNA sequencing libraries were constructed and sequenced on an Illumina HiSeq 4000 or on a Novaseq machine at the Vincent J. Coates Genomics Sequencing Laboratory at UC

Berkeley, with paired 150 bp reads with a 400-800 bp target insert size (10 GB of sequencing was targeted per sample). Publicly available metagenomic sequencing data were downloaded from the NCBI SRA. Sequencing reads were trimmed using BBMap (Bushnell B., sourceforge.net/projects/bbmap/) and assembled with Megahit 11. Open reading frames and protein sequences were predicted with Prodigal. HMM profiles of identified Type V-A CRISPR nucleases were built and searched against all predicted proteins using HMMER3 (hmmer.org) to identify potential effectors. CRISPR arrays on assembled contigs were predicted with Minced (https://github.com/ctSkennerton/minced). Taxonomy was assigned to proteins with Kaiju, and contig taxonomy was determined by finding the consensus of all encoded proteins.

[00375] Predicted and reference (e.g., LbCas12a, AsCas12a, FnCas12a) Type V effector proteins were aligned with MAFFT and a phylogenetic tree was inferred using FasTree2. Novel families were delineated by identifying clades composed of sequences recovered from this study. From within families, candidates were selected if they contained the components for laboratory analysis (i.e., they were found on a well-assembled and annotated contig with a CRISPR array) in a manner that sampled as much phylogenetic diversity as possible. Priority was given to small effectors from diverse families (that is, families with representatives sharing a wider range of protein sequences). Selected representative and reference sequences were aligned using MUSCLE and Clustal W to identify catalytic and PAM interacting residues. CRISPR array repeats were searched for a motif associated with Type V-A systems, TCTAC-N-GTAGA (containing between one and eight N residues). From this analysis, families were putatively classified as V-A if representative CRISPR arrays contained one of these motif sequences. This dataset was used to identify HMM profiles associated with V-A families, which were in turn used to classify additional families (see **FIG. 33-FIG. 37**). Although the convention is to name novel Cas12 nucleases on the basis of the organism that encodes them, it is not possible to do so for the nucleases described herein. Therefore, in order to best adhere to the convention, the systems described herein are named with the prefix MG to indicate they are derived from assembled metagenomic fragments.

[00376] 140,867 Mbp of assembled metagenomic sequencing data was mined from diverse environments (soil, thermophilic, sediments, human and non-human microbiomes). In total, 119 genomic fragments encoded CRISPR effectors distantly related to Type V-A nucleases next to a CRISPR array (see **FIG. 4B**). Type V-A effectors were classified into 14 novel families sharing less than 30% average pairwise amino acid identity between each other, and with reference sequences (e.g., LbCas12a, AsCas12a, FnCas12a). Some effectors contained RuvC and alpha-helical recognition domains, as well as conserved DED nuclease catalytic residues from the

RuvCI/CII/CIII domains (identified in multiple sequence alignments, see e.g. Table 1A below), suggesting that these effectors were active nucleases (**FIG. 5-FIG. 7**). The novel Type V-A nucleases range in size from <800 to1,400 amino acids in length (see **FIG. 5A**) and their taxonomic classification spanned a diverse array of phyla (see **FIG. 4A**) suggesting possible horizontal transfer.

[00377] Some genomic fragments carrying a Type V-A CRISPR system also encoded a second effector, referred to here as Type V-A prime (V-A', **FIG. 7A**). For example, Type V-A' MG26-2, which shared 16.6% amino acid identity with the Type V-A MG26-1, was encoded in the same CRISPR Cas operon, and may share the same crRNA with MG26-1 (**FIG. 7B**). Although no nuclease domains were predicted, MG26-2 contained three RuvC catalytic residues identified from multiple sequence alignments (**FIG. 7B**).

Table 1A: Catalytic residues of Enzymes Described Herein Identified by Alignment

MGID	RuvC-I (D)	RuvC-II (E)	RuvC-III (D)
MG84-16	238	337	413
MG84-15	238	337	413
MG84-3	230	329	405
MG84-2	230	329	405
MG84-1	230	329	405
MG84-13	233	332	408
MG84-14	233	332	408
MG84-12	233	332	408
MG84-11	233	332	408
MG84-10	233	332	408
MG84-9	233	332	408
MG84-8	233	332	408
MG84-7	233	332	408
MG84-4	233	332	408
MG84-5	233	332	408
MG84-6	233	332	408
MG81-18	296	399	497
MG81-17	296	399	497
MG81-9	297	400	498
MG81-6	297	400	498
MG81-11	297	400	498
MG81-7	297	400	498
MG81-8	297	400	498
MG81-13	297	400	498
MG81-5	300	403	501
MG81-12	300	403	501
MG81-1	300	403	502

MGID	RuvC-I (D)	RuvC-II (E)	RuvC-III (D)
MG81-4	310	413	501
MG81-3	310	413	511
MG81-15	388	491	589
MG81-10	310	413	511
MG81-2	306	409	507
MG90-2	388	548	661
MG91-1	444	560	653
MG91-2	245	358	453
MG91-3	297	404	499
MG37-1	763	1167	1335
MG37-2	169	538	689
MG37-3	745	1202	1350
MG37-4	811	1230	1377
MG37-5	775	1173	1319
MG37-6	698	1058	1229
MG37-7	752	1135	1273
MG53-1	--	775	920
MG54-1	--	612	722

Example 3 - (General protocol) PAM Sequence identification/confirmation

[00378] PAM sequences that can be cleaved *in vitro* by a CRISPR effector were identified by incubating an effector with a crRNA and a plasmid library having 8 randomized nucleotides located adjacent to the 5' end of a sequence complementary to the spacer of the crRNA. The plasmid is configured such that if the 8 randomized nucleotides formed a functional PAM sequence, the plasmid was cleaved. Functional PAM sequences were then identified by ligating adapters to the ends of cleaved plasmids and then sequencing DNA fragments comprising the adapters. Putative endonucleases were expressed in an *E. coli* lysate-based expression system (myTXTL, Arbor Biosciences). An *E. coli* codon optimized nucleotide sequence encoding the putative nuclease was transcribed and translated *in vitro* from a PCR fragment under control of a T7 promoter. A second PCR fragment with a minimal CRISPR array composed of a T7 promoter followed by a repeat-spacer-repeat sequence was transcribed in the same reaction. Successful expression of the endonuclease and repeat-spacer-repeat sequence followed by CRISPR array processing provided active *in vitro* CRISPR nuclease complexes.

[00379] A library of target plasmids containing a spacer sequence matching that in the minimal array preceded by 8N (degenerate) bases (potential PAM sequences) was incubated with the output of the TXTL reaction. After 1-3 hours, the reaction was stopped and the DNA was recovered via a DNA clean-up kit, e.g., Zymo DCC, AMPure XP beads, QiaQuick etc. Adapter sequences were blunt-end ligated to DNA fragments with active PAM sequences that had been

cleaved by the endonuclease, whereas DNA that had not been cleaved was inaccessible for ligation. DNA segments comprising active PAM sequences were then amplified by PCR with primers specific to the library and the adapter sequence. The PCR amplification products were resolved on a gel to identify amplicons that corresponded to cleavage events. The amplified segments of the cleavage reaction were also used as templates for preparation of an NGS library or as a substrate for Sanger sequencing. Sequencing this resulting library, which was a subset of the starting 8N library, revealed sequences with PAM activity compatible with the CRISPR complex. For PAM testing with a processed RNA construct, the same procedure was repeated except that an *in vitro* transcribed RNA was added along with the plasmid library and the minimal CRISPR array template was omitted. The following sequences were used as targets in these assays: CGTGAGCCACCACGTCGCAAGCCT (SEQ ID NO: 3860); GTCGAGGCTTGCGACGTGGTGGCT (SEQ ID NO: 3861); GTCGAGGCTTGCGACGTGGTGGCT (SEQ ID NO: 3858); and TGGAGATATCTTGAACCTTGCATC (SEQ ID NO: 3859).

Example 4 – PAM Sequence identification/confirmation for endonucleases described herein

[00380] PAM requirements were determined via an *E. coli* lysate-based expression system (myTXTL, Arbor Biosciences), with modifications. Briefly, the *E. coli* codon optimized effector protein sequences were expressed under control of a T7 promoter at 29˚C for 16 hours. This crude protein stock was then used in an *in vitro* digest reaction at a concentration of 20% of the total reaction volume. The reaction was incubated for 3 hours at 37˚C with 5 nM of a plasmid library comprising a constant target sequence preceded by 8N mixed bases, and 50 nM of *in vitro* transcribed crRNA derived from the same CRISPR locus as the effector linked to a sequence complementary to the target sequence in NEB buffer 2.1 (New England Biolabs; NEB buffer 2.1 was selected in order to compare candidates with commercially available proteins). Protein concentration was not normalized in PAM discovery assays (PCR amplification signal provides high sensitivity for low expression or activity). The cleavage products from the TXTL reactions were recovered via clean up with AMPure SPRI beads (Beckman Coulter). The DNA was blunted via addition of Klenow fragments and dNTPs (New England Biolabs). Blunt-end products were ligated with a 100-fold excess of double stranded adapter sequences and used as template for the preparation of an NGS library, from which PAM requirements were determined from sequence analysis.

[00381] Raw NGS reads were filtered by Phred quality score > 20. The 28 bp representing the identified DNA sequence from the backbone adjacent to the PAM was used as a reference to find the PAM-proximal region and the 8 bp adjacent were identified as the putative PAM. The

distance between the PAM and the ligated adapter was also measured for each read. Reads that did not have an exact match to the reference sequence or adapter sequence were excluded. PAM sequences were filtered by cut site frequency such that PAMs with the most frequent cut site ±2 bp were included in the analysis. This correction removed low levels of background cleavage that may occur at random positions due to the use of crude *E. coli* lysate. This filtering stage can remove between 2% and 40% of the reads depending on the signal to noise ratio of the candidate protein, where less active proteins have more background signal. For reference MG29-1, 2% of reads were filtered out at this stage. The filtered list of PAMs was used to generate a sequence logo using Logomaker. These sequence logo depictions of PAMs are presented in FIGs. 20-24.

Example 5 - tracrRNA prediction and guide design

[00382] The crystal structure of a ternary complex of AacC2c1 (Cas12b) bound to a sgRNA and a target DNA reveals two separate repeat-anti-repeat (R-AR) motifs in the bound sgRNA, denoted R-AR duplex 1 and R-AR duplex 2 (see **FIG. 8** and **FIG. 9** herein and Yang, Hui, Pu Gao, Kanagalaghatta R. Rajashankar, and Dinshaw J. Patel. 2016. "PAM-Dependent Target DNA Recognition and Cleavage by C2c1 CRISPR-Cas Endonuclease." Cell 167 (7): 1814–28.e12 and Liu, Liang, Peng Chen, Min Wang, Xueyan Li, Jiuyu Wang, Maolu Yin, and Yanli Wang. 2017. "C2c1-sgRNA Complex Structure Reveals RNA-Guided DNA Cleavage Mechanism." Molecular Cell 65 (2): 310–22, each of which is incorporated by reference herein in its entirety). Putative tracrRNA sequences for the CRISPR effectors disclosed herein were identified by searching for anti-repeat sequences in the surrounding genomic context of native CRISPR arrays, where the R-AR duplex 2 anti-repeat sequence occurs ~20 - 90 nucleotides upstream of (closer to the 5' end of the tracrRNA than) the R-AR duplex 1 anti-repeat sequence. Following tracrRNA sequence identification, two guide sequences were designed for each enzyme. The first included both R-AR duplexes 1 & 2 (see for example SEQ ID NOs: 3636, 3640, 3644, 3648, 3652, 3656, 3660, 3671, and 3672), and the second was a shorter guide sequence with the R-AR duplex 1 region deleted (see e.g., SEQ ID NOs: 3637, 3641, 3645, 3649, 3653, 3657, and 3661), as this region may not be essential for cleavage.

Example 6 – Protocol for predicted RNA folding

[00383] Predicted RNA folding of RNA sequences at 37°C was computed using the method of Andronescu 2007 (which is entirely incorporated by reference herein).

Example 7 - RNA Guide Identification

[00384] For contigs that encoded a Type V-A effector and a CRISPR array, secondary structure folding of repeats indicated that the novel Type V-A systems require a single guide crRNA (sgRNA, **FIGs. 10A-D**). No tracrRNA sequences were identified. The sgRNA contained ~19-22 nt from the 3' end of the CRISPR repeat. A multiple sequence alignment of CRISPR repeats from six of the Type V-A candidates that were tested for in-vitro activity shows a highly conserved motif at the 3' end of the repeat, which formed the stem-loop structure of the sgRNA (**FIG. 10C**). The motif, UCUAC[N3-5]GUAGAU, comprised short palindromic repeats (the stem) separated by between three and five nucleotides (the loop).

[00385] The conservation of the sgRNA motif was used to uncover novel effectors that may not show similarity to classified Type V-A nucleases. Motifs were searched in repeats from 69,117 CRISPR arrays. The most common motif contained a 4-nucleotide loop, while 3- and 5-nucleotide loops were less common (see **FIG. 12**, **FIG. 13**, **FIG. 14**, **FIG. 15**, and **FIG. 16**). Inspection of the genomic context surrounding the CRISPR arrays containing the repeat motif revealed numerous effectors of varying lengths. For example, effectors of the family MG57 were the largest of the Type V-A nucleases identified (average ~1400 aa), and encoded a repeat with a 4-bp loop. Another family identified from HMM analysis contained a different repeat motif, CCUGC[N_{3-4}]GCAGG (see FIGs. 5C,5D). Although differing in sequence, the structure was predicted to fold into a highly similar stem-loop structure.

Example 8 – *In vitro* cleavage efficiency of MG CRISPR Complexes

[00386] Endonucleases are expressed as His-tagged fusion proteins from an inducible T7 promoter in a protease deficient *E. coli* B strain. Cells expressing the His-tagged proteins are lysed by sonication and the His-tagged proteins purified by Ni-NTA affinity chromatography on a HisTrap FF column (GE Lifescience) on an AKTA Avant FPLC (GE Lifescience). The eluate is resolved by SDS-PAGE on acrylamide gels (Bio-Rad) and stained with InstantBlue Ultrafast coomassie (Sigma-Aldrich). Purity is determined using densitometry of the protein band with ImageLab software (Bio-Rad). Purified endonucleases are dialyzed into a storage buffer composed of 50 mM Tris-HCl, 300 mM NaCl, 1 mM TCEP, 5% glycerol; pH 7.5 and stored at -80°C. Target DNAs containing spacer sequences and PAM sequences (determined for example as in either Example 3 or Example 4) are constructed by DNA synthesis. A single representative PAM is chosen for testing when the PAM has degenerate bases. The target DNAs are comprised of 2200 bp of linear DNA derived from a plasmid via PCR amplification with a PAM and spacer located 700 bp from one end. Successful cleavage results in fragments of 700 and 1500 bp. The target DNA, *in vitro* transcribed single RNA, and purified recombinant protein are combined in cleavage buffer (10 mM Tris, 100 mM NaCl, 10 mM $MgCl_2$) with an excess of protein and RNA

and are incubated for 5 minutes to 3 hours, usually 1 hr. The reaction is stopped via addition of RNAse A and incubation at 60 minutes. The reaction is then resolved on a 1.2% TAE agarose gel and the fraction of cleaved target DNA is quantified in ImageLab software.

Example 9 - Testing of Genome Cleavage Activity of MG CRISPR Complexes in *E. coli*

[00387] *E. coli* lacks the capacity to efficiently repair double-stranded DNA breaks. Thus, cleavage of genomic DNA can be a lethal event. Exploiting this phenomenon, endonuclease activity is tested in *E. coli* by recombinantly expressing an endonuclease and a guide RNA (determined for example as in Example 6) in a target strain with spacer/target and PAM sequences integrated into its genomic DNA (determined for example as in Example 4) integrated into their genomic DNA are transformed with DNA encoding the endonuclease. Transformants are then made chemocompetent and are transformed with 50 ng of guide RNAs (e.g., crRNAs) either specific to the target sequence ("on target"), or non-specific to the target ("non target"). After heat shock, transformations were recovered in SOC for 2 hours at 37℃. Nuclease efficiency is then determined by a 5-fold dilution series grown on induction media. Colonies are quantified from the dilution series in triplicate. A reduction in the number of colonies transformed with an on-target guide RNA compared to the number of colonies transformed with an off-target guide RNA indicates specific genome cleavage by the endonuclease.

Example 10 - Generic Procedure: Testing of Genome Cleavage Activity of MG CRISPR Complexes in Mammalian Cells

[00388] Two types of mammalian expression vectors are used to detected targeting and cleavage activity in mammalian cells. In the first, the MG Cas effector is fused to a C-terminal SV40 NLS and a viral 2A consensus cleavable peptide sequence linked to a GFP tag (the 2A-GFP tag to monitor expression of the protein). In the second, the MG Cas effector is fused to two SV40 NLS sequences, one on the N-terminus and the other on the C-terminus. The NLS sequences comprise any of the NLS sequences described herein (for example SEQ ID NOs: 3938-3953). In some instances, nucleotide sequences encoding the endonucleases are codon-optimized for expression in mammalian cells.

[00389] A single guide RNA with a crRNA sequence fused to a sequence complementary to a mammalian target DNA is cloned into a second mammalian expression vector. The two plasmids are co-transfected into HEK293T cells. 72 hours after co-transfection, DNA is extracted from the transformed HEK293T cells and used for the preparation of an NGS-library. Percent NHEJ is measured by quantifying indels at the target site to demonstrate the targeting efficiency of the

enzyme in mammalian cells. At least 10 different target sites are chosen to test each protein's activity.

Example 11 – Testing of Genome Cleavage Activity of MG CRISPR Complexes in Mammalian Cells

[00390] To show targeting and cleavage activity in mammalian cells, the MG Cas effector protein sequences were cloned into a mammalian expression vector with flanking N and C-terminal SV40 NLS sequences, a C-terminal His tag, and a 2A-GFP (e.g. a viral 2A consensus cleavable peptide sequence linked to a GFP) tag at the C terminus after the His tag (Backbone 1). In some instances, nucleotide sequences encoding the endonucleases were the native sequence, codon-optimized for expression in *E. coli* cells or codon-optimized for expression in mammalian cells.

[00391] The single guide RNA sequence (sgRNA) with a gene target of interest was also cloned into a mammalian expression vector. The two plasmids are co-transfected into HEK293T cells. 72 hours after co-transfection of the expression plasmid and a sgRNA targeting plasmid into HEK293T cells, the DNA was extracted and used for the preparation of an NGS-library. Percent NHEJ was measured via indels in the sequencing of the target site to demonstrate the targeting efficiency of the enzyme in mammalian cells. 7-12 different target sites were chosen for testing each protein's activity. An arbitrary threshold of 5% indels was used to identify active candidates. Genome editing efficiency in human cells was assessed from the NGS reads with CRISPResso using parameters: cleavage offset = -4 and window = 10. All post cleavage events from the CRISPResso output were summed for ±1 bp indels/mutations, and ≥2 bp deletions, insertions, and mutations. All outcomes were normalized to total sequences aligned to the expected amplicon (see **FIGs. 18A-E**)

Example 12 - Characterization of MG29 family

PAM Specificity, tracrRNA/sgRNA Validation

[00392] The targeted endonuclease activity of MG29 family endonuclease systems was confirmed using the myTXTL system described in Example 3 and Example 8. In this assay, PCR amplification of cleaved target plasmids yields a product that migrates at approximately 170 bp in the gel, as shown in **FIGs**. **17A-B**. Amplification products were observed for MG29-1 with crRNA corresponding to SEQ ID NO: 3609 (see **FIG**. **17A**, lane 7). Sequencing the PCR products revealed active PAM sequences for these enzymes as shown in **Table 2** below.

Table 2 : Activity of MG29-1 at various target sites

target ID	target sequence	5' sequence including PAM	locus	%NHEJ (mean ± std)
target1	TGTCAGAAGCAAATGTAAGCAATA (SEQ ID NO: 3914)	AACACAGTTG (SEQ ID NO: 3890)	HBB	2.185±0.007
target2	CTGAAAGGTTATTGTTGTGTTTGT (SEQ ID NO: 3915)	TACAGTTTTG (SEQ ID NO: 3891)	Fibrinogen	10.5±8.74
target3	CTAGTGAACACAGTTGTGTCAGAA (SEQ ID NO: 3916)	TTTGAGGTTG (SEQ ID NO: 3892)	HBB	2.14±2.83
target4	TGAAGTCTTACAAGGTTATCTTAT (SEQ ID NO: 3917)	TTTGTATTTG (SEQ ID NO: 3893)	Albumin	13.757±5.46
target5	CACTTTCCTTAGTGCGCAAAAGAA (SEQ ID NO: 3918)	AGTTACTTTG (SEQ ID NO: 3894)	Albumin	17.937±8.27
target6	GTGGTGAGGCCCTGGGCAGGTTGG (SEQ ID NO: 3919)	GATGAAGTTG (SEQ ID NO: 3895)	HBB	12.545±1.73
target7	GGAGGTCAGAAATAGGGGGTCCAG (SEQ ID NO: 3920)	TAGCTGTTTG (SEQ ID NO: 3896)	VEGFA	23.56±7.04
target8	GAAAGGGGGTGGGGGGAGTTTGCT (SEQ ID NO: 3921)	ATGGGCTTTG (SEQ ID NO: 3897)	VEGFA	30.147±10.17
target9	GTATCAAGGTTACAAGACAGGTTT (SEQ ID NO: 3922)	GGGCAGGTTG (SEQ ID NO: 3898)	HBB	10.935±1.56
target10	TGTGAGGGAGCACCGTTCTCTAGA (SEQ ID NO: 3923)	TACATAGTTG (SEQ ID NO: 3899)	Apolipoprotein	30.43±1.57
target11	GGTAGTTTTCTGTGGTCCTATTAT (SEQ ID NO: 3924)	TACGCATTTG (SEQ ID NO: 3900)	Apolipoprotein	18.173±6.28
target12	CCAGGAAAGTTGATGTGGTCTGCG (SEQ ID NO: 3925)	CCGCAAGTTG (SEQ ID NO: 3901)	Apolipoprotein	7.47±10.52

Targeted endonuclease activity in mammalian cells

[00393] MG29-1 target loci were chosen to test locations in the genome with the PAM YYn (SEQ ID NO: 3871). The spacers corresponding to the chosen target sites were cloned into the sgRNA scaffold in the mammalian vector system backbone 1 described in Example 9. The sites are listed in **Table 3** below. The activity of MG29-1 at various target sites is shown in **Table 2** and **FIG. 19**.

Table 3: 5′ PAM Sequences and crRNAs for Enzymes Described Herein

Enzyme	Enzyme SEQ ID NO:	5′ PAM	PAM SEQ ID NO:	crRNA SEQ ID NO:
MG29-1	215	KTTG	3870	3608

Example 13 – High-replicate PAM determination via NGS

[00394] Type V endonucleases (e.g. MG28, MG29, MG30, MG31 endonucleases) were tested for cleavage activity using E coli lysate-based expression in the myTXTL kit as described in Example 3 and Example 8. Upon incubation with a crRNA and a plasmid library containing a

spacer sequencing matching the crRNA preceded by 8 degenerated ("N") bases (a 5' PAM library), the subset of the plasmid library with a functional PAM was cleaved. Ligation to this cut site and PCR amplification provided evidence of activity, demonstrated by the bands observed in the gel at 170 bp (**FIG. 17B**). Gel 1 (top panel, A) lanes are as follows: 1 (ladder; darkest band corresponds to 200 bp); 2: positive control (previously verified library); 3 (n/a); 4 (n/a); 5 (MG28-1); 6 (MG29-1); 7 (MG30-1); 8 (MG31-1); 9 (MG32-1); and 10 (Ladder). Gel 2 (bottom panel, B) lanes are as follows: 1 (ladder; darkest band corresponds to 200 bp); 2 (LbCpf1 positive control); 3 (LbCpf1 positive control); 4 (negative control); 5 (n/a); 6 (n/a); 7 (MG28-1); 8 (MG29-1); 9 (MG30-1); 10 (MG31-1); 11 (MG32-1).

[00395] The PCR products were further subjected to NGS sequencing and the PAMs were collated into seqLogo (see e.g., Huber et al. Nat Methods. 2015 Feb;12(2):115-21, which is incorporated by reference herein) representations (**FIG. 20**). The seqLogo representation shows the 8 bp which are upstream (5′) of the spacer labeled as positions 0-7. As shown in the **FIG. 20**, the PAMs are pyrimidine rich (C and T), with most sequence requirements 2-4 bp upstream of the spacer (positions 4-6 in the SeqLogo).

[00396] The PAMs for the MG candidates are shown in **Table 4** below.

Table 4 : 5′ PAM Sequences and crRNAs for Enzymes Described Herein

Enzyme	Enzyme SEQ ID NO:	5′ PAM	PAM SEQ ID NO:	crRNA SEQ ID NO:
MG28-1	141	TTTn	3868	3609
MG29-1	215	YYn	3871	3609
MG31-1	229	YTTn	3875	3609
MG32-1	261	TTTn	3877	3609

[00397] In some cases, the position immediately adjacent to the spacer may have a weaker specificity, e.g. for "m" or "v" instead of "n".

Example 14 - Targeted endonuclease activity in mammalian cells with MG31 nucleases

Targeted endonuclease activity in mammalian cells

[00398] MG31-1 target loci were chosen to test locations in the genome with the PAM TTTR (SEQ ID NO: 3875). The spacers corresponding to the chosen target sites were cloned into the sgRNA scaffold in the mammalian vector system backbone 1 described in Example 11. The sites are listed in **Table 5** below. The activity of MG31-1 at various target sites is shown in **Table 5** and **FIG. 25**.

Table 5: Activity of MG31-1 at various target sites

target ID	target sequence	PAM	locus	%NHEJ (mean ± std)
target1	GTTATTAATTTCTTGCTACTTGTC (SEQ ID NO: 3926)	GTTTTCTTTA (SEQ ID NO: 3902)	Fibrinogen	1.005±0.516
target2	CTGAAAGGTTATTGTTGTGTTTGT (SEQ ID NO: 3927)	TACAGTTTTG (SEQ ID NO: 3903)	Fibrinogen	2.417±1.47
target3	GTGTTAGTACAGTTTTGCTGAAAG (SEQ ID NO: 3928)	AGAACTTTTA (SEQ ID NO: 3904)	Fibrinogen	2.925±0.516
target4	TGAAGTCTTACAAGGTTATCTTAT (SEQ ID NO: 3929)	TTTGTATTTG (SEQ ID NO: 3905)	Albumin	7.053±2.72
target5	CACTTTCCTTAGTGCGCAAAAGAA (SEQ ID NO: 3930)	AGTTACTTTG (SEQ ID NO: 3906)	Albumin	0.927±0.50
target6	CCTAGGATGTTTGAATTTTATTAA (SEQ ID NO: 3931)	TTTTTTTTTA (SEQ ID NO: 3907)	Albumin	1.125±0.43
target7	GGAGGTCAGAAATAGGGGGTCCAG (SEQ ID NO: 3932)	TAGCTGTTTG (SEQ ID NO: 3908)	VEGFA	17.39±8.67
target8	GAAAGGGGGTGGGGGGAGTTTGCT (SEQ ID NO: 3933)	ATGGGCTTTG (SEQ ID NO: 3909)	VEGFA	4.01±1.29
target9	GCCAGAGCCGGGGTGTGCAGACGG (SEQ ID NO: 3934)	TCCCTCTTTA (SEQ ID NO: 3910)	VEGFA	6.72±1.92
target10	CTTGGACCTTGTTTTGCTTACTGT (SEQ ID NO: 3935)	ACAAATTTTA (SEQ ID NO: 3911)	Apolipoprotein	-0.32±0.75
target11	GGTAGTTTTCTGTGGTCCTATTAT (SEQ ID NO: 3936)	TACGCATTTG (SEQ ID NO: 3912)	Apolipoprotein	2.593±1.33
target12	ATCATAAGAAGTTAGCTTGACGCA (SEQ ID NO: 3937)	GAAAAATTTA (SEQ ID NO: 3913)	Apolipoprotein	3.095

Example 3 – *In Vitro* activity

[00399] Promising candidates from the bioinformatic analysis and preliminary screens were selected for further biochemical analysis as described in this example. Using the conserved 3' sgRNA structure, a "universal" sgRNA was designed comprising the 3' 20 nt of the CRISPR repeat and a 24 nt spacer (**FIGs. 10A-D**). Of the seven tested candidates, six showed activity *in vitro* against the 8N PAM library (**FIG. 26A**). The remaining inactive candidate (30-1) showed activity when tested with its predicted endogenous trimmed CRISPR repeat (SEQ ID NO: 3608, see **FIG. 26B**), but was not included in NGS library assays. (**FIG. 26C**)

[00400] The majority of identified PAMs are thymine-rich sequences of 2-3 bases (**FIG. 18A**). However, two enzymes, MG26-1 (PAM YYn) and MG29-1 (PAM YYn), had PAM specificity for either pyrimidine base, thymine or cytosine, allowing for broader sequence targeting. Analysis of putative PAM-interacting residues indicated that the active Type V-A nucleases contain a conserved Lysine and a GWxxxK motif, which were shown to be important in recognition and interaction with different PAMs in FnCas12a.

[00401] As our PAM detection assay required ligation to create blunt-end fragments before PAM enrichment, this suggested that these enzymes created a staggered double strand DNA break, similar to reported Type V-A nucleases. The cut site on the target strand can be identified by analysis of the NGS reads used for indel detection (**FIG. 18B**) and showed cleavage after the 22nd PAM-distal base

[00402] *In vitro* cleavage by MG29-1 was further investigated by sequencing the cleavage products. The cut position on the target strand was 22 nucleotides away from the PAM in most sequences, and 21 or 23 nucleotides less frequently (**FIGs. 56A-C**). The cut position on the non-target strand was 17 to 19 nucleotides from the PAM. In combination, these results indicate a 3-5 bp overhang.

Example 4 – Genome Editing

[00403] After confirmation of the PAM, novel proteins described herein were tested in HEK293T cells for gene targeting activity. All candidates showed activity of over 5% NHEJ (background corrected) on at least one of ten tested target loci. MG29-1 showed the highest overall activity in NHEJ modification outcomes (**FIG. 18B**) and was active on the highest number of targets. Thus, this nuclease was selected for purified ribonucleoprotein complex (RNP) testing in HEK293 cells. RNP transfection of MG29-1 holoenzyme showed higher editing levels with RNP than plasmid-based transfection on 4 out of 9 targets, in some cases over 80% editing efficiency (**FIG. 18C**). Analysis of editing profiles for MG29-1 indicates that this nuclease produces deletions of more than two bp more frequently than other types of edits at their target site (**FIG. 18D**). At some targets (5 and 8) the indel frequency for MG29-1 was twice that of AsCpf1 (**FIG. 18E**).

Example 17 – Discussion

[00404] Type V-A CRISPR were identified from metagenomes collected from a variety of complex environments and arranged into families. These novel Type V-A nucleases had diverse sequences and phylogenetic origins within and across families and cleaved targets with diverse PAM sites. Similar to other Type V-A nucleases (e.g. LbCas12a, AsCas12a, and FnCas12a), the

effectors described herein utilized a single guide CRISPR RNA (sgRNA) to target staggered double stranded cleavage of DNA, simplifying guide design and synthesis, which will facilitate multiplexed editing. Analysis of CRISPR repeat motifs that formed the stem-loop structure of the crRNA suggested that the Type V-A effectors described herein have a 4-nt loop guide more frequently than shorter or longer loops. The sgRNA motif of LbCpf1 has a less common 5-nt, although the 4-nt loop was also observed for 16 Cpf1 orthologs already identified. An unusual stem-loop CRISPR repeat motif sequence, CCUGC[N_{3-4}]GCAGG, was identified for the MG61 family of Type V-A effectors. The high degree of conservation of the sgRNA with variable loop lengths in Type V-A may afford flexible levels of activity, as shown for proteins described herein. Taken together, these effectors are not close homologs to previously studied enzymes, and greatly expand the diversity of Type V-A-like sgRNA nucleases.

[00405] Additional Type V effectors described herein may have evolved from duplications of Type V-A-like nucleases, referred to here as Type V-A prime effectors (V-A') which may be encoded next to Cas12a nucleases. Both Type V-A and these Type V-A' systems may share a CRISPR sgRNA but the Type V-A' systems are divergent from Cas12a (**FIG. 4**). The CRISPR repeat associated with these prime effectors also folded into single guide crRNA with the UCUAC[N_{3-5}]GUAGAU motif. One report identified a Type V cms1 effector encoded next to a Type V-A nuclease, which required a single guide crRNA for cleavage activity in plant cells. Different CRISPR arrays were reported for each effector, while the Type V-A' system described herein suggested that both Type V-A and V-A' may require the same crRNA for DNA targeting and cleavage. As described recently in Roizmanbacterial genomes (see e.g., Chen et al. Front Microbiol. 2019 May 3;10:928), both Type V-A and V-A' effectors are distantly related based on sequence homology and phylogenetic analysis. Therefore, the prime effectors do not belong within the Type V-A classification, and warrant a separate Type V sub-classification

[00406] PAMs determined for active Type V-A nucleases were generally thymine-rich, similar to PAMs described for other Type V-A nucleases. In contrast, MG29-1 requires a shorter YYN PAM sequence, which increases target flexibility compared to the four nucleotide TTTV PAM of LbCpf1. Additionally, RNPs containing MG29-1 had higher activity in HEK293 cells compared to sMbCas12a, which has a three-nucleotide PAM. .

[00407] When testing the novel nucleases for in-vitro editing activity, MG29-1 exhibited comparable or better activity to other reported enzymes of the class. Reports of plasmid transfection editing efficiencies in mammalian cells using Cas12a orthologs indicate between 21% and 26% indel frequencies for guides with T-rich PAMs, and one out of 18 guides with CCN PAMs showed ~10% activity in Mb3Cas12a (*Moraxella bovoculi* AAX11_00205 Cas12a,

see e.g. Wang et al. Journal of Cell Science 2020 133: jcs240705). Notably, MG29-1 activity in plasmid transfections appears greater than that reported for Mb3Cas12a for targets with TTN and CCN PAMs (see e.g. **FIGs. 18A-E**). Because the target sites for plasmid transfections have the same TTG PAM on all experiments, the difference in editing efficiency may be attributed to genomic accessibility differences at different target genes. MG29-1 editing as RNP is much more efficient than via plasmid and is more efficient than AsCas12a on two of seven target loci. Therefore, MG29-1 may be a highly active and efficient gene editing nuclease. These findings increase the diversity of identified single guide Type V-A CRISPR nucleases, and demonstrate the genome editing potential of novel enzymes from uncultivated microbes. Seven novel nucleases showed in-vitro activity with diverse PAM requirements, and RNP data showed editing efficiency surpassing 80% for therapeutically relevant targets in human cell lines. These novel nucleases expand the toolkit of CRISPR-associated enzymes and enable diverse genome engineering applications.

Example 18 – MG29-1 induced editing of TRAC locus in T-cells

[00408] The three exons of the T cell receptor alpha chain constant region (TRACA) were scanned for sequences matching an initial predicted 5'-TTN-3' PAM specificity of MG29-1 and single-guide RNAs with proprietary Alt-R modifications were ordered from IDT. All guide spacer sequences were 22 nt long. Guides (80 pmol) were mixed with purified MG29-1 protein (63 pmol), incubated for 15 minutes at room temperature. T cells were purified from PBMCs by negative selection using (Stemcell Technologies Human T cell Isolation Kit #17951) and activated by CD2/3/28 beads (Miltenyi T cell Activation/Expansion Kit #130-091-441). After four days of cell growth, each MG29-1/guide RNA mixture was electroporated into 200,000 T cells with a Lonza 4-D Nucleofector, using program EO-115 and P3 buffer. The cells were harvested seventy-two hours post-transfection, genomic DNA was isolated, and PCR amplified for analysis using high-throughput DNA sequencing using primers targeting the TRACA locus. The creation of insertions and deletions characteristic of NHEJ-based gene editing was quantified using a proprietary Python script (see **FIG. 39**).

Table 5A: Guide sequences used in Example 18

Entity Name	Sequence	SEQ ID NO:
MG29-1 Guide 1 Target Sequence	ACCGATTTTGATTCTCAAACAA	4316
MG29-1 Guide 2 Target Sequence	TGATTCTCAAACAAATGTGTCA	4317
MG29-1 Guide 3 Target Sequence	GATTCTCAAACAAATGTGTCAC	4318

Entity Name	Sequence	SEQ ID NO:
MG29-1 Guide 4 Target Sequence	ATTCTCAAACAAATGTGTCACA	4319
MG29-1 Guide 5 Target Sequence	TCAAACAAATGTGTCACAAAGT	4320
MG29-1 Guide 6 Target Sequence	TGATGTGTATATCACAGACAAA	4321
MG29-1 Guide 7 Target Sequence	AAGAGCAACAGTGCTGTGGCCT	4322
MG29-1 Guide 8 Target Sequence	GCATGTGCAAACGCCTTCAACA	4323
MG29-1 Guide 9 Target Sequence	CATGTGCAAACGCCTTCAACAA	4324
MG29-1 Guide 10 Target Sequence	AACAACAGCATTATTCCAGAAG	4325
MG29-1 Guide 11 Target Sequence	TTCCAGAAGACACCTTCTTCCC	4326
MG29-1 Guide 12 Target Sequence	CAGAAGACACCTTCTTCCCCAG	4327
MG29-1 Guide 13 Target Sequence	TGGAATAATGCTGTTGTTGAAG	4328
MG29-1 Guide 14 Target Sequence	TTGAAGGCGTTTGCACATGCAA	4329
MG29-1 Guide 15 Target Sequence	AAGGCGTTTGCACATGCAAAGT	4330
MG29-1 Guide 16 Target Sequence	GCACATGCAAAGTCAGATTTGT	4331
MG29-1 Guide 17 Target Sequence	CACATGCAAAGTCAGATTTGTT	4332
MG29-1 Guide 18 Target Sequence	GTTGCTCCAGGCCACAGCACTG	4333
MG29-1 Guide 19 Target Sequence	TTGCTCCAGGCCACAGCACTGT	4334
MG29-1 Guide 20 Target Sequence	CTCCAGGCCACAGCACTGTTGC	4335
MG29-1 Guide 21 Target Sequence	CTCTTGAAGTCCATAGACCTCA	4336
MG29-1 Guide 22 Target Sequence	AAGTCCATAGACCTCATGTCTA	4337
MG29-1 Guide 23 Target Sequence	TGTCTGTGATATACACATCAGA	4338
MG29-1 Guide 24 Target Sequence	GTCTGTGATATACACATCAGAA	4339
MG29-1 Guide 25 Target Sequence	TCTGTGATATACACATCAGAAT	4340
MG29-1 Guide 26 Target Sequence	CTTTGTGACACATTTGTTTGAG	4341
MG29-1 Guide 27 Target Sequence	GTGACACATTTGTTTGAGAATC	4342
MG29-1 Guide 28 Target Sequence	TGACACATTTGTTTGAGAATCA	4343
MG29-1 Guide 29 Target Sequence	GTTTGAGAATCAAAATCGGTGA	4344
MG29-1 Guide 30 Target Sequence	TTTGAGAATCAAAATCGGTGAA	4345
MG29-1 Guide 31 Target Sequence	GAGAATCAAAATCGGTGAATAG	4346

Entity Name	Sequence	SEQ ID NO:
MG29-1 Guide 32 Target Sequence	AGAATCAAAATCGGTGAATAGG	4347
MG29-1 Guide 33 Target Sequence	TCACTGGATTTAGAGTCTCTCA	4348
MG29-1 Guide 34 Target Sequence	AGAGTCTCTCAGCTGGTACACG	4349
MG29-1 Guide 35 Target Sequence	GAGTCTCTCAGCTGGTACACGG	4350
MG29-1 Guide 36 Target Sequence	CTGTGATGTCAAGCTGGTCGAG	4351
MG29-1 Guide 37 Target Sequence	CAAAGCTTTTCTCGACCAGCTT	4352
MG29-1 Guide 38 Target Sequence	AAAGCTTTTCTCGACCAGCTTG	4353
MG29-1 Guide 39 Target Sequence	TCTCGACCAGCTTGACATCACA	4354
MG29-1 Guide 40 Target Sequence	CTCGACCAGCTTGACATCACAG	4355
MG29-1 Guide 41 Target Sequence	TCGACCAGCTTGACATCACAGG	4356
MG29-1 Guide 42 Target Sequence	CAAAACCTGTCAGTGATTGGGT	4357
MG29-1 Guide 43 Target Sequence	AAAACCTGTCAGTGATTGGGTT	4358
MG29-1 Guide 44 Target Sequence	GGTTCCGAATCCTCCTCCTGAA	4359
MG29-1 Guide 45 Target Sequence	CGAATCCTCCTCCTGAAAGTGG	4360
MG29-1 Guide 46 Target Sequence	AATCTGCTCATGACGCTGCGGC	4361
MG29-1 Guide 47 Target Sequence	ATCTGCTCATGACGCTGCGGCT	4362
MG29-1 Guide 48 Target Sequence	AACCCGGCCACTTTCAGGAGGA	4363
MG29-1 Guide 49 Target Sequence	CAGGAGGAGGATTCGGAACCCA	4364
MG29-1 Guide 50 Target Sequence	AGGAGGAGGATTCGGAACCCAA	4365
MG29-1 Guide 51 Target Sequence	GGAACCCAATCACTGACAGGTT	4366
MG29-1 Guide 52 Target Sequence	TGAAAGTTTAGGTTCGTATCTG	4367
MG29-1 Guide 53 Target Sequence	GAAAGTTTAGGTTCGTATCTGT	4368
MG29-1 Guide 54 Target Sequence	AAAGTTTAGGTTCGTATCTGTA	4369
MG29-1 Guide 1 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrCrGrArUrUrUrUrGrArUrUrCrUrCrArArArCrArA/AltR2/	4370
MG29-1 Guide 2 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrArUrUrCrUrCrArArArCrArArArUrGrUrGrUrCrA/AltR2/	4371
MG29-1 Guide 3 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArUrUrCrUrCrArArArCrArArArUrGrUrGrUrCrArC/AltR2/	4372

Entity Name	Sequence	SEQ ID NO:
MG29-1 Guide 4 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrUrCrUrCrArArArCrArArArUrGrUrGrUrCrArCrA/AltR2/	4373
MG29-1 Guide 5 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrArArArCrArArArUrGrUrGrUrCrArCrArArArGrU/AltR2/	4374
MG29-1 Guide 6 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrArUrGrUrGrUrArUrArUrCrArCrArGrArCrArArA/AltR2/	4375
MG29-1 Guide 7 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArGrArGrCrArArCrArGrUrGrCrUrGrUrGrGrCrCrU/AltR2/	4376
MG29-1 Guide 8 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrArUrGrUrGrCrArArArCrGrCrCrUrUrCrArArCrA/AltR2/	4377
MG29-1 Guide 9 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArUrGrUrGrCrArArArCrGrCrCrUrUrCrArArCrArA/AltR2/	4378
MG29-1 Guide 10 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArCrArArCrArGrCrArUrUrArUrUrUrCrCrArGrArArG/AltR2/	4379
MG29-1 Guide 11 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrCrCrArGrArArGrArArCrCrUrUrCrUrUrCrCrC/AltR2/	4380
MG29-1 Guide 12 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrArArGrArCrArCrCrUrUrCrUrUrCrCrCrCrArG/AltR2/	4381
MG29-1 Guide 13 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrGrArArUrArArUrGrCrUrGrUrUrGrUrUrGrArArG/AltR2/	4382
MG29-1 Guide 14 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrArGrGrCrGrUrUrUrGrCrArCrArUrGrCrArA/AltR2/	4383
MG29-1 Guide 15 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArGrGrCrGrUrUrUrGrCrArCrArUrGrCrArArArGrU/AltR2/	4384
MG29-1 Guide 16 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrArCrArUrGrCrArArArGrUrCrArGrArUrUrUrGrU/AltR2/	4385
MG29-1 Guide 17 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArCrArCrArUrGrCrArArArGrUrCrArGrArUrUrUrGrU/AltR2/	4386
MG29-1 Guide 18 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrUrGrCrCrUrCrCrArGrGrCrCrArArCrArCrUrG/AltR2/	4387
MG29-1 Guide 19 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrCrUrCrCrArGrGrCrCrArCrArGrCrArCrUrGrU/AltR2/	4388
MG29-1 Guide 20 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrCrCrCrArGrGrCrCrArCrArGrCrUrGrUrGrC/AltR2/	4389
MG29-1 Guide 21 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrCrUrGrArArGrUrCrCrArUrArGrArCrCrUrCrA/AltR2/	4390
MG29-1 Guide 22 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArGrUrCrCrArUrArGrArCrCrUrCrUrGrUrCrUrA/AltR2/	4391

Entity Name	Sequence	SEQ ID NO:
MG29-1 Guide 23 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrUrCrUrGrUrGrArUrArUrArCrArCrArUrCrArGrA/AltR2/	4392
MG29-1 Guide 24 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrCrUrGrUrGrArUrArUrArCrArCrArUrCrArGrArA/AltR2/	4393
MG29-1 Guide 25 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrGrUrGrArUrArUrArCrArCrArUrCrArGrArArU/AltR2/	4394
MG29-1 Guide 26 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrUrGrUrGrArCrArCrArUrUrUrGrUrUrUrGrArG/AltR2/	4395
MG29-1 Guide 27 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrGrArCrArCrArUrUrUrGrUrUrUrGrArGrArArUrC/AltR2/	4396
MG29-1 Guide 28 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrArCrArCrArUrUrUrGrUrUrUrGrArGrArArUrCrA/AltR2/	4397
MG29-1 Guide 29 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrUrUrGrArGrArArArUrCrGrGrUrGrA/AltR2/	4398
MG29-1 Guide 30 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrGrArGrArArArUrCrGrGrUrGrArA/AltR2/	4399
MG29-1 Guide 31 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArGrArArArUrCrGrGrUrGrArArUrArG/AltR2/	4400
MG29-1 Guide 32 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrArArArUrCrGrGrUrGrArArUrArGrG/AltR2/	4401
MG29-1 Guide 33 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrArCrUrGrGrArArGrUrCrUrCrUrCrA/AltR2/	4402
MG29-1 Guide 34 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrArGrUrCrUrCrArGrCrUrGrGrUrArCrArCrG/AltR2/	4403
MG29-1 Guide 35 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArGrUrCrUrCrUrCrArGrCrUrGrGrUrArCrArCrGrG/AltR2/	4404
MG29-1 Guide 36 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrUrGrUrGrArUrGrUrCrArArGrCrUrGrGrUrCrGrArG/AltR2/	4405
MG29-1 Guide 37 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArArArGrCrUrUrUrUrCrUrCrGrArCrCrArGrCrUrU/AltR2/	4406
MG29-1 Guide 38 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArArGrCrUrUrUrUrCrUrCrGrArCrCrArGrCrUrUrG/AltR2/	4407
MG29-1 Guide 39 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrCrGrArCrCrArGrCrUrUrGrArCrUrCrArCrA/AltR2/	4408
MG29-1 Guide 40 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrCrGrArCrCrArGrCrUrUrGrArCrUrCrArCrArG/AltR2/	4409
MG29-1 Guide 41 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrGrArCrCrArGrCrUrUrGrArCrUrCrArCrArGrG/AltR2/	4410

Entity Name	Sequence	SEQ ID NO:
MG29-1 Guide 42 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArArArArCrCrUrGrUrCrArGrUrGrArUrUrGrGrGrU/AltR2/	4411
MG29-1 Guide 43 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArArArCrCrUrGrUrCrArGrUrGrArUrUrGrGrGrUrU/AltR2/	4412
MG29-1 Guide 44 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrUrUrCrCrGrArArUrCrCrUrCrCrUrCrUrGrArA/AltR2/	4413
MG29-1 Guide 45 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrGrArArUrCrCrUrCrCrUrCrCrUrGrArArArGrUrGrG/AltR2/	4414
MG29-1 Guide 46 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArUrCrUrGrCrUrCrArUrGrArCrGrCrUrGrCrGrGrC/AltR2/	4415
MG29-1 Guide 47 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrCrUrGrCrUrCrArUrGrArCrGrCrUrGrCrGrGrCrU/AltR2/	4416
MG29-1 Guide 48 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArCrCrCrGrGrCrCrArCrUrUrUrCrArGrGrArArGrGrA/AltR2/	4417
MG29-1 Guide 49 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrGrArGrArGrGrArGrGrArArUrUrCrGrGrArArCrCrCrA/AltR2/	4418
MG29-1 Guide 50 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrGrArGrGrArGrGrArArUrUrCrGrGrArArCrCrCrArA/AltR2/	4419
MG29-1 Guide 51 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrArArCrCrCrArArUrCrArCrArGrGrUrU/AltR2/	4420
MG29-1 Guide 52 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrArArArGrGrUrUrCrGrUrArUrCrUrG/AltR2/	4421
MG29-1 Guide 53 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArArArGrGrUrUrCrGrUrArUrCrUrGrU/AltR2/	4422
MG29-1 Guide 54 sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArArGrGrUrUrCrGrUrArUrCrUrGrUrA/AltR2/	4423

Example 19 – Re-testing of lead guides of MG29-1

[00409] An experiment retesting the lead guides for MG29-1 was performed. The three exons of the T cell receptor alpha chain constant region were scanned for sequences matching 5'-TTN-3' and single-guide RNAs ordered from IDT using Alt-R modifications. All guide spacer sequences 22 nt long. Guides were mixed with purified MG29-1 protein (80 pmol gRNA + 63 pmol MG29-1; or 160 pmol gRNA with 126 pmol MG29-1), incubated for 15 minutes at room temperature. T cells were purified from PBMCs by negative selection using (Stemcell Technologies Human T cell Isolation Kit #17951) and activated by CD2/3/28 beads (Miltenyi T cell Activation/Expansion Kit #130-091-441). After four days of cell growth, each MG29-1/guide

RNA mixture was electroporated into 200,000 T cells with a Lonza 4-D Nucleofector, using program EO-115 and P3 buffer. Seventy-two hours post-transfection, genomic DNA was harvested, and PCR amplified for analysis using high-throughput DNA sequencing. The creation of insertions and deletions characteristic of NHEJ-based gene editing was quantified using a proprietary Python script (see **FIG. 40**).

Example 20 – Testing length of guide spacer for MG29-1

[00410] An experiment was performed to determine the optimal guide spacer length. The three exons of the T cell receptor alpha chain constant region were scanned for sequences matching 5'-TTN-3' and single-guide RNAs ordered from IDT using Alt-R modifications. Guides were mixed with purified MG29-1 protein (80 pmol gRNA + 60 pmol effector; 160 pmol gRNA + 120 pmol effector; or 320 pmol gRNA + 240 pmol effector), incubated for 15 minutes at room temperature. T cells were purified from PBMCs by negative selection using (Stemcell Technologies Human T cell Isolation Kit #17951) and activated by CD2/3/28 beads (Miltenyi T cell Activation/Expansion Kit #130-091-441). After four days of cell growth, each MG29-1/guide RNA mixture was electroporated into 200,000 T cells with a Lonza 4-D Nucleofector, using program EO-115 and P3 buffer. Seventy-two hours post-transfection, genomic DNA was harvested, and PCR amplified for analysis using high-throughput DNA sequencing. The creation of insertions and deletions characteristic of NHEJ-based gene editing was quantified using a proprietary Python script. The results are shown in **FIG. 41**, which demonstrates that guide spacer lengths of 20-24 nt work well, with a dropoff at 19 nt.

Example 21 – Determination of MG29-1 indel generation versus TCR expression

[00411] Cells from **FIG. 41** were analyzed for TCR expression by flow cytometry using the APC-labeled anti-human TCRα/β Ab (Biolegend #306718, clone IP26) and an Attune NxT flow cytometer (Thermo Fisher). Indel data are taken from **FIG. 41**.

Example 22 – Targeted CAR integration with MG29-1

[00412] The three exons of the T cell receptor alpha chain constant region were scanned for sequences matching 5'-TTN-3' and single-guide RNAs ordered from IDT using IDT's proprietary Alt-R modifications. Guides (80 pmol) were mixed with purified MG29-1 protein (63 pmol), incubated for 15 minutes at room temperature. T cells were purified from PBMCs by negative selection using (Stemcell Technologies Human T cell Isolation Kit #17951) and activated by CD2/3/28 beads (Miltenyi T cell Activation/Expansion Kit #130-091-441). After four days of cell growth, each MG29-1/guide RNA mixture was electroporated into 200,000 T

cells with a Lonza 4-D Nucleofector, using program EO-115 and P3 buffer. 100,000 vector genomes of a serotype 6 adeno-associated virus (AAV-6) containing the coding sequence for a customized chimeric antigen receptor flanked by 5' and 3' homology arms (5' arm SEQ ID NO: 4424 being about 500 nt in length and 3' arm SEQ ID NO: 4425 being about 500 nt in length) targeting the TRAC gene were added to the cells immediately following transfection. Replicates were analyzed for TCR expression versus TRAC indels (**FIG. 42**), showing that indels in the TRAC gene correlated with loss of expression of TCR. Cells were also analyzed by flow cytometry simultaneously for TCR expression as in Example 21 (**FIG. 42**) and for binding of the target antigen to the CAR (**FIG. 43**, in which the plots are gated on single, live cells). The results of the flow analysis in FIG. 43 indicated that while the guide RNAs alone were effective in eliminating TCR expression ("RNP only"), addition of guide RNA plus AAV resulted in a new population of cells binding the CAR antigen (top left of plots "AAV+MG29-1-19-22" and "AAV+MG29-1-35-22"). The sgRNA 35 (SEQ ID NO: 4404) was somewhat more effective in inducing integration of the CAR than sgRNA 19 (SEQ ID NO: 4388). One possible explanation for the difference is that the predicted nuclease cut site for Guide 19 is ~160 bp away from the end of the right homology arm.

Table 5B: Guide RNAs used in Example 22

Entity Name	Sequence	SEQ ID NO:
MG29-1 TRAC Guide 19 Target-binding Sequence	TTGCTCCAGGCCACAGCACTGT	4334
MG29-1 TRAC Guide 19 full sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrCrUrCrCrArGrGrCrCrArCrArGrCrArCrUrGrU/AltR2/	4388
MG29-1 TRAC Guide 35 Target-binding Sequence	GAGTCTCTCAGCTGGTACACGG	4350
MG29-1 TRAC Guide 35 full sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArGrUrCrUrCrUrCrArGrCrUrGrGrUrArCrArCrGrG/AltR2/	4404
/AltR1/ and /AltR2/ refer to IDT's proprietary Alt-R 5' and 3' modifications; m; 2'-O-methyl base (for example an A base with 2'-O-methyl modification is written as mA), i2F; internal 2'-flourine base (for example an internal C with 2'-flourine modification is written as /i2FC/), 52F; 2'-flourine base at the 5' end of the sequence (for example a 5' C with 2'-flourine modification is written as /52FC/), 32F; 2'-flourine base at the 3' end of the sequence (for example a 3' A base with 2'-flourine modification is written as /32FA/), r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *; between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur		

Example 5 – MG29-1 TRAC editing in HSCs

[00413] Hematopoietic stem cells were purchased from Allcells and thawed per the supplier's instructions, washed in DMEM + 10% FBS, and resuspended in Stemspan II medium plus CC110 cytokines. One million cells were cultured for 72 hours in a 6-well dish in 4 mL medium. MG29-1 RNPs were made, transfected, and gene editing analyzed as in Example 18 except for use of the EO-100 nucleofection program. The results are shown in **FIG. 61**, which shows gene editing at TRAC in hematopoietic stem cells using the #19 (SEQ ID NO:4388) and #35 (SEQ ID NO: 4404) sgRNAs in Table 5B below. The results again indicate that the #35 sgRNA is highly effective at targeting the TRAC locus.

Table 5C: Guide RNAs used in Example 23

Entity Name	Sequence	SEQ ID NO:
MG29-1 TRAC Guide 19 Target-binding Sequence	TTGCTCCAGGCCACAGCACTGT	4334
MG29-1 TRAC Guide 19 full sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrCrUrCrCrArGrGrCrCrArCrArGrCrArCrUrGrU/AltR2/	4388
MG29-1 TRAC Guide 35 Target-binding Sequence	GAGTCTCTCAGCTGGTACACGG	4350
MG29-1 TRAC Guide 35 full sgRNA synthesized	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArGrUrCrUrCrUrCrArGrCrUrGrGrUrArCrArCrGrG/AltR2/	4404
/AltR1/ and /AltR2/ refer to IDT's proprietary Alt-R 5' and 3' modifications; m; 2'-O-methyl base (for example a A base with 2'-O-methyl modification is written as mA), i2F; internal 2'-flourine base (for example an internal C with 2'-flourine modification is written as /i2FC/), 52F; 2'-flourine base at the 5' end of the sequence (for example a 5' C with 2'-flourine modification is written as /52FC/), 32F; 2'-flourine base at the 3' end of the sequence (for example a 3' A base with 2'-flourine modification is written as /32FA/), r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *; between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur		

Example 6 – Further analysis of PAM specificity associated with MG29-1

[00414] Further analysis was performed to determine more precisely the PAM specificity of MG29-1. Guide RNAs were designed using a 5'-NTTN-3' PAM sequence and then sorted according to the gene editing activity observed (**FIG. 45**, in which the identity of the underlined base--the 5'-proximal N is shown for each bin). All of the guides with activity greater than 10% had a T at this position in the genomic DNA indicating that the MG29-1 PAM may be better described as 5'-TTTN-3'. The statistical significance of the over-representation of T at this position is shown for each bin. In **FIG. 45**, the various bins (High, medium, low, >1%, <1%) signify:

- High: >50% indels (N = 4)
- Medium: 10-50% indels (N = 15)

- Low: 5-10% indels (N = 5)
- >1%: 1-5% indels (N = 12)
- <1% (N = 82)

Table 2 – p-values for nucleotide specificity analysis in Example 24

	chi^2 p-value
high/med	**0.000005**
low	**0.035110**
>1%	**0.005416**
<1%	**0.126751**

Example 7 – Determining MG29-1 indel induction ability vs spacer base composition

[00415] Further analysis was conducted of gene editing activity versus the base composition of MG29-1 spacer sequences. The correlation was modest (R^2 = 0.23) but there is a trend towards better activity with higher GC content (see **FIG. 46**, in which correlation between indels induced in cultured cells versus GC content of spacer sequences is presented as a dot plot).

Example 26 – MG29-1 guide chemistry modifications

[00416] An experiment to optimize chemical modifications for targeting of the VEGF-A locus using MG29-1 was performed, using the procedure of Example 18 but with the indicated guide RNAs targeting VEGF-A (see Table 7 below). The experiment used 126 pmol MG29-1 and 160 pmol guide RNA. The results are presented in **FIG. 47**. Guides #4, 5, 6, 7, and 8 showed improved activity versus the unmodified guide #1, indicating that the corresponding modifications in these sequences improved the activity of these guide RNAs versus an unmodified RNA sequence.

Table 7 – MG29-1 guide modifications

MG29-1 Test No.	MG29-1 guide with targeting sequence in bold and modifications per legend below	SEQ ID NO:
1	UAAUUUCUACUCUUGUAGAUGAAAGGGGGTGGGGGGAGTTTGCT	3985
2	mU*mA*mA*UUUCUACUCUUGUAGAUGAAAGGGGGTGGGGGGAGTTT*mG*mC*mT	3986
3	mU*mA*AUUUCUACUCUUGUAGAUGAAAGGGGGTGGGGGGAGTTT*mG*mC*mT	3987
4	mU*AAUUUCUACUCUUGUAGAUGAAAGGGGGTGGGGGGAGTTT*mG*mC*mT	3988

MG29-1 Test No.	MG29-1 guide with targeting sequence in bold and modifications per legend below	SEQ ID NO:
5	mU*AAUUUCUACUCUUGUAGAUGAAAGGGGGTGGGGGGAGTTTGC*mT	3989
6	mC*UAAUUUCUACUCUUGUAGAUGAAAGGGGGTGGGGGGAGTTT*mG*mC*mT	3990
7	mC*U*AAUUUCUACUCUUGUAGAUGAAAGGGGGTGGGGGGAGTTTG*C*mT	3991
8	/AltR1/UAAUUUCUACUCUUGUAGAUGAAAGGGGGTGGGGGGAGTTTGCT/AltR2/	3992
Legend: /AltR1/ and /AltR2/ refer to IDT's proprietary Alt-R 5' and 3' modifications; m; 2'-O-methyl base (for example a base with 2'-O-methyl modification is written as mA), i2F; internal 2'-flourine base (for example an internal C with 2'-flourine modification is written as /i2FC/), 52F; 2'-flourine base at the 5' end of the sequence (for example a 5' C with 2'-flourine modification is written as /52FC/), 32F; 2'-flourine base at the 3' end of the sequence (for example a 3' A base with 2'-flourine modification is written as /32FA/), r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *; between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur		

Example 27 – Titration of modified MG29-1 guides from Example 26

[00417] A further experiment was performed to determine the dose dependence of the activity of the modified guides used in Example 26 to identify possible dose-dependent toxicity effects. The experiment was performed as in Example 26 but with 1/4th (B), 1/8th (C), 1/16th (D), and 1/32nd (E) of the starting dose (A, 126 pmol MG29-1 and 160 pmol guide RNA). The results are presented in **FIG. 48**).

Example 28 – Large Scale Synthesis of nucleases described herein

Project Overview

[00418] Production of Metagenomi's Type V-A CRISPR nuclease, MG29-1, is scaled up to an initial culture volume of 10L. An expression screen, scaled-up expression, downstream development, a formulation study, and delivery of purified protein >=90% by SDS-PAGE are performed.

Expression and purification screen

Expression:

[00419] Expression of MG29-1 from the pMG450 vector depicted in **FIG. 49** is tested in a screen varying the following conditions: host strain, expression media, inducer, induction time, and temperature.

[00420] After screening, *E. coli* is transformed with the appropriate expression plasmid, the culture is grown to a suitable density shake flasks, and the culture is induced using materials and methods according to the optimal expression conditions identified during the expression screen.

The cell paste is harvested and expression is verified by SDS-PAGE. For these experiments, the cell culture volume is limited to 20L. Up to 1 gram of protein is purified using the following method, formulated into storage buffer, and yield and concentration by A280 and purity by SDS-PAGE is assessed.

Purification:

[00421] Total soluble protein extracted from *E. coli* cell paste is analyzed by SDS-PAGE for all conditions. Immobilized metal affinity chromatography (IMAC) pull-down followed by SDS-PAGE is performed on the top three expression conditions to estimate yield and purity and to identify the optimal expression condition. A scaled-up method is developed for lysis. Critical parameters are identified for purification by IMAC and subtractive IMAC (including tobacco etch virus protease (TEV) cleavage). Column fractions are tested using SDS-PAGE. Elution pools are tested using SDS- PAGE and photometric absorbance at 280 nm (A280). A method for buffer exchange and concentration by tangential flow filtration (TFF) is developed.

[00422] An additional chromatography stage is developed to achieve ≥90% purity, if purity is lower than 90%. One chromatography mode is tested (e.g., ceramic hydroxyapatite chromatography). Up to 8 unique conditions are tested (e.g., 2-6 resins each with 2-3 buffer systems). Column fractions are tested using SDS-PAGE. Elution pools are tested using SDS-PAGE and A280. One condition is selected, and a three-condition load study is performed. Column fractions and elution pools are analyzed as described above. A method for buffer exchange may be developed and concentration by TFF.

Formulation Study

[00423] Using purified protein, a formulation study is conducted to determine the optimal storage conditions for the purified protein. Study may explore concentration, storage buffer, storage temperature, maximum freeze/thaw cycles, storage time, or other conditions.

Example 29 – Demonstration of the ability of nucleases described herein to edit an intronic region in cultured mouse liver cells

[00424] Intronic regions of expressed genes are attractive genomic targets to integrate a coding sequence of a therapeutic protein of interest with the goal of expressing that protein to treat or cure a disease. Integration of a protein coding sequence may be accomplished by creating a double strand break within the intron using a sequence specific nuclease in the presence of an exogenously supplied donor template. The donor template may be integrated into the double strand break via one of two main cellular repair pathways called homology directed repair (HDR) and non-homologous end joining (NHEJ) resulting in targeted integration of the donor template.

The NHEJ pathway is dominant in non-dividing cells while the HDR pathway is primarily active in dividing cells. The liver is a particularly attractive tissue for targeted integration of a protein coding sequence due to the availability of *in vivo* delivery systems and the ability of the liver to express and secrete proteins with high efficiency.

[00425] To evaluate the potential of MG29-1 to create double strand breaks at intronic regions the intron 1 of serum albumin was selected as the target locus. Single guide RNA (sgRNA) with a spacer length of 22 nt targeted to mouse albumin intron 1 were identified using the guide finding algorithm in the Geneious Prime nucleic acid analysis software (https://www.geneious.com/prime/). Using a PAM of KTTG (SEQ ID NO: 3870) located 5' to the spacer, a total of 112 potential sgRNA were identified within mouse albumin intron 1. Guides that spanned the intron/exon boundaries were excluded. Using Geneious Prime the spacer sequences of these 112 guides were searched against the mouse genome and a specificity score was assigned by the software based on the alignment to additional sites in the genome. Spacer sequences with 4 or more contiguous bases of the same base were excluded due to concerns about specificity. A total of 12 spacers with the highest specificity scores were selected for testing. To create the sgRNA the backbone sequence of "TAATTTCTACTGTTGTAGAT" was added to the 3' end of the spacer sequence. The sgRNA was chemically synthesized incorporating chemically modified bases identified to improve the performance of sgRNA for cpf1 guides (AltR1/AltR2 chemistry available from Integrated DNA Technologies). The spacer sequences of these guides are listed in Table 8 below.

Table 8 - Activity of MG29-1 sgRNA targeting mouse albumin intron 1 in Hepa1-6 cells nucleofected with MG29-1/sgRNA RNP or transfected with MG29-1 mRNA and sgRNA using Messenger Max

sgRNA name	Spacer (DNA sequence, no PAM)	SEQ ID NO:	PAM	SEQ ID NO:	Specificity score	Activity (INDEL %) in Hepa1-6 cells	
						RNP nucleofection	mRNA/sgRNA lipid transfection
mAlb29-1-1	GTATAGCATGGTCGAGCAGGCA	3993	TTTA	4012	98.5	86.5	43
mAlb29-1-2	CCGATCGTTACAGGAAAATCTG	3994	GTTC	4013	98.4	0	0
mAlb29-1-3	AATTTATTACGGTCTCATAGGG	3995	GTTG	4014	98.2	0	0
mAlb29-1-4	TTACGGTCTCATAGGGCCTGCC	3996	TTTA	4015	97.6	43.5	44

sgRNA name	Spacer (DNA sequence, no PAM)	SEQ ID NO:	PAM	SEQ ID NO:	Specificity score	Activity (INDEL %) in Hepa1-6 cells	
						RNP nucleofection	mRNA/sgRNA lipid transfection
mAlb29-1-5	CCTGTAACGATCGGGAACTGGC	3997	TTTT	4016	97.2	3	0
mAlb29-1-7	AGTATAGCATGGTCGAGCAGGC	3998	TTTT	4017	96.8	11	15
mAlb29-1-8	CTGTAACGATCGGGAACTGGCA	3999	TTTC	4018	95.9	77	45
mAlb29-1-9	GATACAGTTGAATTTATTACGG	4000	GTTG	4019	95.3	0	0
mAlb29-1-10	TAGTATAGCATGGTCGAGCAGG	4001	TTTT	4020	95.2	18	35
mAlb29-1-11	CATCTGAGAACCCTTAGGTGGT	4002	TTTG	4021	95.0	7	2
mAlb29-1-12	AGTGTAGCAGAGAGGAACCATT	4003	TTTG	4022	93.8	NT	47
mAlb29-1-13	CTAGTAATGGAAGCCTGGTATT	4004	TTTT	4023	92.4	8	24
mAlb29-1-14	GGTATCTTTGATGACAATAATG	4005	TTTT	4024	91.8	0	13
mAlb29-1-15	TCTAGTAATGGAAGCCTGGTAT	4006	TTTT	4025	91.8	0	0
mAlb29-1-16	TAGTAATGGAAGCCTGGTATTT	4007	TTTC	4026	89.8	90.5	51
mAlb29-1-17	GTATCTTTGATGACAATAATGG	4008	TTTG	4027	87.8	10	NT
mAlb29-1-18	AAGATTGATGAAGACAACTAAC	4009	TTTA	4028	87.4	76	NT
mAlb29-1-19	CTCTCTGCTACACTCAAAGTTA	4010	GTTC	4029	85.7	0	0
mAlb29-1-20	AAACCCGTTAAGTGTTTATATC	4011	TTTA	4030	87.3	0	4

[00426] Hepa1-6 cells, a transformed mouse liver cell line, were cultured under standard conditions (DMEM media with 10% FBS in 5% CO2 incubator) and nucleofected with ribonuclear proteins formed by mixing the sgRNA and purified MG29-1 protein in PBS buffer. Hepa1-6 cells ($1x10^5$) in suspension in complete SF nucleofection reagent (Lonza) were nucleofected using a 4D nucleofection device (Lonza) with RNP formed by mixing 50 pmol of MG29-1 protein and 100 pmol of sgRNA. After nucleofection the cells were plated in 24 well

plates in DMEM plus 10% FBS and incubated in a 5% CO2 incubator for 48 to 72 h. Genomic DNA was then extracted from the cells using a column-based purification kit (Purelink genomic DNA mini kit, ThermoFisher Scientific) and quantified by absorbance at 260 nm. The albumin intron 1 region was PCR amplified from 50 ng of the genomic DNA in a reaction containing 0.5 micro molar each of the primers mAlb90F (CTCCTCTTCGTCTCCGGC) (SEQ ID NO: 4031) and mAlb1073R (CTGCCACATTGCTCAGCAC) (SEQ ID NO: 4032) and 1 x Pfusion Flash PCR Master Mix.

[00427] The resulting 984 bp PCR product which spans the entire intron 1 of mouse albumin was purified using a column-based purification kit (DNA Clean and Concentrator, Zymo Research) and sequenced using primers located within 150 to 350 bp of the predicted target site for each sgRNA. A PCR product generated using primers mAlb90F (SEQ ID NO: 4031) and mAlb1073R (SEQ ID NO: 4032) from un-transfected Hepa1-6 cells was sequenced in parallel as a control. The Sanger sequencing chromatograms were analyzed using Inference of CRISPR Edits (ICE) that determines the frequency of INDELS as well as the INDEL profile (Hsiau et. al, Inference of CRISPR Edits from Sanger Trace Data. BioArxiv. 2018 https://www.biorxiv.org/content/early/2018/01/20/251082).

[00428] When a nuclease creates a double strand break (DSB) in DNA inside a living cell the DSB is repaired by the cellular DNA repair machinery. In actively dividing cells such as transformed mammalian cells in culture, and in the absence of a repair template, this repair occurs by the NHEJ pathway. The NHEJ pathway is an error prone process that introduces insertions or deletions of bases at the site of the double strand break (Lieber, M.R, Annu Rev Biochem. 2010; 79: 181–211). These insertions and deletions are therefore a hallmark of a double strand break that occurred and was subsequently repaired, is widely used as a readout of the editing or cutting efficiency of the nuclease. The profile of insertions and deletions depends on the characteristics of the nuclease that created the double strand break but also upon the sequence context at the cleavage site. Based on *in vitro* assays, the MG29-1 nuclease creates a staggered cut located 3' of the PAM. Staggered cuts will often lead to larger deletions due to the trimming of the single stranded ends before end-joining. Table 8 lists the total INDEL frequency generated by each of the 19 sgRNA targeting mouse albumin intron 1 that were tested in Hepa1-6 cells. Eleven of the 18 sgRNA resulted in detectable INDELS at the target site with 5 sgRNA resulting in INDEL frequencies greater than 50% and 4 sgRNA resulted in indel frequencies greater than 75%. These data demonstrate that the MG29-1 nuclease can edit the genome of cultured mouse liver cells at the predicted target site for the sgRNA with efficiencies greater than 75%.

[00429] The editing efficiencies of the same set of sgRNA were evaluated by co-transfection of the sgRNA and a mRNA encoding the MG29-1 nuclease using a commercial lipid-based transfection reagent (Lipofectamine MessengerMAX, Invitrogen). The mRNA encoding MG29-1 was generated by *in vitro* transcription using T7 polymerase from a plasmid in which the coding sequence of MG29-1 was cloned. The MG29-1 coding sequence was codon optimized using human codon usage tables and flanked by nuclear localization signals derived from SV40 at the N-terminus and from Nucleoplasmin at the C-terminus. In addition, a UTR was included at the 3' end of the coding sequence to improve translation. A 3' UTR followed by an approximately 90 to 110 nucleotide poly A tract was included at the 3' end of the coding sequence to improve mRNA stability *in vivo* (see e.g. SEQ ID NO: 4426 for wild-type MG29-1 and SEQ ID NO: 3327 for the S168R variant). The *in vitro* transcription reaction included the Clean Cap® capping reagent (Trilink BioTechnologies) and the resulting RNA was purified using the MEGAClear™ Transcription Clean-Up kit (Invitrogen) and purity was evaluated using the TapeStation (Agilent) and found to be composed of >90% full length RNA. As seen in **Table 1**, the editing efficiencies after mRNA/sgRNA lipid transfection of Hepa1-6 cells were similar but not identical to those seen with nucleofection of RNP but confirm that the MG29-1 nuclease is active in cultured liver cells when delivered in the form of an mRNA.

[00430] FIG. 50 is a representative example of the indel profile of MG29-1 as determined by ICE analysis using mALb29-1-8 as the guide (SEQ ID NO: 3999) and demonstrates that deletion of 4 bases was the most frequent event (25% of total sequences) and deletions of 1, 5, 6, or 7 bases each accounting for about 10 to 15% of the sequences. Longer deletions of up to 13 bases were also detected, but insertions were undetectable. By contrast, spCas9 with a guide targeting mouse albumin intron 1 generated primarily 1 base insertions or deletions.

[00431] FIG. 51 is a representative example of the indel profile of MG29-1 and sgRNA mAlb29-1-8 as determined by next generation sequencing (NGS) of the PCR product of the mouse albumin intron 1 region. In total approximately 15,000 sequence reads were obtained. By NGS deletion of 4 bases was the most frequent indel (about 20% of total) with deletions of 1, 5, 6 and 7 bases each accounting for about 10% of the indels. Larger deletions of up to 19bp were also detected. The profile observed by NGS analysis matches closely that measured by ICE. These results demonstrate that MG29-1 generates large deletions at the target site consistent with the staggered cleavage observed *in vitro*.

Example 8 – Demonstration of the ability of a nuclease described herein to target an intronic region in cultured human liver cells (HepG2)

[00432] To evaluate the potential of MG29-1 to create double strand breaks at intronic regions in human cells, the intron 1 of human serum albumin was selected as the target locus. Single guide RNA (sgRNA) with a spacer length of 22 nt targeted to human albumin intron 1 were identified using the guide finding algorithm in the Geneious Prime nucleic acid analysis software (https://www.geneious.com/prime/). Using a PAM of KTTG (SEQ ID NO: 3870) located 5' to the spacer, a total of 90 potential sgRNA were identified within human albumin intron 1. Guides that spanned the intron/exon boundaries were excluded. Using Geneious Prime the spacer sequences of these guides were searched against the mouse genome and a specificity score was assigned by the software based on the alignment to additional sites in the genome. Spacer sequences with 4 or more contiguous bases of the same base were excluded due to concerns about specificity. A total of 23 spacers with the highest specificity scores were selected for testing. To create the sgRNA the backbone sequence of "TAATTTCTACTGTTGTAGAT" was added to the 3' end of the spacer sequence. The sgRNA was chemically synthesized incorporating chemically modified bases identified to improve the performance of sgRNA for cpf1 guides (AltR1/AltR2 chemistry available from Integrated DNA Technologies). The spacer sequences of these guides are listed in **Table 9**.

Table 9 - Spacer sequences of MG29-1 sgRNA targeting human albumin intron 1 and activity in HepG2 cells nucleofected with MG29-1/sgRNA RNP

sgRNA name	Spacer (DNA sequence, no PAM)	SEQ ID NO:	PAM	SEQ ID NO:	Specificity score	Activity (INDEL %) in HepG2 cells
hAlb_g63	GTAAACTCTGCATCTTTAAAGA	4033	TTTA	4056	91.25%	0
hAlb_g59	TTTCAAAATATTGGGCTCTGAT	4034	TTTG	4057	90.64%	0
hAlb_g58	AGTAAACTCTGCATCTTTAAAG	4035	TTTT	4058	90.46%	0
hAlb_g56	AAGATGCAGAGTTTACTAAAAC	4036	TTTA	4059	90.23%	0
hAlb_g72	AAAATATTGGGCTCTGATTCCT	4037	TTTC	4060	93.26%	0
hAlb_g70	AAATAAAGCATAGTGCAATGGA	4038	TTTT	4061	92.31%	63
hAlb_g74	AATAAAGCATAGTGCAATGGAT	4039	TTTA	4062	95.41%	93
hAlb_g83	TGAGATCAACAGCACAGGTTTT	4040	TTTA	4063	98.39%	93
hAlb_g85	TGTAGGAATCAGAGCCCAATAT	4041	TTTC	4064	99.20%	55
hAlb_g89	CTGTAGGAATCAGAGCCCAATA	4042	TTTT	4065	100.00%	43
hAlb_g88	TCTGTAGGAATCAGAGCCCAAT	4043	TTTT	4066	100.00%	45
hAlb_g77	GTGACTGTAATTTTCTTTTGCG	4044	TTTA	4067	96.77%	0
hAlb_g69	CTTTTGCGCACTAAGGAAAGTG	4045	TTTT	4068	92.18%	3
hAlb_g66	TGAAGTCTTACAAGGTTATCTT	4046	TTTG	4069	91.80%	19
hAlb_g75	AGTGTCTATCAACAGCAACCAA	4047	TTTT	4070	95.96%	13

sgRNA name	Spacer (DNA sequence, no PAM)	SEQ ID NO:	PAM	SEQ ID NO:	Specificity score	Activity (INDEL %) in HepG2 cells
hAlb_g79	CTTAGTGCGCAAAAGAAAATTA	4048	TTTC	4071	97.45%	60
hAlb_g82	TAGCCTTATATTCAAACTTAGA	4049	TTTA	4072	98.32%	0
hAlb_g80	GGATAGTTATGAATTCAATCTT	4050	TTTG	4073	97.46%	23
hAlb_g84	CACTTTCCTTAGTGCGCAAAAG	4051	TTTG	4074	98.85%	96
hAlb_g81	GTATTTGTGAAGTCTTACAAGG	4052	TTTT	4075	98.07%	17
hAlb_g90	GTGTCTATCAACAGCAACCAAG	4053	TTTA	4076	100.00%	91
hAlb_g87	CGCACTAAGGAAAGTGCAAAGT	4054	TTTG	4077	100.00%	97
hAlb_g86	GCGCACTAAGGAAAGTGCAAAG	4055	TTTT	4078	100.00%	42

[00433] HepG2 cells, a transformed human liver cell line, were cultured under standard conditions (MEM media with 10% FBS in 5% CO2 incubator) and nucleofected with ribonuclear proteins formed by mixing the sgRNA and purified MG29-1 protein in PBS buffer. A total of 1 e5 HepG2 cells in suspension in complete SF nucleofection reagent (Lonza) were nucleofected using a 4D nucleofection device (Lonza) with RNP formed by mixing 80 pmol of MG29-1 protein and 160 pmol of sgRNA. After nucleofection the cells were plated in 24 well plates in DMEM plus 10% FBS and incubated in a 5% CO_2 incubator for 48 to 72 h. Genomic DNA was then extracted from the cells using a column-based purification kit (Purelink genomic DNA mini kit, ThermoFisher Scientific) and quantified by absorbance at 260 nm. The albumin intron 1 region was PCR amplified from 50 ng of the genomic DNA in a reaction containing 0.5 micro molar each of the primers hAlb 11F (TCTTCTGTCAACCCCACACGCC) (SEQ ID NO: 4079) and hAlb834R (CTTGTCTGGGCAAGGGAAGA) (SEQ ID NO: 4080) and 1 x Pfusion Flash PCR Master Mix. The resulting 826 bp PCR product which spans the entire intron 1 of mouse albumin was purified using a column-based purification kit (DNA Clean and Concentrator, Zymo Research) and sequenced using primers located within 150 to 350 bp of the predicted target site for the sgRNA.

[00434] The PCR product generated using primers hAlb 11F (TCTTCTGTCAACCCCACACGCC) (SEQ ID NO: 4079) and hAlb834R (CTTGTCTGGGCAAGGGAAGA) (SEQ ID NO: 4080) from un-transfected HepG2 cells was sequenced in parallel as a control. The Sanger sequencing chromatograms were analyzed using Inference of CRISPR Edits (ICE) that determines the frequency of INDELS as well as the INDEL profile. When a nuclease creates a double strand break (DSB) in DNA inside a living cell the DSB is repaired by the cellular DNA repair machinery. In actively dividing cells such as transformed mammalian cells in culture, and in the absence of a repair template, this repair

occurs by the NHEJ pathway. The NHEJ pathway is an error prone process that introduces insertions or deletions of bases at the site of the double strand break (Lieber, M.R, Annu Rev Biochem. 2010; 79: 181–211).

[00435] These insertions and deletions are therefore a hallmark of a double strand break that occurred and was subsequently repaired, and is widely used as a readout of the editing or cutting efficiency of the nuclease. The profile of insertions and deletions depends on the characteristics of the nuclease that created the double strand break but also upon the sequence context at the cleavage site. Based on *in vitro* assays, the MG29-1 nuclease cleaves the target strand at 22 nucleotides from the PAM (less frequently at 21 nucleotides from the PAM) and cleaves the non-target strand at 18 nucleotides from the PAM which therefore creates 4 nucleotide staggered end located 3' of the PAM. Staggered cuts will often lead to larger deletions due to the trimming of the single stranded ends before end-joining.

[00436] **Table 9** lists the total indel frequency generated by each of the 23 sgRNA targeting human albumin intron 1 that were tested in HepG2 cells. Sixteen of the 23 sgRNA resulted in detectable indel at the target site with 8 sgRNA resulting in INDELS greater than 50% and 5 sgRNA resulted in indel frequencies than 90%. These data demonstrate that the MG29-1 nuclease can edit the genome of a cultured human liver cell line at the predicted target site for the sgRNA with efficiencies greater than 90%.

Example 9 – Demonstration of the ability of nucleases described herein to edit exonic regions in cultured mouse liver cells

[00437] Sequence specific nucleases can be used to disrupt the coding sequences of genes and thereby create a functional knockout of a protein of interest. This can be of therapeutic use when the knockdown of a specific protein has a beneficial effect in a particular disease. One way to disrupt the coding sequence of a gene is to make a double strand break within the exonic regions of the gene using a sequence specific nuclease. These double strand breaks will be repaired via error prone repair pathways to generate insertions or deletions which can result in either frameshift mutations or changes to the amino acid sequence which disrupt the function of the protein.

[00438] To evaluate the potential of MG29-1 to create double strand breaks at exonic regions of a gene expressed in liver cells the gene encoding glycolate oxidase (hao-1) was selected as the target locus. Single guide RNA (sgRNA) with a spacer length of 22 nt targeted to exons 1 to 4 of mouse hao-1 were identified using the guide finding algorithm in the Geneious Prime nucleic acid analysis software (https://www.geneious.com/prime/). The first 4 exons of the hao-1 gene comprise approximately the N-terminal 50% of the hao-1 coding sequence. The first 4 exons

were chosen because INDELS created towards the N-terminus of the coding sequence of a gene are more likely to create a frameshift or missense mutation that disrupts the activity of the protein. Using a PAM of KTTG (SEQ ID NO: 3870) located 5' to the spacer, a total of 45 potential sgRNAs were identified within mouse hao-1 exons 1 through 4. Guides that spanned the intron/exon boundaries were included because such guides may create INDELS that interfere with splicing. Using Geneious Prime, the spacer sequences of these 45 guides were searched against the mouse genome and a specificity score was assigned by the software based on the alignment to additional sites in the mouse genome. Spacer sequences with 4 or more contiguous bases of the same base were excluded due to concerns about specificity. A total of 45 spacers with the highest specificity scores were selected for testing.

[00439] To create the sgRNA the backbone sequence of "TAATTTCTACTGTTGTAGAT" was added to the 3' end of the spacer sequence. The sgRNA was chemically synthesized incorporating chemically modified bases identified to improve the performance of sgRNA for cpf1 guides (AltR1/AltR2 chemistry available from Integrated DNA Technologies). The spacer sequences of these guides are listed in Table 3.

Table 3 - Spacer sequences of MG29-1 sgRNA targeting mouse hao-1 exons 1 to 4 and activity in Hepa1-6 cells nucleofected with MG29-1/sgRNA RNP

sgRNA name	Spacer (DNA sequence, no PAM)	SEQ ID NO:	PAM	SEQ ID NO:	Specificity score	Activity (INDEL %) in Hepa1-6 cells
mH29-1	CCCCAGACCTGTAATAGTCATA	4081	TTTG	4126	100.00%	92.2
mH29-2	AGGACAGAGAGTCAGCATGCCA	4082	TTTT	4127	100.00%	0
mH29-3	GGAGACAACAGTGGACTTGCTG	4083	TTTT	4128	100.00%	0
mH29-4	CCCTACCCTGCCACAATGTTGC	4084	GTTG	4129	100.00%	0
mH29-5	CTTACCTAGAAAATGCTTGGAT	4085	GTTT	4130	100.00%	0
mH29-6	ACAGATCGATATCAGCAACGTT	4086	GTTG	4131	100.00%	0
mH29-7	CGAAGCATCCGTGGATAGAGCT	4087	GTTG	4132	100.00%	0
mH29-8	TTGGGCTACCTCCTCAATAGAA	4088	GTTC	4133	100.00%	0
mH29-9	AAGCTGCCACCACAACTCAGGT	4089	GTTC	4134	100.00%	0
mH29-10	TGGTGGCAGCTTGAACCTGTTC	4090	GTTG	4135	100.00%	0
mH29-11	CGCACGTCATCAATGCGGTTGC	4091	GTTC	4136	100.00%	0
mH29-12	CCCAGGTAAGGGGTGTCCACAG	4092	GTTG	4137	100.00%	0
mH29-13	CATCCAGCGAAGTGCCTCTGGG	4093	GTTG	4138	100.00%	0
mH29-14	AAATTCCAGATGGAAGCTCTAT	4094	TTTT	4139	99.04%	0
mH29-15	TGACTGTGGACACCCCTTACCT	4095	TTTG	4140	99.04%	97.5
mH29-16	ATTACAGCCTGTCAGACCATGG	4096	TTTC	4141	99.30%	91
mH29-17	GAGACAACAGTGGACTTGCTGA	4097	TTTG	4142	98.37%	27.5
mH29-18	CAACAATAGGCAGTGATGTCAA	4098	TTTA	4143	99.22%	85
mH29-19	CCTCGACTGGTCTGCATCAGTG	4099	GTTG	4144	97.69%	0
mH29-20	ATAATCACTGATGCAGACCAGT	4100	GTTC	4145	98.07%	0

sgRNA name	Spacer (DNA sequence, no PAM)	SEQ ID NO:	PAM	SEQ ID NO:	Specificity score	Activity (INDEL %) in Hepa1-6 cells
mH29-21	TCAGCTAACGTCTCCTGATCAT	4101	GTTA	4146	99.33%	0
mH29-22	CTGATATCGATCTGTCAACTTC	4102	GTTG	4147	99.22%	0
mH29-23	TAAAGGGCATTTTGAGAGGTTT	4103	GTTG	4148	99.22%	0
mH29-24	AATAGCAAAGTTTCTTACCTAG	4104	TTTA	4149	95.75%	0
mH29-25	GGACAGAGAGTCAGCATGCCAA	4105	TTTA	4150	95.73%	47
mH29-26	TCCATTTCATTACAGCCTGTCA	4106	TTTC	4151	94.08%	79
mH29-27	AGTCTGTGAGATCATACTGACC	4107	TTTG	4152	96.85%	65
mH29-28	TAGATGTACAGTTGCATCCAGC	4108	TTTG	4153	98.20%	29
mH29-29	CCTTAGGAGAAAATGCCAAATC	4109	TTTC	4154	96.18%	94.5
mH29-30	CTCCTAAGGGAAATTTTGGAGA	4110	TTTT	4155	98.06%	0
mH29-31	GCTGATAACATCCAAGCATTTT	4111	GTTA	4156	93.30%	0
mH29-32	AAATAGCAAAGTTTCTTACCTA	4112	GTTT	4157	97.31%	0
mH29-33	TAGGACAGAGAGTCAGCATGCC	4113	GTTT	4158	95.84%	0
mH29-34	GGGCTACTGCCATGCAGTGCAT	4114	GTTG	4159	97.53%	0
mH29-35	TCTCCAAAATTTCCCTTAGGAG	4115	GTTG	4160	96.94%	0
mH29-36	AATTCCAGATGGAAGCTCTATC	4116	TTTA	4161	96.74%	0
mH29-37	TCCTAAGGGAAATTTTGGAGAC	4117	TTTC	4162	97.59%	53.5
mH29-38	TTACCTAGAAAATGCTTGGATG	4118	TTTC	4163	94.99%	21.8
mH29-39	CAAGGCCATATTTGTGACTGTG	4119	GTTA	4164	93.57%	0
mH29-40	CTCCATTTCATTACAGCCTGTC	4120	TTTT	4165	93.71%	0
mH29-41	GCATTTTCTCCTAAGGGAAATT	4121	TTTG	4166	92.30%	59
mH29-42	TTACCTCGCACAGTGGCCAGCT	4122	TTTC	4167	77.16%	32
mH29-43	TCTCTCTTTTCTTACCTCGCAC	4123	TTTG	4168	87.70%	0
mH29-44	CTTACCTCGCACAGTGGCCAGC	4124	TTTT	4169	95.19%	0
mH29-45	AAACCAATGATTTGGCATTTTC	4125	TTTG	4170	91.09%	0

[00440] Hepa1-6 cells, a transformed mouse liver cell line, were cultured under standard conditions (DMEM media with 10% FBS in 5% CO_2 incubator) and nucleofected with ribonuclear proteins formed by mixing the sgRNA and purified MG29-1 protein in PBS buffer. A total of 1 e^5 Hepa1-6 cells in suspension in complete SF nucleofection reagent (Lonza) were nucleofected using a 4D nucleofection device (Lonza) with RNP formed by mixing 50 pmol of MG29-1 protein and 100 pmol of sgRNA. After nucleofection the cells were plated in 24 well plates in DMEM plus 10% FBS and incubated in a 5% CO2 incubator for 48 to 72 h. Genomic DNA was then extracted from the cells using a column-based purification kit (Purelink genomic DNA mini kit, ThermoFisher Scientific) and quantified by absorbance at 260 nm. Exons 1 through 4 of the mouse hao-1 gene 1 were PCR amplified from 40 ng of the genomic DNA in a reaction containing 0.5 micro molar pairs of the primers specific for each exon. The PCR primers used for exon 1 were PCR_mHE1_F_+233 (GTGACCAACCCTACCCGTTT) (SEQ ID NO: 4171), PCR_mHE1_R_-553 (GCAAGCACCTACTGTCTCGT) (SEQ ID NO: 4172). The PCR

primers used for exon 2 were HAO1_E2_F5721 (CAACGAAGGTTCCCTCCAGG) (SEQ ID NO: 4173), HAO1_E2_R6271 (GGAAGGGTGTTCGAGAAGGA) (SEQ ID NO: 4174). The PCR primers used for exon 3 were HAO1_E3_F23198 (TGCCCTAGACAAGCTGACAC) (SEQ ID NO: 4175), HAO1_E3_R23879 (CAGATTCTGGAAGTGGCCCA) (SEQ ID NO: 4176). The PCR primers used for exon 4 were HAO1_E4_F31087 (CCTGTAGGTGGCTGAGTACG) (SEQ ID NO: 4177), HAO1_E4_R31650 (AGGTTTGGTTCCCCTCACCT) (SEQ ID NO: 4178).

[00441] In addition to primers and genomic DNA the PCR reactions contained 1 x Pfusion Flash PCR Master Mix (Thermo Fisher). The resulting PCR products comprised single bands when analyzed on agarose gels demonstrating that the PCR reaction was specific, and were purified using a column-based purification kit (DNA Clean and Concentrator, Zymo Research). For sequencing, primers complementary to sequences at least 100nt from each cut site were used. The primer to sequence Exon 1 was Seq_mHE1_F_+139 (GTCTAGGCATACAATGTTTGCTCA) (SEQ ID NO: 4179). The primer to sequence Exon 2 was 5938F Seq_HAO1_E2 (CTATGCAAGGAAAAGATTTGGCC) (SEQ ID NO: 4180). The primers to sequence Exon 3 were HAO1_E3_F23476 (TCTTCCCCCTTGAATGAAACACT) (SEQ ID NO: 4181) and the reverse PCR primer, HAO1_E3_R23879 (CAGATTCTGGAAGTGGCCCA) (SEQ ID NO: 4182). The primer to sequence Exon 4 was the reverse PCR primer, HAO1_E4_R31650 (AGGTTTGGTTCCCCTCACCT) (SEQ ID NO: 4183).

[00442] Sequencing of the PCR products showed that they contained the expected sequences of the hao-1 exons. PCR products derived from Hepa-16 cells nucleofected with different RNP or untreated controls were sequenced using primers located within 100 to 350 bp of the predicted target site for each sgRNA. The Sanger sequencing chromatograms were analyzed using Inference of CRISPR Edits (ICE) that determines the frequency of INDELS as well as the INDEL profile (Hsiau et. al, Inference of CRISPR Edits from Sanger Trace Data. BioArxiv. 2018 https://www.biorxiv.org/content/early/2018/01/20/251082). When a nuclease creates a double strand break (DSB) in DNA inside a living cell the DSB is repaired by the cellular DNA repair machinery. In actively dividing cells such as transformed mammalian cells in culture, and in the absence of a repair template, this repair occurs by the NHEJ pathway. The NHEJ pathway is an error prone process that introduces insertions or deletions of bases at the site of the double strand break (Lieber, M.R, Annu Rev Biochem. 2010; 79: 181–211). These insertions and deletions are therefore a hallmark of a double strand break that occurred and was subsequently repaired, and is widely used in the art as a readout of the editing or cutting efficiency of the nuclease. As presented in Table 3, 14 guides demonstrated detectable editing at their predicted target sites.

Four guides exhibited editing activity greater than 90%. All 14 of the active guides had PAM sequences of TTTN demonstrating that this PAM is more efficient *in vivo*. However not all guides utilizing a TTTN PAM were active. These data demonstrate that the MG29-1 nuclease can generate RNA guided, sequence specific, double strand breaks in exonic regions in cultured liver cells with high efficiency.

Example 10 – Design of further sgRNAs for disruption of Hao-1 gene

[00443] Further sgRNAs were designed to target exonic parts of the hao-1 gene. These are designed to target the first 4 exons because these comprise approximately 50% of the coding sequence and indels created towards the N-terminus of the coding sequence of a gene are more likely to create a frameshift or missense mutation that disrupts the activity of the protein. Using the more restrictive PAM of KTTG (SEQ ID NO: 3870) which was shown in Example 9 to be more active in mammalian cells, a total of 42 potential sgRNA were identified within human hao-1 exons 1 through 4 (Table 4).

Table 4 – Spacer sequences for MG29-1 identified in exons 1 to 4 of the human hao-1 gene

sgRNA name	Spacer (DNA sequence, no PAM)	SEQ ID NO:	PAM	SEQ ID NO:	Specificity score
hH29-1	GCATGTTGTTCATAATCATTGA	4184	TTTA	4226	96.25%
hH29-2	GAAGTACTGATTTAGCATGTTG	4185	TTTG	4227	98.37%
hH29-3	TATCAATGATTATGAACAACAT	4186	TTTG	4228	87.44%
hH29-4	CCCCAGACCTGTAATAGTCATA	4187	TTTG	4229	99.04%
hH29-5	TTCATCATTTGCCCCAGACCTG	4188	TTTC	4230	95.59%
hH29-6	TTACCTGGAAAATGCTGCAATA	4189	TTTC	4231	80.36%
hH29-7	CTTACCTGGAAAATGCTGCAAT	4190	TTTT	4232	79.67%
hH29-8	GCTGATAATATTGCAGCATTTT	4191	TTTG	4233	92.20%
hH29-9	AAAAATAAATTTTCTTACCTGG	4192	TTTA	4234	58.56%
hH29-10	AAAAAATAAATTTTCTTACCTG	4193	TTTT	4235	44.93%
hH29-11	ATTTTATTTTTTAATTCTAGAT	4194	TTTT	4236	10.22%
hH29-12	TTTTATTTTTTAATTCTAGATG	4195	TTTA	4237	10.64%
hH29-13	ATTTTTTAATTCTAGATGGAAG	4196	TTTT	4238	70.62%
hH29-14	TTTTTTAATTCTAGATGGAAGC	4197	TTTA	4239	44.69%
hH29-15	TTAATTCTAGATGGAAGCTGTA	4198	TTTT	4240	99.13%
hH29-16	TAATTCTAGATGGAAGCTGTAT	4199	TTTT	4241	97.06%
hH29-17	AATTCTAGATGGAAGCTGTATC	4200	TTTT	4242	96.74%
hH29-18	ATTCTAGATGGAAGCTGTATCC	4201	TTTA	4243	98.94%
hH29-19	AGCAACATTCCGGAGCATCCTT	4202	TTTC	4244	97.81%
hH29-20	AGGACAGAGGGTCAGCATGCCA	4203	TTTT	4245	97.75%
hH29-21	GGACAGAGGGTCAGCATGCCAA	4204	TTTA	4246	100.00%
hH29-22	TTTCTCAGCCTGTCAGTCCCTG	4205	TTTC	4247	88.19%
hH29-23	TCAGCCTGTCAGTCCCTGGGAA	4206	TTTC	4248	100.00%

sgRNA name	Spacer (DNA sequence, no PAM)	SEQ ID NO:	PAM	SEQ ID NO:	Specificity score
hH29-24	TGACAGTGGACACACCTTACCT	4207	TTTG	4249	100.00%
hH29-25	AATCTGTTACGCACATCATCCA	4208	TTTG	4250	100.00%
hH29-26	ATGCATTTCTTATTTTAGGATG	4209	TTTT	4251	80.79%
hH29-27	TGCATTTCTTATTTTAGGATGA	4210	TTTA	4252	76.81%
hH29-28	TTATTTTAGGATGAAAAATTTT	4211	TTTC	4253	52.38%
hH29-29	AGGATGAAAAATTTTGAAACCA	4212	TTTT	4254	90.56%
hH29-30	GGATGAAAAATTTTGAAACCAG	4213	TTTA	4255	89.17%
hH29-31	CTCAGGAGAAAATGATAAAGTA	4214	TTTC	4256	90.51%
hH29-32	CCTCAGGAGAAAATGATAAAGT	4215	TTTT	4257	88.16%
hH29-33	GAAACCAGTACTTTATCATTTT	4216	TTTT	4258	86.74%
hH29-34	AAACCAGTACTTTATCATTTTC	4217	TTTG	4259	91.02%
hH29-35	TCATTTTCTCCTGAGGAAAATT	4218	TTTA	4260	83.29%
hH29-36	CTCCTGAGGAAAATTTTGGAGA	4219	TTTT	4261	91.88%
hH29-37	TCCTGAGGAAAATTTTGGAGAC	4220	TTTC	4262	96.24%
hH29-38	GCCACATATGCAGCAAGTCCAC	4221	TTTA	4263	100.00%
hH29-39	GGAGACGACAGTGGACTTGCTG	4222	TTTT	4264	90.43%
hH29-40	GAGACGACAGTGGACTTGCTGC	4223	TTTG	4265	99.01%
hH29-41	ATATCTTCCCAGCTGATAGATG	4224	TTTG	4266	99.18%
hH29-42	CAACAATTGGCAATGATGTCAG	4225	TTTG	4267	95.26%

[00444] Guides that spanned the intron/exon boundaries were included because such guides may create indels that interfere with splicing. Using Geneious Prime the spacer sequences of these 42 guides were searched against the human genome and a specificity score was assigned by the software based on the alignment to the human genome. A higher specificity score indicates a lower probability of that guide recognizing 1 or more sequences in the human genome other than the site to which the spacer was designed. The specificity scores ranged from 10% to 100% with 25 guides having a specificity score greater than 90% and 33 guides having a specificity score greater than 80%. This analysis demonstrates that guides targeting exonic regions of a human gene with high specificity scores can be readily identified and it is expected that a number of highly active guides are to be identified.

Example 11 – Comparison of the editing potency of nucleases described herein to that of spCas9 in mouse liver cells

[00445] The CRISPR Cas9 nuclease from the bacterial species Streptococcus pyogenes (spCas9) is widely used for genome editing and is among the most active RNA guided nucleases identified. The relative potency of MG29-1 compared to spCas9 was evaluated by nucleofection of different doses of RNP in the mouse liver cell line Hepa1-6. sgRNA targeting intron 1 of mouse albumin were used for both nucleases. For MG29-1, the sgRNA mAlb29-1-8 identified in

Example 29 was selected. Guide mAlb29-1-8 (see Example 29) was chemically synthesized incorporating chemically modifications called AltR1/AltR2 (Integrated DNA Technologies) designed to improve the potency of guides for the Type V nuclease cpf1 that has a similar sgRNA structure as MG29-1. For spCas9 a sgRNA that efficiently edited mouse albumin intron 1 was identified by testing 3 guides selected from an *in-silico* screen. The spCas9 protein used in these studies was obtained from a commercial supplier (Integrated DNA technologies AltR-sPCas9).

[00446] The sgRNA mAlbR1 (spacer sequence TTAGTATAGCATGGTCGAGC) was chemically synthesized and incorporated chemical modifications comprised of 2' O methyl bases and phosphorothioate (PS) linkages on the 3 bases on both ends of the guide that improve potency in cells. The mAlbR1 sgRNA generated INDELS at a frequency of 90% when RNP comprised of 20 pmol spCas9 protein/50 pmol of guide was nucleofected into Hepa1-6 cells indicating that this is a highly active guide. RNP formed with a range of nuclease protein from 20 pmoles to 1 pmole and a constant ratio of protein to sgRNA of 1:2.5 were nucleofected into Hepa1-6 cells. INDELS at the target site in mouse albumin intron 1 were quantified using Sanger sequencing of the PCR amplified genomic DNA and ICE analysis. The results shown in **FIG. 52** demonstrate that MG29-1 generated a higher percentage of INDELS than spCas9 at lower RNP doses when the editing was not saturating. These data indicate that MG29-1 is at least as active and potentially more active than spCas9 in liver-derived mammalian cells.

Example 34 – Engineering sequence variants of nucleases described herein and evaluation in mouse liver cells

[00447] In order to improve the editing efficiency of MG29-1 a set of mutations substituting one or two amino acids was introduced in the MG29-1 coding region. The set of amino acid substitutions was determined by alignment to *Acidaminococcus sp.* Cas12a (AsCas12a). Structured-guide engineering (Kleinstiver, et al, Nat Biotechnol. 2019, 37 276-282) substituted different amino acid in AsCas12a with the goal of altering or improving PAM binding. Four amino acid substitutions in AsCas12a: S170R, E174R, N577R and K583R, showed higher editing efficiencies with canonical and non-canonical PAMs. Sites matching these substitutions were identified in MG29-1 by multiple alignment and correspond to: S168R, E172R, N577R and K583R in MG29-1.

[00448] In order to test the single amino acid substitutions a 2-plasmid delivery system was used. Expression plasmids encoding MG29-1 with single amino acid substitutions were constructed using standard molecular cloning techniques. One plasmid encoded for MG29-1 under CMV

promoter, the second plasmid contained the mAlb29-1-8 sgRNA (see Table 8), which has high editing efficiency in Hepa 1-6 cells. Transcription of the guide was driven by a human U6 promoter. Confirmation of initial results from single amino acid substitutions using the 2-plasmid system and testing of double amino acid substitutions was done using *in vitro* transcribed (IVT) mRNA encoding MG29-1 (see Example 11 for details of how the IVT mRNA was made) and chemically synthesized guides incorporating the AltR1/AltR2 chemical modifications that had been optimized by Integrated DNA Technologies for Cpf1 (synthesized at Integrated DNA technologies). For delivery of the 2-plasmid system 100ng of plasmid encoding MG29-1 and 400ng of plasmid encoding the guide were mixed with Lipofectamine 3000, added to Hepa1-6 cells and incubated for 3 days before to genomic DNA isolation.

[00449] For delivery of IVT mRNA and synthetic guides, 300ng of mRNA and 120ng of synthetic guide were mixed with Lipofectamine Messenger Max, added to cells and incubated for 2 days before to genomic DNA isolation. Synthetic guides screened using IVT mRNA correspond to guides detailed in Table 8 but for simplicity the names of the guides in Figure 53 have been shortened so that guide "mAlb29-1-1" is represented as g1-1, "mAlb29-1-8" is represented as g1-8 and so on. One guide targeting the human T cell receptor locus (TRAC) was also tested (35 TRAC on **FIG. 53D**). Guide 35 TRAC spacer is: GAGTCTCTCAGCTGGTACACGG (SEQ ID NO: 4268) with a TTTG PAM. Guide 35 TRAC was ordered with the same modifications as mentioned before. Genomic DNA and PCR amplification was performed as described in the previous example for MG29-1 editing of mouse albumin intron 1. For guide 35 TRAC, the human TRAC locus was amplified with Primer F: TGCTTTGCTGGGCCTTTTTC (SEQ ID NO: 4269), Primer R: ACAGTCTGAGCAAAGGCAGG (SEQ ID NO: 4270). The resulting 957bp PCR product was purified as described previously. Editing was assessed by Sanger sequencing using primer ATCACGAGCAGCTGGTTTCT (SEQ ID NO: 4271).

[00450] Editing efficiency for mouse albumin intron 1 and human TRAC locus was quantified using Sanger sequencing of the PCR products followed by Inference of CRISPR Edits (ICE). Data representing up to 4 biological replicates are plotted in **FIGs. 53A-D**. The single amino acid substitution S168R demonstrated improved editing efficiency when using guide mAlb29-1-8 in the 2-plasmid system (**FIGs. 53A**). Mutation E172R did not provide a major improvement with guide mAlb29-1-8 while the mutation K583R completely prevented editing with the mAlb29-1-8 guide. Transfection with MG29-1 mRNA and synthetic guide mAlb29-1-8 confirmed the results from plasmid transfection (**FIGs. 53B**). The single amino acid substitution S168R conferred higher editing efficiency across the different concentrations of mRNA tested with guide mAlb29-

1-8 (**FIGs. 53B**). The double amino acid substitutions of S168R with E172R (substitution that did not impair activity alone as seen in **FIGs. 53A**), or N577R (a substitution not tested in MG29-1 plasmid transfection but conferred higher editing efficiency of cpf1) and Y170R (which it was hypothesized might improve editing efficiency based on the predicted MG29-1 protein structure) were tested and compared to the single S168R mutant.

[00451] None of the double mutations conferred improved editing efficiencies under the conditions tested (**FIGs. 53C**). The editing efficiencies of the S168R variant of MG29-1 and MG29-1 WT were compared in parallel with 12 guides targeting mouse albumin intron 1 and 1 guide targeting the human T cell receptor locus (TRAC). The S168R variant of MG29-1 exhibited improved editing efficiency with all 13 guides with some guides benefiting more than others (Figure 4d). Importantly S168R did not impair mammalian editing efficiency for any of the guides tested. These results demonstrate that the S168R (serine at amino acid position 168 changed to arginine) variant of MG29-1 has improved editing activity and which is advantageous in identifying highly active guides for therapeutic use.

Example 35 – Identification of chemical modifications of the sgRNA of nucleases described herein that improve guide stability and improve editing efficiency in mammalian cells

[00452] RNA molecules are inherently unstable in biological systems due to their sensitivity to cleavage by nucleases. Modification of the native chemical structure of RNA has been widely used to improve the stability RNA molecules used for RNA interference (RNAi) in the context for therapeutic drug development (Corey, J Clin Invest. 2007 Dec 3; 117(12): 3615–3622, J.B. Bramsen, J. Kjems Frontiers in Genetics, 3 (2012), p. 154). The introduction of chemical modifications to the nucleobases or the phosphodiester backbone of RNA have been pivotal in improving the stability and thus the potency of short RNA molecules *in vivo*. A wide range of chemical modifications with different properties in terms of stability against nucleases and affinity to complementary DNA or RNA have been developed.

[00453] Similar chemical modifications have been applied to the guide RNA for CRISPR Cas9 nucleases (Hendel et al, Nat Biotechnol. 2015 Sep; 33(9): 985–989, Ryan et al Nucleic Acids Res 2018 Jan 25;46(2):792-803., Mir et al Nature Communications volume 9, Article number: 2641 (2018), O' Reilly et al Nucleic Acids Res 2019 47, 546-558, Yin et al Nature Biotechnology volume 35, pages 1179–1187(2017), each of which is incorporated by reference herein in its entirety).

[00454] The MG29-1 nuclease is a novel nuclease with limited amino acid sequence similarity to identified Type V CRISPR enzymes such as cpf1. While the sequence of the structural (backbone) component of the guide RNA identified for MG29-1 is similar to that of cpf1

chemical modifications to the MG29-1 guide that enable improved stability while retaining activity had not been identified. A series of chemical modifications of the MG29-1 sgRNA were designed in order to evaluate their impact on sgRNA activity in mammalian cells and stability in the presence of mammalian cell protein extracts.

[00455] We selected the sgRNA mAlb29-1-8 which was highly active in the mouse liver cell line Hepa1-6 when the guide contained a set of proprietary chemical modifications developed by IDT called AltR1/AltR2 that were designed to improve the activity of the guide RNA for cpf1 and are available commercially (IDT). We selected to test 2 chemical modifications of the nucleobase; 2'-O-Methyl in which the 2' hydroxyl group is replaced with a methyl group, and 2'-fuoro in which the 2' hydroxyl group is replaced with a fluorine. Both 2'-O-Methyl and 2'-fluoro modifications improve resistance to nucleases. The 2'-O-methyl modification is a naturally occurring post-transcriptional modification of RNA and improves the binding affinity of RNA:RNA duplexes but has little impact on RNA:DNA stability. 2'-fluoro modified bases have reduced immunostimulatory effects and increase the binding affinity of both RNA: RNA and RNA:DNA hybrids (see e.g. Pallan et al Nucleic Acids Res 2011 Apr;39(8):3482-95, Chen et al Scientific Reports volume 9, Article number: 6078 (2019), Kawasaki, A. M. et al. J Med Chem 36, 831–841 (1993)).

[00456] The inclusion of phosphorothioate (PS) linkages in place of phosphodiester linkages between the bases was also evaluated. PS linkages improve resistance to nucleases (Monia et al Nucleic Acids, Protein Synthesis, and Molecular Genetics| Volume 271, ISSUE 24, P14533-14540, June 14, 1996).

[00457] The predicted secondary structure of the MG29-1 sgRNA with the spacer targeting mouse albumin intron 1 (mAlb29-1-8) is shown in **FIG. 54**. The stem-loop in the backbone portion of the guide was presumed to be critical for interaction with the MG29-1 protein based on sequence organization of other CRISPR-cas systems. Based on the secondary structure a series of chemical modifications was designed in different structural and functional regions of the guide. A modular approach was taken that allowed initial testing of guides with fewer chemical modifications that inform which structural and functional regions of the guide may tolerate different chemical modifications without significant loss of activity. The structural and functional regions were defined as follows. The 3' end and 5' end of the guide are targets for exonucleases and can be protected by various chemical modifications including 2'-O-methyl and PS linkages, an approach that has been used to improve the stability of guides for spCas9 (Hendel et al, Nat Biotechnol. 2015 Sep; 33(9): 985–989).

[00458] The sequences comprising both halves of the stem and the loop in the backbone region of the guide were selected for modification. The spacer was divided into the seed region (first 6 nucleotides closest to the PAM) and the remaining 16 nucleotides of the spacer (referred as the non-seed region). In total 43 guides were designed and 39 were synthesized. All 43 guides contain the same nucleotide sequence but with different chemical modifications. The editing activity of 39 of the guides was evaluated in Hepa1-6 cells by nucleofection of RNP or by co-transfection of mRNA encoding MG29-1 and guide or by both methods. These two methods of transfection may impact the observed activity of the guide due to differences in the delivery to the cell.

[00459] When nucleofection of a RNP is used the guide and the MG29-1 protein are pre-complexed in a tube and then delivered to the cell using nucleofection in which an electric current is applied to the cells' suspension in the presence of the RNP. The electric current transiently opens pores in the cell membrane (and possibly the nuclear membrane as well) enabling cellular entry of the RNP driven by the charge on the RNP. Whether the RNP enters the nucleus via pores created by the electric current or via the nuclear localization signals engineered in the protein component of the RNP, or a combination of the two is unclear.

[00460] When co-transfection of mRNA and guide with a lipid transfection reagent such as Messenger MAX is used, the mixture of the two RNA forms a complex with the positively charged lipid and the complex enters the cells via endocytosis and eventually reaches the cytoplasm. In the cytoplasm the mRNA is translated into protein. In the case of an RNA guided nucleases such as MG29-1 the resulting MG29-1 protein will presumably form a complex with the guide RNA in the cytoplasm before entering the nucleus in a process mediated by the nuclear localization signals that were engineered into the MG29-1 protein.

[00461] Because translation of the mRNA into sufficient amounts of MG29-1 protein followed by the binding of the MG29-1 protein to the guide RNA takes a finite amount of time, the guide RNA may require increased stability in the cytoplasm for longer than is the case when pre-formed RNP is delivered by nucleofection. Thus lipid-based mRNA/sgRNA co-transfection may require a more stable guide than is the case for RNP nucleofection which may result in some guide chemistries being active as RNP but inactive when co transfected with mRNA using cationic lipid reagents.

[00462] Guides mAlb298-1 to mAlb298-5 contain chemical modifications limited to the 5' and 3' ends of the sequence using a mixture of 2'-O-methyl and 2' fluoro bases plus PS linkages. In comparison to the sgRNA without chemical modifications these guides were 7 to 11-fold more active when delivered via RNP demonstrating that end modifications to the guide improved

guide activity, presumably through improved resistance to exonucleases. sgRNA mAlb298-1 to mAlb298-5 exhibited 64 to 114% of the editing activity of the guide containing the commercial chemical modifications (AltR1/AltR2). Guide 4, which contains the largest number of chemical modifications, was the least active of the end modified guides but was still 7-fold more active than the un-modified guide. Guide mALB298-30 contains three 2'-O methyl bases and 2 PS linkages at the 5' end and 4 2'-O methyl bases and 3 PS linkages at the 5' end and also exhibited activity about 10-fold higher than the unmodified guide and similar or slightly improved in the case of RNA co-transfection compared to mAlb298-1. These data demonstrate that 2'O-methyl combined with PS linkages on both ends of the MG29-1 guide significantly enhanced guide activity compared to an unmodified guide.

[00463] A combination of 2'-fluoro bases and PS linkages were also tolerated at the 3' end of the guide. Guide mALb298-28 contains three 2'-fluoro bases and 2 PS linkages on the 5' end and four 2'-fluoro bases and three PS linkages on the 3' end. This end modified guide retained good editing activity similar to the guides with 2'-O methyl and PS modifications on both ends demonstrating that 2'-fluoro can be used in place of 2'-O methyl to improve guide stability and retain editing activity.

[00464] The sgRNAs mALb298-6, mALb298-7, and mALb298-8 contain the same minimal chemical modifications on the both 5' and 3' ends present in mAlb298-1 plus PS linkages in different regions of the stem. PS linkages in the 3' stem (mALb298-6) and the 5' stem (mALb298-7) reduced activity by about 30% compared to mAlb298-1 in the RNP nucleofection assay, indicating that these modifications may be tolerated. Larger reductions in activity were observed by lipid-based transfection.

[00465] Introducing PS linkages in both the 3' and 5' stems (mALb298-8) reduced activity by about 80% compared to mAlb298-1 in the RNP nucleofection assay and by more than 95% in the lipid transfection assay, indicating that the combination of two PS linkage modifications significantly impaired the function of the guide.

[00466] The sgRNA mAlb298-9 contains the same minimal chemical modifications on the both 5' and 3' ends present in mAlb298-1 plus PS linkages in the loop and exhibited similar activity as mAlb298-1 indicating that PS linkages in the loop were well tolerated.

[00467] The sgRNAs mAlb298-10, mAlb298-11, and mAlb298-12 contain the same minimal chemical modifications on the both 5' and 3' ends present in mAlb298-1 plus 2'-O methyl bases in different regions of the stem. Including 2'-O methyl bases in either the 3' stem (mAlb298-11) or the 5' stem (mAlb298-12) or both halves of the stem (mAlb298-10) was generally well

tolerated with small reductions in activity compared to mAlb298-1 with guide mAlb298-12 (5' stem modified) being the most active.

[00468] Guide mAlb298-14 contains the same minimal chemical modifications on the both 5' and 3' ends present in mAlb298-1 plus a combination of 2'-O-methyl bases and PS linkages in both halves of the stem and had no editing activity by RNP nucleofection or by lipid-based RNA co-transfection. This confirms and extends the result with mAlb298-8 that contained only PS linkages in both stems had retained low levels of activity and shows that extensive chemical modification of both halves of the stem makes the guide inactive.

[00469] The sgRNA mAlb298-13 contains the same minimal chemical modifications on the both 5' and 3' ends present in mAlb298-1 plus PS linkages spaced every other base throughout the remainder of the backbone and spacer except for in the seed region of the spacer. These modifications resulted in a dramatic loss of editing activity to close to background levels. While the purity of this guide was about 50% compared to >75% for most of the guides, this alone may not account for the complete loss of editing activity. Thus, distributing PS linkages in an essentially random fashion throughout the guide is not an effective approach to improve guide stability while retaining editing activity.

[00470] Guides mALb298-15 and mALb298-16 contain the same minimal chemical modifications on the both 5' and 3' ends present in mAlb298-1 plus extensive PS linkages in the backbone. While both guides retained about 35% of the activity of mAlb298-1 by RNP nucleofection they retained 3% of the activity of mAlb298-1 by lipid-based RNA co-transfection indicating that extensive PS modification of the backbone significantly reduced editing activity. Combining the PS linkages in the backbone with PS linkages in the spacer region as in mAlb298-17 and mAlb298-18 resulted in further loss of activity consistent with the observation the random inclusion of PS linkages is blocks the ability of the guide to direct editing by MG29-1.

[00471] Guide mAlb298-19 contains the same chemical modifications in the spacer as mALb298-1 but in the backbone region the 5' end has additional 4 2'O-methyl bases and an additional 14 PS linkages. The activity of mAlb298-19 was about 40% of that of mAlb298-1 by RNP nucleofection but 22% by RNA co-transfection demonstrating again that extensive chemical modifications in the backbone region of the guide are not well tolerated.

[00472] Guides mAlb298-20, mAlb298-21, mAlb298-22, and mAlb298-23 have identical chemical modifications in the backbone region comprised of a single 2'-O methyl and 2 PS linkages at the 5' end which are the same 5' end modifications as in mAlb298-1. The spacer regions of Guides mAlb298-20, mAlb298-21, mAlb298-22, and mAlb298-23 contain combinations of 2'-O-methyl and 2'-fluoro bases as well as PS linkages. The most active of these

4 guides was mAlb298-2 in which 2'-fluoro modifications were made on all bases in the spacer except for the 7 bases closest to the PAM (seed region) and the terminal base at the 3' end which was modified with a 2'-O-methyl and 2 PS linkages. This demonstrates that including 2'-fluoro modifications on most of the spacer except for the seed region did not significantly reduce activity and thus represents a good strategy to enhance guide stability.

[00473] Guides mAlb298-24, mAlb298-25, mAlb298-26, and mAlb298-8 have identical chemical modifications in the backbone. mAlb298-8 which has PS linkages in both halves on the stem had significantly reduced editing activity with 24% and 2% of guide mAlb298-1 demonstrating that these PS linkages impaired activity. Interestingly, while mALb298-24 and mALb298-25 also had low editing activity the activity of mALb298-26 was improved compared to mAlb298-8 indicating that the additional modifications in mALb298-26 which comprise 2'-fluoro bases in 14 of the bases in the spacer (excluding the seed region) at least partially rescued the reduced editing activity caused by the PS linkages in the stem. This provides additional evidence of the beneficial impact of 2'-fluoro bases in the spacer upon editing activity.

[00474] Guides mAlb298-27 and mAlb298-29 contain extensive base and PS modifications throughout the backbone and spacer regions had no activity again indicating that not all chemical modifications of the guide retain editing activity.

[00475] Based on the structure activity relationships obtained from the analysis of guides mALb298-1 to mALb298-30, an additional set of seven guides were designed and tested by RNP nucleofection and lipid-based RNA co-transfection of Hepa1-6 cells. These guides combined chemical modifications that were observed to retain good editing activity in guides mALb298-1 to mALb298-30. Guides mALb298-31 to mALb298-37 all contain end modifications comprised of at least one 2'-O methyl and 2 PS linkages at the 5' end and one 2'-O methyl and 1 PS linkage at the 5' end. In addition to the end modifications, combining 2'-O methyl bases in both halves of the stem with 2'fluoro bases in 14 bases of the spacer (excluding the seed region) as in mAlb298-31 resulted in editing activity that was slightly improved or similar to end modifications alone and 10-fold improved compared to the unmodified guide. Combining 2'-O methyl bases in just the 5' stem with 2'fluoro bases in 14 bases of the spacer (excluding the seed region) as in mALb298-32 resulted in a guide that was among the most active tested.

[00476] Similarly, combining PS linkages in just the loop with 2'fluoro bases in 14 bases of the spacer (excluding the seed region) as in mALb298-33 resulted in potent activity up to 15-fold higher than the unmodified guide. Guide mAlb298-37 combines more extensive 3' end modifications with 2'-O methyl bases in the 5' stem, PS linkages in the loop and 14 2'fluoro bases and 3 PS linkages in the spacer (excluding the seed region) and still retained editing

activity similar to that of the AltR1/R2 modifications and significantly improved compared to the unmodified guide. mALb298-37 thus represents a heavily modified MG29-1 guide that retains potent editing activity in mammalian cells. Guide mALb298-38 exhibited potent editing activity when delivered as a RNP but was completely inactive when delivered to cells by lipid-based RNA co-transfection suggesting that thus guide may have some unexpected sensitivity to nucleases. Guide mALb298-39 which is identical to guide mAlb298-37 except that it has 11 fewer 2'-fluoro bases and 1 less PS linkage in the spacer had the highest editing activity when considering both RNP and mRNA transfection methods but has fewer chemical modifications than some of the other guide designs which might be detrimental in terms of performance *in vivo*.

[00477] Additional combinations of chemical modifications were designed to create mAlb298-40 to mALb298-43 that might also retain good editing activity while having more extensive chemical modifications. For example, in mAlb298-41 which also incorporates some DNA bases, 6 of the bases are un-modified ribonucleotides. Similarly, mAlb298-42 contains 2'-fluorogroups throughout the entire spacer and has 5 un-modified ribonucleotides. We envisage that testing of these and other guide chemical modifications will lead to one or more optimized designs. Nevertheless, within the set of guides mALb298-1 to mALb298-39 and particularly among the set of guides mALb298-31 to mALb298-39 we have identified guides with extensive chemical modifications that retain editing activity similar or superior to that of unmodified guides or guides with just end modifications.

[00478] In order to test the stability of the chemically modified guides compared to the guide with no chemical modification (native RNA), a stability assay using cell crude extracts was used. Crude cell extracts from mammalian cells were selected because they should contain the mixture of nucleases that a guide RNA will be exposed to when delivered to mammalian cells *in vitro* or *in vivo*. Hepa 1-6 cells were collected by adding 3ml of cold PBS per 15cm dish of confluent cells and releasing the cells from the surface of the dish using a cell scraper. The cells were pelleted at 200g for 10min and frozen at -80°C for future use. For the stability assays, cells were resuspended in 4 volumes of cold PBS (e.g. for a 100mg pellet cells were resuspended in 400µl of cold PBS). Triton X-100 was added to a ending concentration of 0.2% (v/v), cells were vortexed for 10 seconds, put on ice for 10 minutes and vortexed again for 10 seconds. Triton X-100 is a mild non-ionic detergent that disrupts cell membranes but does not inactivate or denature proteins at the concentration used.

[00479] Stability reactions were set up on ice and comprised 20 µl of cell crude extract with 100 fmoles of each guide (1µl of a 100nM stock). Six reactions were set up per guide comprising: input, 15min, 30min, 60min, 240min and 540min (The time in minutes referring to the length of

time each sample was incubated). Samples were incubated at 37°C from 15 minutes up to 540 min while the input control was left on ice for 5 minutes. After each incubation period the reaction was stopped by adding 300µl of a mixture of phenol and guanidine thiocyanate (Tri reagent, Zymo Research) which immediately denatures all proteins and efficiently inhibits ribonucleases and facilitates the subsequent recovery of RNA. After adding Tri Reagent the samples were vortexed for 15seconds and stored at -20°C. RNA was extracted from the samples using Direct-zol RNA miniprep kit (Zymo Research) and eluted in 100µl of nuclease-free water. Detection of the modified guide was performed using Taqman RT - qPCR using the Taqman miRNA Assay technology (Thermo Fisher) and primers and probes designed to specifically detect the sequence in the mAlb298 sgRNA which is the same for all of the guides. Data was plotted as a function of percentage of sgRNA remaining in relation to the input sample. The guide with no chemical modifications was the most rapidly eliminated when incubated with the cell extract (**FIG. 55**) with more than 90% of the guide degraded within 30 minutes. The guide with the AltR1/AltR2 (AltR in **FIG. 55**) chemical modifications was slightly more stable in the presence of cell extract than the un-modified guide with about 80% of the guide degraded in 30 minutes. Guide mALb298-31 that contains chemical modifications at both ends as well as 2' O-methyl bases in both stems and 2'-fluoro bases at all positions of the spacer except for the seed region was significantly more stable than either unmodified guide or the AltR guide.

[00480] Guide mAlb298-34 exhibited improved stability compared to guide mALb298-31. Guide mALb298-34 differs to guide mALb298-31 in the chemical modifications within the spacer. mALb298-34 has 9 fewer 2'-Fluoro bases in the spacer than mALb298-31 but contains 4 PS linkages in the spacer compared to 2 PS linkages in mALb298-31. Because 2'-fluoro bases improve the stability of RNA this suggests that the additional PS linkages in the spacer were responsible for the improved stability of mALb298-34 compared to mALb298-31.

[00481] Guide mALb298-37 was the most stable of all the guides tested and was significantly more stable than mALb298-34 with 80% of the guide remaining after 240 min (4 h) compared to 30% for mALb298-34. The chemical modifications of mALb298-37 differ from guide mALb298-34 in both the spacer and backbone regions. mALb298-37 has an additional two 2'-O-methyl groups and 2 additional PS linkages at the 5' end. In addition, the loop region of mALb298-37 contains PS linkages and does not contain the 2'-O-methyl groups present in the second half of the stem in mALb298-34. In addition, the spacer of mALb298-37 contains 9 more 2'-fluoro bases but the same number of PS linkages as mALb298-34 albeit in different locations.

[00482] Overall, these data suggest that additional PS linkages at the 5' end of the spacer and in the loop of the backbone region significantly improve stability of the guide RNA. Guide

mALb298-37 which exhibited the greatest stability in the cell extracts among the guides tested also exhibited potent editing activity in Hepa1-6 cells that was similar or improved compared to the AltR1/Altr2 modifications and improved compared to chemical modifications of the 5' and 3' ends only.

Table 5 - Impact of chemical modifications of the MG29-1 sgRNA sequence upon editing activity in mammalian cells

sgRNA name	sgRNA sequence	SEQ ID NO:	Editing activity (% of AltR1/AltR2 control)	
			RNP	mRNA
mAlb298-1_AltR1/R2	/AltR1/rCrUrUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrUrGrGrCrA/AltR2/	N/4272	100	100
mAlb298-0	rCrUrUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrUrGrGrCrA	4272	13.5	NT
mAlb298-1	mC*rU*rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrUrGrG*rC*mA	4273	114.7	76.2
mAlb298-2	mC*rU*rU*rArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrUrG*rG*rC*mA	4274	111.7	70.2
mAlb298-3	mC*mU*rU*rArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrUrG*rG*mC*mA	4275	100.2	63.7
mAlb298-4	mC*mU*mU*rArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrUrG*mG*mC*mA	4276	72.5	69.6
mAlb298-5	mC*rU*rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrUrG*/i2FG//i2FC/*/32FA/	4277	76.9	87.5
mAlb298-6	mC*rU*rUrArArUrUrUrCrUrArCrUrGrUrUrG*rU*rA*rG*rArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrUrGrG*rC*mA	4278	89.4	40.4
mAlb298-7	mC*rU*rUrArArUrU*rU*rC*rU*rArCrUrGrUrUrGrUrArGrArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrUrGrG*rC*mA	4279	83.2	24.5
mAlb298-8 #	mC*rU*rUrArArUrU*rU*rC*rU*rArCrUrGrUrUrG*rU*rA*rG*rArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrUrGrG*rC*mA	4280	28.4	2.6

sgRNA name	sgRNA sequence	SEQ ID NO:	Editing activity (% of AltR1/AltR2 control)	
			RNP	mRNA
mAlb298-9	mC*rU*rUrArArUrUrUrCrUrArCrU*rG*rU*rU*rGrUrArGrArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrUrGrG*rC*mA	4281	110.9	87.5
mAlb298-10	mC*rU*rUrArArUrUmUmCmUmArCrUrGrUrUrGmUmAmGmArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrUrGrG*rC*mA	4282	87.5	61.6
mAlb298-11	mC*rU*rUrArArUrUrUrCrUrArCrUrGrUrUrGmUmAmGmArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrUrGrG*rC*mA	4283	94.5	63.5
mAlb298-12	mC*rU*rUrArArUrUmUmCmUmArCrUrGrUrUrGrUrArGrArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrUrGrG*rC*mA	4284	121.8	84.0
mAlb298-13 #	mC*rU*rU*rArA*rUrU*rUrC*rUrA*rCrU*rGrU*rUrG*rUrA*rGrA*rUrCrUrGrUrArA*rCrG*rArU*rCrG*rGrG*rArA*rCrU*rGrG*mC*mA	4285	1.0	0.0
mAlb298-14	mC*rU*rUrArArUrU*mU*mC*mU*mArCrUrGrUrUrG*mU*mA*mG*mArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrUrGrG*rC*mA	4286	0.0	0.0
mAlb298-15 #	mC*rU*rU*rArArU*rU*rU*rC*rU*rA*rC*rU*rG*rU*rU*rG*rU*rA*rG*rA*rUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrUrGrG*rC*mA	4287	39.3	2.4
mAlb298-16	mC*rU*rU*rArArU*rU*rU*rC*rUrArCrU*rG*rU*rU*rG*rU*rA*rG*rA*rUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrUrGrG*rC*mA	4288	41.6	3.7
mAlb298-17 #	mC*rU*rU*rArArU*rU*rU*rC*rUrArCrU*rG*rU*rU*rG*rU*rA*rG*rA*rUrCrUrGrUrArA*rCrG*rArU*rCrG*rGrG*rArA*rCrU*rGrG*mC*mA	4289	0.0	1.2
mAlb298-18	mC*rU*rU*rA*rA*rU*rU*rU*rC*rU*rA*rC*rU*rG*rU*rU*rG*rU*rA*rG*rA*rUrCrUrGrUrArA*rCrG*rArU*rCrG*rGrG*rArA*rCrU*rGrG*mC*mA	4290	5.2	1.2
mAlb298-19	mG*mU*mA*mG*mC*rU*rU*rArA*rUrU*rUrC*rUrA*rCrU*rGrU*rUrG*rUrA*rGrA*rUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrUrGrG*rC*mA	4291	50.1	17.4
mAlb298-20	mC*rU*rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArC*/i2FU//i2FG/*/i2FG//i2FC/*/32FA/	4292	33.6	86.3

sgRNA name	sgRNA sequence	SEQ ID NO:	Editing activity (% of AltR1/AltR2 control)	
			RNP	mRNA
mAlb298-21	mC*rU*rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrArArC/i2FG//i2FA//i2FU//i2FC//i2FG//i2FG//i2FG//i2FA//i2FA//i2FC//i2FU//i2FG//i2FG/*/i2FC/*mA	4293	119.0	80.6
mAlb298-22	mC*rU*rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrArArCrGrArUrCrGrGrGrA*rA/i2FC/*/i2FU//i2FG/*/i2FG//i2FC/*mA	4294	25.1	98.8
mAlb298-23	mC*rU*rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrArArCrGrArUrCrGrGrGrA/i2FA//i2FC//i2FU//i2FG//i2FG/*mC*mA	4295	22.6	61.9
mAlb298-24	mC*rU*rUrArArUrU*rU*rC*rU*rArCrUrGrUrUrG*rU*rA*rG*rArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArC*/i2FU//i2FG/*/i2FG//i2FC/*/32FA/	4296	7.4	12.2
mAlb298-25	mC*rU*rUrArArUrU*rU*rC*rU*rArCrUrGrUrUrG*rU*rA*rG*rArUrCrUrGrUrArArCrGrArUrCrGrG/i2FG//i2FA/*/i2FA//i2FC/*/i2FU//i2FG/*rGrC*mA	4297	0.0	0.0
mAlb298-26	mC*rU*rUrArArUrU*rU*rC*rU*rArCrUrGrUrUrG*rU*rA*rG*rArUrCrUrGrUrArArC/i2FG//i2FA//i2FU//i2FC//i2FG//i2FG//i2FG//i2FA//i2FA//i2FC//i2FU//i2FG//i2FG/*/i2FC/*mA	4298	55.5	29.8
mAlb298-27	/52FC/*/i2FU/*/i2FU/*rUrUrArArU/i2FU//i2FU/rC*rU/i2FA/*/i2FC/rU/i2FG/*/i2FU//i2FU/rG/i2FU//i2FA//i2FG//i2FA/rU/i2FC/rUrG*rUrA/i2FA/rC/i2FG/*/i2FA/*/i2FU/*/i2FC/*/i2FG/rG*rGrA*rA/i2FC/*rU*/i2FG//i2FG/*/i2FC/*/32FA	4299	NT	0
mAlb298-28	/52FC/*/i2FU/*/i2F/rUrUrUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrU*/i2FG//i2FG/*/i2FC/*/52FA/	4300	NT	84.8
mAlb298-29	mC*mU*mU*rUrUrArArUmUmUrC*rUmA*mCrUmG*mUmUrGmUmAmGmArUmCrUrG*rUrAmArCmG*mA*mU*mC*mGrG*rGrA*rAmC*rU*mGmG*mC*mA	4301	0.0	0.0
mAlb298-30	mC*mU*mUrUrUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrU*mGmG*mC*mA	4302	101.1	105.4

sgRNA name	sgRNA sequence	SEQ ID NO:	Editing activity (% of AltR1/AltR2 control)	
			RNP	mRNA
mAlb298-31	mC*rU*rUrArArUrUmUmCmUmArCrUrGrUrUrGmUmAmGmArUrCrUrGrUrArArC/i2FG//i2FA//i2FU//i2FC//i2FG//i2FG//i2FG//i2FA//i2FA//i2FC//i2FU//i2FG//i2FG/*/i2FC/*mA	4303	140.5	74.4
mAlb298-32	mC*rU*rUrArArUrUmUmCmUmArCrUrGrUrUrGrUrArGrArUrCrUrGrUrArArC/i2FG//i2FA//i2FU//i2FC//i2FG//i2FG//i2FG//i2FA//i2FA//i2FC//i2FU//i2FG//i2FG*/i2FC/*mA	4304	170.3	93.1
mAlb298-33	mC*rU*rUrArArUrUrUrCrUrArCrU*rG*rU*rU*rGrUrArGrArUrCrUrGrUrArArC/i2FG//i2FA//i2FU//i2FC//i2FG//i2FG//i2FG//i2FA//i2FA//i2FC//i2FU//i2FG//i2FG/*/i2FC/*mA	4305	202.7	64.4
mAlb298-34	mC*rU*rUrArArUrUmUmCmUmArCrUrGrUrUrGmUmAmGmArUrCrUrGrUrArArCrGrArUrCrGrGrGrA*rA/i2FC/*/i2FU//i2FG/*/i2FG//i2FC/*mA	4306	83.8	107.0
mAlb298-35	mC*rU*rUrArArUrUrUrCrUrArCrU*rG*rU*rU*rGrUrArGrArUrCrUrGrUrArArCrGrArUrCrGrGrGrA*rA/i2FC/*/i2FU//i2FG/*/i2FG//i2FC/*mA	4307	24.3	67.9
mAlb298-36	mC*rU*rUrArArUrUmUmCmUmArCrUrGrUrUrGrUrArGrArUrCrUrGrUrArArCrGrArUrCrGrGrGrA*rA/i2FC/*/i2FU//i2FG/*/i2FG//i2FC/*mA	4308	43.2	116.2
mAlb298-37	mC*mU*mU*U*rUrArArUrUmUmCmUmArCrU*rG*rU*rU*rGrUrArGrArUrCrUrGrUrArArC/i2FG//i2FA//i2FU//i2FC//i2FG//i2FG//i2FG//i2FA//i2FA//i2FC/*/i2FU/*/i2FG//i2FG/*/i2FC/*mA	4309	164.9	84.5
mAlb298-38	mC*mU*mU*rU*rUrArArUrUmUmCmUmArCrU*rG*rU*rU*rGmUmAmGmArUrCrUrGrUrArArC/i2FG//i2FA//i2FU//i2FC//i2FG//i2FG//i2FG//i2FA//i2FA//i2FC/*/i2FU/*/i2FG//i2FG/*/i2FC/*mA	4310	140.5	0.0
mAlb298-39	mC*mU*mU*rU*rUrArArUrUmUmCmUmArCrU*rG*rU*rU*rGrUrArGrArUrCrUrGrUrArArCrGrArUrCrGrGrGrArArCrU*/i2FG//i2FG/*/i2FC/*mA	4311	135.1	114.0
mAlb298-40	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAU/i2FC//i2FU//i2FG//i2FU//i2FA//i2FA//i2FC//i2FG//i2FA//i2FU//i2FC//i2FG//i2FG//i2FG//i2FA//i2FA//	4312	NT	NT

sgRNA name	sgRNA sequence	SEQ ID NO:	Editing activity (% of AltR1/AltR2 control)	
			RNP	mRNA
	i2FC/*/i2FU/*/i2FG//i2FG/*/i2FC/*mA			
mAlb298-41	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*dGdTdAdGdAdT/i2FC//i2FU//i2FG//i2FU//i2FA//i2FA//i2FC//i2FG//i2FA//i2FU//i2FC//i2FG//i2FG//i2FG//i2FA//i2FA//i2FC/*/i2FU/*/i2FG//i2FG/*/i2FC/*mA	4313	NT	NT
mAlb298-42	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*/i2FG//i2FU//i2FA//i2FG//i2FA//i2FU//i2FC//i2FU//i2FG//i2FU//i2FA//i2FA//i2FC//i2FG//i2FA//i2FU//i2FC//i2FG//i2FG//i2FG//i2FA//i2FA//i2FC/*/i2FU/*/i2FG//i2FG/*/i2FC/*mA	4314	NT	NT
mAlb298-43	mC*mU*mU*U*UAAUUmUmCmUmAmCU*G*U*U*GUAGAU/i2FC//i2FU//i2FG//i2FU//i2FA//i2FA//i2FC//i2FG//i2FA//i2FU//i2FC//i2FG//i2FG//i2FG//i2FA//i2FA//i2FC/*/i2FU/*/i2FG//i2FG/*/i2FC/*mA	4315	NT	NT

#: these guides had less than 75% purity based on analytical HPLC with purity ranging from 54 to 64%. All other guides exceeded 75% purity
NT: not tested
Nomenclature of chemical modifications: a "/" is used to separate bases with 2'-flourine modifications, m; 2'-O-methyl base (for example a A base with 2'-O-methyl modification is written as mA), i2F; internal 2'-flourine base (for example an internal C with 2'-flourine modification is written as /i2FC/), 52F; 2'-flourine base at the 5' end of the sequence (for example a 5' C with 2'-flourine modification is written as /52FC/), 32F; 2'-flourine base at the 3' end of the sequence (for example a 3' A base with 2'-flourine modification is written as /32FA/), r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *; between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur; AltR1 and AltR2 refer to IDT technologies' proprietary 5' and 3' AltR modifications

Example 12 – Therapeutic gene editing in mice using nucleases described herein

[00483] Gene editing platforms described herein have the potential to effect reparative alterations *in vivo*. Liver tissue is an example of a tissue that can be advantageously targeted using the gene editing compositions and systems described herein for *in vivo* gene editing, for example by introduction of indels that function to knock down expression of deleterious genes or that are used to replace defective genes. For example, several inherited diseases arise from defects in proteins expressed primarily in the liver, and *in vivo* delivery to the liver has been proven safe and effective in clinical trials of adeno-associated virus (AAV) vectors. Lipid nanoparticles have also been shown to deliver nucleic acids and approved drugs for RNAi strategies. Liver tissue

also includes appropriate cellular machinery for efficient secretion of proteins into the systemic circulation.

[00484] Subjects having a condition in Table 13 or Table 14 are selected for gene editing therapy. For example, a human or mouse model subject having hemophilia A is identified for treatment with gene replacement therapy using a gene editing platform.

Table 6 – Some Indications for Subject Selection in Therapeutic Gene Replacement

Hemophilia A	**Factor VIII**		**1 in 5,000 males**
Hemophilia A	**Factor VIII**	**Secreted**	**1 in 5,000 males**
Hemophilia B	**Factor IX**	**Secreted**	**1 in 20,000**
Hereditary Angioedema	**C1 inhibitor protein**	**Secreted**	**1 in 25,000**
Argininosuccinate Lyase deficiency	**Argininosuccinate Lyase**	**Intracellular**	**1 in 70,000**
Mucopolysaccharidosis type IV (MPS IV),	**Arylsulfatase B**	**Intracellular**	**1 in 200,000**
Progressive familial intrahepatic cholestasis type 2	**ATP binding cassette family B**	**Intracellular**	**1 in 50,000**
Classical galactosemia	**Galactose-1-phosphate uridyltransferase**	**Intracellular**	**1 in 50,000**

Table 7 – Some Indications for Subject Selection in Therapeutic Gene Knockdown

Indication	Target Gene	Prevalence
Primary Hyperoxluria type I	**Glyoxylate oxidase (HAO1)**	**Est 1 in 100,000, up to 5,000 patient in US+EU**
Familial ATTR Amyloidoisis	**Transthyretin**	**1 in 100,000 in US, more frequent in Japan, Sweden**
Acute Hepatic Porphoryia	**Aminolevulinic Acid Synthase (ALAS1)**	**1 in 50,000**
Cardiovascular disease without adequate LDL lowering by statins	**PCSK9**	**High (1 in 3 deaths in US due to CVD)**

Indication	Target Gene	Prevalence
Rare Hyperlipidemias	Angiopoietin like 3	Various, approx 1 in 500,000
Homozygous Familial Hypercholersterolemia	ApoB100	1 in 1 million
Hereditary Angioedema	Kallikrein	1 in 25,000

[00485] A gene editing platform comprising a lipid nanoparticle (LNP) encapsulating an sgRNA and an mRNA encoding an MG nuclease described herein and an AAV (e.g., AAV serotype 8) comprising a donor template nucleic acid encoding a therapeutic gene are introduced into the liver intravenously to the subject. The LNP is targeted to hepatocytes via surface functionalization of the LNPs.

[00486] For example, the subject having hemophilia A is treated with a gene replacement platform comprising LNPs containing mRNA encoding a MG29-1 nuclease described herein (SEQ ID NO: 214). LNPs also contain sgRNA specific for albumin I, which is highly expressed in the liver (e.g., albumin can be expressed at about 5 g/dL in the liver, whereas factor VIII can be expressed at about 10 μg/dL in the liver, or 1 million times less than albumin). In addition to the LNPs, AAV8 (AAV serotype 8) viral particles comprising plasmids, which encode replacement template DNA encoding a replacement factor VIII nucleotide sequence, are delivered to the subject as well. Once inside the cell, the mRNA, sgRNA, and template DNA are transiently expressed. The MG29-1 nuclease targets the target locus of the host hepatocyte DNA using the sgRNA and then cleaves the host DNA. The donor template DNA transcribed from the plasmid delivered to the host hepatocyte in the AAV8 is spliced into the cell and stably integrated into the host DNA at the target site of the albumin I gene, and the inserted factor VIII DNA is expressed under the albumin promoter.

[00487] The gene editing platform is also used in subjects selected for gene knockdown therapy. For instance, a subject presenting with familial ATTR amyloidosis is treated with LNPs containing mRNA encoding an MG29-1 nuclease described herein (SEQ ID NO: 214) and a sgRNA specific to a target site in the transthyretin gene. The MG29-1 nuclease and sgRNA are delivered to and expressed in hepatocytes of the subject. In some embodiments, the sgRNA is targeted to a stop codon of the transthyretin gene, and the MG29-1 nuclease's activity removes the endogenous stop codon, effectively knocking down the expression of the gene. In some embodiments, the gene knockdown platform comprises an AAV8 containing a plasmid encoding a polynucleotide comprising a stop codon. When the AAV8 is delivered to the same cell that is expressing the nuclease and sgRNA, an exogenous stop codon is spliced into the tranthyretin

gene, leading to knockdown of the gene's expression as a result of premature truncation of proteins translated from RNA produced from the edited DNA.

Example 13 – Gene editing outcomes at the DNA level for CD38

[00488] Primary NK cells were expanded using the NK Cloudz system (R&D Systems) according to the manufacturer's recommendations. Nucleofection of MG29-1 RNPs (212 pmol protein/320 pmol guide) (guide SEQ ID NOs: 4428-4465) was performed into NK cells (500,000) using the Lonza 4D electroporator. Cells were harvested and genomic DNA prepared five days post-transfection. PCR primers appropriate for use in NGS-based DNA sequencing were generated, optimized, and used to amplify the individual target sequences for each guide RNA (SEQ ID NOs: 4466-4503). The amplicons were sequenced on an Illumina MiSeq machine and analyzed with a proprietary Python script to measure gene editing (**FIG. 57**).

Table 14A: Sequences of Guide RNAs and Sequences Targeted Thereby for Example 37

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting CD38	4428	MG29-1-CD38-sgRNA-A1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrGrArGrArCrCrGrUrCrCrUrGrGrCrGrCrGrArUrG/AltR2/
MG29-1 sgRNA targeting CD38	4429	MG29-1-CD38-sgRNA-B1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrUrGrUrArCrUrUrGrArCrGrCrArUrCrGrCrGrCrC/AltR2/
MG29-1 sgRNA targeting CD38	4430	MG29-1-CD38-sgRNA-C1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrCrGrGrArCrArCrCrGrGrGrCrUrGrArArCrU/AltR2/
MG29-1 sgRNA targeting CD38	4431	MG29-1-CD38-sgRNA-D1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArGrGrGrUrGrCrArUrUrUrArUrUrUrCrArArArArC/AltR2/
MG29-1 sgRNA targeting CD38	4432	MG29-1-CD38-sgRNA-E1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrCrArArArArCrArUrCrCrUrUrGrCrArArCrArU/AltR2/
MG29-1 sgRNA targeting CD38	4433	MG29-1-CD38-sgRNA-F1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArArCrArUrCrCrUrUrGrCrArArCrArUrUrArCrU/AltR2/
MG29-1 sgRNA targeting CD38	4434	MG29-1-CD38-sgRNA-G1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArArArArUrArArUrGrCrArCrCrCrUrUrGrArArArG/AltR2/
MG29-1 sgRNA targeting CD38	4435	MG29-1-CD38-sgRNA-H1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArArUrArArUrGrCrArCrCrCrUrUrGrArArArGrC/AltR2/
MG29-1 sgRNA targeting CD38	4436	MG29-1-CD38-sgRNA-A2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrArGrUrCrUrArCrArUrGrUrCrUrGrArGrArUrArA/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting CD38	4437	MG29-1-CD38-sgRNA-B2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArGrCrArGrArArUrArArArArGrArUrCrUrGrGrCrC/AltR2/
MG29-1 sgRNA targeting CD38	4438	MG29-1-CD38-sgRNA-C2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrUrCrUrGrCrUrCrCrArArArGrArArGrArArUrCrU/AltR2/
MG29-1 sgRNA targeting CD38	4439	MG29-1-CD38-sgRNA-D2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrUrGrCrUrCrCrArArArGrArArGrArArUrCrUrA/AltR2/
MG29-1 sgRNA targeting CD38	4440	MG29-1-CD38-sgRNA-E2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrUrArUrUrCrUrGrGrArArArArCrGrGrUrUrUrCrC/AltR2/
MG29-1 sgRNA targeting CD38	4441	MG29-1-CD38-sgRNA-F2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrGrCrArGrGrGrUrArArGrUrArCrCrArArGrUrArG/AltR2/
MG29-1 sgRNA targeting CD38	4442	MG29-1-CD38-sgRNA-G2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrArGrArArUrArCrUrGrArArArCrArGrGrGrUrUrG/AltR2/
MG29-1 sgRNA targeting CD38	4443	MG29-1-CD38-sgRNA-H2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrArArUrArCrUrGrArArArCrArGrGrGrUrUrGrU/AltR2/
MG29-1 sgRNA targeting CD38	4444	MG29-1-CD38-sgRNA-A3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrArGrUrCrUrGrGrGrCrArArGrArUrUrGrArUrA/AltR2/
MG29-1 sgRNA targeting CD38	4445	MG29-1-CD38-sgRNA-B3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrCrUrArArArArGrArCrArUrArArGrUrUrUrGrUrA/AltR2/
MG29-1 sgRNA targeting CD38	4446	MG29-1-CD38-sgRNA-C3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrArArGrCrUrGrCrCrUrGrUrGrArUrGrUrGrUrGrU/AltR2/
MG29-1 sgRNA targeting CD38	4447	MG29-1-CD38-sgRNA-D3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrArArArArArCrArGrGrUrArCrArCrArUrUrUrUrArU/AltR2/
MG29-1 sgRNA targeting CD38	4448	MG29-1-CD38-sgRNA-E3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrUrCrArArArGrArUrUrUrUrArCrUrGrCrGrGrGrA/AltR2/
MG29-1 sgRNA targeting CD38	4449	MG29-1-CD38-sgRNA-F3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrCrArArArGrArUrUrUrUrArCrUrGrCrGrGrGrArU/AltR2/
MG29-1 sgRNA targeting CD38	4450	MG29-1-CD38-sgRNA-G3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrArArArGrArUrUrUrUrArCrUrGrCrGrGrGrArUrC/AltR2/
MG29-1 sgRNA	4451	MG29-1-CD38-sgRNA-H3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrUrGrCrGrGrGrArUrCrCrArUrUrGrArGrCrCrArUrC/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
targeting CD38			
MG29-1 sgRNA targeting CD38	4452	MG29-1-CD38-sgRNA-A4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrCrGrGrGrArUrCrCrArUrUrGrArGrCrArUrCrA/AltR2/
MG29-1 sgRNA targeting CD38	4453	MG29-1-CD38-sgRNA-B4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrGrArGrUrGrUrGrGrArArGrUrCrCrArUrArArArUrU/AltR2/
MG29-1 sgRNA targeting CD38	4454	MG29-1-CD38-sgRNA-C4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrArGrUrGrUrGrGrArArGrUrCrCrArUrArArUrUrU/AltR2/
MG29-1 sgRNA targeting CD38	4455	MG29-1-CD38-sgRNA-D4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArArCrCrArGrArGrArArGrGrUrUrCrArGrArCrArC/AltR2/
MG29-1 sgRNA targeting CD38	4456	MG29-1-CD38-sgRNA-E4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrUrGrCrArArGrArArUrArUrCrUrArCrArGrGrUrA/AltR2/
MG29-1 sgRNA targeting CD38	4457	MG29-1-CD38-sgRNA-F4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrCrArArGrArArUrArUrCrUrArCrArGrGrUrArA/AltR2/
MG29-1 sgRNA targeting CD38	4458	MG29-1-CD38-sgRNA-G4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrUrUrArUrArArUrCrGrArUrUrCrCrArGrCrUrCrU/AltR2/
MG29-1 sgRNA targeting CD38	4459	MG29-1-CD38-sgRNA-H4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrArUrArArUrCrGrArUrUrCrCrArGrCrUrCrUrU/AltR2/
MG29-1 sgRNA targeting CD38	4460	MG29-1-CD38-sgRNA-A5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrGrGrUrGrGrArUrCrCrUrGrGrCrArUrArArArGrU/AltR2/
MG29-1 sgRNA targeting CD38	4461	MG29-1-CD38-sgRNA-B5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrGrUrGrGrArUrCrCrUrGrGrCrArUrArArGrUrC/AltR2/
MG29-1 sgRNA targeting CD38	4462	MG29-1-CD38-sgRNA-C5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrArGrUrGrUrGrUrGrArArArArArUrCrCrUrGrA/AltR2/
MG29-1 sgRNA targeting CD38	4463	MG29-1-CD38-sgRNA-D5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrArCrArCrArCrUrGrArArGrArArArCrUrUrGrUrC/AltR2/
MG29-1 sgRNA targeting CD38	4464	MG29-1-CD38-sgRNA-E5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArCrArCrArCrUrGrArArGrArArArCrUrUrGrUrCrA/AltR2/
MG29-1 sgRNA targeting CD38	4465	MG29-1-CD38-sgRNA-F5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrArCrArCrUrGrArArGrArArArCrUrUrGrUrCrArG/AltR2/
DNA sequence	4466	MG29-1-	CCGAGACCGTCCTGGCGCGATG

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
of CD38 target site		CD38-target site-A1	
DNA sequence of CD38 target site	4467	MG29-1-CD38-target site-B1	AGTGTACTTGACGCATCGCGCC
DNA sequence of CD38 target site	4468	MG29-1-CD38-target site-C1	TCCCCGGACACCGGGCTGAACT
DNA sequence of CD38 target site	4469	MG29-1-CD38-target site-D1	AAGGGTGCATTTATTTCAAAAC
DNA sequence of CD38 target site	4470	MG29-1-CD38-target site-E1	TTTCAAAACATCCTTGCAACAT
DNA sequence of CD38 target site	4471	MG29-1-CD38-target site-F1	AAAACATCCTTGCAACATTACT
DNA sequence of CD38 target site	4472	MG29-1-CD38-target site-G1	GAAATAAATGCACCCTTGAAAG
DNA sequence of CD38 target site	4473	MG29-1-CD38-target site-H1	AAATAAATGCACCCTTGAAAGC
DNA sequence of CD38 target site	4474	MG29-1-CD38-target site-A2	GCAGTCTACATGTCTGAGATAA
DNA sequence of CD38 target site	4475	MG29-1-CD38-target site-B2	GAGCAGAATAAAAGATCTGGCC
DNA sequence of CD38 target site	4476	MG29-1-CD38-target site-C2	ATTCTGCTCCAAAGAAGAATCT
DNA sequence of CD38 target site	4477	MG29-1-CD38-target site-D2	TTCTGCTCCAAAGAAGAATCTA
DNA sequence of CD38 target site	4478	MG29-1-CD38-target site-E2	AGTATTCTGGAAAACGGTTTCC
DNA sequence of CD38 target site	4479	MG29-1-CD38-target site-F2	CCGCAGGGTAAGTACCAAGTAG
DNA sequence of CD38 target site	4480	MG29-1-CD38-target site-G2	CCAGAATACTGAAACAGGGTTG
DNA sequence of CD38 target site	4481	MG29-1-CD38-target site-H2	CAGAATACTGAAACAGGGTTGT
DNA sequence of CD38 target site	4482	MG29-1-CD38-target site-A3	TCCAGTCTGGGCAAGATTGATA
DNA sequence of CD38 target site	4483	MG29-1-CD38-target site-B3	TTTCTAAAAGACATAGTTTGTA
DNA sequence of CD38 target site	4484	MG29-1-CD38-target site-C3	CAGAAGCTGCCTGTGATGTGGT
DNA sequence of CD38 target site	4485	MG29-1-CD38-target site-D3	ACAAAAACAGGTACACATTTAT

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of CD38 target site	4486	MG29-1-CD38-target site-E3	TGTCAAAGATTTTACTGCGGGA
DNA sequence of CD38 target site	4487	MG29-1-CD38-target site-F3	GTCAAAGATTTTACTGCGGGAT
DNA sequence of CD38 target site	4488	MG29-1-CD38-target site-G3	TCAAAGATTTTACTGCGGGATC
DNA sequence of CD38 target site	4489	MG29-1-CD38-target site-H3	ACTGCGGGATCCATTGAGCATC
DNA sequence of CD38 target site	4490	MG29-1-CD38-target site-A4	CTGCGGGATCCATTGAGCATCA
DNA sequence of CD38 target site	4491	MG29-1-CD38-target site-B4	GGGAGTGTGGAAGTCCATAATT
DNA sequence of CD38 target site	4492	MG29-1-CD38-target site-C4	GGAGTGTGGAAGTCCATAATTT
DNA sequence of CD38 target site	4493	MG29-1-CD38-target site-D4	CAACCAGAGAAGGTTCAGACAC
DNA sequence of CD38 target site	4494	MG29-1-CD38-target site-E4	CCTGCAAGAATATCTACAGGTA
DNA sequence of CD38 target site	4495	MG29-1-CD38-target site-F4	CTGCAAGAATATCTACAGGTAA
DNA sequence of CD38 target site	4496	MG29-1-CD38-target site-G4	GCTTATAATCGATTCCAGCTCT
DNA sequence of CD38 target site	4497	MG29-1-CD38-target site-H4	CTTATAATCGATTCCAGCTCTT
DNA sequence of CD38 target site	4498	MG29-1-CD38-target site-A5	ATGGTGGGATCCTGGCATAAGT
DNA sequence of CD38 target site	4499	MG29-1-CD38-target site-B5	TGGTGGGATCCTGGCATAAGTC
DNA sequence of CD38 target site	4500	MG29-1-CD38-target site-C5	TTCAGTGTGTGAAAAATCCTGA
DNA sequence of CD38 target site	4501	MG29-1-CD38-target site-D5	TCACACACTGAAGAAACTTGTC
DNA sequence of CD38 target site	4502	MG29-1-CD38-target site-E5	CACACACTGAAGAAACTTGTCA
DNA sequence of CD38 target site	4503	MG29-1-CD38-target site-F5	ACACACTGAAGAAACTTGTCAG
r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *; between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur; AltR1 and AltR2 refer to IDT technologies' proprietary 5' and 3' AltR modifications			

Example 38 – Gene editing outcomes at the phenotypic level for CD38

[00489] Edited cells were assayed for CD38 expression as follows: primary NK cells (500,000) were washed with phosphate buffered saline (PBS) and stained with LIVE/DEAD™ Fixable Aqua Dead Cell Stain Kit (ThermoFisher) according to the manufacturer specifications. Cells were then washed and incubated with FC Block - Human TruStain FcX™ (Biolegend) for 20 minutes on ice. Next, cells were stained with anti-CD56 PE and anti-38 PerCP eFluor 710 antibodies (ThermoFisher) according to the manufacturer recommendations and analyzed by flow cytometry by collecting 25,000 total events per specimen (**FIG. 58**).

Example 39 – Gene editing outcomes at the DNA level for TIGIT

[00490] Primary T cells were purified from PMBCs using a negative selection kit (Miltenyi) according to the manufacturer's recommendations. Nucleofection of MG29-1 RNPs (106 pmol protein/160 pmol guide) (SEQ ID NOs: 4504-4520) was performed into T cells (200,000) using the Lonza 4D electroporator. Cells were harvested and genomic DNA prepared five days post-transfection. PCR primers appropriate for use in NGS-based DNA sequencing were generated, optimized, and used to amplify the individual target sequences for each guide RNA (SEQ ID NOs: 4521-4537). The amplicons were sequenced on an Illumina MiSeq machine and analyzed with a proprietary Python script to measure gene editing (**FIG. 59**).

Table 14B: Sequences of Guide RNAs and Sequences Targeted for Example 39

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting TIGIT	**4504**	**MG29-1-TIGIT-sgRNA-A1**	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrGrCrCrUrUrArCrCrUrGrArGrGrCrGrArGrGrGrG/AltR2/
MG29-1 sgRNA targeting TIGIT	4505	MG29-1-TIGIT-sgRNA-B1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArUrUrGrUrGrCrCrUrGrUrCrArUrCrArUrUrCrCrU/AltR2/
MG29-1 sgRNA targeting TIGIT	4506	MG29-1-TIGIT-sgRNA-C1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrGrCrArGrArArArUrGrUrUrCrCrCrCrGrUrUrG/AltR2/
MG29-1 sgRNA targeting TIGIT	4507	MG29-1-TIGIT-sgRNA-D1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrCrArGrArGrArArArGrGrUrGrGrCrUrCrUrArUrC/AltR2/
MG29-1 sgRNA targeting TIGIT	4508	MG29-1-TIGIT-sgRNA-E1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArArUrGrCrUrGrArCrUrUrGrGrGrGrGrCrArC/AltR2/
MG29-1 sgRNA	4509	MG29-1-TIGIT-sgRNA-F1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrGrArCrCrUrCrCrArGrGrArArGrArUrUrCrUrC/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
targeting TIGIT			
MG29-1 sgRNA targeting TIGIT	**4510**	**MG29-1-TIGIT-sgRNA-G1**	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrCrCrUrCrCrCrUrCrUrArGrUrGrGrCrUrGrArGrC/AltR2/
MG29-1 sgRNA targeting TIGIT	**4511**	**MG29-1-TIGIT-sgRNA-H1**	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrUrCrCrCrUrCrUrArGrUrGrGrCrUrGrArGrCrA/AltR2/
MG29-1 sgRNA targeting TIGIT	4512	MG29-1-TIGIT-sgRNA-A2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrUrCrArArCrGrCrGrArCrCrArCrCrArCrGrArU/AltR2/
MG29-1 sgRNA targeting TIGIT	4513	MG29-1-TIGIT-sgRNA-B2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrUrUrUrGrUrUrUrUrUrArGrArArGrA/AltR2/
MG29-1 sgRNA targeting TIGIT	4514	MG29-1-TIGIT-sgRNA-C2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrGrUrUrUrArGrArArGrArArArGrCrCrCrU/AltR2/
MG29-1 sgRNA targeting TIGIT	4515	MG29-1-TIGIT-sgRNA-D2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrArGrGrArArGrArArArGrCrCrCrUrCrArGrA/AltR2/
MG29-1 sgRNA targeting TIGIT	4516	MG29-1-TIGIT-sgRNA-E2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrArArGrArArArGrCrCrCrUrCrArGrArArUrCrC/AltR2/
MG29-1 sgRNA targeting TIGIT	4517	MG29-1-TIGIT-sgRNA-F2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrArArArGrCrCrCrUrCrArGrArArUrCrCrA/AltR2/
MG29-1 sgRNA targeting TIGIT	**4518**	**MG29-1-TIGIT-sgRNA-G2**	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArArArGrArArArGrCrCrCrUrCrArGrArArUrCrCrArU/AltR2/
MG29-1 sgRNA targeting TIGIT	**4519**	**MG29-1-TIGIT-sgRNA-H2**	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrUrGrArGrGrUrCrArCrCrUrUrCrCrArCrArGrA/AltR2/
MG29-1 sgRNA targeting TIGIT	4520	MG29-1-TIGIT-sgRNA-A3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrCrCrUrGrArGrGrUrCrArCrCrUrUrCrCrArCrArG/AltR2/
DNA sequence of TIGIT target site	**4521**	**MG29-1-TIGIT-target site-A1**	AGGCCTTACCTGAGGCGAGGGG
DNA sequence of TIGIT target site	4522	MG29-1-TIGIT-target site-B1	TATTGTGCCTGTCATCATTCCT
DNA sequence of TIGIT target site	4523	MG29-1-TIGIT-target site-C1	TCTGCAGAAATGTTCCCCGTTG

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of TIGIT target site	4524	MG29-1-TIGIT-target site-D1	TGCAGAGAAAGGTGGCTCTATC
DNA sequence of TIGIT target site	4525	MG29-1-TIGIT-target site-E1	TAATGCTGACTTGGGGTGGCAC
DNA sequence of TIGIT target site	4526	MG29-1-TIGIT-target site-F1	TAGGACCTCCAGGAAGATTCTC
DNA sequence of TIGIT target site	**4527**	**MG29-1-TIGIT-target site-G1**	GTCCTCCCTCTAGTGGCTGAGC
DNA sequence of TIGIT target site	**4528**	**MG29-1-TIGIT-target site-H1**	TCCTCCCTCTAGTGGCTGAGCA
DNA sequence of TIGIT target site	4529	MG29-1-TIGIT-target site-A2	TAGTCAACGCGACCACCACGAT
DNA sequence of TIGIT target site	4530	MG29-1-TIGIT-target site-B2	TAGTTTGTTTGTTTTTAGAAGA
DNA sequence of TIGIT target site	4531	MG29-1-TIGIT-target site-C2	TTTGTTTTTAGAAGAAAGCCCT
DNA sequence of TIGIT target site	4532	MG29-1-TIGIT-target site-D2	TTTTTAGAAGAAAGCCCTCAGA
DNA sequence of TIGIT target site	4533	MG29-1-TIGIT-target site-E2	TAGAAGAAAGCCCTCAGAATCC
DNA sequence of TIGIT target site	4534	MG29-1-TIGIT-target site-F2	AGAAGAAAGCCCTCAGAATCCA
DNA sequence of TIGIT target site	**4535**	**MG29-1-TIGIT-target site-G2**	GAAGAAAGCCCTCAGAATCCAT
DNA sequence of TIGIT target site	**4536**	**MG29-1-TIGIT-target site-H2**	CTCCTGAGGTCACCTTCCACAG
DNA sequence of TIGIT target site	4537	MG29-1-TIGIT-target site-A3	TCCTGAGGTCACCTTCCACAGA
r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *;			

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur; AltR1 and AltR2 refer to IDT technologies' proprietary 5' and 3' AltR modifications			

Example 14 – Gene editing outcomes at the DNA level for AAVS1

[00491] Primary T cells were purified from PMBCs using a negative selection kit (Miltenyi) according to the manufacturer's recommendations. Nucleofection of MG29-1 RNPs (106 pmol protein/160 pmol guide) (SEQ ID NOs: 4538-4568) was performed into T cells (200,000) using the Lonza 4D electroporator. Cells were harvested and genomic DNA prepared five days post-transfection. PCR primers appropriate for use in NGS-based DNA sequencing were generated, optimized, and used to amplify the individual target sequences for each guide RNA (SEQ ID NOs: 4569-4599). The amplicons were sequenced on an Illumina MiSeq machine and analyzed with a proprietary Python script to measure gene editing (**FIG. 60**).

Table 14C: Sequences of Guide RNAs and Sequences Targeted for Example 40

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting AAVS1	4538	MG29-1-AAVS1-sgRNA-A1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrUrUrUrUrCrCrArArArCrUrGrCrUrUrCrUrCrCrU/AltR2/
MG29-1 sgRNA targeting AAVS1	4539	MG29-1-AAVS1-sgRNA-B1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrCrCrArArArCrUrGrCrUrUrCrUrCrCrUrC/AltR2/
MG29-1 sgRNA targeting AAVS1	4540	MG29-1-AAVS1-sgRNA-C1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrArArArCrUrGrCrUrUrCrUrCrUrUrGrG/AltR2/
MG29-1 sgRNA targeting AAVS1	4541	MG29-1-AAVS1-sgRNA-D1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrArArArCrUrGrCrUrUrCrUrCrUrUrGrGrG/AltR2/
MG29-1 sgRNA targeting AAVS1	4542	MG29-1-AAVS1-sgRNA-E1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArArArCrUrGrCrUrUrCrUrCrUrUrGrGrGrA/AltR2/
MG29-1 sgRNA targeting AAVS1	4543	MG29-1-AAVS1-sgRNA-F1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrCrArCrCrArArUrCrCrUrGrUrCrCrCrUrArG/AltR2/
MG29-1 sgRNA targeting AAVS1	4544	MG29-1-AAVS1-sgRNA-G1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrUrCrArCrCrArArUrCrCrUrGrUrCrCrCrUrArGrU/AltR2/
MG29-1 sgRNA targeting AAVS1	4545	MG29-1-AAVS1-sgRNA-H1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrGrUrUrGrUrCrCrArGrArArArArArCrGrGrUrGrA/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting AAVS1	4546	MG29-1-AAVS1-sgRNA-A2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrUrUrCrUrCrCrUrUrCrUrGrGrGrGrCrCrUrGrUrG/AltR2/
MG29-1 sgRNA targeting AAVS1	4547	MG29-1-AAVS1-sgRNA-B2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrCrUrCrCrUrUrCrUrGrGrGrGrCrCrUrGrUrGrC/AltR2/
MG29-1 sgRNA targeting AAVS1	4548	MG29-1-AAVS1-sgRNA-C2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrArGrGrArUrGrGrCrCrUrUrCrUrCrCrGrArCrGrG/AltR2/
MG29-1 sgRNA targeting AAVS1	4549	MG29-1-AAVS1-sgRNA-D2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrGrGrArCrArArCrCrCrCrArArArGrUrArCrCrC/AltR2/
MG29-1 sgRNA targeting AAVS1	4550	MG29-1-AAVS1-sgRNA-E2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrGrArCrArArCrCrCrCrArArArGrUrArCrCrC/AltR2/
MG29-1 sgRNA targeting AAVS1	4551	MG29-1-AAVS1-sgRNA-F2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrGrArCrArArCrCrCrCrArArArGrUrArCrCrCrC/AltR2/
MG29-1 sgRNA targeting AAVS1	4552	MG29-1-AAVS1-sgRNA-G2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrCrArCrCrUrCrUrCrCrArUrCrCrUrCrUrUrGrCrU/AltR2/
MG29-1 sgRNA targeting AAVS1	4553	MG29-1-AAVS1-sgRNA-H2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrGrCrCrUrGrGrArCrArCrCrCrCrGrUrUrCrUrC/AltR2/
MG29-1 sgRNA targeting AAVS1	4554	MG29-1-AAVS1-sgRNA-A3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrUrGrGrArCrArCrCrCrCrGrUrUrCrUrCrCrUrGrU/AltR2/
MG29-1 sgRNA targeting AAVS1	4555	MG29-1-AAVS1-sgRNA-B3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrUrGrGrGrCrArGrCrUrCrCrCrCrUrArCrCrCrC/AltR2/
MG29-1 sgRNA targeting AAVS1	4556	MG29-1-AAVS1-sgRNA-C3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrCrArGrCrUrCrCrCrCrCrUrArCrCrCrCrCrCrUrUrA/AltR2/
MG29-1 sgRNA targeting AAVS1	4557	MG29-1-AAVS1-sgRNA-D3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrCrCrUrCrCrArGrGrGrArUrCrCrUrGrUrGrUrC/AltR2/
MG29-1 sgRNA targeting AAVS1	4558	MG29-1-AAVS1-sgRNA-E3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrCrUrGrUrCrCrCrCrUrCrCrArCrCrCrCrArCrCrArG/AltR2/
MG29-1 sgRNA targeting AAVS1	4559	MG29-1-AAVS1-sgRNA-F3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrGrUrCrCrCrCrUrCrCrArCrCrCrCrArCrCrArGrU/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting AAVS1	4560	MG29-1-AAVS1-sgRNA-G3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrGrArGrCrCrArUrCrUrCrUrCrUrCrCrUrUrGrC/AltR2/
MG29-1 sgRNA targeting AAVS1	4561	MG29-1-AAVS1-sgRNA-H3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrArCrGrArUrGrGrArGrCrCrArGrArGrArGrGrA/AltR2/
MG29-1 sgRNA targeting AAVS1	4562	MG29-1-AAVS1-sgRNA-A4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrUrGrArUrCrCrUrGrGrUrGrCrUrGrCrArGrCrUrU/AltR2/
MG29-1 sgRNA targeting AAVS1	4563	MG29-1-AAVS1-sgRNA-B4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArArArArArCrArArArArUrCrArGrArArUrArArGrU/AltR2/
MG29-1 sgRNA targeting AAVS1	4564	MG29-1-AAVS1-sgRNA-C4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrUrCrUrCrArCrCrUrUrUrCrUrArGrUrCrCrCrC/AltR2/
MG29-1 sgRNA targeting AAVS1	4565	MG29-1-AAVS1-sgRNA-D4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrUrCrCrCrArArUrUrUrArUrArUrUrGrUrUrC/AltR2/
MG29-1 sgRNA targeting AAVS1	4566	MG29-1-AAVS1-sgRNA-E4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArUrUrGrUrUrCrCrUrCrCrGrUrGrCrGrUrCrArGrU/AltR2/
MG29-1 sgRNA targeting AAVS1	4567	MG29-1-AAVS1-sgRNA-F4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrCrUrGrUrGrArGrArUrArArGrGrCrCrArGrUrArG/AltR2/
MG29-1 sgRNA targeting AAVS1	4568	MG29-1-AAVS1-sgRNA-G4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrUrGrUrGrArGrArUrArArGrGrCrCrArGrUrArGrC/AltR2/
DNA sequence of AAVS1 target site	4569	MG29-1-AAVS1-target site-A1	GTTTTTCCAAACTGCTTCTCCT
DNA sequence of AAVS1 target site	4570	MG29-1-AAVS1-target site-B1	TTTTTCCAAACTGCTTCTCCTC
DNA sequence of AAVS1 target site	4571	MG29-1-AAVS1-target site-C1	TCCAAACTGCTTCTCCTCTTGG
DNA sequence of AAVS1 target site	4572	MG29-1-AAVS1-target site-D1	CCAAACTGCTTCTCCTCTTGGG
DNA sequence of AAVS1 target site	4573	MG29-1-AAVS1-target site-E1	CAAACTGCTTCTCCTCTTGGGA
DNA sequence of AAVS1 target site	4574	MG29-1-AAVS1-target site-F1	CTGTCACCAATCCTGTCCCTAG
DNA sequence of AAVS1 target site	4575	MG29-1-AAVS1-target site-G1	TGTCACCAATCCTGTCCCTAGT

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of AAVS1 target site	4576	MG29-1-AAVS1-target site-H1	GGGTTGTCCAGAAAAACGGTGA
DNA sequence of AAVS1 target site	4577	MG29-1-AAVS1-target site-A2	CCTTCTCCTTCTGGGGCCTGTG
DNA sequence of AAVS1 target site	4578	MG29-1-AAVS1-target site-B2	CTTCTCCTTCTGGGGCCTGTGC
DNA sequence of AAVS1 target site	4579	MG29-1-AAVS1-target site-C2	TTAGGATGGCCTTCTCCGACGG
DNA sequence of AAVS1 target site	4580	MG29-1-AAVS1-target site-D2	TCTGGACAACCCCAAAGTACCC
DNA sequence of AAVS1 target site	4581	MG29-1-AAVS1-target site-E2	CTGGACAACCCCAAAGTACCCC
DNA sequence of AAVS1 target site	4582	MG29-1-AAVS1-target site-F2	TGGACAACCCCAAAGTACCCCG
DNA sequence of AAVS1 target site	4583	MG29-1-AAVS1-target site-G2	GCCACCTCTCCATCCTCTTGCT
DNA sequence of AAVS1 target site	4584	MG29-1-AAVS1-target site-H2	TTTGCCTGGACACCCCGTTCTC
DNA sequence of AAVS1 target site	4585	MG29-1-AAVS1-target site-A3	CCTGGACACCCCGTTCTCCTGT
DNA sequence of AAVS1 target site	4586	MG29-1-AAVS1-target site-B3	ATTTGGGCAGCTCCCCTACCCC
DNA sequence of AAVS1 target site	4587	MG29-1-AAVS1-target site-C3	GGCAGCTCCCCTACCCCCCTTA
DNA sequence of AAVS1 target site	4588	MG29-1-AAVS1-target site-D3	CTGCCTCCAGGGATCCTGTGTC
DNA sequence of AAVS1 target site	4589	MG29-1-AAVS1-target site-E3	ATCTGTCCCCTCCACCCCACAG
DNA sequence of AAVS1 target site	4590	MG29-1-AAVS1-target site-F3	TCTGTCCCCTCCACCCCACAGT
DNA sequence of AAVS1 target site	4591	MG29-1-AAVS1-target site-G3	CTGGAGCCATCTCTCTCCTTGC
DNA sequence of AAVS1 target site	4592	MG29-1-AAVS1-target site-H3	CTTACGATGGAGCCAGAGAGGA
DNA sequence of AAVS1 target site	4593	MG29-1-AAVS1-target site-A4	ACTGATCCTGGTGCTGCAGCTT
DNA sequence of AAVS1 target site	4594	MG29-1-AAVS1-target site-B4	GAAAAACAAAATCAGAATAAGT

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of AAVS1 target site	4595	MG29-1-AAVS1-target site-C4	GCTCTTCACCTTTCTAGTCCCC
DNA sequence of AAVS1 target site	4596	MG29-1-AAVS1-target site-D4	TAGTCCCCAATTTATATTGTTC
DNA sequence of AAVS1 target site	4597	MG29-1-AAVS1-target site-E4	TATTGTTCCTCCGTGCGTCAGT
DNA sequence of AAVS1 target site	4598	MG29-1-AAVS1-target site-F4	ACCTGTGAGATAAGGCCAGTAG
DNA sequence of AAVS1 target site	4599	MG29-1-AAVS1-target site-G4	CCTGTGAGATAAGGCCAGTAGC
r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *; between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur; AltR1 and AltR2 refer to IDT technologies' proprietary 5' and 3' AltR modifications			

Example 15 – Gene editing outcomes at the DNA level for B2M

[00492] Primary T cells were purified from PMBCs using a negative selection kit (Miltenyi) according to the manufacturer's recommendations. Nucleofection of MG29-1 RNPs (106 pmol protein/160 pmol guide) (SEQ ID NOs: 4600-4675) was performed into T cells (200,000) using the Lonza 4D electroporator. Cells were harvested and genomic DNA prepared five days post-transfection. PCR primers appropriate for use in NGS-based DNA sequencing were generated, optimized, and used to amplify the individual target sequences for each guide RNA (SEQ ID NOs: 4676-4751). The amplicons were sequenced on an Illumina MiSeq machine and analyzed with a proprietary Python script to measure gene editing (**FIG. 61**).

Table 14D: Sequences of Guide RNAs and Sequences Targeted for Example 41

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting B2M	4600	MG29-1-B2M-sgRNA-A1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrGrCrCrUrGrGrArGrGrCrUrArUrCrCrArGrCrGrU/AltR2/
MG29-1 sgRNA targeting B2M	4601	MG29-1-B2M-sgRNA-B1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrCrGrArUrArUrUrCrCrUrCrArGrGrUrArCrUrCrC/AltR2/
MG29-1 sgRNA targeting B2M	4602	MG29-1-B2M-sgRNA-C1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrGrArUrArUrUrCrCrUrCrArGrGrUrArCrUrCrCrA/AltR2/
MG29-1 sgRNA targeting B2M	4603	MG29-1-B2M-sgRNA-D1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrCrArUrCrCrArGrCrArGrArGrArArU/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting B2M	4604	MG29-1-B2M-sgRNA-E1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrArArUrUrGrCrUrArUrGrUrGrUrCrUrGrGrGrU/AltR2/
MG29-1 sgRNA targeting B2M	4605	MG29-1-B2M-sgRNA-F1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrCrCrArUrCrCrGrArCrArUrUrGrArArGrUrUrGrA/AltR2/
MG29-1 sgRNA targeting B2M	4606	MG29-1-B2M-sgRNA-G1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrCrArArGrGrArCrUrGrUrCrUrUrUrCrUrArUrC/AltR2/
MG29-1 sgRNA targeting B2M	4607	MG29-1-B2M-sgRNA-H1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArUrCrUrCrUrUrGrUrArCrUrArCrArCrUrGrArArU/AltR2/
MG29-1 sgRNA targeting B2M	4608	MG29-1-B2M-sgRNA-A2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrArCrArGrCrCrCrArArGrArUrArGrUrUrArArGrU/AltR2/
MG29-1 sgRNA targeting B2M	4609	MG29-1-B2M-sgRNA-B2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrArGrUrGrGrGrGrGrUrGrArArUrUrCrArGrUrGrU/AltR2/
MG29-1 sgRNA targeting B2M	4610	MG29-1-B2M-sgRNA-C2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrUrGrGrGrGrGrUrGrArArUrUrCrArGrUrGrUrA/AltR2/
MG29-1 sgRNA targeting B2M	4611	MG29-1-B2M-sgRNA-D2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrUrGrGrGrGrGrUrGrArArUrUrCrArGrUrGrUrArG/AltR2/
MG29-1 sgRNA targeting B2M	4612	MG29-1-B2M-sgRNA-E2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrArArUrUrCrUrCrUrCrUrCrCrArUrUrCrUrUrCrA/AltR2/
MG29-1 sgRNA targeting B2M	4613	MG29-1-B2M-sgRNA-F2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArArUrUrCrUrCrUrCrCrArUrUrCrUrUrCrArG/AltR2/
MG29-1 sgRNA targeting B2M	4614	MG29-1-B2M-sgRNA-G2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArUrUrCrUrCrUrCrUrCrCrArUrUrCrUrUrCrArGrU/AltR2/
MG29-1 sgRNA targeting B2M	4615	MG29-1-B2M-sgRNA-H2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrUrUrUrCrCrArUrUrCrUrCrUrGrCrUrGrGrArUrG/AltR2/
MG29-1 sgRNA targeting B2M	4616	MG29-1-B2M-sgRNA-A3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArUrUrCrUrCrUrCrUrGrCrUrGrGrArArUrGrGrUrGrA/AltR2/
MG29-1 sgRNA targeting B2M	4617	MG29-1-B2M-sgRNA-B3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrCrCrArCrUrGrUrCrUrUrUrUrUrCrArUrArGrGrA/AltR2/
MG29-1 sgRNA targeting B2M	4618	MG29-1-B2M-sgRNA-C3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrCrCrArCrUrGrUrCrUrUrUrUrUrCrArUrArGrArU/AltR2/
MG29-1 sgRNA targeting B2M	4619	MG29-1-B2M-sgRNA-D3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrArCrUrGrUrCrUrUrUrUrUrCrArUrArGrArUrC/AltR2/
MG29-1 sgRNA targeting B2M	4620	MG29-1-B2M-sgRNA-E3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrArUrArGrArArCrArUrGrUrArArArGrC/AltR2/
MG29-1 sgRNA targeting B2M	4621	MG29-1-B2M-sgRNA-F3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArUrArGrArArCrArUrGrUrArArArGrCrA/AltR2/
MG29-1 sgRNA targeting B2M	4622	MG29-1-B2M-sgRNA-G3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrGrCrCrUrGrGrArGrGrCrUrArUrCrCrArGrCrGrU/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting B2M	4623	MG29-1-B2M-sgRNA-H3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrCrGrGrGrArGrCrArGrGrGrGrArGrArCrCrUrU/AltR2/
MG29-1 sgRNA targeting B2M	4624	MG29-1-B2M-sgRNA-A4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrCrUrArCrGrGrCrGrArCrGrGrGrArGrGrGrUrCrG/AltR2/
MG29-1 sgRNA targeting B2M	4625	MG29-1-B2M-sgRNA-B4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrGrCrGrUrCrGrArUrArArGrCrGrUrCrArGrArGrC/AltR2/
MG29-1 sgRNA targeting B2M	4626	MG29-1-B2M-sgRNA-C4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrUrCrCrGrCrUrCrUrUrUrCrGrCrGrGrGrGrCrC/AltR2/
MG29-1 sgRNA targeting B2M	4627	MG29-1-B2M-sgRNA-D4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrGrGrGrCrCrUrCrUrGrGrCrUrCrCrCrCrCrArG/AltR2/
MG29-1 sgRNA targeting B2M	4628	MG29-1-B2M-sgRNA-E4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrArArCrGrCrGrUrGrGrArGrGrGrGrCrGrCrUrUrG/AltR2/
MG29-1 sgRNA targeting B2M	4629	MG29-1-B2M-sgRNA-F4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrCrCrCrArCrGrGrUrGrUrGrGrCrCrCrCrArCrA/AltR2/
MG29-1 sgRNA targeting B2M	4630	MG29-1-B2M-sgRNA-G4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrArArGrCrGrArGrCrCrGrUrCrCrCrCrCrCrA/AltR2/
MG29-1 sgRNA targeting B2M	4631	MG29-1-B2M-sgRNA-H4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrCrGrArCrCrCrUrCrCrCrGrUrCrGrCrCrGrUrA/AltR2/
MG29-1 sgRNA targeting B2M	4632	MG29-1-B2M-sgRNA-A5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrGrGrGrArGrCrGrCrArUrGrCrCrUrUrUrUrGrGrC/AltR2/
MG29-1 sgRNA targeting B2M	4633	MG29-1-B2M-sgRNA-B5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrGrGrGrArGrCrGrCrArUrGrCrCrUrUrUrUrGrGrCrU/AltR2/
MG29-1 sgRNA targeting B2M	4634	MG29-1-B2M-sgRNA-C5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrCrUrGrUrArArUrUrCrGrUrGrCrArUrUrUrUrUrU/AltR2/
MG29-1 sgRNA targeting B2M	4635	MG29-1-B2M-sgRNA-D5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrUrGrUrArArUrUrCrGrUrGrCrArUrUrUrUrUrU/AltR2/
MG29-1 sgRNA targeting B2M	4636	MG29-1-B2M-sgRNA-E5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrArArGrArArArArArCrGrCrCrUrGrCrCrU/AltR2/
MG29-1 sgRNA targeting B2M	4637	MG29-1-B2M-sgRNA-F5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrArArGrArArArArArCrGrCrCrUrGrCrCrUrU/AltR2/
MG29-1 sgRNA targeting B2M	4638	MG29-1-B2M-sgRNA-G5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrArArGrArArArArArCrGrCrCrUrGrCrCrUrU/AltR2/
MG29-1 sgRNA targeting B2M	4639	MG29-1-B2M-sgRNA-H5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrArArGrArArArArArCrGrCrCrUrGrCrCrUrU/AltR2/
MG29-1 sgRNA targeting B2M	4640	MG29-1-B2M-sgRNA-A6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrArArGrArArArArArCrGrCrCrUrGrCrCrUrU/AltR2/
MG29-1 sgRNA targeting B2M	4641	MG29-1-B2M-sgRNA-B6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArGrArArArArArCrGrCrCrUrGrCrCrUrU/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting B2M	4642	MG29-1-B2M-sgRNA-C6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrArArArArCrGrCrCrUrGrCrCrUrUrCrUrGrCrG/AltR2/
MG29-1 sgRNA targeting B2M	4643	MG29-1-B2M-sgRNA-D6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrGrUrArCrArGrArGrGrGrCrUrUrCrCrUrCrUrUrUrG/AltR2/
MG29-1 sgRNA targeting B2M	4644	MG29-1-B2M-sgRNA-E6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrArCrArGrArArGrGrGrCrUrUrCrCrUrCrUrUrGrG/AltR2/
MG29-1 sgRNA targeting B2M	4645	MG29-1-B2M-sgRNA-F6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrUrCrUrUrUrGrCrCrUrGrGrUrUrGrUrUrUrCrCrA/AltR2/
MG29-1 sgRNA targeting B2M	4646	MG29-1-B2M-sgRNA-G6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrUrGrGrUrUrUrCrCrArArGrArUrGrUrArArC/AltR2/
MG29-1 sgRNA targeting B2M	4647	MG29-1-B2M-sgRNA-H6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArArGrArUrGrUrArCrUrGrUrGrCrCrUrCrUrUrArC/AltR2/
MG29-1 sgRNA targeting B2M	4648	MG29-1-B2M-sgRNA-A7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArArArArCrCrGrArArArGrUrArArArGrArGrGrCrArC/AltR2/
MG29-1 sgRNA targeting B2M	4649	MG29-1-B2M-sgRNA-B7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArArArCrCrGrArArArGrUrArArArGrArGrGrCrArCrA/AltR2/
MG29-1 sgRNA targeting B2M	4650	MG29-1-B2M-sgRNA-C7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrCrUrGrGrArGrArArUrCrUrCrArCrGrCrCrArGrArArA/AltR2/
MG29-1 sgRNA targeting B2M	4651	MG29-1-B2M-sgRNA-D7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrUrArArArArArArArArArArUrGrCrArCrGrArArU/AltR2/
MG29-1 sgRNA targeting B2M	4652	MG29-1-B2M-sgRNA-E7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrArArArArArArArArArArUrGrCrArCrGrArArUrU/AltR2/
MG29-1 sgRNA targeting B2M	4653	MG29-1-B2M-sgRNA-F7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrArArArArArArArArArArUrGrCrArCrGrArArUrUrA/AltR2/
MG29-1 sgRNA targeting B2M	4654	MG29-1-B2M-sgRNA-G7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrUrArArArArArUrGrGrArGrGrUrGrGrCrUrU/AltR2/
MG29-1 sgRNA targeting B2M	4655	MG29-1-B2M-sgRNA-H7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrCrArGrArGrUrGrGrGrArArArUrUrUrGrG/AltR2/
MG29-1 sgRNA targeting B2M	4656	MG29-1-B2M-sgRNA-A8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArGrUrArCrCrUrGrArGrGrArArUrArUrArUrCrGrGrGrA/AltR2/
MG29-1 sgRNA targeting B2M	4657	MG29-1-B2M-sgRNA-B8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrUrUrUrGrGrGrGrCrUrArArCrUrUrGrGrUrGrUrCrA/AltR2/
MG29-1 sgRNA targeting B2M	4658	MG29-1-B2M-sgRNA-C8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArUrUrCrGrArUrUrUrCrUrCrCrC/AltR2/
MG29-1 sgRNA targeting B2M	4659	MG29-1-B2M-sgRNA-D8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrCrArUrCrGrArUrUrUrCrCrCrArArUrUrCrC/AltR2/
MG29-1 sgRNA targeting B2M	4660	MG29-1-B2M-sgRNA-E8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrCrArArUrUrUrCrCrArCrUrCrUrGrG/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting B2M	4661	MG29-1-B2M-sgRNA-F8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArCrUrCrUrGrGrCrCrArArArUrGrArGrCrUrUrCrC/AltR2/
MG29-1 sgRNA targeting B2M	4662	MG29-1-B2M-sgRNA-G8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArArGrArArUrArArArCrCrGrUrGrArCrUrUrGrGrU/AltR2/
MG29-1 sgRNA targeting B2M	4663	MG29-1-B2M-sgRNA-H8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArGrArArUrArArArCrCrGrUrGrArCrUrUrGrGrUrA/AltR2/
MG29-1 sgRNA targeting B2M	4664	MG29-1-B2M-sgRNA-A9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrUrCrArUrArArUrUrCrCrUrCrUrArUrArUrCrArUrG/AltR2/
MG29-1 sgRNA targeting B2M	4665	MG29-1-B2M-sgRNA-B9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrUrUrUrUrUrUrUrCrUrArGrCrArGrArUrUrU/AltR2/
MG29-1 sgRNA targeting B2M	4666	MG29-1-B2M-sgRNA-C9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrUrUrUrUrUrUrUrCrUrArGrCrArGrArUrUrC/AltR2/
MG29-1 sgRNA targeting B2M	4667	MG29-1-B2M-sgRNA-D9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrUrUrUrUrUrUrUrCrUrArGrCrArGrArUrUrUrCrU/AltR2/
MG29-1 sgRNA targeting B2M	4668	MG29-1-B2M-sgRNA-E9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrUrUrUrCrUrArGrCrArGrArUrUrUrCrUrA/AltR2/
MG29-1 sgRNA targeting B2M	4669	MG29-1-B2M-sgRNA-F9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrCrUrArGrCrArGrArUrUrUrCrUrArGrCrArG/AltR2/
MG29-1 sgRNA targeting B2M	4670	MG29-1-B2M-sgRNA-G9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrCrUrArGrCrArGrArUrUrUrCrUrArGrCrArGrU/AltR2/
MG29-1 sgRNA targeting B2M	4671	MG29-1-B2M-sgRNA-H9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrUrArGrCrArGrArUrUrUrCrUrArGrCrArGrUrA/AltR2/
MG29-1 sgRNA targeting B2M	4672	MG29-1-B2M-sgRNA-A10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrArGrCrArGrArUrUrUrCrUrArGrCrArGrUrArU/AltR2/
MG29-1 sgRNA targeting B2M	4673	MG29-1-B2M-sgRNA-B10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrArGrCrArGrArUrUrUrCrUrArGrCrArGrUrArUrC/AltR2/
MG29-1 sgRNA targeting B2M	4674	MG29-1-B2M-sgRNA-C10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrCrArGrArUrUrUrCrUrArGrCrArGrUrArUrCrU/AltR2/
MG29-1 sgRNA targeting B2M	4675	MG29-1-B2M-sgRNA-D10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrCrArGrUrArUrCrUrUrCrUrGrUrCrArCrUrGrG/AltR2/
DNA sequence of B2M target site	4676	MG29-1-B2M-target site-A1	TGGCCTGGAGGCTATCCAGCGT
DNA sequence of B2M target site	4677	MG29-1-B2M-target site-B1	CCCGATATTCCTCAGGTACTCC
DNA sequence of B2M target site	4678	MG29-1-B2M-target site-C1	CCGATATTCCTCAGGTACTCCA
DNA sequence of B2M target site	4679	MG29-1-B2M-target site-D1	CTCACGTCATCCAGCAGAGAAT

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of B2M target site	4680	MG29-1-B2M-target site-E1	CTGAATTGCTATGTGTCTGGGT
DNA sequence of B2M target site	4681	MG29-1-B2M-target site-F1	ATCCATCCGACATTGAAGTTGA
DNA sequence of B2M target site	4682	MG29-1-B2M-target site-G1	AGCAAGGACTGGTCTTTCTATC
DNA sequence of B2M target site	4683	MG29-1-B2M-target site-H1	TATCTCTTGTACTACACTGAAT
DNA sequence of B2M target site	4684	MG29-1-B2M-target site-A2	TCACAGCCCAAGATAGTTAAGT
DNA sequence of B2M target site	4685	MG29-1-B2M-target site-B2	TCAGTGGGGGTGAATTCAGTGT
DNA sequence of B2M target site	4686	MG29-1-B2M-target site-C2	CAGTGGGGGTGAATTCAGTGTA
DNA sequence of B2M target site	4687	MG29-1-B2M-target site-D2	AGTGGGGGTGAATTCAGTGTAG
DNA sequence of B2M target site	4688	MG29-1-B2M-target site-E2	TCAATTCTCTCTCCATTCTTCA
DNA sequence of B2M target site	4689	MG29-1-B2M-target site-F2	CAATTCTCTCTCCATTCTTCAG
DNA sequence of B2M target site	4690	MG29-1-B2M-target site-G2	AATTCTCTCTCCATTCTTCAGT
DNA sequence of B2M target site	4691	MG29-1-B2M-target site-H2	ACTTTCCATTCTCTGCTGGATG
DNA sequence of B2M target site	4692	MG29-1-B2M-target site-A3	CATTCTCTGCTGGATGACGTGA
DNA sequence of B2M target site	4693	MG29-1-B2M-target site-B3	TCTCCACTGTCTTTTTCATAGA
DNA sequence of B2M target site	4694	MG29-1-B2M-target site-C3	CTCCACTGTCTTTTTCATAGAT
DNA sequence of B2M target site	4695	MG29-1-B2M-target site-D3	TCCACTGTCTTTTTCATAGATC
DNA sequence of B2M target site	4696	MG29-1-B2M-target site-E3	TCATAGATCGAGACATGTAAGC
DNA sequence of B2M target site	4697	MG29-1-B2M-target site-F3	CATAGATCGAGACATGTAAGCA
DNA sequence of B2M target site	4698	MG29-1-B2M-target site-G3	TGGCCTGGAGGCTATCCAGCGT

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of B2M target site	4699	MG29-1-B2M-target site-H3	GGCGGGGAGCAGGGGAGACCTT
DNA sequence of B2M target site	4700	MG29-1-B2M-target site-A4	GCCTACGGCGACGGGAGGGTCG
DNA sequence of B2M target site	4701	MG29-1-B2M-target site-B4	GGGCGTCGATAAGCGTCAGAGC
DNA sequence of B2M target site	4702	MG29-1-B2M-target site-C4	TCTTCCGCTCTTTCGCGGGGCC
DNA sequence of B2M target site	4703	MG29-1-B2M-target site-D4	GCGGGGCCTCTGGCTCCCCCAG
DNA sequence of B2M target site	4704	MG29-1-B2M-target site-E4	TGAACGCGTGGAGGGGCGCTTG
DNA sequence of B2M target site	4705	MG29-1-B2M-target site-F4	CTCCCCACGGTGTGGCCCCACA
DNA sequence of B2M target site	4706	MG29-1-B2M-target site-G4	GGACGAGCCTACCCGTCCCCCA
DNA sequence of B2M target site	4707	MG29-1-B2M-target site-H4	TCCCGACCCTCCCGTCGCCGTA
DNA sequence of B2M target site	4708	MG29-1-B2M-target site-A5	GCGGGAGCGCATGCCTTTTGGC
DNA sequence of B2M target site	4709	MG29-1-B2M-target site-B5	CGGGAGCGCATGCCTTTTGGCT
DNA sequence of B2M target site	4710	MG29-1-B2M-target site-C5	GGCTGTAATTCGTGCATTTTTT
DNA sequence of B2M target site	4711	MG29-1-B2M-target site-D5	GCTGTAATTCGTGCATTTTTTT
DNA sequence of B2M target site	4712	MG29-1-B2M-target site-E5	TTTTTAAGAAAAACGCCTGCCT
DNA sequence of B2M target site	4713	MG29-1-B2M-target site-F5	TTTTAAGAAAAACGCCTGCCTT
DNA sequence of B2M target site	4714	MG29-1-B2M-target site-G5	TTTAAGAAAAACGCCTGCCTTC
DNA sequence of B2M target site	4715	MG29-1-B2M-target site-H5	TTAAGAAAAACGCCTGCCTTCT
DNA sequence of B2M target site	4716	MG29-1-B2M-target site-A6	TAAGAAAAACGCCTGCCTTCTG
DNA sequence of B2M target site	4717	MG29-1-B2M-target site-B6	AAGAAAAACGCCTGCCTTCTGC

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of B2M target site	4718	MG29-1-B2M-target site-C6	AGAAAAACGCCTGCCTTCTGCG
DNA sequence of B2M target site	4719	MG29-1-B2M-target site-D6	CGTACAGAGGGCTTCCTCTTTG
DNA sequence of B2M target site	4720	MG29-1-B2M-target site-E6	GTACAGAGGGCTTCCTCTTTGG
DNA sequence of B2M target site	4721	MG29-1-B2M-target site-F6	GCTCTTTGCCTGGTTGTTTCCA
DNA sequence of B2M target site	4722	MG29-1-B2M-target site-G6	CCTGGTTGTTTCCAAGATGTAC
DNA sequence of B2M target site	4723	MG29-1-B2M-target site-H6	CAAGATGTACTGTGCCTCTTAC
DNA sequence of B2M target site	4724	MG29-1-B2M-target site-A7	CAAAACCGAAAGTAAGAGGCAC
DNA sequence of B2M target site	4725	MG29-1-B2M-target site-B7	AAAACCGAAAGTAAGAGGCACA
DNA sequence of B2M target site	4726	MG29-1-B2M-target site-C7	CTCTGGAGAATCTCACGCAGAA
DNA sequence of B2M target site	4727	MG29-1-B2M-target site-D7	TCTTAAAAAAAAATGCACGAAT
DNA sequence of B2M target site	4728	MG29-1-B2M-target site-E7	CTTAAAAAAAAATGCACGAATT
DNA sequence of B2M target site	4729	MG29-1-B2M-target site-F7	TTAAAAAAAAATGCACGAATTA
DNA sequence of B2M target site	4730	MG29-1-B2M-target site-G7	TTCTTCAAAATGGAGGTGGCTT
DNA sequence of B2M target site	4731	MG29-1-B2M-target site-H7	GCCAGAGTGGAAATGGAATTGG
DNA sequence of B2M target site	4732	MG29-1-B2M-target site-A8	GAGTACCTGAGGAATATCGGGA
DNA sequence of B2M target site	4733	MG29-1-B2M-target site-B8	ACTTGGGGCTAACTTGGTGTCA
DNA sequence of B2M target site	4734	MG29-1-B2M-target site-C8	CATTTGGTCATCGATTTCTCCC
DNA sequence of B2M target site	4735	MG29-1-B2M-target site-D8	GTCATCGATTTCTCCCAATTCC
DNA sequence of B2M target site	4736	MG29-1-B2M-target site-E8	TCCCAATTCCATTTCCACTCTG

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of B2M target site	4737	MG29-1-B2M-target site-F8	CACTCTGGCCAAATGAGCTTCC
DNA sequence of B2M target site	4738	MG29-1-B2M-target site-G8	GAAGAATAAACCGTGACTTGGT
DNA sequence of B2M target site	4739	MG29-1-B2M-target site-H8	AAGAATAAACCGTGACTTGGTA
DNA sequence of B2M target site	4740	MG29-1-B2M-target site-A9	CCTCATAATTCCTCTATACATG
DNA sequence of B2M target site	4741	MG29-1-B2M-target site-B9	TTGTTTTTTTTCTAGCAGATTT
DNA sequence of B2M target site	4742	MG29-1-B2M-target site-C9	TGTTTTTTTTCTAGCAGATTTC
DNA sequence of B2M target site	4743	MG29-1-B2M-target site-D9	GTTTTTTTTCTAGCAGATTTCT
DNA sequence of B2M target site	4744	MG29-1-B2M-target site-E9	TTTTTTTTCTAGCAGATTTCTA
DNA sequence of B2M target site	4745	MG29-1-B2M-target site-F9	TTTTCTAGCAGATTTCTAGCAG
DNA sequence of B2M target site	4746	MG29-1-B2M-target site-G9	TTTCTAGCAGATTTCTAGCAGT
DNA sequence of B2M target site	4747	MG29-1-B2M-target site-H9	TTCTAGCAGATTTCTAGCAGTA
DNA sequence of B2M target site	4748	MG29-1-B2M-target site-A10	TCTAGCAGATTTCTAGCAGTAT
DNA sequence of B2M target site	4749	MG29-1-B2M-target site-B10	CTAGCAGATTTCTAGCAGTATC
DNA sequence of B2M target site	4750	MG29-1-B2M-target site-C10	TAGCAGATTTCTAGCAGTATCT
DNA sequence of B2M target site	4751	MG29-1-B2M-target site-D10	TAGCAGTATCTTCTGTCACTGG
r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *; between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur; AltR1 and AltR2 refer to IDT technologies' proprietary 5' and 3' AltR modifications			

Example 16 – Gene editing outcomes at the DNA level for CD2

[00493] Primary T cells were purified from PMBCs using a negative selection kit (Miltenyi) according to the manufacturer's recommendations. Nucleofection of MG29-1 RNPs (126 pmol

protein/160 pmol guide) (SEQ ID NOs: 4752-4836) was performed into T cells (200,000) using the Lonza 4D electroporator. Cells were harvested and genomic DNA prepared five days post-transfection. PCR primers appropriate for use in NGS-based DNA sequencing were generated, optimized, and used to amplify the individual target sequences for each guide RNA (SEQ ID NOs: 4837-4921). The amplicons were sequenced on an Illumina MiSeq machine and analyzed with a proprietary Python script to measure gene editing (**FIG. 62**).

Table 14E: Sequences of Guide RNAs and Sequences Targeted for Example 42

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting CD2	4752	MG29-1-CD2-sgRNA-A1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArUrGrUrArArArUrUrUrGrUrArGrCrCrArGrCrUrU/AltR2/
MG29-1 sgRNA targeting CD2	4753	MG29-1-CD2-sgRNA-B1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrCrCrArGrCrUrUrCrCrUrUrCrUrGrArUrUrUrU/AltR2/
MG29-1 sgRNA targeting CD2	4754	MG29-1-CD2-sgRNA-C1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrUrCrUrArUrGrCrUrUrArCrCrUrUrUrGrGrArA/AltR2/
MG29-1 sgRNA targeting CD2	4755	MG29-1-CD2-sgRNA-D1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArArGrArArArCrArUrUrGrArArArArUrCrArGrArA/AltR2/
MG29-1 sgRNA targeting CD2	4756	MG29-1-CD2-sgRNA-E1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrUrUrUrUrArUrArGrGrUrGrCrArGrUrCrUrCrC/AltR2/
MG29-1 sgRNA targeting CD2	4757	MG29-1-CD2-sgRNA-F1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrUrUrArUrArGrGrUrGrCrArGrUrCrUrCrCrA/AltR2/
MG29-1 sgRNA targeting CD2	4758	MG29-1-CD2-sgRNA-G1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArUrArGrGrUrGrCrArGrUrCrUrCrCrArArArGrArG/AltR2/
MG29-1 sgRNA targeting CD2	4759	MG29-1-CD2-sgRNA-H1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrArGrGrUrGrCrArGrUrCrUrCrCrArArArGrArGrA/AltR2/
MG29-1 sgRNA targeting CD2	4760	MG29-1-CD2-sgRNA-A2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrGrUrGrCrArGrUrCrUrCrCrArArArGrArGrArU/AltR2/
MG29-1 sgRNA targeting CD2	4761	MG29-1-CD2-sgRNA-B2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArArArUrGrArGrUrGrArUrArUrUrGrArCrG/AltR2/
MG29-1 sgRNA targeting CD2	4762	MG29-1-CD2-sgRNA-C2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArArUrGrArGrUrGrArUrArUrUrGrArCrGrA/AltR2/
MG29-1 sgRNA targeting CD2	4763	MG29-1-CD2-sgRNA-D2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArGrGrArArArArGrArUrArCrArUrArUrArArGrC/AltR2/
MG29-1 sgRNA targeting CD2	4764	MG29-1-CD2-sgRNA-E2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArArArUrGrGrArArCrUrCrUrGrArArArArUrUrArA/AltR2/
MG29-1 sgRNA targeting CD2	4765	MG29-1-CD2-sgRNA-F2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrCrArArCrArCrArUrUrUrUrUrUrCrCrUrUrUrU/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting CD2	4766	MG29-1-CD2-sgRNA-G2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrArArCrArCrArUrUrUrUrUrUrCrCrUrUrUrUrG/AltR2/
MG29-1 sgRNA targeting CD2	4767	MG29-1-CD2-sgRNA-H2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrArArCrArCrArUrUrUrUrUrUrCrCrUrUrUrUrGrU/AltR2/
MG29-1 sgRNA targeting CD2	4768	MG29-1-CD2-sgRNA-A3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArArCrArCrArUrUrUrUrUrUrCrCrUrUrUrUrGrUrA/AltR2/
MG29-1 sgRNA targeting CD2	4769	MG29-1-CD2-sgRNA-B3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrCrUrUrUrUrGrUrArUrCrArUrArUrArUrUrGrA/AltR2/
MG29-1 sgRNA targeting CD2	4770	MG29-1-CD2-sgRNA-C3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrUrUrUrUrGrUrArUrCrArUrArUrArUrUrGrArU/AltR2/
MG29-1 sgRNA targeting CD2	4771	MG29-1-CD2-sgRNA-D3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrUrUrUrUrGrUrArUrCrArUrArUrArUrUrGrArUrA/AltR2/
MG29-1 sgRNA targeting CD2	4772	MG29-1-CD2-sgRNA-E3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrUrUrGrUrArUrCrArUrArUrArUrUrGrArUrArC/AltR2/
MG29-1 sgRNA targeting CD2	4773	MG29-1-CD2-sgRNA-F3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrArUrCrArUrArUrArUrUrUrGrArUrArCrCrUrUrGrU/AltR2/
MG29-1 sgRNA targeting CD2	4774	MG29-1-CD2-sgRNA-G3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArUrCrArUrArUrArUrUrUrGrArUrArCrCrUrUrGrUrA/AltR2/
MG29-1 sgRNA targeting CD2	4775	MG29-1-CD2-sgRNA-H3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrArGrUrUrCrCrArUrUrUrUrUrUrArArArUrArGrC/AltR2/
MG29-1 sgRNA targeting CD2	4776	MG29-1-CD2-sgRNA-A4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrArGrUrUrCrCrArUrUrUrUrUrUrArArArUrArGrCrU/AltR2/
MG29-1 sgRNA targeting CD2	4777	MG29-1-CD2-sgRNA-B4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArArArUrArGrCrUrArUrUrUrArUrArUrGrUrArUrCrUrUrU/AltR2/
MG29-1 sgRNA targeting CD2	4778	MG29-1-CD2-sgRNA-C4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArArUrArGrCrUrArUrUrUrArUrArUrGrUrArUrCrUrUrU/AltR2/
MG29-1 sgRNA targeting CD2	4779	MG29-1-CD2-sgRNA-D4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArArUrArGrCrUrArUrUrUrArUrArUrGrUrArUrCrUrUrU/AltR2/
MG29-1 sgRNA targeting CD2	4780	MG29-1-CD2-sgRNA-E4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrUrUrGrArArArGrUrCrUrCrUrUrUrCrUrCrUrU/AltR2/
MG29-1 sgRNA targeting CD2	4781	MG29-1-CD2-sgRNA-F4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrUrUrGrArArArGrUrCrUrCrUrUrUrCrUrCrUrU/AltR2/
MG29-1 sgRNA targeting CD2	4782	MG29-1-CD2-sgRNA-G4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrGrArArArGrUrCrUrCrUrUrUrCrUrCrUrU/AltR2/
MG29-1 sgRNA targeting CD2	4783	MG29-1-CD2-sgRNA-H4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrUrUrCrUrGrArArUrUrGrUrGrCrArArUrCrU/AltR2/
MG29-1 sgRNA targeting CD2	4784	MG29-1-CD2-sgRNA-A5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrArArUrUrGrUrGrCrArArUrCrUrUrUrUrUrCrU/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting CD2	4785	MG29-1-CD2-sgRNA-B5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrArArUrUrGrUrGrCrArArUrCrUrUrUrUrUrCrUrU/AltR2/
MG29-1 sgRNA targeting CD2	4786	MG29-1-CD2-sgRNA-C5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrUrGrUrCrUrGrArArGrUrUrUrUrUrCrCrCrA/AltR2/
MG29-1 sgRNA targeting CD2	4787	MG29-1-CD2-sgRNA-D5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrGrUrCrUrGrArArGrUrUrUrUrUrCrCrCrArU/AltR2/
MG29-1 sgRNA targeting CD2	4788	MG29-1-CD2-sgRNA-E5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrUrCrUrGrArArGrUrUrUrUrUrCrCrCrArUrU/AltR2/
MG29-1 sgRNA targeting CD2	4789	MG29-1-CD2-sgRNA-F5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrCrCrArUrUrUrUrArArUrArUrCrGrUrCrArArUrA/AltR2/
MG29-1 sgRNA targeting CD2	4790	MG29-1-CD2-sgRNA-G5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrCrArUrUrUrUrArArUrArUrCrGrUrCrArArUrArU/AltR2/
MG29-1 sgRNA targeting CD2	4791	MG29-1-CD2-sgRNA-H5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrCrArUrUrUrUrArUrArUrCrGrUrCrArArUrArUrC/AltR2/
MG29-1 sgRNA targeting CD2	4792	MG29-1-CD2-sgRNA-A6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrArUrUrUrUrArUrArUrCrGrUrCrArArUrArUrCrA/AltR2/
MG29-1 sgRNA targeting CD2	4793	MG29-1-CD2-sgRNA-B6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrArUrCrGrUrCrArArUrArUrCrArArUrCrArCrUrCrA/AltR2/
MG29-1 sgRNA targeting CD2	4794	MG29-1-CD2-sgRNA-C6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArUrCrGrUrCrArArUrArUrCrArArUrCrArCrUrCrArU/AltR2/
MG29-1 sgRNA targeting CD2	4795	MG29-1-CD2-sgRNA-D6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArArArCrUrArGrGrArArUrGrUrCrCrArArGrUrUrG/AltR2/
MG29-1 sgRNA targeting CD2	4796	MG29-1-CD2-sgRNA-E6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArArGrGrCrArUrUrCrGrUrArArUrCrUrCrUrUrUrG/AltR2/
MG29-1 sgRNA targeting CD2	4797	MG29-1-CD2-sgRNA-F6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrUrCrUrUrUrUrArGrArGrArGrGrGrUrCrUrC/AltR2/
MG29-1 sgRNA targeting CD2	4798	MG29-1-CD2-sgRNA-G6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrArGrArGrArGrGrGrUrCrUrCrArArArArC/AltR2/
MG29-1 sgRNA targeting CD2	4799	MG29-1-CD2-sgRNA-H6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrArGrArGrGrGrUrCrUrCrArArArArCrCrArArA/AltR2/
MG29-1 sgRNA targeting CD2	4800	MG29-1-CD2-sgRNA-A7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrArGrArGrGrGrUrCrUrCrArArArArCrCrArArArG/AltR2/
MG29-1 sgRNA targeting CD2	4801	MG29-1-CD2-sgRNA-B7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArGrArGrArGrGrUrCrUrCrArArArCrCrArArArArGrA/AltR2/
MG29-1 sgRNA targeting CD2	4802	MG29-1-CD2-sgRNA-C7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrArGrArGrGrGrUrCrArUrCrArCrArCrArCrArArG/AltR2/
MG29-1 sgRNA targeting CD2	4803	MG29-1-CD2-sgRNA-D7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrCrCrUrGrCrUrGrUrGrCrArCrUrUrGrArArUrU/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting CD2	4804	MG29-1-CD2-sgRNA-E7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrArCrUrCrArGrGrCrUrGrGrUrGrUrCrCrArCrU/AltR2/
MG29-1 sgRNA targeting CD2	4805	MG29-1-CD2-sgRNA-F7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArCrUrCrArGrGrCrUrGrGrUrGrUrCrCrArCrUrU/AltR2/
MG29-1 sgRNA targeting CD2	4806	MG29-1-CD2-sgRNA-G7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrArUrGrUrUrUrCrCrCrArUrCrUrUrGrArUrArCrA/AltR2/
MG29-1 sgRNA targeting CD2	4807	MG29-1-CD2-sgRNA-H7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrUrUrCrCrCrArUrCrUrUrGrArUrArCrArG/AltR2/
MG29-1 sgRNA targeting CD2	4808	MG29-1-CD2-sgRNA-A8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrArUrCrUrUrGrArUrArCrArGrGrUrUrUrArArUrU/AltR2/
MG29-1 sgRNA targeting CD2	4809	MG29-1-CD2-sgRNA-B8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrUrCrGrGrGrGrUrCrArGrUrUrCrCrArUrUrCrArU/AltR2/
MG29-1 sgRNA targeting CD2	4810	MG29-1-CD2-sgRNA-C8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrArGrArArArArGrGrUrCrUrGrGrArCrArUrCrU/AltR2/
MG29-1 sgRNA targeting CD2	4811	MG29-1-CD2-sgRNA-D8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrArGrArArArGrGrUrCrUrGrGrArCrArUrCrUrA/AltR2/
MG29-1 sgRNA targeting CD2	4812	MG29-1-CD2-sgRNA-E8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrGrCrArCrUrGrCrUrCrUrUrCrUrArUrArUrA/AltR2/
MG29-1 sgRNA targeting CD2	4813	MG29-1-CD2-sgRNA-F8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrArUrArUrCrArCrCrArArArArGrGrArArArArArA/AltR2/
MG29-1 sgRNA targeting CD2	4814	MG29-1-CD2-sgRNA-G8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArUrArUrCrArCrCrArArArArGrGrArArArArArA/AltR2/
MG29-1 sgRNA targeting CD2	4815	MG29-1-CD2-sgRNA-H8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrGrArCrUrCrCrUrCrUrGrUrUrUrUrUrUrCrCrU/AltR2/
MG29-1 sgRNA targeting CD2	4816	MG29-1-CD2-sgRNA-A9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrCrUrUrUrUrGrGrUrGrArUrArUrArGrArArA/AltR2/
MG29-1 sgRNA targeting CD2	4817	MG29-1-CD2-sgRNA-B9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrUrUrUrGrGrUrGrArUrArUrArGrArArA/AltR2/
MG29-1 sgRNA targeting CD2	4818	MG29-1-CD2-sgRNA-C9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrUrUrUrGrGrUrGrArUrArUrArGrArArArA/AltR2/
MG29-1 sgRNA targeting CD2	4819	MG29-1-CD2-sgRNA-D9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrUrGrGrUrGrArUrArUrArGrArArArArCrGrA/AltR2/
MG29-1 sgRNA targeting CD2	4820	MG29-1-CD2-sgRNA-E9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrUrGrArUrArUrArGrArArArArCrGrArGrCrArGrU/AltR2/
MG29-1 sgRNA targeting CD2	4821	MG29-1-CD2-sgRNA-F9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrGrArUrArUrArGrArArArArCrGrArGrCrArGrUrG/AltR2/
MG29-1 sgRNA targeting CD2	4822	MG29-1-CD2-sgRNA-G9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrGrCrArArArArGrGrArArGrArArGrUrGrG/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting CD2	4823	MG29-1-CD2-sgRNA-H9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrUrGrUrUrGrCrArGrArUrGrArGrGrArGrCrUrGrG/AltR2/
MG29-1 sgRNA targeting CD2	4824	MG29-1-CD2-sgRNA-A10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrUrUrGrCrArGrArUrGrArGrGrArGrCrUrGrGrA/AltR2/
MG29-1 sgRNA targeting CD2	4825	MG29-1-CD2-sgRNA-B10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArUrCrUrUrUrUrUrUrArArUrUrArGrArGrGrArArG/AltR2/
MG29-1 sgRNA targeting CD2	4826	MG29-1-CD2-sgRNA-C10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrArArUrUrArGrArGrGrArArGrGrGrGrArCrArArU/AltR2/
MG29-1 sgRNA targeting CD2	4827	MG29-1-CD2-sgRNA-D10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArArUrUrArGrArGrGrArArGrGrGrArCrArArUrG/AltR2/
MG29-1 sgRNA targeting CD2	4828	MG29-1-CD2-sgRNA-E10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArUrUrArGrArGrGrArArGrGrGrArCrArArUrGrA/AltR2/
MG29-1 sgRNA targeting CD2	4829	MG29-1-CD2-sgRNA-F10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrUrArGrArGrGrArArGrGrGrArCrArArUrGrArG/AltR2/
MG29-1 sgRNA targeting CD2	4830	MG29-1-CD2-sgRNA-G10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrCrCrCrCrArUrGrGrGrGrArGrGrUrUrU/AltR2/
MG29-1 sgRNA targeting CD2	4831	MG29-1-CD2-sgRNA-H10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrCrUrGrCrCrCrCrArUrGrGrGrGrArGrGrUrUrU/AltR2/
MG29-1 sgRNA targeting CD2	4832	MG29-1-CD2-sgRNA-A11	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrCrUrGrArArCrUrCrGrArGrGrUrCrUrGrGrGrA/AltR2/
MG29-1 sgRNA targeting CD2	4833	MG29-1-CD2-sgRNA-B11	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrUrGrArArCrUrCrGrArGrGrUrCrUrGrGrGrGrA/AltR2/
MG29-1 sgRNA targeting CD2	4834	MG29-1-CD2-sgRNA-C11	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrCrUrGrUrGrArArCrUrUrGrUrGrGrUrGrArCrCrG/AltR2/
MG29-1 sgRNA targeting CD2	4835	MG29-1-CD2-sgRNA-D11	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrGrGrGrCrUrUrCrCrGrGrCrCrCrCrUrUrUrCrU/AltR2/
MG29-1 sgRNA targeting CD2	4836	MG29-1-CD2-sgRNA-E11	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrArGrUrArGrCrUrArCrUrCrUrCrUrGrUrGrGrGrCrU/AltR2/
DNA sequence of CD2 target site	4837	MG29-1-CD2-target site-A1	CATGTAAATTTGTAGCCAGCTT
DNA sequence of CD2 target site	4838	MG29-1-CD2-target site-B1	TAGCCAGCTTCCTTCTGATTTT
DNA sequence of CD2 target site	4839	MG29-1-CD2-target site-C1	ACTCTTATGCTTACCTTTGGAA
DNA sequence of CD2 target site	4840	MG29-1-CD2-target site-D1	GAAGAAACATTGAAAATCAGAA
DNA sequence of CD2 target site	4841	MG29-1-CD2-target site-E1	GCTTTTTATAGGTGCAGTCTCC

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of CD2 target site	4842	MG29-1-CD2-target site-F1	CTTTTTATAGGTGCAGTCTCCA
DNA sequence of CD2 target site	4843	MG29-1-CD2-target site-G1	TATAGGTGCAGTCTCCAAAGAG
DNA sequence of CD2 target site	4844	MG29-1-CD2-target site-H1	ATAGGTGCAGTCTCCAAAGAGA
DNA sequence of CD2 target site	4845	MG29-1-CD2-target site-A2	TAGGTGCAGTCTCCAAAGAGAT
DNA sequence of CD2 target site	4846	MG29-1-CD2-target site-B2	CAAATGAGTGATGATATTGACG
DNA sequence of CD2 target site	4847	MG29-1-CD2-target site-C2	AAATGAGTGATGATATTGACGA
DNA sequence of CD2 target site	4848	MG29-1-CD2-target site-D2	AAGGAAAAAGATACATATAAGC
DNA sequence of CD2 target site	4849	MG29-1-CD2-target site-E2	AAAATGGAACTCTGAAAATTAA
DNA sequence of CD2 target site	4850	MG29-1-CD2-target site-F2	TTCCAACACATTTTTTCCTTTT
DNA sequence of CD2 target site	4851	MG29-1-CD2-target site-G2	TCCAACACATTTTTTCCTTTTG
DNA sequence of CD2 target site	4852	MG29-1-CD2-target site-H2	CCAACACATTTTTTCCTTTTGT
DNA sequence of CD2 target site	4853	MG29-1-CD2-target site-A3	CAACACATTTTTTCCTTTTGTA
DNA sequence of CD2 target site	4854	MG29-1-CD2-target site-B3	TTCCTTTTGTATCATATATTGA
DNA sequence of CD2 target site	4855	MG29-1-CD2-target site-C3	TCCTTTTGTATCATATATTGAT
DNA sequence of CD2 target site	4856	MG29-1-CD2-target site-D3	CCTTTTGTATCATATATTGATA
DNA sequence of CD2 target site	4857	MG29-1-CD2-target site-E3	CTTTTGTATCATATATTGATAC
DNA sequence of CD2 target site	4858	MG29-1-CD2-target site-F3	GTATCATATATTGATACCTTGT
DNA sequence of CD2 target site	4859	MG29-1-CD2-target site-G3	TATCATATATTGATACCTTGTA
DNA sequence of CD2 target site	4860	MG29-1-CD2-target site-H3	CAGAGTTCCATTTTTAAATAGC

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of CD2 target site	4861	MG29-1-CD2-target site-A4	AGAGTTCCATTTTTAAATAGCT
DNA sequence of CD2 target site	4862	MG29-1-CD2-target site-B4	TAAATAGCTTATATGTATCTTT
DNA sequence of CD2 target site	4863	MG29-1-CD2-target site-C4	AAATAGCTTATATGTATCTTTT
DNA sequence of CD2 target site	4864	MG29-1-CD2-target site-D4	AATAGCTTATATGTATCTTTTT
DNA sequence of CD2 target site	4865	MG29-1-CD2-target site-E4	TCCTTGAAAGTCTCTTTCTCTT
DNA sequence of CD2 target site	4866	MG29-1-CD2-target site-F4	CCTTGAAAGTCTCTTTCTCTTT
DNA sequence of CD2 target site	4867	MG29-1-CD2-target site-G4	CTTGAAAGTCTCTTTCTCTTTT
DNA sequence of CD2 target site	4868	MG29-1-CD2-target site-H4	TCTTTTCTGAATTGTGCAATCT
DNA sequence of CD2 target site	4869	MG29-1-CD2-target site-A5	CTGAATTGTGCAATCTTTTTCT
DNA sequence of CD2 target site	4870	MG29-1-CD2-target site-B5	TGAATTGTGCAATCTTTTTCTT
DNA sequence of CD2 target site	4871	MG29-1-CD2-target site-C5	TCTTGTCTGAAGTTTTTTCCCA
DNA sequence of CD2 target site	4872	MG29-1-CD2-target site-D5	CTTGTCTGAAGTTTTTTCCCAT
DNA sequence of CD2 target site	4873	MG29-1-CD2-target site-E5	TTGTCTGAAGTTTTTTCCCATT
DNA sequence of CD2 target site	4874	MG29-1-CD2-target site-F5	TTCCCATTTTATATCGTCAATA
DNA sequence of CD2 target site	4875	MG29-1-CD2-target site-G5	TCCCATTTTATATCGTCAATAT
DNA sequence of CD2 target site	4876	MG29-1-CD2-target site-H5	CCCATTTTATATCGTCAATATC
DNA sequence of CD2 target site	4877	MG29-1-CD2-target site-A6	CCATTTTATATCGTCAATATCA
DNA sequence of CD2 target site	4878	MG29-1-CD2-target site-B6	ATATCGTCAATATCATCACTCA
DNA sequence of CD2 target site	4879	MG29-1-CD2-target site-C6	TATCGTCAATATCATCACTCAT

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of CD2 target site	4880	MG29-1-CD2-target site-D6	AAAACTAGGAATGTCCAAGTTG
DNA sequence of CD2 target site	4881	MG29-1-CD2-target site-E6	CAAGGCATTCGTAATCTCTTTG
DNA sequence of CD2 target site	4882	MG29-1-CD2-target site-F6	CTTTCTTTTTAGAGAGGGTCTC
DNA sequence of CD2 target site	4883	MG29-1-CD2-target site-G6	TTTTTAGAGAGGGTCTCAAAAC
DNA sequence of CD2 target site	4884	MG29-1-CD2-target site-H6	TAGAGAGGGTCTCAAAACCAAA
DNA sequence of CD2 target site	4885	MG29-1-CD2-target site-A7	AGAGAGGGTCTCAAAACCAAAG
DNA sequence of CD2 target site	4886	MG29-1-CD2-target site-B7	GAGAGGGTCTCAAAACCAAAGA
DNA sequence of CD2 target site	4887	MG29-1-CD2-target site-C7	TCAGAGGGTCATCACACACAAG
DNA sequence of CD2 target site	4888	MG29-1-CD2-target site-D7	TTCCCTGCTGTGCACTTGAATT
DNA sequence of CD2 target site	4889	MG29-1-CD2-target site-E7	GCACTCAGGCTGGTGGTCCACT
DNA sequence of CD2 target site	4890	MG29-1-CD2-target site-F7	CACTCAGGCTGGTGGTCCACTT
DNA sequence of CD2 target site	4891	MG29-1-CD2-target site-G7	AGATGTTTCCCATCTTGATACA
DNA sequence of CD2 target site	4892	MG29-1-CD2-target site-H7	GATGTTTCCCATCTTGATACAG
DNA sequence of CD2 target site	4893	MG29-1-CD2-target site-A8	CCATCTTGATACAGGTTTAATT
DNA sequence of CD2 target site	4894	MG29-1-CD2-target site-B8	ATTCGGGGTCAGTTCCATTCAT
DNA sequence of CD2 target site	4895	MG29-1-CD2-target site-C8	CCATCTTGATACAGGTTTAATT
DNA sequence of CD2 target site	4896	MG29-1-CD2-target site-D8	CAGAGAAAGGTCTGGACATCTA
DNA sequence of CD2 target site	4897	MG29-1-CD2-target site-E8	TGGCACTGCTCGTTTTCTATAT
DNA sequence of CD2 target site	4898	MG29-1-CD2-target site-F8	CTATATCACCAAAAGGAAAAAA

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of CD2 target site	4899	MG29-1-CD2-target site-G8	TATATCACCAAAAGGAAAAAAC
DNA sequence of CD2 target site	4900	MG29-1-CD2-target site-H8	TCCGACTCCTCTGTTTTTTCCT
DNA sequence of CD2 target site	4901	MG29-1-CD2-target site-A9	TTCCTTTTGGTGATATAGAAAA
DNA sequence of CD2 target site	4902	MG29-1-CD2-target site-B9	TCCTTTTGGTGATATAGAAAAC
DNA sequence of CD2 target site	4903	MG29-1-CD2-target site-C9	CCTTTTGGTGATATAGAAAACG
DNA sequence of CD2 target site	4904	MG29-1-CD2-target site-D9	CTTTTGGTGATATAGAAAACGA
DNA sequence of CD2 target site	4905	MG29-1-CD2-target site-E9	GGTGATATAGAAAACGAGCAGT
DNA sequence of CD2 target site	4906	MG29-1-CD2-target site-F9	GTGATATAGAAAACGAGCAGTG
DNA sequence of CD2 target site	4907	MG29-1-CD2-target site-G9	TCTGCAAAAGGAAGAGAAGTGG
DNA sequence of CD2 target site	4908	MG29-1-CD2-target site-H9	GTTGTTGCAGATGAGGAGCTGG
DNA sequence of CD2 target site	4909	MG29-1-CD2-target site-A10	TTGTTGCAGATGAGGAGCTGGA
DNA sequence of CD2 target site	4910	MG29-1-CD2-target site-B10	TATCTTTTTTAATTAGAGGAAG
DNA sequence of CD2 target site	4911	MG29-1-CD2-target site-C10	TTAATTAGAGGAAGGGGACAAT
DNA sequence of CD2 target site	4912	MG29-1-CD2-target site-D10	TAATTAGAGGAAGGGGACAATG
DNA sequence of CD2 target site	4913	MG29-1-CD2-target site-E10	AATTAGAGGAAGGGGACAATGA
DNA sequence of CD2 target site	4914	MG29-1-CD2-target site-F10	ATTAGAGGAAGGGGACAATGAG
DNA sequence of CD2 target site	4915	MG29-1-CD2-target site-G10	CTGCTGCCCCATGGGGAGGTTT
DNA sequence of CD2 target site	4916	MG29-1-CD2-target site-H10	TGCTGCCCCATGGGGAGGTTTT
DNA sequence of CD2 target site	4917	MG29-1-CD2-target site-A11	GGCTGAACTCGAGGTCTGGGGA

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of CD2 target site	4918	MG29-1-CD2-target site-B11	GCTGAACTCGAGGTCTGGGGAG
DNA sequence of CD2 target site	4919	MG29-1-CD2-target site-C11	TGCTGGTGAACTTGTGTGCCCG
DNA sequence of CD2 target site	4920	MG29-1-CD2-target site-D11	GTGGGGCTTCCGGCCCCTTTCT
DNA sequence of CD2 target site	4921	MG29-1-CD2-target site-E11	TTCAGTAGCTACTCTGTGGGCT
r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *; between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur; AltR1 and AltR2 refer to IDT technologies' proprietary 5' and 3' AltR modifications			

Example 17 – Gene editing outcomes at the DNA level for CD5

[00494] Primary T cells were purified from PMBCs using a negative selection kit (Miltenyi) according to the manufacturer's recommendations. Nucleofection of MG29-1 RNPs (126 pmol protein/160 pmol guide) (SEQ ID NOs: 4922-4945) was performed into T cells (200,000) using the Lonza 4D electroporator. Cells were harvested and genomic DNA prepared five days post-transfection. PCR primers appropriate for use in NGS-based DNA sequencing were generated, optimized, and used to amplify the individual target sequences for each guide RNA (SEQ ID NOs: 4946-4969). The amplicons were sequenced on an Illumina MiSeq machine and analyzed with a proprietary Python script to measure gene editing (**FIG. 63**).

Table 14F: Sequences of Guide RNAs and Sequences Targeted for Example 43

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting CD5	4922	MG29-1-CD5-sgRNA-A1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrCrArGrUrCrGrCrUrUrCrCrUrGrCrCrUrCrGrGrA/AltR2/
MG29-1 sgRNA targeting CD5	4923	MG29-1-CD5-sgRNA-B1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrArUrGrUrGrGrCrUrCrUrUrCrCrUrUrArCrCrU/AltR2/
MG29-1 sgRNA targeting CD5	4924	MG29-1-CD5-sgRNA-C1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrArCrCrCrCrCrArGrArUrUrUrCrCrArGrGrCrArA/AltR2/
MG29-1 sgRNA targeting CD5	4925	MG29-1-CD5-sgRNA-D1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrGrCrArArGrGrCrUrCrArCrCrCrGrUrUrCrCrA/AltR2/
MG29-1 sgRNA	4926	MG29-1-CD5-sgRNA-E1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrCrCrArGrArArGrCrUrGrGrGrCrCrGrGrArGrC/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
targeting CD5			
MG29-1 sgRNA targeting CD5	4927	MG29-1-CD5-sgRNA-F1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrCrUrGrUrGrGrCrUrGrCrArGrUrUrGrGrArGrArA/AltR2/
MG29-1 sgRNA targeting CD5	4928	MG29-1-CD5-sgRNA-G1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArCrGrCrUrUrGrArCrUrGrGrGrUrCrCrUrCrCrC/AltR2/
MG29-1 sgRNA targeting CD5	4929	MG29-1-CD5-sgRNA-H1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrGrCrUrUrGrArCrUrGrGrGrUrCrCrUrCrCrCrA/AltR2/
MG29-1 sgRNA targeting CD5	4930	MG29-1-CD5-sgRNA-A2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrGrGrCrUrUrGrArUrUrUrUrCrCrUrGrArArG/AltR2/
MG29-1 sgRNA targeting CD5	4931	MG29-1-CD5-sgRNA-B2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrGrGrGrCrUrUrGrArUrUrUrUrCrCrUrGrArArGrC/AltR2/
MG29-1 sgRNA targeting CD5	4932	MG29-1-CD5-sgRNA-C2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrUrGrArArGrCrArArUrGrCrUrCrCrArGrGrGrArG/AltR2/
MG29-1 sgRNA targeting CD5	4933	MG29-1-CD5-sgRNA-D2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrArArGrCrArArUrGrCrUrCrCrArGrGrGrArGrG/AltR2/
MG29-1 sgRNA targeting CD5	4934	MG29-1-CD5-sgRNA-E2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrArUrUrGrCrUrUrCrCrCrCrUrCrUrCrArGrGrGrUrU/AltR2/
MG29-1 sgRNA targeting CD5	4935	MG29-1-CD5-sgRNA-F2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrCrCrCrArArGrGrUrGrCrArGrArArGrCrCrGrUrC/AltR2/
MG29-1 sgRNA targeting CD5	4936	MG29-1-CD5-sgRNA-G2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrArGrArGrArArArUrUrCrCrUrArCrUrGrCrArArG/AltR2/
MG29-1 sgRNA targeting CD5	4937	MG29-1-CD5-sgRNA-H2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrCrCrCrArArArGrUrUrCrGrUrGrGrCrArCrUrG/AltR2/
MG29-1 sgRNA targeting CD5	4938	MG29-1-CD5-sgRNA-A3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrArCrUrArGrCrUrUrCrUrUrGrUrArGrGrCrCrArA/AltR2/
MG29-1 sgRNA targeting CD5	4939	MG29-1-CD5-sgRNA-B3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrUrCrUrCrUrUrGrCrCrCrArGrUrCrCrGrCrCrA/AltR2/
MG29-1 sgRNA targeting CD5	4940	MG29-1-CD5-sgRNA-C3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrUrUrCrArUrUrCrCrCrGrUrUrGrGrGrCrCrCrArArU/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting CD5	4941	MG29-1-CD5-sgRNA-D3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrUrCrArUrUrCrCrCrGrUrUrGrGrGrCrCrArArUrC/AltR2/
MG29-1 sgRNA targeting CD5	4942	MG29-1-CD5-sgRNA-E3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArUrCrGrCrArArCrCrArCrArCrGrGrCrArArCrCrG/AltR2/
MG29-1 sgRNA targeting CD5	4943	MG29-1-CD5-sgRNA-F3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrCrCrCrCrArGrCrUrCrUrGrGrArArGrGrGrGrC/AltR2/
MG29-1 sgRNA targeting CD5	4944	MG29-1-CD5-sgRNA-G3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrCrArGrCrUrCrUrGrGrArArGrGrGrCrUrCrUrG/AltR2/
MG29-1 sgRNA targeting CD5	4945	MG29-1-CD5-sgRNA-H3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrCrCrUrCrUrGrArGrCrCrCrCrArUrGrCrArGrA/AltR2/
DNA sequence of CD5 target site	4946	MG29-1-CD5-target site-A1	TGCAGTCGCTTCCTGCCTCGGA
DNA sequence of CD5 target site	4947	MG29-1-CD5-target site-B1	TCCATGTGGCTCTTCCTTACCT
DNA sequence of CD5 target site	4948	MG29-1-CD5-target site-C1	TGACCCCCAGATTTCCAGGCAA
DNA sequence of CD5 target site	4949	MG29-1-CD5-target site-D1	CAGGCAAGGCTCACCCGTTCCA
DNA sequence of CD5 target site	4950	MG29-1-CD5-target site-E1	CAGCCAGAGCTGGGGCCGGAGC
DNA sequence of CD5 target site	4951	MG29-1-CD5-target site-F1	TGCTGTGGCTGCAGTTGGAGAA
DNA sequence of CD5 target site	4952	MG29-1-CD5-target site-G1	GACGCTTGACTGGGGTCCTCCC
DNA sequence of CD5 target site	4953	MG29-1-CD5-target site-H1	ACGCTTGACTGGGGTCCTCCCA
DNA sequence of CD5 target site	4954	MG29-1-CD5-target site-A2	CTGGGGCTTGATTTTCCTGAAG
DNA sequence	4955	MG29-1-CD5-target site-B2	TGGGGCTTGATTTTCCTGAAGC

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
of CD5 target site			
DNA sequence of CD5 target site	4956	MG29-1-CD5-target site-C2	CCTGAAGCAATGCTCCAGGGAG
DNA sequence of CD5 target site	4957	MG29-1-CD5-target site-D2	CTGAAGCAATGCTCCAGGGAGG
DNA sequence of CD5 target site	4958	MG29-1-CD5-target site-E2	CCATTGCTTCCCCTCTCAGGTT
DNA sequence of CD5 target site	4959	MG29-1-CD5-target site-F2	CAGCCCAAGGTGCAGAGCCGTC
DNA sequence of CD5 target site	4960	MG29-1-CD5-target site-G2	GGAGAGAAATTCCTACTGCAAG
DNA sequence of CD5 target site	4961	MG29-1-CD5-target site-H2	TCTCCCAAAGTTCGTGGCACTG
DNA sequence of CD5 target site	4962	MG29-1-CD5-target site-A3	TTCACTAGCTTCTTGTAGGCAA
DNA sequence of CD5 target site	4963	MG29-1-CD5-target site-B3	TTGTCTCTTGCCCAGTCCGCCA
DNA sequence of CD5 target site	4964	MG29-1-CD5-target site-C3	GGTTCATTCCCGTTGGGCCAAT
DNA sequence of CD5 target site	4965	MG29-1-CD5-target site-D3	GTTCATTCCCGTTGGGCCAATC
DNA sequence of CD5 target site	4966	MG29-1-CD5-target site-E3	CATCGCAACCACACGGCAACCG
DNA sequence of CD5 target site	4967	MG29-1-CD5-target site-F3	TTTCCCCAGCTCTGGAAGGGGC
DNA sequence of CD5 target site	4968	MG29-1-CD5-target site-G3	CCCAGCTCTGGAAGGGGCTCTG
DNA sequence of CD5 target site	4969	MG29-1-CD5-target site-H3	CAGCCTCTGAGCCCCATGCAGA
r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written			

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *; between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur; AltR1 and AltR2 refer to IDT technologies' proprietary 5' and 3' AltR modifications			

Example 18 – Gene editing outcomes at the DNA level for mouse TRAC

[00495] Primary T cells were purified from C57BL/6 mouse spleens. Nucleofection of MG29-1 RNPs (126 pmol protein/160 pmol guide) (SEQ ID NOs: 5056-5125) was performed into T cells (200,000) using the Lonza 4D electroporator and 100 pmol transfection enhancer (IDT). Cells were harvested and genomic DNA prepared five days post-transfection. PCR primers appropriate for use in NGS-based DNA sequencing were generated, optimized, and used to amplify the individual target sequences for each guide RNA (SEQ ID NOs: 5126-5195). The amplicons were sequenced on an Illumina MiSeq machine and analyzed with a proprietary Python script to measure gene editing (**FIG. 64A**).

Table 14G: Sequences of Guide RNAs and Sequences Targeted for Example 44

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting mouse TRAC	5056	MG29-1-mTRAC-sgRNA-A1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrUrCrCrArArArUrCrArArUrGrUrGrCrCrGrArA/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5057	MG29-1-mTRAC-sgRNA-B1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArGrArUrArUrCrUrUrGrGrCrArGrGrUrGrArArGrC/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5058	MG29-1-mTRAC-sgRNA-C1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrGrUrCrCrArGrCrArCrArGrUrUrUrUrGrUrCrArG/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5059	MG29-1-mTRAC-sgRNA-D1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrCrArGrUrGrArUrGrArArCrGrUrUrCrCrArGrArU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5060	MG29-1-mTRAC-sgRNA-E1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrArGrUrGrArUrGrArArCrGrUrUrCrCrArGrArUrU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5061	MG29-1-mTRAC-sgRNA-F1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrGrGrCrArCrArUrUrGrArUrUrUrUrGrGrGrArGrUrCrA/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5062	MG29-1-mTRAC-sgRNA-G1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrCrArCrArUrUrGrArUrUrUrUrGrGrGrArGrUrCrArA/AltR2/
MG29-1 sgRNA	5063	MG29-1-mTRAC-sgRNA-H1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrArGrUrCrArArArGrUrCrGrUrGrArArCrArArGrG/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
targeting mouse TRAC			
MG29-1 sgRNA targeting mouse TRAC	5064	MG29-1-mTRAC-sgRNA-A2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArCrUrGrGrUrArCrArCrArGrCrArGrGrUrUrCrUrG/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5065	MG29-1-mTRAC-sgRNA-B2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrUrGrGrUrArCrArCrArGrCrArGrGrUrUrCrUrGrG/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5066	MG29-1-mTRAC-sgRNA-C2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrCrCrCrUrUrUrArGrArCrGrUrUrCrCrCrU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5067	MG29-1-mTRAC-sgRNA-D2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrCrUrUrUrArGrArCrGrUrUrCrCrCrUrGrUrGrA/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5068	MG29-1-mTRAC-sgRNA-E2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrUrUrUrArGrArCrGrUrUrCrCrCrUrGrUrGrArU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5069	MG29-1-mTRAC-sgRNA-F2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrArCrGrUrUrCrCrCrUrGrUrGrArUrGrCrCrArCrG/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5070	MG29-1-mTRAC-sgRNA-G2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArCrGrUrUrCrCrCrUrGrUrGrArUrGrCrCrArCrGrU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5071	MG29-1-mTRAC-sgRNA-H2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArArGrCrUrUrUrUrCrUrCrArGrUrCrArArCrGrUrG/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5072	MG29-1-mTRAC-sgRNA-A3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrCrArGrUrCrArArCrGrUrGrGrCrArUrCrArCrArG/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5073	MG29-1-mTRAC-sgRNA-B3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrArGrUrCrArArCrGrUrGrGrCrArUrCrArCrArGrG/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5074	MG29-1-mTRAC-sgRNA-C3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrArGrArUrArUrGrArArCrCrUrArArArCrUrUrUrC/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5075	MG29-1-mTRAC-sgRNA-D3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrArUrArUrGrArArCrCrUrArArArCrUrUrUrCrA/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5076	MG29-1-mTRAC-sgRNA-E3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArArCrCrUrGrUrCrArGrUrUrArUrGrGrGrArCrU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5077	MG29-1-mTRAC-sgRNA-F3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrCrUrGrCrUrCrArUrGrArCrGrCrUrGrArGrGrCrU/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting mouse TRAC	5078	MG29-1-mTRAC-sgRNA-G3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrGrCrCrUrUrCrUrUrArCrCrUrCrArArCrUrGrGrA/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5079	MG29-1-mTRAC-sgRNA-H3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrCrArGrGrArGrGrArUrUrCrGrGrArGrUrCrCrCrA/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5080	MG29-1-mTRAC-sgRNA-A4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrArGrGrArArArGrGrUrUrGrCrUrGrGrArGrArGrCrU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5081	MG29-1-mTRAC-sgRNA-B4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrUrUrUrUrUrUrUrUrUrUrUrUrUrGrCrGrGrG/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5082	MG29-1-mTRAC-sgRNA-C4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrUrUrUrUrUrUrUrUrUrUrUrUrUrGrCrGrGrGrU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5083	MG29-1-mTRAC-sgRNA-D4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrUrUrUrUrUrUrUrUrUrUrUrUrGrCrGrGrGrUrU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5084	MG29-1-mTRAC-sgRNA-E4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrUrUrUrUrUrUrUrUrGrCrGrGrGrUrUrU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5085	MG29-1-mTRAC-sgRNA-F4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrUrUrUrUrGrCrGrGrGrUrUrArUrUrU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5086	MG29-1-mTRAC-sgRNA-G4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrUrUrUrGrCrGrGrGrUrUrArUrUrUrU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5087	MG29-1-mTRAC-sgRNA-H4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrUrUrUrGrCrGrGrGrUrUrArUrUrUrU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5088	MG29-1-mTRAC-sgRNA-A5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrUrUrUrGrCrGrGrGrUrUrArUrUrUrUrU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5089	MG29-1-mTRAC-sgRNA-B5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrGrCrGrGrGrUrUrArUrUrUrUrUrU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5090	MG29-1-mTRAC-sgRNA-C5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrGrCrGrGrGrUrUrArUrUrUrUrUrUrA/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5091	MG29-1-mTRAC-sgRNA-D5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrGrCrGrGrGrUrUrArUrUrUrUrUrUrArA/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting mouse TRAC	5092	MG29-1-mTRAC-sgRNA-E5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrGrCrGrGrGrUrUrUrArUrUrUrUrUrUrUrArArG/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5093	MG29-1-mTRAC-sgRNA-F5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrGrCrGrGrGrUrUrUrArUrUrUrUrUrUrUrArArGrC/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5094	MG29-1-mTRAC-sgRNA-G5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrCrGrGrGrUrUrUrArUrUrUrUrUrUrUrArArGrCrA/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5095	MG29-1-mTRAC-sgRNA-H5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrGrGrGrUrUrUrArUrUrUrUrUrUrUrArArGrCrArU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5096	MG29-1-mTRAC-sgRNA-A6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrGrGrGrUrUrUrArUrUrUrUrUrUrUrArArGrCrArUrC/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5097	MG29-1-mTRAC-sgRNA-B6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrUrUrArArGrCrArUrCrCrArUrGrArArGrA/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5098	MG29-1-mTRAC-sgRNA-C6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrArArGrCrArUrCrCrArUrGrArArGrArArArUrG/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5099	MG29-1-mTRAC-sgRNA-D6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrArArGrCrArUrCrCrArUrGrArArGrArArArUrGrC/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5100	MG29-1-mTRAC-sgRNA-E6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArArGrCrArUrCrCrArUrGrArArGrArArArUrGrCrA/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5101	MG29-1-mTRAC-sgRNA-F6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArGrCrArUrCrCrArUrGrArArGrArArArUrGrCrArU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5102	MG29-1-mTRAC-sgRNA-G6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrCrArUrCrCrArUrGrArArGrArArArUrGrCrArUrA/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5103	MG29-1-mTRAC-sgRNA-H6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrCrArArGrGrUrGrUrArGrArArArUrUrArUrCrUrC/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5104	MG29-1-mTRAC-sgRNA-A7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrUrUrUrUrUrCrUrGrArArUrCrArCrCrUrUrUrArA/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5105	MG29-1-mTRAC-sgRNA-B7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrUrUrUrUrCrUrGrArArUrCrArCrCrUrUrUrArArU/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting mouse TRAC	5106	MG29-1-mTRAC-sgRNA-C7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrGrArArUrCrArCrCrUrUrUrArArUrGrArUrGrU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5107	MG29-1-mTRAC-sgRNA-D7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrArArUrCrArCrCrUrUrUrArArUrGrArUrGrUrC/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5108	MG29-1-mTRAC-sgRNA-E7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrArArUrCrArCrCrUrUrUrArArUrGrArUrGrUrCrA/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5109	MG29-1-mTRAC-sgRNA-F7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrGrArArUrGrUrCrArCrArGrCrCrArGrArArU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5110	MG29-1-mTRAC-sgRNA-G7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArCrArCrCrUrUrGrArUrGrArArArGrArGrUrArArU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5111	MG29-1-mTRAC-sgRNA-H7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrArUrGrGrArUrGrCrUrUrArArArArArArArUrA/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5112	MG29-1-mTRAC-sgRNA-A8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrArCrArArCrArUrUrCrUrCrCrArArGrArGrA/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5113	MG29-1-mTRAC-sgRNA-B8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrArCrArArCrArUrUrCrUrCrCrArArGrArGrArU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5114	MG29-1-mTRAC-sgRNA-C8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrArCrArArCrArUrUrCrUrCrCrArArGrArGrArUrU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5115	MG29-1-mTRAC-sgRNA-D8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArCrArArCrArUrUrCrUrCrCrArArGrArGrArUrUrU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5116	MG29-1-mTRAC-sgRNA-E8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrArArCrArUrUrCrUrCrCrArArGrArGrArUrUrUrU/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5117	MG29-1-mTRAC-sgRNA-F8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArArCrArUrUrCrUrCrCrArArGrArGrArUrUrUrUrA/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5118	MG29-1-mTRAC-sgRNA-G8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrArGrGrGrGrArGrUrCrUrGrCrCrArUrGrGrGrGrG/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5119	MG29-1-mTRAC-sgRNA-H8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrGrGrGrArGrUrCrUrGrCrCrArUrGrGrGrGrGrA/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting mouse TRAC	5120	MG29-1-mTRAC-sgRNA-A9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrUrGrArArUrGrGrUrCrArGrCrArGrCrArGrUrG/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5121	MG29-1-mTRAC-sgRNA-B9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrUrGrArArUrGrGrUrCrArGrCrArGrCrArGrUrGrA/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5122	MG29-1-mTRAC-sgRNA-C9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrGrArArUrGrGrUrCrArGrCrArGrCrArGrUrGrArG/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5123	MG29-1-mTRAC-sgRNA-D9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrArArUrGrGrUrCrArGrCrArGrCrArGrUrGrArGrG/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5124	MG29-1-mTRAC-sgRNA-E9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrCrUrUrGrArArGrGrArArGrGrCrGrGrArGrGrG/AltR2/
MG29-1 sgRNA targeting mouse TRAC	5125	MG29-1-mTRAC-sgRNA-F9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrUrUrGrArArGrGrArArGrCrGrGrArGrGrGrG/AltR2/
DNA sequence of TRAC target site	5126	MG29-1-mTRAC-target site-A1	ACTCCCAAATCAATGTGCCGAA
DNA sequence of TRAC target site	5127	MG29-1-mTRAC-target site-B1	AAGATATCTTGGCAGGTGAAGC
DNA sequence of TRAC target site	5128	MG29-1-mTRAC-target site-C1	ATGTCCAGCACAGTTTTGTCAG
DNA sequence of TRAC target site	5129	MG29-1-mTRAC-target site-D1	GTCAGTGATGAACGTTCCAGAT
DNA sequence of TRAC target site	5130	MG29-1-mTRAC-target site-E1	TCAGTGATGAACGTTCCAGATT
DNA sequence of TRAC target site	5131	MG29-1-mTRAC-target site-F1	CGGCACATTGATTTGGGAGTCA
DNA sequence of TRAC target site	5132	MG29-1-mTRAC-target site-G1	GGCACATTGATTTGGGAGTCAA
DNA sequence of TRAC target site	5133	MG29-1-mTRAC-target site-H1	GGAGTCAAAGTCGGTGAACAGG
DNA sequence of TRAC target site	5134	MG29-1-mTRAC-target site-A2	AACTGGTACACAGCAGGTTCTG
DNA sequence of TRAC target site	5135	MG29-1-mTRAC-target site-B2	ACTGGTACACAGCAGGTTCTGG
DNA sequence of TRAC target site	5136	MG29-1-mTRAC-target site-C2	TTTTCCCTTTTAGACGTTCCCT

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of TRAC target site	5137	MG29-1-mTRAC-target site-D2	CCCTTTTAGACGTTCCCTGTGA
DNA sequence of TRAC target site	5138	MG29-1-mTRAC-target site-E2	CCTTTTAGACGTTCCCTGTGAT
DNA sequence of TRAC target site	5139	MG29-1-mTRAC-target site-F2	AGACGTTCCCTGTGATGCCACG
DNA sequence of TRAC target site	5140	MG29-1-mTRAC-target site-G2	GACGTTCCCTGTGATGCCACGT
DNA sequence of TRAC target site	5141	MG29-1-mTRAC-target site-H2	AAAGCTTTTCTCAGTCAACGTG
DNA sequence of TRAC target site	5142	MG29-1-mTRAC-target site-A3	CTCAGTCAACGTGGCATCACAG
DNA sequence of TRAC target site	5143	MG29-1-mTRAC-target site-B3	TCAGTCAACGTGGCATCACAGG
DNA sequence of TRAC target site	5144	MG29-1-mTRAC-target site-C3	ACAGATATGAACCTAAACTTTC
DNA sequence of TRAC target site	5145	MG29-1-mTRAC-target site-D3	CAGATATGAACCTAAACTTTCA
DNA sequence of TRAC target site	5146	MG29-1-mTRAC-target site-E3	AAAACCTGTCAGTTATGGGACT
DNA sequence of TRAC target site	5147	MG29-1-mTRAC-target site-F3	ACCTGCTCATGACGCTGAGGCT
DNA sequence of TRAC target site	5148	MG29-1-mTRAC-target site-G3	ATGCCTTCTTACCTCAACTGGA
DNA sequence of TRAC target site	5149	MG29-1-mTRAC-target site-H3	AGCAGGAGGATTCGGAGTCCCA
DNA sequence of TRAC target site	5150	MG29-1-mTRAC-target site-A4	TGAGGAAGGTTGCTGGAGAGCT
DNA sequence of TRAC target site	5151	MG29-1-mTRAC-target site-B4	TTGTTTTTTTTTTTTTTGCGGG
DNA sequence of TRAC target site	5152	MG29-1-mTRAC-target site-C4	TGTTTTTTTTTTTTTTGCGGGT
DNA sequence of TRAC target site	5153	MG29-1-mTRAC-target site-D4	GTTTTTTTTTTTTTTGCGGGTT
DNA sequence of TRAC target site	5154	MG29-1-mTRAC-target site-E4	TTTTTTTTTTTTTTGCGGGTTT
DNA sequence of TRAC target site	5155	MG29-1-mTRAC-target site-F4	TTTTTTTTTTGCGGGTTTATTT

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of TRAC target site	5156	MG29-1-mTRAC-target site-G4	TTTTTTTTTGCGGGTTTATTTT
DNA sequence of TRAC target site	5157	MG29-1-mTRAC-target site-H4	TTTTTTTTGCGGGTTTATTTTT
DNA sequence of TRAC target site	5158	MG29-1-mTRAC-target site-A5	TTTTTTTGCGGGTTTATTTTTT
DNA sequence of TRAC target site	5159	MG29-1-mTRAC-target site-B5	TTTTTTGCGGGTTTATTTTTTT
DNA sequence of TRAC target site	5160	MG29-1-mTRAC-target site-C5	TTTTTGCGGGTTTATTTTTTTA
DNA sequence of TRAC target site	5161	MG29-1-mTRAC-target site-D5	TTTTGCGGGTTTATTTTTTTAA
DNA sequence of TRAC target site	5162	MG29-1-mTRAC-target site-E5	TTTGCGGGTTTATTTTTTTAAG
DNA sequence of TRAC target site	5163	MG29-1-mTRAC-target site-F5	TTGCGGGTTTATTTTTTTAAGC
DNA sequence of TRAC target site	5164	MG29-1-mTRAC-target site-G5	TGCGGGTTTATTTTTTTAAGCA
DNA sequence of TRAC target site	5165	MG29-1-mTRAC-target site-H5	GCGGGTTTATTTTTTTAAGCAT
DNA sequence of TRAC target site	5166	MG29-1-mTRAC-target site-A6	CGGGTTTATTTTTTTAAGCATC
DNA sequence of TRAC target site	5167	MG29-1-mTRAC-target site-B6	TTTTTTTAAGCATCCATGAAGA
DNA sequence of TRAC target site	5168	MG29-1-mTRAC-target site-C6	TTTAAGCATCCATGAAGAAATG
DNA sequence of TRAC target site	5169	MG29-1-mTRAC-target site-D6	TTAAGCATCCATGAAGAAATGC
DNA sequence of TRAC target site	5170	MG29-1-mTRAC-target site-E6	TAAGCATCCATGAAGAAATGCA
DNA sequence of TRAC target site	5171	MG29-1-mTRAC-target site-F6	AAGCATCCATGAAGAAATGCAT
DNA sequence of TRAC target site	5172	MG29-1-mTRAC-target site-G6	AGCATCCATGAAGAAATGCATA
DNA sequence of TRAC target site	5173	MG29-1-mTRAC-target site-H6	ATCAAGGTGTAGAAATTATCTC
DNA sequence of TRAC target site	5174	MG29-1-mTRAC-target site-A7	GGTTTTTCTGAATCACCTTTAA

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of TRAC target site	5175	MG29-1-mTRAC-target site-B7	GTTTTTCTGAATCACCTTTAAT
DNA sequence of TRAC target site	5176	MG29-1-mTRAC-target site-C7	TCTGAATCACCTTTAATGATGT
DNA sequence of TRAC target site	5177	MG29-1-mTRAC-target site-D7	CTGAATCACCTTTAATGATGTC
DNA sequence of TRAC target site	5178	MG29-1-mTRAC-target site-E7	TGAATCACCTTTAATGATGTCA
DNA sequence of TRAC target site	5179	MG29-1-mTRAC-target site-F7	ATGATGTCATGGACAGCAGAAT
DNA sequence of TRAC target site	5180	MG29-1-mTRAC-target site-G7	TACACCTTGATGAAAGAGTAAT
DNA sequence of TRAC target site	5181	MG29-1-mTRAC-target site-H7	TTCATGGATGCTTAAAAAAATA
DNA sequence of TRAC target site	5182	MG29-1-mTRAC-target site-A8	TTTTACAACATTCTCCAAGAGA
DNA sequence of TRAC target site	5183	MG29-1-mTRAC-target site-B8	TTTACAACATTCTCCAAGAGAT
DNA sequence of TRAC target site	5184	MG29-1-mTRAC-target site-C8	TTACAACATTCTCCAAGAGATT
DNA sequence of TRAC target site	5185	MG29-1-mTRAC-target site-D8	TACAACATTCTCCAAGAGATTT
DNA sequence of TRAC target site	5186	MG29-1-mTRAC-target site-E8	ACAACATTCTCCAAGAGATTTT
DNA sequence of TRAC target site	5187	MG29-1-mTRAC-target site-F8	CAACATTCTCCAAGAGATTTTA
DNA sequence of TRAC target site	5188	MG29-1-mTRAC-target site-G8	ACAGGGGAGTCTGCCATGGGGG
DNA sequence of TRAC target site	5189	MG29-1-mTRAC-target site-H8	CAGGGGAGTCTGCCATGGGGGA
DNA sequence of TRAC target site	5190	MG29-1-mTRAC-target site-A9	TTGTGAATGGTCAGCAGCAGTG
DNA sequence of TRAC target site	5191	MG29-1-mTRAC-target site-B9	TGTGAATGGTCAGCAGCAGTGA
DNA sequence of TRAC target site	5192	MG29-1-mTRAC-target site-C9	GTGAATGGTCAGCAGCAGTGAG
DNA sequence of TRAC target site	5193	MG29-1-mTRAC-target site-D9	TGAATGGTCAGCAGCAGTGAGG

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of TRAC target site	5194	MG29-1-mTRAC-target site-E9	GGCTTGAAGAAGGAGCGGAGGG
DNA sequence of TRAC target site	5195	MG29-1-mTRAC-target site-F9	GCTTGAAGAAGGAGCGGAGGGG
r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *; between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur; AltR1 and AltR2 refer to IDT technologies' proprietary 5' and 3' AltR modifications			

[00496] For analysis by flow cytometry, 3 days post-nucleofection, 100,000 mouse T cells were stained with anti-mouse CD3 antibody (Clone 17A2, Invitrogen 11-0032-82) for 30 minutes at 4C and analyzed on an Attune Nxt flow cytometer. Guides at the 3' end of the mouse TRAC gene have high levels of indels but fail to produce a knock-out of TRAC. Surprisingly and unexpectedly, the phenotype does not correlate with the genotype near the end of the gene, likely due to the retention of function of truncated alleles and the failure of nonsense-mediated decay pathways to remove prematurely truncated out-of-frame mRNAs (**FIG. 65**).

Example 19 – Gene editing outcomes at the DNA level for mouse TRBC1 and TRBC2

[00497] Primary T cells were purified from C57BL/6 mouse spleens. Nucleofection of MG29-1 RNPs (104 pmol protein/120 pmol guide) (TRBC1: SEQ ID NOs: 5196-5210; TRBC2: SEQ ID NOs: 5226-5246) was performed into T cells (200,000) using the Lonza 4D electroporator and 100 pmol transfection enhancer (IDT). Cells were harvested and genomic DNA prepared five days post-transfection. PCR primers appropriate for use in NGS-based DNA sequencing were generated, optimized, and used to amplify the individual target sequences for each guide RNA (TRBC1: SEQ ID NOs: 5211-5225; TRBC2: SEQ ID NOs: 5247-5267). The amplicons were sequenced on an Illumina MiSeq machine and analyzed with a proprietary Python script to measure gene editing (**FIG. 66A** and **66B**). For analysis by flow cytometry, 3 days post-nucleofection, 100,000 mouse T cells were stained with anti-mouse CD3 antibody (Clone 17A2, Invitrogen 11-0032-82) for 30 minutes at 4C and analyzed on an Attune Nxt flow cytometer.

Table 14H: Sequences of Guide RNAs and Sequences Targeted for Example 45

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting mouse TRBC1	5196	MG29-1-mTRBC1-sgRNA-A1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrCrCrArGrArGrGrArUrCrUrGrArGrArArU/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting mouse TRBC1	5197	MG29-1-mTRBC1-sgRNA-B1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrArGrGrArUrCrUrGrArGrArArArUrGrUrGrArC/AltR2/
MG29-1 sgRNA targeting mouse TRBC1	5198	MG29-1-mTRBC1-sgRNA-C1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrCrCrArUrCrArArArArGrCrArGrArGrArUrUrGrC/AltR2/
MG29-1 sgRNA targeting mouse TRBC1	5199	MG29-1-mTRBC1-sgRNA-D1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrArGrGrArCrArArGrUrGrGrCrCrArGrArG/AltR2/
MG29-1 sgRNA targeting mouse TRBC1	5200	MG29-1-mTRBC1-sgRNA-E1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrUrGrArGrCrCrCrUrCrUrGrGrCrCrArCrUrUrGrU/AltR2/
MG29-1 sgRNA targeting mouse TRBC1	5201	MG29-1-mTRBC1-sgRNA-F1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrUrUrGrUrUrUrGrCrArArUrCrUrCrUrGrCrUrUrU/AltR2/
MG29-1 sgRNA targeting mouse TRBC1	5202	MG29-1-mTRBC1-sgRNA-G1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrGrUrUrUrGrCrArArUrCrUrCrUrGrCrUrUrU/AltR2/
MG29-1 sgRNA targeting mouse TRBC1	5203	MG29-1-mTRBC1-sgRNA-H1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrGrCrArArUrCrUrCrUrGrCrUrUrUrUrGrArUrG/AltR2/
MG29-1 sgRNA targeting mouse TRBC1	5204	MG29-1-mTRBC1-sgRNA-A2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArArUrCrUrCrUrGrCrUrUrUrUrGrArUrGrGrCrUrC/AltR2/
MG29-1 sgRNA targeting mouse TRBC1	5205	MG29-1-mTRBC1-sgRNA-B2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArUrGrGrCrUrCrArArArCrArArGrGrArGrArCrCrU/AltR2/
MG29-1 sgRNA targeting mouse TRBC1	5206	MG29-1-mTRBC1-sgRNA-C2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrGrGrCrUrCrArArArCrArArGrGrArGrArCrCrUrU/AltR2/
MG29-1 sgRNA targeting mouse TRBC1	5207	MG29-1-mTRBC1-sgRNA-D2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrCrUrUrCrUrUrUrCrArGrArCrUrGrUrGrGrGrA/AltR2/
MG29-1 sgRNA targeting mouse TRBC1	5208	MG29-1-mTRBC1-sgRNA-E2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrCrArArCrArGrCrArUrCrCrUrArUrCrArArArC/AltR2/
MG29-1 sgRNA targeting mouse TRBC1	5209	MG29-1-mTRBC1-sgRNA-F2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrUrCrArArCrArGrCrArUrCrCrUrArUrCrArArCrA/AltR2/
MG29-1 sgRNA targeting mouse TRBC1	5210	MG29-1-mTRBC1-sgRNA-G2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrUrArGrCrArGrGrArUrCrUrCrArUrArGrArGrGrGrA/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of TRBC1 target site	5211	MG29-1-mTRBC1-target site-A1	CTTTCCAGAGGATCTGAGAAAT
DNA sequence of TRBC1 target site	5212	MG29-1-mTRBC1-target site-B1	CAGAGGATCTGAGAAATGTGAC
DNA sequence of TRBC1 target site	5213	MG29-1-mTRBC1-target site-C1	AGCCATCAAAAGCAGAGATTGC
DNA sequence of TRBC1 target site	5214	MG29-1-mTRBC1-target site-D1	AGAGGAGGACAAGTGGCCAGAG
DNA sequence of TRBC1 target site	5215	MG29-1-mTRBC1-target site-E1	GGTGAGCCCTCTGGCCACTTGT
DNA sequence of TRBC1 target site	5216	MG29-1-mTRBC1-target site-F1	GTTTGTTTGCAATCTCTGCTTT
DNA sequence of TRBC1 target site	5217	MG29-1-mTRBC1-target site-G1	TTTGTTTGCAATCTCTGCTTTT
DNA sequence of TRBC1 target site	5218	MG29-1-mTRBC1-target site-H1	TTTGCAATCTCTGCTTTTGATG
DNA sequence of TRBC1 target site	5219	MG29-1-mTRBC1-target site-A2	CAATCTCTGCTTTTGATGGCTC
DNA sequence of TRBC1 target site	5220	MG29-1-mTRBC1-target site-B2	GATGGCTCAAACAAGGAGACCT
DNA sequence of TRBC1 target site	5221	MG29-1-mTRBC1-target site-C2	ATGGCTCAAACAAGGAGACCTT
DNA sequence of TRBC1 target site	5222	MG29-1-mTRBC1-target site-D2	TCTCTTCTTTCAGACTGTGGGA
DNA sequence of TRBC1 target site	5223	MG29-1-mTRBC1-target site-E2	CTGTCAACAGCATCCTATCAAC
DNA sequence of TRBC1 target site	5224	MG29-1-mTRBC1-target site-F2	TGTCAACAGCATCCTATCAACA
DNA sequence of TRBC1 target site	5225	MG29-1-mTRBC1-target site-G2	CCTAGCAGGATCTCATAGAGGA
MG29-1 sgRNA targeting mouse TRBC2	5226	MG29-1-mTRBC2-sgRNA-A1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrUrCrCrArGrArGrGrArUrCrUrGrArGrArArArU/AltR2/
MG29-1 sgRNA targeting mouse TRBC2	5227	MG29-1-mTRBC2-sgRNA-B1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrArGrGrArUrCrUrGrArGrArArArUrGrUrGrArC/AltR2/
MG29-1 sgRNA targeting mouse TRBC2	5228	MG29-1-mTRBC2-sgRNA-C1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrCrCrArUrCrArArArArGrCrArGrArGrArUrUrGrC/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting mouse TRBC2	5229	MG29-1-mTRBC2-sgRNA-D1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrArGrGrArGrGrArCrArArGrUrGrGrCrCrArGrArG/AltR2/
MG29-1 sgRNA targeting mouse TRBC2	5230	MG29-1-mTRBC2-sgRNA-E1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrCrArUrGrArGrCrUrCrCrGrCrArCrUrUrArCrCrU/AltR2/
MG29-1 sgRNA targeting mouse TRBC2	5231	MG29-1-mTRBC2-sgRNA-F1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrUrGrArGrCrCrCrUrCrUrGrGrCrCrArCrUrUrGrU/AltR2/
MG29-1 sgRNA targeting mouse TRBC2	5232	MG29-1-mTRBC2-sgRNA-G1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArGrGrArUrUrGrUrGrCrCrArGrArArGrGrUrArGrC/AltR2/
MG29-1 sgRNA targeting mouse TRBC2	5233	MG29-1-mTRBC2-sgRNA-H1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrUrUrGrUrUrUrGrCrArArUrCrUrCrUrGrCrUrUrU/AltR2/
MG29-1 sgRNA targeting mouse TRBC2	5234	MG29-1-mTRBC2-sgRNA-A2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrGrUrUrUrGrCrArArUrCrUrCrUrGrCrUrUrUrU/AltR2/
MG29-1 sgRNA targeting mouse TRBC2	5235	MG29-1-mTRBC2-sgRNA-B2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrGrCrArArUrCrUrCrUrGrCrUrUrUrUrGrArUrG/AltR2/
MG29-1 sgRNA targeting mouse TRBC2	5236	MG29-1-mTRBC2-sgRNA-C2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArArUrCrUrCrUrGrCrUrUrUrUrGrArUrGrGrCrUrC/AltR2/
MG29-1 sgRNA targeting mouse TRBC2	5237	MG29-1-mTRBC2-sgRNA-D2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArUrGrCrUrCrArArArCrArArGrGrArGrArCrCrU/AltR2/
MG29-1 sgRNA targeting mouse TRBC2	5238	MG29-1-mTRBC2-sgRNA-E2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrGrGrCrUrCrArArArCrArArGrGrArGrArCrCrUrU/AltR2/
MG29-1 sgRNA targeting mouse TRBC2	5239	MG29-1-mTRBC2-sgRNA-F2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrCrCrCrUrCrUrCrCrUrUrUrCrUrUrUrUrCrArGrArC/AltR2/
MG29-1 sgRNA targeting mouse TRBC2	5240	MG29-1-mTRBC2-sgRNA-G2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrCrUrCrUrCrCrUrUrUrCrUrUrUrUrCrArGrArCrU/AltR2/
MG29-1 sgRNA targeting mouse TRBC2	5241	MG29-1-mTRBC2-sgRNA-H2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrGrCrArGrArCrUrGrUrGrGrArArUrCrArCrUrUrC/AltR2/
MG29-1 sgRNA targeting mouse TRBC2	5242	MG29-1-mTRBC2-sgRNA-A3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrArCrUrGrUrGrGrArArUrCrArCrUrUrCrArGrGrU/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting mouse TRBC2	5243	MG29-1-mTRBC2-sgRNA-B3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrCrArArCrArGrCrArUrCrCrUrArUrCrArUrC/AltR2/
MG29-1 sgRNA targeting mouse TRBC2	5244	MG29-1-mTRBC2-sgRNA-C3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrArUrCrCrUrArCrCrArUrUrCrUrUrArCrCrArU/AltR2/
MG29-1 sgRNA targeting mouse TRBC2	5245	MG29-1-mTRBC2-sgRNA-D3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrArGrGrUrCrArArGrArArArArArArArArUrUrCrC/AltR2/
MG29-1 sgRNA targeting mouse TRBC2	5246	MG29-1-mTRBC2-sgRNA-E3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrCrArGrGrArArUrUrUrUrUrUrUrUrCrUrUrGrA/AltR2/
DNA sequence of TRBC2 target site	5247	MG29-1-mTRBC2-target site-A1	CTTTCCAGAGGATCTGAGAAAT
DNA sequence of TRBC2 target site	5248	MG29-1-mTRBC2-target site-B1	CAGAGGATCTGAGAAATGTGAC
DNA sequence of TRBC2 target site	5249	MG29-1-mTRBC2-target site-C1	AGCCATCAAAAGCAGAGATTGC
DNA sequence of TRBC2 target site	5250	MG29-1-mTRBC2-target site-D1	AGAGGAGGACAAGTGGCCAGAG
DNA sequence of TRBC2 target site	5251	MG29-1-mTRBC2-target site-E1	CTCATGAGCTCCGCACTTACCT
DNA sequence of TRBC2 target site	5252	MG29-1-mTRBC2-target site-F1	GGTGAGCCCTCTGGCCACTTGT
DNA sequence of TRBC2 target site	5253	MG29-1-mTRBC2-target site-G1	GAGGATTGTGCCAGAAGGTAGC
DNA sequence of TRBC2 target site	5254	MG29-1-mTRBC2-target site-H1	GTTTGTTTGCAATCTCTGCTTT
DNA sequence of TRBC2 target site	5255	MG29-1-mTRBC2-target site-A2	TTTGTTTGCAATCTCTGCTTTT
DNA sequence of TRBC2 target site	5256	MG29-1-mTRBC2-target site-B2	TTTGCAATCTCTGCTTTTGATG
DNA sequence of TRBC2 target site	5257	MG29-1-mTRBC2-target site-C2	CAATCTCTGCTTTTGATGGCTC
DNA sequence of TRBC2 target site	5258	MG29-1-mTRBC2-target site-D2	GATGGCTCAAACAAGGAGACCT
DNA sequence of TRBC2 target site	5259	MG29-1-mTRBC2-target site-E2	ATGGCTCAAACAAGGAGACCTT
DNA sequence of TRBC2 target site	5260	MG29-1-mTRBC2-target site-F2	CTCCCTCTCCTTTCTTTCAGAC

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of TRBC2 target site	5261	MG29-1-mTRBC2-target site-G2	TCCCTCTCCTTTCTTTCAGACT
DNA sequence of TRBC2 target site	5262	MG29-1-mTRBC2-target site-H2	TTTCAGACTGTGGAATCACTTC
DNA sequence of TRBC2 target site	5263	MG29-1-mTRBC2-target site-A3	AGACTGTGGAATCACTTCAGGT
DNA sequence of TRBC2 target site	5264	MG29-1-mTRBC2-target site-B3	CTGTCAACAGCATCCTATCATC
DNA sequence of TRBC2 target site	5265	MG29-1-mTRBC2-target site-C3	TCCATCCTACCATTCTTACCAT
DNA sequence of TRBC2 target site	5266	MG29-1-mTRBC2-target site-D3	TCAGGTCAAGAAAAAAAATTCC
DNA sequence of TRBC2 target site	5267	MG29-1-mTRBC2-target site-E3	TCTCAGGAATTTTTTTCTTGA
r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *; between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur; AltR1 and AltR2 refer to IDT technologies' proprietary 5' and 3' AltR modifications			

Example 20 – Gene editing outcomes at the DNA level for human TRBC1/2

[00498] Primary T cells were purified from PMBCs using a negative selection kit (Miltenyi) according to the manufacturer's recommendations. Nucleofection of MG29-1 RNPs (106 pmol protein/160 pmol guide) (SEQ ID Nos: 5642-5660) was performed into T cells (200,000) using the Lonza 4D electroporator. For analysis by flow cytometry, 3 days post-nucleofection, 100,000 T cells were stained with anti-CD3 antibody for 30 minutes at 4C and analyzed on an Attune Nxt flow cytometer (**FIG. 66C**).

Table 14I: Sequences of Guide RNAs and Sequences Targeted for Example 46

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting human TRBC1/2	5642	MG29-1-TRBC1/2-sgRNA-A1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrCrCrArUrCrArGrArArGrCrArGrArGrArUrC/AltR2/
MG29-1 sgRNA targeting human TRBC1/2	5643	MG29-1-TRBC1/2-sgRNA-B1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrUrGrUrGrGrGrArGrArUrCrUrCrUrGrCrUrU/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting human TRBC1/2	5644	MG29-1-TRBC1/2-sgRNA-C1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrGrUrGrUrGrGrGrArGrArUrCrUrCrUrGrCrU/AltR2/
MG29-1 sgRNA targeting human TRBC1/2	5645	MG29-1-TRBC1/2-sgRNA-D1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrCrCrUrArUrCrCrUrGrGrGrUrCrCrArCrUrC/AltR2/
MG29-1 sgRNA targeting human TRBC1/2	5646	MG29-1-TRBC1/2-sgRNA-E1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrCrArGrArCrUrGrUrGrGrCrUrUrUrArCrC/AltR2/
MG29-1 sgRNA targeting human TRBC1/2	5647	MG29-1-TRBC1/2-sgRNA-F1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrArCrUrGrUrGrGrCrUrUrUrArCrCrUrCrGrG/AltR2/
MG29-1 sgRNA targeting human TRBC1/2	5648	MG29-1-TRBC1/2-sgRNA-G1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrUrCrUrGrCrArGrGrUrCrArArGrArGrArA/AltR2/
MG29-1 sgRNA targeting human TRBC1/2	5649	MG29-1-TRBC1/2-sgRNA-H1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrUrGrArCrCrUrGrCrArGrArArGrArGrArA/AltR2/
MG29-1 sgRNA targeting human TRBC1/2	5650	MG29-1-TRBC1/2-sgRNA-A2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrGrUrCrCrUrCrUrGrGrGrArArArGrGrGrArArG/AltR2/
MG29-1 sgRNA targeting human TRBC1/2	5651	MG29-1-TRBC1/2-sgRNA-B2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrGrUrCrCrUrCrUrGrGrGrArArArGrGrGrArA/AltR2/
MG29-1 sgRNA targeting human TRBC1/2	5652	MG29-1-TRBC1/2-sgRNA-C2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrArGrGrUrCrCrUrCrUrGrGrArArArGrGrGrA/AltR2/
MG29-1 sgRNA targeting human TRBC1/2	5653	MG29-1-TRBC1/2-sgRNA-D2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrCrCrArUrCrArGrArArGrCrArGrArArGrArUrC/AltR2/
MG29-1 sgRNA targeting human TRBC1/2	5654	MG29-1-TRBC1/2-sgRNA-E2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrUrGrUrGrGrGrArGrArUrCrUrCrUrGrCrUrU/AltR2/
MG29-1 sgRNA targeting	5655	MG29-1-TRBC1/2-sgRNA-F2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrGrUrGrUrGrGrGrArGrArUrCrUrCrUrGrCrU/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
human TRBC1/2			
MG29-1 sgRNA targeting human TRBC1/2	5656	MG29-1-TRBC1/2-sgRNA-G2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrCrCrUrArUrCrCrUrGrGrGrUrCrCrArCrUrC/AltR2/
MG29-1 sgRNA targeting human TRBC1/2	5657	MG29-1-TRBC1/2-sgRNA-H2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrUrCrArGrArCrUrGrUrGrGrCrUrUrCrArC/AltR2/
MG29-1 sgRNA targeting human TRBC1/2	5658	MG29-1-TRBC1/2-sgRNA-A3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrCrArGrArCrUrGrUrGrGrCrUrUrCrArCrC/AltR2/
MG29-1 sgRNA targeting human TRBC1/2	5659	MG29-1-TRBC1/2-sgRNA-B3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrArCrUrGrUrGrGrCrUrUrCrArCrCrUrCrG/AltR2/
MG29-1 sgRNA targeting human TRBC1/2	5660	MG29-1-TRBC1/2-sgRNA-C3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrUrGrArCrCrUrGrUrGrGrArArArGrArArG/AltR2/
DNA sequence of TRBC1/2 target site	5661	MG29-1-TRBC1/2-target site-A1	AGCCATCAGAAGCAGAGATC
DNA sequence of TRBC1/2 target site	5662	MG29-1-TRBC1/2-target site-B1	GGTGTGGGAGATCTCTGCTT
DNA sequence of TRBC1/2 target site	5663	MG29-1-TRBC1/2-target site-C1	GGGTGTGGGAGATCTCTGCT
DNA sequence of TRBC1/2 target site	5664	MG29-1-TRBC1/2-target site-D1	GCCCTATCCTGGGTCCACTC
DNA sequence of TRBC1/2 target site	5665	MG29-1-TRBC1/2-target site-E1	TTTCAGACTGTGGCTTTACC
DNA sequence of TRBC1/2 target site	5666	MG29-1-TRBC1/2-target site-F1	AGACTGTGGCTTTACCTCGG
DNA sequence of TRBC1/2 target site	5667	MG29-1-TRBC1/2-target site-G1	TCTTCTGCAGGTCAAGAGAA
DNA sequence of TRBC1/2 target site	5668	MG29-1-TRBC1/2-target site-H1	TCTTGACCTGCAGAAGAGAA
DNA sequence of TRBC1/2 target site	5669	MG29-1-TRBC1/2-target site-A2	AGGTCCTCTGGAAAGGGAAG
DNA sequence of TRBC1/2 target site	5670	MG29-1-TRBC1/2-target site-B2	CAGGTCCTCTGGAAAGGGAA

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of TRBC1/2 target site	5671	MG29-1-TRBC1/2-target site-C2	TCAGGTCCTCTGGAAAGGGA
DNA sequence of TRBC1/2 target site	5672	MG29-1-TRBC1/2-target site-D2	AGCCATCAGAAGCAGAGATC
DNA sequence of TRBC1/2 target site	5673	MG29-1-TRBC1/2-target site-E2	GGTGTGGGAGATCTCTGCTT
DNA sequence of TRBC1/2 target site	5674	MG29-1-TRBC1/2-target site-F2	GGGTGTGGGAGATCTCTGCT
DNA sequence of TRBC1/2 target site	5675	MG29-1-TRBC1/2-target site-G2	GCCCTATCCTGGGTCCACTC
DNA sequence of TRBC1/2 target site	5676	MG29-1-TRBC1/2-target site-H2	CTTTCAGACTGTGGCTTCAC
DNA sequence of TRBC1/2 target site	5677	MG29-1-TRBC1/2-target site-A3	TTTCAGACTGTGGCTTCACC
DNA sequence of TRBC1/2 target site	5678	MG29-1-TRBC1/2-target site-B3	AGACTGTGGCTTCACCTCCG
DNA sequence of TRBC1/2 target site	5679	MG29-1-TRBC1/2-target site-C3	TCTTGACCTGTGGAAGAGAG
r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *; between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur; AltR1 and AltR2 refer to IDT technologies' proprietary 5' and 3' AltR modifications			

Example 21 – Gene editing outcomes at the DNA level for HPRT

[00499] Primary T cells were purified from PMBCs using a negative selection kit (Miltenyi) according to the manufacturer's recommendations. Nucleofection of MG29-1 RNPs (126 pmol protein/160 pmol guide) (SEQ ID NOs: 5482-5561) was performed into T cells (200,000) using the Lonza 4D electroporator. Cells were harvested and genomic DNA prepared five days post-transfection. PCR primers appropriate for use in NGS-based DNA sequencing were generated, optimized, and used to amplify the individual target sequences for each guide RNA (SEQ ID NOs: 5562-5641). The amplicons were sequenced on an Illumina MiSeq machine and analyzed with a proprietary Python script to measure gene editing (**FIG. 67**).

Table 14J: Sequences of Guide RNAs and Sequences Targeted for Example 47

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA	5482	MG29-1-HPRT-sgRNA-A1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrCrCrCrUrGrGrCrUrUrArCrCrUrUrUrArGrGrA/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
targeting HPRT			
MG29-1 sgRNA targeting HPRT	5483	MG29-1-HPRT-sgRNA-B1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArArCrArCrArArGrCrCrCrArCrCrArUrUrArArArA/AltR2/
MG29-1 sgRNA targeting HPRT	5484	MG29-1-HPRT-sgRNA-C1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArUrGrGrUrGrGrGrCrUrUrGrUrGrUrUrCrUrArArA/AltR2/
MG29-1 sgRNA targeting HPRT	5485	MG29-1-HPRT-sgRNA-D1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrGrGrUrGrGrCrUrUrGrUrGrUrUrCrUrArArArG/AltR2/
MG29-1 sgRNA targeting HPRT	5486	MG29-1-HPRT-sgRNA-E1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrCrUrArArCrArArArGrArUrGrGrGrUrUrUrGrUrU/AltR2/
MG29-1 sgRNA targeting HPRT	5487	MG29-1-HPRT-sgRNA-F1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArGrGrCrArCrCrCrArArArUrUrArArArUrArArCrGrC/AltR2/
MG29-1 sgRNA targeting HPRT	5488	MG29-1-HPRT-sgRNA-G1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrCrGrUrUrArUrUrArArArUrUrUrGrGrGrUrGrCrC/AltR2/
MG29-1 sgRNA targeting HPRT	5489	MG29-1-HPRT-sgRNA-H1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrGrArGrUrGrUrArGrUrCrUrGrUrUrArGrCrCrArC/AltR2/
MG29-1 sgRNA targeting HPRT	5490	MG29-1-HPRT-sgRNA-A2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrArGrGrGrArGrUrArUrArUrGrArUrGrUrUrGrUrCrC/AltR2/
MG29-1 sgRNA targeting HPRT	5491	MG29-1-HPRT-sgRNA-B2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrGrUrGrCrArGrGrArUrCrArArUrGrArCrArGrC/AltR2/
MG29-1 sgRNA targeting HPRT	5492	MG29-1-HPRT-sgRNA-C2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrArGrGrUrGrCrArGrGrArArUrCrArArUrGrArCrArG/AltR2/
MG29-1 sgRNA targeting HPRT	5493	MG29-1-HPRT-sgRNA-D2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrCrArUrUrGrArUrCrCrUrGrCrArCrCrUrArC/AltR2/
MG29-1 sgRNA targeting HPRT	5494	MG29-1-HPRT-sgRNA-E2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrCrArCrArGrUrArArUrUrCrUrCrUrArCrCrArArA/AltR2/
MG29-1 sgRNA targeting HPRT	5495	MG29-1-HPRT-sgRNA-F2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrUrGrArGrArArUrUrArCrUrGrUrGrCrUrGrArArA/AltR2/
MG29-1 sgRNA targeting HPRT	5496	MG29-1-HPRT-sgRNA-G2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrUrGrArArUrArGrCrCrArUrGrGrCrArGrArGrGrA/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting HPRT	5497	MG29-1-HPRT-sgRNA-H2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArArGrGrGrGrCrCrCrArArArArUrCrCrUrCrUrG/AltR2/
MG29-1 sgRNA targeting HPRT	5498	MG29-1-HPRT-sgRNA-A3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrArArGrGrGrGrCrCrCrArArArArUrCrCrUrCrU/AltR2/
MG29-1 sgRNA targeting HPRT	5499	MG29-1-HPRT-sgRNA-B3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrGrCrCrCrCrCrUrUrGrCrArArArArUrUrArArGrA/AltR2/
MG29-1 sgRNA targeting HPRT	5500	MG29-1-HPRT-sgRNA-C3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrCrCrCrCrCrUrUrGrCrArArArArUrUrArArGrArA/AltR2/
MG29-1 sgRNA targeting HPRT	5501	MG29-1-HPRT-sgRNA-D3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrUrArCrArGrArCrArCrArArArGrArArGrArUrGrC/AltR2/
MG29-1 sgRNA targeting HPRT	5502	MG29-1-HPRT-sgRNA-E3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrArUrUrArArGrUrCrGrGrCrCrUrCrArCrCrUrCrC/AltR2/
MG29-1 sgRNA targeting HPRT	5503	MG29-1-HPRT-sgRNA-F3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArUrArArGrUrCrGrGrCrCrUrCrArCrCrUrCrU/AltR2/
MG29-1 sgRNA targeting HPRT	5504	MG29-1-HPRT-sgRNA-G3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrUrArArGrUrCrGrGrCrCrUrCrArCrCrUrCrCrU/AltR2/
MG29-1 sgRNA targeting HPRT	5505	MG29-1-HPRT-sgRNA-H3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrArArGrUrCrGrGrCrCrUrCrArCrCrUrCrCrU/AltR2/
MG29-1 sgRNA targeting HPRT	5506	MG29-1-HPRT-sgRNA-A4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArUrGrUrArGrGrGrGrUrCrArGrGrUrArArUrGrUrU/AltR2/
MG29-1 sgRNA targeting HPRT	5507	MG29-1-HPRT-sgRNA-B4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrGrUrArGrGrGrGrUrCrArGrGrUrArArUrGrUrUrC/AltR2/
MG29-1 sgRNA targeting HPRT	5508	MG29-1-HPRT-sgRNA-C4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrUrArGrGrGrGrUrCrArGrGrUrArArUrGrUrUrCrU/AltR2/
MG29-1 sgRNA targeting HPRT	5509	MG29-1-HPRT-sgRNA-D4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArArArArArArUrCrArCrGrGrUrArUrCrUrGrUrCrG/AltR2/
MG29-1 sgRNA targeting HPRT	5510	MG29-1-HPRT-sgRNA-E4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArUrCrArGrArGrUrArArGrCrCrUrUrCrUrArGrUrG/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting HPRT	5511	MG29-1-HPRT-sgRNA-F4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrArCrArGrGrUrArCrUrArUrGrArGrArGrUrArUrA/AltR2/
MG29-1 sgRNA targeting HPRT	5512	MG29-1-HPRT-sgRNA-G4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArArUrGrUrCrArArCrCrUrArCrUrGrUrGrGrCrArU/AltR2/
MG29-1 sgRNA targeting HPRT	5513	MG29-1-HPRT-sgRNA-H4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArUrGrUrCrArArCrCrUrArCrUrGrUrGrGrCrArUrA/AltR2/
MG29-1 sgRNA targeting HPRT	5514	MG29-1-HPRT-sgRNA-A5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrUrGrUrUrCrUrUrUrCrCrUrGrGrUrArUrArUrGrC/AltR2/
MG29-1 sgRNA targeting HPRT	5515	MG29-1-HPRT-sgRNA-B5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrGrUrArUrArUrGrCrUrGrUrGrGrArArUrUrGrA/AltR2/
MG29-1 sgRNA targeting HPRT	5516	MG29-1-HPRT-sgRNA-C5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrCrArUrGrUrCrCrUrArCrCrUrGrUrGrGrCrArArC/AltR2/
MG29-1 sgRNA targeting HPRT	5517	MG29-1-HPRT-sgRNA-D5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrGrUrGrCrCrArCrArGrGrUrArGrGrArCrArUrG/AltR2/
MG29-1 sgRNA targeting HPRT	5518	MG29-1-HPRT-sgRNA-E5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrUrUrGrCrCrArCrArGrGrUrArGrGrArCrArUrGrC/AltR2/
MG29-1 sgRNA targeting HPRT	5519	MG29-1-HPRT-sgRNA-F5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArUrArCrUrGrUrArArArUrGrGrUrArArCrCrGrU/AltR2/
MG29-1 sgRNA targeting HPRT	5520	MG29-1-HPRT-sgRNA-G5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrArUrArCrUrGrUrArArArUrGrGrUrArArCrCrG/AltR2/
MG29-1 sgRNA targeting HPRT	5521	MG29-1-HPRT-sgRNA-H5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrArArArGrCrUrCrCrArUrCrUrCrUrArArGrGrCrA/AltR2/
MG29-1 sgRNA targeting HPRT	5522	MG29-1-HPRT-sgRNA-A6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrGrArUrArCrCrUrUrUrUrCrUrGrGrArGrCrArUrU/AltR2/
MG29-1 sgRNA targeting HPRT	5523	MG29-1-HPRT-sgRNA-B6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrGrArGrCrArUrUrCrCrUrGrArGrUrUrCrArGrG/AltR2/
MG29-1 sgRNA targeting HPRT	5524	MG29-1-HPRT-sgRNA-C6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrGrArGrCrArUrUrCrCrUrGrArGrUrUrCrArGrGrU/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting HPRT	5525	MG29-1-HPRT-sgRNA-D6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrCrUrGrUrGrArUrUrGrGrCrUrUrGrUrUrArUrGrU/AltR2/
MG29-1 sgRNA targeting HPRT	5526	MG29-1-HPRT-sgRNA-E6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrUrGrUrGrArUrUrGrGrCrUrUrGrUrUrArUrGrUrU/AltR2/
MG29-1 sgRNA targeting HPRT	5527	MG29-1-HPRT-sgRNA-F6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrGrArUrUrGrGrCrUrUrGrUrUrArUrGrUrUrC/AltR2/
MG29-1 sgRNA targeting HPRT	5528	MG29-1-HPRT-sgRNA-G6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrGrArArGrArGrUrCrArUrGrArGrGrGrArCrArU/AltR2/
MG29-1 sgRNA targeting HPRT	5529	MG29-1-HPRT-sgRNA-H6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrArUrGrUrUrArArArGrGrCrArGrUrCrUrCrArA/AltR2/
MG29-1 sgRNA targeting HPRT	5530	MG29-1-HPRT-sgRNA-A7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrGrArGrArCrUrGrCrCrUrUrArArCrArUrCrU/AltR2/
MG29-1 sgRNA targeting HPRT	5531	MG29-1-HPRT-sgRNA-B7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrArGrArCrUrGrCrCrUrUrArArCrArUrCrUrG/AltR2/
MG29-1 sgRNA targeting HPRT	5532	MG29-1-HPRT-sgRNA-C7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArArCrCrCrArArArUrGrCrUrGrCrCrUrGrUrUrGrA/AltR2/
MG29-1 sgRNA targeting HPRT	5533	MG29-1-HPRT-sgRNA-D7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrArArCrCrCrArArArUrGrCrUrGrCrCrUrGrUrUrG/AltR2/
MG29-1 sgRNA targeting HPRT	5534	MG29-1-HPRT-sgRNA-E7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrArArCrArGrGrCrArGrCrArUrUrUrGrGrGrUrUrA/AltR2/
MG29-1 sgRNA targeting HPRT	5535	MG29-1-HPRT-sgRNA-F7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArArCrArGrGrCrArGrCrArUrUrUrGrGrGrUrUrArU/AltR2/
MG29-1 sgRNA targeting HPRT	5536	MG29-1-HPRT-sgRNA-G7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArCrArGrGrCrArGrCrArUrUrUrGrGrGrUrUrArUrA/AltR2/
MG29-1 sgRNA targeting HPRT	5537	MG29-1-HPRT-sgRNA-H7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArArCrArArArGrCrUrGrGrArGrGrUrGrGrUrArU/AltR2/
MG29-1 sgRNA targeting HPRT	5538	MG29-1-HPRT-sgRNA-A8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrUrArGrArGrUrUrGrArCrUrUrArUrArCrCrArCrC/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting HPRT	5539	MG29-1-HPRT-sgRNA-B8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrArGrArGrUrUrGrArCrUrUrArUrArCrCrArCrCrU/AltR2/
MG29-1 sgRNA targeting HPRT	5540	MG29-1-HPRT-sgRNA-C8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrArArCrArArArArGrCrUrGrGrArGrGrUrGrGrUrA/AltR2/
MG29-1 sgRNA targeting HPRT	5541	MG29-1-HPRT-sgRNA-D8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrGrArArCrArArArArGrCrUrGrGrArGrGrUrGrGrU/AltR2/
MG29-1 sgRNA targeting HPRT	5542	MG29-1-HPRT-sgRNA-E8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrGrGrArArCrArArArArGrCrUrGrGrArGrG/AltR2/
MG29-1 sgRNA targeting HPRT	5543	MG29-1-HPRT-sgRNA-F8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrGrUrArArArArCrArArCrUrArGrUrGrUrGrCrCrA/AltR2/
MG29-1 sgRNA targeting HPRT	5544	MG29-1-HPRT-sgRNA-G8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrCrArCrArCrUrArGrUrUrGrUrUrUrUrArCrCrCrU/AltR2/
MG29-1 sgRNA targeting HPRT	5545	MG29-1-HPRT-sgRNA-H8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrArCrArCrUrArGrUrUrGrUrUrUrUrArCrCrCrUrA/AltR2/
MG29-1 sgRNA targeting HPRT	5546	MG29-1-HPRT-sgRNA-A9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrCrUrArArArGrUrUrCrUrUrUrUrGrUrArArArG/AltR2/
MG29-1 sgRNA targeting HPRT	5547	MG29-1-HPRT-sgRNA-B9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrGrArUrUrArUrUrUrCrCrArArGrGrUrUrUrC/AltR2/
MG29-1 sgRNA targeting HPRT	5548	MG29-1-HPRT-sgRNA-C9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArArGrGrUrUrUrCrUrArGrArCrUrGrArGrArGrCrC/AltR2/
MG29-1 sgRNA targeting HPRT	5549	MG29-1-HPRT-sgRNA-D9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrArGrArGrCrCrCrUrUrUrUrUrCrArUrC/AltR2/
MG29-1 sgRNA targeting HPRT	5550	MG29-1-HPRT-sgRNA-E9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArUrCrUrUrUrGrCrUrCrArUrUrGrArCrArCrUrCrU/AltR2/
MG29-1 sgRNA targeting HPRT	5551	MG29-1-HPRT-sgRNA-F9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrCrUrUrUrGrCrUrCrArUrUrGrArCrArCrUrCrUrG/AltR2/
MG29-1 sgRNA targeting HPRT	5552	MG29-1-HPRT-sgRNA-G9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrCrArUrUrGrArCrArCrUrCrUrGrUrArCrCrCrArU/AltR2/

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting HPRT	5553	MG29-1-HPRT-sgRNA-H9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrArCrArCrCrCrArArGrGrArArArGrArCrUrArUrG/AltR2/
MG29-1 sgRNA targeting HPRT	5554	MG29-1-HPRT-sgRNA-A10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArCrArCrArCrCrCrArArGrGrArArArGrArCrUrArU/AltR2/
MG29-1 sgRNA targeting HPRT	5555	MG29-1-HPRT-sgRNA-B10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrUrCrUrCrCrArUrUrUrCrArUrArGrUrCrUrUrUrC/AltR2/
MG29-1 sgRNA targeting HPRT	5556	MG29-1-HPRT-sgRNA-C10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrArGrUrCrUrUrUrCrCrUrUrGrGrGrUrGrUrU/AltR2/
MG29-1 sgRNA targeting HPRT	5557	MG29-1-HPRT-sgRNA-D10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrGrGrGrUrGrUrUrArArArArGrUrGrArCrC/AltR2/
MG29-1 sgRNA targeting HPRT	5558	MG29-1-HPRT-sgRNA-E10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrCrCrGrUrGrCrUrGrArGrUrGrUrArCrCrArUrGrG/AltR2/
MG29-1 sgRNA targeting HPRT	5559	MG29-1-HPRT-sgRNA-F10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrUrUrCrArUrCrCrGrUrGrCrUrGrArGrUrGrUrArC/AltR2/
MG29-1 sgRNA targeting HPRT	5560	MG29-1-HPRT-sgRNA-G10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArArArCrGrUrCrArGrUrCrUrUrUrCrUrCrUrUrUrG/AltR2/
MG29-1 sgRNA targeting HPRT	5561	MG29-1-HPRT-sgRNA-H10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrArGrArGrCrArCrArUrUrUrGrCrCrArGrUrA/AltR2/
DNA sequence of HPRT target site	5562	MG29-1-HPRT-target site-A1	TCTCCCTGGCTTACCTTTAGGA
DNA sequence of HPRT target site	5563	MG29-1-HPRT-target site-B1	GAACACAAGCCCACCATTAAAA
DNA sequence of HPRT target site	5564	MG29-1-HPRT-target site-C1	AATGGTGGGCTTGTGTTCTAAA
DNA sequence of HPRT target site	5565	MG29-1-HPRT-target site-D1	ATGGTGGGCTTGTGTTCTAAAG
DNA sequence of HPRT target site	5566	MG29-1-HPRT-target site-E1	CCCTAACAAAGATGGGTTTGTT
DNA sequence of HPRT target site	5567	MG29-1-HPRT-target site-F1	AAGGCACCCAAATTAATAACGC
DNA sequence of HPRT target site	5568	MG29-1-HPRT-target site-G1	TAGCGTTATTAATTTGGGTGCC

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of HPRT target site	5569	MG29-1-HPRT-target site-H1	TCGAGTGTAGTCTGTTAGCCAC
DNA sequence of HPRT target site	5570	MG29-1-HPRT-target site-A2	TTAGGGAGTTATGATGTTGTCC
DNA sequence of HPRT target site	5571	MG29-1-HPRT-target site-B2	TAGGTGCAGGATCAATGACAGC
DNA sequence of HPRT target site	5572	MG29-1-HPRT-target site-C2	GTAGGTGCAGGATCAATGACAG
DNA sequence of HPRT target site	5573	MG29-1-HPRT-target site-D2	CTGTCATTGATCCTGCACCTAC
DNA sequence of HPRT target site	5574	MG29-1-HPRT-target site-E2	CAGCACAGTAATTCTCACCAAA
DNA sequence of HPRT target site	5575	MG29-1-HPRT-target site-F2	GGTGAGAATTACTGTGCTGAAA
DNA sequence of HPRT target site	5576	MG29-1-HPRT-target site-G2	TCCTGAATAGCATGGCAGAGGA
DNA sequence of HPRT target site	5577	MG29-1-HPRT-target site-H2	CAAGGGGGCCCAAAATCCTCTG
DNA sequence of HPRT target site	5578	MG29-1-HPRT-target site-A3	GCAAGGGGGCCCAAAATCCTCT
DNA sequence of HPRT target site	5579	MG29-1-HPRT-target site-B3	GGGCCCCCTTGCAAAATTAAGA
DNA sequence of HPRT target site	5580	MG29-1-HPRT-target site-C3	GGCCCCCTTGCAAAATTAAGAA
DNA sequence of HPRT target site	5581	MG29-1-HPRT-target site-D3	ACTACAGACACAAAGAAGATGC
DNA sequence of HPRT target site	5582	MG29-1-HPRT-target site-E3	TTATTAAGTCGGCCTCACCTCC
DNA sequence of HPRT target site	5583	MG29-1-HPRT-target site-F3	TATTAAGTCGGCCTCACCTCCT
DNA sequence of HPRT target site	5584	MG29-1-HPRT-target site-G3	ATTAAGTCGGCCTCACCTCCTC
DNA sequence of HPRT target site	5585	MG29-1-HPRT-target site-H3	TTAAGTCGGCCTCACCTCCTCA
DNA sequence of HPRT target site	5586	MG29-1-HPRT-target site-A4	TATGTAGGGGTCAGGTAATGTT
DNA sequence of HPRT target site	5587	MG29-1-HPRT-target site-B4	ATGTAGGGGTCAGGTAATGTTC

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of HPRT target site	5588	MG29-1-HPRT-target site-C4	TGTAGGGGTCAGGTAATGTTCT
DNA sequence of HPRT target site	5589	MG29-1-HPRT-target site-D4	GAAAAAATCACGGTATCTGTCG
DNA sequence of HPRT target site	5590	MG29-1-HPRT-target site-E4	AATCAGAGTAAGCCTTCTAGTG
DNA sequence of HPRT target site	5591	MG29-1-HPRT-target site-F4	GGACAGGTACTATGAGAGTATA
DNA sequence of HPRT target site	5592	MG29-1-HPRT-target site-G4	AAATGTCAACCTACTGTGGCAT
DNA sequence of HPRT target site	5593	MG29-1-HPRT-target site-H4	AATGTCAACCTACTGTGGCATA
DNA sequence of HPRT target site	5594	MG29-1-HPRT-target site-A5	GTTGTTCTTTCCTGGTATATGC
DNA sequence of HPRT target site	5595	MG29-1-HPRT-target site-B5	CTGGTATATGCTGTGGAATTGA
DNA sequence of HPRT target site	5596	MG29-1-HPRT-target site-C5	AGCATGTCCTACCTGTGGCAAC
DNA sequence of HPRT target site	5597	MG29-1-HPRT-target site-D5	GTGTTGCCACAGGTAGGACATG
DNA sequence of HPRT target site	5598	MG29-1-HPRT-target site-E5	TGTTGCCACAGGTAGGACATGC
DNA sequence of HPRT target site	5599	MG29-1-HPRT-target site-F5	CATACTGTAAATGGGTAACCGT
DNA sequence of HPRT target site	5600	MG29-1-HPRT-target site-G5	CCATACTGTAAATGGGTAACCG
DNA sequence of HPRT target site	5601	MG29-1-HPRT-target site-H5	ATAAAGCTCCATCTCTAAGGCA
DNA sequence of HPRT target site	5602	MG29-1-HPRT-target site-A6	GTGATACCTTTTCTGGAGCATT
DNA sequence of HPRT target site	5603	MG29-1-HPRT-target site-B6	CTGGAGCATTCCTGAGTTCAGG
DNA sequence of HPRT target site	5604	MG29-1-HPRT-target site-C6	TGGAGCATTCCTGAGTTCAGGT
DNA sequence of HPRT target site	5605	MG29-1-HPRT-target site-D6	TGCTGTGATTGGCTTGTTATGT
DNA sequence of HPRT target site	5606	MG29-1-HPRT-target site-E6	GCTGTGATTGGCTTGTTATGTT

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of HPRT target site	5607	MG29-1-HPRT-target site-F6	CTGTGATTGGCTTGTTATGTTC
DNA sequence of HPRT target site	5608	MG29-1-HPRT-target site-G6	TTGGAAGAGTCATGAGGGACAT
DNA sequence of HPRT target site	5609	MG29-1-HPRT-target site-H6	CAGATGTTAAAGGCAGTCTCAA
DNA sequence of HPRT target site	5610	MG29-1-HPRT-target site-A7	CTTGAGACTGCCTTTAACATCT
DNA sequence of HPRT target site	5611	MG29-1-HPRT-target site-B7	TTGAGACTGCCTTTAACATCTG
DNA sequence of HPRT target site	5612	MG29-1-HPRT-target site-C7	TAACCCAAATGCTGCCTGTTGA
DNA sequence of HPRT target site	5613	MG29-1-HPRT-target site-D7	ATAACCCAAATGCTGCCTGTTG
DNA sequence of HPRT target site	5614	MG29-1-HPRT-target site-E7	TCAACAGGCAGCATTTGGGTTA
DNA sequence of HPRT target site	5615	MG29-1-HPRT-target site-F7	CAACAGGCAGCATTTGGGTTAT
DNA sequence of HPRT target site	5616	MG29-1-HPRT-target site-G7	AACAGGCAGCATTTGGGTTATA
DNA sequence of HPRT target site	5617	MG29-1-HPRT-target site-H7	GAACAAAAGCTGGAGGTGGTAT
DNA sequence of HPRT target site	5618	MG29-1-HPRT-target site-A8	GGTAGAGTTGACTTATACCACC
DNA sequence of HPRT target site	5619	MG29-1-HPRT-target site-B8	GTAGAGTTGACTTATACCACCT
DNA sequence of HPRT target site	5620	MG29-1-HPRT-target site-C8	GGAACAAAAGCTGGAGGTGGTA
DNA sequence of HPRT target site	5621	MG29-1-HPRT-target site-D8	TGGAACAAAAGCTGGAGGTGGT
DNA sequence of HPRT target site	5622	MG29-1-HPRT-target site-E8	TTTTTGGAACAAAAGCTGGAGG
DNA sequence of HPRT target site	5623	MG29-1-HPRT-target site-F8	GGGTAAAACAACTAGTGTGCCA
DNA sequence of HPRT target site	5624	MG29-1-HPRT-target site-G8	GGCACACTAGTTGTTTTACCCT
DNA sequence of HPRT target site	5625	MG29-1-HPRT-target site-H8	GCACACTAGTTGTTTTACCCTA

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
DNA sequence of HPRT target site	5626	MG29-1-HPRT-target site-A9	CCCTAAAGTTCCTCTTTGTAAG
DNA sequence of HPRT target site	5627	MG29-1-HPRT-target site-B9	GGGATTGTATTTCCAAGGTTTC
DNA sequence of HPRT target site	5628	MG29-1-HPRT-target site-C9	CAAGGTTTCTAGACTGAGAGCC
DNA sequence of HPRT target site	5629	MG29-1-HPRT-target site-D9	TAGACTGAGAGCCCTTTTCATC
DNA sequence of HPRT target site	5630	MG29-1-HPRT-target site-E9	CATCTTTGCTCATTGACACTCT
DNA sequence of HPRT target site	5631	MG29-1-HPRT-target site-F9	ATCTTTGCTCATTGACACTCTG
DNA sequence of HPRT target site	5632	MG29-1-HPRT-target site-G9	CTCATTGACACTCTGTACCCAT
DNA sequence of HPRT target site	5633	MG29-1-HPRT-target site-H9	ACACACCCAAGGAAAGACTATG
DNA sequence of HPRT target site	5634	MG29-1-HPRT-target site-A10	AACACACCCAAGGAAAGACTAT
DNA sequence of HPRT target site	5635	MG29-1-HPRT-target site-B10	GCTCTCCATTTCATAGTCTTTC
DNA sequence of HPRT target site	5636	MG29-1-HPRT-target site-C10	ATAGTCTTTCCTTGGGTGTGTT
DNA sequence of HPRT target site	5637	MG29-1-HPRT-target site-D10	CTTGGGTGTGTTAAAAGTGACC
DNA sequence of HPRT target site	5638	MG29-1-HPRT-target site-E10	ATCCGTGCTGAGTGTACCATGG
DNA sequence of HPRT target site	5639	MG29-1-HPRT-target site-F10	ATTTCATCCGTGCTGAGTGTAC
DNA sequence of HPRT target site	5640	MG29-1-HPRT-target site-G10	GAAACGTCAGTCTTCTCTTTTG
DNA sequence of HPRT target site	5641	MG29-1-HPRT-target site-H10	TAGAGAGGCACATTTGCCAGTA
r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *; between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur; AltR1 and AltR2 refer to IDT technologies' proprietary 5' and 3' AltR modifications			

Example 22 – Additional MG29-1 guide chemistry optimization

[00500] The editing activity of 5 guides with the same base sequence but different chemical modifications was evaluated in Hepa1-6 cells by co-transfection of mRNA encoding MG29-1 and the guide; the results are shown in Table 15 and **FIG. 68**. A guide with the same base sequence and a commercially available chemical modification called AltR1/AltR2 was used as a control. The spacer sequence in these guides targets a 22 nucleotide region in albumin intron1 of the mouse genome. Guide mAlb298-44 exhibited 67.5 % of the editing activity of the control AltR1/AltR2 guide while the other 4 guides did not result in measurable editing. When co-transfection of mRNA and guide with a lipid transfection reagent such as Messenger MAX is used, the mixture of the two RNA forms a complex with the positively charged lipid and the complex enters the cells via endocytosis and eventually reaches the cytoplasm, where the mRNA is translated into protein. In the case of an RNA guided nuclease such as MG29-1, the resulting MG29-1 protein presumably forms a complex with the guide RNA in the cytoplasm before entering the nucleus in a process mediated by the nuclear localization signals that were engineered into the MG29-1 protein. Because translation of the mRNA into sufficient amounts of MG29-1 protein followed by the binding of the MG29-1 protein to the guide RNA takes a finite amount of time, the guide RNA may require increased stability in the cytoplasm for longer than is the case when pre-formed RNP is delivered by nucleofection. Thus, lipid-based mRNA/sgRNA co-transfection may require a more stable guide than is the case for RNP nucleofection, which may result in some guide chemistries being active as RNP but inactive when co-transfected with mRNA using cationic lipid reagents.

Table 15: Activity of chemically modified MG29-1 guides in Hepa1-6 cells transfected with MG29-1 mRNA and the guide RNA

sgRNA name	SEQ ID No.	sgRNA sequence and chemical modifications	Editing activity (% of AltR1/AltR2 control)
mAlb298-40	5745	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUi2FCi2FUi2FGi2FUi2FAi2FAi2FCi2FGi2FAi2FUi2FCi2FGi2FGi2FGi2FAi2FAi2FC*i2FU*i2FGi2FG*i2FC*mA	0
mAlb298-41	5746	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*dGdTdAdGdAdTi2FCi2FUi2FGi2FUi2FAi2FAi2FCi2FGi2FAi2FUi2FCi2FGi2FGi2FGi2FAi2FAi2FC*i2FU*i2FGi2FG*i2FC*mA	0

sgRNA name	SEQ ID No.	sgRNA sequence and chemical modifications	Editing activity (% of AltR1/AltR2 control)
mAlb298-42	5747	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*i2FGi2FUi2FAi2FGi2FAi2FUi2FCi2FUi2FGi2FUi2FAi2FAi2FCi2FGi2FAi2FUi2FCi2FGi2FGi2FGi2FAi2FAi2FC*i2FU*i2FGi2FG*i2FC*mA	0
mAlb298-43	5748	mC*mU*mU*U*UAAUUmUmCmUmAmCU*G*U*U*GUAGAUi2FCi2FUi2FGi2FUi2FAi2FAi2FCi2FGi2FAi2FUi2FCi2FGi2FGi2FGi2FAi2FAi2FC*i2FU*i2FGi2FG*i2FC*mA	0
mAlb298-44	5749	mC*mU*mU*U*rU*rArArUrUmUmCmUmArCrU*rG*rU*rU*rGrUrArGrArUrCrUrGrUrArArCi2FGi2FAi2FUi2FCi2FGi2FGi2FGi2FA*i2FAi2FC*i2FU*i2FG*i2FG*i2FC*mA	67.5
Nomenclature of chemical modifications: a "/" is used to separate bases with 2'-flourine modifications, m; 2'-O-methyl base (for example a A base with 2'-O-methyl modification is written as mA), i2F; internal 2'-flourine base (for example an internal C with 2'-flourine modification is written as /i2FC/), 52F; 2'-flourine base at the 5' end of the sequence (for example a 5' C with 2'-flourine modification is written as /52FC/), 32F; 2'-flourine base at the 3' end of the sequence (for example a 3' A base with 2'-flourine modification is written as /32FA/), r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *; between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur; AltR1 and AltR2 refer to IDT technologies' proprietary 5' and 3' AltR modifications			

[00501] In order to test the stability of these chemically modified guides compared to the guide with no chemical modification (native RNA), a stability assay using crude cell extracts was used. Crude cell extracts from mammalian cells were selected because they contain the mixture of nucleases that a guide RNA will be exposed to when delivered to mammalian cells *in vitro* or *in vivo*. Hepa1-6 cells were collected by adding 3ml of cold PBS per 15cm dish of confluent cells and releasing the cells from the surface of the dish using a cell scraper. The cells were pelleted at 200g for 10min and frozen at -80°C for future use. For the stability assays, cells were resuspended in 4 volumes of cold PBS (i.e. for a 100mg pellet cells were resuspended in 400ul of cold PBS). Triton X-100 was added to a ending concentration of 0.2% (v/v), cells were vortexed for 10 seconds, put on ice for 10 minutes and vortexed again for 10 seconds. Triton X-100 is a mild non-ionic detergent that disrupts cell membranes but does not inactivate or denature proteins at the concentration used. Stability reactions were set up on ice and comprised 20 ul of cell crude extract with 2 pmoles of each guide (1ul of a 2uM stock). Six reactions were set up per

guide comprising: input, 0.5hour, 1hour, 4hours, 9hours and 21hours (the time in hours referring to the length of time each sample was incubated). Samples were incubated at 37°C from 0.5hours up to 21hours while the input control was left on ice for 5 minutes. After each incubation period, the reaction was stopped by adding 300ul of a mixture of phenol and guanidine thiocyanate (Tri reagent, Zymo Research) which immediately denatures all proteins and efficiently inhibits ribonucleases and facilitates the subsequent recovery of RNA. After adding Tri Reagent the samples were vortexed for 15seconds and stored at -20°C. RNA was extracted from the samples using Direct-zol RNA miniprep kit (Zymo Research) and eluted in 100ul of nuclease-free water. Detection of the modified guide was performed using Taqman RT - qPCR using the Taqman miRNA Assay technology (Thermo Fisher) and primers and probes designed to specifically detect the sequence in the mAlb298 sgRNA, which is the same for all of the guides. Data was plotted as a function of percentage of sgRNA remaining in relation to the input sample (Table 16 and **FIG. 69**).

Table 16: Stability of Chemically Modified MG29-1 Guides

	Percentage of guide left		
Time (H)	**mAlb298 unmodified**	**mAlb298-AltR**	**mAlb298-44**
0.5	**31.8640157**	**46.1691**	**89.5025**
1	**19.6826994**	**18.8809**	**68.7771**
4	**3.80753207**	**7.99366**	**55.0953**
9	**1.67464604**	**2.14928**	**50**
21	**1.17190546**	**0.4044**	**44.1351**

[00502] Guide mAlb289-44 exhibited significantly improved stability in the cell lysate compared to both un-modified guide and the guide with AltR1/AltR2 modifications. Thus, the chemical modifications present in the mAlb289-44 guide may be useful for optimizing editing *in vivo*. The chemical modifications present on the mAlb289-44 guide are detailed in Table 15. The mAlb289-44 guide chemistry differs from another highly stable guide chemistry called the mAlb289-37 by the presence of 3 additional phosphorothioate linkages

Example 23 – Improving the stability of the guide RNA for MG29-1 by addition of a stem-loop at the 5' end

[00503] A comparison of the stability in cell lysates *in vitro* of the guide RNA for MG29-1 to that of the guide RNA for a type II CRISPR nuclease called MG3-6/3-4 shows that the MG29-1 guide is inherently less stable (Table 17 and **FIGs. 70A-B**).

Table 17: Stability of Type II Versus Type V Guides

	Percentage of guide left			
	Unmodified		Only 5' and 3' End Modifications	
Time (H)	Type II	Type V	Type II	Type V
0.5	68.30201	31.86402	61.98539	46.16912
1	51.05061	19.6827	59.66679	18.88091
4	9.672281	3.807532	51.05061	7.99366
9	1.757904	1.674646	40.47211	2.149284
21	0.034051	1.171905	1.447794	0.4044

[00504] As shown in **FIG. 70A**, when comparing guide RNA without chemical modifications, the Type V guide (MG29-1 guide) was degraded more rapidly than the Type II guide (MG3-6/3-4 guide). As shown in **FIG. 70B**, when comparing guide RNA with chemical modifications of both ends of the RNA, the difference in stability between the Type V guide (MG29-1 guide) and the Type II guide (MG3-6/3-4 guide) was even more pronounced. The end-modified type V guide was almost completely degraded in 10 h while 40% of the Type II guide remained at the same time points. The secondary structures of the backbone (CRISPR repeat and tracr) of the MG29-1 (Type V) guide and the backbone of the MG3-6/3-4 (Type II) guide were predicted using the folding algorithm in Geneious Prime and are shown in **FIG. 71**. The backbone of the MG29-1 guide is 24 nucleotides long while that of MG3-6/3-4 is 88 nucleotides long. The backbone (CRISPR repeat) of the MG29-1 guide is predicted to form a single stem loop with a stem comprised of 5 nucleotides and a free energy of -1.22 kcal/mol while the backbone (CRISPR repeat and tracr) of MG3-6/3-4 is predicted to form 3 stem loops with a free energy of -14.8 kcal/mol. The three stem-loops of the MG3-6/3-4 guide RNA are comprised of stem 1 (at the 5' end of the TRACR) that has a 10 nucleotide stem, stem 2 (in the middle) that has a 5 nucleotide stem, and stem 3 (at the 3' end of the backbone) that has a 11 nucleotide stem. The larger size and more extensive stem structures in the MG3-6/3-4 guide RNA may contribute to the greater stability of this guide compared to the MG29-1 guide. The addition of a stem-loop forming RNA sequence at the 5' end of the MG29-1 guide may improve its stability and thereby potentially improve the efficiency of editing *in vivo*. In one embodiment, stem 1 of the MG3-6/3-

4 guide comprises the sequence (GUUGAGAAUCGAAAGAUUCUUAAU), wherein the underlined bases are predicted to form non-canonical G-U base pairs. To improve the stability of the stem, the underlined bases were changed from U to C to convert these to G-C base pairs in the predicted stem (GUUGAGAAUCGAAAGAUUCUCAAC). In one embodiment, this stem-loop forming sequence is added at the 5' end of the MG29-1 guide RNA with chemistry #37. In addition, the chemical modifications on the 5'-most 4 nucleotides of the #37 chemistry (comprised of 2' O-methyl and phosphorothioate linkages) were replicated at the new 5' end of the guide in order to protect the 5' end of the guide from nuclease attack. This gave rise to the guide RNA sequence called mAlb29-8-50 (Table 18). In an alternative guide design, the chemically modified bases at the original 5' end of the mAlb298-37 were moved to the new 5' end of the guide after the addition of the stem-loop sequence as in mAlb29-8-49. In another design for a potentially more stable MG29-1 guide, the RNA sequence from the MG3-6/3-4 backbone that encompasses stem-loop 1 and stem-loop 2 was added to the 5' end of the MG29-1 guide to create mAlb29-8-48 and mAlb29-8-47, which differ in the chemically modified bases included. In another version, guide mAlb29-8-48 is further chemically modified by inclusion of phosphorothioate and 2' O-methyl bases in the loop 1 (mALb29-8-47) or loop 1 and loop 2 ((mALb29-8-46). The activity of these guides can be tested in mammalian cells by transfection of mRNA and guide RNA mixtures using MessengerMax lipid reagent or other methodologies. The stability of these guides can be tested in the same mammalian cell lysate assay system as described above. Guides that retain editing activity and exhibit improvements in stability are candidates for testing *in vivo* in mice.

Table 18: Activity of chemically modified MG29-1 guides in Hepa1-6 cells transfected with MG29-1 mRNA and the guide RNA

sgRNA name	SEQ ID NO.	sgRNA sequence and chemical modifications
mAlb298-37	5750	mC*mU*mU*U*rUrArArUrUmUmCmUmArCrU*rG*rU*rU*rGrUrArGrArUrCrUrGrUrArArCi2FGi2FAi2FUi2FCi2FGi2FGi2FGi2FAi2FAi2FC*i2FU*i2FGi2FG*i2FC*mA
mAlb29-8-50	5751	mG*mU*mU*GrArGrArArUrC*mG*mA*mA*mArGrArUrUrCrUrCrArArC*mC*mU*mU*U*UrArArUrUmUmCmUmArCrU*G*U*U*GrUrArGrGrArUrCrUrGrUrArArCfGfAfUfCfGfGfGfAfAfC*fU*fGfG*fC*mA
mAlb29-8-49	5752	mG*mU*mU*GrArGrArArUrCrGrArArArGrArUrUrCrUrCrArArC*mC*mU*mU*U*UrArArUrUmUmCmUmArCrU

sgRNA name	SEQ ID NO.	sgRNA sequence and chemical modifications
		***G*U*U*GrUrArGrArUrCrUrGrUrArArCfGfAfUfCfGfGf GfAfAfC*fU*fGfG*fC*mA**
mAlb29-8-48	**5753**	**mG*mU*mU*GAGAAUCGAAAGAUUCUUAAUAAGGCA UCCUUCCGAUGCmC*mU*mU*U*rUrArArUrUmUmCm UmArCrU*rG*rU*rU*rGrUrArGrArUrCrUrGrUrArArCi2 FGi2FAi2FUi2FCi2FGi2FGi2FGi2FAi2FAi2FC*i2FU*i2FGi 2FG*i2FC*mA**
mAlb29-8-47	**5754**	**mG*mU*mU*GAGAAUCGAAAGAUUCUUAAUAAGGCA UCCUUCCGAUGCCUUUrUrArArUrUmUmCmUmArCrU *rG*rU*rU*rGrUrArGrArUrCrUrGrUrArArCi2FGi2FAi2 FUi2FCi2FGi2FGi2FGi2FAi2FAi2FC*i2FU*i2FGi2FG*i2FC *mA**
mAlb29-8-46	**5755**	**mG*mU*mU*GAGAAUCmG*mA*mA*mAGAUUCUUAAU AAGGCAUCmC*mU*mU*mC*mCGAUGCmC*mU*mU*U *rUrArArUrUmUmCmUmArCrU*rG*rU*rU*rGrUrArGrA rUrCrUrGrUrArArCi2FGi2FAi2FUi2FCi2FGi2FGi2FGi2FA i2FAi2FC*i2FU*i2FGi2FG*i2FC*mA**
Nomenclature of chemical modifications: a "/" is used to separate bases with 2'-flourine modifications, m; 2'-O-methyl base (for example a A base with 2'-O-methyl modification is written as mA), i2F; internal 2'-flourine base (for example an internal C with 2'-flourine modification is written as /i2FC/), 52F; 2'-flourine base at the 5' end of the sequence (for example a 5' C with 2'-flourine modification is written as /52FC/), 32F; 2'-flourine base at the 3' end of the sequence (for example a 3' A base with 2'-flourine modification is written as /32FA/), r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *; between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur; AltR1 and AltR2 refer to IDT technologies' proprietary 5' and 3' AltR modifications		

Example 24 – *In vivo* genome editing with the MG29-1 nuclease

[00505] To evaluate the ability of the MG29-1 Type V nuclease to edit the genome *in vivo* in a living animal, we used a lipid nanoparticle to deliver a mRNA encoding the MG29-1 nuclease and one of four guide RNA. The four guide RNA tested are mAlb298-37, mAlb2912-37, mAlb2918-37, and mAlb298-34, the sequences of which are shown in Table 19. Guides mAlb298-37 and mAlb298-34 have the same nucleotide sequence but different chemical modifications while guides mAlb298-37, mAlb2912-37, and mAlb2918-37 have different spacer sequences but the same chemical modifications.

Table 19: Sequences and chemical modifications of guide RNA tested *in vivo* in mice

Guide name	SEQ ID NO.	Sequence
mAlb298-37	**5756**	**mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGA UCUGUAACfGfAfUfCfGfGfGfAfAfC*fU*fGfG*fC*mA**
mAlb2912-37	**5757**	**mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGA UAGUGUAGfCfAfGfAfGfAfGfGfAfA*fC*fCfA*fU*mU**
mAlb2918-37	**5758**	**mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGA UAAGAUUGfAfUfGfAfAfGfAfCfAfA*fC*fUfA*fA*mC**
mAlb298-34	**5759**	**mC*rU*rUrArArUrUmUmCmUmArCrUrGrUrUrGmUmAm GmArUrCrUrGrUrArArCrGrArUrCrGrGrGrA*rAfC*fUfG* fGfC*mA**

Nomenclature of chemical modifications: a "/" is used to separate bases with 2'-flourine modifications, m; 2'-O-methyl base (for example a A base with 2'-O-methyl modification is written as mA), i2F; internal 2'-flourine base (for example an internal C with 2'-flourine modification is written as /i2FC/), 52F; 2'-flourine base at the 5' end of the sequence (for example a 5' C with 2'-flourine modification is written as /52FC/), 32F; 2'-flourine base at the 3' end of the sequence (for example a 3' A base with 2'-flourine modification is written as /32FA/), r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *; between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur; AltR1 and AltR2 refer to IDT technologies' proprietary 5' and 3' AltR modifications

[00506] In the *in vitro* stability assay in Hepa1-6 cell lysates, the mAlb298-37 guide was more stable than the mAlb298-34 guide (**FIG. 72**), demonstrating that the chemical modifications on the mAlb298-37 guide were more effective at protecting the guide RNA against degradation.

[00507] The mRNA encoding MG29-1 was generated by *in vitro* transcription of a linearized plasmid template using T7 RNA polymerase and standard conditions using nucleotides and enzymes purchased from New England Biolabs or Trilink Biotechnologies. The sequence of the MG29-1 coding sequence is shown in SEQ ID No. 5680. The protein coding sequence of the MG29-1 cassette comprises the following elements from 5' to 3': the nuclear localization signal from SV40, a five amino acid linker (GGGS), the protein coding sequence of the MG29-1 nuclease from which the initiating methionine codon was removed, a 3 amino acid linker (SGG) and the nuclear localization signal from nucleoplasmin. The DNA sequence of this cassette was codon optimized for human using a commercially available algorithm. An approximately 100 nucleotide polyA tail was encoded in the plasmid used for *in vitro* transcription, and the mRNA was co-transcriptionally capped using the CleanCAP (™) reagent purchased from Trilink Biotechnologies. Uridine in the mRNA was replaced with N1-methyl pseudouridine.

[00508] The lipid nanoparticle (LNP) formulation used to deliver the MG29-1 mRNA and the guide RNA is based on LNP formulations described in the literature including Kauffman *et al.* (see e.g. *Nano Lett.* 2015, 15, 11, 7300–7306, which is incorporated by reference herein). The

four lipid components were dissolved in ethanol and mixed in an appropriate molar ratio to make the lipid working mix. The mRNA and the guide RNA were either mixed before formulation at a 1:1 mass ratio or formulated in separate LNP that were later co-injected into mice at a 1:1 mass ratio of the two RNA's. In either case, the RNA was diluted in 100 mM Sodium Acetate (pH4.0) to make the RNA working stock. The lipid working stock and the RNA working stock were mixed in a microfluidics device (Ignite NanoAssembler, Precision Nanosystems) at a flow rate ratio of 1:3, respectively, and a flow rate of 12 mls/min. The LNP were dialyzed against phosphate buffered saline (PBS) for 2 to 16 hours and then concentrated using Amicon spin concentrators (Millipore) until the ending volume was achieved. The concentration of RNA in the LNP formulation was measured using the Ribogreen reagent (Thermo Fisher). The diameter and polydispersity (PDI) of the LNP were determined by dynamic light scattering. Example LNP had diameters ranging from 65 nm to 120 nm and PDI of 0.05 to 0.20. LNP were injected intravenously into 8 to 12 week old C57Bl6 wild type mice via the tail vein (0.1 ml per mouse) at a total RNA dose of 1 mg RNA per kg body weight. The mice were sacrificed three days post dosing, and the liver was collected and homogenized using a bead beater (Omni International) in a digestion buffer supplied in the PureLink Genomic DNA Isolation Kit (Thermo Fisher Scientific). Genomic DNA was purified from the resulting homogenate using the PureLink Genomic DNA Isolation Kit (Thermo Fisher Scientific) and quantified by measuring the absorbance at 260 nm. Genomic DNA purified from mice injected with buffer alone was used as a control. The liver genomic DNA was then PCR amplified using primers flanking the region targeted by the guides. The PCR primers used are shown in Table 20. PCR was performed using Pfusion flash high fidelity PCR master mix (Thermo Fisher Scientific) on 50ng of genomic DNA and an annealing temperature of 64°C.

Table 20: Sequences of PCR primers and Sequencing primers used to analyze *in vivo* genome editing in mice

Primer name	SEQ ID NO.	Sequence
mAlb90F	5760	CTCCTCTTCGTCTCCGGC
mAlb1073	5761	CTGCCACATTGCTCAGCAC
mALb460F	5762	GCCTGCTCGACCATGCTATA

[00509] The resulting PCR product was a single band by agarose gel electrophoresis and was purified using the DNA Clean & Concentrator-5 kit (Zymo Research), then subjected to Sanger

sequencing with the primer mAlb460F that is located between 100 and 300 bases from the target sites of the different guides. The Sanger sequencing chromatograms were analyzed for insertions and deletions (INDELS) at the predicted target site for each guide by Tracking of Indels by DEcomposition (TIDE) as described by Brinkman et al (Nucleic Acids Res. 2014 Dec 16; 42(22): e168). The presence of INDELS at the target site is the consequence of the generation of double strand breaks in the DNA, which are then repaired by the error prone cellular repair machinery which introduces insertions and deletions.

[00510] The results of the TIDE analysis are shown in **FIG. 73** and Table 21. Group A mice received LNP encapsulating guide RNA mAlb298-37. Group B mice received LNP encapsulating guide RNA mAlb2912-37. Group C mice received LNP encapsulating guide RNA mAlb2918-37. Group D mice received LNP encapsulating guide RNA mAlb298-34. All mice also received LNP encapsulating the MG29-1 mRNA. The average INDEL frequency in group A that received guide mAlb298-37 was 21%. The average INDEL frequency in group B that received guide mAlb2912-37 was 20%. The average INDEL frequency in group C that received guide mAlb2918-37 was 15%. The average INDEL frequency in group D that received guide mAlb298-34 was 0%. This data demonstrates that the MG29-1 nuclease together with a guide RNA comprised of chemical modified bases (chemistry #37) was active *in vivo* in the liver of mice. Guide mAlb298-34 that has the same nucleotide sequence as guide mAlb298-37, but with different chemical modifications, was not active. Guide mAlb298-34 exhibited less stability in cell lysate than guide mAlb298-37, which correlates to *in vivo* activity.

Table 21: Gene editing at the on target site in the liver of mice at 3 days after IV injection of nuclease mRNA and guide RNA packaged in LNP

Group	Animal	Editing Efficiency (%)	R-Squared
A	1281	20	0.98
	1282	19	0.98
	1283	25	0.98
	1284	24	0.98
	1285	17	0.99
B	1286	24	0.97
	1287	12	0.98

Group	Animal	Editing Efficiency (%)	R-Squared
	1288	16	0.98
	1289	22	0.97
	1290	27	0.97
C	1291	12	0.99
	1292	19	0.99
	1293	Not analyzed	
	1294	13	0.99
	1295	17	0.99
D	1296	0	1
	1297	0	1
	1298	0	1
	1299	0	1
	1300	0	1

Example 25 – MG29-1 guide screen for mouse HAO-1 gene using mRNA transfection

[00511] From a guide screen of exons 1 to 4 of the mouse HAO-1 gene that was performed using MG29-1 protein complexed to the guide RNA that was nucleofected into Hepa1-6 cells, 5 highly active guides were selected for further evaluation by transfection of mRNA encoding MG29-1 mixed with the guide RNA. 300ng mRNA and 120ng single guide RNA were transfected into Hepa1-6 cells as follows. One day before to transfection, Hepa1-6 cells that have been cultured for less than 10 days in DMEM, 10% FBS, 1xNEAA media, without Pen/Strep, were seeded into a TC-treated 24 well plate. Cells were counted, and the equivalent volume to 60,000 viable cells were added to each well. Additional pre-equilibrated media was added to each well to bring the total volume to 500µL. On the day of transfection, 25µL of OptiMEM media and 1.25ul of Lipofectamine Messenger Max Solution (Thermo Fisher) were mixed in a mastermix solution, vortexed, and allowed to sit for at least 5 minutes at room temperature. In separate tubes, 300ng of the MG29-1 mRNA and 120ng of the sgRNA were mixed together with 25µL of OptiMEM media, and vortexed briefly. The appropriate volume of MessengerMax solution was added to

each RNA solution, mixed by flicking the tube and briefly spun down at a low speed. The complete editing reagent solutions were allowed to incubate for 10 minutes at room temperature, then added directly to the Hepa1-6 cells. Two days post transfection, the media was aspirated off of each well of Hepa1-6 cells and genomic DNA was purified by automated magnetic bead purification, via the KingFisher Flex with the MagMAX™ DNA Multi-Sample Ultra 2.0 Kit. The activity of the guides is summarized in Table 22, while the primers used are summarized in Table 23.

Table 22: Average Activity of MG29-1 guides at mouse HAO1 delivered by mRNA Transfection

Guide Name	PAM	SEQ ID NO.	Spacer Sequence	Editing Activity (Average % INDELs)
mH29-1	TTTG	5763	CCCCAGACCTGTAATAGTCATA	26.0
mH29-15	TTTG	5764	TGACTGTGGACACCCCTTACCT	32.7
mH29-16	TTTC	5765	ATTACAGCCTGTCAGACCATGG	15.0
mH29-26	TTTC	5766	TCCATTTCATTACAGCCTGTCA	10.0
mH29-29	TTTC	5767	CCTTAGGAGAAAATGCCAAATC	36.7

Table 23: Primers designed for the mouse HAO1 gene, used for PCR at each of the first four exons, and for sanger sequencing

Target Exon	Use	Primer Name	SEQ ID NO:	Primer Sequence
Mouse HAO1 Exon 1	Fwd PCR	PCR_mHE1_F_+233	5768	GTGACCAACCCTACCCGTTT
	Rev PCR	PCR_mHE1_R_-553	5769	GCAAGCACCTACTGTCTCGT
	Sequencing	Seq_mHE1_F_+139	5770	GTCTAGGCATACAATGTTTGCTCA
Mouse HAO1 Exon 2	Fwd PCR	HAO1_E2_F5721	5771	CAACGAAGGTTCCCTCCAGG
	Rev PCR	HAO1_E2_R6271	5772	GGAAGGGTGTTCGAGAAGGA

Target Exon	Use	Primer Name	SEQ ID NO:	Primer Sequence
	Sequencing	5938F Seq_HAO1_E2	5773	CTATGCAAGGAAAAGATTTGGCC
Mouse HAO1 Exon 3	Fwd PCR	HAO1_E3_F23198	5774	TGCCCTAGACAAGCTGACAC
	Rev PCR	HAO1_E3_R23879	5775	CAGATTCTGGAAGTGGCCCA
	Sequencing	HAO1_E3_F23198	5774	Same as Fwd PCR Primer
Mouse HAO1 Exon 4	Fwd PCR	PCR_mHE4_F_+300	5776	GGCTGGCTGAAAATAGCATCC
	Rev PCR	HAO1_E4_R31650	5777	AGGTTTGGTTCCCCTCACCT
	Sequencing	PCR_mHE4_R_-149	5778	TCTGCCATGAAGGCATATGGAC

Example 52 – Efficiency of mRNA electroporation in T cells

[00512] Primary T cells were purified from PMBCs using a negative selection kit (Miltenyi) according to the manufacturer's recommendations. Nucleofection of mRNA was performed as follow: 200,000 cells were co-transfected with 500ng of mRNA and the indicated amount of guide RNA using a Lonza 4D electroporator (DS-120). Cells were harvested and genomic DNA prepared three days post initial transfection. For conditions labeled "+gRNA": 15h post initial transfection, cells were nucleofected with indicated amount of additional guide. PCR primers appropriate for use in NGS-based DNA sequencing were generated, optimized, and used to amplify the individual target sequences for each guide RNA. The amplicons were sequenced on an Illumina MiSeq machine and analyzed with a proprietary Python script to measure gene editing (**FIG. 74**).

Example 53 – Editing with chemically modified guides

[00513] Primary T cells were purified from PBMCs using a negative selection kit (Miltenyi) according to the manufacturer's recommendations. Nucleofection of mRNA was performed as follow: 200,000 cells were co-transfected with 500ng of mRNA and the indicated amount of guide using a Lonza 4D electroporator (DS-120). Cells were harvested and genomic DNA prepared three days post initial transfection. Nucleofection of RNPs was performed by combining 120 pmol protein and 160 pmol guide RNA. PCR primers appropriate for use in NGS-based DNA sequencing were generated, optimized, and used to amplify the individual

target sequences for each guide RNA. The amplicons were sequenced on an Illumina MiSeq machine and analyzed with a proprietary Python script to measure gene editing (**FIG. 75**).

Example 54 – ELISA assay to assess pre-existing antibody response

[00514] MG29-1 was expressed in and purified from human HEK293 cells using the Expi293™ Expression System Kit (ThermoFisher Scientific). Briefly, 293 cells were lipofected with plasmids encoding the nucleases driven by a strong viral promoter. Cells were grown in suspension culture with agitation and harvested two days post-transfection. The nuclease proteins were fused to a Six-His affinity tag and purified by metal-affinity chromatography to between 50-60% purity. Parallel lysates were made from mock-transfected cells and were subjected to an identical metal-affinity chromatography process. Cas9 was purchased from IDT and is >95% pure.

[00515] MaxiSorp® ELISA plates (Thermo Scientific) were coated with 0.5 µg of nucleases or control proteins diluted in 1X phosphate buffered saline (PBS) and incubated overnight at room temperature. Plates were then washed and incubated with a 1% (w/v) bovine serum albumin (BSA) (Sigma-Aldrich)/1X PBS solution (1% BSA-PBS) for an hour at room temperature. After another washing operation, wells were incubated for 1 h at room temperature with more than 50 separate serum samples taken from randomly selected human donors (1:50 dilution in 1% BSA-PBS). Plates were then washed and incubated for an hour at room temperature with a peroxidase-labeled goat anti-human (Fcγ fragment-specific) secondary antibody (Jackson Immuno Research), diluted 1:50,000 in 1% BSA-PBS. The assay was developed using a 3,3′,5,5′-Tetramethylbenzidine (TMB) Liquid Substrate System kit (Sigma-Aldrich), according to the manufacturer's specifications. Antibody titers are reported as absorbance values measured at 450 nm (**FIG. 76**). Tetanus toxoid was used as the positive control due to wide-spread vaccination against this antigen and was purchased from Sigma Aldrich. The data indicated that, in contrast to SpCas9 and tetanus toxoid (positive control), MG29-1 had similar antibody response to albumin and 293T cell extract, indicating that the donors did not have existing exposure to antigenic epitopes of MG29-1. This suggests this enzyme may be more efficacious for *in vivo* editing as it would be less susceptible to inactivation *in vivo* by existing antibody responses.

Example 55 – *In Vivo* Editing of Mouse HAO-1 with MG29-1 Delivered via Lipid Nanoparticle (LNP)

[00516] The human genetic disease Primary Hyperoxaluria Type I (PH1) is caused by mutations in the alanine-glyoxylate aminotransferase gene (AGXT) that disrupt glycolate metabolism in the liver and result in the overproduction of oxalate. Oxalate is an insoluble metabolite that is cleared

from the body by the kidney and excreted in the urine. Elevated levels of oxalate production result in the accumulation of oxalate in the kidney and other organs which results in kidney failure as well as damage to other organs. The available curative treatment for PH1 is a liver transplant which is often combined with a kidney transplant to replace the defective kidney function. The HAO-1 gene encodes the enzyme glycolate oxidase (GO) which lies upstream of AGXT in the glycolate metabolic pathway. Reduction in the amount of GO protein reduces the production of oxalate and is thus an effective approach for the treatment of PH1 as demonstrated in a mouse model of PH1 (*see* Martin-Higueras et al. Molecular Therapy vol. 24 no. 4, 719–725 (2016) doi: 10.1038/mt.2015.224, which is incorporated by reference in its entirety herein) and in clinical studies with a RNAi drug that targets HAO-1 (see Frishberg et al. CJASN July 2021, 16 (7) 1025-1036 doi: 10.2215/CJN.14730920, which is incorporated by reference in its entirety herein).

[00517] A genome editing approach that knocks down the HAO-1 gene is an attractive approach for a curative therapy for PH1 patients. One approach for a genome editing therapy for PH1 is to create a double strand break within the coding region of the HAO-1 gene in hepatocytes which is repaired by the non-homologous end joining (NHEJ) DNA repair pathway. Hepatocytes are the cell type in the liver that express the HAO-1 gene and the NHEJ pathway is the dominant DNA repair pathway in these cells. The NHEJ repair pathway is error-prone and introduces insertions or deletions at the site of the double strand break which can lead to frame shifts (if the insertions or deletions are not multiples of 3 nucleotides) or to deletions or insertions of amino acids. The introduction of a frame shift can lead to the induction of nonsense-mediated mRNA decay which reduces the level of the mRNA, which further contributes to protein knockdown.

[00518] Double-strand breaks can be generated in a sequence specific manner by RNA guided CRISPR-Cas nucleases. MG29-1 is a type V CRISPR nuclease that utilizes a short guide RNA of between 38 and 42 nucleotides. MG29-1 primarily generates deletions at the cut site when tested in cultured mammalian cells which makes it attractive for the purposes of knocking down a gene. In order to be useful for *in vivo* therapeutics, MG29-1 activity is ideally preserved in living mammals when delivered using a clinically appropriate delivery system. Lipid nanoparticles represent an attractive delivery system for *in vivo* genome editing of hepatocytes in the liver because they efficiently deliver mRNA and sgRNA to hepatocytes after intravenous administration in rodents and primates. We therefore evaluated MG29-1 as a genome editing system for use in knocking down the HAO-1 gene as a potential therapy for PH1.

[00519] Messenger RNA encoding the MG29-1 nuclease was generated by *in vitro* transcription of a linearized plasmid template using T7 RNA polymerase and a mixture of ribonucleotides

rATP, rCTP and rGTP and N1-methyl pseudouridine in place of rUTP and CleanCAP (Trilink). The plasmid also encoded an approximately 100 nt polyA tail at the 3' end of the coding sequence. In addition to an mRNA encoding the wild type MG29-1 protein, a second mRNA encoding the MG29-1 protein with a single amino acid change of S168R was synthesized. The mRNA was purified on commercial spin columns, the concentration was determined by absorbance at 260 nM, and the purity was determined by Tape Station (Agilent).

[00520] Guide RNAs were selected based on editing efficiency evaluations of multiple guides spanning exons 1 to 4 of the mouse HAO-1 gene in the mouse liver cell line Hepa1-6. Three guides were chemically synthesized incorporating a combination of chemical modifications of bases at specific positions referred to collectively as chemical modification #37; these guide RNAs are shown below in **Table 25**.

Table 25: Sequences and chemical modifications of guide RNA tested *in vivo* in mice

guide RNA Name	SEQ ID No.	Sequence
mH29-1_37	5779	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGA UCCCCAGAfCfCfUfGfUfAfAfUfAfG*fU*fCfA*fU*mA
mH29-15_37	5780	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGA UUGACUGUfGfGfAfCfAfCfCfCfCfU*fU*fAfC*fC*mU
mH29-29_37	5781	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGA UCCUUAGGfAfGfAfAfAfAfUfGfCfC*fA*fAfA*fU*mC

Code for modified bases and linkages: m: 2'-O-methyl modified base, f: 2'-fluoro modifed base, *: phosphorothioate linkage

[00521] The MG29-1 mRNA and the guide RNA were separately packaged inside lipid nanoparticles (LNP) using a process essentially as described by Kaufmann et al (Nano Lett. 2015, 15, 11, 7300–7306, PMID: 26469188, DOI:10.1021/acs.nanolett.5b02497, which is incorporated by reference herein in its entirety). Lipids were purchased from Avanti Polar Lipids or from Corden Pharma and dissolved in ethanol. The mRNA or sgRNA was prepared in water then diluted in 100 mM sodium acetate (pH 4.0) to make the RNA working stock. The four lipid components were combined in ethanol at the specific ratios to make the lipid working stock. An example lipid mixture comprised cholesterol, 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), 1,1'-((2-(4-(2-((2-(bis(2-hydroxydodecyl)amino)ethyl)(2-hydroxydodecyl)amino)ethyl)piperazin-1-yl)ethyl)azanediyl)bis(dodecan-2-ol) (C12-200), and DMG-PEG-2000 at molar ratios of 47.5:16:35:1.5. The lipid working stock and the RNA working stock were combined in a microfluidics mixing device (Precision Nanosystems) at a flow rate of 12 mL/min and a ratio of 1 volume of lipid working stock to 3 volumes of RNA working stock. The mass ratio of C12-200 to RNA in the formulation was 10 to 1. The formulated LNPs were diluted 1:1 with 1x PBS then dialyzed twice in 1x PBS for 1 hour each

followed by concentration in Amicon spin concentrators. The resultant LNPs were formulated in 1x PBS buffer, filter sterilized through a 0.2 uM filter, and stored at 4 °C. The concentration of the RNA inside and outside of the LNP was measured using the Ribogreen reagent (Thermo Fisher). The average diameter and polydispersity of the LNPs were measured in the resultant concentrated LNP by dynamic light scattering using a NanoBrook 90Plus (Brookhaven Instruments).

[00522] LNPs encapsulating a guide RNA and the MG29-1 mRNA or MG29-1_S168R mRNA were mixed at a RNA mass ratio of 1:1, then injected intravenously into wild type C57Bl/6 mice via the tail vein at a dose of 1 mg of RNA per kg in a total volume of 0.1 ml per mouse. Mice were sacrificed at 10 days post dosing and the 3 lobes of the liver (left lateral, right lateral, medial) were collected, flash frozen, and stored at -80 °C. The entire left lateral lobe of the liver was homogenized in Genomic Digestion Buffer (Purelink Genomic DNA Purification Kit, Thermo Fisher) using 0.4 mL of buffer per 100 mg of tissue weight in a Bead Mill. Genomic DNA was purified from an aliquot of the homogenate using the Purelink Genomic DNA Purification Kit (Thermo Fisher). The region of the HAO-1 gene targeted by each specific guide RNA was PCR amplified using gene specific primers with adapters complementary to the barcoded primers used for next generation sequencing (NGS) in a PCR reaction comprised of the Q5 high fidelity DNA polymerase and a total of 29 cycles. The product of this first PCR reaction was PCR amplified using the barcoded primers for NGS using a total of 10 cycles. The resulting product was subjected to NGS on an Illumina MiSeq instrument and the results were processed using a custom script to generate the percentage of sequencing reads that contain insertions or deletions (INDELS) at the targeted site in the HAO-1 gene. The genomic DNA from livers of mice injected with PBS buffer were used as controls. The average sequencing read count was 142,000 reads (range 54,000 to 205,000). The NGS data also enabled a prediction of the percentage of INDELS that generate a frame shift as well as a determination of the INDEL profile (**FIG. 80**).

[00523] Total protein was extracted from the entire right lateral lobe of the liver from the same mice by homogenization in PBS in a bead mill followed by three rounds of freeze-thaw at -80 °C and room temperature. The lysate was centrifuged to remove tissue debris and the supernatant was collected. The concentration of total protein in the supernatant was determined using the BCA assay. Equal amounts of total protein were fractionated on SDS-PAGE gels and transferred to nitrocellulose membranes which were then probed with an anti-glycolate oxidase antibody (R&D Systems AF6197, sheep anti-HAO1). Detection utilized an HRP-conjugated secondary antibody (R&D Systems HAF016, Donkey anti-Sheep IgG) followed by detection with

SuperSignal West Dura Chemiluminescent substrate (ThermoFisher Cat. #34076) and visualization with the Bio-Rad ChemiDoc MP imager.

[00524] Editing activity of the tested guides is summarized below in **Table 26**.

Table 26: Editing activity of guide RNA tested *in vivo* in mice

guide RNA name	Exon of HAO-1 targeted	Editing activity in Hepa1-6 cells (% INDELS) by mRNA transfection	Editing activity in Hepa1-6 cells (% INDELS) by RNP nucleofection
mH29-1_37	Exon 1	17 %	92.6 %
mH29-15_37	Exon 3	34 %	96.8 %
mH29-29_37	Exon 4	35 %	93.8 %

[00525] By nucleofection of ribonuclear protein complexes, these guide RNA all had editing activity of 90% or greater in Hepa1-6 cells. The level of editing in Hepa1-6 cells when the MG29-1 mRNA and the guide RNA were co-transfected using lipofection (MessengerMax reagent, Thermo Fisher) was lower due to the lower transfection efficiency, and guides mH29-15 and mH29-29 were significantly more potent than mH29-1 when tested by this transfection method.

[00526] The characteristics of the LNP encapsulating MG29-1 mRNA and the guide RNA are summarized in **Table 27** below.

Table 27: Characteristics of lipid nanoparticles (LNP)

LNP	Payload	LNP Diameter	LNP PDI	% Recovery of RNA (yield)	% of RNA Encapsulated
L012-A	mH29-1_37	74	0.15	71	92.7
L012-B	mH29-15_37	85	0.08	87	93.5
L012-C	mH29-29_37	83	0.121	78	92.1
L012-M	MG29-1 mRNA	98	0.13	97	92.9
L012-S	MG29-1_S168R mRNA	76	0.10	84	92.8

[00527] The LNP encapsulating the 3 guide RNAs had diameters between 74 nm and 85 nm with polydispersity (PDI) of 0.08 to 0.15. The LNP encapsulating the MG29-1 mRNA and the MG29-1_S168R mRNA had diameters of 98 nm and 76 nm, respectively, and PDI values less

than 0.15 indicative of low polydispersity. The percentage of input RNA recovered in the resultant LNP ranged from 71% to 97 % and the percentage of the total RNA that was encapsulated inside the LNP was 92% or greater for all the LNPs. These data demonstrate that both the guide RNA and the MG29-1 mRNA can be efficiently encapsulated in LNPs and that the resulting LNPs were of small size (less than 100 nM) and low polydispersity.

[00528] The level of editing at the target site in the HAO-1 gene 10 days after intravenous injection of LNP encapsulating MG29-1 mRNA and each of the guide RNAs is shown in **FIG. 77**. As expected, a mouse injected with PBS buffer had no editing. Mice injected with LNP encapsulating MG29-1 mRNA and the guide RNA mH29-1_37 exhibited variable levels of editing that ranged from 1% to 52%. Mice injected with LNP encapsulating MG29-1 mRNA and the guide RNA mH29-15_37 exhibited consistent levels of editing with a mean of 50.4% (range 45 %to 54%). Mice injected with LNP encapsulating MG29-1 mRNA and the guide RNA mH29-29_37 exhibited consistent levels of editing with a mean of 57.7% (range 54% to 63%). Mice injected with LNP encapsulating MG29-1_S168R mRNA and the guide RNA mH29-15_37 exhibited consistent levels of editing with a mean of 50.8% (range 38 % to 61%). These results demonstrate that mRNA encoding the MG29-1 nuclease together with a guide RNA with an optimized chemistry can edit a target locus in the liver of mice when delivered in an LNP. Among the 3 guide RNAs with different spacer sequences that were tested, two exhibited consistently high levels of editing while one exhibited more variable editing. Because LNP of this type deliver their payload almost exclusively to hepatocytes, and because hepatocytes make up 60 to 65% of the total number of cells in the rat liver (Bale et al, Scientific Reports volume 6, Article number: 25329 (2016) doi: 10.1038/srep25329, which is incorporated by reference in its entirety herein) and 52% of the total cells in the mouse liver (Barratta et al, Histochemistry and Cell Biology volume 131, pages713–726 (2009) doi: 10.1007/s00418-009-0577-1, which is incorporated by reference in its entirety herein), the maximum level of editing achievable if every hepatocyte were edited at each copy of the HAO-1 gene is predicted to be 60% to 65%. Thus, 50% editing measured in total genomic DNA purified from the liver represents editing in approximately 75 to 80% of the hepatocytes. The inclusion of the S168R amino acid change in MG29-1, which can improve editing efficiency in cultured cells, did not improve editing efficiency in this study with the one guide RNA tested. The improvement in editing efficiencies with the S168R amino acid variant of MG29-1 was previously observed with guide RNAs that exhibited low editing, which may explain why no improvement was observed here with a guide RNA that was already selected for high levels of editing. The INDEL profile (**FIG. 80**) was composed entirely of deletions. The majority of the deletions were between 1 and 11 nucleotides

with a small number of larger deletions. The frequency of predicted frame shift creation among the mice treated with guides mH29-15 and mH29-29 ranged from 70 to 80% of the total INDELS with a mean of 75%. Thus, on average, 75% of the observed INDELS are predicted to create a frame shift in the HAO-1 coding sequence which will result in disruption of the amino acid sequence downstream of the editing site and a high chance of creating a stop codon.

[00529] The other main lobe of the liver from the same mice was analyzed for the level of the protein glycolate oxidase (GO) that is the product of the HAO-1 gene to determine if the INDELS introduced into the HAO-1 gene resulted in a reduction in GO protein levels in the liver. Western blot analysis using an antibody to the GO protein detected a band of the expected size (**Table 28** and **FIGs. 78A-B**).

Table 28: Editing of glycolate oxidase in mice

Mouse	Treatment	Editing %
1	PBS	0
2		0
3	mH29-1_37 MG29-1 mRNA	23
4		1
5		52
6		ND
8	mH29-15_37 MG29-1 mRNA	54
9		54
10		49
11		45
12		51
13	mH29-29_37 MG29-1 mRNA	56
14		63
15		58
16		58
17		54
18	mH29-15_37 MG29-1_S168R mRNA	61
19		39
20		51
21		52
22		ND

[00530] In comparison to mice that received PBS buffer, mice that received LNP encapsulating MG29-1 mRNA and the various guide RNAs exhibited reduced levels of the GO protein. In general, the magnitude of the reduction in GO protein correlated to the editing efficiency at the HAO-1 gene as measured in the same mouse by NGS. Liver protein from 2 of the mice that showed reductions in GO protein were further tested by loading 3 different amounts of total protein on the gel and repeating the Western blot. As shown in **FIG. 79**, the reduction of GO protein in the 2 mice treated with LNP encapsulating MG29-1 mRNA and either guide RNA

mH29-1_37 or mH29-15_37 was clearly observed at the different loadings of total protein. These data demonstrate that the editing of the HAO-1 gene resulted in reductions in the level of GO protein in the liver of the mice.

Example 56 – Gene editing outcomes at the DNA level for TRAC in human peripheral blood B cells

[00531] Human Peripheral Blood B cells were purchased from STEMCELL Technologies and expanded using ImmunoCult™ Human B Cell Expansion Kit for 2 days prior to nucleofection. Nucleofection of MG29-1 RNPs (126 pmol protein/160 pmol guide) was performed into B cells (200,000) using the Lonza 4D electroporator. Cells were harvested and genomic DNA prepared three days post-transfection. For NGS analysis PCR primers appropriate for use in NGS-based DNA sequencing were used to amplify the target sequence for the TRAC 35 guide RNA (SEQ ID NO: 5681). The amplicon was sequenced on an Illumina MiSeq machine and analyzed with a proprietary Python script to measure gene editing (**FIG. 81**).

Table 28B: Sequences of Guide RNAs and Sequences Targeted for Example 56

Guide Target	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting TRAC	5681	MG29-1-TRAC-sgRNA-35	mU*rArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArGrUrCrUrCrUrCrArGrCrUrGrGrUrArCrArCrG*mG
DNA sequence of TRAC target site	5682	MG29-1-TRAC-target site-35	GAGTCTCTCAGCTGGTACACGG
MG29-1 sgRNA targeting TRAC	5683	MG29-1-TRAC-sgRNA-35-AltR	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArGrUrCrUrCrUrCrArGrCrUrGrGrUrArCrArCrGrG/AltR2/
DNA sequence of TRAC target site	5684	MG29-1-TRAC-target site-35-AltR	GAGTCTCTCAGCTGGTACACGG
r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *; between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur; AltR1 and AltR2 refer to IDT technologies' proprietary 5' and 3' AltR modifications			

Example 57 – Gene editing outcomes at the DNA level for TRAC in hematopoietic stem cells (HSCs)

[00532] Mobilized peripheral blood CD34+ cells were acquired from AllCells and cultured in STEMCELL StemSpan™ SFEM II media supplemented with StemSpan™ CC110 cytokine cocktail for 48 hours prior to nucleofection. Nucleofection of MG29-1 RNPs (126 pmol protein/160 pmol guide) was performed into HSCs (200,000) using the Lonza 4D electroporator.

Cells were harvested and genomic DNA prepared three days post-transfection. PCR primers appropriate for use in NGS-based DNA sequencing used to amplify the individual target sequences for MG29-1 TRAC 35 gRNA (SEQ ID NO: 5681). The NGS amplicons were sequenced on an Illumina MiSeq machine and analyzed with a proprietary Python script to measure gene editing (**FIG. 82**).

Example 58 – Gene editing outcomes at the DNA level for TRAC in induced pluripotent stem cells (iPSCs)

[00533] ATCC-BXS0116 Human [Non-Hispanic Caucasian Female] Induced Pluripotent Stem (IPS) Cells are cultured on Corning Matrigel-coated plasticware in mTESR Plus (STEMCELL Technologies) containing 10 μM ROCK inhibitor Y-27632 for 24 hr prior to nucleofection. Nucleofection of MG29-1 RNP (126 pmol protein/160 pmol guide) was performed into iPSCs (200,000) using the Lonza 4D electroporator. Cells were harvested with Accutase for genomic DNA extraction five days post-transfection. PCR primers appropriate for use in NGS-based DNA sequencing were used to amplify the individual target sequences for the TRAC 35 gRNA (SEQ ID NO: 5681). The amplicons were sequenced on an Illumina MiSeq machine and analyzed with a proprietary Python script to measure gene editing (**FIG. 83**).

Example 59 – *In vivo* genome editing with the MG29-1 nuclease quantified by next generation sequencing (NGS)

[00534] To evaluate the ability of the MG29-1 Type V nuclease to edit the genome *in vivo* in a living animal, an mRNA encoding the MG29-1 nuclease and one of four guide RNAs were delivered in a lipid nanoparticle. The four guide RNAs tested were mAlb298-37, mAlb2912-37, mAlb2918-37, and mAlb298-34, the sequences of which are shown below in **Table 29**. Guides mAlb298-37 and mAlb298-34 have the same nucleotide sequence but different chemical modifications, while guides mAlb298-37, mAlb2912-37, and mAlb2918-37 have different spacer sequences but the same chemical modifications.

Table 29: Sequences and chemical modifications of guide RNA tested *in vivo* in mice

Guide name	SEQ ID NO.	Sequence
mAlb298-37	5756	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAU CUGUAACfGfAfUfCfGfGfGfAfAfC*fU*fGfG*fC*mA
mAlb2912-37	5757	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAU AGUGUAGfCfAfGfAfGfAfGfGfAfA*fC*fCfA*fU*mU
mAlb2918-37	5758	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAU AAGAUUGfAfUfGfAfAfGfAfCfAfA*fC*fUfA*fA*mC

Guide name	SEQ ID NO.	Sequence
mAlb298-34	5759	mC*rU*rUrArArUrUmUmCmUmArCrUrGrUrUrGmUmAmGm ArUrCrUrGrUrArArCrGrArUrCrGrGrGrA*rAfC*fUfG*fGfC*m A

M: 2'-O methyl modified base; f: 2'-fluorine modified base; *: phosphorothioate backbone

[00535] In an *in vitro* stability assay in Hepa1-6 cell lysates, the mAlb298-37 guide was more stable than the mAlb298-34 guide, demonstrating that the chemical modifications on the mAlb298-37 guide were more effective at protecting the guide RNA against degradation.

[00536] The mRNA encoding MG29-1 (SEQ ID NO: 5687) was generated by *in vitro* transcription of a linearized plasmid template using T7 RNA polymerase using nucleotides and enzymes purchased from New England Biolabs or Trilink Biotechnologies. The protein coding sequence of the MG29-1 cassette comprises the following elements from 5' to 3': the nuclear localization signal from SV40, a five amino acid linker (GGGS), the protein coding sequence of the MG29-1 nuclease from which the initiating methionine codon was removed, a 3 amino acid linker (SGG), and the nuclear localization signal from nucleoplasmin. The DNA sequence of this cassette was codon optimized for human using a commercially available algorithm. An approximately 100 nucleotide polyA tail was encoded in the plasmid used for *in vitro* transcription and the mRNA was co-transcriptionally capped using the CleanCAP (™) reagent purchased from Trilink Biotechnologies. Uridine in the mRNA was replaced with N1-methyl pseudouridine.

[00537] To generate the lipid nanoparticle (LNP) formulation used to deliver the MG29-1 mRNA and the guide RNA, the four lipid components were dissolved in ethanol and mixed in an appropriate molar ratio to make the lipid working mix. The mRNA and the guide RNA were either mixed prior to formulation at a 1:1 mass ratio or formulated in separate LNP that were later co-injected into mice at a 1:1 mass ratio of the two RNA's. In either case, the RNA was diluted in 100 mM Sodium Acetate (pH 4.0) to make the RNA working stock. The lipid working stock and the RNA working stock were mixed in a microfluidics device (Ignite NanoAssembler, Precision Nanosystems) at a flow rate ratio of 1:3, respectively, and a flow rate of 12 mL/min. The LNP were dialyzed against phosphate buffered saline (PBS) for 2 to 16 hours and then concentrated using Amicon spin concentrators (Milipore) until the pre-determined volume was achieved. The concentration of RNA in the LNP formulation was measured using the Ribogreen reagent (Thermo Fisher). The diameter and polydispersity (PDI) of the LNP were determined by dynamic light scattering. Example LNP had diameters ranging from 65 nm to 120 nm with PDI of 0.05 to 0.20. LNP were injected intravenously into 8 to 12 week old C57Bl6 wild type mice

via the tail vein (0.1 mL per mouse) at a total RNA dose of 1 mg RNA per kg body weight. Three days post dosing, the mice were sacrificed and the liver was collected and homogenized using a bead beater (Omni International) in a digestion buffer supplied in the PureLink Genomic DNA Isolation Kit (Thermo Fisher Scientific). Genomic DNA was purified from the resulting homogenate using the PureLink Genomic DNA Isolation Kit (Thermo Fisher Scientific) and quantified by measuring the absorbance at 260 nm. Genomic DNA purified from mice injected with buffer alone was used as a control.

[00538] The liver genomic DNA was then PCR amplified using a first set of primers flanking the region targeted by the guides. The PCR primers used are shown below in **Table 30**. The 5' end of these primers comprise conserved regions complementary to the PCR primers used in the second PCR, followed by 5 Ns in order to give sequence diversity and improve MiSeq sequencing quality, and end with sequences complementary to the target region in the mouse genome. PCR was performed using Q5® Hot Start High-Fidelity 2X Master Mix (New England Biolabs) on 100ng of genomic DNA and an annealing temperature of 60 °C for a total of 30 cycles. This was followed by a 2nd round of 10 cycles of PCR using primers designed to add unique dual Illumina barcodes (IDT) for next generation sequencing on a MiSeq instrument. Each sample was sequenced to a depth of greater than 10,000 reads using 150bp paired end reads. Reads were merged to generate a single 250 bp sequence from which Indel percentage and INDEL profile was calculated using a proprietary Python Script.

Table 30: Sequences of PCR primers and Next Generation Sequencing primers used to analyze *in vivo* genome editing in mice

Primer name	Purpose	SEQ ID NO.	Sequence
mAlb298_NGS_F	Amplify the target site in albumin intron 1	5782	GCTCTTCCGATCTNNNNNCTTG AGTTTGAATGCACAGATA
mAlb298_NGS_R	Amplify the target site in albumin intron 1	5783	GCTCTTCCGATCTNNNNNTGG AAACAGGGAGAGAAAAACC
mAlb2912_NGS_F	Amplify the target site in albumin intron 1	5784	GCTCTTCCGATCTNNNNNTACA AACATGACAGAAACACTAA
mAlb2912_NGS_R	Amplify the target site in albumin intron 1	5785	GCTCTTCCGATCTNNNNNGATT GATGAAGACAACTAACTGT
mAlb2918_NGS_F	Amplify the target	5786	GCTCTTCCGATCTNNNNNCTTT

Primer name	Purpose	SEQ ID NO.	Sequence
	site in albumin intron 1		GAGTGTAGCAGAGAGG
mAlb2918_NGS_R	Amplify the target site in albumin intron 1	5787	GCTCTTCCGATCTNNNNNCATTATACCGATGGGCGATC

[00539] The results of the NGS analysis are shown in **FIG. 84** and **Table 31**. Group A mice received LNP encapsulating guide RNA mAlb298-37. Group B mice received LNP encapsulating guide RNA mAlb2912-37. Group C mice received LNP encapsulating guide RNA mAlb2918-37. Group D mice received LNP encapsulating guide RNA mAlb298-34. All mice in groups A to D also received LNP encapsulating the MG29-1 mRNA that was mixed with the guide RNA containing LNP at a 1:1 RNA mass ratio prior to injection. Two mice were injected with PBS as controls (Group E).

Table 31: Gene editing at the on target site in the liver of mice at 3 days after IV injection of nuclease mRNA and guide RNA packaged in LNP

Group	Animal	Editing Efficiency (%) by NGS	Mean per group	Standard deviation per group
A	1281	50.65	53.9	2.8
	1282	53.80		
	1283	53.50		
	1284	58.43		
	1285	53.14		
B	1286	53.19	52.3	4.5
	1287	44.49		
	1288	53.73		

Group	Animal	Editing Efficiency (%) by NGS	Mean per group	Standard deviation per group
	1289	55.90		
	1290	54.27		
C	1291	22.46	26.5	4.9
	1292	31.98		
	1293	31.59		
	1294	21.74		
	1295	24.88		
D	1296	12.94	12.7	1.1
	1297	14.35		
	1298	11.98		
	1299	12.78		
	1300	11.52		
E	1209	5.6		
	1209	0.45		

[00540] The average INDEL frequency in group A that received guide mAlb298-37 was 53.9%. The average INDEL frequency in group B that received guide mAlb2912-37 was 52.3%. The average INDEL frequency in group C that received guide mAlb2918-37 was 26.5%. The average

INDEL frequency in group D that received guide mAlb298-34 was 12.7%. These data demonstrate that the MG29-1 nuclease together with a guide RNA comprised of chemical modified bases (chemistry #37 or chemistry #34) was active *in vivo* in the liver of mice. Guide mAlb298-34, which resulted in about 50% of the editing as guide mAlb298-37, has the same nucleotide sequence as mAlb298-37 but different chemical modifications, demonstrating that chemical modifications #37 enable significantly more editing activity *in vivo* than chemical modifications #34. The improved *in vivo* editing observed with chemistry #37 compared to chemistry #34 is consistent with the superior *in vitro* stability of chemistry #37.

[00541] An example INDEL profile generated by the MG29-1 nuclease and guide 298-37 as measured by NGS is shown in **FIG. 85**. The INDEL profiles from the other 4 mice treated with the same LNP were essentially identical. The majority of the INDELS were deletions, with very few insertions detectable. This INDEL profile is distinct from that seen with spCas9, which commonly generates a mixture of insertions and deletions with a tendency to generate +1 and -1 INDELS. The deletions resulting from *in vivo* cleavage by MG29-1 range from -1 to -30, with the majority of deletions between -1 and -10 nucleotides.

Example 59 – Spacer length optimization for MG29-1 single guide RNA

[00542] The guides tested comprised 5 different spacers targeting different regions of human albumin intron 1 (spacers 74, 83, 84, 78, and 87) with chemical modifications called "AltR1/AltR2" provided by Integrated DNA Technologies. The spacer length was titrated from 22 nucleotides (nt) to 17 nt by removal of nucleotides from the 3' end of the guide RNA. Each of these guides (6 per spacer sequence) were evaluated for their editing efficiency in the human liver cell line Hep3B. Hep3B cells (1x 10^5 cells/sample) were electroporated using an Amaxa nucleofection device and program EH-100 with pre-formed ribonucleoprotein complex made by mixing 120 pmol MG29-1 protein and 160 pmol guide RNA. After electroporation, the cells were plated in 24 well plates and cultured for 3 days after which genomic DNA was purified from the cells using a commercial kit (Purelink, Invitrogen). The genomic DNA was analyzed for editing at the on target site (human albumin intron 1) by next generation sequencing (NGS). The NGS data was analyzed by a custom Python script (IndelCalculator v1.3.1). As shown in **FIG. 86**, editing activity was unchanged for all 5 guides when the spacer length was titrated from 22 nucleotides to 20 nucleotides. Shortening the spacer to 19 nucleotides reduced the editing activity of all 5 guides, with more pronounced 75% reductions in activity for 3 of the 5 guides. Reduction of the spacer length to 18 or 17 nucleotides further reduced the editing activity such that a 17 nucleotide spacer was inactive for all 5 guides. Similar data were obtained for 4 different guide RNA targeting a different genomic locus, the human HAO-1 gene (**FIG. 86**). In the case of the

HAO-1 guides, editing activity dropped by 75 to 95% when the guide length was reduced from 20 nt to 22 nt. These data demonstrate that across a range of different guide spacer sequences targeting different genomic loci, the minimal MG29-1 spacer length that retained maximal editing activity was 20 nt. Because longer spacer sequences may increase the risk of off-target editing, identification of the shortest spacers that retain full activity is beneficial.

[00543] A guide RNA (" guide 29") was identified as a highly active guide in a screen for guides with 22 nt spacers for MG29-1 that target the human HAO-1 locus. The spacer length of this guide was reduced to 20 nt by removing the 3' most 2 nt from the spacer to create guides designated as mH29-29.1_37 (SEQ ID NO: 5710) and mH29-29.2_37 (SEQ ID NO: 5711) which differ in their chemical modifications.

[00544] These guides contain the same chemical modifications as chemistry #37 which was based on a 22nt spacer, except that the modifications on the 5' end where the spacer was shortened had to be adjusted. In mH29-29.1_37, the number of fluoro bases was reduced by 2, but the 4 PS bonds and 1 2'-O-methyl base at the 5' remained the same as in the original #37 chemistry. In mH29-29.2_37, the number of fluoro bases was reduced by 2 and the 2'-O-methyl on the last base was retained, but the number of PS bonds at the 5' end was increased from 4 to 5. The relative potency of these guides will be tested in cells in culture and in mice.

Example 60 – Design of a guide RNA for MG29-1 with an 24 nucleotide stem-loop structure at the 5' end to improve stability (chemistry #50)

[00545] An *in vitro* stability assay for MG29-1 and MG3-6/3-4 guides demonstrated that MG29-1 guides can be less stable (**FIG. 87**). In this assay, the guide RNA was incubated in a crude extract from mammalian cells (Hepa1-6) that contains nucleases that can degrade RNA. An MG29-1 guide with chemical modifications on the 5' and 3' ends (Alt-R) was degraded in about 200 mins while about 50% of a MG3-6/3-4 guide with chemical modifications of the 5' and 3' ends (Mod 1) remained intact after 500 mins (**FIG. 87**).

[00546] The structures of the MG29-1 and MG3-6/3-4 guide RNAs (**FIG. 88**) were predicted using the Geneious Prime Software (Turner 2004 algorithm: https://rna.urmc.rochester.edu/NNDB/index.html) and were noted to be significantly different. The MG29-1 guide is about one third of the length of the guide RNA for MG3-6/3-4. In addition, the MG29-1 guide contains minimal secondary structure comprising one stem of 5 nucleotides in length. In contrast, the guide RNA for MG3-6/3-4 contains 3 stem-loops with stem lengths of 10, 6, and 10 nucleotides (**FIG. 88**). The highly active MG3-6/3-4 guide containing a spacer targeting mouse albumin that was used to generate the data in **FIG. 86** was also predicted to

contain a stem of 10 nucleotides (Stem-loop 1 in **FIG. 89**) identical to the 10 nt stem-loop predicted for the backbone alone.

[00547] It was hypothesized that the minimal secondary structure of the MG29-1 guide made it less stable in a cellular milieu that is mimicked by the *in vitro* stability assay. Given the significantly greater *in vitro* stability of the MG3-6/3-4 guide RNA, a modified MG29-1 sgRNA backbone was designed in which the largest stem loop of MG3-6/3-4 (Stem-loop 1) was added at the 5' end of the MG29-1 guide. The predicted structure of this modified MG29-1 guide is shown in **FIG. 89**. The chemical modifications designated chemistry #37 that had been previously demonstrated to significantly improve the stability and activity of the standard MG29-1 guide RNA were incorporated in the design of chemistry #50; specifically, the same phosphorothioate, 2'-O-methyl, and 2'-fluoro modifications present in the backbone and spacer of chemistry #37 were included in chemistry #50. To further stabilize this new design, the 3 nucleotides at the 5' end were modified with phosphorothioate linkages and 2'-O-methyl bases. In addition, phosphorothioate linkages and 2'-O-methyl bases were included in the loop of the added stem loop (stem loop 1 in **FIG. 90**).

[00548] Additional variants of chemistry #50 were designed as shown in **Table 32** below. The designs for chemistries #44, #50, #51, #52, #53, and #54 for any spacer sequence are shown in SEQ ID NOs: 5695-5701, in which N is any ribonucleotide base in the spacer.

Table 32: Summary of chemistry #50 and additional variants

Guide name	SEQ ID	Spacer length (nt)	Extra stem loop	Chemistry on spacer / stem 2	Additional changes
mAlb29-8-50	5689	22	Yes	#37	
mAlb29-8-50b	5690	20	Yes	#37	
mAlb29-8-51b	5691	20	Yes	No 2'-fluoro in spacer	
mAlb29-8-52b	5692	20	Yes	Reduced 2'-fluoro in spacer	
mAlb29-8-53b	5693	20	Yes	#37	PS and methyl on every base in stem loop 1
mAlb29-8-54b	5694	20	Yes	No 2'-fluoro in spacer	PS and methyl on every base in stem loop 1

Example 61 – *In vitro* editing with MG29-1 guide chemistry #50 containing a 24 nucleotide stem-loop structure at the 5' end

[00549] The mouse liver cell line Hepa1-6 (1x 10^5 cells/sample) was electroporated using an Amaxa nucleofection device and program: EH-100 with either pre-formed ribonucleoprotein complex (120 pmol MG29-1 protein mixed with 160 pmol guide RNA) or with a mixture of 500 ng MG29-1 mRNA and 210 pmol guide RNA. The guides tested were mAlb29-8-44 (spacer 8, chemistry 44) and mAlb29-8-50 (spacer 8, chemistry 50). Chemistry 44 comprises the MG29-1 backbone plus the 22 nt spacer and a specific set of chemical modifications of either the bases or the backbone that had been optimized for activity and stability. The sequence of mALb29-8-44 with chemical modifications is shown in SEQ ID NO: 5688. The sequence of mAlb29-8-50 is shown in SEQ ID NO: 5689. Both of these guides contain 22 nucleotide spacers. After electroporation, the cells were plated in 24 well plates and cultured for 3 days after which genomic DNA was purified from the cells using a commercial kit (Purelink, Invitrogen). The genomic DNA was analyzed for editing at the on target site (albumin intron 1) by next generation sequencing (NGS). The NGS data was analyzed by a custom Python script (IndelCalculator v1.3.1). As shown in **FIG. 91**, when the MG29-1 nuclease was delivered by mRNA transfection, chemistry 50 improved editing efficiency from 46% to 94%. When the MG29-1 nuclease was delivered as a protein that was pre-complexed to the guide, both chemistries exhibited editing of 94%. When the nuclease is delivered as a mRNA, the guide is ideally stable enough to survive inside the cell until the mRNA is translated into protein, after which the nuclease can complex with the guide and transit to the nucleus. Thus guide stability is likely more critical when the nuclease is delivered as mRNA. In contrast, the guide may be stabilized when complexed with the nuclease as a RNP prior to electroporation into the cells, in line with the observation that both chemistry 44 and 50 resulted in 94% editing by RNP electroporation (**FIG. 91**). These data suggest that chemistry 50 on the MG29-1 guide RNA containing an additional stem loop mediates improved editing in cells in culture.

Example 62 – *In vivo* gene editing in the liver of mice with MG29-1 guide chemistry 50

[00550] To evaluate the impact of different guide chemistries on gene editing in the liver of mice, we delivered the MG29-1 mRNA and the guide RNA using a lipid nanoparticle (LNP). Messenger RNA encoding the MG29-1 nuclease was generated by *in vitro* transcription of a linearized plasmid template using T7 RNA polymerase and a mixture of ribonucleotides rATP, rCTP, and rGTP and N1-methyl pseudouridine in place of rUTP and CleanCAP (Trilink). The plasmid also encoded an approximately 100 nt polyA tail at the 3' end of the coding sequence. The mRNA was purified on commercial spin columns, the concentration was determined by absorbance at 260 nM, and the purity was determined by Tape Station (Agilent). Three different guide RNAs that comprise the same spacer sequence but with different chemistries / backbones

were evaluated in a single mouse study: mAlb29-8-44 (SEQ ID NO: 5688), mAlb29-8-50 (SEQ ID NO: 5689), and mAlb29-8-37 (SEQ ID NO: 5702). Guide mAlb29-12-44 (SEQ ID NO: 5703), which contains a different spacer (spacer 12) with chemistry 44, was also tested. The MG29-1 mRNA and the guide RNA were separately packaged inside lipid nanoparticles (LNP) using a process essentially as described by Kaufmann et al (PMID: 26469188, DOI:10.1021/acs.nanolett.5b02497, which is incorporated by reference herein in its entirety). Lipids were purchased from Avanti Polar Lipids or from Corden Pharma and dissolved in ethanol. The mRNA or sgRNA was prepared in water then diluted in 100 mM sodium acetate (pH 4.0) to make the RNA working stock. The four lipid components were combined in ethanol at specified ratios to make the lipid working stock. An example lipid mixture comprised cholesterol, DOPE, C12-200, and DMG-PEG-2000 at molar ratios of 47.5:16:35:1.5. The lipid working stock and the RNA working stock were combined in a microfluidics mixing device (Precision Nanosystems) at a flow rate of 12 mL/min and a ratio of 1 volume of lipid working stock to 3 volumes of RNA working stock. The mass ratio of C12-200 to RNA in the formulation was 10 to 1. The formulated LNP were diluted 1:1 with 1x PBS then dialyzed twice in 1x PBS for 1 hour each followed by concentration in Amicon spin concentrators. The resultant LNPs were formulated into 1x PBS buffer, filter sterilized through a 0.2 uM filter, and stored at 4 °C. The concentration of the RNA inside and outside of the LNP was measured using the Ribogreen reagent (Thermo Fisher). The average diameter and polydispersity of the LNPs were measured in the resultant concentrated LNPs by dynamic light scattering using a NanoBrook 90Plus (Brookhaven Instruments). Representative LNPs ranged in size from 80 to 100 nanometers with a PDI <0.15 and an RNA encapsulation ratio of greater than 90%. LNP encapsulating a guide RNA and the MG29-1 mRNA were mixed at an RNA mass ratio of 1:1 then injected intravenously into wild type C57Bl/6 mice via the tail vein at a dose of 0.5 mg of RNA per kg in a total volume of 0.1 ml per mouse (N=5 mice per LNP). Mice were sacrificed at 4 days post dosing and the 3 lobes of the liver (left lateral, right lateral, medial) were collected, flash frozen, and stored at -80 °C. The entire left lateral lobe of the liver was homogenized in Genomic Digestion Buffer (Purelink Genomic DNA Purification Kit, Thermo Fisher) using 0.4 mL of buffer per 100 mg of tissue weight in a Bead Mill. Genomic DNA was purified from an aliquot of the homogenate using the Purelink Genomic DNA Purification Kit (Thermo Fisher). The albumin intron 1 region was PCR amplified from 50 ng of the genomic DNA in a reaction containing 0.5 micro molar each of the primers mAlb90F (CTCCTCTTCGTCTCCGGC) and mAlb1073R (CTGCCACATTGCTCAGCAC) and 1x Pfusion Flash PCR Master Mix. The resulting 984 bp PCR product which spans the entire intron 1 of mouse albumin was purified using a column

based purification kit (DNA Clean and Concentrator, Zymo Research) and sequenced using primers located within 150 to 350 bp of the predicted target site for each guide RNA. The PCR product generated using primers mAlb90F and mAlb1073R from a PBS buffer injected mouse was sequenced in parallel as a control. The Sanger sequencing chromatograms were analyzed using Inference of CRISPR Edits (ICE) that determines the frequency of INDELS as well as the INDEL profile (Hsiau et. al, Inference of CRISPR Edits from Sanger Trace Data. BioArxiv. 2018 https://www.biorxiv.org/content/early/2018/01/20/251082).

[00551] Without wishing to be bound by theory, it is understood that when a nuclease creates a double strand break (DSB) in DNA inside a living cell, the DSB is repaired by the cellular DNA repair machinery. In actively dividing cells, such as transformed mammalian cells in culture, and in the absence of a repair template, it is understood that this repair occurs by the NHEJ pathway. Without wishing to be bound by theory, it is understood that the NHEJ pathway is an error prone process that introduces insertions or deletions of bases at the site of the double strand break (Lieber, M.R, Annu Rev Biochem. 2010; 79: 181–211). These insertions and deletions are understood to be a hallmark of a double strand break that occurred and was subsequently repaired, and are thus widely used as a readout of the editing or cutting efficiency of the nuclease.

[00552] The editing efficiency in the liver of the mice is shown in **FIG. 92**. Comparing guides with the same spacer sequence (guide 8) but different chemistries or backbones, the guide with chemistry #44 (SEQ ID NO: 5688) was the least active (mean 1.6% editing) while the guide with chemistry #37 (SEQ ID NO: 5702) was more active (mean 4% editing). Chemistry #44 differs from chemistry #37 by the addition of 2 additional PS bonds in the spacer region (chemistry #44 has 6 PS bonds in the spacer region while chemistry #37 has 4 PS bonds in the spacer region). The guide with chemistry #50 (SEQ ID NO: 5689) was significantly more active with mean editing of 22%, approximately 5-fold higher than the guide with chemistry #37 and 10-fold higher than the guide with chemistry #44. When the same genomic DNA was analyzed for editing by next generation sequencing (NGS), the levels of editing with guide spacer 8 were determined to be 7%, 7%, and 42% for chemistries 44, 37, and 50, respectively. For guide spacer 12 with chemistry 44, the level of editing was 4% and 9% when measured by ICE and NGS, respectively confirming that chemistry 44 has similar activity to chemistry 37. These data demonstrate that chemistry #50, which contains a stem-loop from the MG3-6/3-4 guide added to the 5' end of the normal MG29-1 guide backbone, exhibits significantly improved editing in the liver of mice after systemic delivery in an LNP. Thus, guide chemistry 50 provides improved *in vivo* potency of the MG29-1 nuclease.

Example 63 – Further improvements to MG29-1 guide chemistry 50

[00553] Further improvements to the MG29-1 guide chemistry #50 are contemplated. In one potentially improved version, all of the nucleotides in the stem-loop 1 that was added to the 5' end of the standard MG29-1 guide backbone are chemically modified with both 2'-O-methyl on the bases and phosphorothioate linkages as in chemistries 53 (SEQ ID NO: 5700) and 54 (SEQ ID NO: 5701). In one potentially improved version, all of the 2'-flouro bases in the spacer are changed to standard nucleotides as in chemistries 51 (SEQ ID NO: 5698) and 54 (SEQ ID NO: 5701). In another potentially improved version, the number of 2'-flouro bases in the spacer are reduced by 2-fold as in chemistry 52 (SEQ ID NO: 5699). The reduction in the number of 2'-fluoro bases may have impacts on guide specificity.

Example 64 – Design of a MG29-1 single guide RNA comprising the native guide array

[00554] An alternative approach to improving the stability, and thus the potency, of the MG29-1 single guide RNA is to design a native like CRISPR array for MG29-1, mimicking the documented process in which MG29-1 nuclease cleaves its own CRISPR array to generate a mature guide. The array was designated as mAlb29-g8-37-array (SEQ ID NO: 5712) and it comprises two copies of a 22 nt spacer targeting mouse albumin (spacer 8) embedded in the native CRISPR array for MG29-1.

[00555] The designed array is 126nt long and it comprises a repeat, followed by a spacer, followed by a repeat, followed by a spacer. The predicted secondary structure of mAlb29-g8-37-array is shown in **FIG. 93** in which the 5' end is circled in blue and the 3' end is circled in red. This RNA is designed to be cleaved inside mammalian cells by an expressed MG29-1 nuclease to generate two functional sgRNAs. Chemical modifications were included in mAlb29-g8-37-array to promote stability. The modifications in the spacer and the MG29-1 backbone portions are based on those used in chemistry #37 but with additional modifications and some changes.

Example 65 – Guides for MG29-1 nuclease with 20 nt spacers targeting human albumin intron 1 with chemistry #37

[00556] A guide screen against human albumin intron 1 using the MG29-1 nuclease and single guide RNA with 22 nt spacers identified 5 guides with high editing activity in human Hep3B cells. These guides were designated as spacer numbers 87, 78, 74, 83, and 84. Versions of these single guide RNA's with 20 nt spacers were designed incorporating the chemistry #37 chemical modifications and these were designated as hA29-87-37B (SEQ ID NO: 5713), hA29-78-37B (SEQ ID NO: 5714), hA29-74-37B (SEQ ID NO: 5715), hA29-83-37B (SEQ ID NO: 5716), and hA29-84-37B (SEQ ID NO: 5717).

Example 66 – Guides for MG29-1 nuclease with 20 nt spacers targeting human HAO1 with chemistry #37

[00557] A guide screen against human HAO-1 (encoding glycolate oxidase) using the MG29-1 nuclease and single guide RNA with 22 nt spacers identified 4 guides with high editing activity in human Hep3B cells. These guides were designated as spacer numbers 4, 21, 23, and 41. Versions of these single guide RNA's with 20 nt spacers were designed incorporating the chemistry #37 chemical modifications and these were designated as hH29-4_37b (SEQ ID NO: 5718), hH29-21_37b (SEQ ID NO: 5719), hH29-23_37b (SEQ ID NO: 5720), and hH29-41_37b (SEQ ID NO: 5721).

Example 67 – Guides for MG29-1 nuclease with 22 nt spacers targeting human HAO1 with chemistry #50

[00558] A guide screen against human HAO-1 (encoding glycolate oxidase) using the MG29-1 nuclease and single guide RNA with 22 nt spacers identified 4 guides with high editing activity in human Hep3B cells. These guides were designated as spacer numbers 4, 21, 23, and 41. Versions of these single guide RNA's with 22 nt spacers were designed incorporating the chemistry #50 chemical modifications and these were designated as hH29-4_50 (SEQ ID NO: 5722), hH29-21_50 (SEQ ID NO: 5723), hH29-23_50 (SEQ ID NO: 5724), and hH29-41_50 (SEQ ID NO: 5725).

Example 68 – Guides for MG29-1 with 20 nt spacers targeting human HAO1 with chemistry #50

[00559] A guide screen against human HAO-1 (encoding glycolate oxidase) using the MG29-1 nuclease and single guide RNA with 22 nt spacers identified 4 guides with high editing activity in human Hep3B cells. These guides were designated as spacer numbers 4, 21, 23, and 41. Versions of these single guide RNA's with 20 nt spacers were designed incorporating the chemistry #50 chemical modifications and these were designated as hH29-4_50b (SEQ ID NO: 5726), hH29-21_50b (SEQ ID NO: 5727), hH29-23_50b (SEQ ID NO: 5728), and hH29-41_50b (SEQ ID NO: 5729).

Example 69 – Guides for MG29-1 with 22 nt spacers targeting mouse HAO1 with chemistry #50

[00560] A guide screen against mouse HAO-1 (encoding glycolate oxidase) using the MG29-1 nuclease and single guide RNA with 22 nt spacers identified 3 guides with high editing activity in mouse Hepa1-6 cells. These guides were designated as spacer numbers 1, 15, and 29. Versions of these single guide RNA's with 22 nt spacers were designed incorporating the chemistry #50

chemical modifications and these were designated as mH29-1-50 (SEQ ID NO: 5730), mH29-15-50 (SEQ ID NO: 5731), and mH29-29-50 (SEQ ID NO: 5704).

Example 70 – Guides for MG29-1 with 20 nt spacers targeting mouse HAO1 with chemistry #50

[00561] A guide screen against mouse HAO-1 (encoding glycolate oxidase) using the MG29-1 nuclease and single guide RNA with 22 nt spacers identified 4 guides with high editing activity in mouse Hepa1-6 cells. These guides were designated as spacer numbers 1, 15, and 29. Versions of these single guide RNA's with 20 nt spacers were designed incorporating the chemistry #50 chemical modifications and these were designated as mH29-1-50b (SEQ ID NO: 5732), mH29-15-50b (SEQ ID NO: 5733), and mH29-29-50b (SEQ ID NO: 5705).

Example 71 – Comparison of the *in vivo* editing efficiency of MG29-1 to spCas9

[00562] To compare the *in vivo* editing efficiency of the MG29-1 nuclease to that of spCas9, a dose response was performed in wild type C57Bl6 mice. Albumin intron 1 was selected as a genomic target locus for both spCas9 and MG29-1. An in silico search for spCas9 guide target sites in mouse intron 1 using the Chop-Chop algorithm (see e.g. Labun et al doi: 10.1093/nar/gkz365, which is incorporated by reference in its entirety herein) identified a total of 39 potential guides, which were ranked according to their efficiency score and off-target prediction. In addition, guide target sites located within 50 bp of exon 1 or exon 2 were excluded. The top 3 guides from this ranking were designated mAlbR1 (SEQ ID NO: 5734), mAlbR2 (SEQ ID NO: 5735), and mAlbR3 (SEQ ID NO: 5736), and were chemically synthesized with chemical modifications at both the 5' and 3' ends comprising methylated bases (represented by the nomenclature mA, mC, mG, and mU) and phosphorothioate backbone linkages (represented by the nomenclature A*, C*, G*, and U*). The editing efficiencies of these 3 guides were evaluated in the mouse liver cell line Hepa1-6 by nucleofection of ribonucleoprotein complexes formed by mixing the guide RNA and commercially sourced spCas9 protein (purchased from Integrated DNA technologies) at a molar ratio of 1:2.5 (protein to guide RNA). 20 moles of spCas9 protein was mixed with 50 moles of guide RNA and subsequently nucleofected into $2x10^5$ Hepa1-6 cells using an Amaxa electroporation device with program setting EH100. The nucleofected cells were each transferred to a well of a 48 well plate in fresh growth media and cultured for 48 h in a 5% CO_2/37 °C humidified incubator. Genomic DNA was purified from the cells using the Purelink kit (Invitrogen, ThermoFisher) and analyzed for editing at the target site in albumin intron 1 by PCR amplification of the target locus using primers mAlb90F and mAlb1073R (SEQ ID NOs: 5737 and 5738) and a high fidelity PCR enzyme mix. The PCR product was subjected to Sanger sequencing using primers mAlb282F or mAlb460F. The Sanger

sequencing chromatograms were analyzed for insertions and deletions ("indels") at the predicted target site for each guide by Tracking of Indels by DEcomposition (TIDE) as described by Brinkman et al (Nucleic Acids Res. 2014 Dec 16; 42(22), doi: 10.1093/nar/gku936, which is incorporated by reference in its entirety herein). The presence of indels at the target site is the consequence of the generation of double strand breaks in the DNA, which are then repaired by the error prone cellular repair machinery which introduces insertions and deletions. The results of the TIDE analysis are shown in **Table 33**. All three guides generated indel frequencies of greater than 90%, demonstrating that all three guides are highly active.

Table 33: INDEL frequencies in Hepa1-6 cells nucleofected with guide RNA for spCas9 targeting mouse albumin intron 1 and spCas9 protein as a RNP

Sample ID	Guide	INDEL %	R2
1	mAlbR1	92	0.95
2	mAlbR2	91	0.91
3	mAlbR3	96	0.96

[00563] Guide mALbR2 was synthesized with extensive chemical modifications as described previously (Yin et al. doi:10.1038/nbt.4005, and WO 2019/079527 Al, each of which are incorporated by reference in their entirety herein). The chemical modifications include modifications of the 3 bases at the 5' end and 3 bases at the 3' end with 2'-O-methyl bases and phosphorothioate linkages between the 3 bases at the 5' end and the 3 bases at the 3' end. In addition, 33 of the internal bases are modified with 2'-O-methyl (SEQ ID NO: 5741). These chemical modifications of the guide RNA for spCas9 were reported to enable efficient editing *in vivo* in mouse liver after delivery of the mRNA for spCas9 and the guide RNA in a lipid nanoparticle (see e.g. WO 2019/079527 A1, which is incorporated by reference in its entirety herein).

[00564] A guide screen for guides that target the MG29-1 nuclease to mouse albumin intron 1 and promote cleavage and indel formation was performed. The two guides with the highest editing activity in Hepa1-6 cells when the nuclease was delivered as a mRNA were mALb29-8 and mAlb29-12. Guide mALb29-8 was selected for comparison to spCas9 guide mAlbR2 *in vivo* in mice. Chemical and structural modifications to the guide RNA for MG29-1 were optimized by

evaluating the impact of different chemical modifications including 2'O-methyl and 2'-fluoro modified bases, phosphorothioate linkages, as well as an additional stem loop upon the stability and editing activity of the guide.

[00565] Experiments on guide chemistry optimization indicated that guide chemistry #50 was the most active guide chemistry among those tested. When delivered *in vivo* to mice using a LNP encapsulating MG29-1 mRNA and the same guide RNA sequence targeting mouse albumin intron 1, but with two different guide chemistries (#37 and #50), chemistry #50 was about 4-fold more potent than chemistry #37 at a dose of 0.5 mg/kg. Therefore, MG29-1 guide chemistry #50 was selected to test *in vivo* in comparison to spCas9 with its cognate guide mALbR2 (SEQ ID NO: 5741).

[00566] Messenger RNA encoding the MG29-1 nuclease or the spCas9 nuclease was generated by *in vitro* transcription of a linearized plasmid template using T7 RNA polymerase and a mixture of ribonucleotides rATP, rCTP, and rGTP, N1-methyl pseudouridine, and the CleanCAP capping reagent (Trilink Biotechnologies). The SV40-derived nuclear localization sequence (PKKKRKVGGGGS) followed by a short linker was included at the N terminus of the coding sequence of both spCas9 and MG29-1. The nuclear localization signal from nucleoplasmin preceded by a short linker (SGGKRPAATKKAGQAKKKK) was added to the C-terminus of the coding sequence for both spCas9 and MG29-1. Thus, the same nuclear localization signals were used for both MG29-1 and spCas9. The plasmids also encoded an approximately 100 nt polyA tail at the 3' end of both spCas9 and MG29-1 coding sequences, which generates a polyA tail in the mRNA. The coding sequences for both spCas9 and MG29-1 were codon optimized using the same algorithm (see e.g. Raab et al , DOI 10.1007/s11693-010-9062-3, which is incorporated by reference in its entirety herein). The DNA sequence encoding the spCas9 mRNA is in SEQ ID NO: 5742 and the amino acid sequence encoded by the spCas9 mRNA is in SEQ ID NO: 5743. The mRNA was purified on commercial spin columns, the concentration was determined by absorbance at 260 nM, and the purity was determined by Tape Station (Agilent); the purity was found to be equivalent for both spCas9 mRNA and MG29-1 mRNA. For *in vivo* delivery to mice, the spCas9 mRNA/mAlbR2 guide or the MG29-1 mRNA/mAlb29-8-50 guide were packaged inside lipid nanoparticles (LNP) using a process essentially as described by Kaufmann *et al.* (PMID: 26469188, DOI:10.1021/acs.nanolett.5b02497, which is incorporated by reference herein). The guide RNA and the mRNA were separately packaged for both spCas9 and for MG29-1. Lipids (purchased from Avanti Polar Lipids or from Corden Pharma) were dissolved in ethanol. The mRNA or guide RNA was prepared in water, then diluted in 100 mM sodium acetate (pH 4.0) to make the RNA working stock. The four lipid components were combined in

ethanol at the specified ratios to make the lipid working stock. An example lipid mixture comprised cholesterol, a neutral lipid such as 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), a cationic lipid such as 1,1'-((2-(4-(2-((2-(bis(2-hydroxydodecyl)amino)ethyl)(2-hydroxydodecyl)amino)ethyl)piperazin-1-yl)ethyl)azanediyl)bis(dodecan-2-ol) (C12-200), and a PEG-linked lipid such as 1,2-dimyristoyl-rac-glycero-3-methoxypolyethylene glycol-2000 (DMG-PEG-2000) at molar ratios of 47.5:16:35:1.5. The lipid working stock and the RNA working stock were combined in a microfluidics mixing device (Precision Nanosystems) at a flow rate of 12 mL/min and a ratio of 1 volume of lipid working stock to 3 volumes of RNA working stock. The mass ratio of C12-200 to RNA in the formulation was 10 to 1. The formulated LNP were diluted 1:1 with 1x PBS then dialyzed twice in 1x PBS for 1 hour each, followed by concentration in Amicon spin concentrators. The resultant LNP were formulated into 1x PBS buffer, filter sterilized through a 0.2 μM filter and stored at 4 °C. The concentration of the RNA inside and outside of the LNP was measured using the Ribogreen reagent (Thermo Fisher). The average diameter and polydispersity of the LNP were measured in the resultant concentrated LNP by dynamic light scattering using a NanoBrook 90Plus (Brookhaven Instruments). Representative LNP ranged in size from 80 to 100 nanometers with a PDI <0.15 and an RNA encapsulation ratio of greater than 90%. The average diameter, polydispersity, and RNA encapsulation efficiency is shown below in **Table 34**.

Table 34: Summary of LNP characteristics

LNP	RNA	Average Diameter (nm)	Polydispersity	Percent encapsulation of RNA
LNP-A	MG29-1 mRNA	57	0.082	94.9
LNP-B	mAlb29-8-50 guide (SEQ ID NO: 5744)	47	0.082	93.7
LNP-C	Cas9 mRNA	60	0.114	94.5
LNP-D	mAlbR2 guide (SEQ ID NO: 5741)	50	0.059	94.5

[00567] LNP encapsulating the guide RNA mAlb29-8-50 and the MG29-1 mRNA were mixed at an RNA mass ratio of 1:1. LNP encapsulating the guide RNA mAlbR1 and the spCas9 mRNA were mixed at an RNA mass ratio of 1:1. Both LNP mixtures were injected intravenously into wild type C57Bl/6 mice via the tail vein with total RNA doses of 1 mg/kg, 0.5 mg/kg, or 0.25 mg/kg of RNA in a total volume of 0.1 mL per mouse (N=5 mice per LNP dose). Mice were sacrificed at 5 days post-dosing and the whole liver was flash frozen and stored at -80 °C. The entire left lateral lobe of the liver was homogenized in Genomic Digestion Buffer (Purelink Genomic DNA Purification Kit, Thermo Fisher) using 0.4 mL of buffer per 100 mg of tissue weight in a Bead Mill. Genomic DNA was purified from an aliquot of the homogenate using the Purelink Genomic DNA Purification Kit (Thermo Fisher). The albumin intron 1 region was PCR amplified from 50 ng of the genomic DNA in a reaction containing 0.5 micromolar each of the primers mAlb90F (SEQ ID NO: 5737, CTCCTCTTCGTCTCCGGC) and mAlb1073R (SEQ ID NO: 5738, CTGCCACATTGCTCAGCAC) and 1x Pfusion Flash high fidelity PCR Master Mix. The resulting 984 bp PCR product, which spans the entire intron 1 of mouse albumin, was purified using a column-based purification kit (DNA Clean and Concentrator, Zymo Research). The PCR product was sequenced by next generation sequencing (NGS), analyzing for creation of indels in the target sequence, which were used as indicators of creation of double-strand breaks by the Cas enzymes and engagement of the NHEJ pathway. This detection method is based in the concept that when a nuclease creates a double strand break (DSB) in DNA inside a living cell, the DSB is believed to be repaired by the cellular DNA repair machinery which, in the absence of a repair template, occurs by the NHEJ pathway. As the NHEJ pathway is known to be an error prone process that introduces insertions or deletions of bases at the site of the double strand break (Lieber, M.R, Annu Rev Biochem. 2010; 79: 181–211), these insertions and deletions (indels) are therefore used as a hallmark of a double strand break that occurred and was subsequently repaired, and thus as a readout of the editing or cutting efficiency of the nuclease.

[00568] The sequencing reads were analyzed with a custom Python script (IndelCalculator v1.3.1) that aligns each sequence read to the wild-type target sequence (in this case Albumin intron 1) and calculates the number of reads that contain at least one indel irrespective of the indel size within a window that spans 10 base pairs either side of the predicted on-target cut site for the nuclease. The editing efficiency (indel frequency) in each of the 5 mice in each group, as well as the mean and standard deviation for the group, are summarized in **FIG. 94**. No editing was detected in the control mice injected with PBS buffer. Both spCas9 mRNA/mAlbR2 LNP and the MG29-1 mRNA/mAlb29-8-50 LNP resulted in dose-dependent editing. At all 3 doses, the editing efficiency was higher for the MG29-1 mRNA/mAlb29-8-50 LNP than for the spCas9

mRNA/mAlbR2 LNP. The mean editing efficiencies at the 3 doses are summarized in **Table 35**. At a dose of 1 mg/kg (0.5 mg/kg mRNA and 0.5 mg/kg guide RNA), MG29-1 was slightly more potent than spCas9, resulting in about 15% more indels. At a dose of 0.5 mg/kg (0.25 mg/kg mRNA and 0.25 mg/kg guide RNA), MG29-1 resulted in about 50% more indels. At a dose of 0.25 mg/kg (0.125 mg/kg mRNA and 0.125 mg/kg guide RNA), MG29-1 resulted in 100% more indels. These data demonstrate that using the same LNP for delivery and mRNA produced using an identical process, the MG29-1 nuclease combined with an appropriately optimized guide RNA is more potent than the spCas9 nuclease and an appropriately modified guide RNA. The superior *in vivo* editing efficiency of MG29-1 was especially evident at the lowest dose tested, where MG29-1 was 2-fold more potent than spCas9 at the same dose. These results suggest that the MG29-1 nuclease and an appropriately modified guide RNA exemplified by chemistry #50 may have an advantage for *in vivo* gene editing using LNP delivery.

Table 35. Mean editing efficiency in the whole liver of mice at 5 days after intravenous injection of LNP encapsulating either MG29-1 mRNA and guide mAlb29-8-50 (mA29-8-50) or spCas9 mRNA and guide mAlbR2 at three doses, or PBS buffer (Control).

mRNA	Guide RNA	Dose (mg/kg)	Mean editing (%)	Standard deviation
MG29-1	mAlb29-8-50	1	72.3	2.3
MG29-1	mAlb29-8-50	0.5	62.7	1.4
MG29-1	mAlb29-8-50	0.25	33.2	5.4
spCas9	mAlbR2	1	60.0	2.3
spCas9	mAlbR2	0.5	40.3	6.3
spCas9	mAlbR2	0.25	15.6	2.3
PBS control		-	**0.18**	**0.1**

Example 72 – Identification of active single guide RNA for MG29-1 that target the exonic regions of human HAO-1 by mRNA-based transfection in Hep3B cells

[00569] Sequence-specific nucleases can be used to disrupt the coding sequence of a gene of interest, thereby creating a functional knockout of the protein encoded by that gene. This can be of therapeutic use when the knockdown of the protein has a beneficial effect in a particular human disease. One way to disrupt the coding sequence of a gene is to make a double strand break within the exonic regions of the gene using a sequence-specific nuclease. The double

strand break is repaired via error-prone repair pathways, primarily non-homologous end joining (NHEJ) to generate insertions or deletions which can result in either frameshift mutations or changes to the amino acid sequence which disrupt the function of the protein.

[00570] To identify guide RNAs for MG29-1 that efficiently create double strand breaks at exonic regions of the gene encoding human glycolate oxidase (GO), single guide RNAs (sgRNAs) with a spacer length of 22 nt targeted to exons 1 to 4 of the human hydroxyacid oxidase (HAO-1) gene (GenBank RefSeq accession number NG_046733) were identified using the guide finding algorithm in the Geneious Prime nucleic acid analysis software (https://www.geneious.com/prime/).

[00571] The first four exons of the human HAO-1 gene encode amino acids comprising approximately the N-terminal 50% of the HAO-1 coding sequence. The first 4 exons were chosen because indels created towards the N-terminus of the coding sequence of a gene are more likely to create a frameshift or missense mutation that disrupts the activity of the protein. Using the more restrictive PAM for MG29-1 of TTTN, a total of 42 potential sgRNAs were identified within human HAO-1 exons 1 through 4. Guides that spanned the intron/exon boundaries were included because such guides may create INDELS that interfere with splicing. To create the full sgRNAs, the backbone sequence (UAAUUUCUACUGUUGUAGAU) was added to the 3' end of the spacer sequence. The sgRNAs were chemically synthesized incorporating chemically modified bases documented to improve the performance of sgRNAs for the type V nuclease cpf1 ("AltR1/AltR2" chemistry, commercially available at Integrated DNA Technologies). The spacer sequences of these guides are shown in **Table 36**.

Table 36: Sequences of 42 single guide RNA targeting human HAO-1 for testing in human Hep3B cells

Guide Name	SEQ ID	Exon	PAM	Spacer Sequence (as DNA)	sgRNA Sequence
hH29-1	4184	1	TTTA	GCATGTTGTTCATAATC ATTGA	UAAUUUCUACUGUUGUA GAUGCAUGUUGUUCAUA AUCAUUGA
hH29-2	4185	1	TTTG	GAAGTACTGATTTAGCA TGTTG	UAAUUUCUACUGUUGUA GAUGAAGUACUGAUUUA GCAUGUUG
hH29-3	4186	1	TTTG	TATCAATGATTATGAAC AACAT	UAAUUUCUACUGUUGUA GAUUAUCAAUGAUUAUG AACAACAU

Guide Name	SEQ ID	Exon	PAM	Spacer Sequence (as DNA)	sgRNA Sequence
hH29-4	4187	1	TTTG	CCCCAGACCTGTAATAGTCATA	UAAUUUCUACUGUUGUAGAUCCCCAGACCUGUAAUAGUCAUA
hH29-5	4188	1	TTTC	TTCATCATTTGCCCCAGACCTG	UAAUUUCUACUGUUGUAGAUUUCAUCAUUUGCCCCAGACCUG
hH29-6	4189	1	TTTC	TTACCTGGAAAATGCTGCAATA	UAAUUUCUACUGUUGUAGAUUUACCUGGAAAAUGCUGCAAUA
hH29-7	4190	1	TTTT	CTTACCTGGAAAATGCTGCAAT	UAAUUUCUACUGUUGUAGAUCUUACCUGGAAAAUGCUGCAAU
hH29-8	4191	1	TTTG	GCTGATAATATTGCAGCATTTT	UAAUUUCUACUGUUGUAGAUGCUGAUAAUAUUGCAGCAUUUU
hH29-9	4192	1	TTTA	AAAAATAAATTTTCTTACCTGG	UAAUUUCUACUGUUGUAGAUAAAAAUAAAUUUUCUUACCUGG
hH29-10	4193	1	TTTT	AAAAAATAAATTTTCTTACCTG	UAAUUUCUACUGUUGUAGAUAAAAAAUAAAUUUUCUUACCUG
hH29-11	4194	2	TTTT	ATTTTATTTTTTAATTCTAGAT	UAAUUUCUACUGUUGUAGAUAUUUUAUUUUUUAAUUCUAGAU
hH29-12	4195	2	TTTA	TTTTATTTTTTAATTCTAGATG	UAAUUUCUACUGUUGUAGAUUUUUAUUUUUUAAUUCUAGAUG
hH29-13	4196	2	TTTT	ATTTTTTAATTCTAGATGGAAG	UAAUUUCUACUGUUGUAGAUAUUUUUUAAUUCUAGAUGGAAG
hH29-14	4197	2	TTTA	TTTTTTAATTCTAGATGGAAGC	UAAUUUCUACUGUUGUAGAUUUUUUUAAUUCUAGAUGGAAGC
hH29-15	4198	2	TTTT	TTAATTCTAGATGGAAG	UAAUUUCUACUGUUGUA

Guide Name	SEQ ID	Exon	PAM	Spacer Sequence (as DNA)	sgRNA Sequence
				CTGTA	GAUUUAAUUCUAGAUGG AAGCUGUA
hH29-16	4199	2	TTTT	TAATTCTAGATGGAAGC TGTAT	UAAUUUCUACUGUUGUA GAUUAAUUCUAGAUGGA AGCUGUAU
hH29-17	4200	2	TTTT	AATTCTAGATGGAAGCT GTATC	UAAUUUCUACUGUUGUA GAUAAUUCUAGAUGGAA GCUGUAUC
hH29-18	4201	2	TTTA	ATTCTAGATGGAAGCTG TATCC	UAAUUUCUACUGUUGUA GAUAUUCUAGAUGGAAG CUGUAUCC
hH29-19	4202	2	TTTC	AGCAACATTCCGGAGCA TCCTT	UAAUUUCUACUGUUGUA GAUAGCAACAUUCCGGAG CAUCCUU
hH29-20	4203	2	TTTT	AGGACAGAGGGTCAGCA TGCCA	UAAUUUCUACUGUUGUA GAUAGGACAGAGGGUCA GCAUGCCA
hH29-21	4204	2	TTTA	GGACAGAGGGTCAGCAT GCCAA	UAAUUUCUACUGUUGUA GAUGGACAGAGGGUCAG CAUGCCAA
hH29-22	4205	3	TTTC	TTTCTCAGCCTGTCAGTC CCTG	UAAUUUCUACUGUUGUA GAUUUUCUCAGCCUGUCA GUCCCUG
hH29-23	4206	3	TTTC	TCAGCCTGTCAGTCCCT GGGAA	UAAUUUCUACUGUUGUA GAUUCAGCCUGUCAGUCC CUGGGAA
hH29-24	4207	3	TTTG	TGACAGTGGACACACCT TACCT	UAAUUUCUACUGUUGUA GAUUGACAGUGGACACAC CUUACCU
hH29-25	4208	3	TTTG	AATCTGTTACGCACATC ATCCA	UAAUUUCUACUGUUGUA GAUAAUCUGUUACGCACA UCAUCCA
hH29-26	4209	4	TTTT	ATGCATTTCTTATTTTAG GATG	UAAUUUCUACUGUUGUA GAUAUGCAUUUCUUAUU

Guide Name	SEQ ID	Exon	PAM	Spacer Sequence (as DNA)	sgRNA Sequence
					UUAGGAUG
hH29-27	4210	4	TTTA	TGCATTTCTTATTTTAGGATGA	UAAUUUCUACUGUUGUAGAUUGCAUUUCUUAUUUUAGGAUGA
hH29-28	4211	4	TTTC	TTATTTTAGGATGAAAAATTTT	UAAUUUCUACUGUUGUAGAUUUAUUUUAGGAUGAAAAUUUU
hH29-29	4212	4	TTTT	AGGATGAAAAATTTTGAAACCA	UAAUUUCUACUGUUGUAGAUAGGAUGAAAAAUUUUGAAACCA
hH29-30	4213	4	TTTA	GGATGAAAAATTTTGAAACCAG	UAAUUUCUACUGUUGUAGAUGGAUGAAAAAUUUUGAAACCAG
hH29-31	4214	4	TTTC	CTCAGGAGAAAATGATAAAGTA	UAAUUUCUACUGUUGUAGAUCUCAGGAGAAAAUGAUAAAGUA
hH29-32	4215	4	TTTT	CCTCAGGAGAAAATGATAAAGT	UAAUUUCUACUGUUGUAGAUCCUCAGGAGAAAAUGAUAAAGU
hH29-33	4216	4	TTTT	GAAACCAGTACTTTATCATTTT	UAAUUUCUACUGUUGUAGAUGAAACCAGUACUUUAUCAUUUU
hH29-34	4217	4	TTTG	AAACCAGTACTTTATCATTTTC	UAAUUUCUACUGUUGUAGAUAAACCAGUACUUUAUCAUUUUC
hH29-35	4218	4	TTTA	TCATTTTCTCCTGAGGAAAATT	UAAUUUCUACUGUUGUAGAUUCAUUUUCUCCUGAGGAAAAUU
hH29-36	4219	4	TTTT	CTCCTGAGGAAAATTTTGGAGA	UAAUUUCUACUGUUGUAGAUCUCCUGAGGAAAAUUUUGGAGA
hH29-37	4220	4	TTTC	TCCTGAGGAAAATTTTGGAGAC	UAAUUUCUACUGUUGUAGAUUCCUGAGGAAAAUUUUGGAGAC

Guide Name	SEQ ID	Exon	PAM	Spacer Sequence (as DNA)	sgRNA Sequence
hH29-38	4221	4	TTTA	GCCACATATGCAGCAAGTCCAC	UAAUUUCUACUGUUGUAGAUGCCACAUAUGCAGCAAGUCCAC
hH29-39	4222	4	TTTT	GGAGACGACAGTGGACTTGCTG	UAAUUUCUACUGUUGUAGAUGGAGACGACAGUGGACUUGCUG
hH29-40	4223	4	TTTG	GAGACGACAGTGGACTTGCTGC	UAAUUUCUACUGUUGUAGAUGAGACGACAGUGGACUUGCUGC
hH29-41	4224	4	TTTG	ATATCTTCCCAGCTGATAGATG	UAAUUUCUACUGUUGUAGAUAUCUUCCCAGCUGAUAGAUG
hH29-42	4225	4	TTTG	CAACAATTGGCAATGATGTCAG	UAAUUUCUACUGUUGUAGAUCAACAAUUGGCAAUGAUGUCAG

[00572] Hep3B cells (ATCC catalog number HB-8064), a transformed human liver cell line (derived from a hepatocellular carcinoma), were cultured under standard conditions in growth media (EMEM media with 10% FBS) in a 5% CO_2 incubator and transfected with a mixture of mRNA encoding MG29-1 and each of the single guide RNAs. The mRNA encoding MG29-1 was generated by T7 polymerase *in vitro* transcription of a plasmid in which the coding sequence of MG29-1 had been cloned. The MG29-1 coding sequence was codon optimized using human codon usage tables and flanked by nuclear localization signals derived from SV40 (at the N-terminus) and from Nucleoplasmin (at the C-terminus). In addition, a 5' untranslated region (5' UTR) was included at the 5' end of the coding sequence to improve translation. A 3' UTR followed by an approximately 90 to 110 nucleotide polyA tract was included in the mRNA (encoded in the plasmid) at the 3' end of the coding sequence to improve mRNA stability *in vivo*. The DNA sequence that encodes the MG29-1 mRNA without the polyA tail is shown in SEQ ID 5830. The *in vitro* transcription reaction included the Clean Cap® capping reagent (Trilink BioTechnologies), the resulting RNA was purified using the MEGAClear™ Transcription Clean-Up kit (Invitrogen), and purity was evaluated using the TapeStation (Agilent) and found to be composed of >90% full length RNA.

[00573] When co-transfection of mRNA and guide with a lipid transfection reagent such as MessengerMAX is used, the mixture of the two RNA molecules forms a complex with the

positively charged lipid, the complex enters the cells via endocytosis, and eventually some of the RNA reaches the cytoplasm. In the cytoplasm, the mRNA is translated into protein. In the case of an RNA-guided nuclease such as MG29-1, the resulting MG29-1 protein will presumably form a complex with the sgRNA in the cytoplasm before entering the nucleus in a process mediated by the nuclear localization signals that were engineered into the MG29-1 protein. This process is similar to that of delivering an mRNA and a guide RNA in a lipid nanoparticle *in vivo*, a method of use that is envisaged for therapeutic applications in which the HAO-1 gene would be functionally inactivated by introduction of INDELS within the coding sequence. Thus, the 42 single guide RNAs for editing activity were screened using co-transfection of mRNA and sgRNA because this method is likely a better representation of the planned therapeutic use and is thus likely to more accurately predict which of the 42 sgRNAs will be most active in a therapeutic application.

[00574] A total of 2 x 10^5 Hep3B cells were plated per well of 24 well plates in growth media (EMEM plus 10% FBS) and incubated overnight in a 5% CO_2/37 °C humidified incubator. The following day, Lipofectamine MessengerMAX (Thermo Fisher) was diluted in OPTIMEM media (1.25 µL MessengerMAX plus 25 µL OPTIMEM per transfection). 300 ng of MG29-1 mRNA (0.22 pmol) and 120 ng (8.4 pmol) of sgRNA were combined in 25 µL OPTIMEM media, then mixed with 26 µL of the diluted MessengerMAX by gently flicking the tube. After incubating for 5 to 10 mins at room temperature, the RNA/MessengerMAX mixture was added to each well of Hep3B cells and mixed by swirling gently. The cells were incubated overnight (16 h), after which the media was exchanged for fresh growth media. At 48 h after addition of the RNA/MessengerMAX mixture, the cells were collected by trypsinization or by on-plate lysis, and genomic DNA was purified using either the Purelink Genomic DNA Extraction kit (Thermo Fisher) or the MagMax DNA Extraction Kit (Thermo Fisher) and the KingFisher robotic system (Thermo Fisher). The purified genomic DNA was quantified by absorbance at 260 nm. HAO-1 gene sequences targeted by the single guides were amplified by PCR from the purified genomic DNA using exon-specific primers (**Table 37**) and Phusion Flash High-Fidelity PCR Master Mix (Thermo Fisher). PCR products were purified and concentrated using the DNA Clean & Concentrator 5 kit (Zymo Research), and 40 ng of PCR product was subjected to Sanger sequencing (at ELIM Biosciences) using primers located within 100 bp to 350 bp of the predicted target site for each sgRNA (**Table 37**). PCR products derived from untreated Hep3B cells were included as controls. The sequences of the PCR products matched the expected sequences of the HAO-1 exons.

Table 37: Primers designed for the human HAO1 gene, used for PCR amplification of the first four exons, and for Sanger sequencing.

Target Exon	Use	Primer Name	Primer Sequence
Human HAO1 Exon 1	Fwd PCR	PCR_hHe1_F_+490	TTTCATGGATGCCCCGTTCA
	Rev PCR	PCR_hHe1_R_-412	ACGAAAAGCCAGCAGGAAGA
	Sequencing	Seq_hHe1_R_-121	AGCCCCAAGAACTTTTCCCT
Human HAO1 Exon 2	Fwd PCR	PCR_hHe2_F_+391	TGCATCAGTGGTTGTCAGGG
	Rev PCR	PCR_hHe2_R_-387	CCTAGCTGTGACTTTGGGCA
	Sequencing	Seq_hHe2_R_-152	TGGAAAGAAGAGGAGCAGGAC
Human HAO1 Exon 3	Fwd PCR	PCR_hHe3_F_+238	AGGCTGGATGTTCAGGTTCTT
	Rev PCR	PCR_hHe3_R_-212	TCCCAAAGCCAAAGCCCTTA
	Sequencing	Seq_hHe3_F_+186	AGCAGAAATAACTCCAGTAGCCA
Human HAO1 Exon 4	Fwd PCR	PCR_hHe4_F_+324	GCTGGCTGAAAATCGTGTCAA
	Rev PCR	PCR_hHe4_R_-348	TCCTTGGGGCTTCTCTTTGG
	Sequencing	Seq_hHe4_F_+174	ACTGATTAAGACCACTAGAGTATCACA

[00575] The Sanger sequencing chromatograms were analyzed for insertions and deletions (indels) at the predicted target site for each sgRNA by an algorithm called Tracking of Indels by DEcomposition (TIDE) as described by Brinkman *et al.* (*See, e.g.*, Nucleic Acids Res. 2014 Dec 16; 42(22): e168.Published online 2014 Oct 9. doi: 10.1093/nar/gku936, which is incorporated by reference herein). As presented in **Table 38**, 14 guides demonstrated detectable editing at their predicted target sites. Ten guides exhibited editing activity of 10% or greater. These data demonstrate that the MG29-1 nuclease can generate RNA-guided, sequence-specific, double strand breaks in exonic regions of the human HAO-1 gene in a cultured liver cell line.

Table 38: Editing activity of 42 sgRNA targeting exons 1 to 4 of the human HAO-1 gene in Hep3B cells.

Guide Name	PAM	Average INDEL (Percentage)*	SD*
hH29-1	TTTA	8.5	4.95
hH29-2	TTTG	15.25	5.56
hH29-3	TTTG	6	5.66
hH29-4	TTTG	39.5	10.47
hH29-5	TTTC	0	0
hH29-6	TTTC	8	8.49
hH29-7	TTTT	0	0
hH29-8	TTTG	2.5	3.54
hH29-9	TTTA	0	0
hH29-10	TTTT	0	0
hH29-11	TTTT	0	0
hH29-12	TTTA	0	0
hH29-13	TTTT	0	0
hH29-14	TTTA	0	0
hH29-15	TTTT	0	0
hH29-16	TTTT	0	0
hH29-17	TTTT	0	0
hH29-18	TTTA	3.5	4.95
hH29-19	TTTC	18.75	12.74
hH29-20	TTTT	6.5	4.95
hH29-21	TTTA	68	8.04
hH29-22	TTTC	6	1.41
hH29-23	TTTC	17.5	2.89
hH29-24	TTTG	10	4.24
hH29-25	TTTG	0	0
hH29-26	TTTT	0	0
hH29-27	TTTA	0	0
hH29-28	TTTC	0	0
hH29-29	TTTT	0	0
hH29-30	TTTA	0	0
hH29-31	TTTC	0	0
hH29-32	TTTT	0	0

Guide Name	PAM	Average INDEL (Percentage)*	SD*
hH29-33	TTTT	0	0
hH29-34	TTTG	0	0
hH29-35	TTTA	0	0
hH29-36	TTTT	0	0
hH29-37	TTTC	0	0
hH29-38	TTTA	0	0
hH29-39	TTTT	0	0
hH29-40	TTTG	0	0
hH29-41	TTTG	34	13.04
hH29-42	TTTG	0	0

*Data are the mean of 2 independent transfections

[00576] Six of the sgRNAs (hH29-2, hH29-4, hH29-19, hH29-21, hH29-23, and hH29-41) with the highest editing activity from this initial screen were re-tested using the same MG29-1 mRNA / sgRNA MessengerMax transfection method in Hep3B cells. Two independent transfections were performed, and indel frequency was determined using the same Sanger sequencing method described above, followed by analysis using TIDE. The mean indel frequencies (**FIG. 95**) ranged from 20% to 75%. The rank order of editing efficiency was hH29-21 > hH29-41 > hH29-4 > hH29-19 > hH29-23 = hH29-2. An evaluation of the frequency of indels that generate a frame shift (Out of Frame Editing) was also made based on the Sanger Chromatograms, and these results are plotted in **FIG. 95**. The ratio of out of frame edits to total indels was different for different sgRNAs, but the rank order of out of frame editing frequency was the same as total indels.

[00577] Four of the sgRNAs (hH29-21, hH29-4, hH29-41, and hH29-23) with the highest editing activity in Hep3B cells were also evaluated for their editing activity by nucleofection of ribonucleoprotein particles (RNP) into both Hep3B cells and another human liver-derived cell line, HuH7. The ribonuclear protein complex was formed by mixing 160 pmol of sgRNA and 126 pmol purified MG29-1 protein in PBS buffer. A total of 2 x 10^5 Hep3B or HuH7 cells in suspension in complete SF nucleofection reagent (Lonza) were nucleofected with the pre-formed nuclease / sgRNA complex using a 4D nucleofection device (Lonza). After nucleofection, the cells were plated in 24 well plates in growth media plus 10% FBS and incubated in a 5% CO_2 incubator for 48 h to 72 h. Genomic DNA was then extracted from the cells using a column

based purification kit (Purelink genomic DNA mini kit, ThermoFisher Scientific) and quantified by absorbance at 260 nm. HAO-1 gene sequences targeted by the single guides were amplified by PCR from the purified genomic DNA using the relevant exon-specific primers (**Table 37**) and Phusion Flash High-Fidelity PCR Master Mix (Thermo Fisher). PCR products were purified and concentrated using DNA Clean & Concentrator 5 (Zymo Research), and 40 ng of PCR product subjected to Sanger sequencing (at ELIM Biosciences) using relevant primers (**Table 37**) located within 100 to 350 bp of the predicted target site for each sgRNA. PCR products derived from untransfected cells were included as controls. The sequence of the PCR products matched the expected sequences of the HAO-1 exons. The Sanger sequencing chromatograms were analyzed for insertions and deletions (indels) at the predicted target site for each guide by Tracking of Indels by DEcomposition (TIDE) as described by Brinkman *et al.* (Nucleic Acids Res. 2014 Dec 16; 42(22): e168.Published online 2014 Oct 9. doi: 10.1093/nar/gku936). The presence of indels at the target site is the consequence of the generation of double strand breaks in the DNA, which are then repaired by the error-prone cellular repair machinery which introduces insertions and deletions. The results (**FIG. 96**) demonstrate that using the more efficient transfection method of RNP nucleofection, the editing frequency for these top 4 guides ranged from 25% to 95%. The rank order of editing activity for these 4 guides was hH29-21 > hH29-4> hH29-41> hH29-23, which was similar but not identical to the rank order of editing activity measured by transfection of mRNA and sgRNA by MessengerMAX in Hep3B cells (**FIG. 95**).

Example 73 – Editing activity of the most active MG29-1 guides targeting exons 1 to 4 of human HAO-1 in primary human hepatocytes

[00578] Primary human hepatocytes (PHH) are hepatocytes that are isolated from the livers of deceased humans using documented methods such as Kegel *et al.* (doi: 10.3791/53069). PHH are the closest cell-based model for human hepatocytes in their native state *in vivo* and thus represent an *in vitro* cell system that can be used to predict the performance of gene editing systems in humans *in vivo*. PHH have undergone minimal manipulation, do not undergo cell division, and have a limited lifespan in culture of about 7 days. PHH were obtained from commercial suppliers (Lonza, Gibco) and cultured according to the protocols provided by the supplier. To evaluate the editing activity of the four most active MG29-1 guides targeting human HAO-1, the sgRNAs with spacers 4, 21, 23, and 41 were chemically synthesized with the incorporation of chemical modification #37 to the backbone and the nucleobases that improve the stability and activity of the sgRNA. The 4 sgRNAs with chemical modifications #37 are designated as hH29-4-37 (SEQ ID 5831), hH29-21-37 (SEQ ID 5832), hH29-23-37 (SEQ ID 5833), and hH29-41-37 (SEQ ID

5834). 1028 ng of MG29-1 mRNA and 222 ng of each single guide RNA (1:20 molar ratio of mRNA:guide RNA) were transfected into primary human hepatocyte (PHH) cells as follows. One day prior to transfection, PHH cells were thawed and seeded in 1.0 ml HBM™- 5% FBS - HCM™ SingleQuot Supplements media into collagen-treated 12 well plates at 1,000,000 viable cells per well. On the day of transfection, 60.4 µL of OptiMEM media and 2.1 µL of Lipofectamine MessengerMax Solution (Thermo Fisher) were mixed in a master mix solution, vortexed, and allowed to sit for at least 10 minutes at room temperature. In separate tubes, 1028 ng of MG29-1 mRNA and 222 ng of the sgRNA were mixed, brought to a volume of 62.5 µL with OptiMEM media, and pipetted briefly. The appropriate volume of MessengerMax solution was added to each RNA solution, mixed by flicking the tube, and briefly spun down at a low speed. The complete editing reagent solutions were allowed to incubate for at least 10 minutes at room temperature, then added directly to the PHH cells. Following the transfection, media was replaced every day until harvest. Three days post-transfection, the culturing media was aspirated from each well of PHH cells and replaced with MagMAX™ Cell and Tissue DNA Extraction Buffer (Thermo Fisher). Cells were scraped and transferred to a 96-well plate, and genomic DNA was purified by automated magnetic bead purification via the KingFisher Flex with the MagMAX™ DNA Multi-Sample Ultra 2.0 Kit (Thermo Fisher).

[00579] The region of the HAO-1 gene targeted by each specific sgRNA was PCR amplified with Q5 high fidelity DNA polymerase and exon specific primers (**Table 37**) but with the addition of adapters complementary to the barcoded primers used for next generation sequencing (NGS) for a total of 29 cycles. The product of this first PCR reaction was PCR amplified using the barcoded primers for NGS using a total of 10 cycles. The resulting product was subjected to NGS on an Illumina MiSeq instrument, and the results were processed using a custom script to generate the percentage of sequencing reads that contain insertions or deletions (indels) at the targeted site in the HAO-1 gene. Two independent transfections of PHH were performed, and in each experiment, each of the sgRNA were tested in duplicate wells that were separately assayed for indels by NGS. The average of the indel frequency in the 2 wells was then calculated. The mean of indel frequency from the 2 independent experiments was then determined (**Table 39**) and also shown in graph format (**FIG. 97**), in which the error bars represent the standard deviation.

Table 39: Editing activity in PHH of 4 MG29-1 sgRNA targeting human HAO-1

sgRNA	Exon	Mean total INDEL %*	Stdev
hH29-4-37	1	58.3	5.6

sgRNA	Exon	Mean total INDEL %*	Stdev
hH29-21-37	2	59.1	10.5
hH29-23-37	3	31.5	10.5
hH29-41-37	4	49.2	9.9

* data are the mean of 2 independent transfection experiments each performed in duplicate wells

[00580] The results indicate that all four guides edited the HAO-1 gene in PHH with mean INDEL frequencies ranging from 20% to 58% (**FIG. 97**). Guides hH29-4-37 and mH29-21-37 exhibited the highest editing activity in PHH. Guide hH29-41-37 was slightly less active than guides hH29-4-37 and mH29-21-37, but the difference was not significant. Guide hH29-23-37 was the least active of the 4 guides in PHH. Guide hH29-23-37 was also the least active of these 4 guides in Hep3B and HuH7 cells (**Table 37**, **FIG. 95**, **FIG. 96**). PHH are a surrogate for editing hepatocytes *in vivo*. These data demonstrate that the MG29-1 nuclease and an appropriate sgRNA have utility in generating insertions and deletions in the coding sequence of the human HAO-1 gene, which are expected to generate a mixture of non-sense, missense, and deletion mutations leading to disruption of GO protein expression and/or activity.

[00581] The indel profile obtained from the same NGS sequence data of the HAO-1 gene in PHH transfected with MG29-1 mRNA and the 4 sgRNAs was used to determine the percentage of INDELS that result in a frame shift. Sequence reads in which the number of bases inserted or deleted at the target site are not 3 bases or a multiple of 3 bases will shift the reading frame of the HAO-1 protein coding sequence. A frame shift will alter the amino acid sequence encoded in the mRNA downstream of the indel and in many cases will introduce an in frame stop codon at some point (this can be predicted for each allele). The NGS data (comprising several thousand reads for each sample) was analyzed using a Python script that calculates the percentage of total sequencing reads in which there is an indel that created a frame shift, and this was designated as the out of frame (OOF) indel percentage. The OOF indel percentage is plotted in **FIG. 97** alongside the total indel percentage. The OOF indel percentage ranged from 20% to 40%, which represents between 70% and 80% of the total INDEL percentage, demonstrating that the majority of indels are predicted to create a frameshift. The in-frame indels will delete 1 or more codons from the mRNA, and thus will remove 1 or more amino acids from the glycolate oxidase protein that is encoded by HAO-1. **FIGs. 98A** and **98B** show representative indel profiles for each of the 4 MGf29-1 sgRNAs from edited PHH. In-frame deletions of 3, 6, 9, 12, 15, 18, and 21 bases are evident at different relative frequencies. An analysis of the frequencies of the in-frame deletions

and their impact on the GO protein sequence can be used to inform selection of sgRNA that will result in the greatest reduction in GO protein function.

Example 74 – *In vivo* editing activity of single guide RNAs for MG29-1 with 22 and 20 nucleotide spacers that target the exonic regions of mouse HAO-1

[00582] To evaluate the ability of the MG29-1 Type V nuclease to edit the genome *in vivo* in a living animal, a lipid nanoparticle was used to deliver an mRNA encoding the MG29-1 nuclease and one of four guide RNAs. The ability of MG29-1 with sgRNA comprising 22 nucleotide (nt) spacers to edit the mouse HAO-1 locus in the liver of mice when delivered in an LNP was demonstrated in Example 55. Experiments in cultured mammalian cells demonstrated that reducing the length of the spacer region in MG29-1 sgRNA from 22 nt to 20 nt did not change editing activity. Reducing the spacer from 22 nt to 20 nt may be advantageous in terms of minimizing off-target activity and in terms of sgRNA manufacture. In order to validate that MG29-1 sgRNA with a reduced spacer length of 20 nt retained potency *in vivo*, four guide RNAs (mH29-29-37, mH29-29-44, mH29-29s-37, and mH29-29s-44) were designed and tested in mice. The sequences of these guides are shown below in Table 40.

Table 40: Sequences and chemical modifications of guide RNAs tested in vivo in mouse study

Guide RNA	Sequence
mH29-29-37	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAU **CCUUAGGfAfGfAfAfAfAfUfGfCfC*fA*fAfA*fU*mC**
mH29-29-44	mC*mU*mU*U*U*AAUUmUmCmUmACU*G*U*U*GUAGAU **CCUUAGGfAfGfAfAfAfAfUfG*fCfC*fA*fA*fA*fU*mC**
mH29-29s-37	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAU **CCUUAGGfAfGfAfAfAfAfUfG*fCfC*fA*fA*mA**
mH29-29s-44	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAU **CCUUAGGfAfGfAfAfAfAfUfG*fC*fC*fA*fA*mA**

Notations for chemical modifications: m = 2'O-Methyl ribonucleotide (e.g mC = cytosine ribonucleotide with 2'-O Methyl in place of 2' hydroxyl); f = 2'Fluorine ribonucleotide (e.g fC = cytosine ribonucleotide with 2' fluorine in place of 2' hydroxyl); * = phosphorothioate bond. All other bases are native ribonucleotides. Backbone sequence in normal type. Spacer sequence in bold type.

[00583] Guides mH29-29-37 and mH29-29-44 have the same nt sequence (spacer 29) but different chemical modifications (chemistries 37 and 44), while guides mH29-29s-37 and mH29-29s-44 have spacer sequences shortened by two nts at the 3' end but are otherwise identical in nt sequence to mH29-29-37 and mH29-29-44, respectively. The locations of the 2'-fluoro, 2'-O-methyl, and phosphorothioate modifications in the spacer region of the guides with a 20 nt spacer are shifted relative to those in the corresponding guides with a 22 nt spacer. Because the

chemical modifications in the sgRNA impact sgRNA stability, which is critical for *in vivo* potency, it was important to evaluate the impact of these changes on in vivo editing.

Preparation of MG29-1 mRNA

[00584] The mRNA encoding MG29-1 (SEQ ID NO: 5846) was generated by *in vitro* transcription of a linearized plasmid template using T7 RNA polymerase and standard conditions using nucleotides and enzymes purchased from New England Biolabs or Trilink Biotechnologies. The protein coding sequence of the MG29-1 cassette comprises the following elements from 5' to 3': the nuclear localization signal from SV40, a five amino acid linker (GGGS), the protein coding sequence of the MG29-1 nuclease from which the initiating methionine codon was removed, a 3 amino acid linker (SGG), and the nuclear localization signal from nucleoplasmin. The DNA sequence of this cassette was codon optimized for human using a commercially available algorithm. An approximately 100 nucleotide polyA tail was encoded in the plasmid used for *in vitro* transcription and the mRNA was co-transcriptionally capped using the CleanCAP™ reagent purchased from Trilink Biotechnologies. Uridine in the mRNA was replaced with N1-methyl pseudouridine.

Preparation of lipid nanoparticles

[00585] The lipid nanoparticle (LNP) formulation used to deliver the MG29-1 mRNA and the guide RNA is based on LNP formulations described in the literature including Kauffman *et al.* (*Nano Lett.* 2015, 15, 11, 7300–7306 https://doi.org/10.1021/acs.nanolett.5b024970). The four lipid components were dissolved in ethanol and mixed in an appropriate molar ratio to make the lipid working mix. The RNAs were diluted in 100 mM Sodium Acetate (pH 4.0) to make the RNA working stocks. The lipid working stock and the RNA working stocks were mixed in a microfluidics device (Ignite NanoAssembler, Precision Nanosystems) at a flow rate ratio of 1:3, respectively, and a flow rate of 12 mL/min. The LNP were dialyzed against phosphate buffered saline (PBS) for 2 to 16 hours and then concentrated using Amicon spin concentrators (Milipore) until the desired volume was achieved. The concentration of RNA in the LNP formulation was measured using the Ribogreen reagent (Thermo Fisher). The diameter and polydispersity (PDI) of the LNP were determined by dynamic light scattering. Typical LNP had diameters ranging from 70 nm to 84 nm and PDI of 0.098 to 0.150.

Mouse dosing and harvesting

[00586] LNP for mRNA and sgRNA were mixed at 1:1 mass ratio and injected intravenously into 7 week-old C57Bl6 wild type mice via the tail vein (0.1 mL per mouse) at a total RNA dose of 0.84 mg RNA per kg body weight. Fourteen days after dosing, the mice were sacrificed, and the left lateral, medial, and right lateral lobes of the liver were collected for preparation of DNA,

RNA, and protein, respectively. Blood was collected by exsanguination via cardiac puncture and collected onto BD microtainer (heparin coated). Samples were kept on wet ice for no longer than 30 minutes prior to centrifugation. Samples were centrifuged at 2,000 G for 10 minutes and the plasma transferred to 1.5 mL Cryotubes and stored at -80 °C.

Genomic DNA preparation and editing analysis by Next-Generation sequencing (NGS)

[00587] The left lateral lobe of the liver (100 mg) was homogenized using a Bead Ruptor (Omni International) in the digestion buffer supplied in the PureLink Genomic DNA Isolation Kit (Thermo Fisher Scientific). Genomic DNA was purified from the resulting homogenate using the MagMAX™ DNA Multi-Sample Ultra 2.0 Kit (Applied Biosystems) and quantified by measuring the absorbance at 260 nm. Genomic DNA purified from mice injected with PBS buffer alone was used as a control. The region of the HAO-1 gene targeted by each specific sgRNA was PCR amplified with Q5 high fidelity DNA polymerase and gene specific primers (Table 41) with adapters complementary to the barcoded primers used for next generation sequencing (NGS) for a total of 29 cycles.

Table 41: Primers used to amplify HAO1 guide target site and for NGS

Guide	Primer Set Name	Fwd with Miseq adapter	Rev with Miseq adapter
mH29-29 (spacer 29)	mHAO1-NGS-P3	GCTCTTCCGATCTNNNNNGTG ATGTCAATCGTCTGAGC	GCTCTTCCGATC TNNNNNTGTAG GTGGCTGAGTA CGTT

[00588] The product of this first PCR reaction was PCR amplified using the barcoded primers for NGS for a total of 10 cycles. The resulting product was subjected to NGS on an Illumina MiSeq instrument, and the results were processed using a custom Python script to generate the percentage of sequencing reads that contain insertions or deletions (INDELS) at the targeted site in the HAO-1 gene.

[00589] NGS analysis showed that editing of the groups dosed with LNP encapsulating MG29-1 mRNA and each of the four sgRNA ranged from 42% to 48% of total liver genomic DNA (**FIG. 99**). There were only slight differences between the groups, indicating that the sgRNA with the 20-nt spacer edited the target locus as efficiently as the 22-nt spacer, and that guide RNAs with chemistries 37 and 44 edited equally well.

RNA preparation and analysis by RT-ddPCR

[00590] The medial lobe of the liver was stored in RNAlater or RNAprotect (Qiagen) to preserve the integrity of the RNA prior to homogenization. A maximum of 10 mg of tissue was transferred to a 2 mL tube containing 1.4 mm ceramic beads and homogenized using the Bead Ruptor Elite

(OMNI International) following the soft tissue homogenization protocol, 5.00 m/s for 10-15 sec. Homogenized tissue was then processed using RNeasy Protect Kit (Qiagen) with an additional 45 minute on-column DNase I digestion treatment (Qiagen). The RNA products were quantified by measuring the absorbance at 260 nm. Isolated RNA samples then underwent an additional DNase treatment using the ezDNase™ Enzyme (Invitrogen) and reverse transcription using the SuperScript™ IV Vilo Master Mix (Invitrogen). The cDNA product was then quantified by measuring the absorbance at 260 nm.

[00591] HAO-1 mRNA levels were quantified using a custom ddPCR probe-based assay that was multiplexed with the housekeeping gene GAPDH. HAO-1 was labeled with a HEX probe while GAPDH was labeled with a FAM probe. Both probes were flanked by primers specific to the respective genes creating amplicons of about 115 bp. The template material for the assay was 100 ng of cDNA product from the process above mixed with 900 nM of each primer and 250 nM of probe per gene assay along with 10 µL of ddPCR Supermix for Probes (No dUTP) (Bio-Rad Laboratories) raised to a final volume of 20 µL with nuclease-free water. The PCR mix was parsed into thousands of oil droplets via the AutoDG ddPCR System (Bio-Rad Laboratories), ran in the C1000 Touch™ deep well thermal cycler (Bio-Rad Laboratories), and analyzed using the QX200 Droplet Reader (Bio-Rad Laboratories). Results were divided into four sections: negative droplets (no fluorescence), single positive droplets (HEX or FAM positive fluorescence), and double positive droplets (both HEX and FAM fluorescence). Using the QuantaSoft Software (Bio-Rad Laboratories), copies per µL of HAO1 and GAPDH were calculated for each sample. The ratio of HAO1 to GAPDH was compared for mice treated with mH29-29 against mice treated with buffer only. The GAPDH-normalized levels of HAO1 mRNA of the mH29-29 treated groups ranged from 52% to 70% of the control group (**FIG. 99**) and correlated with the editing efficiency as defined by INDEL %, indicating little or no effect of spacer length or chemistry upon guide RNA activity.

Table 42: Primers and probes used in the amplification HAO1 and GAPDH ddPCR assays

Target Gene	Probe Fluorescence	Probe Sequence	Fwd Primer Sequence	Rev primer sequence
HAO1	HEX	AGTG+GGTG+C+CA+GAAT+GTGAA	GGGGAGAAAGGTGTCAAGATGT	CTCACCAATGTCTTGTCGATGA
GAPDH	FAM	CATGACCA+CAGT+C+CATGCCATC	GCACCACCAACTGCTTAG	CCATCCACAGTCTTCTGGG

Notations for chemical modifications: + = Locked Nucleic Acid (LNA) base modification

Example 75 – Investigation of different mRNA:sgRNA ratios and LNP formulation procedures on editing efficiency in vivo in mice

[00592] In this *in vivo* mouse study, MG29-1 mRNA and sgRNA mH29-29-50b were formulated separately into LNP and then mixed at different mass ratios of mRNA and sgRNA prior to dosing. The same mRNA and sgRNA were also co-formulated in the same LNP at a 1:1 mass ratio then either kept at 4 °C and dosed within 48 h (fresh LNP), or frozen and stored at -80 °C, then thawed and dosed within 2 h (frozen LNP). The sequence of the mH29-29-50b sgRNA is shown in Table 43. This guide has a 20-nt spacer sequence and the chemical modifications designated chemistry 50.

Table 43: Sequences and chemical modifications of guide RNAs tested in vivo in mouse studies

Guide RNA	Sequence
mH29-29-50	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAU**CCUUAGGfAfGfAfAfAfUfGfCfC*fA*fAfA*fU*mC**
mH29-29-50b	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAU**CCUUAGGfAfGfAfAfAfUfGfCfC*fA*fA*mA**
mH29-29-51b	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAU**CCUUAGGAGAAAAUGCC*A*mA*mA**
mH29-29-52b	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAU**CCUUAGGAfGAfAAfAUfG*C*fCA*fA*mA**
mH29-29-53b	mG*mU*mU*mG*mA*mG*mA*mA*mU*mC*mG*mA*mA*mA*mG*mA*mU*mU*mC*mU*mC*mA*mA*mC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAU**CCUUAGGfAfGfAfAfAfUfGfCfC*fA*fA*mA**
mH29-29-54b	mG*mU*mU*mG*mA*mG*mA*mA*mU*mC*mG*mA*mA*mA*mG*mA*mU*mU*mC*mU*mC*mA*mA*mC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAU**CCUUAGGAGAAAAUGCC*A*mA*mA**

Notations for chemical modifications: m = 2'O-Methyl ribonucleotide (e.g mC = cytosine ribonucleotide with 2'-O Methyl in place of 2' hydroxyl); f = 2'Fluorine ribonucleotide (e.g fC = cytosine ribonucleotide with 2' fluorine in place of 2' hydroxyl); * = phosphorothioate bond. All other bases are native ribonucleotides. Backbone sequence in normal type. Spacer sequence in bold type.

Preparation of lipid nanoparticles

[00593] The LNP formulations used to deliver the MG29-1 mRNA and the guide RNA were prepared by the same procedure as in Example 74 with the following differences:

1. For different ratios of separately-formulated LNP, the mRNA and guide RNA LNP were mixed at 1:2, 1:1.5, 1:1, 1:0.75, and 1:0.5 mRNA:guide RNA mass ratios.

2. For the co-formulated LNP, the mRNA and the guide RNA were mixed prior to formulation at a 1:1 mass ratio and stored overnight at either 4 °C ("Fresh") or -80 °C ("Frozen").

Mouse dosing and harvesting

[00594] mRNA and sgRNA formulated in separate LNP were mixed at different mass ratios as above and injected intravenously into 7 week-old C57Bl6 wild type mice via the tail vein (0.1 mL per mouse) at a total RNA dose of 0.35 mg RNA per kg body weight. Co-formulated LNPs were injected similarly at the same dose. Seven days after dosing, the mice were sacrificed, and the left lateral, medial, and right lateral lobes of the liver were collected for preparation of DNA, RNA, and protein, respectively, by the same procedure as in Example 74. Terminal blood was collected by exsanguination via cardiac puncture as per the procedure of Example 74.

Genomic DNA preparation and editing analysis by NGS

[00595] Genomic DNA was prepared and analyzed by NGS as in Example 74.

RNA preparation and analysis by RT-ddPCR

[00596] RNA was prepared and analyzed by RT-ddPCR as in Example 74.

[00597] The results demonstrated that the ratio between the mRNA and guide in the separately-formulated LNPs does not greatly affect the editing efficiency (**FIG. 100**). Although there was somewhat higher editing efficiency at the middle three ratios (1:1.5, 1:1, and 1:0.75) than at the extremes (1:2 and 1:0.5), the differences were within the variation of the experiment. The co-formulated LNPs resulted in higher editing than the separate formulations, with the fresh and frozen LNPs performing equally well. Similar to the editing efficiency, there was little difference in the amount of HAO-1 mRNA present in the livers of mice treated with separately-formulated LNPs at different MG29-1 mRNA:guide ratios (**FIG. 100**). HAO-1 mRNA levels in the mice treated with the co-formulated LNPs were reduced more than 50%, similar to the groups that received the separately-formulated LNPs. In conclusion, these results demonstrate that the MG29-1 mRNA and sgRNA with chemistry 50 can be co-formulated into LNP with no reduction in editing potency and that a range of mRNA to sgRNA ratios between 1:2 to 1:0.5 can be used.

Example 76 – Investigation of the impact of different guide chemistries upon editing efficiency *in vivo*

Preparation of lipid nanoparticles

[00598] The LNP formulations used to deliver the MG29-1 mRNA and the guide RNA were prepared by the same procedure as in Example 74. Guide RNA sequences and chemistries are listed in Table 40.

Mouse dosing and harvesting

[00599] mRNA and sgRNA were formulated separately into LNP then mixed at 1:1 mass ratios as in Example 74 and injected intravenously into 7 week-old C57Bl6 wild type mice via the tail vein (0.1 mL per mouse) at a total RNA dose of 0.25 mg RNA per kg body weight. Seven days after dosing, the mice were sacrificed, and the left lateral, medial, and right lateral lobes of the liver were collected for preparation of DNA, RNA, and protein, respectively, by the same procedure as in Example 74. Terminal blood was collected by exsanguination via cardiac puncture as per the procedure of Example 74.

Genomic DNA preparation and editing analysis by NGS

[00600] Genomic DNA was prepared and analyzed by NGS as in Example 74.

RNA preparation and analysis by RT-ddPCR

[00601] RNA was prepared and analyzed by RT-ddPCR as in Example 74.

[00602] All guides designated with a "b" in the guide name (e.g. mH29-29-50b) contain 20-nt spacers. Guides mH29-29-37 and mH29-29-50 contain 22-nt spacers. Guides with chemistries 51, 52, 53, and 54 contain the same nucleotide sequence as the guide with chemistry 50 but have differences in the chemical modifications at specific regions of the guide. Chemistry 51 is identical to chemistry 50 with the exception of the removal of the 2'-fluoro modifications in the spacer. Chemistry 52 is identical to chemistry 50 with the exception of the removal of half of the 2'-flouro modifications in the spacer. Chemistry 53 is identical to chemistry 50 with the exception of an additional 21 phosphorothioate linkages and 21 2'-O methyl modifications. Chemistry 53 is identical to chemistry 50 with the exception of an additional 21 phosphorothioate linkages and 21 2'-O methyl modifications in the 5' stem loop. Chemistry 54 is identical to chemistry 50 with the exception of an additional 21 phosphorothioate linkages and 21 2'-O methyl modifications in the 5' stem loop and the removal of the 2'-fluoro modifications in the spacer.

[00603] The results indicate that guide RNAs with chemistry 50 (50b) and chemistry 52 (52b) had the highest editing efficiency (**FIG. 101**). The introduction of INDELS in the HAO-1 gene is expected to reduce the level of HAO-1 mRNA through nonsense-mediated mRNA decay, due to reading frame shifts and the resulting premature stop codons. Treatment with the 22-nt guide with chemistry 50 (mH29-29-50b) resulted in the lowest level (largest reduction) of HAO-1 mRNA, consistent with this mechanism (**FIG. 101**). Chemistry 51 (51b) was 2-fold less potent than chemistry 50 (50b), indicating that the 2'-fluoro modifications in the spacer contributed significantly to potency. Chemistry 52 (52b) had similar or slightly improved editing potency compared to chemistry 50 (50b), indicating that in the context of this guide spacer sequence, it

was possible to remove half of the 2-fluoro modifications in the spacer without significantly reducing potency. Chemistry 53 (53b) had about 2-fold lower editing potency compared to chemistry 50 (50b), indicating that the addition of 2'-O-methyl and PS linkages in the 5' stem-loop had a negative impact on potency, which is surprising given that these additional modifications were expected to improve guide stability. Chemistry 54 (54b) had about 4 -fold lower potency compared to chemistry 50 (50b) and about 2-fold lower editing potency compared to chemistry 53 (53b), confirming that the additional 2'O-methyl and PS bases in the 5' stem loop and the removal of the 2'-fluoro bases in the spacer both had an additive negative effect on guide potency.

Example 77 – Gene editing outcomes at the DNA level for APO-A1 in Hepa1-6 Cells

[00604] Nucleofection of MG29-1 RNPs (126 pmol protein/160 pmol guide) was performed into Hepa1-6 cells (200,000) using the Lonza 4D electroporator. Cells were harvested and genomic DNA prepared five days post-transfection. PCR primers appropriate for use in NGS-based DNA sequencing were generated, optimized, and used to amplify the individual target sequences for each guide RNA. The amplicons were sequenced on an Illumina MiSeq machine and analyzed with a proprietary Python script to measure gene editing (**FIG. 102**).

Table 43A: Sequences of Guide RNAs and Sequences Targeted for Example 77

Type	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting mouse APO-A1	5847	MG29-1-mouse APO-A1-sgRNA-A1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrUrArArUrGrUrGrUrArUrGrUrGrGrArUrGrCrGrG/AltR2/
MG29-1 sgRNA targeting mouse APO-A1	5848	MG29-1-mouse APO-A1-sgRNA-B1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArUrCrCrUrCrCrUrCrCrUrUrGrGrGrCrCrArArCrA/AltR2/
MG29-1 sgRNA targeting mouse APO-A1	5849	MG29-1-mouse APO-A1-sgRNA-C1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrArCrUrUrCrArGrCrUrGrUrUrGrGrCrCrArA/AltR2/
MG29-1 sgRNA targeting mouse APO-A1	5850	MG29-1-mouse APO-A1-sgRNA-D1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrCrGrCrArUrCrCrArCrArUrArCrArCrArUrUrArG/AltR2/
MG29-1 sgRNA targeting mouse APO-A1	5851	MG29-1-mouse APO-A1-sgRNA-E1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrCrArUrUrGrGrGrGrUrUrCrArU/AltR2/
MG29-1 sgRNA targeting mouse APO-A1	5852	MG29-1-mouse APO-A1-sgRNA-F1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrGrArCrCrGrCrArUrGrCrGrCrArCrArCrGrU/AltR2/
MG29-1 sgRNA targeting mouse APO-A1	5853	MG29-1-mouse APO-A1-sgRNA-G1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrGrCrCrArArGrUrGrUrCrUrUrCrArGrGrUrGrGrG/AltR2/
MG29-1 sgRNA targeting mouse APO-A1	5854	MG29-1-mouse APO-A1-sgRNA-H1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrCrCrCrUrGrGrUrGrGrUrArCrUrCrGrUrUrC/AltR2/

Type	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting mouse APO-A1	5855	MG29-1-mouse APO-A1-sgRNA-A2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrCrCrUrGrGrUrGrUrGrGrUrArCrUrCrGrUrUrCrA/AltR2/
MG29-1 sgRNA targeting mouse APO-A1	5856	MG29-1-mouse APO-A1-sgRNA-B2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArUrUrUrCrUrUrCrUrGrGrArArUrUrCrGrUrCrCrA/AltR2/
MG29-1 sgRNA targeting mouse APO-A1	5857	MG29-1-mouse APO-A1-sgRNA-C2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrUrGrGrArArUrUrCrGrUrCrCrArGrGrUrArGrG/AltR2/
MG29-1 sgRNA targeting mouse APO-A1	5858	MG29-1-mouse APO-A1-sgRNA-D2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrUrUrCrCrUrCrCrUrUrGrGrGrCrCrArArCrA/AltR2/
MG29-1 sgRNA targeting mouse APO-A1	5859	MG29-1-mouse APO-A1-sgRNA-E2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrCrUrCrCrUrUrGrGrGrCrCrArArArG/AltR2/
MG29-1 sgRNA targeting mouse APO-A1	5860	MG29-1-mouse APO-A1-sgRNA-F2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrArGrGrUrUrArUrCrCrCrArGrArArGrUrCrCrC/AltR2/
DNA Sequence of APO-A1 Target Site	**5861**	**MG29-1-mouse APO-A1-target site-A1**	GCTAATGTGTATGTGGATGCGG
DNA Sequence of APO-A1 Target Site	**5862**	**MG29-1-mouse APO-A1-target site-B1**	AATCCTCCTCCTTGGGCCAACA
DNA Sequence of APO-A1 Target Site	**5863**	**MG29-1-mouse APO-A1-target site-C1**	CTTACTTCAGCTGTTGGCCCAA
DNA Sequence of APO-A1 Target Site	**5864**	**MG29-1-mouse APO-A1-target site-D1**	ACCGCATCCACATACACATTAG
DNA Sequence of APO-A1 Target Site	**5865**	**MG29-1-mouse APO-A1-target site-E1**	TCCCATTGGGACTGGGGTTCAT
DNA Sequence of APO-A1 Target Site	**5866**	**MG29-1-mouse APO-A1-target site-F1**	GCGACCGCATGCGCACACACGT
DNA Sequence of APO-A1 Target Site	5867	MG29-1-mouse APO-A1-target site-G1	TCGCCAAGTGTCTTCAGGTGGG
DNA Sequence of APO-A1 Target Site	**5868**	**MG29-1-mouse APO-A1-target site-H1**	GGCCCTGGTGTGGTACTCGTTC
DNA Sequence of APO-A1 Target Site	**5869**	**MG29-1-mouse APO-A1-target site-A2**	GCCCTGGTGTGGTACTCGTTCA
DNA Sequence of APO-A1 Target Site	5870	MG29-1-mouse APO-A1-target site-B2	CATTTCTTCTGGAATTCGTCCA
DNA Sequence of APO-A1 Target Site	**5871**	**MG29-1-mouse APO-A1-target site-C2**	TTCTGGAATTCGTCCAGGTAGG
DNA Sequence of APO-A1 Target Site	**5872**	**MG29-1-mouse APO-A1-target site-D2**	ACTTCCTCTAGGTCCTTGTTCA
DNA Sequence of APO-A1 Target Site	**5873**	**MG29-1-mouse APO-A1-target site-E2**	TTTCTCCAGGTTATCCCAGAAG

Type	SEQ ID NO	Guide Name	SEQUENCE
DNA Sequence of APO-A1 Target Site	5874	**MG29-1-mouse APO-A1-target site-F2**	TCCAGGTTATCCCAGAAGTCCC
r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *; between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur; AltR1 and AltR2 refer to IDT technologies' proprietary 5' and 3' AltR modifications			

Example 78 – Gene editing outcomes at the DNA level for ANGPTL3 in Hepa1-6 Cells

[00605] Nucleofection of MG29-1 RNPs (126 pmol protein/160 pmol guide) was performed into Hepa1-6 cells (200,000) using the Lonza 4D electroporator. Cells were harvested and genomic DNA prepared five days post-transfection. PCR primers appropriate for use in NGS-based DNA sequencing were generated, optimized, and used to amplify the individual target sequences for each guide RNA. The amplicons were sequenced on an Illumina MiSeq machine and analyzed with a proprietary Python script to measure gene editing (**FIG. 103**).

Table 43B: Sequences of Guide RNAs and Sequences Targeted for Example 78

Type	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting mouse ANGPTL3	5875	MG29-1-mouse ANGPTL3-sgRNA-A1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrUrUrGrUrUrCrCrUrUrUrArGrUrArArUrUrGrCrA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5876	MG29-1-mouse ANGPTL3-sgRNA-B1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrUrGrUrUrCrCrUrUrUrArGrUrArArUrUrGrCrArU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5877	MG29-1-mouse ANGPTL3-sgRNA-C1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrUrUrCrCrUrUrUrArGrUrArArUrUrGrCrArUrC/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5878	MG29-1-mouse ANGPTL3-sgRNA-D1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrArArUrUrGrCrArUrCrCrArGrArGrUrGrGrArUrC/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5879	MG29-1-mouse ANGPTL3-sgRNA-E1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrCrArUrUrUrGrArUrUrCrUrGrCrArCrCrUrUrCrA/AltR2/
MG29-1 sgRNA targeting mouse	5880	MG29-1-mouse ANGPTL3-sgRNA-F1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrUrCrUrGrCrArCrCrUrUrCrArArA/AltR2/

Type	SEQ ID NO	Guide Name	SEQUENCE
ANGPTL3			
MG29-1 sgRNA targeting mouse ANGPTL3	5881	MG29-1-mouse ANGPTL3-sgRNA-G1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrArUrGrUrUrGrGrArUrGrArUrGrUrCrArArArArU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5882	MG29-1-mouse ANGPTL3-sgRNA-H1	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrCrGrArArUrGrGrCrCrUrCrCrUrGrCrArGrCrUrG/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5883	MG29-1-mouse ANGPTL3-sgRNA-A2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrGrArArUrGrGrCrCrUrCrCrUrGrCrArGrCrUrG/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5884	MG29-1-mouse ANGPTL3-sgRNA-B2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrCrCrArUrArArGrArCrUrArArGrGrGrArCrArArA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5885	MG29-1-mouse ANGPTL3-sgRNA-C2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrArUrArArGrArCrUrArArGrGrGrArCrArArArU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5886	MG29-1-mouse ANGPTL3-sgRNA-D2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrArGrCrUrCrArArCrArUrArUrUrUrGrArUrCrA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5887	MG29-1-mouse ANGPTL3-sgRNA-E2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrCrArGrUrCrUrUrUrUrUrArUrGrArCrCrUrArUrC/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5888	MG29-1-mouse ANGPTL3-sgRNA-F2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArUrGrArCrCrUrArUrCrArCrUrUrCrGrArArCrCrA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5889	MG29-1-mouse ANGPTL3-sgRNA-G2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrGrArCrCrUrArUrCrArCrUrUrCrGrArArCrCrArA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5890	MG29-1-mouse ANGPTL3-sgRNA-H2	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrArCrCrUrArUrCrArCrUrUrCrGrArArCrCrArArU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5891	MG29-1-mouse ANGPTL3-sgRNA-A3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArGrCrArGrCrUrArArCrCrArArCrUrUrArA/AltR2/

Type	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting mouse ANGPTL3	5892	MG29-1-mouse ANGPTL3-sgRNA-B3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArCrUrUrArCrUrUrUrGrArGrUrGrArUrGrUrUrArC/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5893	MG29-1-mouse ANGPTL3-sgRNA-C3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrUrGrArUrGrUrUrArCrUrUrCrUrGrGrGrUrGrCrU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5894	MG29-1-mouse ANGPTL3-sgRNA-D3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrUrUrCrArGrUrUrCrUrArCrUrGrArCrArUrGrUrU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5895	MG29-1-mouse ANGPTL3-sgRNA-E3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArArCrUrUrGrUrArGrUrGrUrArGrArUrGrUrArGrU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5896	MG29-1-mouse ANGPTL3-sgRNA-F3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArCrUrUrGrUrArGrUrGrUrArGrArUrGrUrArGrU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5897	MG29-1-mouse ANGPTL3-sgRNA-G3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrUrUrGrUrArGrUrGrUrArGrArUrGrUrArGrUrC/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5898	MG29-1-mouse ANGPTL3-sgRNA-H3	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrUrCrUrUrCrUrUrUrGrArUrUrUrCrArUrUrGrGrU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5899	MG29-1-mouse ANGPTL3-sgRNA-A4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrCrUrUrCrUrUrUrGrArUrUrUrCrArUrUrGrGrUrU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5900	MG29-1-mouse ANGPTL3-sgRNA-B4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrUrUrCrArUrUrGrGrUrUrCrGrArArGrUrGrArUrA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5901	MG29-1-mouse ANGPTL3-sgRNA-C4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrUrGrGrUrUrCrGrArArGrUrGrArUrArGrGrUrCrA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5902	MG29-1-mouse ANGPTL3-sgRNA-D4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrCrUrUrArGrUrCrUrUrArUrGrGrArCrArArArA/AltR2/
MG29-1 sgRNA targeting	5903	MG29-1-mouse ANGPTL3-sgRNA-E4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrUrCrCrArUrGrArCrCrCrArGrCrUrGrCrArGrGrA/AltR2/

Type	SEQ ID NO	Guide Name	SEQUENCE
mouse ANGPTL3			
MG29-1 sgRNA targeting mouse ANGPTL3	5904	MG29-1-mouse ANGPTL3-sgRNA-F4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArCrArUrCrArUrCrCrArArCrArUrArGrCrArArArU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5905	MG29-1-mouse ANGPTL3-sgRNA-G4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrArUrCrArUrCrCrArArCrArUrArGrCrArArArUrC/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5906	MG29-1-mouse ANGPTL3-sgRNA-H4	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrCrUrCrUrGrArArGrGrUrGrCrArGrArArUrCrArA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5907	MG29-1-mouse ANGPTL3-sgRNA-A5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrUrCrUrGrArArGrGrUrGrCrArGrArArUrCrArArA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5908	MG29-1-mouse ANGPTL3-sgRNA-B5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrArGrArArCrArGrCrArArGrArCrArArCrArGrCrA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5909	MG29-1-mouse ANGPTL3-sgRNA-C5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrArArCrArGrCrArArGrArCrArArCrArGrCrArU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5910	MG29-1-mouse ANGPTL3-sgRNA-D5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrArUrUrUrCrUrUrUrUrArUrCrUrGrCrArUrGrUrG/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5911	MG29-1-mouse ANGPTL3-sgRNA-E5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArUrUrUrCrUrUrUrUrArUrCrUrGrCrArUrGrUrGrC/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5912	MG29-1-mouse ANGPTL3-sgRNA-F5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrArUrCrUrGrCrArUrGrUrGrCrUrGrUrUrGrA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5913	MG29-1-mouse ANGPTL3-sgRNA-G5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrCrUrGrCrArUrGrUrGrCrUrGrUrUrGrArCrUrUrA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5914	MG29-1-mouse ANGPTL3-sgRNA-H5	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrGrCrArUrGrUrGrCrUrGrUrUrGrArCrUrUrArA/AltR2/

Type	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting mouse ANGPTL3	5915	MG29-1-mouse ANGPTL3-sgRNA-A6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArCrUrGrUrUrCrUrUrCrCrArCrArCrUrCrUrGrGrA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5916	MG29-1-mouse ANGPTL3-sgRNA-B6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArUrUrCrUrUrUrUrArUrCrArGrCrUrCrArGrArArA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5917	MG29-1-mouse ANGPTL3-sgRNA-C6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrCrArGrCrUrCrArGrArArArGrArCrUrGrGrUrArU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5918	MG29-1-mouse ANGPTL3-sgRNA-D6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrArGrCrUrCrArGrArArArGrArCrUrGrGrUrArU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5919	MG29-1-mouse ANGPTL3-sgRNA-E6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrUrArArArUrCrArArGrArGrCrArCrCrArArGrA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5920	MG29-1-mouse ANGPTL3-sgRNA-F6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrUrUrCrGrUrUrCrArGrUrUrCrArGrGrUrUrGrArArGrArGrG/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5921	MG29-1-mouse ANGPTL3-sgRNA-G6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrUrUrUrCrGrUrUrCrArGrUrUrCrArGrGrUrUrGrArArGrArGrG/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5922	MG29-1-mouse ANGPTL3-sgRNA-H6	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrUrCrArGrUrUrGrArArGrArGrGrGrGrArGrUrA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5923	MG29-1-mouse ANGPTL3-sgRNA-A7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArArGrArArArUrCrArArArGrArGrCrArCrCrArArGrG/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5924	MG29-1-mouse ANGPTL3-sgRNA-B7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrArGrGrUrUrCrUrUrGrArArUrArArCrCrArGrU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5925	MG29-1-mouse ANGPTL3-sgRNA-C7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrArGrGrGrUrUrCrUrUrGrArArUrArArCrCrArGrUrC/AltR2/
MG29-1 sgRNA targeting	5926	MG29-1-mouse ANGPTL3-sgRNA-D7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArArCrArGrArGrGrCrGrArArCrArUrArCrArArGrU/AltR2/

Type	SEQ ID NO	Guide Name	SEQUENCE
mouse ANGPTL3			
MG29-1 sgRNA targeting mouse ANGPTL3	5927	MG29-1-mouse ANGPTL3-sgRNA-E7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrGrUrCrUrArCrUrGrUrGrArUrArCrCrCrArArUrC/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5928	MG29-1-mouse ANGPTL3-sgRNA-F7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArUrGrGrGrUrUrUrArCrCrUrGrArUrUrGrGrGrUrA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5929	MG29-1-mouse ANGPTL3-sgRNA-G7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrUrGrArUrUrGrGrGrUrArUrCrArCrArGrUrArGrA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5930	MG29-1-mouse ANGPTL3-sgRNA-H7	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrGrUrUrUrArArUrArGrUrGrUrArCrArCrGrCrC/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5931	MG29-1-mouse ANGPTL3-sgRNA-A8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrArGrUrGrUrArCrArCrGrCrCrArCrUrCrUrGrUrArU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5932	MG29-1-mouse ANGPTL3-sgRNA-B8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrArUrCrGrArGrCrCrUrCrCrCrArArArGrCrCrCrU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5933	MG29-1-mouse ANGPTL3-sgRNA-C8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrGrUrArGrUrUrUrUrCrCrCrArUrGrUrUrUrCrGrUrU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5934	MG29-1-mouse ANGPTL3-sgRNA-D8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrArGrUrUrUrUrCrCrCrArUrGrUrUrUrCrGrUrU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5935	MG29-1-mouse ANGPTL3-sgRNA-E8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrCrArUrGrUrUrUrCrGrUrUrUrGrArArGrUrCrCrUrG/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5936	MG29-1-mouse ANGPTL3-sgRNA-F8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrArUrGrUrUrUrCrGrUrUrUrGrArArGrUrCrCrUrG/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5937	MG29-1-mouse ANGPTL3-sgRNA-G8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrUrGrArArGrUrCrCrUrGrUrUrUrCrGrUrUrUrGrArGrCrCrArUrCrU/AltR2/

Type	SEQ ID NO	Guide Name	SEQUENCE
MG29-1 sgRNA targeting mouse ANGPTL3	5938	MG29-1-mouse ANGPTL3-sgRNA-H8	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrGrGrUrGrUrUrGrArArUrUrArArUrGrUrCrCrArUrG/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5939	MG29-1-mouse ANGPTL3-sgRNA-A9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrCrUrCrUrArArCrUrUrUrUrUrUrCrUrUrUrArGrG/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5940	MG29-1-mouse ANGPTL3-sgRNA-B9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrUrCrUrArArCrUrUrUrUrUrUrCrUrUrUrArGrGrA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5941	MG29-1-mouse ANGPTL3-sgRNA-C9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrUrUrUrArGrGrArArGrArArArUrUrUrUrUrGrGrUrUrG/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5942	MG29-1-mouse ANGPTL3-sgRNA-D9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrUrUrArGrGrArGrArGrArArUrUrUrUrUrGrGrUrUrGrG/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5943	MG29-1-mouse ANGPTL3-sgRNA-E9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrUrArGrGrArGrArGrArArUrUrUrUrUrGrGrUrUrGrGrG/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5944	MG29-1-mouse ANGPTL3-sgRNA-F9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrArGrGrArGrArGrArArUrUrUrUrUrGrGrUrUrGrGrC/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5945	MG29-1-mouse ANGPTL3-sgRNA-G9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrArGrArArUrUrUrUrGrGrUrUrGrGrGrCrCrUrArG/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5946	MG29-1-mouse ANGPTL3-sgRNA-H9	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrUrUrGrGrGrCrCrUrArGrArGrArArGrArUrCrUrA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5947	MG29-1-mouse ANGPTL3-sgRNA-A10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrUrGrGrGrCrCrUrArGrArGrArArGrArUrCrUrArU/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5948	MG29-1-mouse ANGPTL3-sgRNA-B10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrGrArCrUrCrGrArGrCrUrArCrArArGrArCrUrGrG/AltR2/
MG29-1 sgRNA targeting	5949	MG29-1-mouse ANGPTL3-sgRNA-C10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrGrArCrUrCrGrArGrCrUrArCrArArGrArCrUrGrGrA/AltR2/

Type	SEQ ID NO	Guide Name	SEQUENCE
mouse ANGPTL3			
MG29-1 sgRNA targeting mouse ANGPTL3	5950	MG29-1-mouse ANGPTL3-sgRNA-D10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrCrUrGrGrGrCrArGrUrCrArCrGrArArArCrCrArA/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5951	MG29-1-mouse ANGPTL3-sgRNA-E10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrGrArCrUrGrCrCrCrArGrGrUrGrArArArGrGrArG/AltR2/
MG29-1 sgRNA targeting mouse ANGPTL3	5952	MG29-1-mouse ANGPTL3-sgRNA-F10	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrUrCrUrUrGrUrArGrCrUrCrGrArGrUrCrGrUrA/AltR2/
DNA Sequence of ANGPTL3 Target Site	5953	MG29-1-mouse ANGPTL3-target site-A1	TGTTGTTCCTTTAGTAATTGCA
DNA Sequence of ANGPTL3 Target Site	5954	MG29-1-mouse ANGPTL3-target site-B1	GTTGTTCCTTTAGTAATTGCAT
DNA Sequence of ANGPTL3 Target Site	**5955**	**MG29-1-mouse ANGPTL3-target site-C1**	TTGTTCCTTTAGTAATTGCATC
DNA Sequence of ANGPTL3 Target Site	**5956**	**MG29-1-mouse ANGPTL3-target site-D1**	GTAATTGCATCCAGAGTGGATC
DNA Sequence of ANGPTL3 Target Site	**5957**	**MG29-1-mouse ANGPTL3-target site-E1**	ATCATTTGATTCTGCACCTTCA
DNA Sequence of ANGPTL3 Target Site	**5958**	**MG29-1-mouse ANGPTL3-target site-F1**	ATTCTGCACCTTCAGAGCCAAA
DNA Sequence of ANGPTL3 Target Site	**5959**	**MG29-1-mouse ANGPTL3-target site-G1**	CTATGTTGGATGATGTCAAAAT
DNA Sequence of ANGPTL3 Target Site	**5960**	**MG29-1-mouse ANGPTL3-target site-H1**	AGCGAATGGCCTCCTGCAGCTG

Type	SEQ ID NO	Guide Name	SEQUENCE
DNA Sequence of ANGPTL3 Target Site	**5961**	**MG29-1-mouse ANGPTL3-target site-A2**	GCGAATGGCCTCCTGCAGCTGG
DNA Sequence of ANGPTL3 Target Site	**5962**	**MG29-1-mouse ANGPTL3-target site-B2**	GTCCATAAGACTAAGGGACAAA
DNA Sequence of ANGPTL3 Target Site	**5963**	**MG29-1-mouse ANGPTL3-target site-C2**	TCCATAAGACTAAGGGACAAAT
DNA Sequence of ANGPTL3 Target Site	5964	MG29-1-mouse ANGPTL3-target site-D2	AGAAGCTCAACATATTTGATCA
DNA Sequence of ANGPTL3 Target Site	5965	MG29-1-mouse ANGPTL3-target site-E2	ATCAGTCTTTTTATGACCTATC
DNA Sequence of ANGPTL3 Target Site	5966	MG29-1-mouse ANGPTL3-target site-F2	TATGACCTATCACTTCGAACCA
DNA Sequence of ANGPTL3 Target Site	5967	MG29-1-mouse ANGPTL3-target site-G2	ATGACCTATCACTTCGAACCAA
DNA Sequence of ANGPTL3 Target Site	**5968**	**MG29-1-mouse ANGPTL3-target site-H2**	TGACCTATCACTTCGAACCAAT
DNA Sequence of ANGPTL3 Target Site	**5969**	**MG29-1-mouse ANGPTL3-target site-A3**	GAGGAGCAGCTAACCAACTTAA
DNA Sequence of ANGPTL3 Target Site	**5970**	**MG29-1-mouse ANGPTL3-target site-B3**	TACTTACTTTGAGTGATGTTAC
DNA Sequence of ANGPTL3 Target Site	**5971**	**MG29-1-mouse ANGPTL3-target site-C3**	AGTGATGTTACTTCTGGGTGCT
DNA Sequence of	**5972**	**MG29-1-mouse ANGPTL3-target site-D3**	AGTTCAGTTCTACTGACATGTT

Type	SEQ ID NO	Guide Name	SEQUENCE
ANGPTL3 Target Site			
DNA Sequence of ANGPTL3 Target Site	**5973**	**MG29-1-mouse ANGPTL3-target site-E3**	TAACTTGTAGTGTAGATGTAGT
DNA Sequence of ANGPTL3 Target Site	**5974**	**MG29-1-mouse ANGPTL3-target site-F3**	AACTTGTAGTGTAGATGTAGTT
DNA Sequence of ANGPTL3 Target Site	**5975**	**MG29-1-mouse ANGPTL3-target site-G3**	ACTTGTAGTGTAGATGTAGTTC
DNA Sequence of ANGPTL3 Target Site	5976	MG29-1-mouse ANGPTL3-target site-H3	CCTCTTCTTTGATTTCATTGGT
DNA Sequence of ANGPTL3 Target Site	5977	MG29-1-mouse ANGPTL3-target site-A4	CTCTTCTTTGATTTCATTGGTT
DNA Sequence of ANGPTL3 Target Site	5978	MG29-1-mouse ANGPTL3-target site-B4	ATTTCATTGGTTCGAAGTGATA
DNA Sequence of ANGPTL3 Target Site	**5979**	**MG29-1-mouse ANGPTL3-target site-C4**	ATTGGTTCGAAGTGATAGGTCA
DNA Sequence of ANGPTL3 Target Site	**5980**	**MG29-1-mouse ANGPTL3-target site-D4**	TCCCTTAGTCTTATGGACAAAA
DNA Sequence of ANGPTL3 Target Site	**5981**	**MG29-1-mouse ANGPTL3-target site-E4**	AGTCCATGACCCAGCTGCAGGA
DNA Sequence of ANGPTL3 Target Site	**5982**	**MG29-1-mouse ANGPTL3-target site-F4**	GACATCATCCAACATAGCAAAT
DNA Sequence of ANGPTL3 Target Site	**5983**	**MG29-1-mouse ANGPTL3-target site-G4**	ACATCATCCAACATAGCAAATC

Type	SEQ ID NO	Guide Name	SEQUENCE
DNA Sequence of ANGPTL3 Target Site	**5984**	**MG29-1-mouse ANGPTL3-target site-H4**	GGCTCTGAAGGTGCAGAATCAA
DNA Sequence of ANGPTL3 Target Site	**5985**	**MG29-1-mouse ANGPTL3-target site-A5**	GCTCTGAAGGTGCAGAATCAAA
DNA Sequence of ANGPTL3 Target Site	**5986**	**MG29-1-mouse ANGPTL3-target site-B5**	GTAGAACAGCAAGACAACAGCA
DNA Sequence of ANGPTL3 Target Site	**5987**	**MG29-1-mouse ANGPTL3-target site-C5**	TAGAACAGCAAGACAACAGCAT
DNA Sequence of ANGPTL3 Target Site	5988	MG29-1-mouse ANGPTL3-target site-D5	CTATTTCTTTTATCTGCATGTG
DNA Sequence of ANGPTL3 Target Site	**5989**	**MG29-1-mouse ANGPTL3-target site-E5**	TATTTCTTTTATCTGCATGTGC
DNA Sequence of ANGPTL3 Target Site	**5990**	**MG29-1-mouse ANGPTL3-target site-F5**	TTTTATCTGCATGTGCTGTTGA
DNA Sequence of ANGPTL3 Target Site	5991	MG29-1-mouse ANGPTL3-target site-G5	ATCTGCATGTGCTGTTGACTTA
DNA Sequence of ANGPTL3 Target Site	**5992**	**MG29-1-mouse ANGPTL3-target site-H5**	TCTGCATGTGCTGTTGACTTAA
DNA Sequence of ANGPTL3 Target Site	**5993**	**MG29-1-mouse ANGPTL3-target site-A6**	TACTGTTCTTCCACACTCTGGA
DNA Sequence of ANGPTL3 Target Site	5994	MG29-1-mouse ANGPTL3-target site-B6	TATTCTTTTATCAGCTCAGAAA
DNA Sequence of	5995	MG29-1-mouse ANGPTL3-target site-C6	ATCAGCTCAGAAAGACTGGTAT

Type	SEQ ID NO	Guide Name	SEQUENCE
ANGPTL3 Target Site			
DNA Sequence of ANGPTL3 Target Site	5996	MG29-1-mouse ANGPTL3-target site-D6	TCAGCTCAGAAAGACTGGTATT
DNA Sequence of ANGPTL3 Target Site	**5997**	**MG29-1-mouse ANGPTL3-target site-E6**	TTCTAAATCAAGAGCACCAAGA
DNA Sequence of ANGPTL3 Target Site	5998	MG29-1-mouse ANGPTL3-target site-F6	CTGTTTCGTTCAGTTGAAGAGG
DNA Sequence of ANGPTL3 Target Site	**5999**	**MG29-1-mouse ANGPTL3-target site-G6**	TGTTTCGTTCAGTTGAAGAGGG
DNA Sequence of ANGPTL3 Target Site	6000	MG29-1-mouse ANGPTL3-target site-H6	GTTCAGTTGAAGAGGGGGAGTA
DNA Sequence of ANGPTL3 Target Site	6001	MG29-1-mouse ANGPTL3-target site-A7	GAAGAAAGAGAATTTTCTGAGG
DNA Sequence of ANGPTL3 Target Site	6002	MG29-1-mouse ANGPTL3-target site-B7	CTGAGGGTTCTTGAATACCAGT
DNA Sequence of ANGPTL3 Target Site	**6003**	**MG29-1-mouse ANGPTL3-target site-C7**	TGAGGGTTCTTGAATACCAGTC
DNA Sequence of ANGPTL3 Target Site	**6004**	**MG29-1-mouse ANGPTL3-target site-D7**	TAACAGAGGCGAACATACAAGT
DNA Sequence of ANGPTL3 Target Site	**6005**	**MG29-1-mouse ANGPTL3-target site-E7**	ATGTCTACTGTGATACCCAATC
DNA Sequence of ANGPTL3 Target Site	**6006**	**MG29-1-mouse ANGPTL3-target site-F7**	CATGGGTTTACCTGATTGGGTA

Type	SEQ ID NO	Guide Name	SEQUENCE
DNA Sequence of ANGPTL3 Target Site	**6007**	**MG29-1-mouse ANGPTL3-target site-G7**	CCTGATTGGGTATCACAGTAGA
DNA Sequence of ANGPTL3 Target Site	**6008**	**MG29-1-mouse ANGPTL3-target site-H7**	TTGGTTTAATAGTGTACACGCC
DNA Sequence of ANGPTL3 Target Site	**6009**	**MG29-1-mouse ANGPTL3-target site-A8**	ATAGTGTACACGCCACTTGTAT
DNA Sequence of ANGPTL3 Target Site	**6010**	**MG29-1-mouse ANGPTL3-target site-B8**	CCATCGAGCCTCCCAAAGCCCT
DNA Sequence of ANGPTL3 Target Site	6011	MG29-1-mouse ANGPTL3-target site-C8	CGTAGTTTTCCCATGTTTCGTT
DNA Sequence of ANGPTL3 Target Site	6012	MG29-1-mouse ANGPTL3-target site-D8	GTAGTTTTCCCATGTTTCGTTG
DNA Sequence of ANGPTL3 Target Site	6013	MG29-1-mouse ANGPTL3-target site-E8	CCCATGTTTCGTTGAAGTCCTG
DNA Sequence of ANGPTL3 Target Site	**6014**	**MG29-1-mouse ANGPTL3-target site-F8**	CCATGTTTCGTTGAAGTCCTGT
DNA Sequence of ANGPTL3 Target Site	**6015**	**MG29-1-mouse ANGPTL3-target site-G8**	GTTGAAGTCCTGTGAGCCATCT
DNA Sequence of ANGPTL3 Target Site	**6016**	**MG29-1-mouse ANGPTL3-target site-H8**	CGGTGTTGAATTAATGTCCATG
DNA Sequence of ANGPTL3 Target Site	6017	MG29-1-mouse ANGPTL3-target site-A9	CCCTCTAACTTTTTTCTTTAGG
DNA Sequence of	6018	MG29-1-mouse ANGPTL3-target site-B9	CCTCTAACTTTTTTCTTTAGGA

Type	SEQ ID NO	Guide Name	SEQUENCE
ANGPTL3 Target Site			
DNA Sequence of ANGPTL3 Target Site	6019	MG29-1-mouse ANGPTL3-target site-C9	TTCTTTAGGAGAATTTTGGTTG
DNA Sequence of ANGPTL3 Target Site	6020	MG29-1-mouse ANGPTL3-target site-D9	TCTTTAGGAGAATTTTGGTTGG
DNA Sequence of ANGPTL3 Target Site	6021	MG29-1-mouse ANGPTL3-target site-E9	CTTTAGGAGAATTTTGGTTGGG
DNA Sequence of ANGPTL3 Target Site	6022	MG29-1-mouse ANGPTL3-target site-F9	TTTAGGAGAATTTTGGTTGGGC
DNA Sequence of ANGPTL3 Target Site	6023	MG29-1-mouse ANGPTL3-target site-G9	GGAGAATTTTGGTTGGGCCTAG
DNA Sequence of ANGPTL3 Target Site	**6024**	**MG29-1-mouse ANGPTL3-target site-H9**	GGTTGGGCCTAGAGAAGATCTA
DNA Sequence of ANGPTL3 Target Site	**6025**	**MG29-1-mouse ANGPTL3-target site-A10**	GTTGGGCCTAGAGAAGATCTAT
DNA Sequence of ANGPTL3 Target Site	6026	MG29-1-mouse ANGPTL3-target site-B10	ACGACTCGAGCTACAAGACTGG
DNA Sequence of ANGPTL3 Target Site	**6027**	**MG29-1-mouse ANGPTL3-target site-C10**	CGACTCGAGCTACAAGACTGGA
DNA Sequence of ANGPTL3 Target Site	**6028**	**MG29-1-mouse ANGPTL3-target site-D10**	ACCTGGGCAGTCACGAAACCAA
DNA Sequence of ANGPTL3 Target Site	**6029**	**MG29-1-mouse ANGPTL3-target site-E10**	GTGACTGCCCAGGTGAAAGGAG T

Type	SEQ ID NO	Guide Name	SEQUENCE
DNA Sequence of ANGPTL3 Target Site	6030	**MG29-1-mouse ANGPTL3-target site-F10**	CAGTCTTGTAGCTCGAGTCGTA
r; native RNA linkage comprising the sugar ribose (for example the ribose or RNA form of the A base is written rA), d; deoxyribose sugar (DNA) linkage (for example a deoxyribose form of the A base is written dA), *; between bases in which one of the oxygen molecules in the phosphodiester bond is replaced with sulfur; AltR1 and AltR2 refer to IDT technologies' proprietary 5' and 3' AltR modifications			

Example 79 – Compact MG55-43 nuclease system

In vitro characterization to identify putative tracrRNAs

[00606] To identify tracrRNA sequences, the nuclease (MG55-43, protein SEQ ID NO: 470), intergenic sequences, and minimal arrays were expressed in transcription-translation reaction mixtures using myTXTL®Sigma 70 Master Mix Kit (Arbor Biosciences). The final reaction mixtures contained 5 nM nuclease DNA template, 12 nM intergenic DNA template, 15 nM minimal array DNA template, 0.1 nM pTXTL-P70a-T7rnap, and 1X of myTXTL®Sigma 70 Master Mix. The reactions were incubated at 29 °C for 16 hours then stored at 4 °C.

[00607] Ribonucleoprotein complexes were tested via *in vitro* cleavage reactions. Plasmid DNA library cleavage reactions were carried out by mixing 5 nM of the target plasmid DNA library representing all possible 8N PAMs, a 5-fold dilution of the TXTL expressions, 10 nM Tris-HCl, 10 nM $MgCl_2$, and 100 mM NaCl at 37 °C for 2 hours. Reactions were stopped and cleaned with HighPrep™ PCR clean up beads (MAGBIO Genomics, Inc.) and eluted in Tris EDTA pH 8.0 buffer. 3 nM of the cleavage product ends were blunted with 3.33 µM dNTPs, 1X T4 DNA ligase buffer, and 0.167 U/µL of Klenow Fragment (New England Biolabs Inc.) at 25 °C for 15 minutes. 1.5 nM of the cleavage products were ligated with 150 nM adapters, 1 X T4 DNA ligase buffer (New England Biolabs Inc.), and 20 U/µL T4 DNA ligase (New England Biolabs Inc.) at room temperature for 20 minutes. The ligated products were amplified by PCR with NGS primers and sequenced by NGS to obtain the PAM.

[00608] To obtain the sequence of the tracrRNA and crRNA, RNA is extracted from TXTL lysate following the Quick-RNA™ Miniprep Kit (Zymo Research) and eluted in 30-50 µL of water. The total concentration of the transcripts was measured on a Nanodrop, Tapestation, and Qubit. RNA libraries were subjected to RNA sequencing.

In silico search for novel tracrRNA sequences

[00609] To identify additional non-coding regions containing potential tracrRNAs, the sequence of the active tracrRNA was mapped to other contigs containing nucleases in the same nuclease

family. The newly identified sequences were used to generate covariance models to predict additional tracrRNAs. Covariance models were built from a multiple sequence alignment (MSA) of the active and predicted tracrRNA sequences. The secondary structure of the MSA was obtained with RNAalifold (Vienna Package), and the covariance models were built with Infernal packages (http://eddylab.org/infernal/). Contigs containing candidate nucleases were searched using the covariance models with the Infernal command 'cmsearch'.

sgRNA design

[00610] The predicted tracrRNA obtained from the covariance models and associated CRISPR repeat sequence were modified to generate an sgRNA (**FIG. 104A**, **SEQ ID NO: 6031**) as follows: the 3' end of the predicted tracrRNA sequence as well as the 5' end of the repeat sequence were trimmed, and then connected with a GAAA tetraloop (**FIG. 104B)**.

In vitro cleavage reactions confirmed MG55-43 activity and enabled PAM determination

[00611] Target plasmid DNA library cleavage reactions described above (***In vitro characterization to identify putative tracrRNAs***) were carried out using the guides. The product of the reactions were amplified by PCR with NGS primers and sequenced by NGS to obtain the PAM. Active proteins that successfully cleaved the PAM library yielded a band around 188 or 205 bp in agarose gel electrophoresis (**FIG. 104C**).

Table 44: Spacer sequences for tested guides

Code	Sequence
U67 spacer	GTCGAGGCTTGCGACGTGGT
U40 spacer	TGGAGATATCTTGAACCTTG

[00612] Sequence logos for PAMs were made using Seqlogo maker, and histograms showing cut-site preferences were made from the counts of reads mapping at each nucleotide position. The PAM recognized by MG55-43 is a 5'-yTn sequence (**FIG. 104D**, **SEQ ID NO: 6032**), and the preferred cut position is nucleotide 23 (**FIG. 104E**).

Example 80 – MG91 family of compact type V nucleases

De novo prediction of tracrRNA sequences encoded in intergenic regions

[00613] Compact type V nuclease proteins from distinct clades were targeted for *in silico* characterization of genomic regions encoding CRISPR Cas systems. To identify intergenic regions potentially encoding tracrRNAs, individual protein clades (with confirmed catalytic residues) were chosen for visual inspection of contigs encoding the compact type nuclease genes

and a CRISPR array. Genomic regions devoid of coding sequence predictions between two genes, or between genes and CRISPR arrays, were manually annotated as intergenic regions (**FIG. 105**). Intergenic regions upstream and downstream from a nuclease gene as well as a CRISPR array, i.e., at the same location relative to a nuclease and the corresponding CRISPR array, were consistently assigned labels across contigs encoding homologous nucleases. Nucleotide sequences of matching intergenic regions were aligned and inspected for conserved motifs across sequences (**FIG. 106**). Similarly, nucleotide sequences from non-matching intergenic regions within clades were aligned and inspected. By comparison, intergenic regions with the highest degree of conservation among them were identified as potentially encoding tracrRNAs.

[00614] Intergenic regions potentially encoding tracrRNAs corresponding to candidate nucleases (SEQ ID NOs: 6040-6049) were identified by the methods described herein, and are the subject of *in vitro* characterization. Results for active nucleases MG91-15 (SEQ ID NO: 2824), MG91-32 (SEQ ID NO: 2841), and MG91-87 (SEQ ID NO: 2896) are presented herein.

Intergenic region secondary structure predictions inform tracrRNA prediction

[00615] Intergenic regions potentially encoding tracrRNAs were folded with the corresponding repeat sequences using different energy models (Turner 2004 or Andronescu 2007) and parameters (for example, 20 °C, 37 °C, dangling ends). The stability of potential secondary RNA structures was visually inspected based on base pairs probabilities. Optimal intergenic region/repeat folds for MG91-15, MG91-32, and MG91-87 were obtained and used to inform the design of single guide RNAs.

MG91-15, MG91-32 and MG91-87 sgRNA design

[00616] Promising folds between intergenic regions potentially containing tracrRNAs with the corresponding repeat sequences were modified as follows: tracrRNA sequences were trimmed on the 3' end and sometimes on the 5' end, and repeat sequences were trimmed on the 5' end. Both RNA sequences were connected via a GAAA tetraloop and folded for secondary structure prediction as described above. The secondary structure for active sgRNAs corresponding to nucleases MG91-15 sgRNA1 (**SEQ ID NO: 6033**), MG91-32 sgRNA1 (**SEQ ID NO: 6034**), and MG91-87 sgRNA1 (**SEQ ID NO: 6035**) are shown in **FIG. 107**.

In vitro activity and PAM sequence determination for MG91-15, MG91-32 and MG91-87 nucleases

[00617] 5 nM of nuclease amplified DNA templates and 25 nM sgRNA amplified DNA templates were expressed at 37 °C for 3 hours with PURExpress® In Vitro Protein Synthesis Kit (New England Biolabs Inc.). Plasmid library DNA cleavage reactions were carried out by mixing 5 nM of the target library representing all possible 8N PAMs, a 5-fold dilution of PURExpress expressions, 10 nM Tris-HCl, 10 nM MgCl2, and 100 mM NaCl at 37 °C for 2 hours. Reactions were stopped and cleaned with HighPrep™ PCR clean up beads (MAGBIO Genomics, Inc.) and eluted in Tris EDTA pH 8.0 buffer. 3 nM of the cleavage product ends were blunted with 3.33 µM dNTPs, 1X T4 DNA ligase buffer, and 0.167 U/µL of Klenow Fragment (New England Biolabs Inc.) at 25 °C for 15 minutes. 1.5 nM of the cleavage products were ligated with 150 nM adapters, 1 X T4 DNA ligase buffer (New England Biolabs Inc.), and 20 U/µL T4 DNA ligase (New England Biolabs Inc.) at room temperature for 20 minutes. The ligated products were amplified by PCR with NGS primers and sequenced by NGS to obtain the PAM.

[00618] Active proteins that successfully cleaved the PAM library yielded a band around 188 or 205 bp in an agarose gel (**FIG. 107D**).

[00619] Sequence logos were made using Seqlogo maker, and histograms were made from the counts of reads at each nucleotide position. The PAM recognized by MG91-15 is a 5'-TtTYn sequence (**FIG. 108A, SEQ ID NO: 6037**), by MG91-32 is a 5'-GnYYn sequence (**FIG. 108B, SEQ ID NO: 6038**), and by MG-87 is a 5'-wCCC sequence (**FIG. 108C, SEQ ID NO: 6039**). The preferred cut position(s) were nucleotides 21 and 22 for MG91-15 (**FIG. 108D**), nucleotide 17 for MG91-32 (**FIG. 108E**) and nucleotide 20 for MG91-87 (**FIG. 108F**).

[00620] Covariance models were built as described above (*In silico* search for novel tracrRNA sequences, Compact type MG55-43 Cas nuclease system) from active tracrRNA sequences, and used to refine the prediction of tracrRNAs from other intergenic regions associated with nucleases in the MG91 family.

In vitro activity of compact type V MG91 nucleases and selected intergenic regions (prophetic)

[00621] Nuclease, intergenic sequences, and minimal arrays are expressed in transcription-translation reaction mixtures using myTXTL®Sigma 70 Master Mix Kit (Arbor Biosciences) as described before. Ribonucleoprotein complexes are tested in *in vitro* cleavage reactions of plasmid target DNA library using the TXTL expressions. The products are amplified by PCR with NGS primers and sequenced by NGS to obtain the PAM sequence.

[00622] Active proteins that successfully cleave the PAM library are expected to yield a band around 188 or 205 bp in agarose gel electrophoresis. Sequence logos are made using Seqlogo maker, and histograms aregenerated from the counts of reads at each nucleotide position.

[00623] To obtain the sequence of active tracrRNA and crRNA, RNA is extracted from TXTL lysate following the Quick-RNA™ Miniprep Kit (Zymo Research). The total concentration of the transcripts is measured on a Nanodrop, Tapestation, and Qubit; and is subjected to next generation sequencing.

[00624] The sequence of active tracrRNAs and crRNAs is used to design sgRNAs, and is subsequently tested in PURExpress with the corresponding MG91 nuclease to confirm the PAM sequence and cut site.

E. coli expressions (compact type V nucleases) (prophetic)

[00625] Plasmids encoding the effector, intergenic sequence from the genomic contig, native repeat, and universal spacer sequences with a T7 promoter are transformed into BL21 DE3 or T7 Express lysY/Iq and cultured at 37 °C in 60 mL terrific broth media supplemented with 100 µg/mL of ampicillin. Expression is induced with 0.4 mM IPTG after cultures reach OD600nm of 0.5 and cells are incubated at 16 °C overnight. 25 mL of cells are pelleted by centrifugation and resuspended in 1.5 mL of lysis buffer (20 mM Tris-HCl, 500 mM NaCl, 1 mM TCEP, 5% glycerol, 10 mM $MgCl_2$ pH 7.5 with Pierce Protease Inhibitor, (Thermo Scientific™)). Cells are then lysed by sonication. Supernatant and cell debris are separated by centrifugation.

In vitro cleavage efficiency with purified protein (compact type V nucleases) (prophetic)

[00626] Proteins are expressed in *E. coli* protease deficient B strain under T7 inducible promoter, the cells are lysed using sonication, and the His-tagged protein of interest is purified using HisTrap FF (GE Lifescience) Ni-NTA affinity chromatography on the AKTA Avant FPLC (GE Lifescience). Purity is determined using densitometry in ImageLab software (Bio-Rad) of the protein bands resolved on SDS-PAGE and InstantBlue Ultrafast (Sigma-Aldrich) coomassie stained acrylamide gels (Bio-Rad). The protein is desalted in storage buffer composed of 50 mM Tris-HCl, 300 mM NaCl, 1 mM TCEP, 5% glycerol; pH 7.5 and stored at -80°C.

[00627] A target DNA is constructed that contains a spacer sequence and the PAM determined via NGS. In the case of degenerate bases in the PAM, a single representative PAM is chosen for testing. The target DNA is 2200 bp of linear DNA derived from a plasmid via PCR amplification. The PAM and spacer are located 700 bp from one end. Successful cleavage results in fragments of 700 and 1500 bp.

[00628] The target DNA, *in vitro* transcribed single RNA, and purified recombinant protein are combined in cleavage buffer (10 mM Tris, 100 mM NaCl, 10 mM $MgCl_2$) with an excess of protein and RNA and incubated for 5' to 3 hours, usually 1 hr. The reaction is stopped via

addition of RNAse A and incubation at 60 °C. The reaction is resolved on a 1.2% TAE agarose gel and the fraction of cleaved target DNA is quantified in ImageLab software.

Activity in E. coli (compact type V nucleases) (prophetic)

[00629] For testing of nuclease activity in bacterial cells, strains are constructed with genome sequences containing a spacer target with corresponding PAM sequence specific to the enzyme of interest. Engineered strains are then transformed with the nuclease of interest and transformants are then subsequently made chemocompetent and transformed with 50 ng of single guides either specific to the target sequence (on target) or non specific to the target (off target). After heat shock, transformations are recovered in SOC for 2 hrs at 37 °C, and nuclease efficiency is determined by a 5-fold dilution series grown on induction media. Colonies are quantified from the dilution series in triplicate. Nuclease, and therefore genome editing capability, is assessed by quantifying the reduction of viable cells in the presence of guides and nucleases targeting the host cell chromosome.

Activity in mammalian cells (compact type V nucleases) (prophetic)

[00630] To show targeting and cleavage activity in mammalian cells, and therefore genome editing potential, the protein sequences are cloned into two mammalian expression vectors, one with a C-terminal SV40 NLS and a 2A-GFP tag and one with no GFP tag and two NLS sequences (one on the N-terminus and one on the C-terminus). Alternative NLS sequences that could also be used are listed in **Table 45**.

Table 45: Alternative nuclear localization sequences (NLS)

Source	NLS amino acid sequence
SV40	PKKKRKV
nucleoplasmin bipartite NLS	KRPAATKKAGQAKKKK
c-myc NLS	PAAKRVKLD
c-myc NLS	RQRRNELKRSP
hRNPA1 M9 NLS	NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY
Importin-alpha IBB domain	RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV
Myoma T protein	VSRKRPRP
Myoma T protein	PPKKARED
p53	PQPKKKPL
mouse c-abl IV	SALIKKKKKMAP
influenza virus NS1	DRLRR

Source	NLS amino acid sequence
influenza virus NS1	PKQKKRK
Hepatitis virus delta antigen	RKLKKKIKKL
mouse Mx1 protein	REKKKFLKRR
human poly(ADP-ribose) polymerase	KRKGDEVDGVDEVAKKKSKK
steroid hormone receptors (human) glucocorticoid	RKCLQAGMNLEARKTKK

[00631] The DNA sequence for the protein can be the native sequence, the *E. coli* codon optimized sequence, or the mammalian codon optimized sequence. The single guide RNA sequence with a gene target of interest is also cloned into a mammalian expression vector. The two plasmids are cotransfected into HEK293T cells. 72 hr after co-transfection of the expression plasmid and a sgRNA targeting plasmid into HEK293T cells, the DNA is extracted and used for the preparation of an NGS library. Percent NHEJ is measured via InDels in the sequencing of the target site to demonstrate the targeting efficiency of the enzyme in mammalian cells. At least 10 different target sites are chosen for testing protein activity.

Table 46 – Listing of additional protein and nucleic acid sequences referred to herein not included in the sequence listing

Cat.	SEQ ID:	Description	Type	Organism	Sequence
	4426	MG29-1 WT mRNA			TAATACGACTCACTATAA GGAAAAGCCAGCTCCAG CAGGCGCTGCTCACTCC TCCCCATCCTCTCCCTCT GTCCCTCTGTCCCTCTGA CCCTGCACTGTCCCAGC ACCATGGCCCCCAAGAA GAAGCGGAAAGTTGGCG GCGGAGGCAGCTTCAAC AACTTCATCAAGAAATAC AGCCTGCAGAAGACCCT GCGGTTCGAACTGAAGC CCGTGGGCGAGACAGCG GACTACATCGAAGACTT CAAGAGCGAATACCTGA AGGACACGGTGCTGAAG GACGAACAGCGGGCAAA AGACTACCAGGAGATCA AAACACTGATCGACGAC TACCACCGGGAGTACAT CGAAGAATGCCTGAGGG AACCCGTGGACAAAAAG ACCGGCGAGATCCTGGA CTTCACACAGGACCTGG AAGACGCATTCAGCTAC TACCAGAAACTGAAAGA AAACCCCACCGAGAACC GAGTGGGGTGGGAGAAA GAGCAGGAGAGCCTGAG AAAGAAGCTGGTGACCA GCTTCGTGGGGAACGAC GGCCTGTTCAAGAAAGA GTTCATCACCCGCGACC TGCCCGAATGGCTGCAG AAAAAGGGGCTGTGGGG CGAATACAAGGACACCG TGGAGAACTTCAAAAAA TTCACCACCTACTTCAGC GGCTTCCACGAGAACAG GAAGAATATGTACACAG CCGAAGCCCAGAGCACA GCCATCGCCAACAGGCT GATGAACGACAACCTGC CCAAGTTCTTCAACAACT ACCTGGCATACCAGACC ATCAAGGAGAAACACCC CGACCTGGTGTTCCGAC TGGACGACGCCCTGCTG CAGGCCGCCGGCGTGGA GCACCTGGACGAGGCAT TCCAGCCCAGATACTTC AGCAGACTGTTCGCACA GAGCGGAATCACGGCCT TCAACGAGCTGATCGGA GGAAGGACCACGGAAAA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					CGGCGAAAAGATCCAGG GCCTGAACGAGCAGATC AACCTGTACAGACAGCA GAACCCCGAGAAGGCCA AGGGCTTCCCAAGATTC ATGCCCCTGTTCAAGCA AATCCTGAGCGACAGGG AGACCCACAGCTTCCTG CCCGACGCATTCGAAAA CGACAAAGAGCTGCTGC AGGCCCTGAGGGACTAC GTGGACGCCGCCACCAG CGAAGAAGGAATGATCA GCCAACTGAACAAGGCC ATGAACCAGTTCGTGAC CGCCGACCTGAAAAGGG TGTACATCAAAAGCGCC GCCCTGACCAGCCTGAG CCAGGAACTGTTCCACT TCTTCGGCGTGATCAGC GACGCCATCGCGTGGTA CGCCGAGAAGAGACTGA GCCCCAAGAAAGCCCAG GAGAGCTTCCTGAAACA GGAAGTGTACGCCATCG AAGAACTGAACCAGGCC GTGGTGGGCTACATCGA CCAGCTGGAAGACCAGA GCGAGCTGCAGCAGCTG CTGGTGGACCTGCCAGA CCCCCAGAAACCAGTGA GCAGCTTCATCCTGACC CACTGGCAAAAAAGCCA GGAGCCGCTGCAGGCCG TGATCGCGAAGGTGGAA CCCCTGTTCGAACTGGA GGAGCTGAGCAAAAACA AACGGGCCCCGAAACAC GACAAGGACCAGGGAGG GGAAGGCTTCCAGCAGG TGGACGCAATCAAGAAC ATGCTGGACGCATTCAT GGAGGTGAGCCACGCCA TCAAGCCCCTGTACCTG GTGAAGGGCCGGAAAGC AATCGACATGCCGGACG TGGACACAGGATTCTAC GCCGACTTCGCGGAGGC ATACAGCGCCTACGAGC AAGTGACGGTGAGCCTG TACAACAAGACCCGAAA CCACCTGAGCAAGAAAC CCTTCAGCAAAGACAAA ATCAAAATCAACTTCGAC GCCCCAACACTGCTGAA CGGCTGGGACCTGAACA AGGAAAGCGACAACAAA AGCATCATCCTGAGAAA AGACGGAAACTTCTACC TGGCCATCATGCACCCC

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					AAACACACAAAGGTGTT CGACTGCTACAGCGCCA GCGAGGCGGCCGGGAAA TGCTACGAGAAAATGAA CTACAAACTGCTGAGCG GCGCCAACAAGATGCTG CCCAAAGTGTTCTTCAG CAAGAAGGGAATCGAAA CCTTCAGCCCACCCCAG GAAATCCTGGACCTGTA CAAGAACAACGAACACA AGAAGGGAGCCACCTTC AAGCTGGAGAGCTGCCA CAAGCTGATCGACTTCTT CAAGCGGAACATCCCCA AGTACAAGGTGCACCCA ACCGACAACTTCGGATG GGACGTCTTCGGATTCC ACTTCAGCCCAACCAGC AGCTACGGCGACCTGAG CGGCTTCTACCGAGAGG TGGAAGCCCAGGGGTAC AAACTGTGGTTCAGCGA CGTGAGCGAGGCATACA TCAACAAGTGCGTGGAA GAGGGCAAACTGTTCCT GTTCCAGATCTACAACA AGGACTTCAGCCCCAAC AGCACCGGGAAGCCAAA CCTGCACACACTGTACT GGAAAGGACTGTTCGAA CCCGAGAACCTGAAGGA CGTGGTGCTGAAACTGA ACGGCGAGGCCGAGATC TTCTACAGGAAACACAG CATCAAGCACGAGGACA AGACGATCCACCGGGCC AAGGACCCAATCGCCAA CAAAAACGCAGACAACC CCAAGAAGCAGAGCGTG TTCGACTACGACATCATC AAGGACAAGCGCTACAC CCAGGACAAATTCTTCTT CCACGTGCCCATCAGCC TGAACTTCAAGAGCCAG GGAGTGGTGCGGTTCAA CGACAAGATCAACGGCC TGCTGGCCGCACAGGAC GACGTGCACGTGATCGG GATCGACCGAGGGGAAC GCCACCTGCTGTACTAC ACCGTGGTGAACGGCAA GGGCGAGGTGGTGGAAC AGGGCAGCCTGAACCAG GTGGCCACAGACCAGGG GTACGTGGTGGACTACC AACAGAAACTGCACGCC AAAGAGAAGGAGAGAGA CCAGGCCAGGAAGAACT GGAGCACCATCGAAAAC

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					ATCAAGGAGCTGAAGGC
					CGGGTACCTGAGCCAGG
					TGGTGCACAAACTGGCC
					CAGCTGATCGTGAAACA
					CAACGCCATCGTGTGCC
					TGGAGGACCTGAACTTC
					GGATTCAAGAGGGGACG
					GTTCAAAGTGGAGAAGC
					AGGTGTACCAGAAGTTC
					GAGAAAGCCCTGATCGA
					CAAGCTGAACTACCTGG
					TGTTCAAGGAACGGGGG
					GCCACCCAGGCAGGCGG
					ATACCTGAACGCCTACC
					AGCTGGCCGCACCATTC
					GAGAGCTTCGAAAAACT
					GGGCAAGCAGACCGGCA
					TCCTGTACTACGTGCGG
					AGCGACTACACCAGCAA
					GATCGACCCCGCCACAG
					GCTTCGTGGACTTCCTG
					AAGCCCAAATACGAAAG
					CATGGCAAAGAGCAAAG
					TGTTCTTCGAGAGCTTC
					GAAAGAATCCAGTGGAA
					CCAGGCCAAAGGCTACT
					TCGAGTTCGAATTCGAC
					TACAAGAAAATGTGCCC
					CAGCAGGAAGTTCGGCG
					ACTACCGCACCCGGTGG
					GTGGTGTGCACATTCGG
					CGACACACGGTACCAGA
					ACAGGCGCAACAAAAGC
					AGCGGCCAATGGGAGAC
					CGAGACAATCGACGTGA
					CCGCCCAGCTGAAGGCC
					CTGTTCGCGGCCTACGG
					CATCACCTACAACCAGG
					AGGACAACATCAAGGAC
					GCCATCGCAGCCGTGAA
					GTACACAAAATTCTACAA
					ACAGCTGTACTGGCTGC
					TGAGACTGACGCTGAGC
					CTGCGGCACAGCGTGAC
					CGGGACCGACGAGGACT
					TCATCCTGAGCCCCGTG
					GCCGACGAGAACGGCGT
					GTTCTTCGACAGCAGGA
					AGGCCACGGACAAACAG
					CCCAAGGACGCAGACGC
					GAACGGCGCCTACCACA
					TCGCCCTGAAGGGACTG
					TGGAACCTGCAGCAGAT
					CAGGCAGCACGACTGGA
					ACGTGGAAAAACCAAAA
					AAGCTGAACCTGGCCAT
					GAAAAACGAAGAGTGGT
					TCGGCTTCGCACAGAAG
					AAGAAATTCAGGGCCTC
					TGGCGGAAAAAGACCTG

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					CCGCCACAAAGAAAGCC GGACAGGCCAAGAAAAA GAAGTGACCACACCCCC ATTCCCCCACTCCAGATA GAACTTCAGTTATATCTC ACGTGTCTGGAGTTGGA TCCAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA AAAAAAAA
	4427	MG29-1 S168R mRNA			TAATACGACTCACTATAA GGAAAAGCCAGCTCCAG CAGGCGCTGCTCACTCC TCCCCATCCTCTCCCTCT GTCCCTCTGTCCCTCTGA CCCTGCACTGTCCCAGC ACCATGGCCCCCAAGAA GAAGCGGAAAGTTGGCG GCGGAGGCAGCTTCAAC AACTTCATCAAGAAATAC AGCCTGCAGAAGACCCT GCGGTTCGAACTGAAGC CCGTGGGCGAGACAGCG GACTACATCGAAGACTT CAAGAGCGAATACCTGA AGGACACGGTGCTGAAG GACGAACAGCGGGCAAA AGACTACCAGGAGATCA AAACACTGATCGACGAC TACCACCGGGAGTACAT CGAAGAATGCCTGAGGG AACCCGTGGACAAAAAG ACCGGCGAGATCCTGGA CTTCACACAGGACCTGG AAGACGCATTCAGCTAC TACCAGAAACTGAAAGA AAACCCCACCGAGAACC GAGTGGGGTGGGAGAAA GAGCAGGAGAGCCTGAG AAAGAAGCTGGTGACCA GCTTCGTGGGGAACGAC GGCCTGTTCAAGAAAGA GTTCATCACCCGCGACC TGCCCGAATGGCTGCAG AAAAAGGGGCTGTGGGG CGAATACAAGGACACCG TGGAGAACTTCAAAAAA TTCACCACCTACTTCAGG GGCTTCCACGAGAACAG GAAGAATATGTACACAG CCGAAGCCCAGAGCACA GCCATCGCCAACAGGCT GATGAACGACAACCTGC CCAAGTTCTTCAACAACT ACCTGGCATACCAGACC ATCAAGGAGAAACACCC CGACCTGGTGTTCCGAC

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					TGGACGACGCCCTGCTG CAGGCCGCCGGCGTGGA GCACCTGGACGAGGCAT TCCAGCCCAGATACTTC AGCAGACTGTTCGCACA GAGCGGAATCACGGCCT TCAACGAGCTGATCGGA GGAAGGACCACGGAAAA CGGCGAAAAGATCCAGG GCCTGAACGAGCAGATC AACCTGTACAGACAGCA GAACCCCGAGAAGGCCA AGGGCTTCCCAAGATTC ATGCCCCTGTTCAAGCA AATCCTGAGCGACAGGG AGACCCACAGCTTCCTG CCCGACGCATTCGAAAA CGACAAAGAGCTGCTGC AGGCCCTGAGGGACTAC GTGGACGCCGCCACCAG CGAAGAAGGAATGATCA GCCAACTGAACAAGGCC ATGAACCAGTTCGTGAC CGCCGACCTGAAAAGGG TGTACATCAAAAGCGCC GCCCTGACCAGCCTGAG CCAGGAACTGTTCCACT TCTTCGGCGTGATCAGC GACGCCATCGCGTGGTA CGCCGAGAAGAGACTGA GCCCCAAGAAAGCCCAG GAGAGCTTCCTGAAACA GGAAGTGTACGCCATCG AAGAACTGAACCAGGCC GTGGTGGGCTACATCGA CCAGCTGGAAGACCAGA GCGAGCTGCAGCAGCTG CTGGTGGACCTGCCAGA CCCCCAGAAACCAGTGA GCAGCTTCATCCTGACC CACTGGCAAAAAAGCCA GGAGCCGCTGCAGGCCG TGATCGCGAAGGTGGAA CCCCTGTTCGAACTGGA GGAGCTGAGCAAAAACA AACGGGCCCCGAAACAC GACAAGGACCAGGGAGG GGAAGGCTTCCAGCAGG TGGACGCAATCAAGAAC ATGCTGGACGCATTCAT GGAGGTGAGCCACGCCA TCAAGCCCCTGTACCTG GTGAAGGGCCGGAAAGC AATCGACATGCCGGACG TGGACACAGGATTCTAC GCCGACTTCGCGGAGGC ATACAGCGCCTACGAGC AAGTGACGGTGAGCCTG TACAACAAGACCCGAAA CCACCTGAGCAAGAAAC

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					CCTTCAGCAAAGACAAA
					ATCAAAATCAACTTCGAC
					GCCCCAACACTGCTGAA
					CGGCTGGGACCTGAACA
					AGGAAAGCGACAACAAA
					AGCATCATCCTGAGAAA
					AGACGGAAACTTCTACC
					TGGCCATCATGCACCCC
					AAACACACAAAGGTGTT
					CGACTGCTACAGCGCCA
					GCGAGGCGGCCGGGAAA
					TGCTACGAGAAAATGAA
					CTACAAACTGCTGAGCG
					GCGCCAACAAGATGCTG
					CCCAAAGTGTTCTTCAG
					CAAGAAGGGAATCGAAA
					CCTTCAGCCCACCCCAG
					GAAATCCTGGACCTGTA
					CAAGAACAACGAACACA
					AGAAGGGAGCCACCTTC
					AAGCTGGAGAGCTGCCA
					CAAGCTGATCGACTTCTT
					CAAGCGGAACATCCCCA
					AGTACAAGGTGCACCCA
					ACCGACAACTTCGGATG
					GGACGTCTTCGGATTCC
					ACTTCAGCCCAACCAGC
					AGCTACGGCGACCTGAG
					CGGCTTCTACCGAGAGG
					TGGAAGCCCAGGGGTAC
					AAACTGTGGTTCAGCGA
					CGTGAGCGAGGCATACA
					TCAACAAGTGCGTGGAA
					GAGGGCAAACTGTTCCT
					GTTCCAGATCTACAACA
					AGGACTTCAGCCCCAAC
					AGCACCGGGAAGCCAAA
					CCTGCACACACTGTACT
					GGAAAGGACTGTTCGAA
					CCCGAGAACCTGAAGGA
					CGTGGTGCTGAAACTGA
					ACGGCGAGGCCGAGATC
					TTCTACAGGAAACACAG
					CATCAAGCACGAGGACA
					AGACGATCCACCGGGCC
					AAGGACCCAATCGCCAA
					CAAAAACGCAGACAACC
					CCAAGAAGCAGAGCGTG
					TTCGACTACGACATCATC
					AAGGACAAGCGCTACAC
					CCAGGACAAATTCTTCTT
					CCACGTGCCCATCAGCC
					TGAACTTCAAGAGCCAG
					GGAGTGGTGCGGTTCAA
					CGACAAGATCAACGGCC
					TGCTGGCCGCACAGGAC
					GACGTGCACGTGATCGG
					GATCGACCGAGGGGAAC
					GCCACCTGCTGTACTAC
					ACCGTGGTGAACGGCAA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					GGGCGAGGTGGTGGAAC
					AGGGCAGCCTGAACCAG
					GTGGCCACAGACCAGGG
					GTACGTGGTGGACTACC
					AACAGAAACTGCACGCC
					AAAGAGAAGGAGAGAGA
					CCAGGCCAGGAAGAACT
					GGAGCACCATCGAAAAC
					ATCAAGGAGCTGAAGGC
					CGGGTACCTGAGCCAGG
					TGGTGCACAAACTGGCC
					CAGCTGATCGTGAAACA
					CAACGCCATCGTGTGCC
					TGGAGGACCTGAACTTC
					GGATTCAAGAGGGGACG
					GTTCAAAGTGGAGAAGC
					AGGTGTACCAGAAGTTC
					GAGAAAGCCCTGATCGA
					CAAGCTGAACTACCTGG
					TGTTCAAGGAACGGGGG
					GCCACCCAGGCAGGCGG
					ATACCTGAACGCCTACC
					AGCTGGCCGCACCATTC
					GAGAGCTTCGAAAAACT
					GGGCAAGCAGACCGGCA
					TCCTGTACTACGTGCGG
					AGCGACTACACCAGCAA
					GATCGACCCCGCCACAG
					GCTTCGTGGACTTCCTG
					AAGCCCAAATACGAAAG
					CATGGCAAAGAGCAAAG
					TGTTCTTCGAGAGCTTC
					GAAAGAATCCAGTGGAA
					CCAGGCCAAAGGCTACT
					TCGAGTTCGAATTCGAC
					TACAAGAAAATGTGCCC
					CAGCAGGAAGTTCGGCG
					ACTACCGCACCCGGTGG
					GTGGTGTGCACATTCGG
					CGACACACGGTACCAGA
					ACAGGCGCAACAAAAGC
					AGCGGCCAATGGGAGAC
					CGAGACAATCGACGTGA
					CCGCCCAGCTGAAGGCC
					CTGTTCGCGGCCTACGG
					CATCACCTACAACCAGG
					AGGACAACATCAAGGAC
					GCCATCGCAGCCGTGAA
					GTACACAAAATTCTACAA
					ACAGCTGTACTGGCTGC
					TGAGACTGACGCTGAGC
					CTGCGGCACAGCGTGAC
					CGGGACCGACGAGGACT
					TCATCCTGAGCCCCGTG
					GCCGACGAGAACGGCGT
					GTTCTTCGACAGCAGGA
					AGGCCACGGACAAACAG
					CCCAAGGACGCAGACGC
					GAACGGCGCCTACCACA
					TCGCCCTGAAGGGACTG

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					TGGAACCTGCAGCAGAT CAGGCAGCACGACTGGA ACGTGGAAAAACCAAAA AAGCTGAACCTGGCCAT GAAAAACGAAGAGTGGT TCGGCTTCGCACAGAAG AAGAAATTCAGGGCCTC TGGCGGAAAAAGACCTG CCGCCACAAAGAAAGCC GGACAGGCCAAGAAAAA GAAGTGACCACACCCCC ATTCCCCCACTCCAGATA GAACTTCAGTTATATCTC ACGTGTCTGGAGTTGGA TCCAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAA AAAAAAAA
MG29-1 sgRNA targeting hRosa26	4970	MG29-1-hRosa26-sgRNA-A1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrGrUrCrArCrUrGrUrCrCrUrArGrCrUrCrUrCrCrA/AltR2/
MG29-1 sgRNA targeting hRosa26	4971	MG29-1-hRosa26-sgRNA-B1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrArGrGrCrCrGrGrGrCrGrCrGrGrUrGrGrC/AltR2/
MG29-1 sgRNA targeting hRosa26	4972	MG29-1-hRosa26-sgRNA-C1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrGrCrCrGrGrGrCrGrCrGrGrUrGrGrCrUrCrArC/AltR2/
MG29-1 sgRNA targeting hRosa26	4973	MG29-1-hRosa26-sgRNA-D1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrGrCrCrGrGrGrCrGrCrGrGrUrGrGrCrUrCrArCrA/AltR2/
MG29-1 sgRNA targeting hRosa26	4974	MG29-1-hRosa26-sgRNA-E1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrCrCrGrGrGrCrGrCrGrGrUrGrGrCrUrCrArCrArC/AltR2/
MG29-1 sgRNA targeting hRosa26	4975	MG29-1-hRosa26-sgRNA-F1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrArGrCrCrGrArGrGrCrArGrGrCrArGrArUrCrA/AltR2/
MG29-1 sgRNA targeting hRosa26	4976	MG29-1-hRosa26-sgRNA-G1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrGrUrCrArGrGrArGrUrUrCrArArGrArCrCrArGrC/AltR2/
MG29-1 sgRNA	4977	MG29-1-hRosa26-sgRNA-H1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrUrGrArGrCrU

Cat.	SEQ ID:	Description	Type	Organism	Sequence
targeting hRosa26					rGrArGrArUrCrGrUrGrCrCrArU/AltR2/
MG29-1 sgRNA targeting hRosa26	4978	MG29-1-hRosa26-sgRNA-A2	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrUrGrGrUrUrGrGrUrGrUrGrUrGrGrUrGrGrC/AltR2/
MG29-1 sgRNA targeting hRosa26	4979	MG29-1-hRosa26-sgRNA-B2	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrGrUrUrGrGrGrUrGrUrGrGrUrGrGrCrUrCrArC/AltR2/
MG29-1 sgRNA targeting hRosa26	4980	MG29-1-hRosa26-sgRNA-C2	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrGrUrUrGrGrGrUrGrUrGrGrUrGrGrCrUrCrArCrA/AltR2/
MG29-1 sgRNA targeting hRosa26	4981	MG29-1-hRosa26-sgRNA-D2	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrUrUrGrGrGrUrGrUrGrGrUrGrCrUrCrArCrArC/AltR2/
MG29-1 sgRNA targeting hRosa26	4982	MG29-1-hRosa26-sgRNA-E2	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrUrGrGrGrUrGrUrGrGrCrUrCrArCrArCrC/AltR2/
MG29-1 sgRNA targeting hRosa26	4983	MG29-1-hRosa26-sgRNA-F2	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArArGrGrArUrGrArGrGrCrArGrArArGrGrArUrCrA/AltR2/
MG29-1 sgRNA targeting hRosa26	4984	MG29-1-hRosa26-sgRNA-G2	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrArGrGrCrCrArArGrGrCrGrGrGrCrGrGrArUrCrA/AltR2/
MG29-1 sgRNA targeting hRosa26	4985	MG29-1-hRosa26-sgRNA-H2	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrGrCrCrArUrUrCrUrCrCrUrGrCrCrUrCrArGrCrCrU/AltR2/
MG29-1 sgRNA targeting hRosa26	4986	MG29-1-hRosa26-sgRNA-A3	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrUrGrUrUrUrUrUrArGrUrArGrArGrArArArGrGrG/AltR2/
MG29-1 sgRNA targeting hRosa26	4987	MG29-1-hRosa26-sgRNA-B3	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrUrGrUrUrUrUrUrArGrUrArGrArGrArArArGrGrG/AltR2/
MG29-1 sgRNA targeting hRosa26	4988	MG29-1-hRosa26-sgRNA-C3	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrGrUrUrUrUrUrArGrUrArGrArGrArArArGrGrGrU/AltR2/
MG29-1 sgRNA	4989	MG29-1-hRosa26-sgRNA-D3	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGr

Cat.	SEQ ID:	Description	Type	Organism	Sequence
targeting hRosa26					ArUrUrGrUrUrUrUrUrArGrUrArGrArGrArArGrGrGrGrUrU/AltR2/
MG29-1 sgRNA targeting hRosa26	4990	MG29-1-hRosa26-sgRNA-E3	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrUrArGrArGrArArGrGrGrGrUrUrUrCrArCrCrG/AltR2/
MG29-1 sgRNA targeting hRosa26	4991	MG29-1-hRosa26-sgRNA-F3	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrUrArGrArGrArArGrGrGrGrUrUrUrCrArCrCrGrU/AltR2/
MG29-1 sgRNA targeting hRosa26	4992	MG29-1-hRosa26-sgRNA-G3	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrArGrArGrArArGrGrGrGrUrUrUrCrArCrCrGrUrG/AltR2/
MG29-1 sgRNA targeting hRosa26	4993	MG29-1-hRosa26-sgRNA-H3	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrCrGrUrGrUrUrArGrCrCrArGrGrArUrGrGrUrCrU/AltR2/
MG29-1 sgRNA targeting hRosa26	4994	MG29-1-hRosa26-sgRNA-A4	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArArCrUrCrCrUrUrGrGrCrUrCrArArGrUrGrArUrC/AltR2/
MG29-1 sgRNA targeting hRosa26	4995	MG29-1-hRosa26-sgRNA-B4	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArCrUrCrCrUrUrGrGrCrUrCrArArGrUrGrArUrCrC/AltR2/
MG29-1 sgRNA targeting hRosa26	4996	MG29-1-hRosa26-sgRNA-C4	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrUrUrUrCrUrUrGrArGrUrCrArGrArArGrUrC/AltR2/
MG29-1 sgRNA targeting hRosa26	4997	MG29-1-hRosa26-sgRNA-D4	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrCrUrUrGrArGrUrCrArGrArGrUrCrUrUrGrC/AltR2/
MG29-1 sgRNA targeting hRosa26	4998	MG29-1-hRosa26-sgRNA-E4	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrCrUrUrGrArGrUrCrArGrArGrUrCrUrUrGrCrU/AltR2/
MG29-1 sgRNA targeting hRosa26	4999	MG29-1-hRosa26-sgRNA-F4	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrUrUrGrArGrUrCrArGrArGrUrCrUrUrGrCrUrC/AltR2/
MG29-1 sgRNA targeting hRosa26	5000	MG29-1-hRosa26-sgRNA-G4	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrUrGrArGrUrCrArGrArGrUrCrUrUrGrCrUrCrC/AltR2/

Cat.	SEQ ID:	Description	Type	Organism	Sequence
MG29-1 sgRNA targeting hRosa26	5001	MG29-1-hRosa26-sgRNA-H4	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrGrArGrUrCrArGrArGrUrCrUrUrGrCrUrCrCrG/AltR2/
MG29-1 sgRNA targeting hRosa26	5002	MG29-1-hRosa26-sgRNA-A5	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrArGrUrCrArGrArGrUrCrUrUrGrCrUrCrCrGrU/AltR2/
MG29-1 sgRNA targeting hRosa26	5003	MG29-1-hRosa26-sgRNA-B5	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrUrArUrUrUrUrUrUrArGrUrArGrArGrArUrGrGrGrGrG/AltR2/
MG29-1 sgRNA targeting hRosa26	5004	MG29-1-hRosa26-sgRNA-C5	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrArUrUrUrUrUrArGrUrArGrArGrArUrGrGrGrGrU/AltR2/
MG29-1 sgRNA targeting hRosa26	5005	MG29-1-hRosa26-sgRNA-D5	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArUrUrUrUrUrArGrUrArGrArGrArUrGrGrGrGrU/AltR2/
MG29-1 sgRNA targeting hRosa26	5006	MG29-1-hRosa26-sgRNA-E5	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrUrArGrArGrArArUrGrGrGrGrUrUrUrCrArCrCrA/AltR2/
MG29-1 sgRNA targeting hRosa26	5007	MG29-1-hRosa26-sgRNA-F5	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrUrArGrArGrArUrGrGrGrGrUrUrUrCrArCrCrArC/AltR2/
MG29-1 sgRNA targeting hRosa26	5008	MG29-1-hRosa26-sgRNA-G5	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrArGrArGrArUrGrGrGrGrUrUrUrCrArCrCrArCrG/AltR2/
MG29-1 sgRNA targeting hRosa26	5009	MG29-1-hRosa26-sgRNA-H5	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrCrArCrGrUrUrGrGrUrCrArGrGrCrUrGrGrUrCrU/AltR2/
MG29-1 sgRNA targeting hRosa26	5010	MG29-1-hRosa26-sgRNA-A6	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrCrUrUrCrCrUrGrUrGrArCrUrUrCrCrUrGrGrArG/AltR2/
MG29-1 sgRNA targeting hRosa26	5011	MG29-1-hRosa26-sgRNA-B6	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrUrUrCrCrUrGrUrGrArCrUrUrCrCrUrGrGrArGrA/AltR2/
MG29-1 sgRNA targeting hRosa26	5012	MG29-1-hRosa26-sgRNA-C6	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrGrGrArArGrG

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					rUrUrCrUrCrUrGrUrCrUrGrCrC/AltR2/
DNA sequence of hRosa26 target site	5013	MG29-1-hRosa26-target site-A1	Nucleotide	N.A.	AGGTCACTGTCCTAGCTCTCCA
DNA sequence of hRosa26 target site	5014	MG29-1-hRosa26-target site-B1	Nucleotide	N.A.	TTTTTAGGCCGGGCGCGGTGGC
DNA sequence of hRosa26 target site	5015	MG29-1-hRosa26-target site-C1	Nucleotide	N.A.	TAGGCCGGGCGCGGTGGCTCAC
DNA sequence of hRosa26 target site	5016	MG29-1-hRosa26-target site-D1	Nucleotide	N.A.	AGGCCGGGCGCGGTGGCTCACA
DNA sequence of hRosa26 target site	5017	MG29-1-hRosa26-target site-E1	Nucleotide	N.A.	GGCCGGGCGCGGTGGCTCACAC
DNA sequence of hRosa26 target site	5018	MG29-1-hRosa26-target site-F1	Nucleotide	N.A.	GGAGGCCGAGGCAGGCAGATCA
DNA sequence of hRosa26 target site	5019	MG29-1-hRosa26-target site-G1	Nucleotide	N.A.	AGGTCAGGAGTTCAAGACCAGC
DNA sequence of hRosa26 target site	5020	MG29-1-hRosa26-target site-H1	Nucleotide	N.A.	CAGTGAGCTGAGATCGTGCCAT
DNA sequence of hRosa26 target site	5021	MG29-1-hRosa26-target site-A2	Nucleotide	N.A.	TTTTTTGGTTGGGTGTGGTGGC
DNA sequence of hRosa26 target site	5022	MG29-1-hRosa26-target site-B2	Nucleotide	N.A.	TTGGTTGGGTGTGGTGGCTCAC
DNA sequence of hRosa26 target site	5023	MG29-1-hRosa26-target site-C2	Nucleotide	N.A.	TGGTTGGGTGTGGTGGCTCACA
DNA sequence	5024	MG29-1-hRosa26-target site-D2	Nucleotide	N.A.	GGTTGGGTGTGGTGGCTCACAC

Cat.	SEQ ID:	Description	Type	Organism	Sequence
of hRosa26 target site					
DNA sequence of hRosa26 target site	5025	MG29-1-hRosa26-target site-E2	Nucleotide	N.A.	GTTGGGTGTGGTGGCTCACACC
DNA sequence of hRosa26 target site	5026	MG29-1-hRosa26-target site-F2	Nucleotide	N.A.	GAAGGATGAGGCAGAAGGATCA
DNA sequence of hRosa26 target site	5027	MG29-1-hRosa26-target site-G2	Nucleotide	N.A.	GGAGGCCAAGGCGGGCGGATCA
DNA sequence of hRosa26 target site	5028	MG29-1-hRosa26-target site-H2	Nucleotide	N.A.	CGCCATTCTCCTGCCTCAGCCT
DNA sequence of hRosa26 target site	5029	MG29-1-hRosa26-target site-A3	Nucleotide	N.A.	TTGTGTTTTTAGTAGAGAAGGG
DNA sequence of hRosa26 target site	5030	MG29-1-hRosa26-target site-B3	Nucleotide	N.A.	TGTGTTTTTAGTAGAGAAGGGG
DNA sequence of hRosa26 target site	5031	MG29-1-hRosa26-target site-C3	Nucleotide	N.A.	GTGTTTTTAGTAGAGAAGGGGT
DNA sequence of hRosa26 target site	5032	MG29-1-hRosa26-target site-D3	Nucleotide	N.A.	TGTTTTTAGTAGAGAAGGGGTT
DNA sequence of hRosa26 target site	5033	MG29-1-hRosa26-target site-E3	Nucleotide	N.A.	TAGTAGAGAAGGGGTTTCACCG
DNA sequence of hRosa26 target site	5034	MG29-1-hRosa26-target site-F3	Nucleotide	N.A.	AGTAGAGAAGGGGTTTCACCGT
DNA sequence of hRosa26 target site	5035	MG29-1-hRosa26-target site-G3	Nucleotide	N.A.	GTAGAGAAGGGGTTTCACCGTG

Cat.	SEQ ID:	Description	Type	Organism	Sequence
DNA sequence of hRosa26 target site	5036	MG29-1-hRosa26-target site-H3	Nucleotide	N.A.	ACCGTGTTAGCCAGGATGGTCT
DNA sequence of hRosa26 target site	5037	MG29-1-hRosa26-target site-A4	Nucleotide	N.A.	GAACTCCTTGGCTCAAGTGATC
DNA sequence of hRosa26 target site	5038	MG29-1-hRosa26-target site-B4	Nucleotide	N.A.	AACTCCTTGGCTCAAGTGATCC
DNA sequence of hRosa26 target site	5039	MG29-1-hRosa26-target site-C4	Nucleotide	N.A.	TTTTTTTTCTTGAGTCAGAGTC
DNA sequence of hRosa26 target site	5040	MG29-1-hRosa26-target site-D4	Nucleotide	N.A.	TTTTCTTGAGTCAGAGTCTTGC
DNA sequence of hRosa26 target site	5041	MG29-1-hRosa26-target site-E4	Nucleotide	N.A.	TTTCTTGAGTCAGAGTCTTGCT
DNA sequence of hRosa26 target site	5042	MG29-1-hRosa26-target site-F4	Nucleotide	N.A.	TTCTTGAGTCAGAGTCTTGCTC
DNA sequence of hRosa26 target site	5043	MG29-1-hRosa26-target site-G4	Nucleotide	N.A.	TCTTGAGTCAGAGTCTTGCTCC
DNA sequence of hRosa26 target site	5044	MG29-1-hRosa26-target site-H4	Nucleotide	N.A.	CTTGAGTCAGAGTCTTGCTCCG
DNA sequence of hRosa26 target site	5045	MG29-1-hRosa26-target site-A5	Nucleotide	N.A.	TTGAGTCAGAGTCTTGCTCCGT
DNA sequence of hRosa26 target site	5046	MG29-1-hRosa26-target site-B5	Nucleotide	N.A.	TGTATTTTTAGTAGAGATGGGG
DNA sequence of	5047	MG29-1-hRosa26-target site-C5	Nucleotide	N.A.	GTATTTTTAGTAGAGATGGGGT

Cat.	SEQ ID:	Description	Type	Organism	Sequence
hRosa26 target site					
DNA sequence of hRosa26 target site	5048	MG29-1-hRosa26-target site-D5	Nucleotide	N.A.	TATTTTTAGTAGAGATGGGGTT
DNA sequence of hRosa26 target site	5049	MG29-1-hRosa26-target site-E5	Nucleotide	N.A.	TAGTAGAGATGGGGTTTCACCA
DNA sequence of hRosa26 target site	5050	MG29-1-hRosa26-target site-F5	Nucleotide	N.A.	AGTAGAGATGGGGTTTCACCAC
DNA sequence of hRosa26 target site	5051	MG29-1-hRosa26-target site-G5	Nucleotide	N.A.	GTAGAGATGGGGTTTCACCACG
DNA sequence of hRosa26 target site	5052	MG29-1-hRosa26-target site-H5	Nucleotide	N.A.	ACCACGTTGGTCAGGCTGGTCT
DNA sequence of hRosa26 target site	5053	MG29-1-hRosa26-target site-A6	Nucleotide	N.A.	ACCTTCCTGTGACTTCCTGGAG
DNA sequence of hRosa26 target site	5054	MG29-1-hRosa26-target site-B6	Nucleotide	N.A.	CCTTCCTGTGACTTCCTGGAGA
DNA sequence of hRosa26 target site	5055	MG29-1-hRosa26-target site-C6	Nucleotide	N.A.	TGGGGAAGGTTCTCTGTCTGCC
MG29-1 sgRNA targeting FAS	5268	MG29-1-FAS-sgRNA-A1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrGrUrArArArCrArArCrCrGrArArCrUrGrArUrGrA/AltR2/
MG29-1 sgRNA targeting FAS	5269	MG29-1-FAS-sgRNA-B1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrUrGrArGrCrArArArGrArCrUrCrUrUrGrCrUrArC/AltR2/
MG29-1 sgRNA targeting FAS	5270	MG29-1-FAS-sgRNA-C1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrArArCrGrUrGrGrCrArUrCrArArCrArUrCrArC/AltR2/
MG29-1 sgRNA	5271	MG29-1-FAS-sgRNA-D1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGr

Cat.	SEQ ID:	Description	Type	Organism	Sequence
targeting FAS					ArUrUrCrArCrCrArGrCrU rCrCrCrArUrGrUrGrArUr GrUrU/AltR2/
MG29-1 sgRNA targeting FAS	5272	MG29-1-FAS-sgRNA-E1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrU rArCrUrGrUrUrGrUrArGr ArUrUrGrUrGrUrCrUrArU rUrArGrArUrGrCrUrCrAr GrArG/AltR2/
MG29-1 sgRNA targeting FAS	5273	MG29-1-FAS-sgRNA-F1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrU rArCrUrGrUrUrGrUrArGr ArUrGrUrGrUrCrUrArUrU rArGrArUrGrCrUrCrArGr ArGrU/AltR2/
MG29-1 sgRNA targeting FAS	5274	MG29-1-FAS-sgRNA-G1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrU rArCrUrGrUrUrGrUrArGr ArUrUrGrUrCrUrArUrUrA rGrArUrGrCrUrCrArGrAr GrUrG/AltR2/
MG29-1 sgRNA targeting FAS	5275	MG29-1-FAS-sgRNA-H1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrU rArCrUrGrUrUrGrUrArGr ArUrGrCrArUrCrUrGrUrC rArCrUrGrCrArCrUrUrAr CrCrA/AltR2/
MG29-1 sgRNA targeting FAS	5276	MG29-1-FAS-sgRNA-A2	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrU rArCrUrGrUrUrGrUrArGr ArUrCrArUrCrUrGrUrCrA rCrUrGrCrArCrUrUrArCr CrArC/AltR2/
MG29-1 sgRNA targeting FAS	5277	MG29-1-FAS-sgRNA-B2	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrU rArCrUrGrUrUrGrUrArGr ArUrArGrUrArGrArCrUrG rUrUrArGrUrGrCrCrArUr GrArG/AltR2/
MG29-1 sgRNA targeting FAS	5278	MG29-1-FAS-sgRNA-C2	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrU rArCrUrGrUrUrGrUrArGr ArUrArUrUrUrUrArCrArG rGrUrUrCrUrUrArCrGrUr CrUrG/AltR2/
MG29-1 sgRNA targeting FAS	5279	MG29-1-FAS-sgRNA-D2	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrU rArCrUrGrUrUrGrUrArGr ArUrUrUrUrUrArCrArGrG rUrUrCrUrUrArCrGrUrCr UrGrU/AltR2/
MG29-1 sgRNA targeting FAS	5280	MG29-1-FAS-sgRNA-E2	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrU rArCrUrGrUrUrGrUrArGr ArUrArCrArGrGrUrUrCrU rUrArCrGrUrCrUrGrUrUr GrCrU/AltR2/
MG29-1 sgRNA targeting FAS	5281	MG29-1-FAS-sgRNA-F2	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrU rArCrUrGrUrUrGrUrArGr ArUrCrArGrGrUrUrCrUrU rArCrGrUrCrUrGrUrUrGr CrUrA/AltR2/
MG29-1 sgRNA targeting FAS	5282	MG29-1-FAS-sgRNA-G2	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrU rArCrUrGrUrUrGrUrArGr ArUrGrUrGrUrArArCrArU rArCrCrUrGrGrArGrGrAr CrArG/AltR2/

Cat.	SEQ ID:	Description	Type	Organism	Sequence
MG29-1 sgRNA targeting FAS	5283	MG29-1-FAS-sgRNA-H2	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrUrArArCrArUrArCrCrUrGrGrArGrGrArCrArGrG/AltR2/
MG29-1 sgRNA targeting FAS	5284	MG29-1-FAS-sgRNA-A3	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrArCrGrArUrArArUrCrUrArGrCrArArCrArGrArC/AltR2/
MG29-1 sgRNA targeting FAS	5285	MG29-1-FAS-sgRNA-B3	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArCrGrArUrArArUrCrUrArGrCrArArCrArGrArG/AltR2/
MG29-1 sgRNA targeting FAS	5286	MG29-1-FAS-sgRNA-C3	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrCrUrUrGrGrGrCrArGrGrUrGrArArArGrGrArA/AltR2/
MG29-1 sgRNA targeting FAS	5287	MG29-1-FAS-sgRNA-D3	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrUrUrGrGrGrCrArGrGrUrGrArArArGrGrArA/AltR2/
MG29-1 sgRNA targeting FAS	5288	MG29-1-FAS-sgRNA-E3	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrUrUrGrGrGrCrArGrGrUrGrArArArGrGrArArArG/AltR2/
MG29-1 sgRNA targeting FAS	5289	MG29-1-FAS-sgRNA-F3	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrGrGrGrCrArGrGrUrGrArArArGrGrArArArC/AltR2/
MG29-1 sgRNA targeting FAS	5290	MG29-1-FAS-sgRNA-G3	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrUrCrCrArArArUrGrCrArGrArArGrArUrGrUrA/AltR2/
MG29-1 sgRNA targeting FAS	5291	MG29-1-FAS-sgRNA-H3	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrCrCrArArArUrGrCrArGrArArGrArUrGrUrArG/AltR2/
MG29-1 sgRNA targeting FAS	5292	MG29-1-FAS-sgRNA-A4	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrCrArArArUrGrCrArGrArArGrArUrGrUrArGrA/AltR2/
MG29-1 sgRNA targeting FAS	5293	MG29-1-FAS-sgRNA-B4	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArGrArCrUrCrUrUrArCrCrArUrGrUrCrCrUrUrCrA/AltR2/
MG29-1 sgRNA targeting FAS	5294	MG29-1-FAS-sgRNA-C4	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrArCrUrCrUrUrA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					rCrCrArUrGrUrCrCrUrUrCrArU/AltR2/
MG29-1 sgRNA targeting FAS	5295	MG29-1-FAS-sgRNA-D4	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArArGrArArArArArUrGrGrGrCrUrUrUrGrUrCrUrG/AltR2/
MG29-1 sgRNA targeting FAS	5296	MG29-1-FAS-sgRNA-E4	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrGrUrGrUrArCrUrCrCrUrUrCrCrCrUrUrCrUrU/AltR2/
MG29-1 sgRNA targeting FAS	5297	MG29-1-FAS-sgRNA-F4	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArArArCrUrGrArUrUrUrUrCrUrArGrGrCrUrUrUrArG/AltR2/
MG29-1 sgRNA targeting FAS	5298	MG29-1-FAS-sgRNA-G4	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrArGrGrCrUrUrArGrArArGrUrGrGrArArArUrArA/AltR2/
MG29-1 sgRNA targeting FAS	5299	MG29-1-FAS-sgRNA-H4	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrGrCrUrUrArGrArArGrUrGrGrArArArUrArA/AltR2/
MG29-1 sgRNA targeting FAS	5300	MG29-1-FAS-sgRNA-A5	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrGrUrArArCrUrCrUrArCrUrGrUrArUrGrUrGrA/AltR2/
MG29-1 sgRNA targeting FAS	5301	MG29-1-FAS-sgRNA-B5	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrUrArArCrUrCrUrArCrUrGrUrArUrGrUrGrArA/AltR2/
MG29-1 sgRNA targeting FAS	5302	MG29-1-FAS-sgRNA-C5	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrUrArArCrUrCrUrArCrUrGrUrArUrGrUrGrArArC/AltR2/
MG29-1 sgRNA targeting FAS	5303	MG29-1-FAS-sgRNA-D5	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrArArCrUrCrUrArCrUrGrUrArUrGrUrGrArArCrA/AltR2/
MG29-1 sgRNA targeting FAS	5304	MG29-1-FAS-sgRNA-E5	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArArCrUrCrUrArCrUrGrUrArUrGrUrGrArArCrArC/AltR2/
MG29-1 sgRNA targeting FAS	5305	MG29-1-FAS-sgRNA-F5	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrUrUrArCrArUrCrUrGrCrArCrUrUrGrGrUrArUrU/AltR2/
MG29-1 sgRNA	5306	MG29-1-FAS-sgRNA-G5	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGr

Cat.	SEQ ID:	Description	Type	Organism	Sequence
targeting FAS					ArUrCrArUrCrUrGrCrArCrUrUrGrGrUrArUrUrCrUrGrGrG/AltR2/
MG29-1 sgRNA targeting FAS	5307	MG29-1-FAS-sgRNA-H5	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrUrGrUrArUrUrUrUrUrUrUrUrUrCrUrArGrArU/AltR2/
MG29-1 sgRNA targeting FAS	5308	MG29-1-FAS-sgRNA-A6	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrCrUrArGrArUrGrUrGrArArCrArUrGrGrA/AltR2/
MG29-1 sgRNA targeting FAS	5309	MG29-1-FAS-sgRNA-B6	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrUrCrUrArGrArUrGrUrGrArArCrArUrGrGrGrA/AltR2/
MG29-1 sgRNA targeting FAS	5310	MG29-1-FAS-sgRNA-C6	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrUrCrUrArGrArUrGrUrGrArArCrArUrGrGrArArU/AltR2/
MG29-1 sgRNA targeting FAS	5311	MG29-1-FAS-sgRNA-D6	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrCrUrArGrArUrGrUrGrArArCrArUrGrGrArArUrC/AltR2/
MG29-1 sgRNA targeting FAS	5312	MG29-1-FAS-sgRNA-E6	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrArGrArUrGrUrGrArArCrArUrGrGrArArUrCrA/AltR2/
MG29-1 sgRNA targeting FAS	5313	MG29-1-FAS-sgRNA-F6	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrArGrArUrGrUrGrArArCrArUrGrGrArArUrCrArU/AltR2/
MG29-1 sgRNA targeting FAS	5314	MG29-1-FAS-sgRNA-G6	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrArUrGrUrGrArArCrArUrGrGrArArUrCrArUrC/AltR2/
MG29-1 sgRNA targeting FAS	5315	MG29-1-FAS-sgRNA-H6	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArCrUrUrGrGrUrGrUrGrCrUrGrGrUrGrArGrUrG/AltR2/
MG29-1 sgRNA targeting FAS	5316	MG29-1-FAS-sgRNA-A7	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrUrCrUrUrCrUrUrUrUrUrGrCrCrArArUrUrCrCrA/AltR2/
MG29-1 sgRNA targeting FAS	5317	MG29-1-FAS-sgRNA-B7	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrCrArArUrUrCrCrArCrUrArArUrUrGrUrUrUrGrG/AltR2/

Cat.	SEQ ID:	Description	Type	Organism	Sequence
MG29-1 sgRNA targeting FAS	5318	MG29-1-FAS-sgRNA-C7	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrArArUrUrCrCrArCrUrArArUrUrGrUrUrUrGrGrG/AltR2/
MG29-1 sgRNA targeting FAS	5319	MG29-1-FAS-sgRNA-D7	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArCrArArArGrCrArArGrArArCrUrUrArCrCrCrCrA/AltR2/
MG29-1 sgRNA targeting FAS	5320	MG29-1-FAS-sgRNA-E7	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrGrUrUrCrUrUrUrCrArGrUrGrArArGrArArGrArA/AltR2/
MG29-1 sgRNA targeting FAS	5321	MG29-1-FAS-sgRNA-F7	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrCrUrUrUrCrArGrUrGrArArGrArGrArArArGrGrA/AltR2/
MG29-1 sgRNA targeting FAS	5322	MG29-1-FAS-sgRNA-G7	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrUrGrArArGrArGrArArGrUrArCrArG/AltR2/
MG29-1 sgRNA targeting FAS	5323	MG29-1-FAS-sgRNA-H7	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArUrArCrCrUrArCrArGrGrArUrUrArArArGrUrU/AltR2/
MG29-1 sgRNA targeting FAS	5324	MG29-1-FAS-sgRNA-A8	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArArGrUrUrGrGrArGrArUrUrCrArUrGrArGrArArCrC/AltR2/
MG29-1 sgRNA targeting FAS	5325	MG29-1-FAS-sgRNA-B8	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrCrUrUrUrCrUrGrUrGrCrUrUrUrCrUrGrCrArUrGrU/AltR2/
MG29-1 sgRNA targeting FAS	5326	MG29-1-FAS-sgRNA-C8	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrUrCrUrGrUrGrCrUrUrUrCrUrGrCrArUrGrUrU/AltR2/
MG29-1 sgRNA targeting FAS	5327	MG29-1-FAS-sgRNA-D8	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrUrGrCrUrUrUrCrUrGrCrArUrGrUrUrUrUrCrUrG/AltR2/
MG29-1 sgRNA targeting FAS	5328	MG29-1-FAS-sgRNA-E8	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrCrArUrGrUrUrUrUrCrUrGrUrArCrUrUrCrCrCrU/AltR2/
MG29-1 sgRNA targeting FAS	5329	MG29-1-FAS-sgRNA-F8	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrUrArCrUrUrC

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					rCrUrUrUrCrUrCrUrUrCrArCrU/AltR2/
MG29-1 sgRNA targeting FAS	5330	MG29-1-FAS-sgRNA-G8	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrCrUrUrUrCrUrArGrGrArArArCrArGrUrGrGrC/AltR2/
MG29-1 sgRNA targeting FAS	5331	MG29-1-FAS-sgRNA-H8	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrCrUrUrUrCrUrArGrGrArArArCrArGrUrGrGrCrA/AltR2/
MG29-1 sgRNA targeting FAS	5332	MG29-1-FAS-sgRNA-A9	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrUrUrUrCrUrArGrGrArArArCrArGrUrGrGrCrArA/AltR2/
MG29-1 sgRNA targeting FAS	5333	MG29-1-FAS-sgRNA-B9	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrUrUrCrUrArGrGrArArArCrArGrUrGrGrCrArArU/AltR2/
MG29-1 sgRNA targeting FAS	5334	MG29-1-FAS-sgRNA-C9	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArGrGrArArArCrArGrUrGrGrCrArArUrArArArArU/AltR2/
MG29-1 sgRNA targeting FAS	5335	MG29-1-FAS-sgRNA-D9	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrArCrUrArUrUrUrUrCrUrArUrUrUrUrUrCrArGrA/AltR2/
MG29-1 sgRNA targeting FAS	5336	MG29-1-FAS-sgRNA-E9	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrArUrUrUrUrUrCrArGrArUrGrUrUrGrArCrUrUrG/AltR2/
MG29-1 sgRNA targeting FAS	5337	MG29-1-FAS-sgRNA-F9	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArUrUrUrUrUrCrArGrArUrGrUrUrGrArCrUrUrGrA/AltR2/
MG29-1 sgRNA targeting FAS	5338	MG29-1-FAS-sgRNA-G9	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrArGrArUrGrUrUrGrArCrUrUrGrArGrUrArArArU/AltR2/
MG29-1 sgRNA targeting FAS	5339	MG29-1-FAS-sgRNA-H9	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrArUrGrUrUrGrArCrUrUrGrArGrUrArArArU/AltR2/
MG29-1 sgRNA targeting FAS	5340	MG29-1-FAS-sgRNA-A10	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrArUrGrUrUrGrArCrUrUrGrArGrUrArArArU/AltR2/
MG29-1 sgRNA	5341	MG29-1-FAS-sgRNA-B10	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGr

Cat.	SEQ ID:	Description	Type	Organism	Sequence
targeting FAS					ArUrUrUrCrGrArArArGrArArUrGrGrUrGrUrCrArArUrGrA/AltR2/
MG29-1 sgRNA targeting FAS	5342	MG29-1-FAS-sgRNA-C10	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrArCrUrCrUrUrGrCrArGrArGrArArArArUrUrCrArG/AltR2/
MG29-1 sgRNA targeting FAS	5343	MG29-1-FAS-sgRNA-D10	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrUrUrUrUrUrCrArCrUrCrUrArGrArCrCrArArG/AltR2/
MG29-1 sgRNA targeting FAS	5344	MG29-1-FAS-sgRNA-E10	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrArCrUrCrUrArGrArCrCrArArGrCrUrUrUrGrGrA/AltR2/
MG29-1 sgRNA targeting FAS	5345	MG29-1-FAS-sgRNA-F10	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArCrUrCrUrArGrArCrCrArArGrCrUrUrUrGrGrArU/AltR2/
MG29-1 sgRNA targeting FAS	5346	MG29-1-FAS-sgRNA-G10	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArCrUrCrUrArGrArCrCrArArGrCrUrUrUrGrGrArUrU/AltR2/
MG29-1 sgRNA targeting FAS	5347	MG29-1-FAS-sgRNA-H10	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArUrUrUrCrArUrUrUrCrUrGrArArGrUrUrUrGrArA/AltR2/
MG29-1 sgRNA targeting FAS	5348	MG29-1-FAS-sgRNA-A11	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrUrUrCrUrGrArArGrUrUrUrGrArArUrUrUrUrCrU/AltR2/
MG29-1 sgRNA targeting FAS	5349	MG29-1-FAS-sgRNA-B11	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrArArGrUrUrUrGrArArUrUrUrCrUrGrArGrUrC/AltR2/
MG29-1 sgRNA targeting FAS	5350	MG29-1-FAS-sgRNA-C11	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrUrUrUrCrUrGrArGrUrCrArCrUrArGrUrArArU/AltR2/
MG29-1 sgRNA targeting FAS	5351	MG29-1-FAS-sgRNA-D11	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrGrArGrUrCrArCrUrArGrUrArArUrGrUrCrCrUrU/AltR2/
MG29-1 sgRNA targeting FAS	5352	MG29-1-FAS-sgRNA-E11	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrArGrUrCrArCrUrArGrUrArArUrGrUrCrCrUrUrG/AltR2/

Cat.	SEQ ID:	Description	Type	Organism	Sequence
MG29-1 sgRNA targeting FAS	5353	MG29-1-FAS-sgRNA-F11	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrCrUrGrCrArArGrArGrUrArCrArArArGrArUrUrG/AltR2/
MG29-1 sgRNA targeting FAS	5354	MG29-1-FAS-sgRNA-G11	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrGrCrArArGrArGrUrArCrArArArGrArUrUrG/AltR2/
MG29-1 sgRNA targeting FAS	5355	MG29-1-FAS-sgRNA-H11	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrGrArGrArUrCrUrUrUrArArUrCrArArUrGrUrGrU/AltR2/
MG29-1 sgRNA targeting FAS	5356	MG29-1-FAS-sgRNA-A12	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrArGrArUrCrUrUrUrArArUrCrArArUrGrUrGrC/AltR2/
MG29-1 sgRNA targeting FAS	5357	MG29-1-FAS-sgRNA-B12	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArGrArUrCrUrUrUrArArUrCrArArUrGrUrGrUrCrA/AltR2/
MG29-1 sgRNA targeting FAS	5358	MG29-1-FAS-sgRNA-C12	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArGrArUrCrUrUrUrArArUrCrArArUrGrUrGrUrCrArU/AltR2/
MG29-1 sgRNA targeting FAS	5359	MG29-1-FAS-sgRNA-D12	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrCrArArUrGrUrGrUrCrArUrArCrGrCrUrUrCrUrU/AltR2/
MG29-1 sgRNA targeting FAS	5360	MG29-1-FAS-sgRNA-E12	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrUrUrCrCrArUrGrArArGrUrUrGrArUrGrCrCrArArU/AltR2/
MG29-1 sgRNA targeting FAS	5361	MG29-1-FAS-sgRNA-F12	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArUrGrArArGrUrUrGrArUrGrCrCrArArUrUrArCrG/AltR2/
MG29-1 sgRNA targeting FAS	5362	MG29-1-FAS-sgRNA-G12	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrUrUrCrUrGrCrUrGrUrGrUrCrUrUrGrGrArCrArU/AltR2/
MG29-1 sgRNA targeting FAS	5363	MG29-1-FAS-sgRNA-H12	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrGrCrUrUrCrArUrUrGrArCrArCrCrArUrUrCrUrUrU/AltR2/
MG29-1 sgRNA targeting FAS	5364	MG29-1-FAS-sgRNA-A13	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrUrUrCrArUrUrG

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					rArCrArCrCrArUrUrCrUrU rUrC/AltR2/
MG29-1 sgRNA targeting FAS	5365	MG29-1-FAS-sgRNA-B13	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrU rArCrUrGrUrUrGrUrArGr ArUrGrArArCrArArArGrC rCrUrUrUrArArCrUrUrGr ArCrU/AltR2/
MG29-1 sgRNA targeting FAS	5366	MG29-1-FAS-sgRNA-C13	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrU rArCrUrGrUrUrGrUrArGr ArUrArCrUrUrGrArCrUrU rArGrUrGrUrCrArUrGrAr CrUrC/AltR2/
DNA sequence of FAS target site	5367	MG29-1-FAS-target site-A1	Nucleotide	N.A.	AGGTAAACAACCGAACT GATGA
DNA sequence of FAS target site	5368	MG29-1-FAS-target site-B1	Nucleotide	N.A.	CCTGAGCAAAGACTCTT GCTAC
DNA sequence of FAS target site	5369	MG29-1-FAS-target site-C1	Nucleotide	N.A.	CAGAACGTGGCATCAAC ATCAC
DNA sequence of FAS target site	5370	MG29-1-FAS-target site-D1	Nucleotide	N.A.	TCACCAGCTCCCATGTG ATGTT
DNA sequence of FAS target site	5371	MG29-1-FAS-target site-E1	Nucleotide	N.A.	TGTGTCTATTAGATGCTC AGAG
DNA sequence of FAS target site	5372	MG29-1-FAS-target site-F1	Nucleotide	N.A.	GTGTCTATTAGATGCTCA GAGT
DNA sequence of FAS target site	5373	MG29-1-FAS-target site-G1	Nucleotide	N.A.	TGTCTATTAGATGCTCAG AGTG
DNA sequence of FAS target site	5374	MG29-1-FAS-target site-H1	Nucleotide	N.A.	GCATCTGTCACTGCACTT ACCA
DNA sequence of FAS target site	5375	MG29-1-FAS-target site-A2	Nucleotide	N.A.	CATCTGTCACTGCACTTA CCAC
DNA sequence of FAS target site	5376	MG29-1-FAS-target site-B2	Nucleotide	N.A.	AGTAGACTGTTAGTGCC ATGAG
DNA sequence of FAS target site	5377	MG29-1-FAS-target site-C2	Nucleotide	N.A.	ATTTTACAGGTTCTTACG TCTG
DNA sequence	5378	MG29-1-FAS-target site-D2	Nucleotide	N.A.	TTTTACAGGTTCTTACGT CTGT

Cat.	SEQ ID:	Description	Type	Organism	Sequence
of FAS target site					
DNA sequence of FAS target site	5379	MG29-1-FAS-target site-E2	Nucleotide	N.A.	ACAGGTTCTTACGTCTGTTGCT
DNA sequence of FAS target site	5380	MG29-1-FAS-target site-F2	Nucleotide	N.A.	CAGGTTCTTACGTCTGTTGCTA
DNA sequence of FAS target site	5381	MG29-1-FAS-target site-G2	Nucleotide	N.A.	GTGTAACATACCTGGAGGACAG
DNA sequence of FAS target site	5382	MG29-1-FAS-target site-H2	Nucleotide	N.A.	TGTAACATACCTGGAGGACAGG
DNA sequence of FAS target site	5383	MG29-1-FAS-target site-A3	Nucleotide	N.A.	GGACGATAATCTAGCAACAGAC
DNA sequence of FAS target site	5384	MG29-1-FAS-target site-B3	Nucleotide	N.A.	GACGATAATCTAGCAACAGACG
DNA sequence of FAS target site	5385	MG29-1-FAS-target site-C3	Nucleotide	N.A.	TTCCTTGGGCAGGTGAAAGGAA
DNA sequence of FAS target site	5386	MG29-1-FAS-target site-D3	Nucleotide	N.A.	TCCTTGGGCAGGTGAAAGGAAA
DNA sequence of FAS target site	5387	MG29-1-FAS-target site-E3	Nucleotide	N.A.	CCTTGGGCAGGTGAAAGGAAAG
DNA sequence of FAS target site	5388	MG29-1-FAS-target site-F3	Nucleotide	N.A.	CTTGGGCAGGTGAAAGGAAAGC
DNA sequence of FAS target site	5389	MG29-1-FAS-target site-G3	Nucleotide	N.A.	TCTTCCAAATGCAGAAGATGTA
DNA sequence of FAS target site	5390	MG29-1-FAS-target site-H3	Nucleotide	N.A.	CTTCCAAATGCAGAAGATGTAG
DNA sequence of FAS target site	5391	MG29-1-FAS-target site-A4	Nucleotide	N.A.	TTCCAAATGCAGAAGATGTAGA
DNA sequence of FAS target site	5392	MG29-1-FAS-target site-B4	Nucleotide	N.A.	AAGACTCTTACCATGTCCTTCA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
DNA sequence of FAS target site	5393	MG29-1-FAS-target site-C4	Nucleotide	N.A.	AGACTCTTACCATGTCCTTCAT
DNA sequence of FAS target site	5394	MG29-1-FAS-target site-D4	Nucleotide	N.A.	GAAGAAAAATGGGCTTTGTCTG
DNA sequence of FAS target site	5395	MG29-1-FAS-target site-E4	Nucleotide	N.A.	TCTGTGTACTCCTTCCCTTCTT
DNA sequence of FAS target site	5396	MG29-1-FAS-target site-F4	Nucleotide	N.A.	CAAACTGATTTTCTAGGCTTAG
DNA sequence of FAS target site	5397	MG29-1-FAS-target site-G4	Nucleotide	N.A.	CTAGGCTTAGAAGTGGAAATAA
DNA sequence of FAS target site	5398	MG29-1-FAS-target site-H4	Nucleotide	N.A.	TAGGCTTAGAAGTGGAAATAAA
DNA sequence of FAS target site	5399	MG29-1-FAS-target site-A5	Nucleotide	N.A.	TTTGTAACTCTACTGTATGTGA
DNA sequence of FAS target site	5400	MG29-1-FAS-target site-B5	Nucleotide	N.A.	TTGTAACTCTACTGTATGTGAA
DNA sequence of FAS target site	5401	MG29-1-FAS-target site-C5	Nucleotide	N.A.	TGTAACTCTACTGTATGTGAAC
DNA sequence of FAS target site	5402	MG29-1-FAS-target site-D5	Nucleotide	N.A.	GTAACTCTACTGTATGTGAACA
DNA sequence of FAS target site	5403	MG29-1-FAS-target site-E5	Nucleotide	N.A.	TAACTCTACTGTATGTGAACAC
DNA sequence of FAS target site	5404	MG29-1-FAS-target site-F5	Nucleotide	N.A.	GTTTACATCTGCACTTGGTATT
DNA sequence of FAS target site	5405	MG29-1-FAS-target site-G5	Nucleotide	N.A.	CATCTGCACTTGGTATTCTGGG
DNA sequence of FAS target site	5406	MG29-1-FAS-target site-H5	Nucleotide	N.A.	TCCTGTATTTTTTTTTCTAGAT
DNA sequence	5407	MG29-1-FAS-target site-A6	Nucleotide	N.A.	TTTTTCTAGATGTGAACATGGA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
of FAS target site					
DNA sequence of FAS target site	5408	MG29-1-FAS-target site-B6	Nucleotide	N.A.	TTTTCTAGATGTGAACATGGAA
DNA sequence of FAS target site	5409	MG29-1-FAS-target site-C6	Nucleotide	N.A.	TTTCTAGATGTGAACATGGAAT
DNA sequence of FAS target site	5410	MG29-1-FAS-target site-D6	Nucleotide	N.A.	TTCTAGATGTGAACATGGAATC
DNA sequence of FAS target site	5411	MG29-1-FAS-target site-E6	Nucleotide	N.A.	TCTAGATGTGAACATGGAATCA
DNA sequence of FAS target site	5412	MG29-1-FAS-target site-F6	Nucleotide	N.A.	CTAGATGTGAACATGGAATCAT
DNA sequence of FAS target site	5413	MG29-1-FAS-target site-G6	Nucleotide	N.A.	TAGATGTGAACATGGAATCATC
DNA sequence of FAS target site	5414	MG29-1-FAS-target site-H6	Nucleotide	N.A.	CACTTGGTGTTGCTGGTGAGTG
DNA sequence of FAS target site	5415	MG29-1-FAS-target site-A7	Nucleotide	N.A.	TCTTCTTCTTTTGCCAATTCCA
DNA sequence of FAS target site	5416	MG29-1-FAS-target site-B7	Nucleotide	N.A.	GCCAATTCCACTAATTGTTTGG
DNA sequence of FAS target site	5417	MG29-1-FAS-target site-C7	Nucleotide	N.A.	CCAATTCCACTAATTGTTTGGG
DNA sequence of FAS target site	5418	MG29-1-FAS-target site-D7	Nucleotide	N.A.	AACAAAGCAAGAACTTACCCCA
DNA sequence of FAS target site	5419	MG29-1-FAS-target site-E7	Nucleotide	N.A.	TTTGTTCTTTCAGTGAAGAGAA
DNA sequence of FAS target site	5420	MG29-1-FAS-target site-F7	Nucleotide	N.A.	TTCTTTCAGTGAAGAGAAAGGA
DNA sequence of FAS target site	5421	MG29-1-FAS-target site-G7	Nucleotide	N.A.	AGTGAAGAGAAAGGAAGTACAG

Cat.	SEQ ID:	Description	Type	Organism	Sequence
DNA sequence of FAS target site	5422	MG29-1-FAS-target site-H7	Nucleotide	N.A.	AATACCTACAGGATTTAAAGTT
DNA sequence of FAS target site	5423	MG29-1-FAS-target site-A8	Nucleotide	N.A.	AAGTTGGAGATTCATGAGAACC
DNA sequence of FAS target site	5424	MG29-1-FAS-target site-B8	Nucleotide	N.A.	CCTTTCTGTGCTTTCTGCATGT
DNA sequence of FAS target site	5425	MG29-1-FAS-target site-C8	Nucleotide	N.A.	CTTTCTGTGCTTTCTGCATGTT
DNA sequence of FAS target site	5426	MG29-1-FAS-target site-D8	Nucleotide	N.A.	TGTGCTTTCTGCATGTTTTCTG
DNA sequence of FAS target site	5427	MG29-1-FAS-target site-E8	Nucleotide	N.A.	TGCATGTTTTCTGTACTTCCTT
DNA sequence of FAS target site	5428	MG29-1-FAS-target site-F8	Nucleotide	N.A.	CTGTACTTCCTTTCTCTTCACT
DNA sequence of FAS target site	5429	MG29-1-FAS-target site-G8	Nucleotide	N.A.	TTGCTTTCTAGGAAACAGTGGC
DNA sequence of FAS target site	5430	MG29-1-FAS-target site-H8	Nucleotide	N.A.	TGCTTTCTAGGAAACAGTGGCA
DNA sequence of FAS target site	5431	MG29-1-FAS-target site-A9	Nucleotide	N.A.	GCTTTCTAGGAAACAGTGGCAA
DNA sequence of FAS target site	5432	MG29-1-FAS-target site-B9	Nucleotide	N.A.	CTTTCTAGGAAACAGTGGCAAT
DNA sequence of FAS target site	5433	MG29-1-FAS-target site-C9	Nucleotide	N.A.	TAGGAAACAGTGGCAATAAATT
DNA sequence of FAS target site	5434	MG29-1-FAS-target site-D9	Nucleotide	N.A.	AGACTATTTTCTATTTTTCAGA
DNA sequence of FAS target site	5435	MG29-1-FAS-target site-E9	Nucleotide	N.A.	CTATTTTTCAGATGTTGACTTG
DNA sequence	5436	MG29-1-FAS-target site-F9	Nucleotide	N.A.	TATTTTTCAGATGTTGACTTGA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
of FAS target site					
DNA sequence of FAS target site	5437	MG29-1-FAS-target site-G9	Nucleotide	N.A.	TCAGATGTTGACTTGAGTAAAT
DNA sequence of FAS target site	5438	MG29-1-FAS-target site-H9	Nucleotide	N.A.	CAGATGTTGACTTGAGTAAATA
DNA sequence of FAS target site	5439	MG29-1-FAS-target site-A10	Nucleotide	N.A.	AGATGTTGACTTGAGTAAATAT
DNA sequence of FAS target site	5440	MG29-1-FAS-target site-B10	Nucleotide	N.A.	TTCGAAAGAATGGTGTCAATGA
DNA sequence of FAS target site	5441	MG29-1-FAS-target site-C10	Nucleotide	N.A.	TACTCTTGCAGAGAAAATTCAG
DNA sequence of FAS target site	5442	MG29-1-FAS-target site-D10	Nucleotide	N.A.	TTGTTTTTCACTCTAGACCAAG
DNA sequence of FAS target site	5443	MG29-1-FAS-target site-E10	Nucleotide	N.A.	TCACTCTAGACCAAGCTTTGGA
DNA sequence of FAS target site	5444	MG29-1-FAS-target site-F10	Nucleotide	N.A.	CACTCTAGACCAAGCTTTGGAT
DNA sequence of FAS target site	5445	MG29-1-FAS-target site-G10	Nucleotide	N.A.	ACTCTAGACCAAGCTTTGGATT
DNA sequence of FAS target site	5446	MG29-1-FAS-target site-H10	Nucleotide	N.A.	GATTTCATTTCTGAAGTTTGAA
DNA sequence of FAS target site	5447	MG29-1-FAS-target site-A11	Nucleotide	N.A.	ATTTCTGAAGTTTGAATTTTCT
DNA sequence of FAS target site	5448	MG29-1-FAS-target site-B11	Nucleotide	N.A.	TGAAGTTTGAATTTTCTGAGTC
DNA sequence of FAS target site	5449	MG29-1-FAS-target site-C11	Nucleotide	N.A.	AATTTTCTGAGTCACTAGTAAT
DNA sequence of FAS target site	5450	MG29-1-FAS-target site-D11	Nucleotide	N.A.	CTGAGTCACTAGTAATGTCCTT

Cat.	SEQ ID:	Description	Type	Organism	Sequence
DNA sequence of FAS target site	5451	MG29-1-FAS-target site-E11	Nucleotide	N.A.	TGAGTCACTAGTAATGTCCTTG
DNA sequence of FAS target site	5452	MG29-1-FAS-target site-F11	Nucleotide	N.A.	CTCTGCAAGAGTACAAAGATTG
DNA sequence of FAS target site	5453	MG29-1-FAS-target site-G11	Nucleotide	N.A.	TCTGCAAGAGTACAAAGATTGG
DNA sequence of FAS target site	5454	MG29-1-FAS-target site-H11	Nucleotide	N.A.	TTGAGATCTTTAATCAATGTGT
DNA sequence of FAS target site	5455	MG29-1-FAS-target site-A12	Nucleotide	N.A.	TGAGATCTTTAATCAATGTGTC
DNA sequence of FAS target site	5456	MG29-1-FAS-target site-B12	Nucleotide	N.A.	GAGATCTTTAATCAATGTGTCA
DNA sequence of FAS target site	5457	MG29-1-FAS-target site-C12	Nucleotide	N.A.	AGATCTTTAATCAATGTGTCAT
DNA sequence of FAS target site	5458	MG29-1-FAS-target site-D12	Nucleotide	N.A.	ATCAATGTGTCATACGCTTCTT
DNA sequence of FAS target site	5459	MG29-1-FAS-target site-E12	Nucleotide	N.A.	TTTCCATGAAGTTGATGCCAAT
DNA sequence of FAS target site	5460	MG29-1-FAS-target site-F12	Nucleotide	N.A.	CATGAAGTTGATGCCAATTACG
DNA sequence of FAS target site	5461	MG29-1-FAS-target site-G12	Nucleotide	N.A.	TGTTCTGCTGTGTCTTGGACAT
DNA sequence of FAS target site	5462	MG29-1-FAS-target site-H12	Nucleotide	N.A.	GGCTTCATTGACACCATTCTTT
DNA sequence of FAS target site	5463	MG29-1-FAS-target site-A13	Nucleotide	N.A.	GCTTCATTGACACCATTCTTTC
DNA sequence of FAS target site	5464	MG29-1-FAS-target site-B13	Nucleotide	N.A.	GAACAAAGCCTTTAACTTGACT
DNA sequence	5465	MG29-1-FAS-target site-C13	Nucleotide	N.A.	ACTTGACTTAGTGTCATGACTC

Cat.	SEQ ID:	Description	Type	Organism	Sequence
of FAS target site					
MG29-1 sgRNA targeting PD-1	5466	MG29-1-PD-1-sgRNA-A1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrUrArGrCrGrGrArArUrGrGrGrCrArCrCrUrCrArUrC/AltR2/
MG29-1 sgRNA targeting PD-1	5467	MG29-1-PD-1-sgRNA-B1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrCrArGrUrGrGrCrGrArGrArGrArArGrArCrCrCrCrGrG/AltR2/
MG29-1 sgRNA targeting PD-1	5468	MG29-1-PD-1-sgRNA-C1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrCrUrCrArArArGrArArGrGrArGrGrArCrCrCrCrU/AltR2/
MG29-1 sgRNA targeting PD-1	5469	MG29-1-PD-1-sgRNA-D1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrCrUrGrCrArGrGrGrArCrArArUrArGrGrArGrCrCrA/AltR2/
MG29-1 sgRNA targeting PD-1	5470	MG29-1-PD-1-sgRNA-E1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArArCrUrGrGrCrCrGrGrCrUrGrGrCrCrUrGrGrU/AltR2/
MG29-1 sgRNA targeting PD-1	5471	MG29-1-PD-1-sgRNA-F1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrUrGrCrCrCrUrUrCrCrArGrArGrArArGrArArGrGrGrCrA/AltR2/
MG29-1 sgRNA targeting PD-1	5472	MG29-1-PD-1-sgRNA-G1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrArUrCrUrGrCrGrCrCrUrUrGrGrGrGrCrCrArGrGrG/AltR2/
MG29-1 sgRNA targeting PD-1	5473	MG29-1-PD-1-sgRNA-H1	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrCrArCrGrArArGrCrUrCrUrCrCrGrArUrGrUrGrUrU/AltR2/
DNA sequence of PD-1 target site	5474	MG29-1-PD-1-target site-A1	Nucleotide	N.A.	ATCTGCGCCTTGGGGGCCAGGG
DNA sequence of PD-1 target site	5475	MG29-1-PD-1-target site-B1	Nucleotide	N.A.	GCACGAAGCTCTCCGATGTGTT
DNA sequence of PD-1 target site	5476	MG29-1-PD-1-target site-C1	Nucleotide	N.A.	GAACTGGCCGGCTGGCCTGGGT
DNA sequence of PD-1 target site	5477	MG29-1-PD-1-target site-D1	Nucleotide	N.A.	TGCCCTTCCAGAGAGAAGGGCA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
DNA sequence of PD-1 target site	5478	MG29-1-PD-1-target site-E1	Nucleotide	N.A.	TCTGCAGGGACAATAGGAGCCA
DNA sequence of PD-1 target site	5479	MG29-1-PD-1-target site-F1	Nucleotide	N.A.	TCCTCAAAGAAGGAGGACCCCT
DNA sequence of PD-1 target site	5480	MG29-1-PD-1-target site-G1	Nucleotide	N.A.	CAGTGGCGAGAGAAGACCCCGG
DNA sequence of PD-1 target site	5481	MG29-1-PD-1-target site-H1	Nucleotide	N.A.	CTAGCGGAATGGGCACCTCATC
MG29-1 containing 5'UTR, NLS, CDS, NLS, 3'UTR, polyA tail	5680	MG29-1 mRNA	Nucleotide	N.A.	GAAAAGCCAGCTCCAGCAGGCGCTGCTCACTCCTCCCCATCCTCTCCCTCTGTCCCTCTGTCCCTCTGACCCTGCACTGTCCCAGCACCATGGCCCCTAAGAAGAAGAGAAAAGTCGGCGGAGGCGGCAGCTTCAACAACTTCATCAAGAAGTACAGCCTGCAGAAAACCCTGCGCTTCGAGCTGAAGCCTGTGGGCGAGACAGCCGACTACATCGAGGACTTCAAGAGCGAGTACCTGAAGGACACCGTGCTGAAGGACGAGCAGAGAGCCAAGGACTACCAAGAGATCAAGACCCTGATCGACGATTACCACCGCGAGTACATCGAAGAGTGCCTGAGAGAACCCGTGGACAAGAAAACCGGCGAGATCCTGGACTTCACCCAGGACCTGGAAGATGCCTTCAGCTACTACCAGAAGCTGAAAGAGAACCCCACCGAGAACAGAGTCGGCTGGGAGAAAGAGCAAGAGAGCCTGAGGAAGAAGCTGGTCACCTCCTTCGTGGGCAACGACGGCCTGTTCAAGAAAGAGTTCATCACCAGGGACCTGCCTGAGTGGCTGCAGAAGAAAGGACTCTGGGGCGAGTACAAGGACACAGTGGAAAACTTCAAGAAGTTCACCACCTACTTCAGCGGCTTCCACGAGAACCGGAAGAACATGTACACCGCCGAGGCTCAGAGCACCGCTATCGCCAACAGACTGATGAACGACAACCTGCCTAAGTTCTTTAACAACTACCTGGCCTACCAGACC

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					ATCAAAGAGAAGCACCC
					CGACCTGGTGTTCAGAC
					TGGATGATGCTCTGCTG
					CAGGCCGCTGGCGTGGA
					ACATCTGGATGAGGCTT
					TCCAGCCTAGATACTTCA
					GCAGACTGTTCGCCCAG
					AGCGGCATCACCGCTTT
					CAACGAGCTGATCGGCG
					GCAGAACCACAGAGAAC
					GGCGAGAAGATCCAGGG
					CCTGAACGAGCAGATCA
					ACCTGTACAGACAGCAG
					AACCCCGAGAAGGCCAA
					GGGCTTCCCCAGATTCA
					TGCCTCTGTTCAAGCAG
					ATCCTGAGCGACAGAGA
					GACACACAGCTTTCTGC
					CCGACGCCTTCGAGAAC
					GACAAAGAGCTGCTCCA
					GGCTCTGAGAGACTACG
					TGGACGCCGCCACATCT
					GAGGAAGGCATGATCAG
					CCAGCTGAACAAGGCCA
					TGAACCAGTTCGTGACC
					GCCGACCTGAAGAGAGT
					GTACATCAAGAGCGCCG
					CTCTGACCAGCCTGAGC
					CAAGAGCTGTTCCACTT
					CTTCGGCGTGATCAGCG
					ACGCTATCGCTTGGTAC
					GCCGAGAAGAGACTGAG
					CCCCAAGAAGGCCCAAG
					AGTCTTTCCTGAAGCAA
					GAGGTGTACGCCATCGA
					GGAACTGAACCAGGCTG
					TCGTGGGCTACATCGAC
					CAGCTGGAAGATCAGAG
					CGAGCTGCAGCAACTGC
					TGGTGGACCTGCCAGAT
					CCTCAGAAACCCGTGTC
					CAGCTTCATCCTGACAC
					ACTGGCAGAAGTCTCAA
					GAGCCCCTGCAGGCAGT
					GATCGCCAAGGTGGAAC
					CTCTGTTCGAACTGGAA
					GAACTGAGCAAGAACAA
					GAGGGCCCCAAAGCACG
					ACAAGGACCAAGGCGGC
					GAGGGATTTCAGCAGGT
					CGACGCCATCAAGAACA
					TGCTGGACGCCTTCATG
					GAAGTGTCCCACGCTAT
					CAAGCCCCTGTACCTGG
					TCAAGGGAAGAAAGGCC
					ATCGACATGCCCGACGT
					GGACACCGGCTTCTACG
					CTGATTTCGCCGAGGCC
					TACAGCGCCTACGAGCA
					AGTGACAGTGTCCCTGT

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					ACAACAAGACCAGAAAC CACCTGTCCAAGAAGCC CTTCAGCAAGGACAAGA TCAAGATCAACTTCGAC GCCCCTACACTGCTGAA CGGCTGGGACCTGAACA AAGAGAGCGACAACAAG TCCATCATCCTGCGGAA GGACGGCAACTTCTACC TGGCAATCATGCACCCC AAGCACACCAAGGTGTT CGACTGCTACTCTGCCT CTGAGGCTGCCGGCAAG TGCTACGAGAAGATGAA CTACAAGCTGCTGAGCG GCGCCAACAAGATGCTG CCTAAGGTGTTCTTTAGC AAGAAGGGCATCGAGAC ATTCAGCCCTCCACAAG AAATCCTGGACCTGTAC AAGAACAACGAGCATAA GAAGGGCGCCACCTTCA AGCTGGAATCCTGCCAC AAGCTGATCGATTTCTTC AAGCGGAACATCCCCAA GTACAAGGTGCACCCTA CCGACAACTTTGGCTGG GACGTGTTCGGCTTTCA CTTCAGCCCTACCAGCA GCTACGGCGACCTGTCT GGCTTCTACAGAGAGGT GGAAGCCCAGGGATACA AGCTGTGGTTCAGCGAC GTGTCCGAGGCTTACAT CAACAAATGCGTGGAAG AGGGCAAGCTGTTCCTG TTCCAAATCTACAACAAG GACTTCTCCCCTAACTCC ACCGGCAAGCCCAACCT GCACACCCTGTATTGGA AGGGCCTGTTCGAGCCC GAGAACCTGAAAGACGT GGTGCTGAAGCTGAATG GCGAGGCCGAGATCTTC TACCGGAAGCACAGCAT CAAGCACGAGGACAAGA CCATCCACAGAGCTAAG GACCCTATCGCTAACAA GAACGCTGACAACCCCA AGAAACAGAGCGTGTTC GATTACGACATCATCAA GGATAAGCGGTATACCC AGGACAAGTTCTTCTTCC ACGTGCCAATCAGCCTG AACTTCAAAAGCCAGGG CGTCGTGCGGTTCAACG ATAAGATCAACGGCCTG CTGGCCGCTCAGGACGA TGTGCATGTGATCGGCA TCGACAGAGGCGAGAGA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					CATCTGCTGTACTACACC GTGGTCAACGGCAAGGG CGAAGTGGTGGAACAGG GCAGCCTGAATCAGGTG GCCACAGATCAGGGCTA CGTGGTGGATTACCAGC AGAAGCTGCACGCCAAA GAGAAAGAACGCGACCA GGCCAGAAAGAACTGGT CCACCATCGAGAACATC AAAGAACTGAAGGCCGG CTACCTGAGCCAGGTGG TGCATAAGCTGGCTCAG CTGATCGTGAAGCACAA CGCCATCGTGTGCCTCG AGGACCTGAATTTCGGC TTCAAGAGGGGCAGATT CAAGGTCGAGAAACAGG TGTACCAGAAGTTCGAG AAGGCTCTGATCGACAA GCTGAACTACCTCGTGT TCAAAGAGAGAGGCGCC ACACAGGCTGGCGGATA CCTGAATGCTTACCAGC TGGCCGCACCTTTCGAG AGCTTTGAGAAGCTGGG CAAGCAGACCGGCATCC TGTACTACGTGCGGAGC GACTACACCAGCAAGAT CGACCCTGCTACCGGCT TCGTGGACTTTCTGAAG CCTAAGTACGAGAGCAT GGCCAAGAGCAAAGTGT TCTTCGAGTCCTTCGAG CGCATCCAGTGGAACCA GGCCAAAGGCTACTTCG AGTTCGAGTTTGACTAC AAGAAGATGTGCCCCAG CAGAAAGTTCGGCGACT ACAGAACCAGATGGGTC GTGTGCACCTTCGGCGA CACCCGCTACCAGAACA GAAGAAACAAGAGCAGC GGCCAGTGGGAGACAGA GACAATCGATGTGACAG CCCAGCTGAAAGCCCTG TTCGCCGCTTACGGCAT CACATACAATCAAGAGG ATAACATCAAGGACGCC ATTGCCGCCGTGAAGTA CACCAAGTTCTACAAGC AGCTGTACTGGCTGCTG AGACTGACCCTGAGCCT GAGACACAGCGTGACAG GCACCGACGAGGATTTC ATCCTGTCTCCAGTGGC CGACGAGAATGGCGTGT TCTTTGACTCTAGGAAG GCCACCGACAAGCAGCC TAAGGACGCTGATGCTA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					ACGGCGCCTACCATATC GCCCTGAAAGGCCTGTG GAATCTCCAGCAGATCA GACAGCACGACTGGAAC GTGGAAAAGCCCAAAAA GCTGAACCTCGCCATGA AGAACGAAGAGTGGTTC GGCTTCGCTCAGAAGAA GAAGTTTAGAGCCAGCG GCGGCAAGAGGCCTGCC GCTACAAAAAAAGCCGG CCAGGCCAAGAAAAAGA AGTGACCACACCCCCAT TCCCCCACTCCAGATAG AACTTCAGTTATATCTCA CGTGTCTGGAGTTGGAT CCCTTGAAGACTAAAAA AAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAA A
MG29-1 sgRNA targeting TRAC	5681	MG29-1-TRAC-sgRNA-35	Nucleotide	N.A.	mU*rArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArGrUrCrUrCrUrCrArGrCrUrGrGrUrArCrArCrG*mG
DNA sequence of TRAC target site	5682	MG29-1-TRAC-target site-35	Nucleotide	N.A.	GAGTCTCTCAGCTGGTACACGG
MG29-1 sgRNA targeting TRAC	5683	MG29-1-TRAC-sgRNA-35-AltR	Nucleotide	N.A.	/AltR1/rUrArArUrUrUrCrUrArCrUrGrUrUrGrUrArGrArUrGrArGrUrCrUrCrUrCrArGrCrUrGrGrUrArCrArCrG/AltR2/
DNA sequence of TRAC target site	5684	MG29-1-TRAC-target site-35-AltR	Nucleotide	N.A.	GAGTCTCTCAGCTGGTACACGG
MG29-1 sgRNA targeting CD38	5685	MG29-1-CD38-sgRNA	Nucleotide	N.A.	mC*mU*mU*rU*rUrArArUrUmUmCmUmArCrU*rG*rU*rU*rGrUrArGrArUrCrCrGrArGrArC/i2FC//i2FG//i2FU//i2FC//i2FC//i2FU//i2FG//i2FG//i2FC//i2FG/*/i2FC/*/i2FG//i2FA/*/i2FU/*mG
DNA sequence of CD38 target site	5686	MG29-1-CD38-target site	Nucleotide	N.A.	CCGAGACCGTCCTGGCGCGATG
MG29-1 coding sequence used for generation of	5687	MG29-1-mRNA coding sequence	Nucleotide	N.A.	ATGGCCCCTAAGAAGAAGAGAAAAGTCGGCGGAGGCGGCAGCTTCAACAACTTCATCAAGAAGTACAGCCTGCAGAAAACCCTGCGCTTCGAGCTGAAGCCT

Cat.	SEQ ID:	Description	Type	Organism	Sequence
mRNA					GTGGGCGAGACAGCCGA CTACATCGAGGACTTCA AGAGCGAGTACCTGAAG GACACCGTGCTGAAGGA CGAGCAGAGAGCCAAGG ACTACCAAGAGATCAAG ACCCTGATCGACGATTA CCACCGCGAGTACATCG AAGAGTGCCTGAGAGAA CCCGTGGACAAGAAAAC CGGCGAGATCCTGGACT TCACCCAGGACCTGGAA GATGCCTTCAGCTACTA CCAGAAGCTGAAAGAGA ACCCCACCGAGAACAGA GTCGGCTGGGAGAAAGA GCAAGAGAGCCTGAGGA AGAAGCTGGTCACCTCC TTCGTGGGCAACGACGG CCTGTTCAAGAAAGAGT TCATCACCAGGGACCTG CCTGAGTGGCTGCAGAA GAAAGGACTCTGGGGCG AGTACAAGGACACAGTG GAAAACTTCAAGAAGTT CACCACCTACTTCAGCG GCTTCCACGAGAACCGG AAGAACATGTACACCGC CGAGGCTCAGAGCACCG CTATCGCCAACAGACTG ATGAACGACAACCTGCC TAAGTTCTTTAACAACTA CCTGGCCTACCAGACCA TCAAAGAGAAGCACCCC GACCTGGTGTTCAGACT GGATGATGCTCTGCTGC AGGCCGCTGGCGTGGAA CATCTGGATGAGGCTTT CCAGCCTAGATACTTCA GCAGACTGTTCGCCCAG AGCGGCATCACCGCTTT CAACGAGCTGATCGGCG GCAGAACCACAGAGAAC GGCGAGAAGATCCAGGG CCTGAACGAGCAGATCA ACCTGTACAGACAGCAG AACCCCGAGAAGGCCAA GGGCTTCCCCAGATTCA TGCCTCTGTTCAAGCAG ATCCTGAGCGACAGAGA GACACACAGCTTTCTGC CCGACGCCTTCGAGAAC GACAAAGAGCTGCTCCA GGCTCTGAGAGACTACG TGGACGCCGCCACATCT GAGGAAGGCATGATCAG CCAGCTGAACAAGGCCA TGAACCAGTTCGTGACC GCCGACCTGAAGAGAGT GTACATCAAGAGCGCCG

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					CTCTGACCAGCCTGAGC CAAGAGCTGTTCCACTT CTTCGGCGTGATCAGCG ACGCTATCGCTTGGTAC GCCGAGAAGAGACTGAG CCCCAAGAAGGCCCAAG AGTCTTTCCTGAAGCAA GAGGTGTACGCCATCGA GGAACTGAACCAGGCTG TCGTGGGCTACATCGAC CAGCTGGAAGATCAGAG CGAGCTGCAGCAACTGC TGGTGGACCTGCCAGAT CCTCAGAAACCCGTGTC CAGCTTCATCCTGACAC ACTGGCAGAAGTCTCAA GAGCCCCTGCAGGCAGT GATCGCCAAGGTGGAAC CTCTGTTCGAACTGGAA GAACTGAGCAAGAACAA GAGGGCCCCAAAGCACG ACAAGGACCAAGGCGGC GAGGGATTTCAGCAGGT CGACGCCATCAAGAACA TGCTGGACGCCTTCATG GAAGTGTCCCACGCTAT CAAGCCCCTGTACCTGG TCAAGGGAAGAAAGGCC ATCGACATGCCCGACGT GGACACCGGCTTCTACG CTGATTTCGCCGAGGCC TACAGCGCCTACGAGCA AGTGACAGTGTCCCTGT ACAACAAGACCAGAAAC CACCTGTCCAAGAAGCC CTTCAGCAAGGACAAGA TCAAGATCAACTTCGAC GCCCCTACACTGCTGAA CGGCTGGGACCTGAACA AAGAGAGCGACAACAAG TCCATCATCCTGCGGAA GGACGGCAACTTCTACC TGGCAATCATGCACCCC AAGCACACCAAGGTGTT CGACTGCTACTCTGCCT CTGAGGCTGCCGGCAAG TGCTACGAGAAGATGAA CTACAAGCTGCTGAGCG GCGCCAACAAGATGCTG CCTAAGGTGTTCTTTAGC AAGAAGGGCATCGAGAC ATTCAGCCCTCCACAAG AAATCCTGGACCTGTAC AAGAACAACGAGCATAA GAAGGGCGCCACCTTCA AGCTGGAATCCTGCCAC AAGCTGATCGATTTCTTC AAGCGGAACATCCCCAA GTACAAGGTGCACCCTA CCGACAACTTTGGCTGG

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					GACGTGTTCGGCTTTCA CTTCAGCCCTACCAGCA GCTACGGCGACCTGTCT GGCTTCTACAGAGAGGT GGAAGCCCAGGGATACA AGCTGTGGTTCAGCGAC GTGTCCGAGGCTTACAT CAACAAATGCGTGGAAG AGGGCAAGCTGTTCCTG TTCCAAATCTACAACAAG GACTTCTCCCCTAACTCC ACCGGCAAGCCCAACCT GCACACCCTGTATTGGA AGGGCCTGTTCGAGCCC GAGAACCTGAAAGACGT GGTGCTGAAGCTGAATG GCGAGGCCGAGATCTTC TACCGGAAGCACAGCAT CAAGCACGAGGACAAGA CCATCCACAGAGCTAAG GACCCTATCGCTAACAA GAACGCTGACAACCCCA AGAAACAGAGCGTGTTC GATTACGACATCATCAA GGATAAGCGGTATACCC AGGACAAGTTCTTCTTCC ACGTGCCAATCAGCCTG AACTTCAAAAGCCAGGG CGTCGTGCGGTTCAACG ATAAGATCAACGGCCTG CTGGCCGCTCAGGACGA TGTGCATGTGATCGGCA TCGACAGAGGCGAGAGA CATCTGCTGTACTACACC GTGGTCAACGGCAAGGG CGAAGTGGTGGAACAGG GCAGCCTGAATCAGGTG GCCACAGATCAGGGCTA CGTGGTGGATTACCAGC AGAAGCTGCACGCCAAA GAGAAAGAACGCGACCA GGCCAGAAAGAACTGGT CCACCATCGAGAACATC AAAGAACTGAAGGCCGG CTACCTGAGCCAGGTGG TGCATAAGCTGGCTCAG CTGATCGTGAAGCACAA CGCCATCGTGTGCCTCG AGGACCTGAATTTCGGC TTCAAGAGGGGCAGATT CAAGGTCGAGAAACAGG TGTACCAGAAGTTCGAG AAGGCTCTGATCGACAA GCTGAACTACCTCGTGT TCAAAGAGAGAGGCGCC ACACAGGCTGGCGGATA CCTGAATGCTTACCAGC TGGCCGCACCTTTCGAG AGCTTTGAGAAGCTGGG CAAGCAGACCGGCATCC

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					TGTACTACGTGCGGAGC GACTACACCAGCAAGAT CGACCCTGCTACCGGCT TCGTGGACTTTCTGAAG CCTAAGTACGAGAGCAT GGCCAAGAGCAAAGTGT TCTTCGAGTCCTTCGAG CGCATCCAGTGGAACCA GGCCAAAGGCTACTTCG AGTTCGAGTTTGACTAC AAGAAGATGTGCCCCAG CAGAAAGTTCGGCGACT ACAGAACCAGATGGGTC GTGTGCACCTTCGGCGA CACCCGCTACCAGAACA GAAGAAACAAGAGCAGC GGCCAGTGGGAGACAGA GACAATCGATGTGACAG CCCAGCTGAAAGCCCTG TTCGCCGCTTACGGCAT CACATACAATCAAGAGG ATAACATCAAGGACGCC ATTGCCGCCGTGAAGTA CACCAAGTTCTACAAGC AGCTGTACTGGCTGCTG AGACTGACCCTGAGCCT GAGACACAGCGTGACAG GCACCGACGAGGATTTC ATCCTGTCTCCAGTGGC CGACGAGAATGGCGTGT TCTTTGACTCTAGGAAG GCCACCGACAAGCAGCC TAAGGACGCTGATGCTA ACGGCGCCTACCATATC GCCCTGAAAGGCCTGTG GAATCTCCAGCAGATCA GACAGCACGACTGGAAC GTGGAAAAGCCCAAAAA GCTGAACCTCGCCATGA AGAACGAAGAGTGGTTC GGCTTCGCTCAGAAGAA GAAGTTTAGAGCCAGCG GCGGCAAGAGGCCTGCC GCTACAAAAAAAGCCGG CCAGGCCAAGAAAAAGA AGTGA
MG29-1 sgRNA targeting mouse albumin	5688	mAlb29-8-44	Nucleotide	N.A.	mC*mU*mU*U*U*AAUUm UmCmUmACU*G*U*U*GU AGAUCUGUAACfGfAfUfCf GfGfGfA*fAfC*fU*fG*fG*f C*mA
MG29-1 sgRNA targeting mouse albumin	5689	mAlb29-8-50	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*m G*mA*mA*mAGAUUCUCA AC*mC*mU*mU*U*UAAU UmUmCmUmACU*G*U*U* GUAGAUCUGUAACfGfAfU fCfGfGfGfAfAfC*fU*fGfG*f C*mA
MG29-1 sgRNA	5690	mAlb29-8-50b	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*m G*mA*mA*mAGAUUCUCA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
targeting mouse albumin					AC*mC*mU*mU*U*UAAU UmUmCmUmACU*G*U*U* GUAGAUCUGUAACfGfAfU fCfGfGfGfAfAfC*fU*fG*mG
MG29-1 sgRNA targeting mouse albumin	5691	mAlb29-8-51b	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*m G*mA*mA*mAGAUUCUCA AC*mC*mU*mU*U*UAAU UmUmCmUmACU*G*U*U* GUAGAUCUGUAACGAUC GGGAAC*U*mG*mG
MG29-1 sgRNA targeting mouse albumin	5692	mAlb29-8-52b	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*m G*mA*mA*mAGAUUCUCA AC*mC*mU*mU*U*UAAU UmUmCmUmACU*G*U*U* GUAGAUCUGUAACGfAUf CGfGGfA*A*fCU*fG*mG
MG29-1 sgRNA targeting mouse albumin	5693	mAlb29-8-53b	Nucleotide	N.A.	mG*mU*mU*mG*mA*mG* mA*mA*mU*mC*mG*mA* mA*mA*mG*mA*mU*mU* mC*mU*mC*mA*mA*mC* mC*mU*mU*U*UAAUUmU mCmUmACU*G*U*U*GUA GAUCUGUAACfGfAfUfCfG fGfGfAfAfC*fU*fG*mG
MG29-1 sgRNA targeting mouse albumin	5694	mAlb29-8-54b	Nucleotide	N.A.	mG*mU*mU*mG*mA*mG* mA*mA*mU*mC*mG*mA* mA*mA*mG*mA*mU*mU* mC*mU*mC*mA*mA*mC* mC*mU*mU*U*UAAUUmU mCmUmACU*G*U*U*GUA GAUCUGUAACGAUCGGG AAC*U*mG*mG
Guide modification chemistry	5695	Chemistry 44 (22 nt spacer)	Nucleotide	N.A.	mC*mU*mU*U*U*AAUUm UmCmUmACU*G*U*U*GU AGAUNNNNNNNfNfNfNfNf NfNfNfN*fNfN*fN*fN*fN*fN *mN
Guide modification chemistry	5696	Chemistry 50 (22 nt spacer)	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*m G*mA*mA*mAGAUUCUCA AC*mC*mU*mU*U*UAAU UmUmCmUmACU*G*U*U* GUAGAUNNNNNNNfNfNfN fNfNfNfNfNfNfN*fN*fNfN*f N*mN
Guide modification chemistry	5697	Chemistry 50 (20 nt spacer)	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*m G*mA*mA*mAGAUUCUCA AC*mC*mU*mU*U*UAAU UmUmCmUmACU*G*U*U* GUAGAUNNNNNNNfNfNfN fNfNfNfNfNfNfN*fN*fN*mN
Guide modification chemistry	5698	Chemistry 51 (20 nt spacer)	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*m G*mA*mA*mAGAUUCUCA AC*mC*mU*mU*U*UAAU UmUmCmUmACU*G*U*U* GUAGAUNNNNNNNNNNN NNNNNN*N*mN*mN
Guide modification	5699	Chemistry 52 (20 nt spacer)	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*m G*mA*mA*mAGAUUCUCA AC*mC*mU*mU*U*UAAU

Cat.	SEQ ID:	Description	Type	Organism	Sequence
chemistry					UmUmCmUmACU*G*U*U*GUAGAUNNNNNNNNfNNfNNfN*N*fNN*fN*mN
Guide modification chemistry	5700	Chemistry 53 (20 nt spacer)	Nucleotide	N.A.	mG*mU*mU*mG*mA*mG*mA*mA*mU*mC*mG*mA*mA*mA*mG*mA*mU*mU*mC*mU*mC*mA*mA*mC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUNNNNNNNfNfNfNfNfNfNfNfNfN*fN*fN*mN
Guide modification chemistry	5701	Chemistry 54 (20 nt spacer)	Nucleotide	N.A.	mG*mU*mU*mG*mA*mG*mA*mA*mU*mC*mG*mA*mA*mA*mG*mA*mU*mU*mC*mU*mC*mA*mA*mC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUNNNNNNNNNNNNNNNNNN*N*mN*mN
MG29-1 sgRNA targeting mouse albumin	5702	mAlb29-8-37	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUCUGUAACfGfAfUfCfGfGfGfAfAfC*fU*fGfG*fC*mA
MG29-1 sgRNA targeting mouse albumin	5703	mAlb29-12-44	Nucleotide	N.A.	mC*mU*mU*U*U*AAUUmUmCmUmACU*G*U*U*GUAGAUAGUGUAGfCfAfGfAfGfAfGfG*fAfA*fC*fC*fA*fU*mU
MG29-1 sgRNA targeting mouse HAO1	5704	mH29-29-50	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUCCUUAGGfAfGfAfAfAfAfUfGfCfC*fA*fAfA*fU*mC
MG29-1 sgRNA targeting mouse HAO1	5705	mH29-29-50b	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUCCUUAGGfAfGfAfAfAfAfUfGfCfC*fA*fA*mA
MG29-1 sgRNA targeting mouse HAO1	5706	mH29-29-51b	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUCCUUAGGAGAAAAUGCC*A*mA*mA
MG29-1 sgRNA targeting mouse HAO1	5707	mH29-29-52b	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUCCUUAGGAfGAfAAfAUfG*C*fCA*fA*mA
MG29-1 sgRNA targeting mouse HAO1	5708	mH29-29-53b	Nucleotide	N.A.	mG*mU*mU*mG*mA*mG*mA*mA*mU*mC*mG*mA*mA*mA*mG*mA*mU*mU*mC*mU*mC*mA*mA*mC*mC*mU*mU*U*UAAUUmU

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					mCmUmACU*G*U*U*GUA GAUCCUUAGGfAfGfAfAfA fAfUfGfCfC*fA*fA*mA
MG29-1 sgRNA targeting mouse HAO1	5709	mH29-29-54b	Nucleotide	N.A.	mG*mU*mU*mG*mA*mG* mA*mA*mU*mC*mG*mA* mA*mA*mG*mA*mU*mU* mC*mU*mC*mA*mA*mC* mC*mU*mU*U*UAAUUmU mCmUmACU*G*U*U*GUA GAUCCUUAGGAGAAAAU GCC*A*mA*mA
MG29-1 sgRNA targeting mouse HAO1	5710	mH29-29.1_37	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmU mCmUmACU*G*U*U*GUA GAUCCUUAGGfAfGfAfAfA fAfUfG*fCfC*fA*fA*mA
MG29-1 sgRNA targeting mouse HAO1	5711	mH29-29.2_37	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmU mCmUmACU*G*U*U*GUA GAUCCUUAGGfAfGfAfAfA fAfUfG*fC*fC*fA*fA*mA
MG29-1 CRISPR array targeting mouse albumin	5712	mAlb29-g8-37-array	Nucleotide	N.A.	mG*mU*mC*UAAGACCU mU*mA*mC*U*U*AAUUm UmCmUmACU*G*U*U*GU AGAUCUGUAACfGfAfUfCf GfGfGfA*fAfC*fU*fG*fG*f C*mAUCUUCAGUCUAAG ACCUmU*mA*mC*U*U*A AUUmUmCmUmACU*G*U* U*GUAGAUCUGUAACfGf AfUfCfGfGfGfA*fAfC*fU*f G*fG*fC*mAUCU*mU*mC* mA
MG29-1 sgRNA targeting human albumin	5713	hA29-87-37B	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmU mCmUmACU*G*U*U*GUA GAUCGCACUAfAfGfGfAfA fAfGfU*fGfC*fA*fA*mA
MG29-1 sgRNA targeting human albumin	5714	hA29-78-37B	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmU mCmUmACU*G*U*U*GUA GAUUUUGCGfCfAfCfUfA fAfGfG*fAfA*fA*fG*mU
MG29-1 sgRNA targeting human albumin	5715	hA29-74-37B	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmU mCmUmACU*G*U*U*GUA GAUAAUAAAGfCfAfUfAfG fUfGfC*fAfA*fU*fG*mG
MG29-1 sgRNA targeting human albumin	5716	hA29-83-37B	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmU mCmUmACU*G*U*U*GUA GAUUGAGAUCfAfAfCfAfG fCfAfC*fAfG*fG*fU*mU
MG29-1 sgRNA targeting human albumin	5717	hA29-84-37B	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmU mCmUmACU*G*U*U*GUA GAUCACUUUCfCfUfUfAfG fUfGfC*fGfC*fA*fA*mA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
MG29-1 sgRNA targeting human HAO1	5718	hH29-4_37b	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUCCCCAGAfCfCfUfGfUfAfAfU*fAfG*fU*fC*mA
MG29-1 sgRNA targeting human HAO1	5719	hH29-21_37b	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUGGACAGAfGfGfGfUfCfAfGfC*fAfU*fG*fC*mC
MG29-1 sgRNA targeting human HAO1	5720	hH29-23_37b	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUUCAGCCUfGfUfCfAfGfUfCfC*fCfU*fG*fG*mG
MG29-1 sgRNA targeting human HAO1	5721	hH29-41_37b	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUAUAUCUUfCfCfCfAfGfCfUfG*fAfU*fA*fG*mA
MG29-1 sgRNA targeting human HAO1	5722	hH29-4_50	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUCCCCAGAfCfCfUfGfUfAfAfUfAfG*fU*fCfA*fU*mA
MG29-1 sgRNA targeting human HAO1	5723	hH29-21_50	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUGGACAGAfGfGfGfUfCfAfGfCfAfU*fG*fCfC*fA*mA
MG29-1 sgRNA targeting human HAO1	5724	hH29-23_50	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUUCAGCCUfGfUfCfAfGfUfCfCfCfU*fG*fGfG*fA*mA
MG29-1 sgRNA targeting human HAO1	5725	hH29-41_50	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUAUAUCUUfCfCfCfAfGfCfUfGfAfU*fA*fGfA*fU*mG
MG29-1 sgRNA targeting human HAO1	5726	hH29-4_50b	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUCCCCAGAfCfCfUfGfUfAfAfU*fAfG*fU*fC*mA
MG29-1 sgRNA targeting human	5727	hH29-21_50b	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*

Cat.	SEQ ID:	Description	Type	Organism	Sequence
HAO1					GUAGAUGGACAGAfGfGfGfUfCfAfGfC*fAfU*fG*fC*mC
MG29-1 sgRNA targeting human HAO1	5728	hH29-23_50b	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUUCAGCCUfGfUfCfAfGfUfCfC*fCfU*fG*fG*mG
MG29-1 sgRNA targeting human HAO1	5729	hH29-41_50b	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUAUAUCUUfCfCfCfAfGfCfUfG*fAfU*fA*fG*mA
MG29-1 sgRNA targeting mouse HAO1	5730	mH29-1-50	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUCCCCAGAfCfCfUfGfUfAfAfUfAfG*fU*fCfA*fU*mA
MG29-1 sgRNA targeting mouse HAO1	5731	mH29-15-50	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUUGACUGUfGfGfAfCfAfCfCfCfCfU*fU*fAfC*fC*mU
MG29-1 sgRNA targeting mouse HAO1	5732	mH29-1-50b	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUCCCCAGAfCfCfUfGfUfAfAfUfAfG*fU*fC*mA
MG29-1 sgRNA targeting mouse HAO1	5733	mH29-15-50b	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUUGACUGUfGfGfAfCfAfCfCfCfCfU*fU*fA*mC
SpCas9 sgRNA targeting mouse albumin	5734	mAlbR1 guide RNA	Nucleotide	N.A.	mU*mU*mA*GUAUAGCAUGGUCGAGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U
SpCas9 sgRNA targeting mouse albumin	5735	mAlbR2 guide RNA	Nucleotide	N.A.	mU*mU*mC*CUGUAACGAUCGGGAACGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U

Cat.	SEQ ID:	Description	Type	Organism	Sequence
SpCas9 sgRNA targeting mouse albumin	5736	mAlbR3 guide RNA	Nucleotide	N.A.	mU*mG*mC*CAGUUCCCGAUCGUUACGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U
PCR primer	5737	mAlb90F PCR primer	Nucleotide	N.A.	CTCCTCTTCGTCTCCGGC
PCR primer	5738	mAlb1073R PCR primer	Nucleotide	N.A.	CTGCCACATTGCTCAGCAC
PCR primer	5739	mAlb282F PCR primer	Nucleotide	N.A.	TTGCATCTGAGAACCCTTAGG
PCR primer	5740	mAlb460F PCR primer	Nucleotide	N.A.	GCCTGCTCGACCATGCTATA
SpCas9 sgRNA targeting mouse albumin	5741	mAlbR2 guide RNA with extensive chemical modifications	Nucleotide	N.A.	mU*mU*mC*CUGUAACGAUCGGGAACGUUUUAGAmGmCmUmAGAAAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCmAmAmCmUmUGAAAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*U
SpCas9 DNA sequence	5742	Template DNA for spCas9 in vitro transcription	Nucleotide	N.A.	GCGGCCGCTAATACGACTCACTATAAGAAAAGCCAGCTCCAGCAGGCGCTGCTCACTCCTCCCCATCCTCTCCCTCTGTCCCTCTGTCCCTCTGACCCTGCACTGTCCCAGCACCATGGCCCCCAAGAAGAAGCGGAAAGTTGGCGGCGGAGGCAGCGACAAGAAGTACTCTATCGGCCTGGACATCGGCACCAACTCTGTTGGATGGGCCGTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAATCTGATCGGCGCCCTGCTGTTCGACTCTGGCGAAACAGCCGAAGCCACCAGACTGAAGAGAACCGCCAGACGGCGGTACACCAGAAGAAAGAACCGGATCTGCTACCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGACAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGATGAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCACCGACAAGGCCGACCTGAGACTGA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					TCTATCTGGCCCTGGCT CACATGATCAAGTTCCG GGGCCACTTCCTGATCG AGGGCGACCTGAATCCT GACAACAGCGACGTGGA CAAGCTGTTCATCCAGC TGGTGCAGACCTACAAC CAGCTGTTCGAGGAAAA CCCCATCAACGCCAGCG GAGTGGATGCCAAGGCC ATCCTGTCTGCCAGACT GAGCAAGAGCAGACGGC TGGAAAACCTGATCGCT CAGCTGCCCGGCGAGAA GAAGAATGGCCTGTTCG GCAACCTGATTGCCCTG AGCCTGGGCCTGACACC TAACTTCAAGAGCAACTT CGACCTGGCCGAGGACG CCAAACTGCAGCTGTCC AAGGACACCTACGACGA CGACCTGGACAATCTGC TGGCCCAGATCGGCGAT CAGTACGCCGACTTGTT TCTGGCCGCCAAGAACC TGTCCGACGCCATCCTG CTGAGCGACATCCTGAG AGTGAACACCGAGATCA CAAAGGCCCCTCTGAGC GCCTCTATGATCAAGAG ATACGACGAGCACCACC AGGATCTGACCCTGCTG AAGGCCCTCGTTAGACA GCAGCTGCCTGAGAAGT ACAAAGAGATTTTCTTCG ACCAGAGCAAGAACGGC TACGCCGGCTACATTGA TGGCGGAGCCAGCCAAG AGGAATTCTACAAGTTC ATCAAGCCCATCCTCGA GAAGATGGACGGCACCG AGGAACTGCTGGTCAAG CTGAACAGAGAGGACCT GCTGCGGAAGCAGCGGA CCTTCGACAATGGCTCT ATCCCTCACCAGATCCA CCTGGGAGAGCTGCACG CCATTCTGCGGAGACAA GAGGACTTTTACCCATTC CTGAAGGACAACCGGGA AAAGATTGAGAAGATCC TGACCTTCAGGATCCCC TACTACGTGGGACCACT GGCCAGAGGCAATAGCA GATTCGCCTGGATGACC AGAAAGAGCGAGGAAAC CATCACACCCTGGAACT TCGAGGAAGTGGTGGAC AAGGGCGCCAGCGCTCA GTCCTTCATCGAGCGGA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					TGACCAACTTCGATAAG
					AACCTGCCTAACGAGAA
					GGTGCTGCCCAAGCACA
					GCCTGCTGTACGAGTAC
					TTCACCGTGTACAACGA
					GCTGACCAAAGTGAAAT
					ACGTGACCGAGGGAATG
					AGAAAGCCCGCCTTTCT
					GAGCGGCGAGCAGAAAA
					AGGCCATTGTGGATCTG
					CTGTTCAAGACCAACCG
					GAAAGTGACCGTGAAGC
					AGCTGAAAGAGGACTAC
					TTCAAGAAAATCGAGTG
					CTTCGACAGCGTGGAAA
					TCAGCGGCGTGGAAGAT
					CGGTTCAATGCCAGCCT
					GGGCACATACCACGACC
					TGCTGAAAATTATCAAG
					GACAAGGACTTCCTGGA
					CAACGAAGAGAACGAGG
					ACATCCTGGAAGATATC
					GTGCTGACCCTGACACT
					GTTTGAGGACAGAGAGA
					TGATCGAGGAACGGCTG
					AAAACATACGCCCACCT
					GTTCGACGACAAAGTGA
					TGAAGCAACTGAAGCGG
					CGGAGATACACCGGCTG
					GGGCAGACTGTCTCGGA
					AGCTGATCAACGGCATC
					CGGGATAAGCAGTCCGG
					CAAGACCATCCTGGACT
					TTCTGAAGTCCGACGGC
					TTCGCCAACAGAAACTT
					CATGCAGCTGATCCACG
					ACGACAGCCTGACCTTT
					AAAGAGGATATCCAGAA
					AGCCCAGGTGTCCGGCC
					AGGGCGATTCTCTGCAT
					GAGCACATTGCCAACCT
					GGCCGGCTCTCCCGCCA
					TTAAGAAGGGCATTCTG
					CAGACAGTGAAGGTGGT
					GGACGAGCTGGTCAAAG
					TCATGGGCAGACACAAG
					CCCGAGAACATCGTGAT
					CGAAATGGCCAGAGAGA
					ACCAGACCACACAGAAG
					GGCCAGAAGAACAGCCG
					CGAGAGAATGAAGCGGA
					TCGAAGAGGGCATCAAA
					GAGCTGGGCAGCCAGAT
					CCTGAAAGAACACCCCG
					TGGAAAACACCCAGCTG
					CAGAACGAGAAGCTGTA
					CCTGTACTACCTCCAGA
					ACGGCCGGGATATGTAC
					GTGGACCAAGAGCTGGA
					CATCAACCGGCTGTCCG

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					ACTACGATGTGGACCAT
					ATCGTGCCCCAGTCTTTT
					CTGAAGGACGACTCCAT
					CGACAACAAGGTCCTGA
					CCAGATCCGACAAGAAT
					CGGGGCAAGAGCGACAA
					CGTGCCCTCCGAAGAGG
					TGGTCAAGAAGATGAAG
					AACTACTGGCGACAGCT
					GCTGAACGCCAAGCTGA
					TTACCCAGCGGAAGTTC
					GATAACCTGACCAAGGC
					CGAGAGAGGCGGCCTGT
					CTGAACTGGATAAGGCC
					GGCTTCATCAAGAGACA
					GCTGGTGGAAACCCGGC
					AGATCACCAAACACGTG
					GCACAGATTCTGGACTC
					CCGGATGAACACTAAGT
					ACGACGAGAATGACAAG
					CTGATCCGGGAAGTGAA
					AGTGATCACCCTGAAGT
					CCAAGCTGGTGTCCGAT
					TTCCGGAAGGATTTCCA
					GTTCTACAAAGTGCGCG
					AGATCAACAACTACCAT
					CACGCCCACGACGCCTA
					CCTGAATGCCGTTGTTG
					GAACAGCCCTGATCAAG
					AAGTATCCCAAGCTGGA
					AAGCGAGTTCGTGTACG
					GCGACTACAAGGTGTAC
					GACGTGCGGAAGATGAT
					CGCCAAGAGCGAGCAAG
					AGATTGGCAAGGCTACC
					GCCAAGTACTTTTTCTAC
					AGCAACATCATGAACTTT
					TTCAAGACCGAGATTAC
					CCTGGCCAACGGCGAGA
					TCAGAAAGCGGCCTCTG
					ATCGAGACAAACGGCGA
					AACCGGCGAGATTGTGT
					GGGATAAGGGCAGAGAC
					TTTGCCACAGTGCGGAA
					AGTGCTGAGCATGCCCC
					AAGTGAATATCGTGAAG
					AAAACCGAGGTGCAGAC
					AGGCGGCTTCAGCAAAG
					AGTCCATTCTGCCCAAG
					AGAAACAGCGATAAGCT
					GATCGCCCGGAAGAAGG
					ACTGGGACCCTAAGAAG
					TACGGCGGCTTCGATAG
					CCCTACCGTGGCCTATT
					CTGTGCTGGTGGTGGCC
					AAAGTGGAAAAGGGCAA
					GTCCAAGAAACTCAAGA
					GCGTGAAAGAGCTGCTG
					GGGATCACCATCATGGA
					AAGAAGCAGCTTCGAGA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					AGAATCCTATCGATTTCC TCGAGGCCAAGGGCTAC AAAGAAGTGAAAAAGGA CCTGATCATCAAGCTCC CCAAGTACTCCCTGTTC GAGCTGGAAAATGGCCG GAAGCGGATGCTGGCTT CTGCTGGCGAACTGCAG AAGGGAAACGAACTGGC CCTGCCTAGCAAATATG TGAACTTCCTGTACCTG GCCAGCCACTATGAGAA GCTGAAGGGCAGCCCCG AGGACAATGAGCAAAAG CAGCTGTTTGTGGAACA GCACAAGCACTACCTGG ACGAGATCATCGAGCAG ATCAGCGAGTTCTCCAA GAGAGTGATCCTGGCCG ACGCTAATCTGGACAAA GTGCTGTCCGCCTACAA CAAGCACCGGGACAAGC CTATCAGAGAGCAGGCC GAGAATATCATCCACCT GTTTACCCTGACCAATCT GGGAGCCCCTGCCGCCT TCAAGTACTTCGACACC ACCATCGACCGGAAGCG CTACACCAGCACCAAAG AGGTGCTGGACGCCACA CTGATCCACCAGTCTATC ACCGGCCTGTACGAGAC ACGGATCGACCTGTCTC AGCTCGGAGGCGATTCT GGCGGAAAAAGACCTGC CGCCACAAAGAAAGCCG GACAGGCCAAGAAAAAG AAGTGACCACACCCCCA TTCCCCCACTCCAGATAA AGCTTCAGTTATATCTCA CGTGTCTGGAGTTAAAA AAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAA AAAAAAAAAAAGAAGAG CCCTGCAGG
SpCas9 protein sequence	5743	Amino acid sequence of spCas9 encoded in mRNA inclusing nuclear localization signals	Protein	N.A.	MAPKKKRKVGGGGSDKK YSIGLDIGTNSVGWAVITD EYKVPSKKFKVLGNTDRH SIKKNLIGALLFDSGETAE ATRLKRTARRRYTRRKN RICYLQEIFSNEMAKVDDS FFHRLEESFLVEEDKKHE RHPIFGNIVDEVAYHEKYP TIYHLRKKLVDSTDKADL RLIYLALAHMIKFRGHFLI EGDLNPDNSDVDKLFIQL VQTYNQLFEENPINASGV

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					DAKAILSARLSKSRRLENL IAQLPGEKKNGLFGNLIA LSLGLTPNFKSNFDLAEDA KLQLSKDTYDDDLDNLLA QIGDQYADLFLAAKNLSD AILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLK ALVRQQLPEKYKEIFFDQ SKNGYAGYIDGGASQEEF YKFIKPILEKMDGTEELL VKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQE DFYPFLKDNREKIEKILTF RIPYYVGPLARGNSRFAW MTRKSEETITPWNFEEVV DKGASAQSFIERMTNFDK NLPNEKVLPKHSLLYEYF TVYNELTKVKYVTEGMR KPAFLSGEQKKAIVDLLF KTNRKVTVKQLKEDYFK KIECFDSVEISGVEDRFNA SLGTYHDLLKIIKDKDFLD NEENEDILEDIVLTLTLFE DREMIEERLKTYAHLFDD KVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTIL DFLKSDGFANRNFMQLIH DDSLTFKEDIQKAQVSGQ GDSLHEHIANLAGSPAIKK GILQTVKVVDELVKVMG RHKPENIVIEMARENQTT QKGQKNSRERMKRIEEGI KELGSQILKEHPVENTQL QNEKLYLYYLQNGRDMY VDQELDINRLSDYDVDHIV PQSFLKDDSIDNKVLTRSD KNRGKSDNVPSEEVVKK MKNYWRQLLNAKLITQR KFDNLTKAERGGLSELDK AGFIKRQLVETRQITKHV AQILDSRMNTKYDENDKL IREVKVITLKSKLVSDFRK DFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKL ESEFVYGDYKVYDVRKMI AKSEQEIGKATAKYFFYS NIMNFFKTEITLANGEIRK RPLIETNGETGEIVWDKG RDFATVRKVLSMPQVNIV KKTEVQTGGFSKESILPK RNSDKLIARKKDWDPKK YGGFDSPTVAYSVLVVAK VEKGKSKKLKSVKELLGI TIMERSSFEKNPIDFLEAK GYKEVKKDLIIKLPKYSLF ELENGRKRMLASAGELQ KGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQL FVEQHKHYLDEIIEQISEF SKRVILADANLDKVLSAY

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					NKHRDKPIREQAENIIHLF TLTNLGAPAAFKYFDTTID RKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGDS GGKRPAATKKAGQAKKK K
MG29-1 sgRNA targeting mouse albumin	5744	mAlb29-8-50 guide RNA	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*m G*mA*mA*mAGAUUCUCA AC*mC*mU*mU*U*UAAU UmUmCmUmACU*G*U*U* GUAGAUCUGUAACfGfAfU fCfGfGfGfAfAfC*fU*fGfG*f C*mA
MG29-1 sgRNA targeting human HAO-1	5788	hH29-1	Nucleotide	N.A.	UAAUUUCUACUGUUGUA GAUGCAUGUUGUUCAUA AUCAUUGA
MG29-1 sgRNA targeting human HAO-1	5789	hH29-2	Nucleotide	N.A.	UAAUUUCUACUGUUGUA GAUGAAGUACUGAUUUA GCAUGUUG
MG29-1 sgRNA targeting human HAO-1	5790	hH29-3	Nucleotide	N.A.	UAAUUUCUACUGUUGUA GAUUAUCAAUGAUUAUG AACAACAU
MG29-1 sgRNA targeting human HAO-1	5791	hH29-4	Nucleotide	N.A.	UAAUUUCUACUGUUGUA GAUCCCCAGACCUGUAA UAGUCAUA
MG29-1 sgRNA targeting human HAO-1	5792	hH29-5	Nucleotide	N.A.	UAAUUUCUACUGUUGUA GAUUUCAUCAUUUGCCC CAGACCUG
MG29-1 sgRNA targeting human HAO-1	5793	hH29-6	Nucleotide	N.A.	UAAUUUCUACUGUUGUA GAUUUACCUGGAAAAUG CUGCAAUA
MG29-1 sgRNA targeting human HAO-1	5794	hH29-7	Nucleotide	N.A.	UAAUUUCUACUGUUGUA GAUCUUACCUGGAAAAU GCUGCAAU
MG29-1 sgRNA targeting human HAO-1	5795	hH29-8	Nucleotide	N.A.	UAAUUUCUACUGUUGUA GAUGCUGAUAAUAUUGC AGCAUUUU
MG29-1 sgRNA targeting human HAO-1	5796	hH29-9	Nucleotide	N.A.	UAAUUUCUACUGUUGUA GAUAAAAAUAAAUUUUC UUACCUGG

Cat.	SEQ ID:	Description	Type	Organism	Sequence
MG29-1 sgRNA targeting human HAO-1	5797	hH29-10	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUAAAAAAUAAAUUUUCUUACCUG
MG29-1 sgRNA targeting human HAO-1	5798	hH29-11	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUAUUUUAUUUUUUAAUUCUAGAU
MG29-1 sgRNA targeting human HAO-1	5799	hH29-12	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUUUUUAUUUUUUAAUUCUAGAUG
MG29-1 sgRNA targeting human HAO-1	5800	hH29-13	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUAUUUUUUAAUUCUAGAUGGAAG
MG29-1 sgRNA targeting human HAO-1	5801	hH29-14	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUUUUUUUAAUUCUAGAUGGAAGC
MG29-1 sgRNA targeting human HAO-1	5802	hH29-15	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUUUAAUUCUAGAUGGAAGCUGUA
MG29-1 sgRNA targeting human HAO-1	5803	hH29-16	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUUAAUUCUAGAUGGAAGCUGUAU
MG29-1 sgRNA targeting human HAO-1	5804	hH29-17	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUAAUUCUAGAUGGAAGCUGUAUC
MG29-1 sgRNA targeting human HAO-1	5805	hH29-18	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUAUUCUAGAUGGAAGCUGUAUCC
MG29-1 sgRNA targeting human HAO-1	5806	hH29-19	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUAGCAACAUUCCGGAGCAUCCUU
MG29-1 sgRNA targeting human HAO-1	5807	hH29-20	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUAGGACAGAGGGUCAGCAUGCCA
MG29-1 sgRNA targeting human	5808	hH29-21	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUGGACAGAGGGUCAGCAUGCCAA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
HAO-1					
MG29-1 sgRNA targeting human HAO-1	5809	hH29-22	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUUUUCUCAGCCUGUCAGUCCCUG
MG29-1 sgRNA targeting human HAO-1	5810	hH29-23	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUUCAGCCUGUCAGUCCCUGGGAA
MG29-1 sgRNA targeting human HAO-1	5811	hH29-24	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUUGACAGUGGACACACCUUACCU
MG29-1 sgRNA targeting human HAO-1	5812	hH29-25	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUAAUCUGUUACGCACAUCAUCCA
MG29-1 sgRNA targeting human HAO-1	5813	hH29-26	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUAUGCAUUUCUUAUUUUAGGAUG
MG29-1 sgRNA targeting human HAO-1	5814	hH29-27	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUUGCAUUUCUUAUUUUAGGAUGA
MG29-1 sgRNA targeting human HAO-1	5815	hH29-28	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUUUAUUUUAGGAUGAAAAAUUUU
MG29-1 sgRNA targeting human HAO-1	5816	hH29-29	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUAGGAUGAAAAAUUUUGAAACCA
MG29-1 sgRNA targeting human HAO-1	5817	hH29-30	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUGGAUGAAAAAUUUUGAAACCAG
MG29-1 sgRNA targeting human HAO-1	5818	hH29-31	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUCUCAGGAGAAAAUGAUAAAGUA
MG29-1 sgRNA targeting human HAO-1	5819	hH29-32	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUCCUCAGGAGAAAAUGAUAAAGU
MG29-1 sgRNA targeting	5820	hH29-33	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUGAAACCAGUACUUUAUCAUUUU

Cat.	SEQ ID:	Description	Type	Organism	Sequence
human HAO-1					
MG29-1 sgRNA targeting human HAO-1	5821	hH29-34	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUAAACCAGUACUUUAUCAUUUUC
MG29-1 sgRNA targeting human HAO-1	5822	hH29-35	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUUCAUUUUCUCCUGAGGAAAAUU
MG29-1 sgRNA targeting human HAO-1	5823	hH29-36	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUCUCCUGAGGAAAAUUUUGGAGA
MG29-1 sgRNA targeting human HAO-1	5824	hH29-37	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUUCCUGAGGAAAAUUUUGGAGAC
MG29-1 sgRNA targeting human HAO-1	5825	hH29-38	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUGCCACAUAUGCAGCAAGUCCAC
MG29-1 sgRNA targeting human HAO-1	5826	hH29-39	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUGGAGACGACAGUGGACUUGCUG
MG29-1 sgRNA targeting human HAO-1	5827	hH29-40	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUGAGACGACAGUGGACUUGCUGC
MG29-1 sgRNA targeting human HAO-1	5828	hH29-41	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUAUCUUCCCAGCUGAUAGAUG
MG29-1 sgRNA targeting human HAO-1	5829	hH29-42	Nucleotide	N.A.	UAAUUUCUACUGUUGUAGAUCAACAAUUGGCAAUGAUGUCAG
DNA sequence encoding the MG29-1 messenger RNA	5830	DNA sequence encoding the MG29-1 messenger RNA	Nucleotide	N.A.	AAAAGCCAGCTCCAGCAGGCGCTGCTCACTCCTCCCCATCCTCTCCCTCTGTCCCTCTGTCCCTCTGACCCTGCACTGTCCCAGCACCATGGCCCCTAAGAAGAAGAGAAAAGTCGGCGGAGGCGGCAGCTTCAACAACTTCATCAAGAAGTACAGCCTGCAGAAAACCCTGCGCTTCGAGCTGAAGCCTGTGGGCGAGACAGCCG

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					ACTACATCGAGGACTTC AAGAGCGAGTACCTGAA GGACACCGTGCTGAAGG ACGAGCAGAGAGCCAAG GACTACCAAGAGATCAA GACCCTGATCGACGATT ACCACCGCGAGTACATC GAAGAGTGCCTGAGAGA ACCCGTGGACAAGAAAA CCGGCGAGATCCTGGAC TTCACCCAGGACCTGGA AGATGCCTTCAGCTACT ACCAGAAGCTGAAAGAG AACCCCACCGAGAACAG AGTCGGCTGGGAGAAAG AGCAAGAGAGCCTGAGG AAGAAGCTGGTCACCTC CTTCGTGGGCAACGACG GCCTGTTCAAGAAAGAG TTCATCACCAGGGACCT GCCTGAGTGGCTGCAGA AGAAAGGACTCTGGGGC GAGTACAAGGACACAGT GGAAAACTTCAAGAAGT TCACCACCTACTTCAGC GGCTTCCACGAGAACCG GAAGAACATGTACACCG CCGAGGCTCAGAGCACC GCTATCGCCAACAGACT GATGAACGACAACCTGC CTAAGTTCTTTAACAACT ACCTGGCCTACCAGACC ATCAAAGAGAAGCACCC CGACCTGGTGTTCAGAC TGGATGATGCTCTGCTG CAGGCCGCTGGCGTGGA ACATCTGGATGAGGCTT TCCAGCCTAGATACTTCA GCAGACTGTTCGCCCAG AGCGGCATCACCGCTTT CAACGAGCTGATCGGCG GCAGAACCACAGAGAAC GGCGAGAAGATCCAGGG CCTGAACGAGCAGATCA ACCTGTACAGACAGCAG AACCCCGAGAAGGCCAA GGGCTTCCCCAGATTCA TGCCTCTGTTCAAGCAG ATCCTGAGCGACAGAGA GACACACAGCTTTCTGC CCGACGCCTTCGAGAAC GACAAAGAGCTGCTCCA GGCTCTGAGAGACTACG TGGACGCCGCCACATCT GAGGAAGGCATGATCAG CCAGCTGAACAAGGCCA TGAACCAGTTCGTGACC GCCGACCTGAAGAGAGT GTACATCAAGAGCGCCG CTCTGACCAGCCTGAGC

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					CAAGAGCTGTTCCACTT CTTCGGCGTGATCAGCG ACGCTATCGCTTGGTAC GCCGAGAAGAGACTGAG CCCCAAGAAGGCCCAAG AGTCTTTCCTGAAGCAA GAGGTGTACGCCATCGA GGAACTGAACCAGGCTG TCGTGGGCTACATCGAC CAGCTGGAAGATCAGAG CGAGCTGCAGCAACTGC TGGTGGACCTGCCAGAT CCTCAGAAACCCGTGTC CAGCTTCATCCTGACAC ACTGGCAGAAGTCTCAA GAGCCCCTGCAGGCAGT GATCGCCAAGGTGGAAC CTCTGTTCGAACTGGAA GAACTGAGCAAGAACAA GAGGGCCCCAAAGCACG ACAAGGACCAAGGCGGC GAGGGATTTCAGCAGGT CGACGCCATCAAGAACA TGCTGGACGCCTTCATG GAAGTGTCCCACGCTAT CAAGCCCCTGTACCTGG TCAAGGGAAGAAAGGCC ATCGACATGCCCGACGT GGACACCGGCTTCTACG CTGATTTCGCCGAGGCC TACAGCGCCTACGAGCA AGTGACAGTGTCCCTGT ACAACAAGACCAGAAAC CACCTGTCCAAGAAGCC CTTCAGCAAGGACAAGA TCAAGATCAACTTCGAC GCCCCTACACTGCTGAA CGGCTGGGACCTGAACA AAGAGAGCGACAACAAG TCCATCATCCTGCGGAA GGACGGCAACTTCTACC TGGCAATCATGCACCCC AAGCACACCAAGGTGTT CGACTGCTACTCTGCCT CTGAGGCTGCCGGCAAG TGCTACGAGAAGATGAA CTACAAGCTGCTGAGCG GCGCCAACAAGATGCTG CCTAAGGTGTTCTTTAGC AAGAAGGGCATCGAGAC ATTCAGCCCTCCACAAG AAATCCTGGACCTGTAC AAGAACAACGAGCATAA GAAGGGCGCCACCTTCA AGCTGGAATCCTGCCAC AAGCTGATCGATTTCTTC AAGCGGAACATCCCCAA GTACAAGGTGCACCCTA CCGACAACTTTGGCTGG GACGTGTTCGGCTTTCA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					CTTCAGCCCTACCAGCA GCTACGGCGACCTGTCT GGCTTCTACAGAGAGGT GGAAGCCCAGGGATACA AGCTGTGGTTCAGCGAC GTGTCCGAGGCTTACAT CAACAAATGCGTGGAAG AGGGCAAGCTGTTCCTG TTCCAAATCTACAACAAG GACTTCTCCCCTAACTCC ACCGGCAAGCCCAACCT GCACACCCTGTATTGGA AGGGCCTGTTCGAGCCC GAGAACCTGAAAGACGT GGTGCTGAAGCTGAATG GCGAGGCCGAGATCTTC TACCGGAAGCACAGCAT CAAGCACGAGGACAAGA CCATCCACAGAGCTAAG GACCCTATCGCTAACAA GAACGCTGACAACCCCA AGAAACAGAGCGTGTTC GATTACGACATCATCAA GGATAAGCGGTATACCC AGGACAAGTTCTTCTTCC ACGTGCCAATCAGCCTG AACTTCAAAAGCCAGGG CGTCGTGCGGTTCAACG ATAAGATCAACGGCCTG CTGGCCGCTCAGGACGA TGTGCATGTGATCGGCA TCGACAGAGGCGAGAGA CATCTGCTGTACTACACC GTGGTCAACGGCAAGGG CGAAGTGGTGGAACAGG GCAGCCTGAATCAGGTG GCCACAGATCAGGGCTA CGTGGTGGATTACCAGC AGAAGCTGCACGCCAAA GAGAAAGAACGCGACCA GGCCAGAAAGAACTGGT CCACCATCGAGAACATC AAAGAACTGAAGGCCGG CTACCTGAGCCAGGTGG TGCATAAGCTGGCTCAG CTGATCGTGAAGCACAA CGCCATCGTGTGCCTCG AGGACCTGAATTTCGGC TTCAAGAGGGGCAGATT CAAGGTCGAGAAACAGG TGTACCAGAAGTTCGAG AAGGCTCTGATCGACAA GCTGAACTACCTCGTGT TCAAAGAGAGAGGCGCC ACACAGGCTGGCGGATA CCTGAATGCTTACCAGC TGGCCGCACCTTTCGAG AGCTTTGAGAAGCTGGG CAAGCAGACCGGCATCC TGTACTACGTGCGGAGC

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					GACTACACCAGCAAGAT CGACCCTGCTACCGGCT TCGTGGACTTTCTGAAG CCTAAGTACGAGAGCAT GGCCAAGAGCAAAGTGT TCTTCGAGTCCTTCGAG CGCATCCAGTGGAACCA GGCCAAAGGCTACTTCG AGTTCGAGTTTGACTAC AAGAAGATGTGCCCCAG CAGAAAGTTCGGCGACT ACAGAACCAGATGGGTC GTGTGCACCTTCGGCGA CACCCGCTACCAGAACA GAAGAAACAAGAGCAGC GGCCAGTGGGAGACAGA GACAATCGATGTGACAG CCCAGCTGAAAGCCCTG TTCGCCGCTTACGGCAT CACATACAATCAAGAGG ATAACATCAAGGACGCC ATTGCCGCCGTGAAGTA CACCAAGTTCTACAAGC AGCTGTACTGGCTGCTG AGACTGACCCTGAGCCT GAGACACAGCGTGACAG GCACCGACGAGGATTTC ATCCTGTCTCCAGTGGC CGACGAGAATGGCGTGT TCTTTGACTCTAGGAAG GCCACCGACAAGCAGCC TAAGGACGCTGATGCTA ACGGCGCCTACCATATC GCCCTGAAAGGCCTGTG GAATCTCCAGCAGATCA GACAGCACGACTGGAAC GTGGAAAAGCCCAAAAA GCTGAACCTCGCCATGA AGAACGAAGAGTGGTTC GGCTTCGCTCAGAAGAA GAAGTTTAGAGCCAGCG GCGGCAAGAGGCCTGCC GCTACAAAAAAAGCCGG CCAGGCCAAGAAAAAGA AGTGACCACACCCCCAT TCCCCCACTCCAGATAG AACTTCAGTTATATCTCA CGTGTCTGGAGTT
MG29-1 sgRNA targeting human HAO-1	5831	hH29-4_37	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmU mCmUmACU*G*U*U*GUA GAUCCCCAGAfCfCfUfGfU fAfAfUfAfG*fU*fCfA*fU*m A
MG29-1 sgRNA targeting human HAO-1	5832	hH29-21-37	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmU mCmUmACU*G*U*U*GUA GAUGGACAGAfGfGfGfUf CfAfGfCfAfU*fG*fCfC*fA* mA
MG29-1 sgRNA	5833	hH29-23-37	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmU mCmUmACU*G*U*U*GUA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
targeting human HAO-1					GAUUCAGCCUfGfUfCfAfGfUfCfCfCfU*fG*fGfG*fA*mA
MG29-1 sgRNA targeting human HAO-1	5834	hH29-41-37	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUAUAUCUUfCfCfCfAfGfCfUfGfAfU*fA*fGfA*fU*mG
Guide modification chemistry	5835	Chemistry 37 (22 nt spacer)	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUNNNNNNNfNfNfNfNfNfNfNfNfN*fN*fNfN*fN*mN
MG29-1 sgRNA targeting mouse HAO-1	5836	mH29-29-37	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUCCUUAGGfAfGfAfAfAfAfUfGfCfC*fA*fAfA*fU*mC
MG29-1 sgRNA targeting mouse HAO-1	5837	mH29-29-44	Nucleotide	N.A.	mC*mU*mU*U*U*AAUUmUmCmUmACU*G*U*U*GUAGAUCCUUAGGfAfGfAfAfAfAfUfG*fCfC*fA*fA*fA*fU*mC
MG29-1 sgRNA targeting mouse HAO-1	5838	mH29-29s-37	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUCCUUAGGfAfGfAfAfAfAfUfG*fCfC*fA*fA*mA
MG29-1 sgRNA targeting mouse HAO-1	5839	mH29-29s-44	Nucleotide	N.A.	mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUCCUUAGGfAfGfAfAfAfAfUfG*fC*fC*fA*fA*mA
MG29-1 sgRNA targeting mouse HAO-1	5840	mH29-29-50	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUCCUUAGGfAfGfAfAfAfAfUfGfCfC*fA*fAfA*fU*mC
MG29-1 sgRNA targeting mouse HAO-1	5841	mH29-29-50b	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUCCUUAGGfAfGfAfAfAfAfUfGfCfC*fA*fA*mA
MG29-1 sgRNA targeting mouse HAO-1	5842	mH29-29-51b	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUCCUUAGGAGAAAAUGCC*A*mA*mA
MG29-1 sgRNA targeting mouse HAO-1	5843	mH29-29-52b	Nucleotide	N.A.	mG*mU*mU*GAGAAUC*mG*mA*mA*mAGAUUCUCAAC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUCCUUAGGAfGAfAAfAUfG*C*fCA*fA*mA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
MG29-1 sgRNA targeting mouse HAO-1	5844	mH29-29-53b	Nucleotide	N.A.	mG*mU*mU*mG*mA*mG*mA*mA*mU*mC*mG*mA*mA*mA*mG*mA*mU*mU*mC*mU*mC*mA*mA*mC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUCCUUAGGfAfGfAfAfAfAfUfGfCfC*fA*fA*mA
MG29-1 sgRNA targeting mouse HAO-1	5845	mH29-29-54b	Nucleotide	N.A.	mG*mU*mU*mG*mA*mG*mA*mA*mU*mC*mG*mA*mA*mA*mG*mA*mU*mU*mC*mU*mC*mA*mA*mC*mC*mU*mU*U*UAAUUmUmCmUmACU*G*U*U*GUAGAUCCUUAGGAGAAAAUGCC*A*mA*mA
DNA sequence encoding the MG29-1 messenger RNA	5846	DNA sequence encoding the MG29-1 messenger RNA	Nucleotide	N.A.	AAAAGCCAGCTCCAGCAGGCGCTGCTCACTCCTCCCCATCCTCTCCCTCTGTCCCTCTGTCCCTCTGACCCTGCACTGTCCCAGCACCATGGCCCCTAAGAAGAAGAGAAAAGTCGGCGGAGGCGGCAGCTTCAACAACTTCATCAAGAAGTACAGCCTGCAGAAAACCCTGCGCTTCGAGCTGAAGCCTGTGGGCGAGACAGCCGACTACATCGAGGACTTCAAGAGCGAGTACCTGAAGGACACCGTGCTGAAGGACGAGCAGAGAGCCAAGGACTACCAAGAGATCAAGACCCTGATCGACGATTACCACCGCGAGTACATCGAAGAGTGCCTGAGAGAACCCGTGGACAAGAAAACCGGCGAGATCCTGGACTTCACCCAGGACCTGGAAGATGCCTTCAGCTACTACCAGAAGCTGAAAGAGAACCCCACCGAGAACAGAGTCGGCTGGGAGAAAGAGCAAGAGAGCCTGAGGAAGAAGCTGGTCACCTCCTTCGTGGGCAACGACGGCCTGTTCAAGAAAGAGTTCATCACCAGGGACCTGCCTGAGTGGCTGCAGAAGAAAGGACTCTGGGGCGAGTACAAGGACACAGTGGAAAACTTCAAGAAGTTCACCACCTACTTCAGCGGCTTCCACGAGAACCGGAAGAACATGTACACCGCCGAGGCTCAGAGCACCGCTATCGCCAACAGACTGATGAACGACAACCTGCCTAAGTTCTTTAACAACTACCTGGCCTACCAGACC

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					ATCAAAGAGAAGCACCC CGACCTGGTGTTCAGAC TGGATGATGCTCTGCTG CAGGCCGCTGGCGTGGA ACATCTGGATGAGGCTT TCCAGCCTAGATACTTCA GCAGACTGTTCGCCCAG AGCGGCATCACCGCTTT CAACGAGCTGATCGGCG GCAGAACCACAGAGAAC GGCGAGAAGATCCAGGG CCTGAACGAGCAGATCA ACCTGTACAGACAGCAG AACCCCGAGAAGGCCAA GGGCTTCCCCAGATTCA TGCCTCTGTTCAAGCAG ATCCTGAGCGACAGAGA GACACACAGCTTTCTGC CCGACGCCTTCGAGAAC GACAAAGAGCTGCTCCA GGCTCTGAGAGACTACG TGGACGCCGCCACATCT GAGGAAGGCATGATCAG CCAGCTGAACAAGGCCA TGAACCAGTTCGTGACC GCCGACCTGAAGAGAGT GTACATCAAGAGCGCCG CTCTGACCAGCCTGAGC CAAGAGCTGTTCCACTT CTTCGGCGTGATCAGCG ACGCTATCGCTTGGTAC GCCGAGAAGAGACTGAG CCCCAAGAAGGCCCAAG AGTCTTTCCTGAAGCAA GAGGTGTACGCCATCGA GGAACTGAACCAGGCTG TCGTGGGCTACATCGAC CAGCTGGAAGATCAGAG CGAGCTGCAGCAACTGC TGGTGGACCTGCCAGAT CCTCAGAAACCCGTGTC CAGCTTCATCCTGACAC ACTGGCAGAAGTCTCAA GAGCCCCTGCAGGCAGT GATCGCCAAGGTGGAAC CTCTGTTCGAACTGGAA GAACTGAGCAAGAACAA GAGGGCCCCAAAGCACG ACAAGGACCAAGGCGGC GAGGGATTTCAGCAGGT CGACGCCATCAAGAACA TGCTGGACGCCTTCATG GAAGTGTCCCACGCTAT CAAGCCCCTGTACCTGG TCAAGGGAAGAAAGGCC ATCGACATGCCCGACGT GGACACCGGCTTCTACG CTGATTTCGCCGAGGCC TACAGCGCCTACGAGCA AGTGACAGTGTCCCTGT

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					ACAACAAGACCAGAAAC
					CACCTGTCCAAGAAGCC
					CTTCAGCAAGGACAAGA
					TCAAGATCAACTTCGAC
					GCCCCTACACTGCTGAA
					CGGCTGGGACCTGAACA
					AAGAGAGCGACAACAAG
					TCCATCATCCTGCGGAA
					GGACGGCAACTTCTACC
					TGGCAATCATGCACCCC
					AAGCACACCAAGGTGTT
					CGACTGCTACTCTGCCT
					CTGAGGCTGCCGGCAAG
					TGCTACGAGAAGATGAA
					CTACAAGCTGCTGAGCG
					GCGCCAACAAGATGCTG
					CCTAAGGTGTTCTTTAGC
					AAGAAGGGCATCGAGAC
					ATTCAGCCCTCCACAAG
					AAATCCTGGACCTGTAC
					AAGAACAACGAGCATAA
					GAAGGGCGCCACCTTCA
					AGCTGGAATCCTGCCAC
					AAGCTGATCGATTTCTTC
					AAGCGGAACATCCCCAA
					GTACAAGGTGCACCCTA
					CCGACAACTTTGGCTGG
					GACGTGTTCGGCTTTCA
					CTTCAGCCCTACCAGCA
					GCTACGGCGACCTGTCT
					GGCTTCTACAGAGAGGT
					GGAAGCCCAGGGATACA
					AGCTGTGGTTCAGCGAC
					GTGTCCGAGGCTTACAT
					CAACAAATGCGTGGAAG
					AGGGCAAGCTGTTCCTG
					TTCCAAATCTACAACAAG
					GACTTCTCCCCTAACTCC
					ACCGGCAAGCCCAACCT
					GCACACCCTGTATTGGA
					AGGGCCTGTTCGAGCCC
					GAGAACCTGAAAGACGT
					GGTGCTGAAGCTGAATG
					GCGAGGCCGAGATCTTC
					TACCGGAAGCACAGCAT
					CAAGCACGAGGACAAGA
					CCATCCACAGAGCTAAG
					GACCCTATCGCTAACAA
					GAACGCTGACAACCCCA
					AGAAACAGAGCGTGTTC
					GATTACGACATCATCAA
					GGATAAGCGGTATACCC
					AGGACAAGTTCTTCTTCC
					ACGTGCCAATCAGCCTG
					AACTTCAAAAGCCAGGG
					CGTCGTGCGGTTCAACG
					ATAAGATCAACGGCCTG
					CTGGCCGCTCAGGACGA
					TGTGCATGTGATCGGCA
					TCGACAGAGGCGAGAGA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					CATCTGCTGTACTACACC GTGGTCAACGGCAAGGG CGAAGTGGTGGAACAGG GCAGCCTGAATCAGGTG GCCACAGATCAGGGCTA CGTGGTGGATTACCAGC AGAAGCTGCACGCCAAA GAGAAAGAACGCGACCA GGCCAGAAAGAACTGGT CCACCATCGAGAACATC AAAGAACTGAAGGCCGG CTACCTGAGCCAGGTGG TGCATAAGCTGGCTCAG CTGATCGTGAAGCACAA CGCCATCGTGTGCCTCG AGGACCTGAATTTCGGC TTCAAGAGGGGCAGATT CAAGGTCGAGAAACAGG TGTACCAGAAGTTCGAG AAGGCTCTGATCGACAA GCTGAACTACCTCGTGT TCAAAGAGAGAGGCGCC ACACAGGCTGGCGGATA CCTGAATGCTTACCAGC TGGCCGCACCTTTCGAG AGCTTTGAGAAGCTGGG CAAGCAGACCGGCATCC TGTACTACGTGCGGAGC GACTACACCAGCAAGAT CGACCCTGCTACCGGCT TCGTGGACTTTCTGAAG CCTAAGTACGAGAGCAT GGCCAAGAGCAAAGTGT TCTTCGAGTCCTTCGAG CGCATCCAGTGGAACCA GGCCAAAGGCTACTTCG AGTTCGAGTTTGACTAC AAGAAGATGTGCCCCAG CAGAAAGTTCGGCGACT ACAGAACCAGATGGGTC GTGTGCACCTTCGGCGA CACCCGCTACCAGAACA GAAGAAACAAGAGCAGC GGCCAGTGGGAGACAGA GACAATCGATGTGACAG CCCAGCTGAAAGCCCTG TTCGCCGCTTACGGCAT CACATACAATCAAGAGG ATAACATCAAGGACGCC ATTGCCGCCGTGAAGTA CACCAAGTTCTACAAGC AGCTGTACTGGCTGCTG AGACTGACCCTGAGCCT GAGACACAGCGTGACAG GCACCGACGAGGATTTC ATCCTGTCTCCAGTGGC CGACGAGAATGGCGTGT TCTTTGACTCTAGGAAG GCCACCGACAAGCAGCC TAAGGACGCTGATGCTA

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					ACGGCGCCTACCATATC GCCCTGAAAGGCCTGTG GAATCTCCAGCAGATCA GACAGCACGACTGGAAC GTGGAAAAGCCCAAAAA GCTGAACCTCGCCATGA AGAACGAAGAGTGGTTC GGCTTCGCTCAGAAGAA GAAGTTTAGAGCCAGCG GCGGCAAGAGGCCTGCC GCTACAAAAAAAGCCGG CCAGGCCAAGAAAAAGA AGTGACCACACCCCCAT TCCCCCACTCCAGATAG AACTTCAGTTATATCTCA CGTGTCTGGAGTT
MG55 active effector sgRNA	6031	MG55 sgRNA	Nucleotide	Unknown	AAATATTTCATTAAGTAC CGAATTTAAAAAATAGG ATTGCAGAAAGTGCAAT TAGGCTGGTTGTGCAGC CTTAATCTGAGGGATTA ATCCACTCGGAAAGTAC CTTTATTGAAAAATGAAA GGTATTCACAAC
MG55 active effector PAM (5')	6032	MG55 PAM (5')	Nucleotide	Unknown	YTn
MG91 active effector sgRNA	6033	MG91-15 sgRNA1	Nucleotide	Unknown	UAGAGAAAAUUAUAUAU UAGGUUUUGUUAAGCCU AACAAUCGUUAAGUGUU CUUUGGAAUAUUGAUUG UAAAUCUAUUUUGGGAA AUGAAAAGGCAAAAAUU ACAGUUAUCAAUUAAUU GAGAAGAGUAUAGAGUC AGUUUUAUAGUACCAAA AUAUACCUUAAUUCUUU AAGAAAUUAAAUUUAAG GUAAUAACAAG
MG91 active effector sgRNA	6034	MG91-32 sgRNA1	Nucleotide	Unknown	UGCACAUCGGGUAUGUG UGGGGUCGAGUAAGGCC GACGUUGUCCGCUACAA CUUAGCCGUCGGGCGGU UGGGCAACCGAUCGGAA GCGGAACCUGGAAUAAG GCCAGGCAGCGGCACUG CCGUCAAGCGGGAAUGA AGUCCAGUAGUACGUAA CCAGUAACUUACAAAGA AAUUUGUAAGGUACUUA CAGG
MG91 active effector sgRNA	6035	MG91-87 sgRNA1	Nucleotide	Unknown	GAAUUGAUGCUUCGUGC AUCUAAAAAUAUACUGG GAAUUGUAUUCCCGAAG UGAGCGUUAAUUGGCAC AGUGGUGUCAUUGCUCA UCAAAAGAAGAAUUGGA AAAACAGCGAACUUCAU CUCGUUUCUUCACCUUU

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					GGUGCAAGCAAAGGUAA UGAAGUGAAGGCUUUUU AGUACAAUCUCAUACCC CUAACUGUGUGAUACUA UGCCCUCGAAAGAGGGG UAAAUACAGG
MG91 active effector sgRNA	6036	MG91-87 sgRNA2	Nucleotide	Unknown	UGGGAAUUGUAUUCCCG AAGUGAGCGUUAAUUGG CACAGUGGUGUCAUUGC UCAUCAAAAGAAGAAUU GGAAAAACAGCGAACUU CAUCUCGUUUCUUCACC UUUGGUGCAAGCAAAGG UAAUGAAGUGAAGGCUU UUUAGUACAAUCUCAUA CCCCUAACUGUGUGAUA CUAUGCCCUCGAAAGAG GGGUAAAUACAGG
MG91 active effector PAM (5')	6037	MG91-15 PAM (5')	Nucleotide	Unknown	TtTYn
MG91 active effector PAM (5')	6038	MG91-32 PAM (5')	Nucleotide	Unknown	GnYYn
MG91 active effector PAM (5')	6039	MG91-87 PAM (5')	Nucleotide	Unknown	wCCC
MG91 intergenic region potentially encoding tracrRNA	6040	MG91-2 IG2	Nucleotide	Unknown	AAATAACATACAGAGGT TTTGTTAAGCCTCACAAT CTTAATAAATAAGTGTTC TTTGAAAATATTTAGTTG ATTGTAAATCTATTTTGG GAAATAAAAAAACAAAA ATTACAGTTATTAGTTAA CTAAGAAGAGTATAGAG TTAGTTTTAAAGTACCAA AATATACCCTAAATTATT GTTTTTCAAAACTTTATA GAATATAT
MG91 intergenic region potentially encoding tracrRNA	6041	MG91-10 IG1	Nucleotide	Unknown	TATGGCGTGTAATCACC TGGGAATGAAATAAAAA GGCCATGATTTGCATAC TCGGAGACCTGCTCTAT ACTCCCATTCCCGAAGG AACTGACTGTTATTTTTG CATTCTACCATTCTGCAT AATCGTTTGTCGCACCA GCCGCCCACTGGCATTA GCTTAACGTACGCAAGG CCTGACAACACCTTTATG TCGCAAATATACGAATTA TTATTGAAAGTGCGTAC AGAATTCAGAATGAAAG TGCTATACTTTAACAGTA TTTATCCAAAACGTTTAA TATCTTACACATCAATTC ATGTTTGTAAATCACATT

Cat.	SEQ ID:	Description	Type	Organism	Sequence
					TCATTTATTAATCAGACC TGTTTAACGTAGTAACAA GAACGAAATTACTGACT TTCAATCTGCCTTCTATA TGATAGCAGAAAAATTC TCTAGAGAATTTCTTCCT GTATAACTAAGTAGCAA ATGCTATATCACTTAGTA AGGACAAAATTCTCTAG AGAATTTCTTTACTACTC TACAGCTAAGAAATGAG GAACAACAGACTGGGTT TTACTTTATAACCGATAG AGATGGAACGGGTGAGT GTCGGAAGATGTTTAAA TCAAGATTATCGAGAAA TAATTGGCATTCTCAGG AGTATTTACTAAATTTGC ACGAAAATACGAAAAGC ATT
MG91 intergenic region potentially encoding tracrRNA	6042	MG91-52 IG1	Nucleotide	Unknown	GATATAAAGAGGCATTG TCAGGCCTCGCGGACGT TAAGCTAATGCCGGTGG GCGGCTGGTCCGACGAA TTGGTGAGTGGAAGAGG AATGCGAAAATGACAGT CACCCTCGTAAGAGGGC TGAGTTATAGAGTATATC CGCGAGTACGCAAGCCG TGACTACATACTGGCATT GAGCCAGTTAGACATTA TGA
MG91 intergenic region potentially encoding tracrRNA	6043	MG91-61 IG2	Nucleotide	Unknown	TAATTAAGGTTGTGATCT TTATGATTGAGATTGCAA CCTTTTTTTAAAAAAATA TGCAACAAAACGCCCCA AAACGCATCATTTAGGG GTGTTAGGGTCTTGAAT TTGTTAAATCAAGACTTT TATTTTGAATATTTTGGC TCCACAAGGGAATTTTC GCAATTTTGCACTGCAC ACCAGCAATGGCGTGTG GGGGCTTGTAGGCCCGA CGCATCCTCGCTTTAAAA CTCAGCCATCGGGCGGT TATGGGGAGCTGAACGG AAACGGAAATGGGAATA AGACAATGCAGCGCAAG CAGAGCGTGGATGAAGT CTCAGAGTACGCGACCA TTGACTTGCAAAAAGTG ATCTTTGACCAACTGAA GGGTGTAATACACACCG CTGTAAGGCAAATACAG GCTGCAGACTGGCCATC CCAGAATACAATAA
MG91 intergenic region	6044	MG91-69 IG2	Nucleotide	Unknown	TAAAAAAAGATATTAGG TTTTGTTAAGCCTAACAA TCGTTAAGTGTTCTTTGG

Cat.	SEQ ID:	Description	Type	Organism	Sequence
potentially encoding tracrRNA					AATATTGATTGTAAATCT ATTTTGGGAAATAAAAA AGCAAAAATTACAGTTAT CAGTTTACTGAGAAGAG TATAGAGTTAGTTTTAAA GTACCAAAATATACCCTA AATTATTGTTTTTCAAAA CTTTATAGAATATAT
MG91 intergenic region potentially encoding tracrRNA	6045	MG91-94 IG1	Nucleotide	Unknown	CCGTAGTGCATAATCAC CTGAGAATGAATGAAAA AGGCCATGATATGCATA CTTGGGGACTTGCTCTA TACTCCCATTCCCGAAG GAATTGACTGTTATTTTT GCATTCCACCATTCTGCA AATCGTTTGTCGCACCA GCCGCCCACTGGCATTA AGCTTAACGTGCGCAAG GCCTGACAACACCTTTAT GGGTGCAAAGATACTAA AAGTTTTCAAAAGAAGA TAGGGAAAAGAGAAAAA ACTTCAAATCCATTTTGG CGTTGCAGCGTTGCAGT GTTGCAGTTCCAAAATA GATTCTGCGATAGTCAA AAAATAACTCTATATTTA ATAAATATAGAAGTATT
MG91 intergenic region potentially encoding tracrRNA	6046	MG91-101 IG2	Nucleotide	Unknown	TTCAGGAATTCAAAGAT CACCTTTTTTGCAAGTTA CTGGTCACGTACTACTG GACTTCATTCCTGCTTGA CGGCAGTGCCGCTGCAT GGCCTTATTCCAGGTTC CGCTTCCGATTGGCTGC CCAACCGCCCGACGGCT AAGTTGTAGCGGACAGC GTTGGCCCTACAAGACC CCACACATTCCCGATGT GCATTGCAAAGTTCGGG AATGATATCGAGAAAGG CAAAATCTGGCAAAACT TTGTCTTGATTTAACACA TTTATGATTTTTCCGGCG AAAAAATGGTCGCGAAA AATGTCATCCAGGATAT ATGAAGACTCCAGCGGA AAAAGGTTTAACGCCGT TGTAAACTTAGTTTTCTT TGGAGTAAACAGGGTAA CCTCTTGTGGAAATTCG GACAGGATGAGTGTTGA TTGCTTTATCCGGATGTT ATTTTGCTATAGTATTT
MG91 intergenic region potentially encoding	6047	MG91-107 IG2	Nucleotide	Unknown	GTATAAACACTATTGGA CTGTGACAAATAAAGTT AGGATAGGATGTTCTTA CTTTAAATAGTATTTGCT ATCGTTCAAAAAGGCAA GATAGCCATTTTTACATG

Cat.	SEQ ID:	Description	Type	Organism	Sequence
tracrRNA					ATAACATTCTGCTTTATTTCATCACATCTTTGTTTTAGCAATGGCAATTGTTTGCCATTAACGCTAAGATTGGATTACAAATCCATATTTCT
MG91 intergenic region potentially encoding tracrRNA	6048	MG91-155 IG1	Nucleotide	Unknown	CCTATGGTTGGTAATCACCTAGGAATGAATAAAAAGACTATGATAAGCATACTCGGGGACTTGCTCTATACTCCCGTTCTCGAAGGAACTGACTGTTATTTTTGCATTCCACCATTCTGCAAATCGTTTGTCGGACCAGCCGCCCACTGGCATTAGCTTGACGTCCGCAAGGCCTAACAACGCCTTTCTTTTGCAAATATACGGATTTATATTTCTTTCACAAAATTAATGAGGAACTTTTTCATATTCCTTTTGTCCATCTGTTACTGCCACATGATTTTTCGTAGGAAAGACCTTGGAGGCGGATGGATTGAAA
MG91 intergenic region potentially encoding tracrRNA	6049	MG91-201 IG2	Nucleotide	Unknown	ATATATTCTATAAAGTTTTGAAAAACAATAATTTAGGGTATATTTTGGTACTTTAAAACTAACTCTATACTCTTCTTAGTTAACTAATAACTGTAATTTTTGTTTTTTTATTTCCCAAAATAGATTTACAATCAACTAAATATTTCAAAGAACACTTAATTATTAAGATTGTGAGGCTTAACAAAACCTCTGTATGTTATTTTTT
MG91 CRISPR repeat	6050	MG91-2 Repeat	Nucleotide	Unknown	GGTAGATATACCTATTAAAGTTAGGGTACTAACAAG
MG91 CRISPR repeat	6051	MG91-10 Repeat	Nucleotide	Unknown	GGTGTTAATCACCCGGAAATGTAGGCTATTGACAGG
MG91 CRISPR repeat	6052	MG91-52 Repeat	Nucleotide	Unknown	GGTGTAAGACACCTGGAAATGTAAGGCATTGACAGG
MG91 CRISPR repeat	6053	MG91-61 Repeat	Nucleotide	Unknown	CCTGTATTTGGCTTACAAATTAGGATGATTAACACC
MG91 CRISPR repeat	6054	MG91-69 Repeat	Nucleotide	Unknown	GGTGTATATACCTAATAAAGTTAGGGTACTAACAAG
MG91 CRISPR repeat	6055	MG91-94 Repeat	Nucleotide	Unknown	GGTGTTAATCACCCGAAAATGTAGGTCATTGACAGGA
MG91 CRISPR repeat	6056	MG91-101 Repeat	Nucleotide	Unknown	GGGTGTTACACACCCGAATTTGCAAGGTACTTACAGG

Cat.	SEQ ID:	Description	Type	Organism	Sequence
MG91 CRISPR repeat	6057	MG91-107 Repeat	Nucleotide	Unknown	CCTGTATGTGACTTTGAAATATAGGGTAGTTACAAG
MG91 CRISPR repeat	6058	MG91-155 Repeat	Nucleotide	Unknown	GGTGTTAATCACCCAGAAATGTAGACTATTAACAGG
MG91 CRISPR repeat	6059	MG91-201 Repeat	Nucleotide	Unknown	GCTGTATATGCCTAATAAAGTTAGGGTACTAACAAGAA

mN: 2'-Omethyl modified base N; fN: 2'-Fluoro modified base N; *: phosphorothioate linkage; N: standard ribonucleotide base

[00632] While preferred embodiments of the present invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. It is not intended that the invention be limited by the specific examples provided within the specification. While the invention has been described with reference to the aforementioned specification, the descriptions and illustrations of the embodiments herein are not meant to be construed in a limiting sense. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention. Furthermore, it shall be understood that all aspects of the invention are not limited to the specific depictions, configurations or relative proportions set forth herein which depend upon a variety of conditions and variables. It should be understood that various alternatives to the embodiments of the invention described herein may be employed in practicing the invention. It is therefore contemplated that the invention shall also cover any such alternatives, modifications, variations or equivalents. It is intended that the following claims define the scope of the invention and that methods and structures within the scope of these claims and their equivalents be covered thereby.

EMBODIMENTS

The following embodiments are not intended to be limiting in any way.

1. An engineered nuclease system comprising:

(a) an endonuclease comprising a RuvC domain, wherein said endonuclease is derived from an uncultivated microorganism, and wherein said endonuclease is a Cas12a endonuclease; and

(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence.

2. An engineered nuclease system comprising:

(a) an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof; and

(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence.

3. An engineered nuclease system comprising:

(a) an endonuclease configured to bind to a protospacer adjacent motif (PAM) sequence comprising any one of SEQ ID NOs: 3862-3913, wherein said endonuclease is a class 2, type V Cas endonuclease; and

(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence.

4. The engineered nuclease system of any one of embodiments 1-3, wherein said endonuclease further comprises a zinc finger-like domain.

5. The engineered nuclease system of any one of embodiments 1-4, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3551-3559, 3608-3609, 3612, 3636-3637,

3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, -3678, 3695-3696, 3729-3730, 3734-3735, and 3851-3857.

6. An engineered nuclease system comprising:

(a) an engineered guide RNA comprising a sequence with at least 80% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3677-3678, 3695-3696, 3729-3730, 3734-3735, and 3851-3857, and

(b) a class 2, type V Cas endonuclease configured to bind to said engineered guide RNA.

7. The engineered nuclease system of any of embodiments 1-6, wherein said endonuclease is configured to bind to a protospacer adjacent motif (PAM) sequence comprising any one of SEQ ID NOs: 3863-3913.

8. The engineered nuclease system of any one of embodiments 1-7, wherein said guide RNA comprises a sequence complementary to a eukaryotic, fungal, plant, mammalian, or human genomic polynucleotide sequence.

9. The engineered nuclease system of any one of embodiments 1-8, wherein said guide RNA is 30-250 nucleotides in length.

10. The engineered nuclease system of any one of embodiments 1-9, wherein said endonuclease comprises one or more nuclear localization sequences (NLSs) proximal to an N- or C-terminus of said endonuclease.

11. The engineered nuclease system of any one of embodiments 1-10, wherein said NLS comprises a sequence at least 80% identical to a sequence from the group consisting of SEQ ID NO: 3938-3953.

12. The engineered nuclease system of any one of embodiments 1-10, wherein said endonuclease comprises at least one of the following mutations: S168R, E172R, N577R, or Y170R when a sequence of said endonuclease is optimally aligned to SEQ ID NO: 215.

13. The engineered nuclease system of any one of embodiments 1-10, wherein said endonuclease comprises the mutations S168R and E172R when a sequence of said endonuclease is optimally aligned to SEQ ID NO: 215.

14. The engineered nuclease system of any one of embodiments 1-10, wherein said endonuclease comprises the mutations N577R or Y170R when a sequence of said endonuclease is optimally aligned to SEQ ID NO: 215.

15. The engineered nuclease system of any one of embodiments 1-10, wherein said endonuclease comprises the mutation S168R when a sequence of said endonuclease is optimally aligned to SEQ ID NO: 215.

16. The engineered nuclease system of embodiment 15, wherein said endonuclease does not comprise a mutation of E172, N577, or Y170.

17. The engineered nuclease system of any one of embodiments 1-16, further comprising
a single- or double-stranded DNA repair template comprising from 5′ to 3′: a first homology arm comprising a sequence of at least 20 nucleotides 5′ to said target deoxyribonucleic acid sequence, a synthetic DNA sequence of at least 10 nucleotides, and a second homology arm comprising a sequence of at least 20 nucleotides 3′ to said target sequence.

18. The engineered nuclease system of embodiment 17, wherein said first or second homology arm comprises a sequence of at least 40, 80, 120, 150, 200, 300, 500, or 1,000 nucleotides.

19. The engineered nuclease system of any one of embodiments 12-18, wherein said first and second homology arms are homologous to a genomic sequence of a prokaryote, bacteria, fungus, or eukaryote.

20. The engineered nuclease system of embodiments 12-19, wherein said single- or double-stranded DNA repair template comprises a transgene donor.

21. The engineered nuclease system of any one of embodiments 1-20, further comprising a DNA repair template comprising a double-stranded DNA segment flanked by one or two single-stranded DNA segments.

22. The engineered nuclease system of embodiment 21, wherein single-stranded DNA segments are conjugated to the 5′ ends of said double-stranded DNA segment.

23. The engineered nuclease system of embodiment 21, wherein said single stranded DNA segments are conjugated to the 3′ ends of said double-stranded DNA segment.

24. The engineered nuclease system of any one of embodiments 21-23, wherein said single-stranded DNA segments have a length from 4 to 10 nucleotide bases.

25. The engineered nuclease system of any one of embodiments 21-24, wherein said single-stranded DNA segments have a nucleotide sequence complementary to a sequence within said spacer sequence.

26. The engineered nuclease system of any one of embodiments 21-25, wherein said double-stranded DNA sequence comprises a barcode, an open reading frame, an enhancer, a promoter, a protein-coding sequence, a miRNA coding sequence, an RNA coding sequence, or a transgene.

27. The engineered nuclease system of any one of embodiments 21-25, wherein said double-stranded DNA sequence is flanked by a nuclease cut site.

28. The engineered nuclease system of embodiment 27, wherein said nuclease cut site comprises a spacer and a PAM sequence.

29. The engineered nuclease system of any one of embodiments 1-28, wherein said system further comprises a source of Mg^{2+}.

30. The engineered nuclease system of any one of embodiments 1-29, wherein said guide RNA comprises a hairpin comprising at least 8, at least 10, or at least 12 base-paired ribonucleotides.

31. The engineered nuclease system of embodiment 30, wherein said hairpin comprises 10 base-paired ribonucleotides.

32. The engineered nuclease system of any one of embodiments 1-31, wherein:

a) said endonuclease comprises a sequence at least 75%, 80%, or 90% identical to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1721 or a variant thereof; and

b) said guide RNA structure comprises a sequence at least 80%, or 90% identical to the non-degenerate nucleotides of any one of SEQ ID NOs: 3608-3609, 3853, or 3851-3857.

33. The engineered nuclease system of any one of embodiments 1-32, wherein said endonuclease is configured to bind to a PAM comprising any one of SEQ ID NOs: 3863-3913.

34. The engineered nuclease system of any one of embodiments 1-32, wherein said endonuclease is configured to bind to a PAM comprising SEQ ID NO: 3871.

35. The engineered nuclease system of any one of embodiments 5-34, wherein said sequence identity is determined by a BLASTP, CLUSTALW, MUSCLE, MAFFT algorithm, or a CLUSTALW algorithm with the Smith-Waterman homology search algorithm parameters.

36. The engineered nuclease system of embodiment 35, wherein said sequence identity is determined by said BLASTP homology search algorithm using parameters of a wordlength (W) of 3, an expectation (E) of 10, and a BLOSUM62 scoring matrix setting gap costs at existence of 11, extension of 1, and using a conditional compositional score matrix adjustment.

37. An engineered guide RNA comprising:

a) a DNA-targeting segment comprising a nucleotide sequence that is complementary to a target sequence in a target DNA molecule; and

b) a protein-binding segment comprising two complementary stretches of nucleotides that hybridize to form a double-stranded RNA (dsRNA) duplex,

wherein said two complementary stretches of nucleotides are covalently linked to one another with intervening nucleotides, and

wherein said engineered guide ribonucleic acid polynucleotide is capable of forming a complex with an endonuclease having at least 75% sequence

identity to any one of SEQ ID NOs: 1-3470, and targeting said complex to said target sequence of said target DNA molecule.

38. The engineered guide ribonucleic acid polynucleotide of embodiment 37, wherein said DNA-targeting segment is positioned 3′ of both of said two complementary stretches of nucleotides.

39. The engineered guide ribonucleic acid polynucleotide of embodiment 37-38, wherein said protein binding segment comprises a sequence having at least 70%, at least 80%, or at least 90% identity to the non-degenerate nucleotides of SEQ ID NO: 3608-3609.

40. The engineered guide ribonucleic acid polynucleotide of any one of embodiments 37-39, wherein said double-stranded RNA (dsRNA) duplex comprises at least 5, at least 8, at least 10, or at least 12 ribonucleotides.

41. A deoxyribonucleic acid polynucleotide encoding the engineered guide ribonucleic acid polynucleotide of any one of embodiments 1-40.

42. A nucleic acid comprising an engineered nucleic acid sequence optimized for expression in an organism, wherein said nucleic acid encodes a class 2, type V Cas endonuclease, and wherein said endonuclease is derived from an uncultivated microorganism, wherein the organism is not said uncultivated organism.

43. The nucleic acid of embodiment 42, wherein said endonuclease comprises a variant having at least 70% or at least 80% sequence identity to any one of SEQ ID NOs: 1-3470.

44. The nucleic acid of embodiment 42 or 43, wherein said endonuclease comprises a sequence encoding one or more nuclear localization sequences (NLSs) proximal to an N- or C-terminus of said endonuclease.

45. The nucleic acid of embodiment 44, wherein said NLS comprises a sequence selected from SEQ ID NOs: 3938-3953.

46. The nucleic acid of embodiment 44or 45, wherein said NLS comprises SEQ ID NO: 3939.

47. The nucleic acid of embodiment 46, wherein said NLS is proximal to said N-terminus of said endonuclease.

48. The nucleic acid of embodiment 44 or 45, wherein said NLS comprises SEQ ID NO: 3938.

49. The nucleic acid of embodiment 48, wherein said NLS is proximal to said C-terminus of said endonuclease.

50. The nucleic acid of any one of embodiments 42-49, wherein said organism is prokaryotic, bacterial, eukaryotic, fungal, plant, mammalian, rodent, or human.

51. An engineered vector comprising a nucleic acid sequence encoding a class 2, type V Cas endonuclease, wherein said endonuclease is derived from an uncultivated microorganism.

52. An engineered vector comprising the nucleic acid of any of embodiments 42-46.

53. An engineered vector comprising the deoxyribonucleic acid polynucleotide of embodiment 41.

54. The engineered vector of any of embodiments 51-53, wherein the vector is a plasmid, a minicircle, a CELiD, an adeno-associated virus (AAV) derived virion, a lentivirus, or an adenovirus.

55. A cell comprising the vector of any of embodiments 51-54.

56. A method of manufacturing an endonuclease, comprising cultivating said cell of embodiment 55.

57. A method for binding, cleaving, marking, or modifying a double-stranded deoxyribonucleic acid polynucleotide, comprising:

(a) contacting said double-stranded deoxyribonucleic acid polynucleotide with a class 2, type V Cas endonuclease in complex with an engineered guide RNA configured to bind to said endonuclease and said double-stranded deoxyribonucleic acid polynucleotide;

wherein said double-stranded deoxyribonucleic acid polynucleotide comprises a protospacer adjacent motif (PAM); and

wherein said PAM comprises a sequence comprising any one of SEQ ID NOs: 3863-3913.

58. The method of embodiment 57, wherein said double-stranded deoxyribonucleic acid polynucleotide comprises a first strand comprising a sequence complementary to a sequence of said engineered guide RNA and a second strand comprising said PAM.

59. The method of embodiment 58, wherein said PAM is directly adjacent to the 5′ end of said sequence complementary to said sequence of said engineered guide RNA.

60. The method of any one of embodiments 57-59, wherein said PAM comprises SEQ ID NO: 3871.

61. The method of any one of embodiments 57-60, wherein said class 2, type V Cas endonuclease is derived from an uncultivated microorganism.

62. The method of any one of embodiments 57-61, wherein said double-stranded deoxyribonucleic acid polynucleotide is a eukaryotic, plant, fungal, mammalian, rodent, or human double-stranded deoxyribonucleic acid polynucleotide.

63. A method of modifying a target nucleic acid locus, said method comprising delivering to said target nucleic acid locus said engineered nuclease system of any one of embodiments 1-36, wherein said endonuclease is configured to form a complex with said engineered guide ribonucleic acid structure, and wherein said complex is configured such that upon binding of said complex to said target nucleic acid locus, said complex modifies said target nucleic acid locus.

64. The method of embodiment 63, wherein modifying said target nucleic acid locus comprises binding, nicking, cleaving, or marking said target nucleic acid locus.

65. The method of embodiment 63 or 64, wherein said target nucleic acid locus comprises deoxyribonucleic acid (DNA) or ribonucleic acid (RNA).

66. The method of embodiment 63, wherein said target nucleic acid comprises genomic DNA, viral DNA, viral RNA, or bacterial DNA.

67. The method of any one of embodiments 63-66, wherein said target nucleic acid locus is *in vitro*.

68. The method of any one of embodiments 63-66, wherein said target nucleic acid locus is within a cell.

69. The method of embodiment 68, wherein said cell is a prokaryotic cell, a bacterial cell, a eukaryotic cell, a fungal cell, a plant cell, an animal cell, a mammalian cell, a rodent cell, a primate cell, a human cell, or a primary cell.

70. The method of embodiment 68 or 69, wherein said cell is a primary cell.

71. The method of embodiment 70, wherein said primary cell is a T cell.

72. The method of embodiment 70, wherein said primary cell is a hematopoietic stem cell (HSC).

73. A method of any one of embodiments 63-72, wherein delivering said engineered nuclease system to said target nucleic acid locus comprises delivering the nucleic acid of any of embodiments 42-46 or the vector of any of embodiments 51-54.

74. The method of any one of embodiments 63-73, wherein delivering said engineered nuclease system to said target nucleic acid locus comprises delivering a nucleic acid comprising an open reading frame encoding said endonuclease.

75. The method of embodiment 74, wherein said nucleic acid comprises a promoter to which said open reading frame encoding said endonuclease is operably linked.

76. The method of any one of embodiments 63-75, wherein delivering said engineered nuclease system to said target nucleic acid locus comprises delivering a capped mRNA containing said open reading frame encoding said endonuclease.

77. The method of any one of embodiments 63-76, wherein delivering said engineered nuclease system to said target nucleic acid locus comprises delivering a translated polypeptide.

78. The method of any one of embodiments 63-76, wherein delivering said engineered nuclease system to said target nucleic acid locus comprises delivering a deoxyribonucleic acid (DNA) encoding said engineered guide RNA operably linked to a ribonucleic acid (RNA) pol III promoter.

79. The method of any one of embodiments 63-78, wherein said endonuclease induces a single-stranded break or a double-stranded break at or proximal to said target locus.

80. The method of embodiment 79, wherein said endonuclease induces a staggered single stranded break within or 3′ to said target locus.

81. A method of editing a TRAC locus in a cell, comprising contacting to said cell

(a) an RNA-guided endonuclease; and

(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said TRAC locus,

wherein said engineered guide RNA comprises a targeting sequence having at least 85% identity at least 18 consecutive nucleotides of any one of SEQ ID NOs: 4316-4369.

82. The method of embodiment 81, wherein said RNA-guided nuclease is a Cas endonuclease.

83. The method of embodiment 82, wherein said Cas endonuclease is a class 2, type V Cas endonuclease.

84. The method of embodiment 83, wherein said class 2, type V Cas endonuclease comprises a RuvC domain comprising a RuvCI subdomain, a RuvCII subdomain, and a RuvCIII subdomain.

85. The method of embodiment 83 or 84, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

86. The method of any one of embodiments 81-85, wherein said engineered guide RNA further comprises a sequence with at least 80% sequence identity to at least 19 of the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3677-3678, 3695-3696, 3729-3730, 3734-3735, and 3851-3857.

87. The method of any one of embodiments 81-85, wherein said endonuclease comprises a sequence at least 75%, 80%, or 90% identical to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1721 or a variant thereof.

88. The method of embodiment 87, wherein said guide RNA structure comprises a sequence at least 80%, or at least 90% identical to at least 19 of the non-degenerate nucleotides of any one of SEQ ID NOs: 3608-3609, 3853, or 3851-3857.

89. The method of any one of embodiments 81-88, wherein said method further comprises contacting to said cell or introducing to said cell a donor nucleic acid comprising a cargo sequence flanked on a 3' or 5' end by sequence having at least 80% identity to any one of SEQ ID NOs: 4424 or 4425.

90. The method of any one of embodiments 81-89, wherein said cell is a peripheral blood mononuclear cell (PBMC).

91. The method of any one of embodiments 81-89, wherein said cell is a T-cell or a precursor thereof or a hematopoietic stem cell (HSC).

92. The method of any one of embodiments 89-91, wherein said cargo sequence comprises a sequence encoding a T-cell receptor polypeptide, a CAR-T polypeptide, or a fragment or derivative thereof.

93. The method of any one of embodiments 81-92, wherein said engineered guide RNA comprises a sequence having at least 80% identity to any one of SEQ ID NOs:4370-4423.

94. The method of embodiment 93, wherein said engineered guide RNA is comprises the nucleotide sequence of sgRNAs 1-54 from Table 5A comprising the corresponding chemical modifications listed in Table 5A.

95. The method of any one of embodiments 81-93, wherein said engineered guide RNA comprises a targeting sequence having at least 80% sequence identity to any one of SEQ ID NOs: 4334, 4350, or 4324.

96. The method of any one of embodiments 81-93, wherein said engineered guide RNA comprises a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 4388, 4404, or 4378.

97. The method of embodiment 96, wherein said engineered guide RNA comprises the nucleotide sequence of sgRNAs 9, 35, or 19 from Table 5A.

98. An engineered nuclease system comprising:

(a) an RNA-guided endonuclease; and

(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence,

wherein said engineered guide RNA comprises at least one of the following modifications:

(i) a 2'-O methyl or a 2'-fluoro base modification of at least one nucleotide within the first 4 bases of the 5' end of said engineered guide RNA or the last 4 bases of a 3' end of said engineered guide RNA;

(ii) a thiophosphate (PS) linkage between at least 2 of the first five bases of a 5' end of said engineered guide RNA, or a thiophosphate linkage between at least two of the last five bases of a 3' end of said engineered guide RNA;

(iii) a thiophosphate linkage within a 3' stem or a 5' stem of said engineered guide RNA;

(iv) a 2'-O methyl or 2'base modification within a 3' stem or a 5' stem of said engineered guide RNA;

(v) a 2'-fluoro base modification of at least 7 bases of a spacer region of said engineered guide RNA; and

(vi) a thiophosphate linkage within a loop region of said engineered guide RNA.

99. The system of embodiment 98, wherein said engineered guide RNA comprises a 2'-O methyl or a 2'-fluoro base modification of at least one nucleotide within the first 5 bases of a 5' end of said engineered guide RNA or the last 5 bases of a 3' end of said engineered guide RNA.

100. The system of embodiment 98, wherein said engineered guide RNA comprises a 2'-O methyl or a 2'-fluoro base modification at a 5' end of said engineered guide RNA or a 3' end of said engineered guide RNA.

101. The system of any one of embodiments 98-100, wherein said engineered guide RNA comprises a thiophosphate (PS) linkage between at least 2 of the first five bases of a 5' end of said engineered guide RNA, or a thiophosphate linkage between at least two of the last five bases of a 3' end of said engineered guide RNA.

102. The system of any one of embodiments 98-101, wherein said engineered guide RNA comprises a thiophosphate linkage within a 3' stem or a 5' stem of said engineered guide RNA.

103. The system of any one of embodiments 98-102, wherein said engineered guide RNA comprises a 2'-O methyl base modification within a 3' stem or a 5' stem of said engineered guide RNA.

104. The system of any one of embodiments 98-103, wherein said engineered guide RNA comprises a 2'-fluoro base modification of at least 7 bases of a spacer region of said engineered guide RNA.

105. The system of any one of embodiments 98-104, wherein said engineered guide RNA comprises a thiophosphate linkage within a loop region of said engineered guide RNA.

106. The system of any one of embodiments 98-105, wherein said engineered guide RNA comprises at least three 2'-O methyl or 2'-fluoro bases at said 5' end of said engineered

guide RNA, two thiophosphate linkages between the first 3 bases of said 5' end of said engineered guide RNA, at least 4 2'-O methyl or 2'-fluoro bases at said 4' end of said engineered guide RNA, and three thiophosphate linkages between the last three bases of said 3' end of said engineered guide RNA.

107. The system of embodiment 98, wherein said engineered guide RNA comprises at least two 2'-O-methyl bases and at least two thiophosphate linkages at a 5' end of said engineered guide RNA and at least one 2'-O-methyl bases and at least one thiophosphate linkage at a 3' end of said engineered guide RNA.

108. The system of embodiment 107, wherein said engineered guide RNA comprises at least one 2'-O-methyl base in both said 3' stem or said 5' stem region of said engineered guide RNA.

109. The system of embodiment 107 or 108, wherein said engineered guide RNA comprises at least one to at least fourteen 2'-fluoro bases in said spacer region excluding a seed region of said engineered guide RNA.

110. The system of embodiment 107, wherein said engineered guide RNA comprises at least one 2'-O-methyl base in said 5' stem region of said engineered guide RNA and at least one to at least fourteen 2'-fluoro bases in said spacer region excluding a seed region of said guide RNA.

111. The system of any one of embodiments 98-110, wherein said guide RNA comprises a spacer sequence targeting a VEGF-A gene.

112. The system of embodiment 111, wherein said guide RNA comprises a spacer sequence having at least 80% identity to SEQ ID NO: 3985.

113. The system of embodiment 111, wherein said guide RNA comprises the nucleotides of guide RNAs 1-7 from Table 7 comprising the chemical modifications listed in Table 7.

114. The method of any one of embodiments 98-113, wherein said RNA-guided nuclease is a Cas endonuclease.

115. The method of embodiment 114, wherein said Cas endonuclease is a class 2, type V Cas endonuclease.

116. The method of embodiment 115, wherein said class 2, type V Cas endonuclease comprises a RuvC domain comprising a RuvCI subdomain, a RuvCII subdomain, and a RuvCIII subdomain.

117. The method of any one of embodiments 115-116, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

118. The method of any one of embodiments 115-116, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1721 or a variant thereof.

119. The method of any one of embodiments 114-118, wherein said engineered guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3677-3678, 3695-3696, 3729-3730, 3734-3735, and 3851-3857.

120. The method of any one of embodiments 111-118, wherein said engineered guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3608 -3609, 3853, or 3851-3857.

121. A host cell comprising an open reading frame encoding a heterologous endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

122. The host cell of embodiment 121, wherein said endonuclease has at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1721, or a variant thereof.

123. The host cell of any one of embodiments 121-122 wherein said host cell is an *E. coli* cell.

124. The host cell of embodiment 123, wherein said *E. coli* cell is a λDE3 lysogen or said *E. coli* cell is a BL21(DE3) strain.

125. The host cell of any one of embodiments 109-110, wherein said *E. coli* cell has an *ompT lon* genotype.

126. The host cell of any one of embodiments 121-125, wherein said open reading frame is operably linked to a T7 promoter sequence, a T7-lac promoter sequence, a lac promoter sequence, a tac promoter sequence, a trc promoter sequence, a ParaBAD promoter sequence, a PrhaBAD promoter sequence, a T5 promoter sequence, a *cspA* promoter sequence, an $araP_{BAD}$ promoter, a strong leftward promoter from phage lambda (pL promoter), or any combination thereof.

127. The host cell of any one of embodiments 121-126, wherein said open reading frame comprises a sequence encoding an affinity tag linked in-frame to a sequence encoding said endonuclease.

128. The method of embodiment 127, wherein said affinity tag is an immobilized metal affinity chromatography (IMAC) tag.

129. The method of embodiment 128, wherein said IMAC tag is a polyhistidine tag.

130. The method of embodiment 127, wherein said affinity tag is a myc tag, a human influenza hemagglutinin (HA) tag, a maltose binding protein (MBP) tag, a glutathione S-transferase (GST) tag, a streptavidin tag, a FLAG tag, or any combination thereof.

131. The host cell of any one of embodiments 127-130, wherein said affinity tag is linked in-frame to said sequence encoding said endonuclease via a linker sequence encoding a protease cleavage site.

132. The host cell of embodiment 131, wherein said protease cleavage site is a tobacco etch virus (TEV) protease cleavage site, a PreScission® protease cleavage site, a Thrombin cleavage site, a Factor Xa cleavage site, an enterokinase cleavage site, or any combination thereof.

133. The host cell of any one of embodiments 121-132, wherein said open reading frame is codon-optimized for expression in said host cell.

134. The host cell of any one of embodiments 121-133, wherein said open reading frame is provided on a vector.

135. The host cell of any one of embodiments 121-133, wherein said open reading frame is integrated into a genome of said host cell.

136. A culture comprising the host cell of any one of embodiments 121-135 in compatible liquid medium.

137. A method of producing an endonuclease, comprising cultivating the host cell of any one of embodiments 121-135 in compatible growth medium.

138. The method of embodiment 137, further comprising inducing expression of said endonuclease by addition of an additional chemical agent or an increased amount of a nutrient.

139. The method of embodiment 138, wherein an additional chemical agent or an increased amount of a nutrient comprises Isopropyl β-D-1-thiogalactopyranoside (IPTG) or additional amounts of lactose.

140. The method of any one of embodiments 137-139, further comprising isolating said host cell after said cultivation and lysing said host cell to produce a protein extract.

141. The method of embodiment 140, further comprising subjecting said protein extract to IMAC, or ion-affinity chromatography.

142. The method of embodiment 141, wherein said open reading frame comprises a sequence encoding an IMAC affinity tag linked in-frame to a sequence encoding said endonuclease.

143. The method of embodiment 142, wherein said IMAC affinity tag is linked in-frame to said sequence encoding said endonuclease via a linker sequence encoding protease cleavage site.

144. The method of embodiment 143, wherein said protease cleavage site comprises a tobacco etch virus (TEV) protease cleavage site, a PreScission® protease cleavage site, a

Thrombin cleavage site, a Factor Xa cleavage site, an enterokinase cleavage site, or any combination thereof.

145. The method of any one of embodiments 143-144, further comprising cleaving said IMAC affinity tag by contacting a protease corresponding to said protease cleavage site to said endonuclease.

146. The method of embodiment 145, further comprising performing subtractive IMAC affinity chromatography to remove said affinity tag from a composition comprising said endonuclease.

147. A system comprising

(a) a class 2, Type V-A Cas endonuclease configured to bind a 3- or 4-nucleotide PAM sequence, wherein said endonuclease has increased cleavage activity relative to sMbCas12a; and

(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said class 2, Type V-A Cas endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid comprising a target nucleic acid sequence.

148. The system of embodiment 147, wherein said cleavage activity is measured *in vitro* by introducing said endonucleases alongside compatible guide RNAs to cells comprising said target nucleic acid and detecting cleavage of said target nucleic acid sequence in said cells.

149. The system of any one of embodiments 147-148, wherein said class 2, Type V-A Cas endonuclease comprises a sequence having at least 75% identity to any one of 215-225 or a variant thereof.

150. The system of embodiment 149, wherein said engineered guide RNA comprises a sequence having at least 80% identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

151. The system of any one of embodiments 149-150, wherein said target nucleic acid further comprises a YYN PAM sequence proximal to said target nucleic acid sequence.

152. The system of any one of embodiments 147-151, wherein said class 2, Type V-A Cas endonuclease has at least about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, or 200%, or more increased activity relative to sMbCas12a.

153. A system comprising:

(a) a class 2, Type V-A' Cas endonuclease; and

(b) an engineered guide RNA, wherein said engineered guide RNA comprises a sequence having at least 80% identity to about 19 to about 25 or about 19 to about 31 consecutive nucleotides of a natural effector repeat sequence of a class 2, Type V Cas endonuclease.

154. The system of embodiment 153, wherein said natural effector repeat sequence is any one of SEQ ID NOs: 3560-3572.

155. The system of any one of embodiments 153-154, wherein said class 2, Type V-A' Cas endonuclease has at least 75% identity to SEQ ID NO: 126.

156. A method of disrupting the VEGF-A locus in a cell, comprising introducing to said cell:

(b) a class 2, type V Cas endonuclease; and

(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said VEGF-A locus,

wherein said engineered guide RNA comprises a targeting sequence having at least 80% identity to SEQ ID NO: 3985; or

wherein said engineered guide RNA comprises the nucleotide sequence of any one of guide RNAs 1-7 from Table 7.

157. The system of embodiment 156, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

158. The system of any one of embodiments 156-157, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1721 or a variant thereof.

159. The system of any one of embodiments 156-158, wherein said engineered guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3677-3678, 3695-3696, 3729-3730, 3734-3735, and 3851-3857.

160. The system of any one of embodiments 156-158, wherein said engineered guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3608 -3609, 3853, or 3851-3857.

161. A method of disrupting a locus in a cell, comprising contacting to said cell a composition comprising:

(a) a class 2, type V Cas endonuclease having at least 75% identity to any one of SEQ ID NOs: 215-225 or a variant thereof; and

(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said locus,

wherein said class 2, type V Cas endonuclease has at least equivalent cleavage activity to spCas9 in said cell.

162. The method of embodiment 161, wherein said cleavage activity is measured *in vitro* by introducing said endonucleases alongside compatible guide RNAs to cells comprising said target nucleic acid and detecting cleavage of said target nucleic acid sequence in said cells.

163. The method of any one of embodiments 161-162, wherein said composition comprises 20 pmoles or less of said class 2, type V Cas endonuclease.

164. The method of embodiment 163, wherein said composition comprises 1 pmol or less of said class 2, type V Cas endonuclease.

CLAIMS

WHAT IS CLAIMED IS:

1. A method of disrupting a CD38 locus in a cell, comprising introducing to said cell:
(a) a class 2, type V Cas endonuclease; and
(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said CD38 locus,

wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4466-4503 and 5686; or

wherein said engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4428-4465 and 5685.

2. The method of claim 1, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

3. The method of claim 1, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

4. The method of any one of claims 1-3, wherein said engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-

3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, and 6033-6036.

5. The method of any one of claims 1-4, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

6. The method of any one of claims 1-5, wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 4466, 4467, 4468, 4479, 4484, 4490, 4492, 4493, 4495, 4498.

7. The method of any one of claims 1-6, wherein said engineered guide RNA comprises a nucleotide sequence having at least 80% identity to any one of SEQ ID NOs: 4428, 4429, 4430, 4436, 4441, 4446, 4452, 4454, 4455, 4460, or 4461.

8. The method of any one of claims 1-7, wherein said cell is a eukaryotic cell, T-cell, hematopoietic stem cell, or precursor thereof.

9. A method of disrupting a TIGIT locus in a cell, comprising introducing to said cell:
(a) a class 2, type V Cas endonuclease; and
(b) an engineered guide RNA, wherein said engineerewhd guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said TIGIT locus,

wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4521-4537; or

wherein said engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4504-4520.

10. The method of claim 9, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

11. The method of claim 9, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

12. The method of any one of claims 9-11, wherein said engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, and 6033-6036.

13. The method of any one of claims 9-12, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

14. The method of any one of claims 9-13, wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 4521, 4527, 4528, 4535, or 4536.

15. The method of any one of claims 9-13, wherein said engineered guide RNA comprises a nucleotide sequence having at least 80% identity to any one of SEQ ID NOs: 4504, 4510, 4511, 4518, or 4519.

16. The method of any one of claims 9-15, wherein said cell is a eukaryotic cell, T-cell, hematopoietic stem cell, or precursor thereof.

17. A method of disrupting an AAVS1 locus in a cell, comprising introducing to said cell: (a) a class 2, type V Cas endonuclease; and

(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said AAVS1 locus,

wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4569-4599; or

wherein said engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4538-4568.

18. The method of claim 17, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

19. The method of claim 18, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

20. The method of any one of claims 17-19, wherein said engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, 6033-6036.

21. The method of of any one of claims 17-20, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

22. The method of any one of claims 17-21, wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 4574, 4577, 4578, 4579, 4582, 4584, 4585, 4586, 4587, 4589, 4590, 4591, 4592, 4593, 4595, 4596, or 4598.

23. The method of any one of claims 17-21, wherein said engineered guide RNA comprises a nucleotide sequence having at least 80% identity to any one of SEQ ID NOs: 4543, 4546, 4547, 4548, 4551, 4553, 4554, 4555, 4556, 4558, 4559, 4560, 4561, 4562, 4565, or 4567.

24. The method of any one of claims 17-23, wherein said cell is a eukaryotic cell, T-cell, hematopoietic stem cell, hepatocyte, or precursor thereof.

25. A method of disrupting a B2M locus in a cell, comprising introducing to said cell:
(a) a class 2, type V Cas endonuclease; and
(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said B2M locus,

wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4676-4751; or

wherein said engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4600-4675.

26. The method of claim 25, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

27. The method of claim 26, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

28. The method of any one of claims 25-27, wherein said engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857 and 6033-6036.

29. The method of any one of claims 25-28, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

30. The method of any one of claims 25-29, wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 4676, 4678-4687, 4690, 4692, 4698-4707, 4720-4723, 4725-4726, 4732-4733, 4736-4737, 4741, or 4750-4751.

31. The method of any one of claims 25-30, wherein said engineered guide RNA comprises a nucleotide sequence having at least 80% identity to any one of SEQ ID NOs: 4600, 4602-4611, 4614, 4616, 4622-4631, 4644-4647, 4649-4650, 4656-4657, 4660-4661, 4665, or 4674-4675.

32. The method of any one of claims 25-31, wherein said cell is a eukaryotic cell, T-cell, hematopoietic stem cell, or precursor thereof.

33. A method of disrupting a CD2 locus in a cell, comprising introducing to said cell:
(a) a class 2, type V Cas endonuclease; and
(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said CD2 locus,

wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4837-4921; or

wherein said engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4752-4836.

34. The method of claim 33, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

35. The method of claim 34, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

36. The method of any one of claims 33-35, wherein said engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, and 6033-6036.

37. The method of claim 36, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

38. The method of any one of claims 33-37, wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at

least 80% identity to any one of SEQ ID NOs: 4837, 4844, 4845, 4848, 4857-4858, 4883, 4887, 4892-4893, 4904-4909, 4914, 4916, or 4918.

39. The method of claim 38, wherein said engineered guide RNA has at least 80% sequence identity to any of the guide RNAs from Table 14E that target any one of SEQ ID NOs: 4837, 4844, 4845, 4848, 4857-4858, 4883, 4887, 4892-4893, 4904-4909, 4914, 4916, or 4918.

40. The method of any one of claims 33-39, wherein said cell is a eukaryotic cell, T-cell, hematopoietic stem cell, or precursor thereof.

41. A method of disrupting a CD5 locus in a cell, comprising introducing to said cell:
(a) a class 2, type V Cas endonuclease; and
(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said CD5 locus,

wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4946-4969; or

wherein said engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4922-4945.

42. The method of claim 41, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

43. The method of claim 42, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

44. The method of any one of claims 41-43, wherein said engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, and 6033-6036.

45. The method of claim 44, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

46. The method of any one of claims 41-45, wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 4946-4947, 4949, 4951, 4957-4960, 4963, 4967, or 4969.

47. The method of any one of claims 41-45, wherein said engineered guide RNA has at least 80% sequence identity to any of the guide RNAs from Table 14F that target any one of SEQ ID NOs: 4946-4947, 4949, 4951, 4957-4960, 4963, 4967, or 4969.

48. The method of any one of claims 41-47, wherein said cell is a eukaryotic cell, T-cell, hematopoietic stem cell, or precursor thereof.

49. A method of disrupting a mouse TRAC locus in a cell, comprising introducing to said cell:

(a) a class 2, type V Cas endonuclease; and

(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said mouse TRAC locus,

wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at

least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5126-5195, 5682, or 5684; or

wherein said engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5056-5125, 5681, or 5683.

50. The method of claim 49, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

51. The method of claim 50, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

52. The method of any one of claims 49-51, wherein said engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3677-3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036.

53. The method of claim 52, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

54. The method of any one of claims 49-53, wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 5126-5130, 5133-5143, 5147-5150, 5172-5173, 5184-5189, or 5192-5194.

55. The method of any one of claims 49-53, wherein said engineered guide RNA has at least 80% sequence identity to any of the guide RNAs from Table 14G that target any one of SEQ ID NOs: 5126-5130, 5133-5143, 5147-5150, 5172-5173, 5184-5189, or 5192-5194.

56. The method of any one of claims 49-55, wherein said cell is a eukaryotic cell, T-cell, hematopoietic stem cell, or precursor thereof.

57. A method of disrupting a mouse TRBC1 or TRBC2 locus in a cell, comprising introducing to said cell:
(a) a class 2, type V Cas endonuclease; and
(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said mouse TRBC1 or TRBC2 locus,

wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5211-5225 or 5247-5267; or

wherein said engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5196-5210 or 5226-5246.

58. The method of claim 57, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

59. The method of claim 58, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

60. The method of any one of claims 57-59, wherein said engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-

3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3677-3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036.

61. The method of any one of claims 57-60, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

62. The method of any one of claims 57-61, wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 5211, 5213-5215, 5217, 5221, 5223, 5247, 5249-5250, 5252-5253, 5258-5259, or 5264.

63. The method of any one of claims 57-61, wherein said engineered guide RNA has at least 80% sequence identity to any of the guide RNAs from Table 14H that target any one of SEQ ID NOs: 5211, 5213-5215, 5217, 5221, 5223, 5247, 5249-5250, 5252-5253, 5258-5259, or 5264.

64. The method of any one of claims 57-63, wherein said cell is a eukaryotic cell, T-cell, hematopoietic stem cell, or precursor thereof.

65. A method of disrupting a human TRBC1 or TRBC2 locus in a cell, comprising introducing to said cell:
(a) a class 2, type V Cas endonuclease; and
(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said human TRBC1 or TRBC2 locus,

wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5661-5679; or

wherein said engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at

least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5642-5660.

66. The method of claim 65, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

67. The method of claim 66, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

68. The method of any one of claims 65-67, wherein said engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036.

69. The method of any one of claims 65-68, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

70. The method of any one of claims 65-69, wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 5661-5663, 5672-5675, or 5678.

71. The method of any one of claims 65-69, wherein said engineered guide RNA has at least 80% sequence identity to any of the guide RNAs from Table 14I that target any one of SEQ ID NOs: 5661-5663, 5672-5675, or 5678.

72. The method of any one of claims 65-71, wherein said cell is a eukaryotic cell, T-cell, hematopoietic stem cell, or precursor thereof.

73. A method of disrupting an HPRT locus in a cell, comprising introducing to said cell:
(a) a class 2, type V Cas endonuclease; and
(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said HPRT locus,

wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5562-5641; or

wherein said engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5482-5561.

74. The method of claim 73, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

75. The method of claim 74, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

76. The method of any one of claims 73-75, wherein said engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036.

77. The method of any one of claims 73-76, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

78. The method of any one of claims 73-77, wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 5562-5564 or 5568.

79. The method of any one of claims 73-77, wherein said engineered guide RNA has at least 80% sequence identity to any of the guide RNAs from Table 14J that target any one of SEQ ID NOs: 5562-5564 or 5568.

80. The method of any one of claims 73-80, wherein said cell is a eukaryotic cell, hepatocyte, T-cell, hematopoietic stem cell, or precursor thereof.

81. A method of disrupting an APO-A1 locus in a cell, comprising introducing to said cell:
(a) a class 2, type V Cas endonuclease; and
(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said APO-A1 locus,

wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5861-5874; or

wherein said engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5847-5860.

82. The method of claim 81, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

83. The method of claim 82, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

84. The method of any one of claims 81-83, wherein said engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036.

85. The method of any one of claims 81-84, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

86. The method of any one of claims 81-85, wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 5861-5866 or 5868-5869.

87. The method of any one of claims 81-85, wherein said engineered guide RNA has at least 80% sequence identity to any of the guide RNAs from Table 43A that target any one of SEQ ID NOs: 5861-5866 or 5868-5869.

88. The method of any one of claims 81-87, wherein said cell is a eukaryotic cell, hepatocyte, T-cell, hematopoietic stem cell, or precursor thereof.

89. A method of disrupting an ANGPTL3 locus in a cell, comprising introducing to said cell:
(a) a class 2, type V Cas endonuclease; and

(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said ANGPTL3 locus,

wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5953-6030; or

wherein said engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5875-5952.

90. The method of claim 89, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

91. The method of claim 90, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

92. The method of any one of claims 89-91, wherein said engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036.

93. The method of any one of claims 89-92, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

94. The method of any one of claims 89-93, wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 5955-5963, 5968-5975, 5979-5987, 5989-5993, 5997, 5999, 6003-6010, 6014-6016, 6024-6025, or 6027-6030.

95. The method of any one of claims 89-93, wherein said engineered guide RNA has at least 80% sequence identity to any of the guide RNAs from Table 43B that target any one of SEQ ID NOs: 5955-5963, 5968-5975, 5979-5987, 5989-5993, 5997, 5999, 6003-6010, 6014-6016, 6024-6025, or 6027-6030.

96. The method of any one of claims 89-95, wherein said cell is a eukaryotic cell, hepatocyte, T-cell, hematopoietic stem cell, or precursor thereof.

97. A method of disrupting a human Rosa26 locus in a cell, comprising introducing to said cell:
(a) a class 2, type V Cas endonuclease; and
(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said human Rosa26 locus,

wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5013-5055; or

wherein said engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 4970-5012.

98. The method of claim 97, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

99. The method of claim 98, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

100. The method of any one of claims 97-99, wherein said engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036.

101. The method of any one of claims 97-100, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

102. The method of any one of claims 97-101, wherein said cell is a eukaryotic cell, hepatocyte, T-cell, hematopoietic stem cell, or precursor thereof.

103. A method of disrupting a FAS locus in a cell, comprising introducing to said cell:
(a) a class 2, type V Cas endonuclease; and
(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said FAS locus,

wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5367-5465; or

wherein said engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about

97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5268-5366.

104. The method of claim 103, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

105. The method of claim 104, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

106. The method of any one of claims 103-105, wherein said engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036.

107. The method of any one of claims 103-106, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

108. The method of any one of claims 103-107, wherein said cell is a eukaryotic cell, hepatocyte, T-cell, hematopoietic stem cell, or precursor thereof.

109. A method of disrupting a PD-1 locus in a cell, comprising introducing to said cell:
(a) a class 2, type V Cas endonuclease; and
(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said PD-1 locus,

wherein said engineered guide RNA is configured to hybridize to a sequence having at least 20-22 consecutive nucleotides complementary to a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at

least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5474-5481; or

wherein said engineered guide RNA comprises a nucleotide sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 5466-5473.

110. The method of claim 109, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

111. The method of claim 110, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

112. The method of any one of claims 109-111, wherein said engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036.

113. The method of any one of claims 109-112, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

114. The method of any one of claims 109-113, wherein said cell is a eukaryotic cell, hepatocyte, T-cell, hematopoietic stem cell, or precursor thereof.

115. An engineered nuclease system comprising:

(a) an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 215 or a variant thereof; and

(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence,
wherein said system has reduced immunogenicity when administered to a human subject compared to an equivalent system comprising a Cas9 enzyme.

116. The system of claim 115, wherein said Cas9 enzyme is an SpCas9 enzyme.

117. The system of claim 115-116, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

118. The system of any one of claims 115-117, wherein said immunogenicity is antibody immunogenicity.

119. A method of disrupting a mouse HAO-1 locus in a cell, comprising introducing to said cell:
(a) a class 2, type V Cas endonuclease; and
(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said mouse HAO-1 locus,

wherein said engineered guide RNA comprises the nucleotides of guide RNAs mH29-1_37, mH29-15_37, mH29-29_37 from Table 25 comprising the nucleotide modifications described in Table 25;

or wherein said engineered guide RNA comprises any one of SEQ ID NOs: 4184-4225.

120. The method of claim 119, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

121. The method of claim 120, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

122. The method of any one of claims 119-121, wherein said engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least

about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036.

123. The method of any one of claims 119-122, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

124. The method of any one of claims 119-123, wherein said cell is a eukaryotic cell, hepatocyte, T-cell, hematopoietic stem cell, or precursor thereof.

125. The method of any one of claims 119-124, wherein said engineered guide RNA comprises the nucleotides of guide RNAs mH29-15_37 or mH29-29_37 from Table 25 comprising the nucleotide modifications described in Table 25.

126. The method of any one of claims 119-125, wherein said method further comprises disrupting expression of glycolate oxidase from said HAO-1 locus.

127. A method of disrupting a human TRAC locus in a cell, comprising introducing to said cell:
(a) a class 2, type V Cas endonuclease; and
(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said human TRAC locus,

wherein said engineered guide RNA comprises the nucleotides of MG29-1-TRAC-sgRNA-35 from Table 28B comprising the nucleotide modifications described in Table 28B.

128. The method of claim 127, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

129. The method of claim 128, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

130. The method of any one of claims 127-129, wherein said engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036.

131. The method of any one of claims 127-130, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

132. The method of any one of claims 127-131, wherein said cell is a eukaryotic cell, hepatocyte, T-cell, hematopoietic stem cell, or precursor thereof.

133. A method of disrupting an albumin locus in a cell, comprising introducing to said cell:
(a) a class 2, type V Cas endonuclease; and
(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a region of said albumin locus,

Wherein said engineered guide RNA comprises the nucleotides of mAlb298-37, mAlb2912-37, mAlb2918-37, or mAlb298-34 from Table 29 comprising the nucleotide modifications described in Table 29; or

wherein said engineered guide RNA comprises the nucleotides of mAlb29-8-44, mAlb29-8-50, mAlb29-8-50b, mAlb29-8-51b, mAlb29-8-52b, mAlb29-8-53b, or mAlb29-8-54b comprising the nucleotide modifications described in Table 46.

134. The method of claim 133, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof.

135. The method of claim 134, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

136. The method of any one of claims 133-135, wherein said engineered guide RNA comprises a sequence with at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the non-degenerate nucleotides of any one of SEQ ID NOs: 3471, 3539, 3551-3559, 3608-3609, 3612, 3636-3637, 3640-3641, 3644-3645, 3648-3649, 3652-3653, 3656-3657, 3660-3661, 3664-3667, 3671-3672, 3678, 3695-3696, 3729-3730, 3734-3735, 3851-3857, or 6033-6036.

137. The method of any one of claims 133-136, wherein said guide RNA comprises a sequence with at least 80% sequence identity to the non-degenerate nucleotides of SEQ ID NO: 3609.

138. The method of any one of claims 133-137, wherein said cell is a eukaryotic cell, hepatocyte, T-cell, hematopoietic stem cell, or precursor thereof.

139. The method of any one of claims 133-138, wherein said engineered guide RNA comprises the nucleotides of mAlb298-37, mAlb2912-37, mAlb2918-37, or mAlb298-34 from Table 29 comprising the nucleotide modifications described in Table 29.

140. An engineered guide RNA comprising:

a) a DNA-targeting segment comprising a nucleotide sequence that is complementary to a target sequence in a target DNA molecule; and

b) a protein-binding segment configured to bind to a class 2, type V Cas endonuclease, and

wherein said guide RNA comprises a nucleotide modification pattern depicted in any one of SEQ ID NOs: 5695-5701 in Table 34.

141. The engineered guide RNA of claim 140, wherein said guide RNA comprises mAlb29-8-44, mAlb29-8-50, mAlb29-8-37, or mAlb29-12-44.

142. The engineered guide RNA of claim 140, wherein said guide RNA comprises hH29-4_50, hH29-21_50, hH29-23_50, hH29-41_50, hH29-4_50b, hH29-21_50b, hH29-23_50b, or hH29-41_50b, mH29-1-50, mH29-15-50, mH29-29-50, mH29-1-50b, mH29-15-50b, or mH29-29-50b.

143. The engineered guide RNA of claim 140, wherein said DNA-targeting segment is configured to hybridize to an HAO-1 gene or an albumin gene.

144. The engineered guide RNA of any one of claims 140-142, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

145. The engineered guide RNA of any one of claims 140-144, wherein said class 2, type V Cas endonuclease comprises an endonuclease having at least 75% sequence identity to SEQ ID NO: 215.

146. An engineered nuclease system comprising:

(a) an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 1-3470 or a variant thereof, or a nucleotide sequence encoding said enonuclease; and

(b) a polynucleotide sequence encoding a CRISPR array, wherein said CRISPR array is configured to be processed by said endonuclease to an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence,

wherein said spacer sequence is configured to hybridize to an albumin gene.

147. The system of claim 146, wheren said polynucleotide sequence comprises a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to SEQ ID NO: 5712.

148. The system of claim 146 or 147, wherein said endonuclease comprises an endonuclease having at least 75% sequence identity to any one of SEQ ID NOs: 141, 215, 229, 261, or 1711-1722 or a variant thereof.

149. The system of claim 148, wherein said endonuclease comprises an endonuclease having at least 75% sequence identity to SEQ ID NO: 215.

150. An engineered nuclease system comprising:

(a) an endonuclease having at least 75% sequence identity to SEQ ID NOs: 470 or a variant thereof; and

(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence.

151. The engineered nuclease system of claim 150, wherein said engineered guide RNA comprises a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to SEQ ID NO: 6031.

152. The engineered nuclease system of claim 151 or 152, wherein said endonuclease is configured to be selective for a 5' PAM sequence comprising SEQ ID NO: 6032.

153. An engineered nuclease system comprising:

(a) an endonuclease having at least at least about 80%, at least about 85%, at least about 86%, at least about 87%, at least about 88%, at least about 89%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs: 2824, 2841, or 2896, or a variant thereof; and

(b) an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence,

wherein said engineered guide RNA comprises a sequence having at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to any one of SEQ ID NOs:6033, 6034, or 6035.

154. The engineered nuclease system of claim 153, wherein said endonuclease has at least 80% sequence identity to SEQ ID NO: 2824 and said engineered guide RNA has at least 80% sequence identity to SEQ ID NO: 6033.

155. The engineered nuclease system of claim 153, wherein said endonuclease has at least 80% sequence identity to SEQ ID NO: 2841 and said engineered guide RNA has at least 80% sequence identity to SEQ ID NO: 6034.

156. The engineered nuclease system of claim 153, wherein said endonuclease has at least 80% sequence identity to SEQ ID NO: 2896 and said engineered guide RNA has at least 80% sequence identity to SEQ ID NO: 6035.

157. The engineered nuclease system of any one of claims 153-156, wherein said endonuclease is configured to be selective for a 5' PAM sequence comprising any one of SEQ ID NOs: 6037-6039.

158. A lipid nanoparticle comprising:

(a) any of the endonucleases described herein;

(b) any of the engineered guide RNAs described herein:

(c) a cationic lipid;

(d) a sterol;

(e) a neutral lipid; and

(f) a PEG-modified lipid.

159. The lipid nanoparticle of claim 158, wherein said cationic lipid comprises C12-200, said sterol comprises cholesterol, said neutral lipid comprises DOPE, or said PEG-modified lipid comprises DMG-PEG2000.

160. The lipid nanoparticle of claim 158, wherein said cationic lipid comprises any of the cationic lipids depicted in FIG. 109.

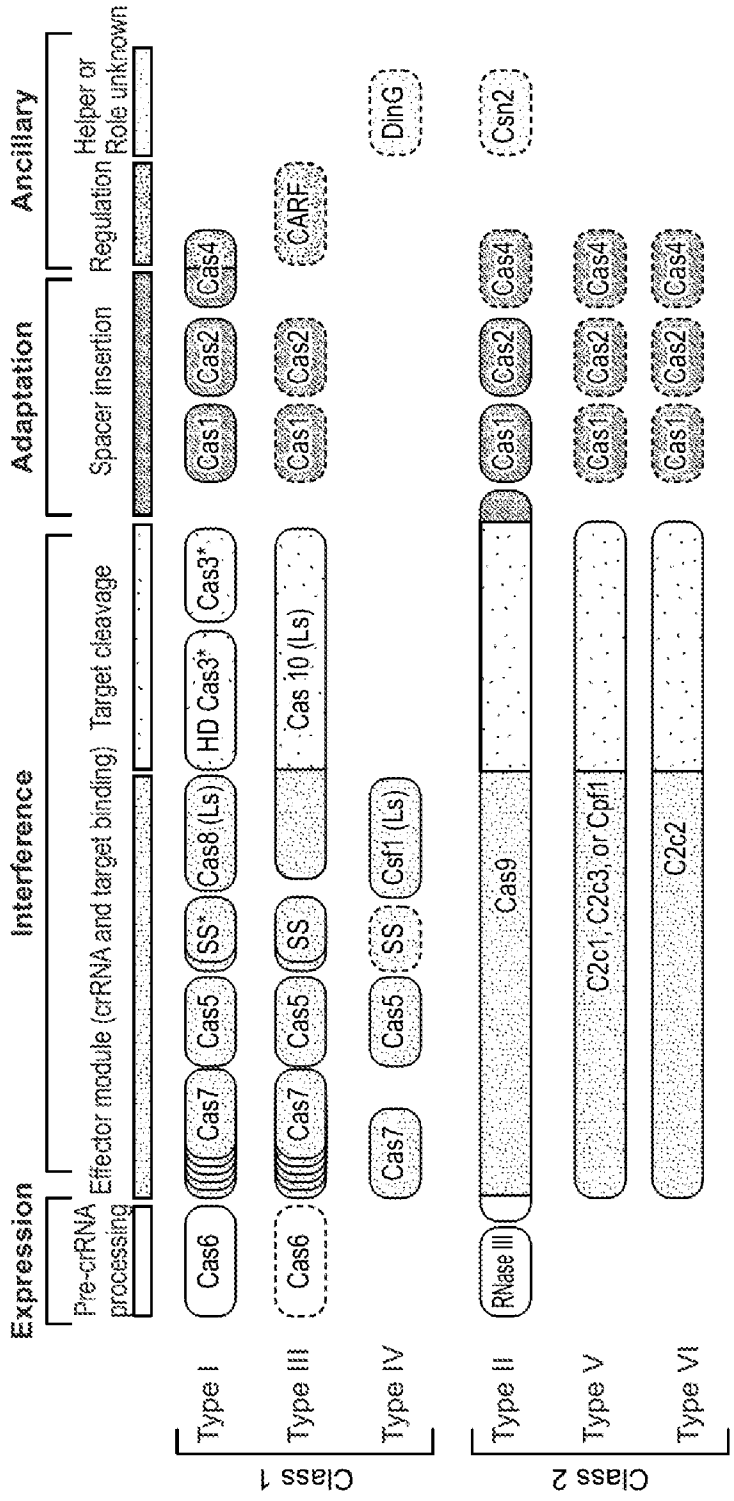

Expression
Interference
Adaptation
Ancillary
Pre-crRNA processing
Effector module (crRNA and target binding)
Target cleavage
Spacer insertion
Regulation
Helper or Role unknown
Class 1
Type I
Cas6
Cas7
Cas5
SS*
Cas8 (Ls)
HD Cas3*
Cas3*
Cas1
Cas2
Cas4
Type III
Cas6
Cas7
Cas5
SS
Cas 10 (Ls)
Cas1
Cas2
CARF
Type IV
Cas7
Cas5
SS
Csf1 (Ls)
DinG
Class 2
Type II
RNase III
Cas9
Cas1
Cas2
Cas4
Csn2
Type V
C2c1, C2c3, or Cpf1
Cas1
Cas2
Cas4
Type VI
C2c2
Cas1
Cas2
Cas4

FIG. 1

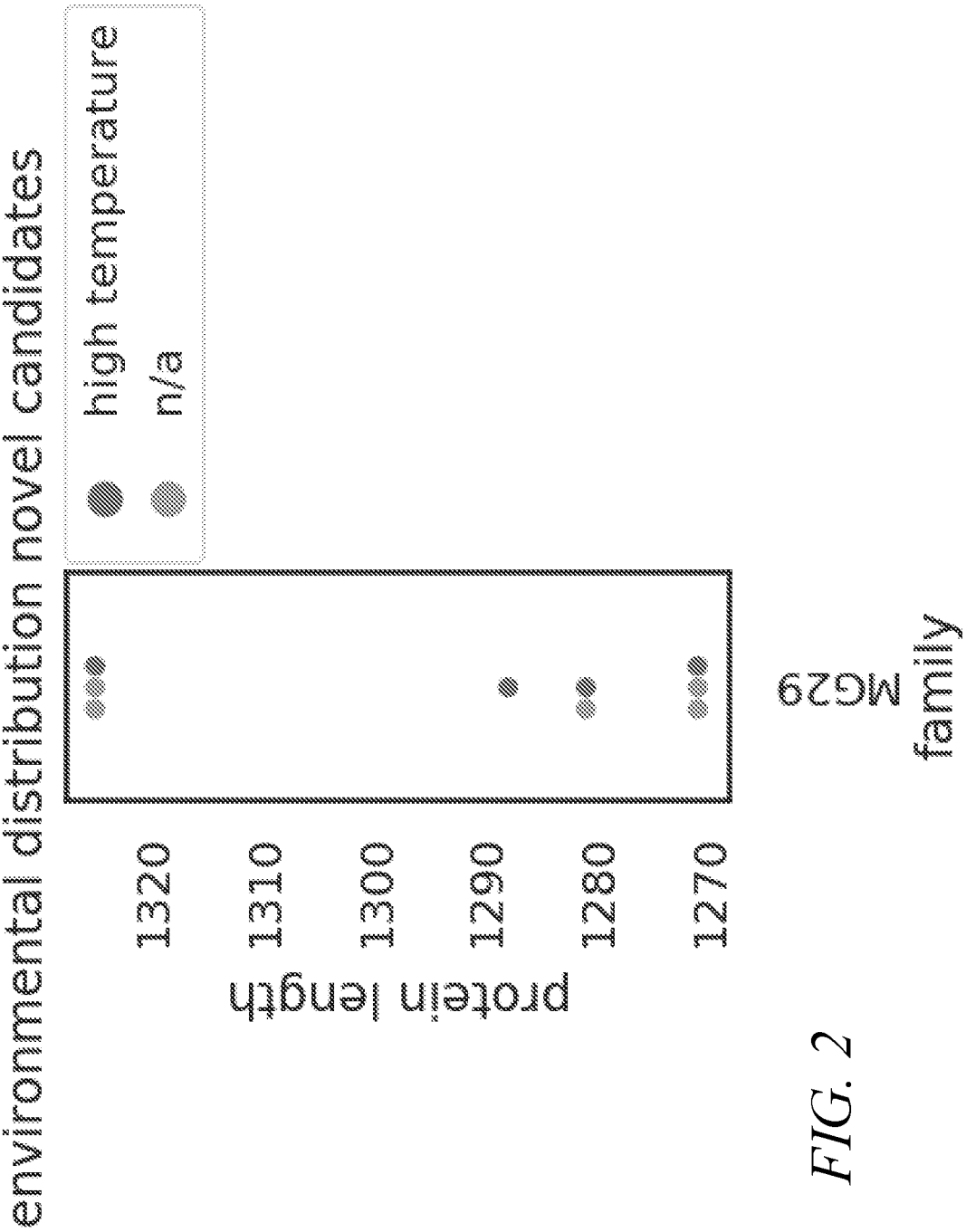

environmental distribution novel candidates
high temperature
n/a
protein length
1320
1310
1300
1290
1280
1270
MG29
family

FIG. 2

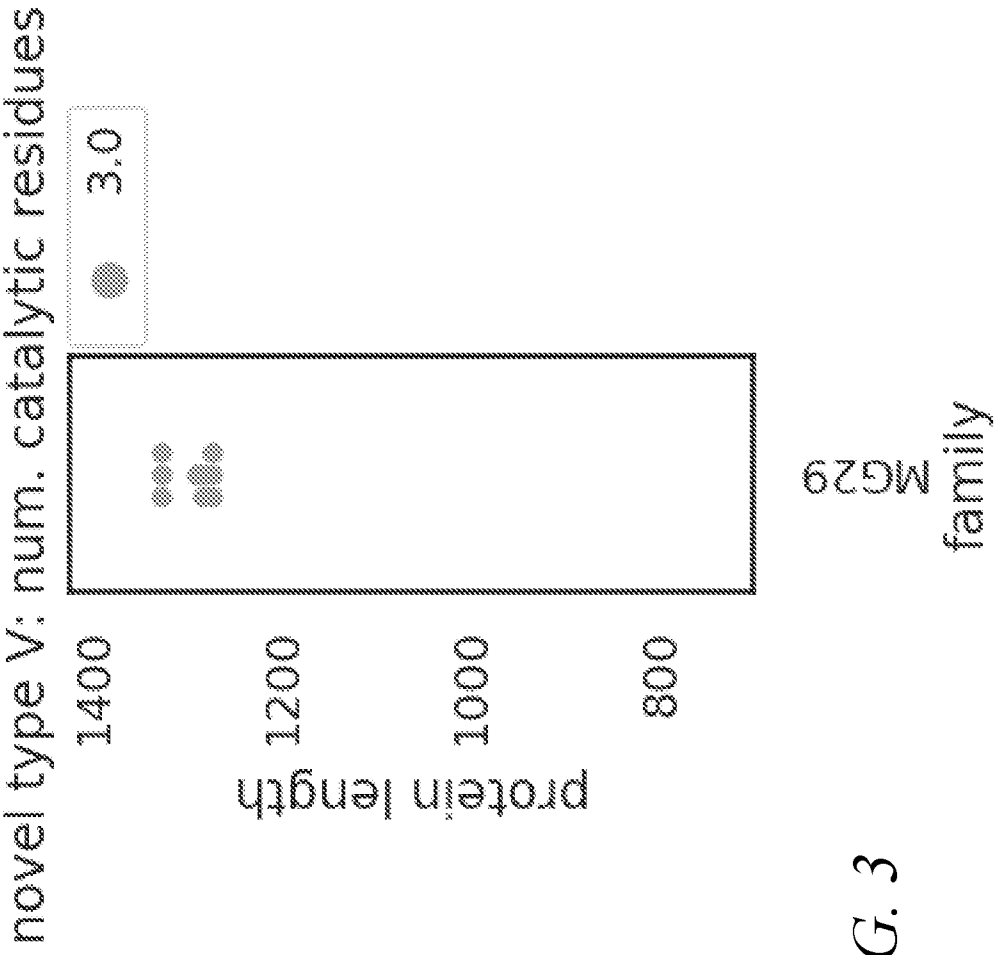

novel type V: num. catalytic residues
3.0
protein length
1400
1200
1000
800
MG29
family

FIG. 3

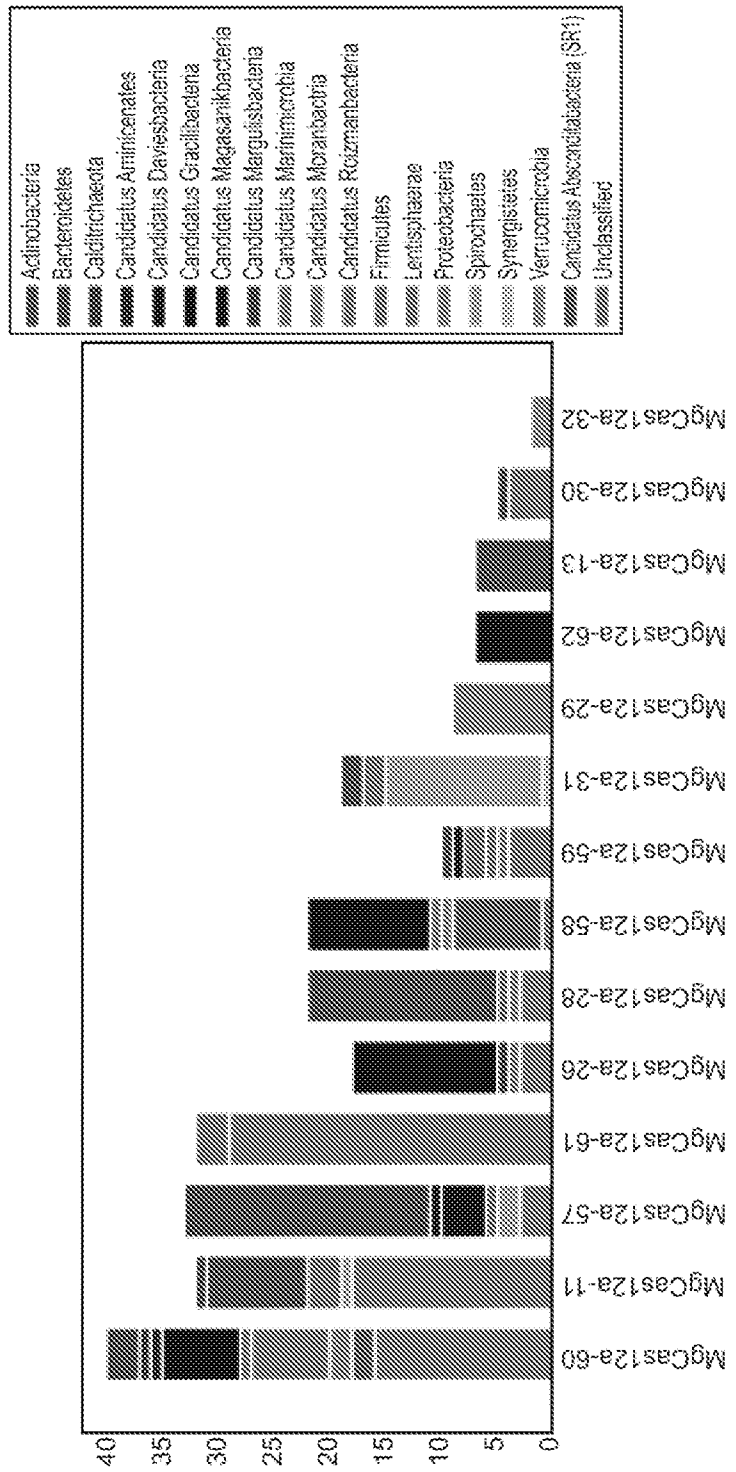

Actinobacteria
Bacteroidetes
Calditrichaeota
Candidatus Aminicenates
Candidatus Daviesbacteria
Candidatus Gracilibacteria
Candidatus Magasanikbacteria
Candidatus Marguisbacteria
Candidatus Marinimicrobia
Candidatus Moranbacteria
Candidatus Roizmanbacteria
Firmicutes
Lentisphaerae
Proteobacteria
Spirochaetes
Synergistetes
Verrucomicrobia
Candidatus Absconditabacteria (SR1)
Unclassified
MgCas12a-60
MgCas12a-11
MgCas12a-57
MgCas12a-61
MgCas12a-26
MgCas12a-28
MgCas12a-58
MgCas12a-59
MgCas12a-31
MgCas12a-29
MgCas12a-62
MgCas12a-13
MgCas12a-30
MgCas12a-32
0
5
10
15
20
25
30
35
40

FIG. 4A

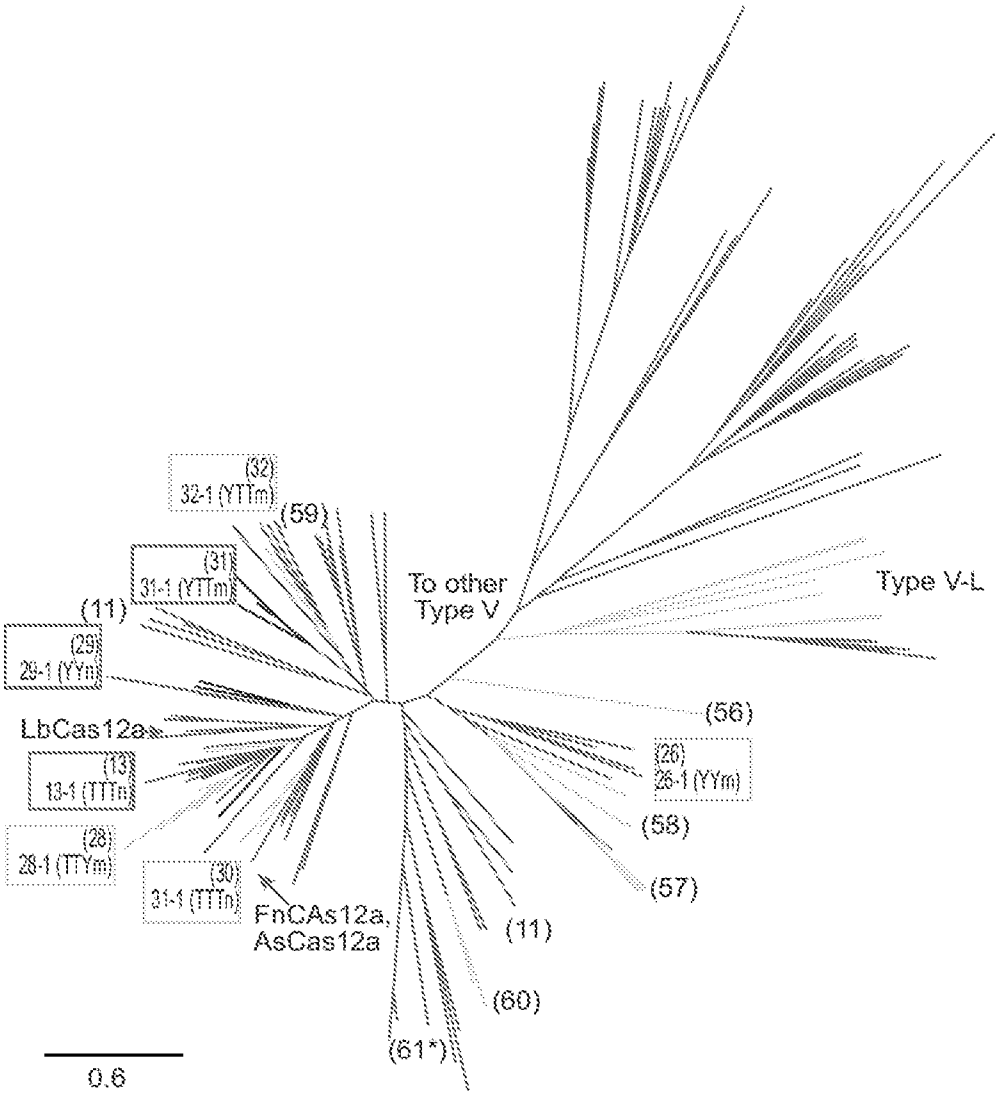

(32)
32-1 (YTTm)
(59)
(31)
31-1 (YTTm)
(11)
(29)
29-1 (YYm)
LbCas12a
(13)
13-1 (TTTn)
(28)
28-1 (TTYm)
(30)
31-1 (TTTn)
FnCAs12a,
AsCas12a
To other
Type V
Type V-L
(56)
(26)
26-1 (YYm)
(58)
(57)
(11)
(60)
(61*)
0.6

FIG. 4B

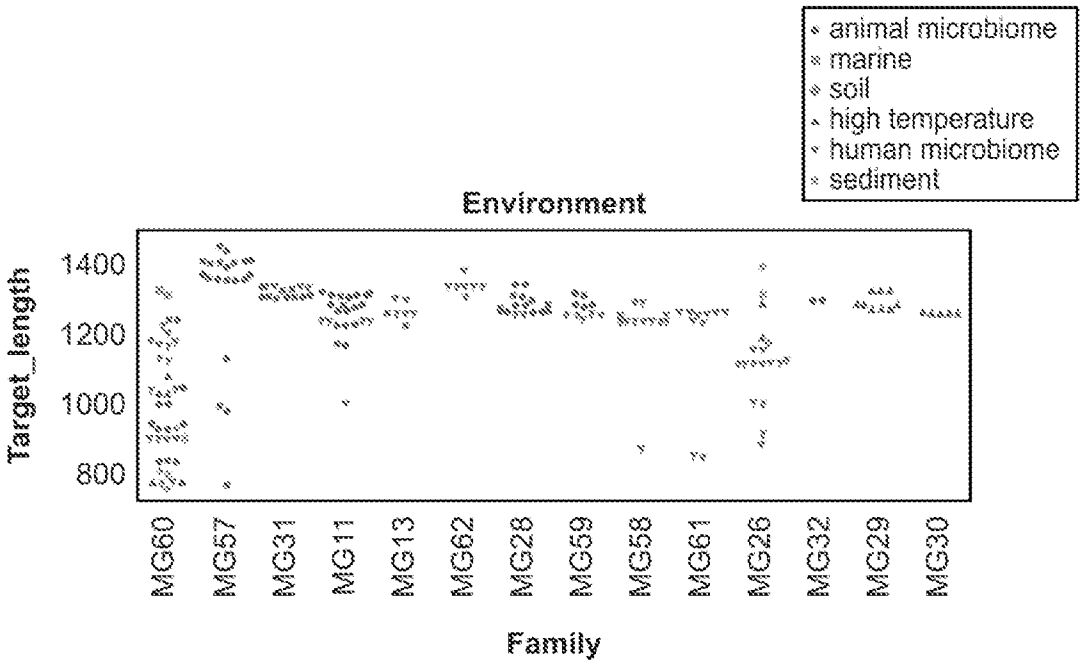

animal microbiome
marine
soil
high temperature
human microbiome
sediment
Environment
Target_length
1400
1200
1000
800
MG60
MG57
MG31
MG11
MG13
MG62
MG28
MG59
MG58
MG61
MG26
MG32
MG29
MG30
Family

FIG. 5A

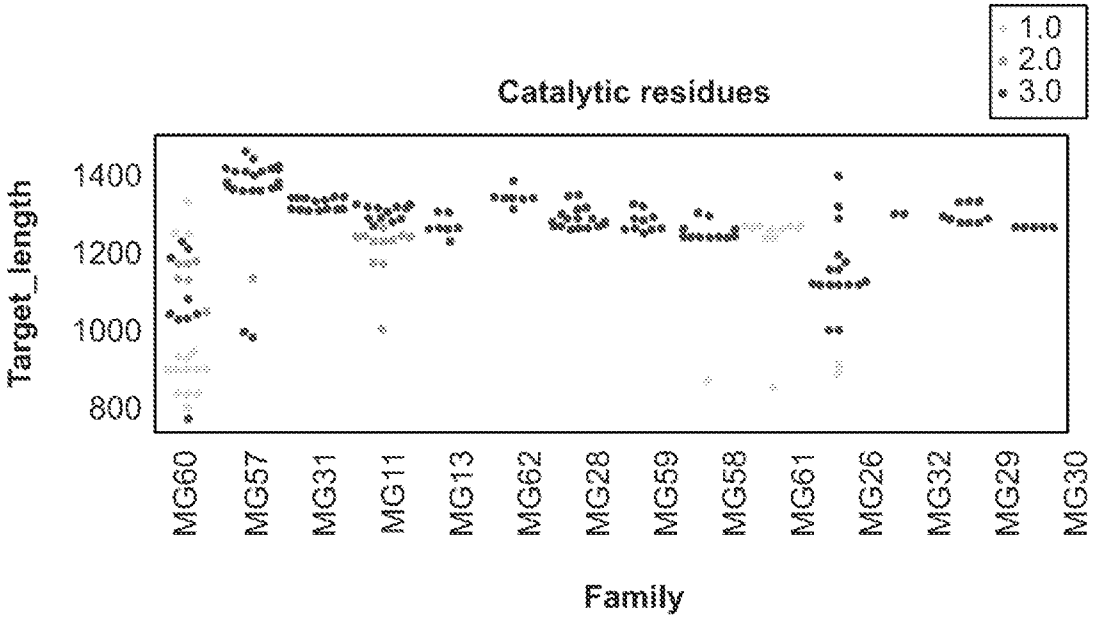

1.0
2.0
3.0
Catalytic residues
Target_length
1400
1200
1000
800
MG60
MG57
MG31
MG11
MG13
MG62
MG28
MG59
MG58
MG61
MG26
MG32
MG29
MG30
Family

FIG. 5B

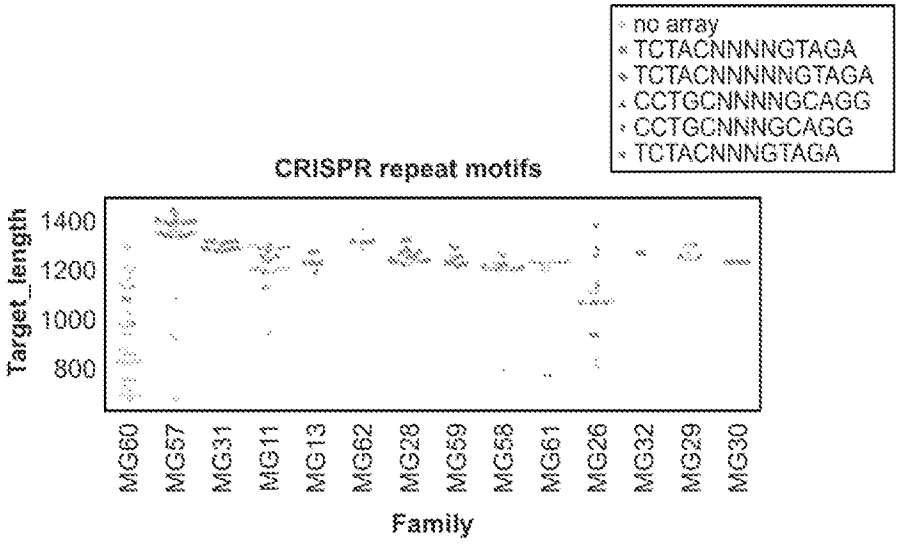

no array
TCTACNNNNGTAGA
TCTACNNNNNGTAGA
CCTGCNNNNGCAGG
CCTGCNNNGCAGG
TCTACNNNGTAGA
CRISPR repeat motifs
Target_length
1400
1200
1000
800
MG60
MG57
MG31
MG11
MG13
MG62
MG28
MG59
MG58
MG61
MG26
MG32
MG29
MG30
Family

FIG. 5C

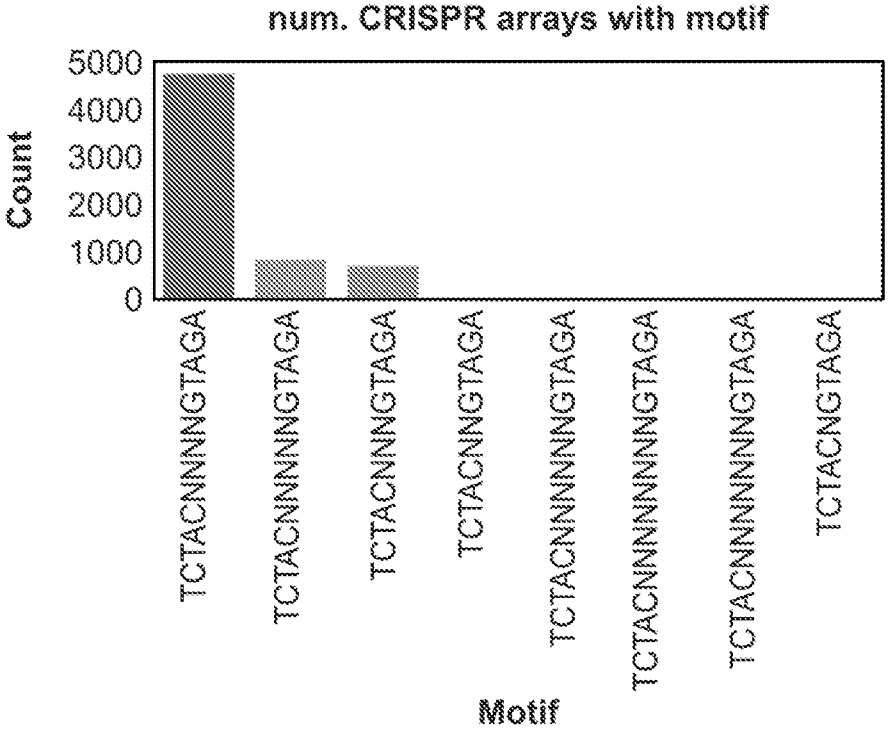

num. CRISPR arrays with motif
Count
5000
4000
3000
2000
1000
0
TCTACNNNNGTAGA
TCTACNNNNNGTAGA
TCTACNNNGTAGA
TCTACNNGTAGA
TCTACNNNNNNGTAGA
TCTACNNNNNNNGTAGA
TCTACNNNNNNNNGTAGA
TCTACNGTAGA
Motif

FIG. 5D

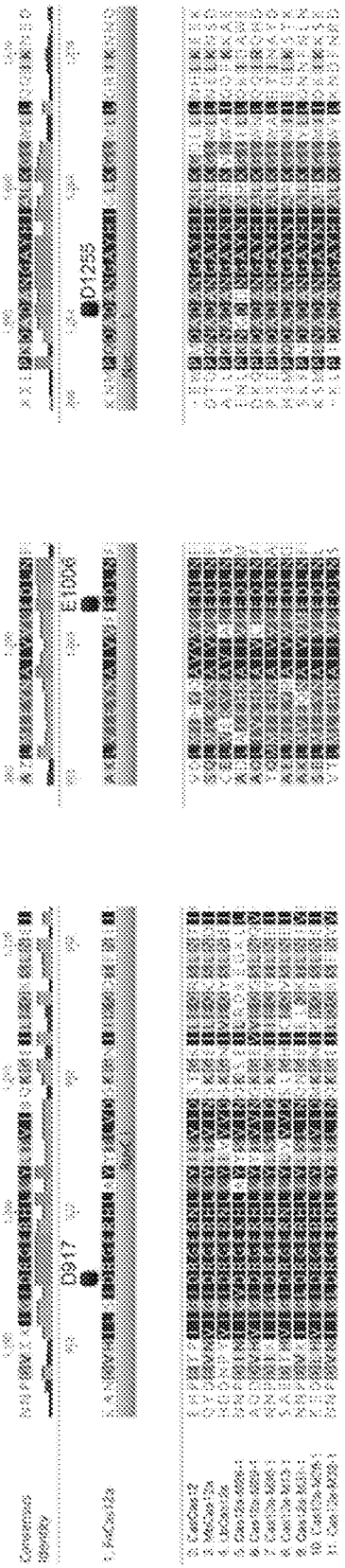


FIG. 6A

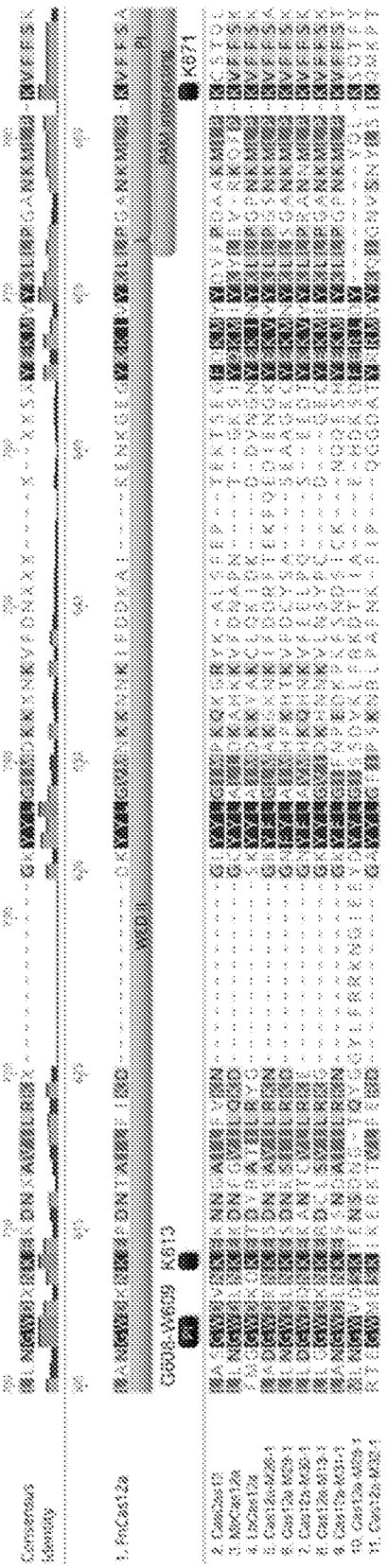


FIG. 6B

FIG. 7A

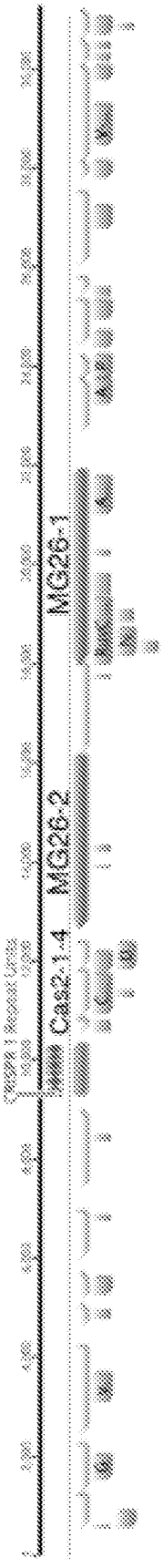


FIG. 7B

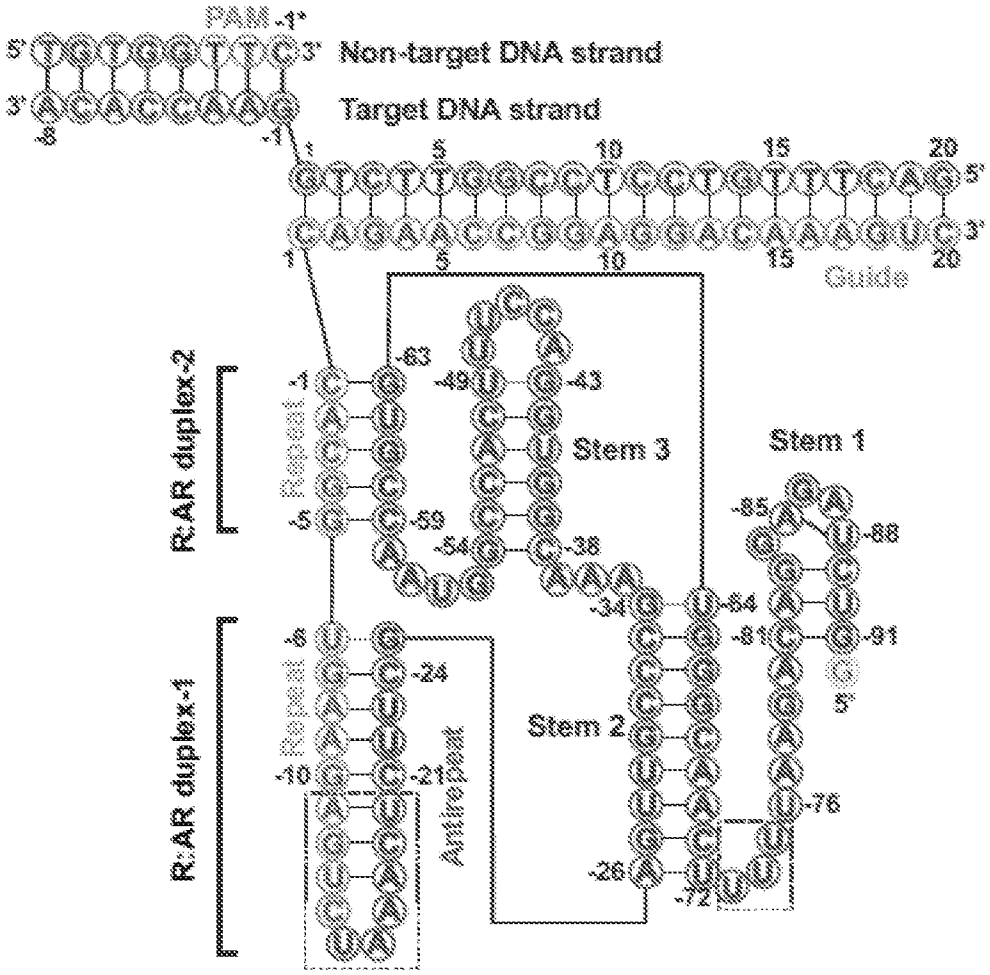

PAM
Non-target DNA strand
Target DNA strand
Guide
R:AR duplex-2
R:AR duplex-1
Repeat 1
Repeat 2
Stem 1
Stem 2
Stem 3
Antirepeat

FIG. 8

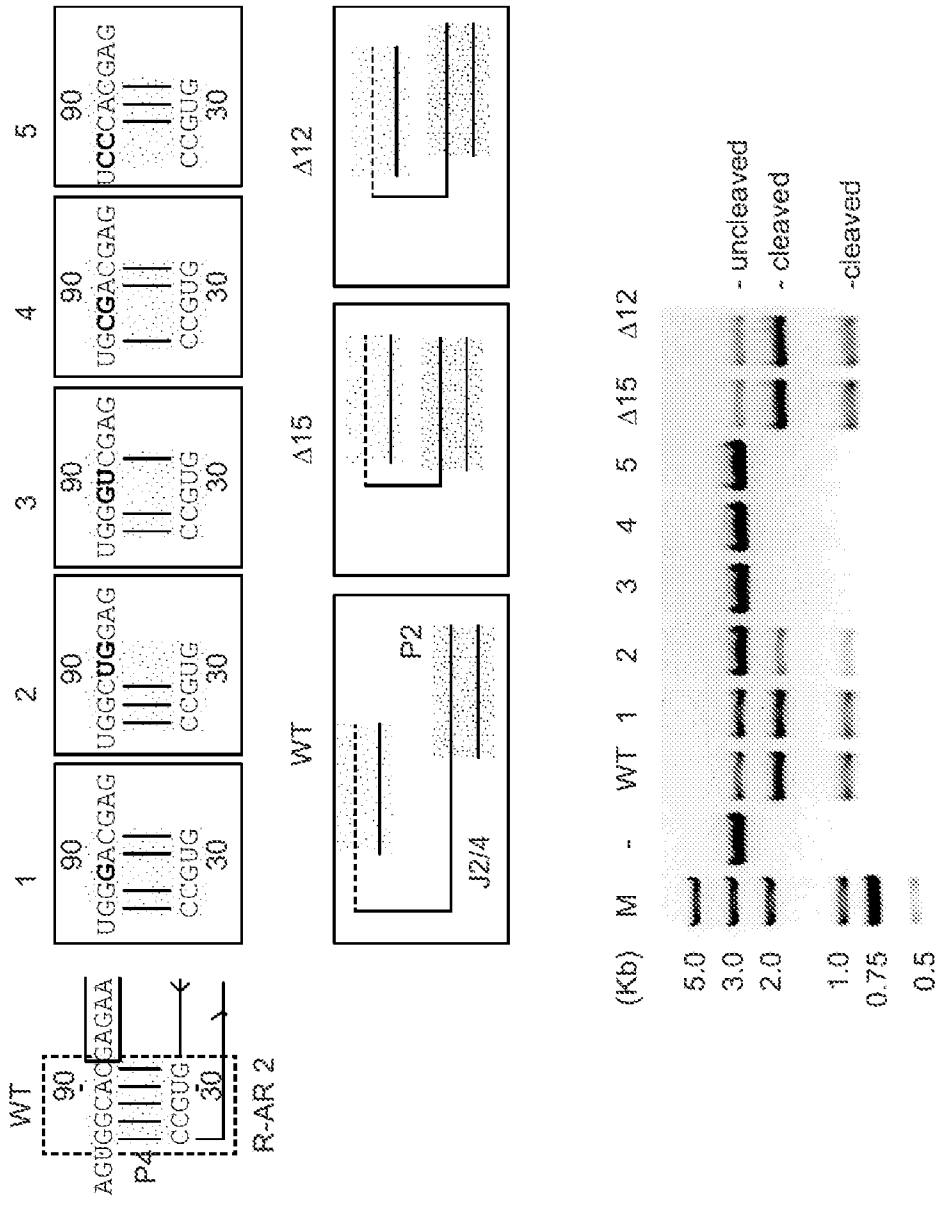

WT
90
AGUGGCACGAGAA
P4
CCGUG
30
R-AR 2
1
2
3
4
5
90
UGGGACGAG
CCGUG
30
90
UGGCUGGAG
CCGUG
30
90
UGGGUCGAG
CCGUG
30
90
UGCGACGAG
CCGUG
30
90
UCCCACGAG
CCGUG
30
WT
Δ15
Δ12
P2
J2/4
(Kb)
M
-
WT
1
2
3
4
5
Δ15
Δ12
5.0
3.0
2.0
1.0
0.75
0.5
- uncleaved
- cleaved
-cleaved

FIG. 9

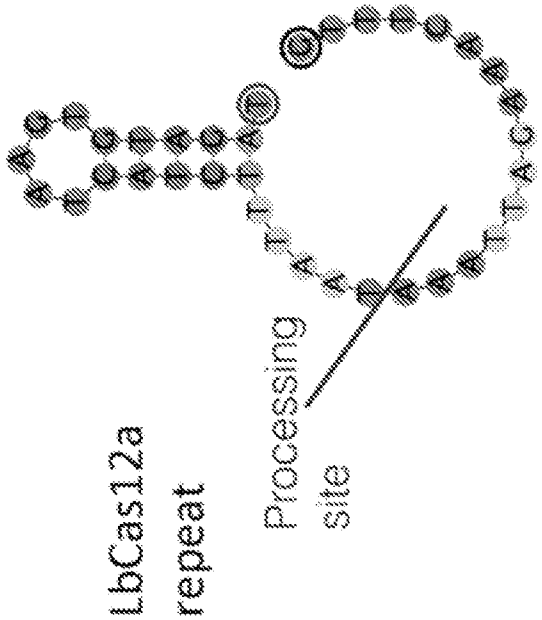


FIG. 10A

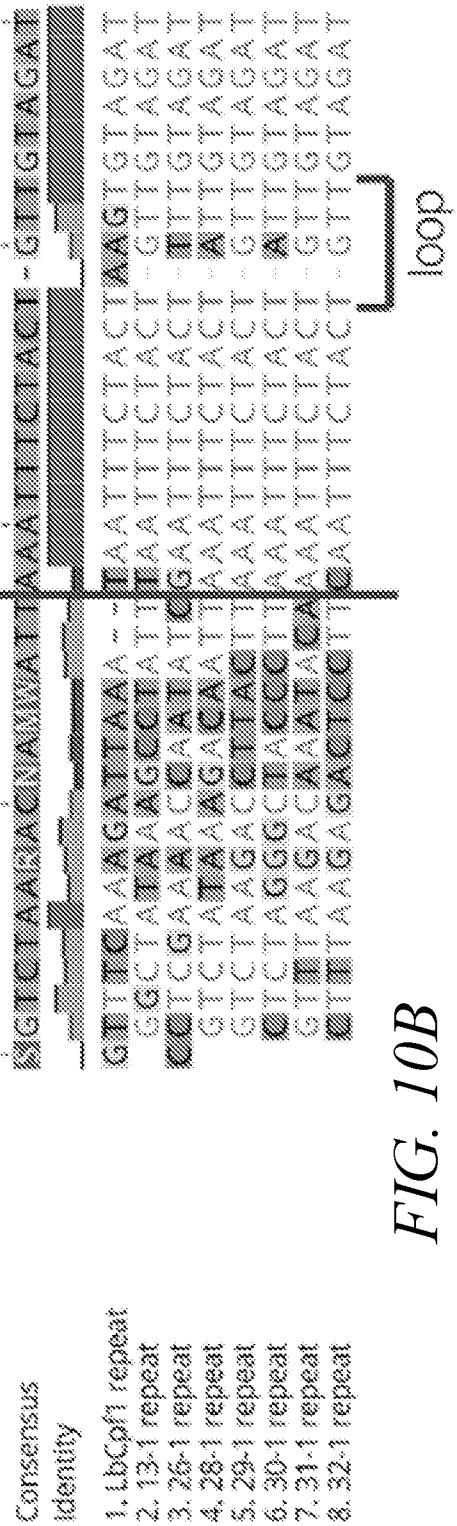


FIG. 10B



FIG. 10C

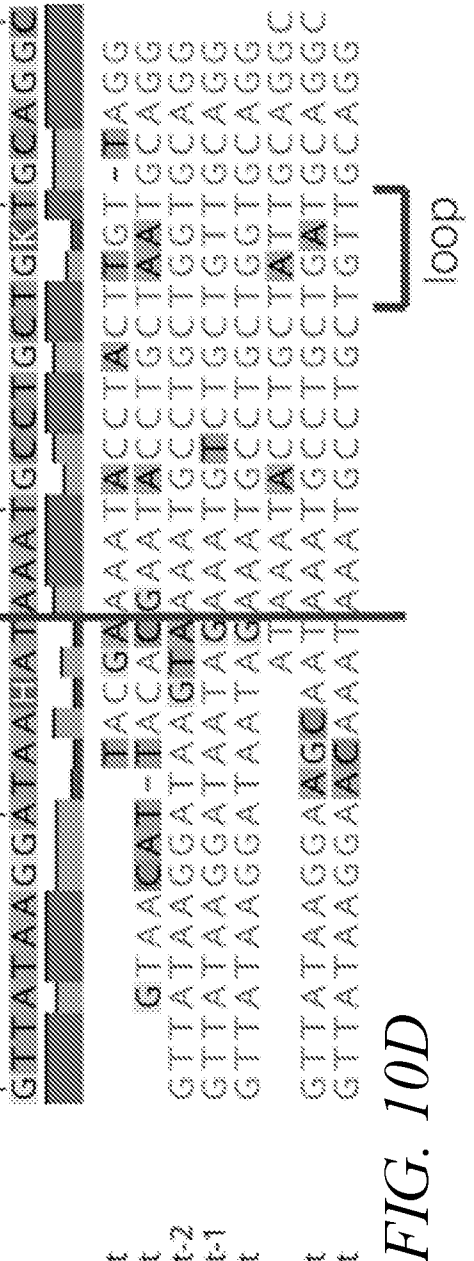


FIG. 10D

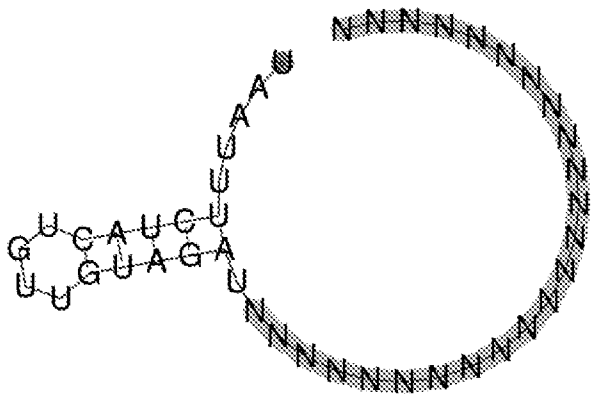


FIG. 11

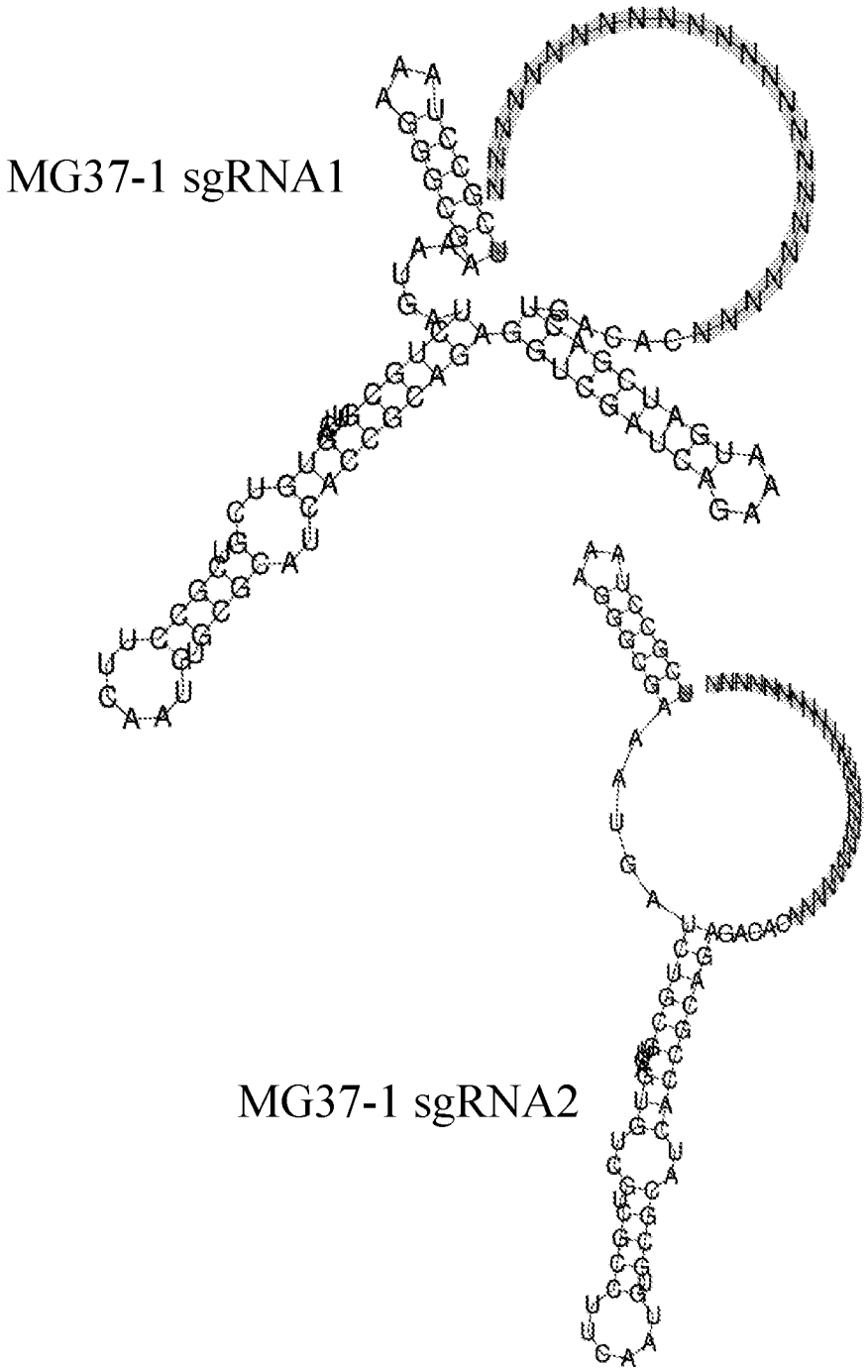
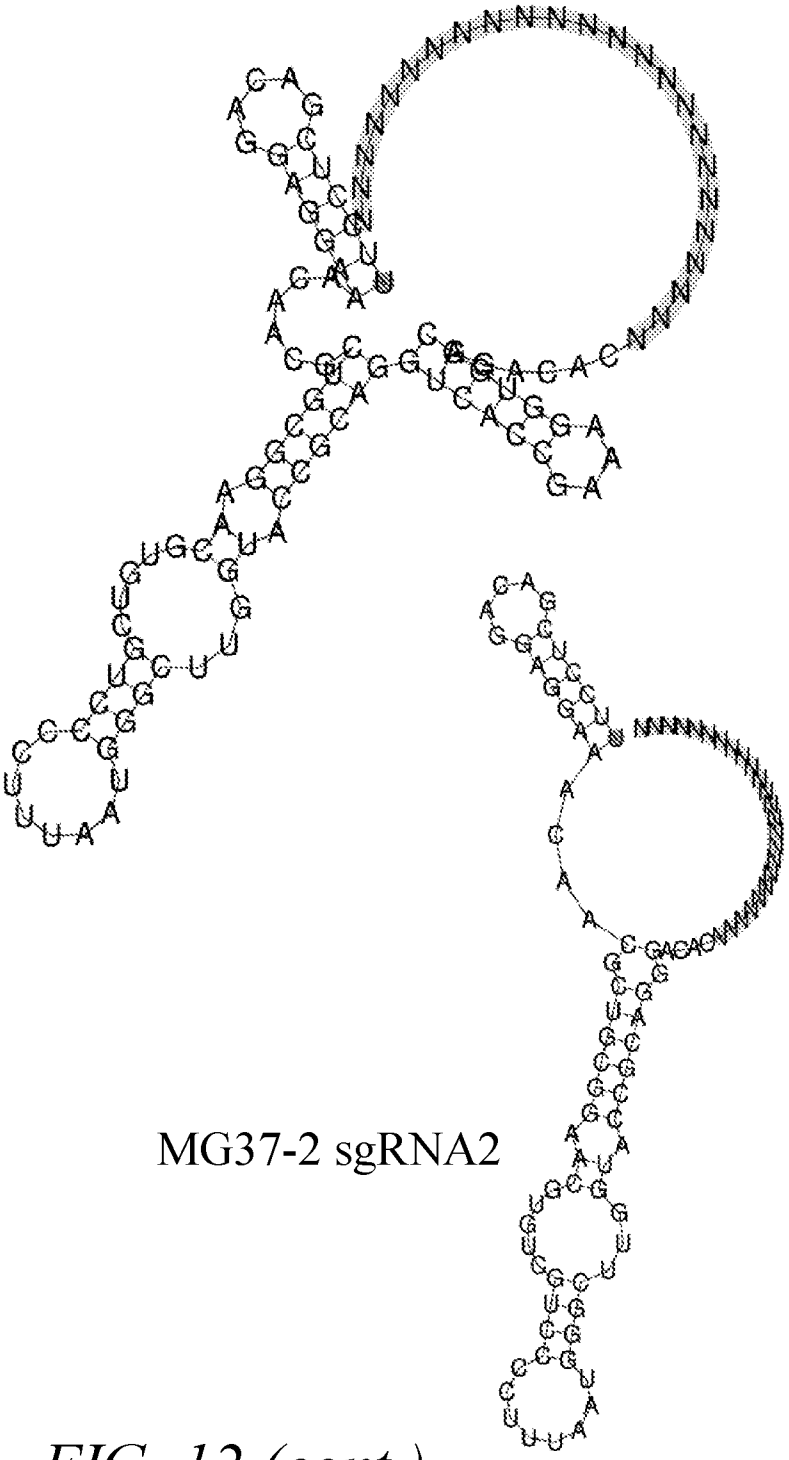


FIG. 12

MG37-2 sgRNA1



MG37-2 sgRNA2

FIG. 12 (cont.)

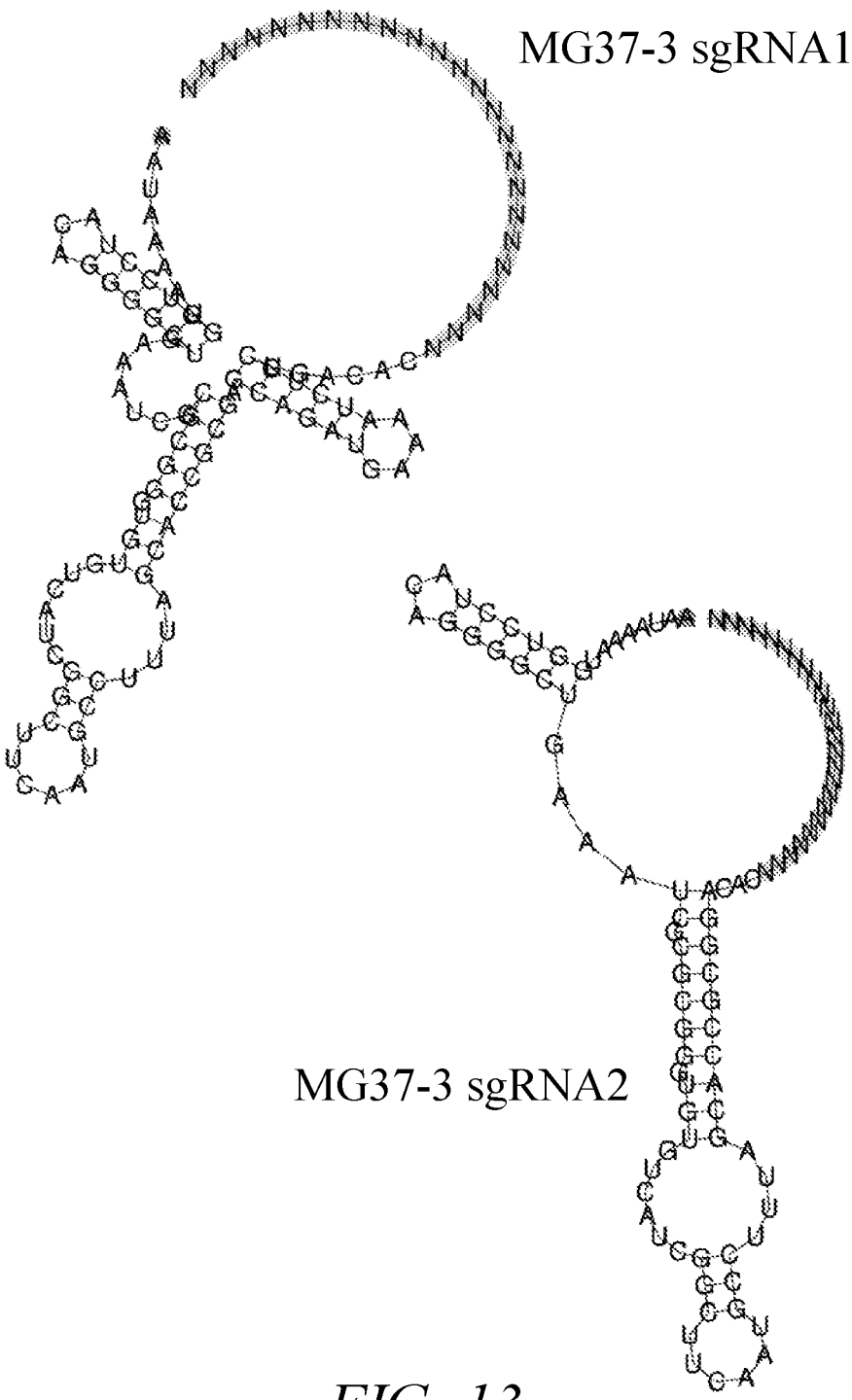


FIG. 13

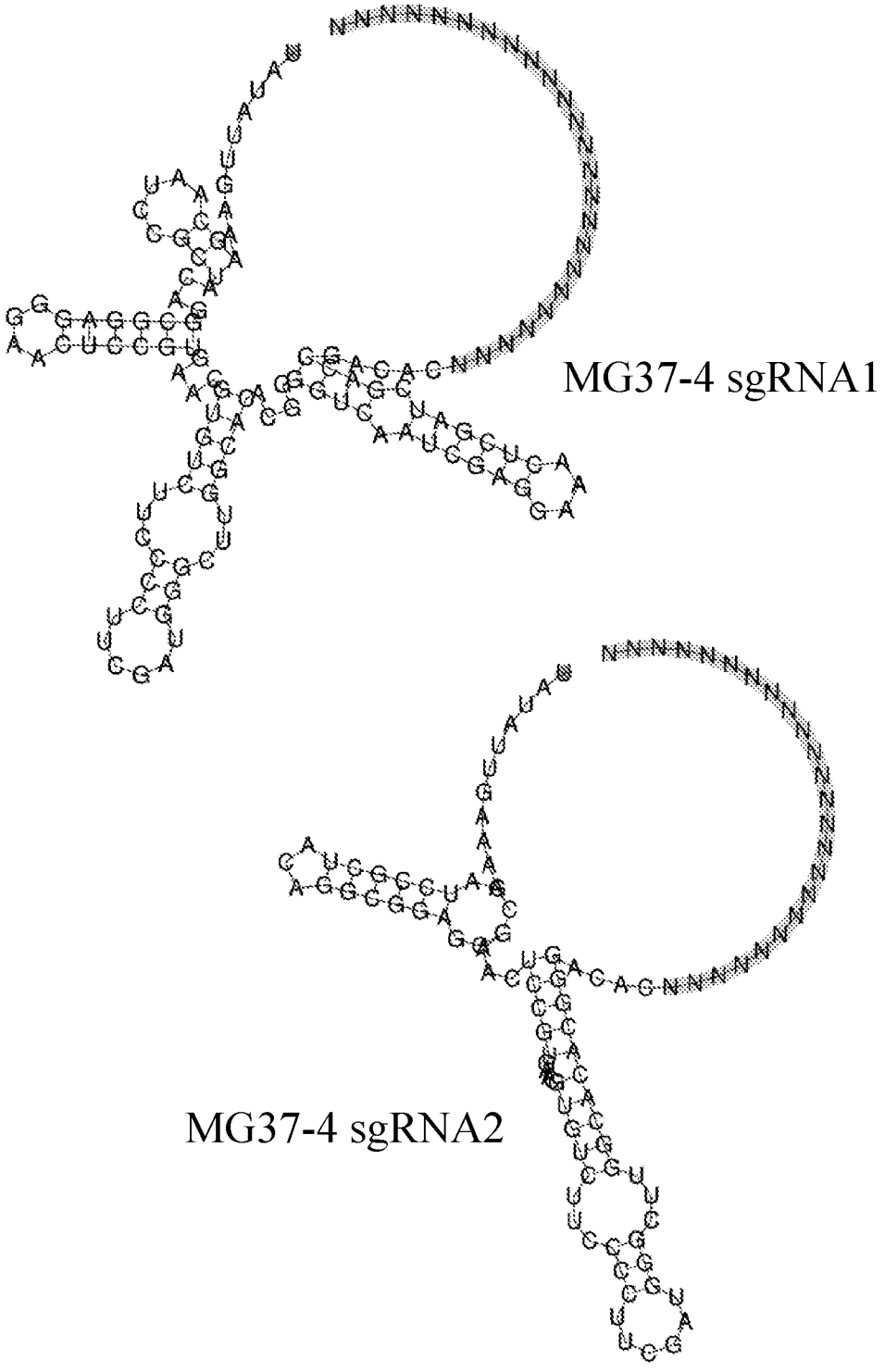


FIG. 13 (cont.)

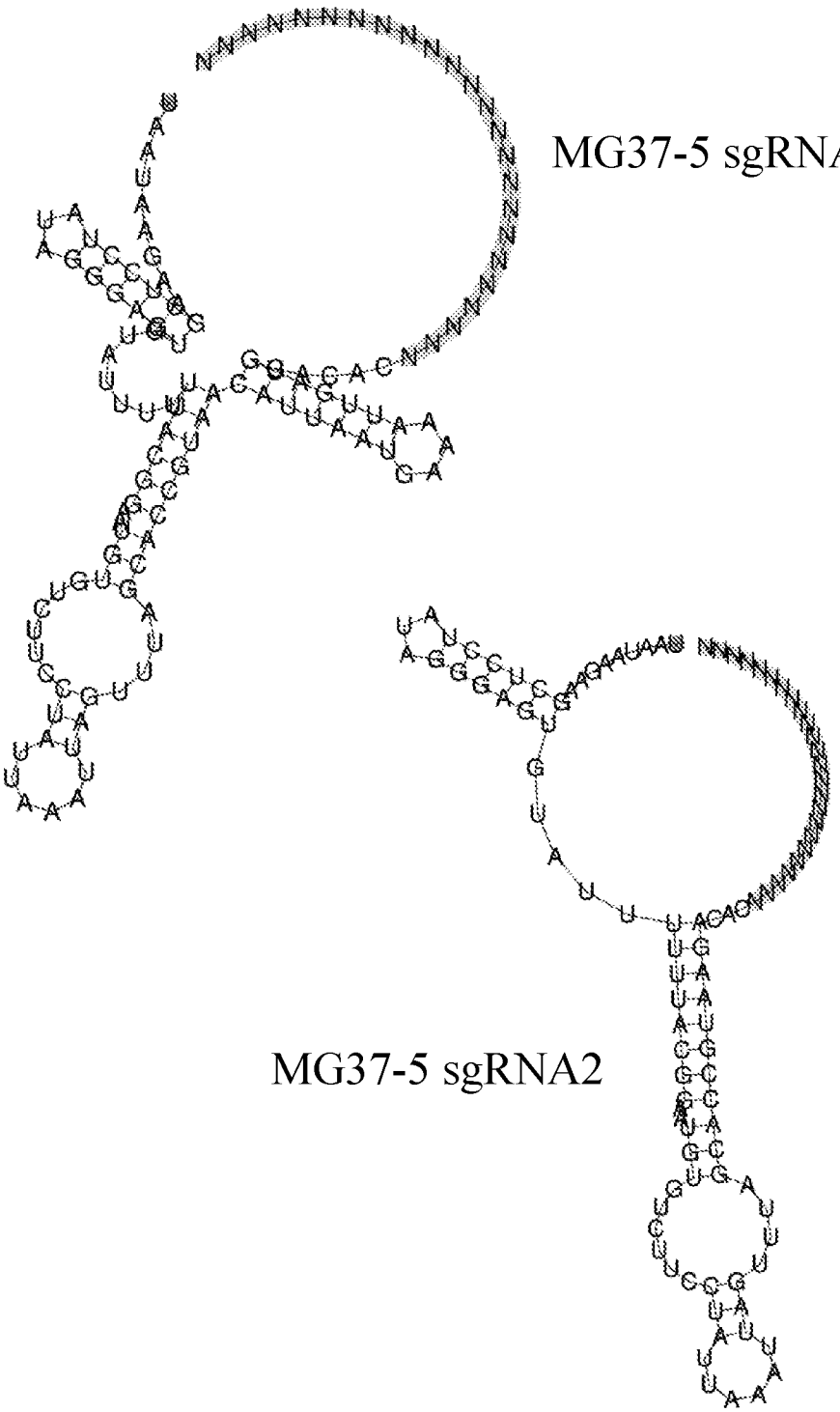


FIG. 14

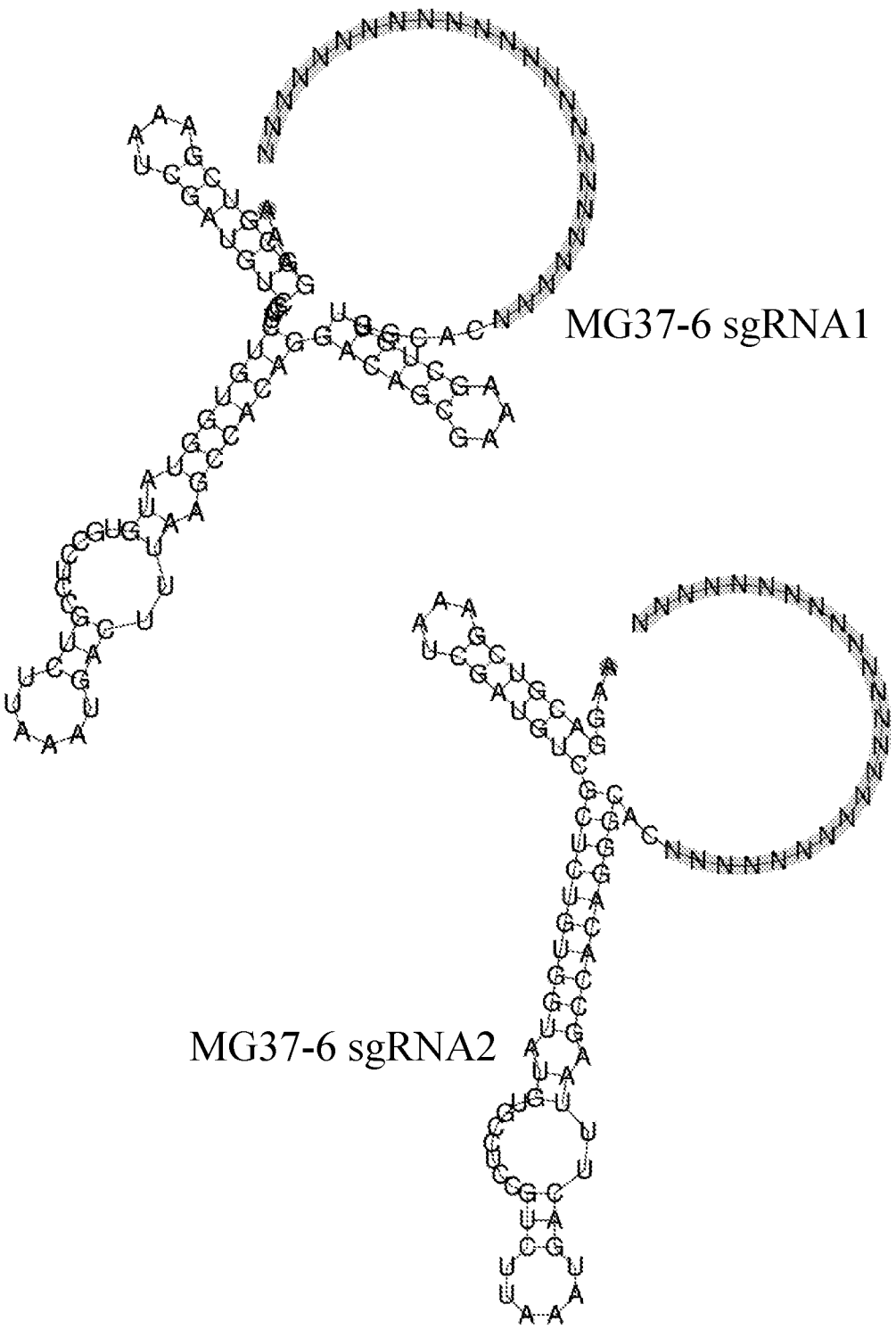


FIG. 14 (cont.)

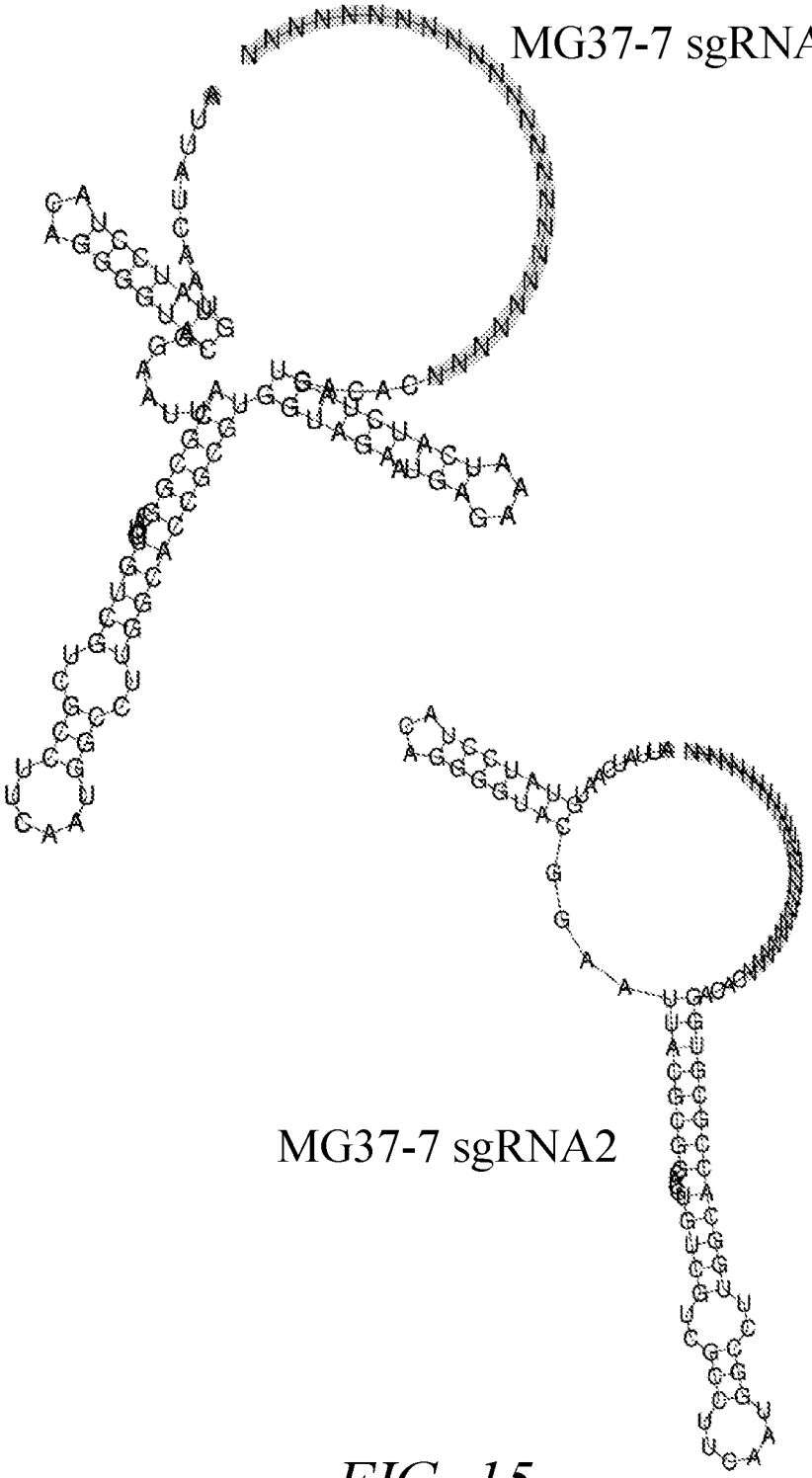


FIG. 15

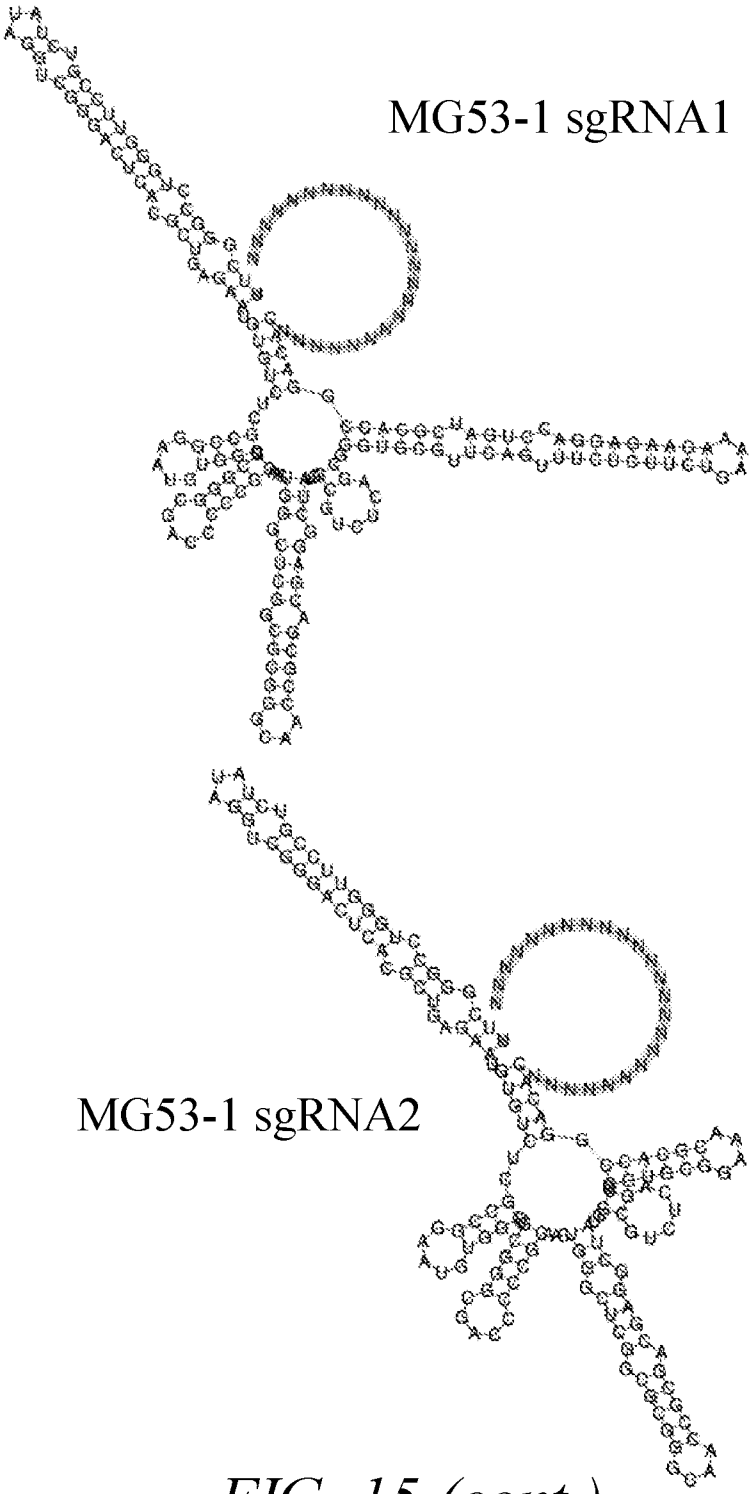
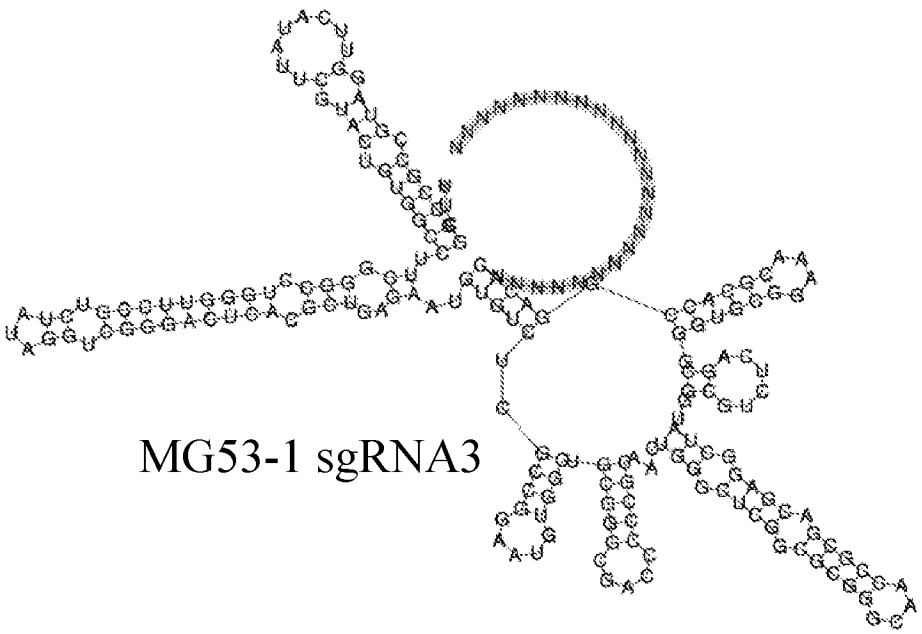
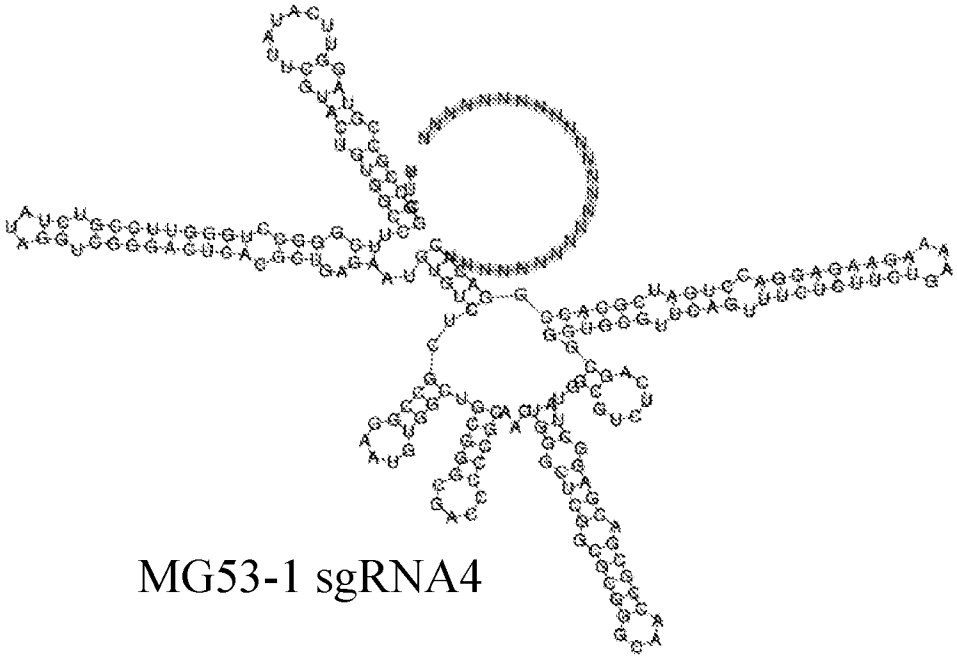


FIG. 15 (cont.)



MG53-1 sgRNA3



MG53-1 sgRNA4

FIG. 16

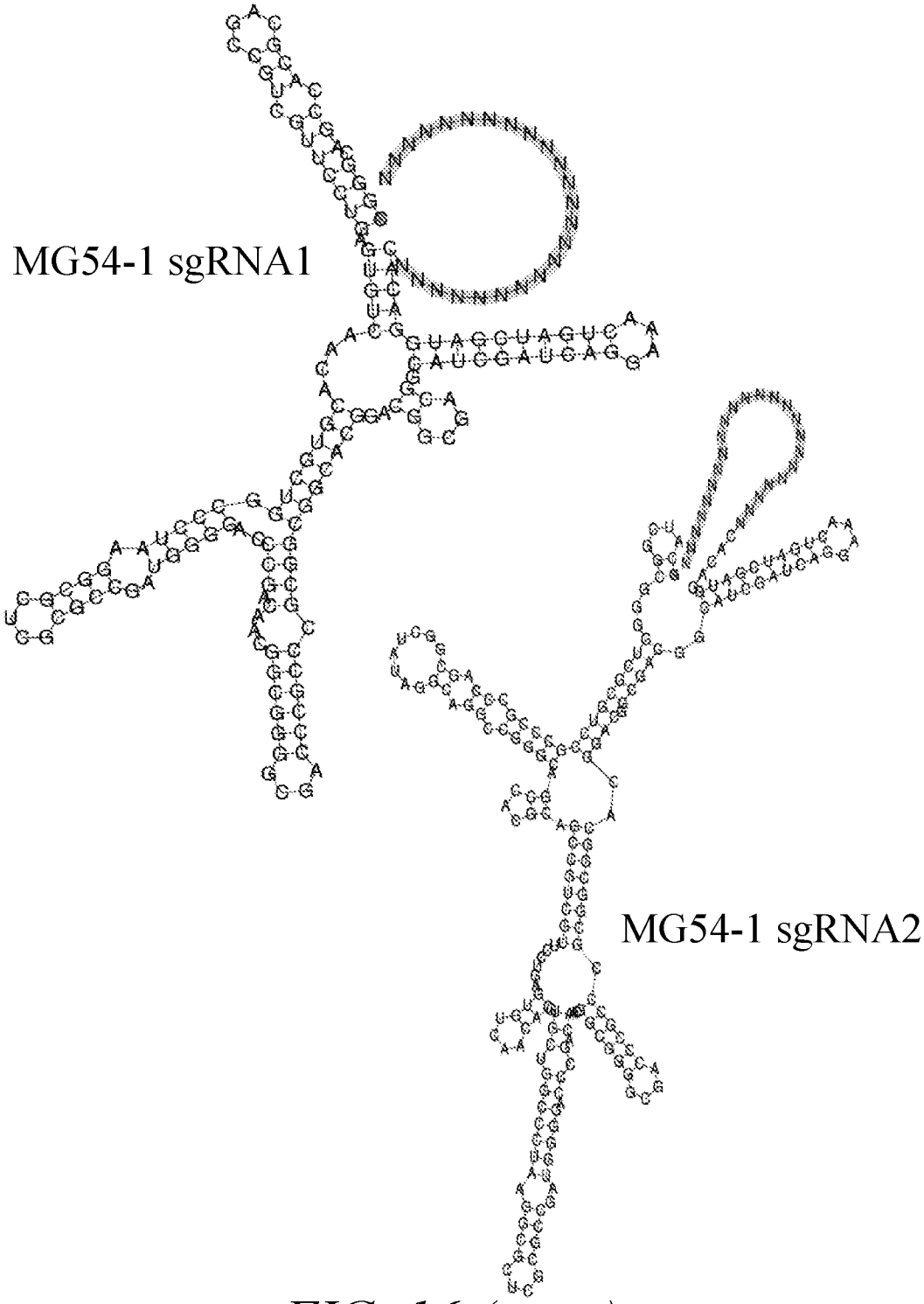


FIG. 16 (cont.)

FIG. 17A

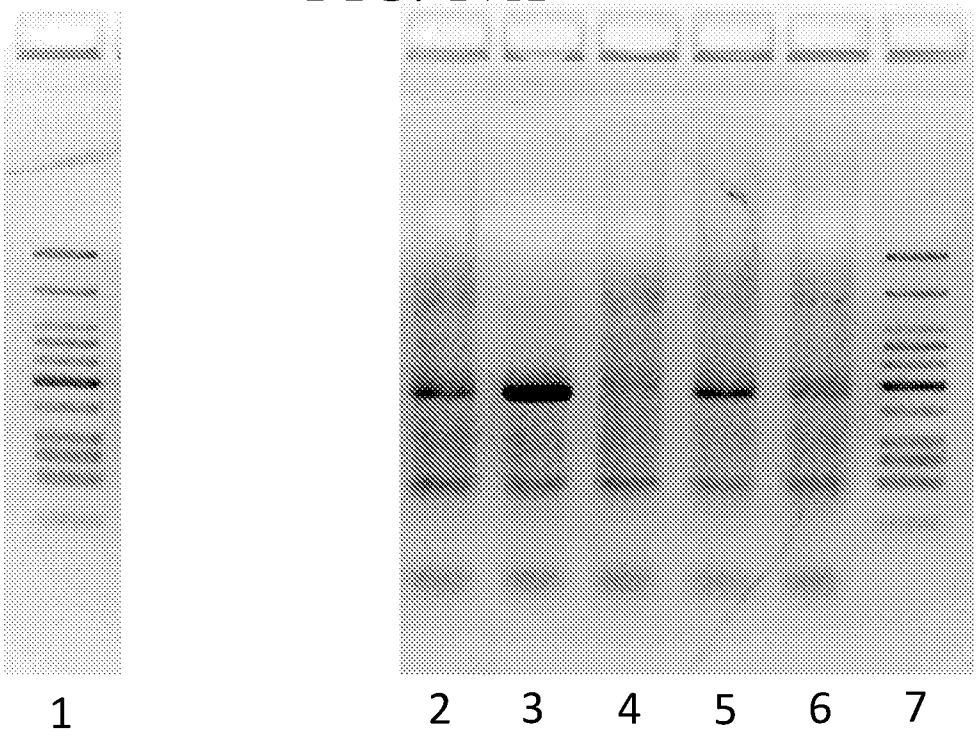
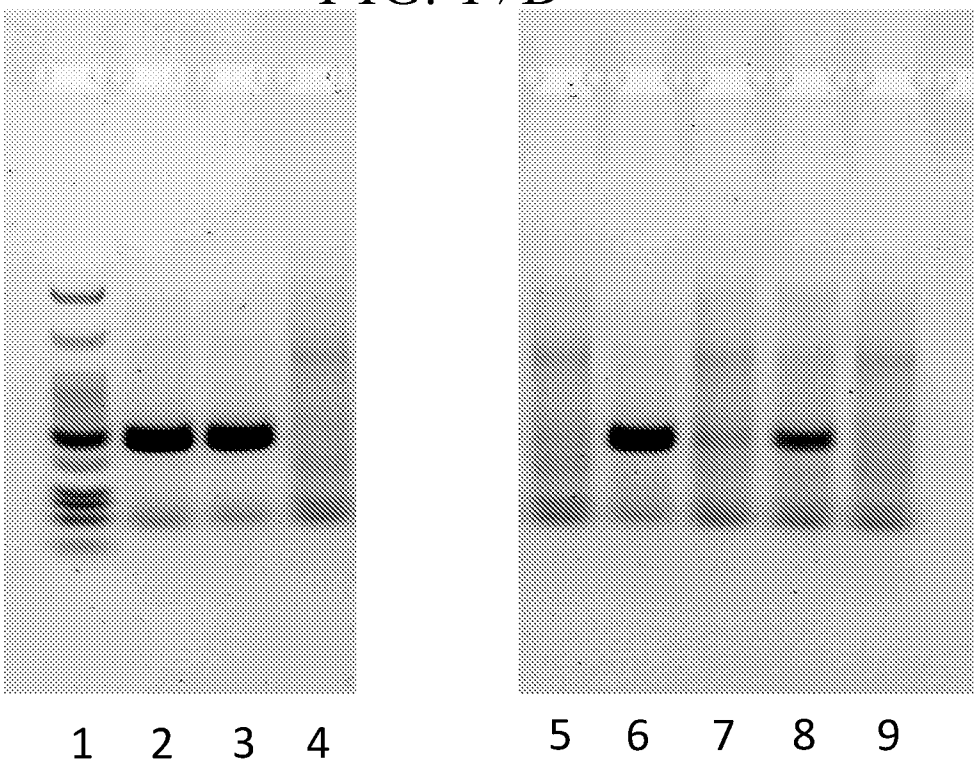


FIG. 17B



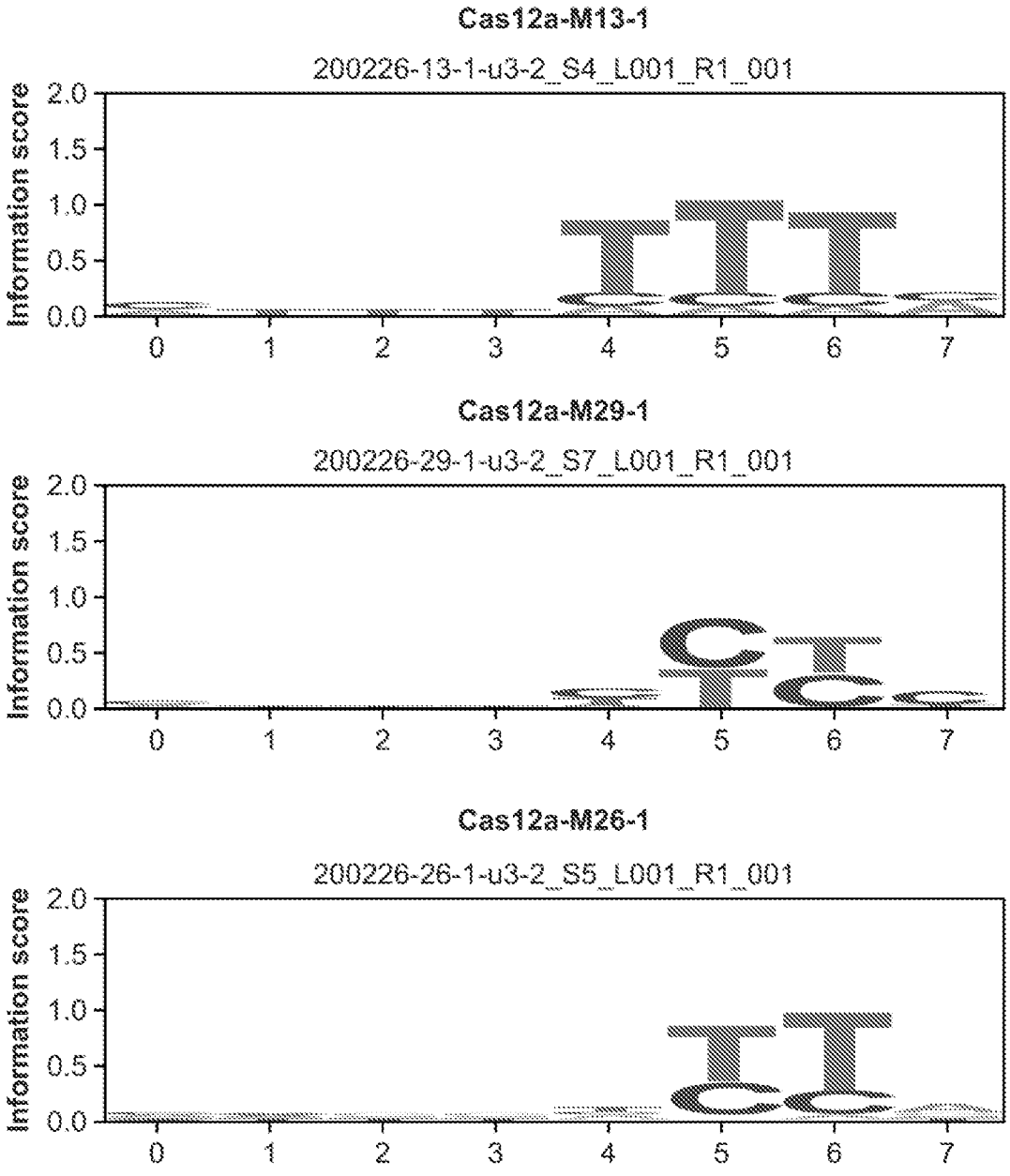

Cas12a-M13-1
200226-13-1-u3-2_S4_L001_R1_001
Information score
2.0
1.5
1.0
0.5
0.0
0
1
2
3
4
5
6
7
Cas12a-M29-1
200226-29-1-u3-2_S7_L001_R1_001
Information score
2.0
1.5
1.0
0.5
0.0
0
1
2
3
4
5
6
7
Cas12a-M26-1
200226-26-1-u3-2_S5_L001_R1_001
Information score
2.0
1.5
1.0
0.5
0.0
0
1
2
3
4
5
6
7

FIG. 18A

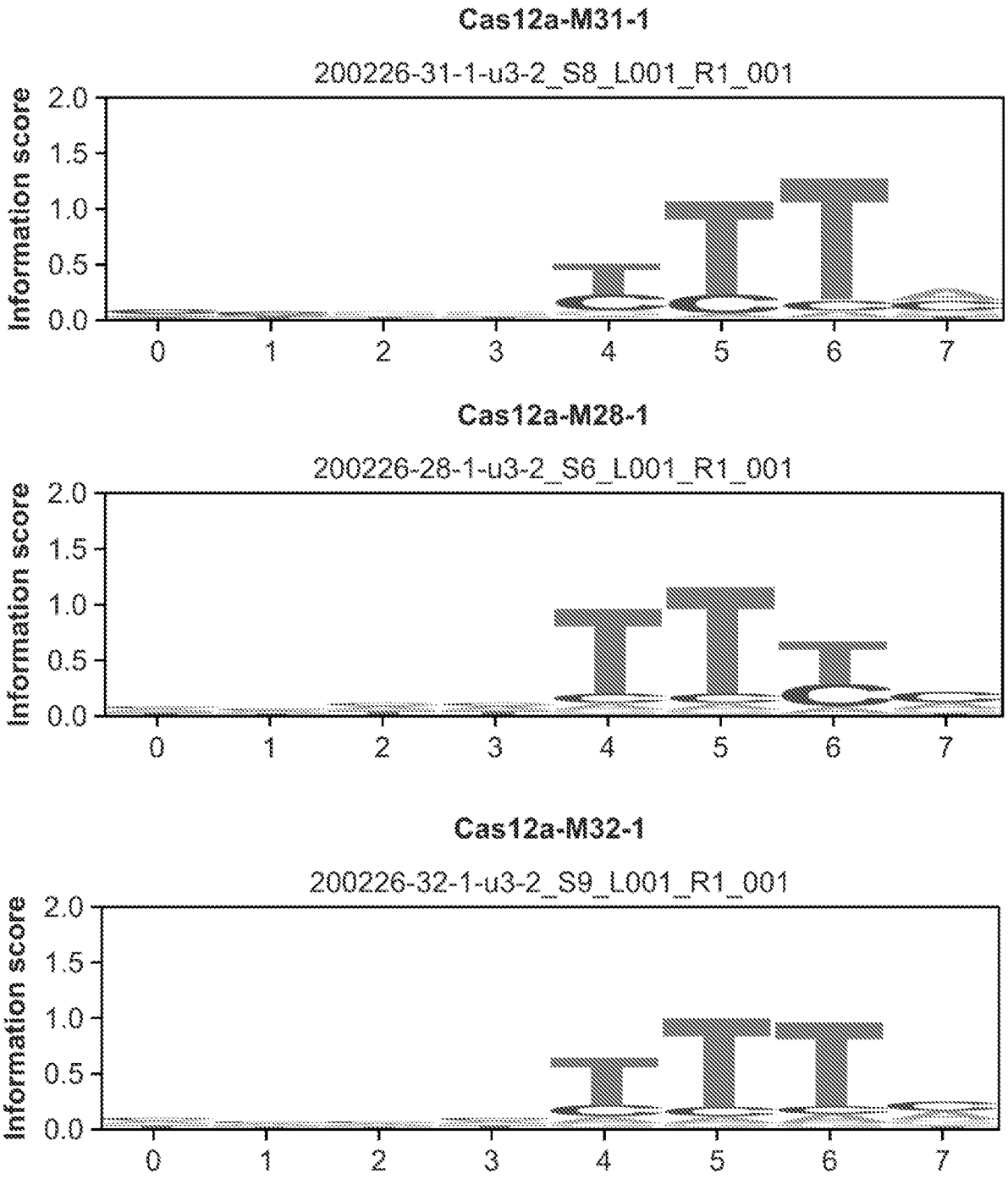

Cas12a-M31-1
200226-31-1-u3-2_S8_L001_R1_001
Information score
2.0
1.5
1.0
0.5
0.0
0
1
2
3
4
5
6
7
Cas12a-M28-1
200226-28-1-u3-2_S6_L001_R1_001
Information score
2.0
1.5
1.0
0.5
0.0
0
1
2
3
4
5
6
7
Cas12a-M32-1
200226-32-1-u3-2_S9_L001_R1_001
Information score
2.0
1.5
1.0
0.5
0.0
0
1
2
3
4
5
6
7

FIG. 18A (cont.)

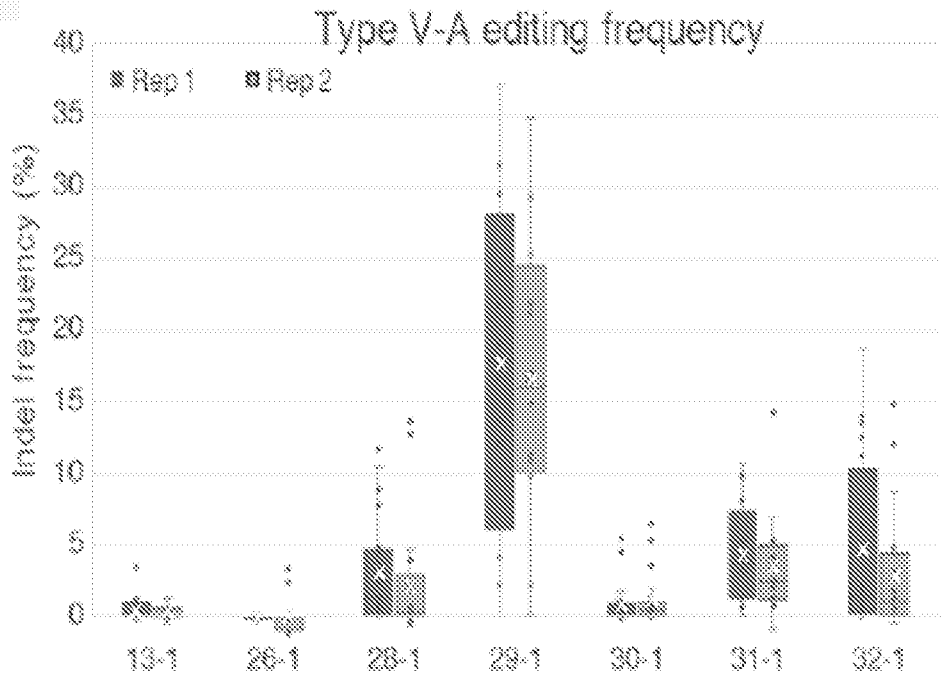


FIG. 18B

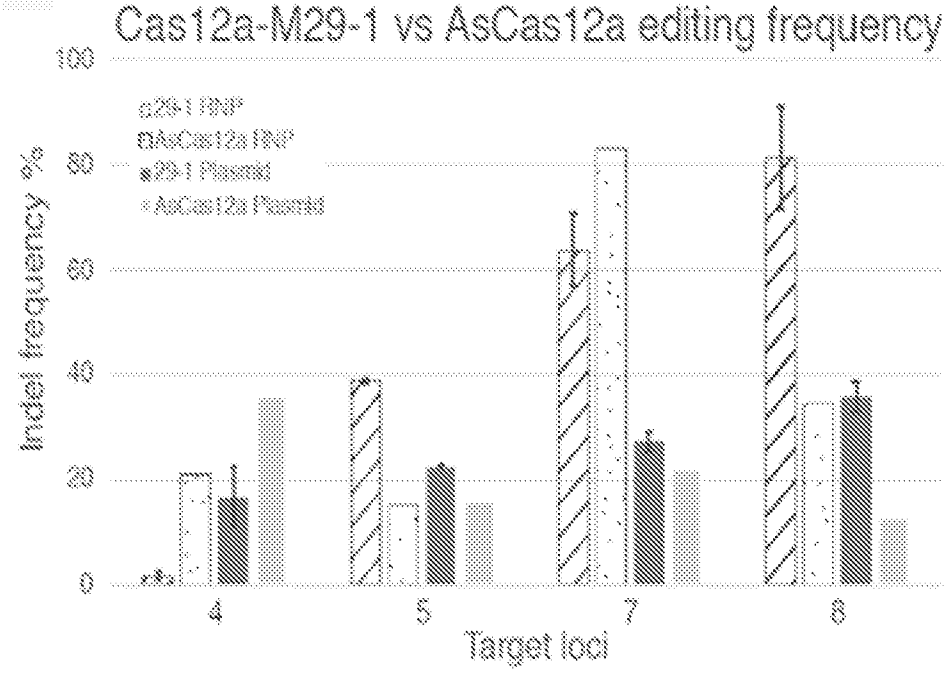


FIG. 18C

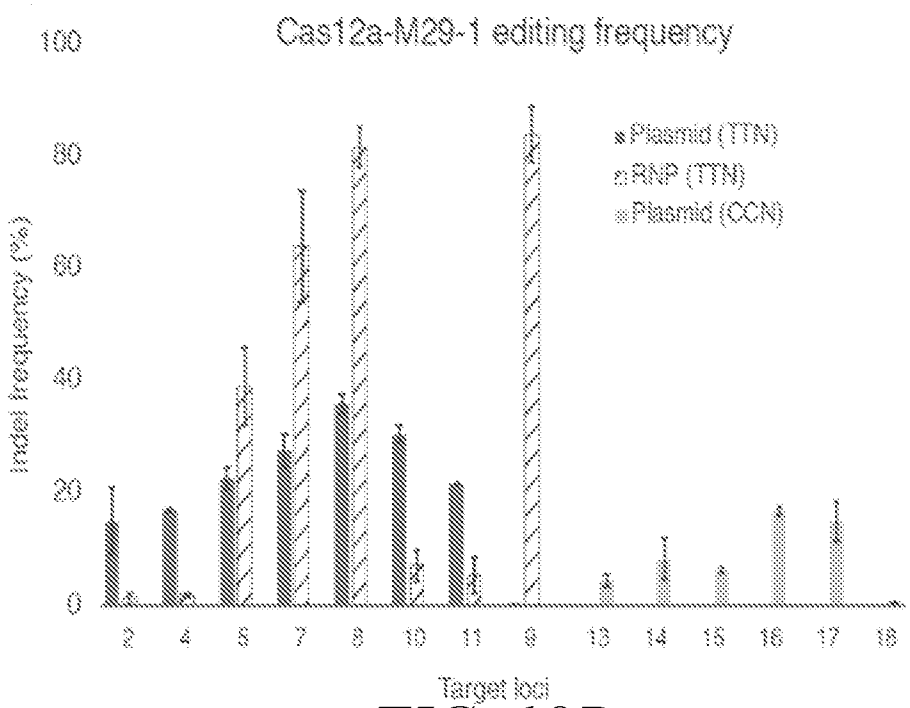


FIG. 18D

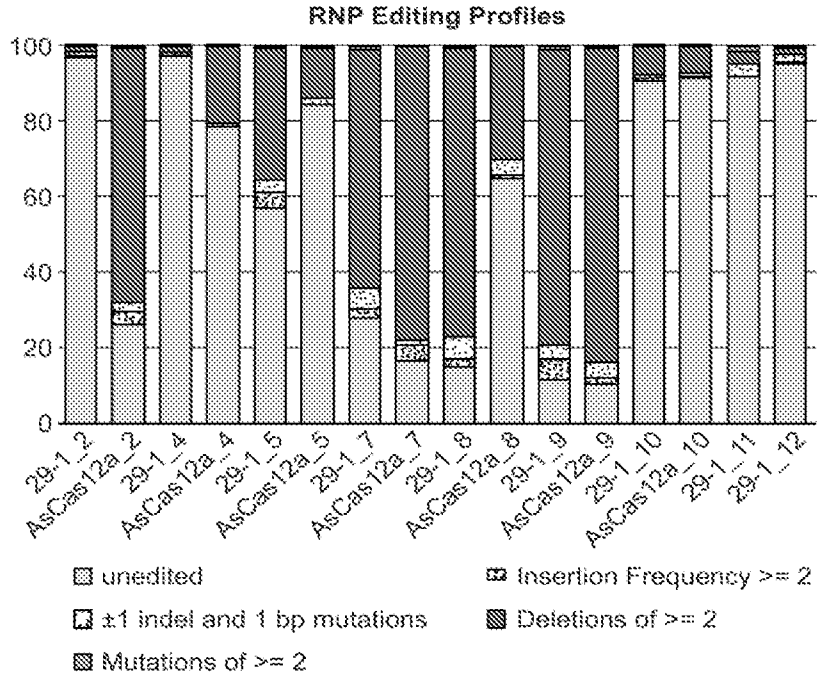


FIG. 18E

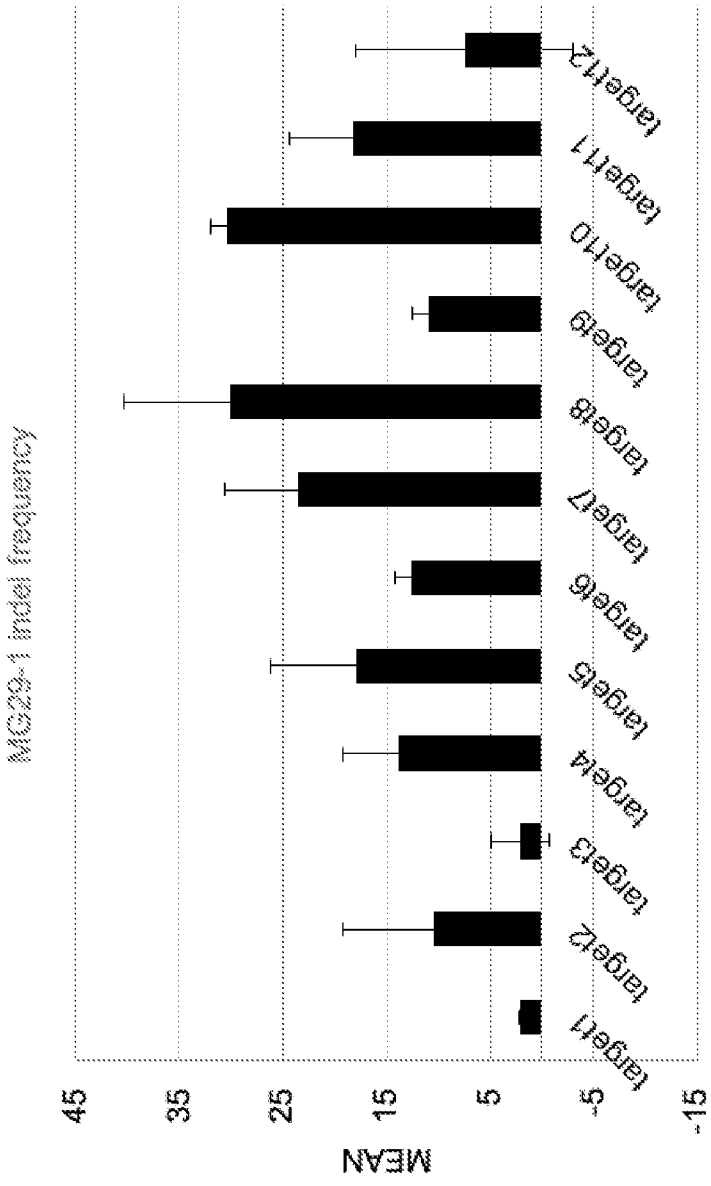
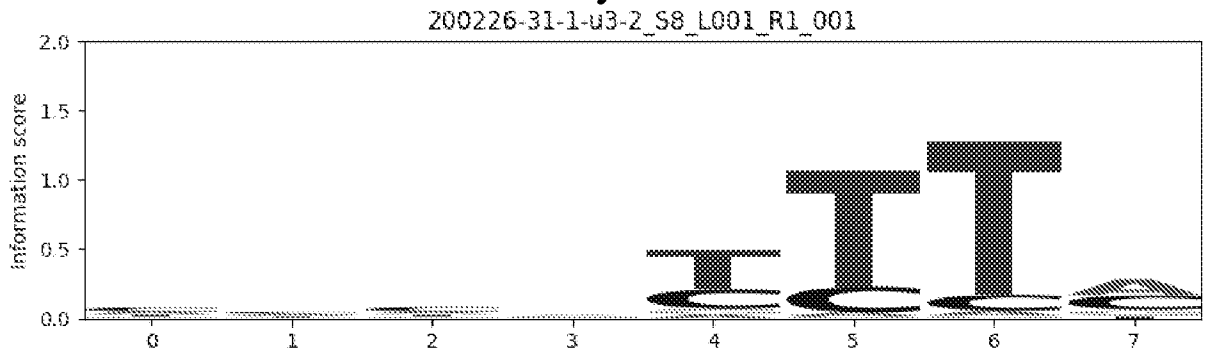
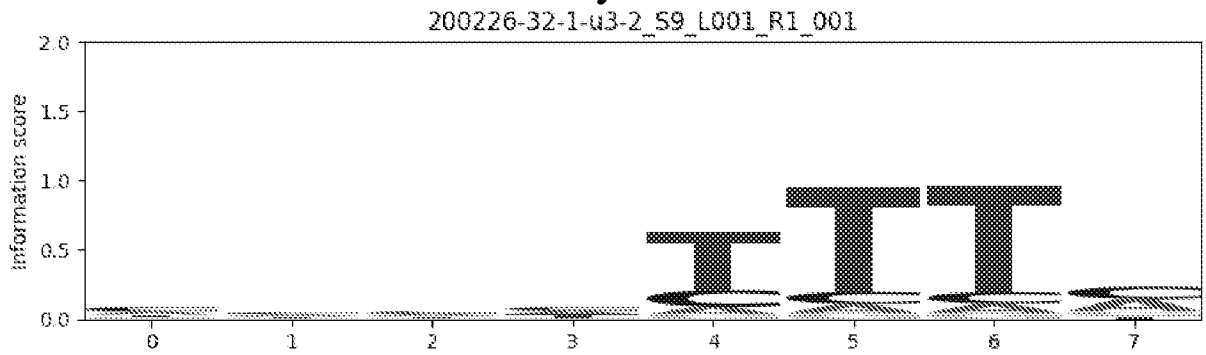

MG29-1 indel frequency
MEAN
45
35
25
15
5
-5
-15
target1
target2
target3
target4
target5
target6
target7
target8
target9
target10
target11
target12

FIG. 19

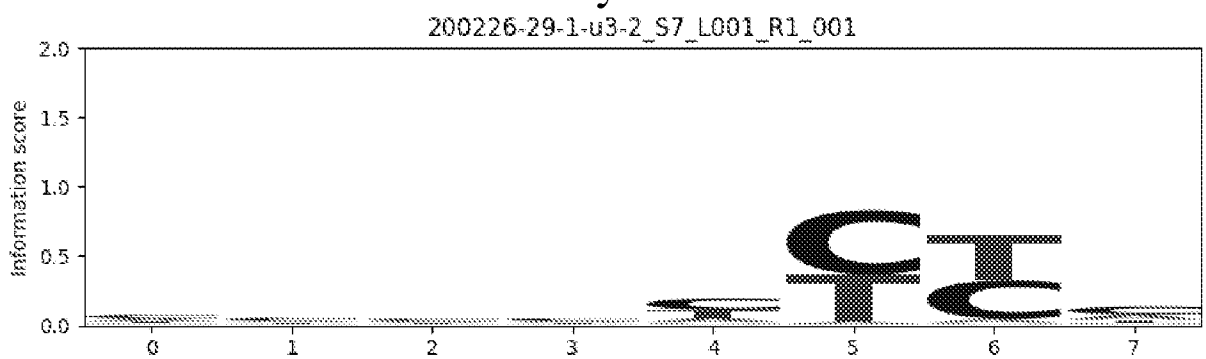
MG31-1 Synthetic PAM



MG32-1 Synthetic PAM



MG29-1 Synthetic PAM



MG28-1 Synthetic PAM

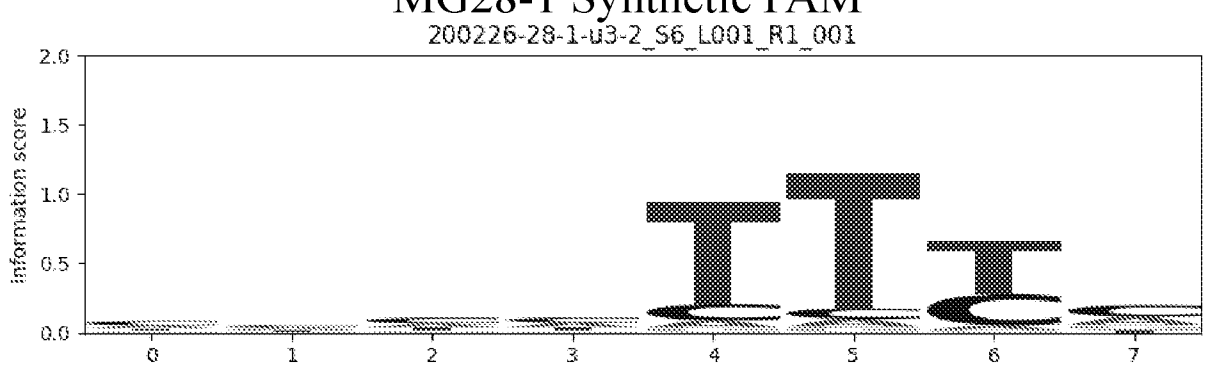
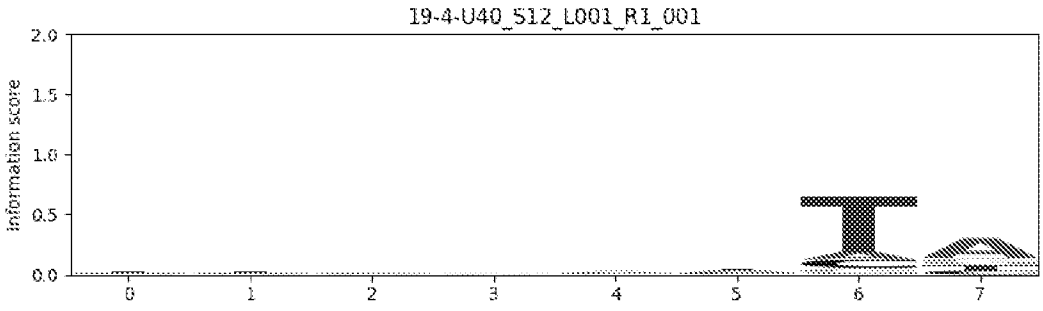
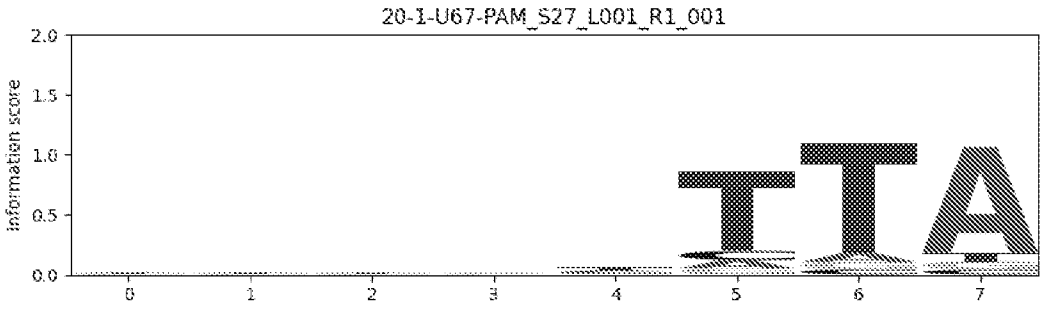


FIG. 20

MG19-4 Synthetic PAM



MG20-1 Synthetic PAM



MG29-5 Synthetic PAM

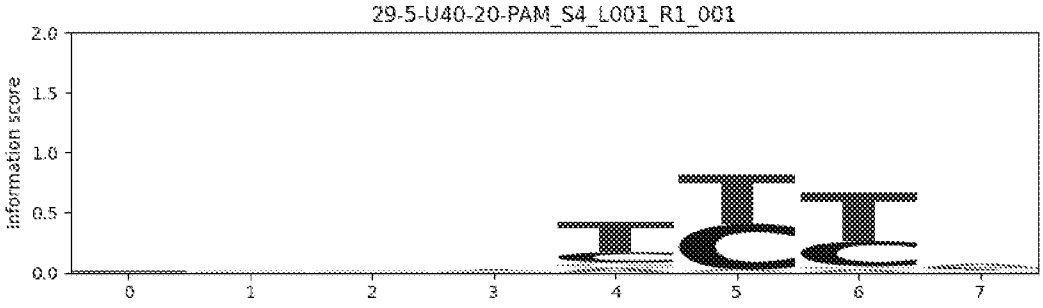
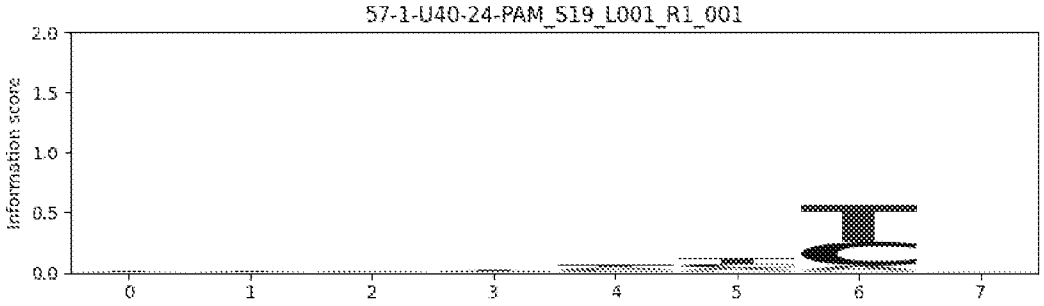
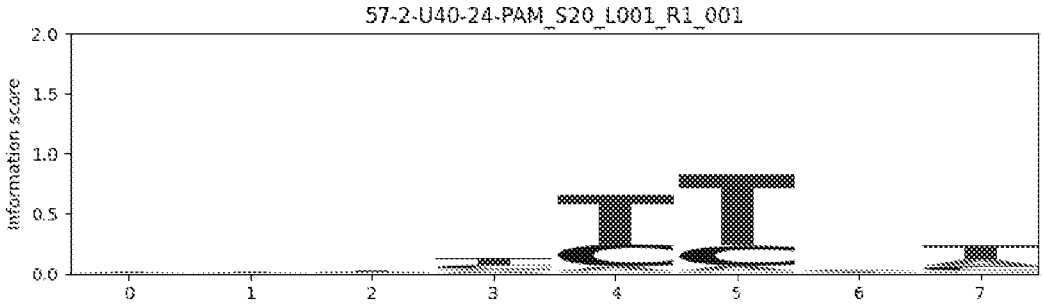


FIG. 21

MG57-1 Synthetic PAM



MG57-2 Synthetic PAM



MG59-1 Synthetic PAM

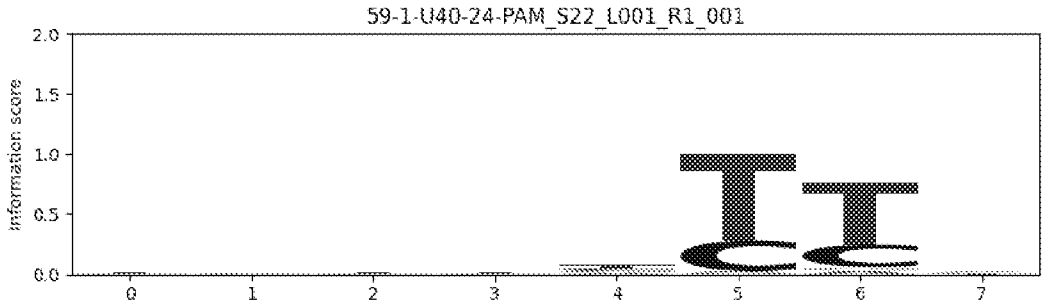
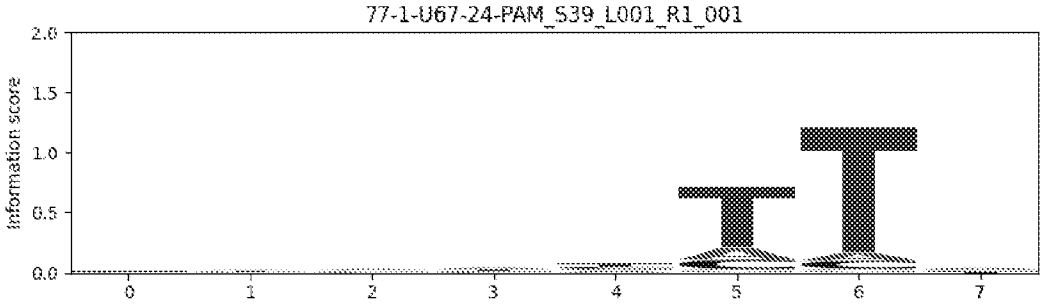
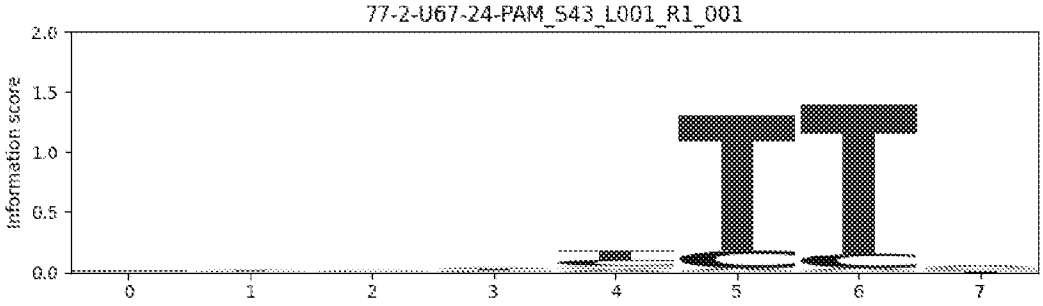


FIG. 22

MG77-1 Synthetic PAM



MG77-2 Synthetic PAM



MG78-1 Synthetic PAM

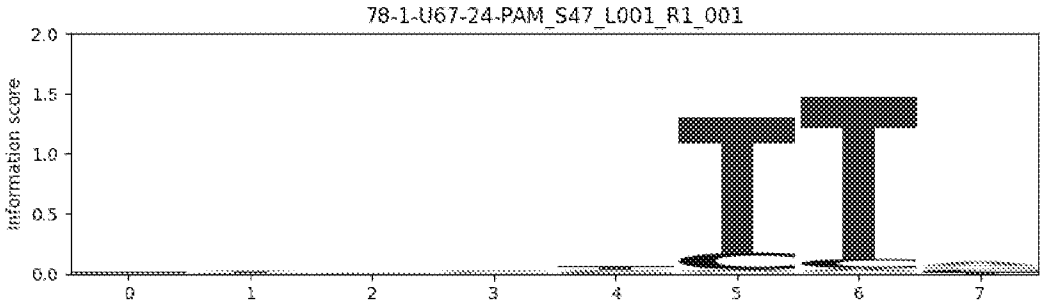


FIG. 23

MG59-2 Synthetic PAM

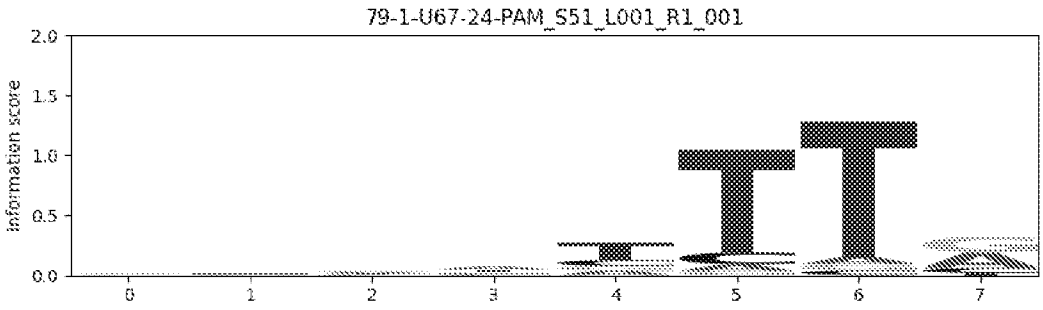


FIG. 24

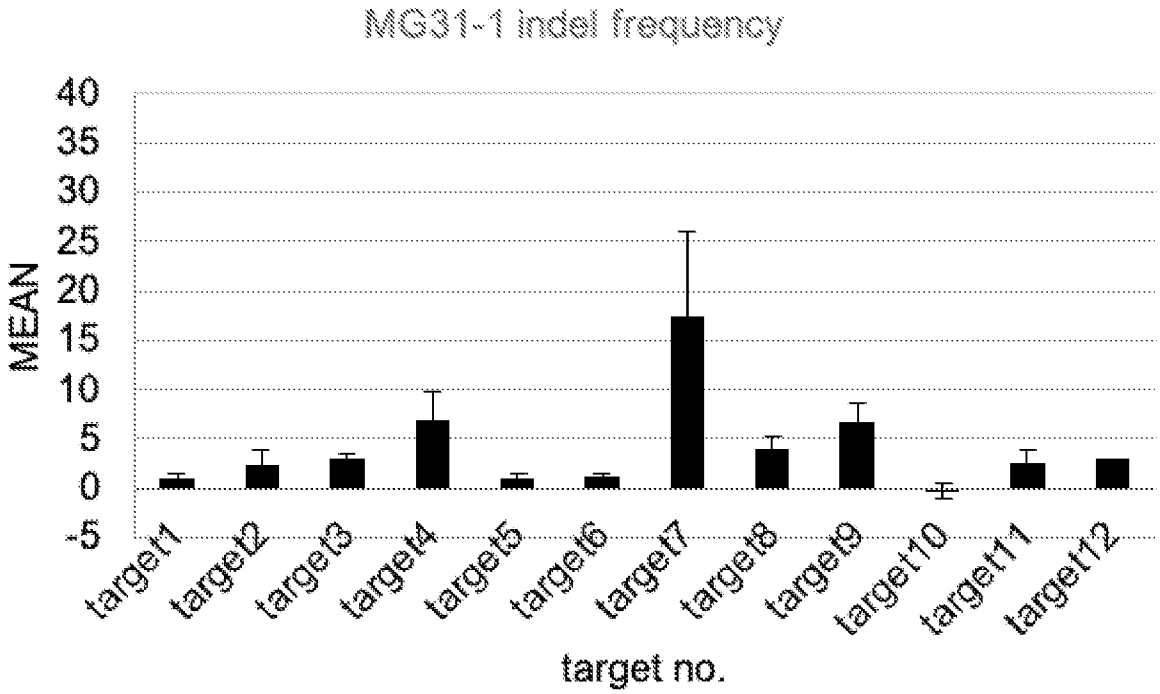

MG31-1 indel frequency
MEAN
40
35
30
25
20
15
10
5
0
-5
target1
target2
target3
target4
target5
target6
target7
target8
target9
target10
target11
target12
target no.

FIG. 25

FIG. 26B

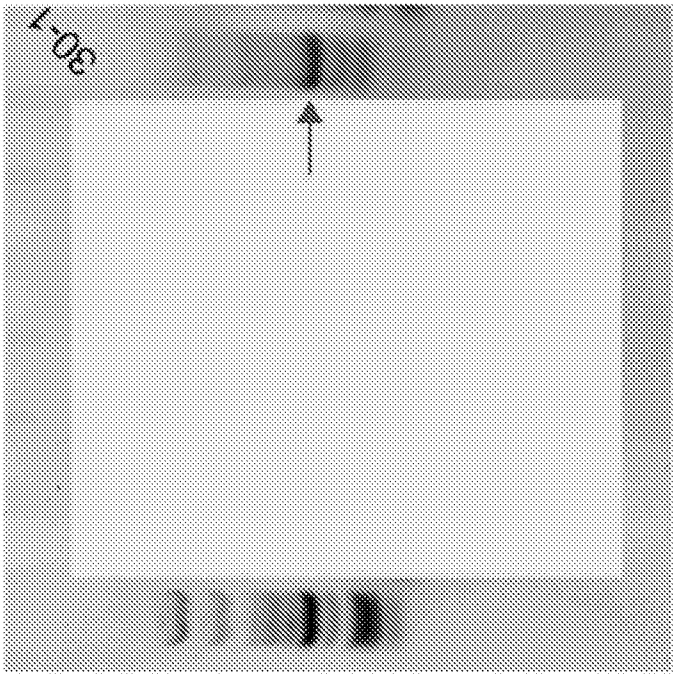
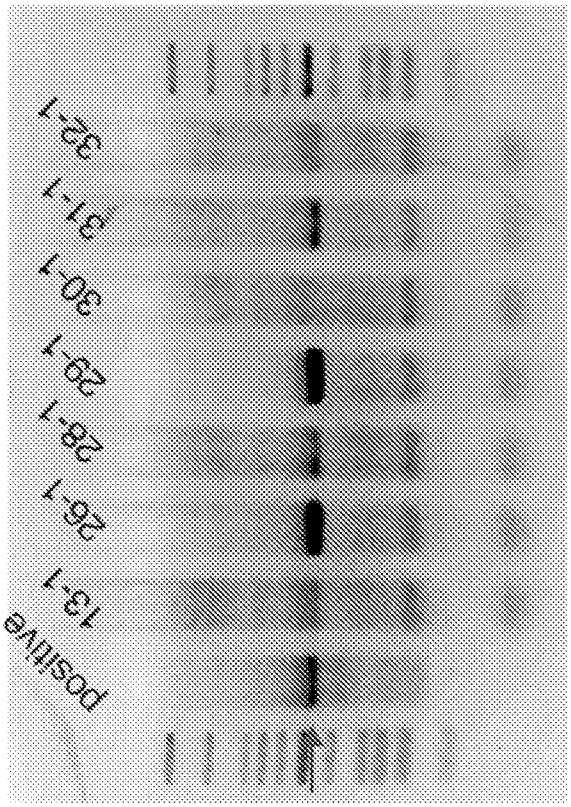


FIG. 26A



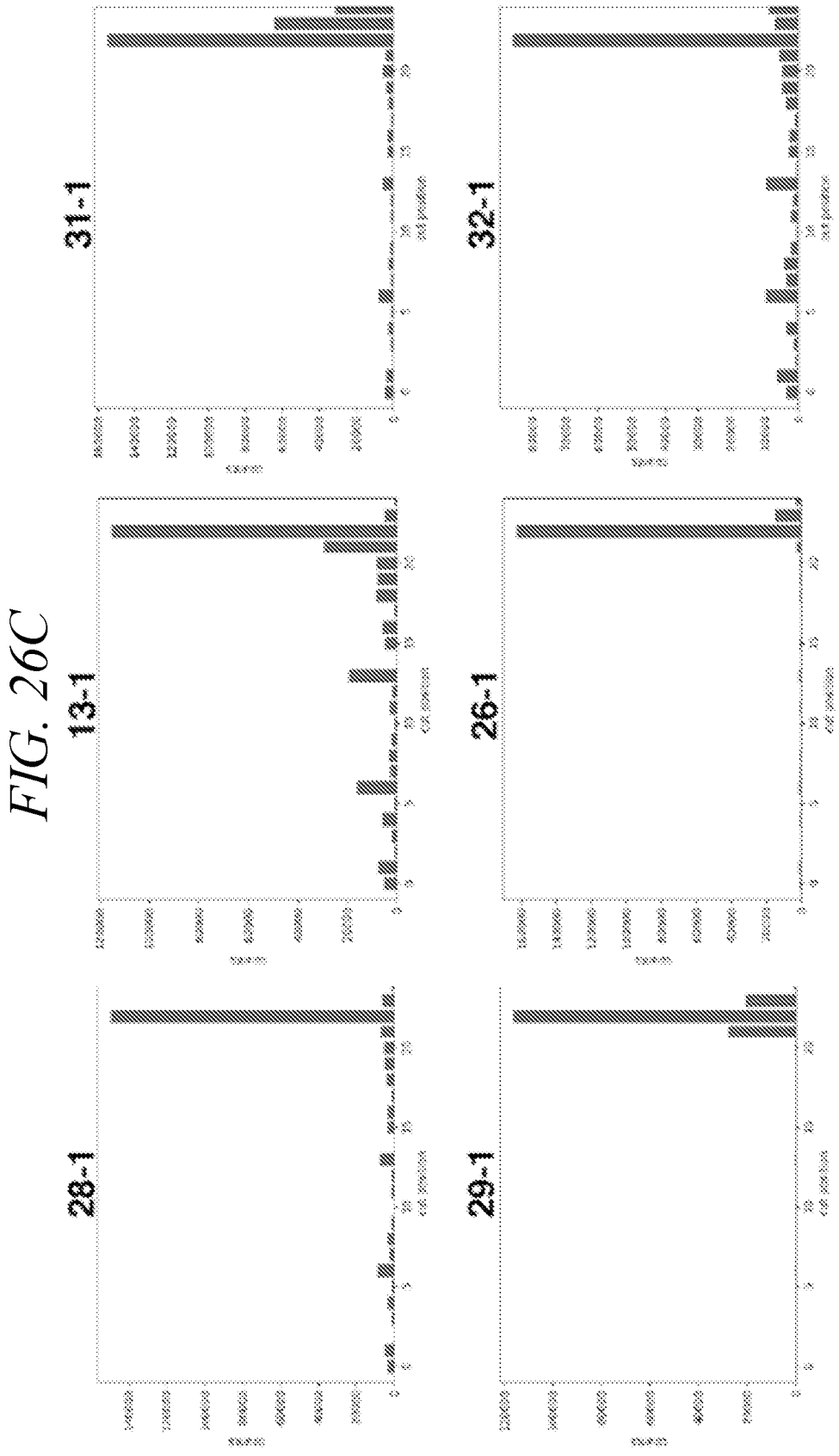


FIG. 26C

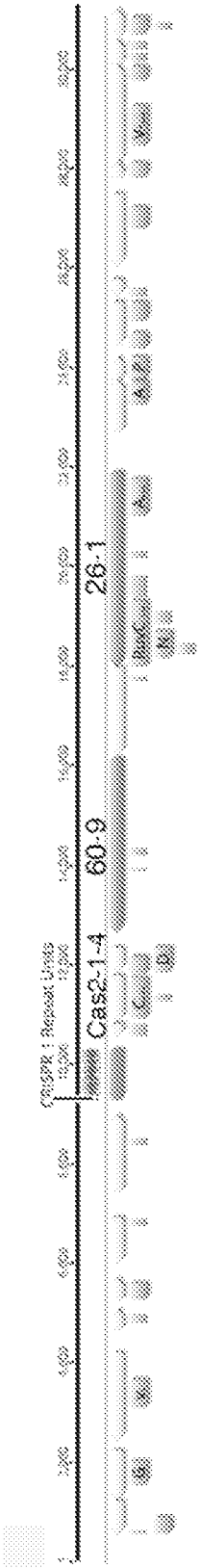


FIG. 27A

FIG. 27B

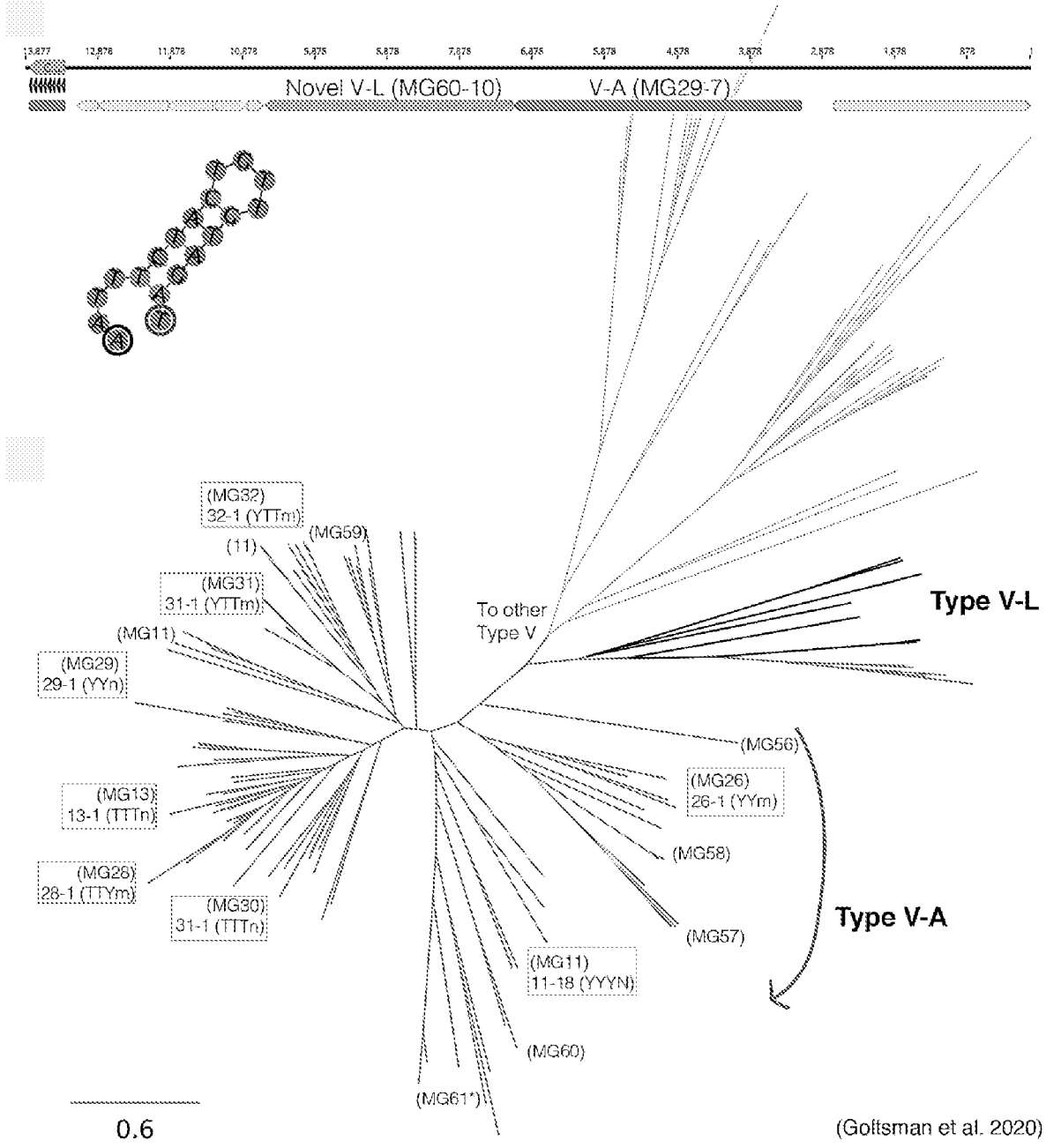

Novel V-L (MG60-10)
V-A (MG29-7)
(MG32)
32-1 (YTTm)
(MG59)
(11)
(MG31)
31-1 (YTTm)
(MG11)
(MG29)
29-1 (YYn)
To other
Type V
Type V-L
(MG56)
(MG26)
26-1 (YYm)
(MG13)
13-1 (TTTn)
(MG58)
(MG28)
28-1 (TTYm)
(MG30)
31-1 (TTTn)
(MG57)
Type V-A
(MG11)
11-18 (YYYN)
(MG60)
(MG61*)
0.6
(Goltsman et al. 2020)

FIG. 28

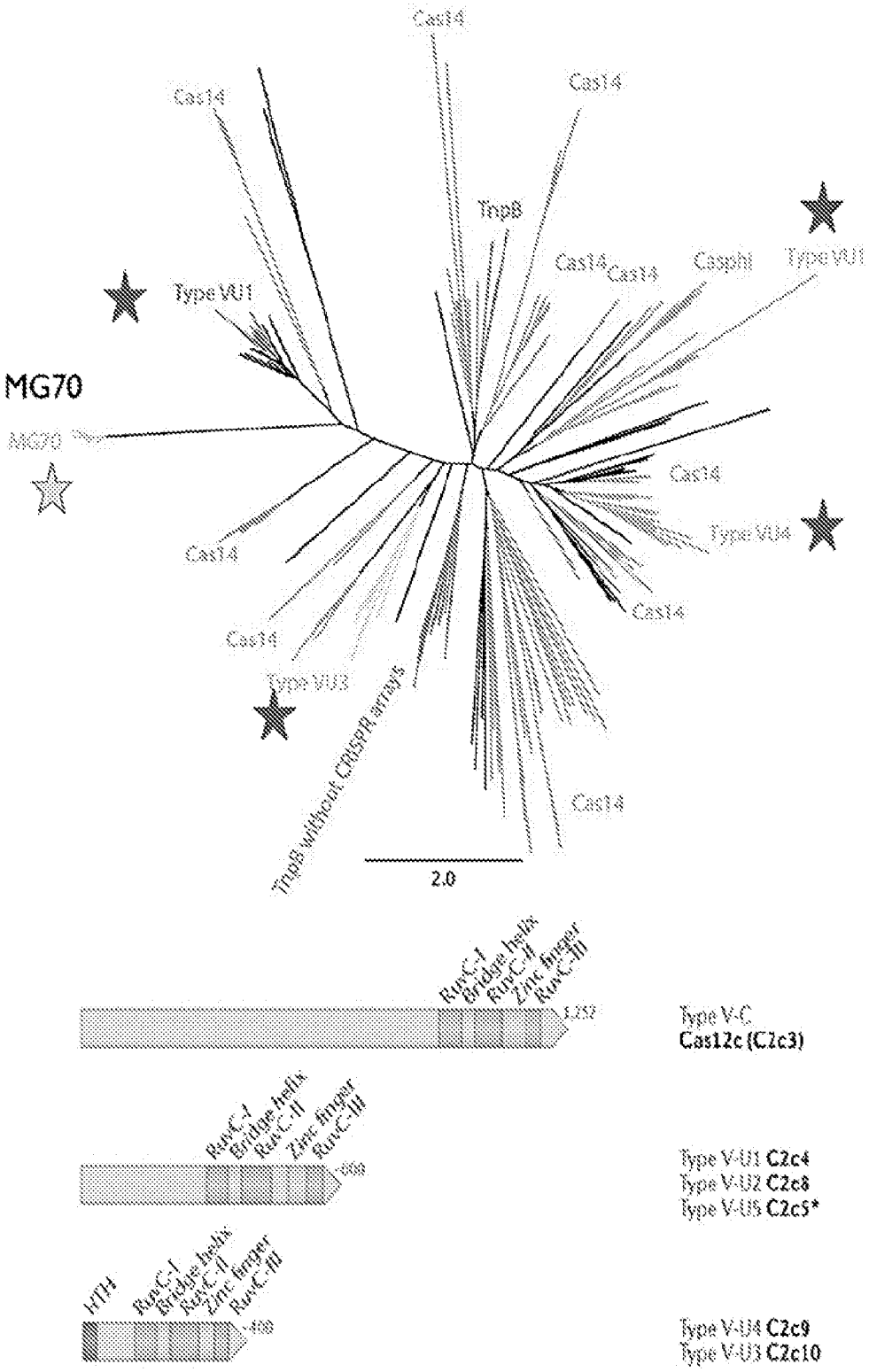

Cas14
Cas14
Cas14
TnpB
Type VU1
Cas14
Cas14
Casphi
Type VU1
MG70
MG70
Cas14
Type VU4
Cas14
Cas14
Cas14
Type VU3
TnpB without CRISPR arrays
Cas14
2.0
RuvC-I
Bridge helix
RuvC-II
Zinc finger
RuvC-III
1257
Type V-C
Cas12c (C2c3)
RuvC-I
Bridge helix
RuvC-II
Zinc finger
RuvC-III
~980
Type V-U1 C2c4
Type V-U2 C2c8
Type V-U5 C2c5*
HTH
RuvC-I
Bridge helix
RuvC-II
Zinc finger
RuvC-III
~700
Type V-U4 C2c9
Type V-U3 C2c10

FIG. 29

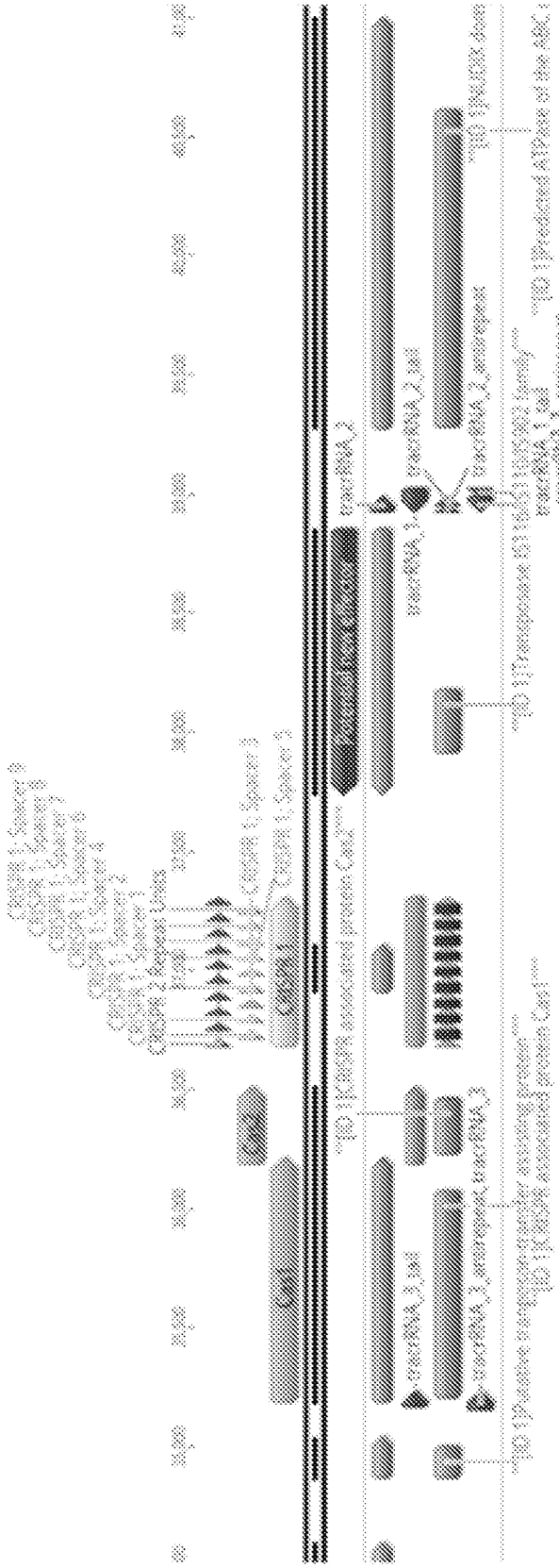


FIG. 30

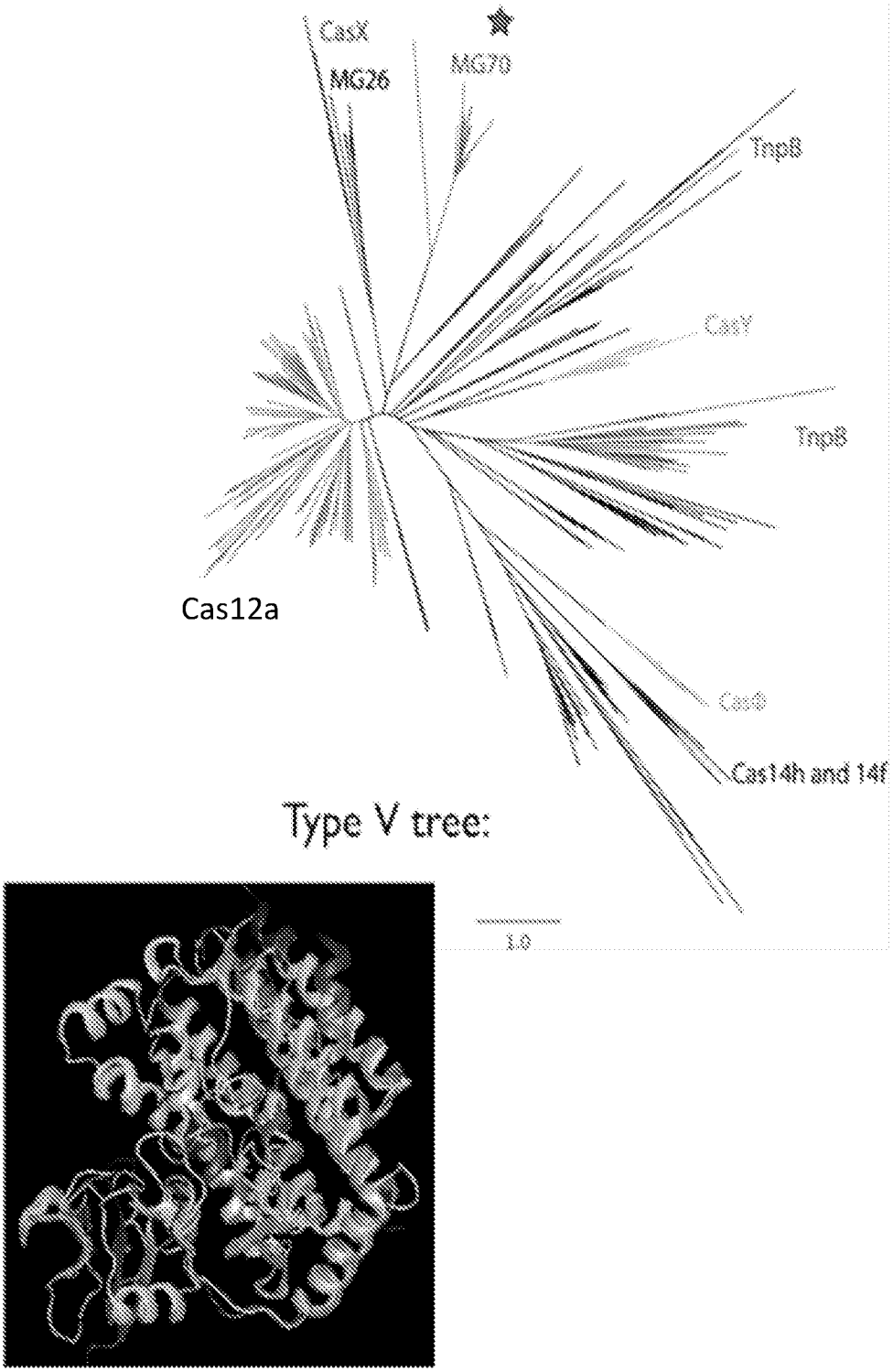


FIG. 30 (cont.)

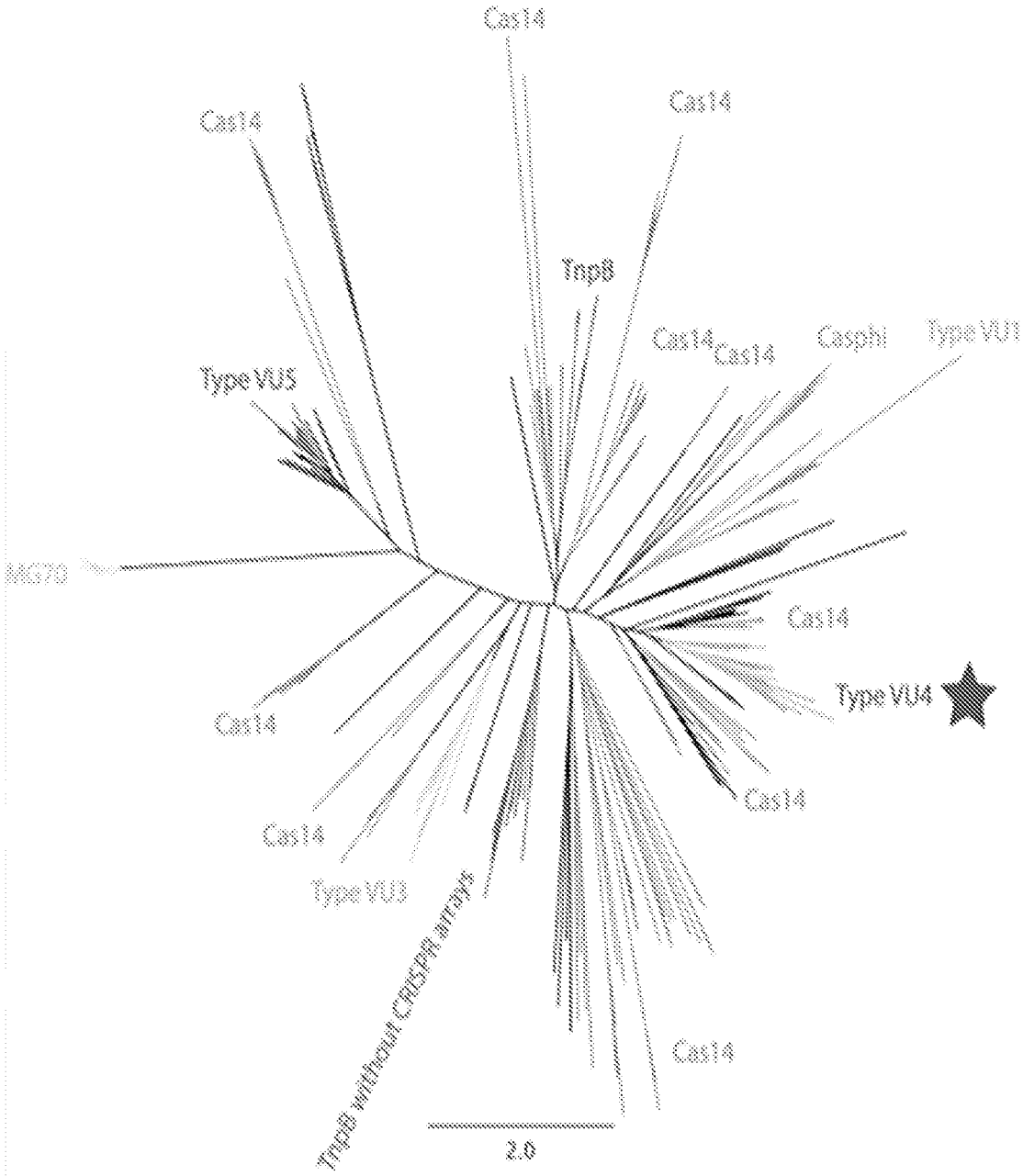

Cas14
Cas14
Cas14
Tnp8
Cas14
Cas14
Casphi
Type VU1
Type VU5
Cas14
Type VU4
Cas14
Cas14
Cas14
Type VU3
Tnp8 without CRISPR arrays
Cas14
2.0

FIG. 31

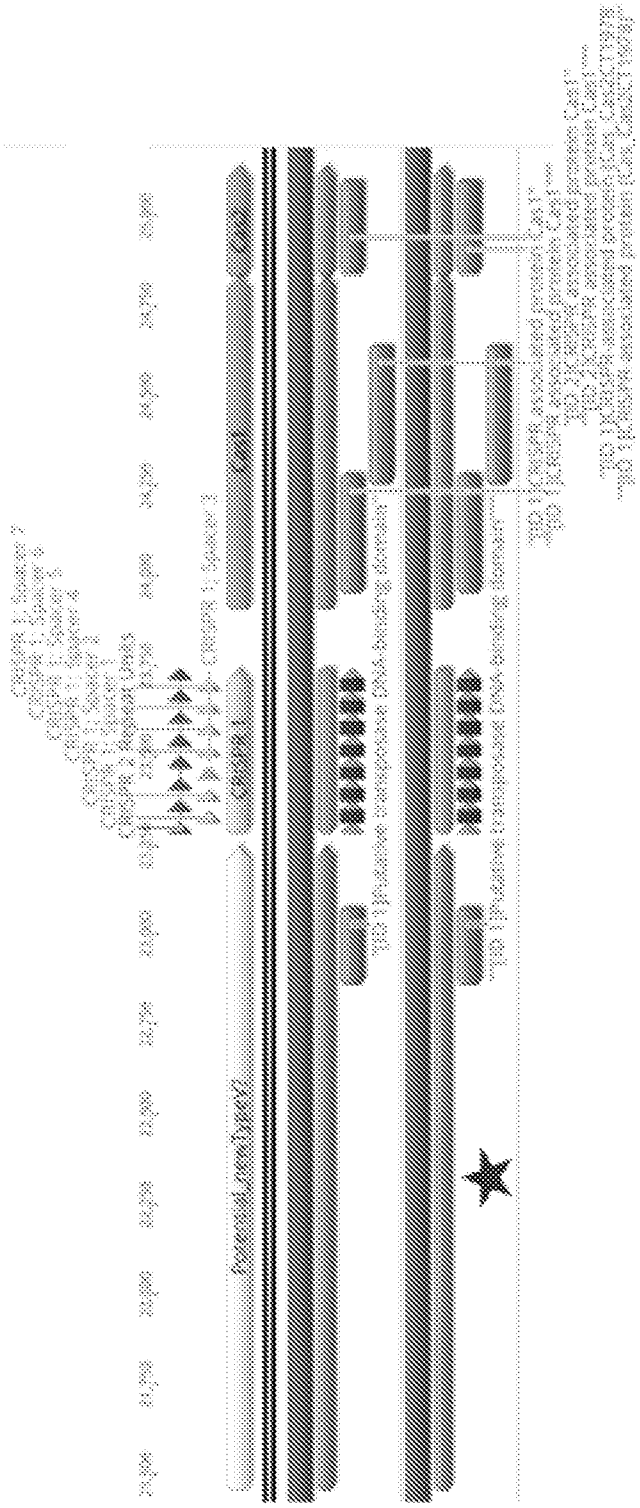


FIG. 31 (cont.)

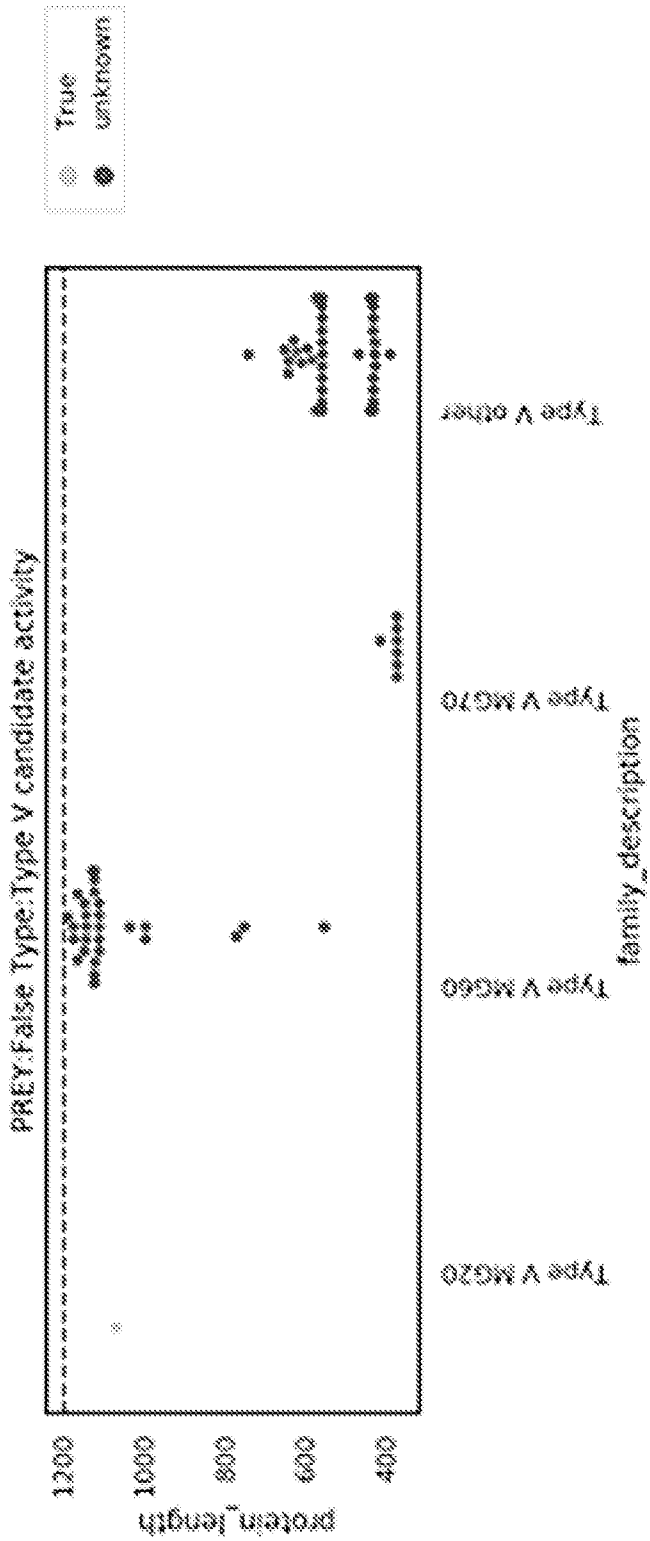

True
unknown
PREY:False Type:Type V candidate activity
protein_length
1200
1000
800
600
400
Type V MG20
Type V MG60
Type V MG70
Type V other
family_description

FIG. 32

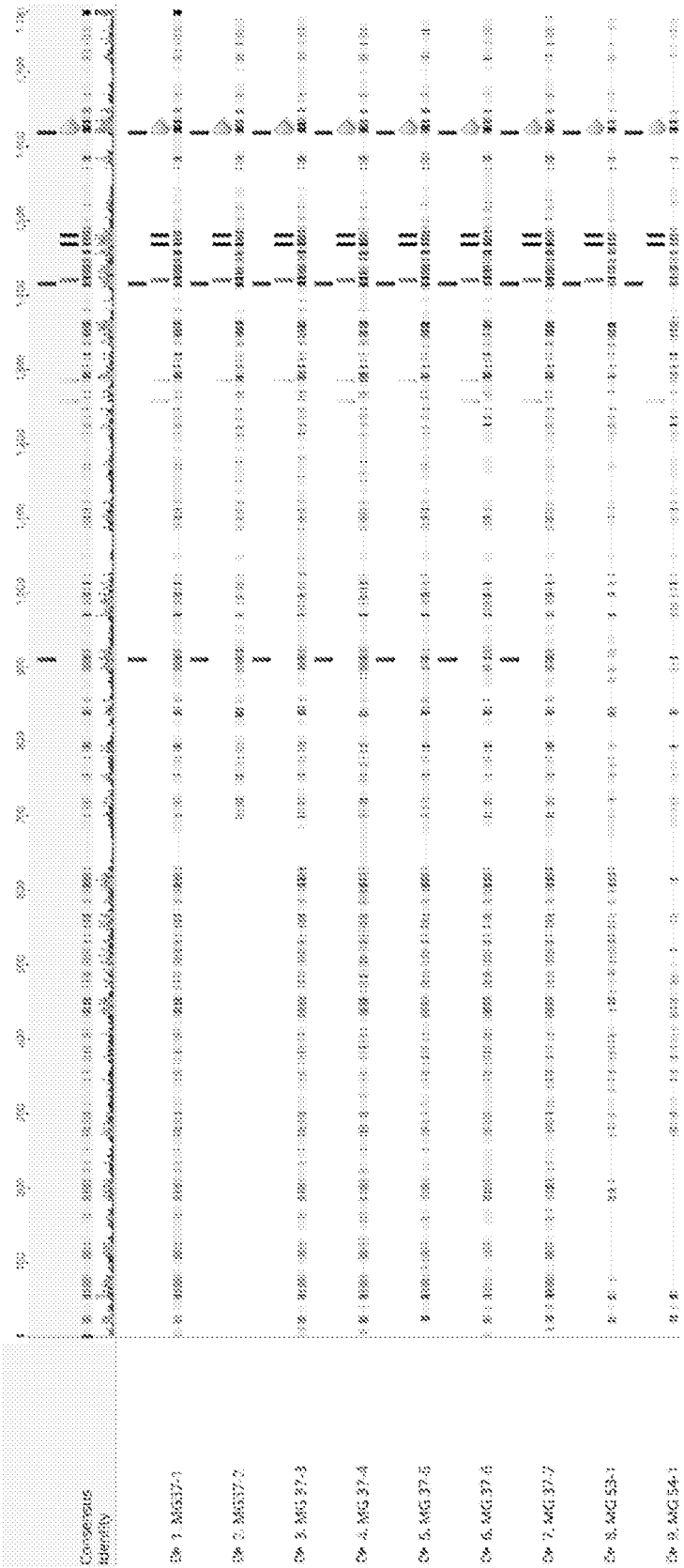
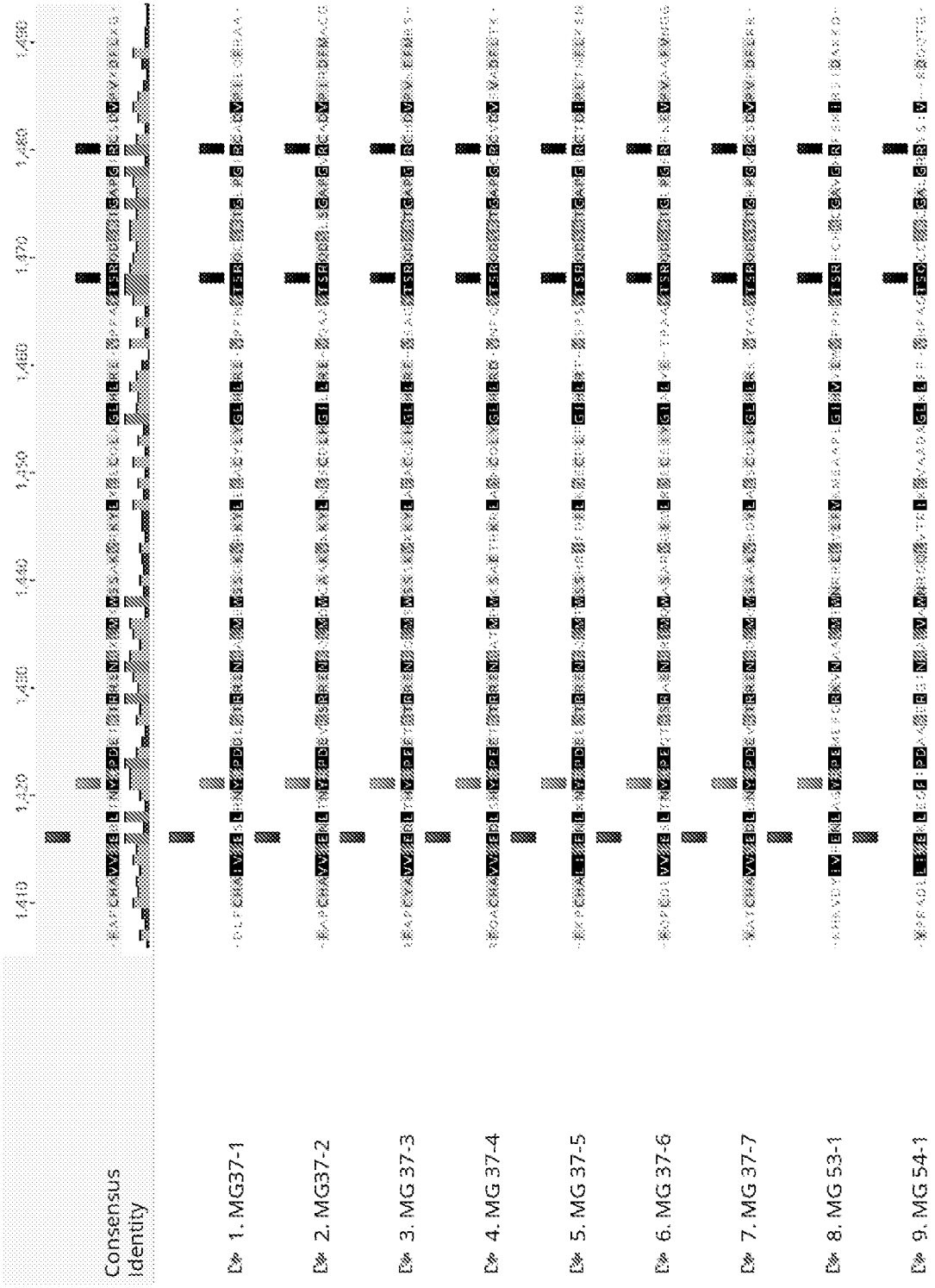


FIG. 33

FIG. 34


890
900
910
920
930
Consensus
Identity
1. MG37-1
2. MG37-2
3. MG37-3
4. MG37-4
5. MG37-5
6. MG37-6
7. MG37-7
8. MG53-1
9. MG54-1

FIG. 35


1,410
1,420
1,430
1,440
1,450
1,460
1,470
1,480
1,490
Consensus identity
1. MG37-1
2. MG37-2
3. MG 37-3
4. MG 37-4
5. MG 37-5
6. MG 37-6
7. MG 37-7
8. MG 53-1
9. MG 54-1

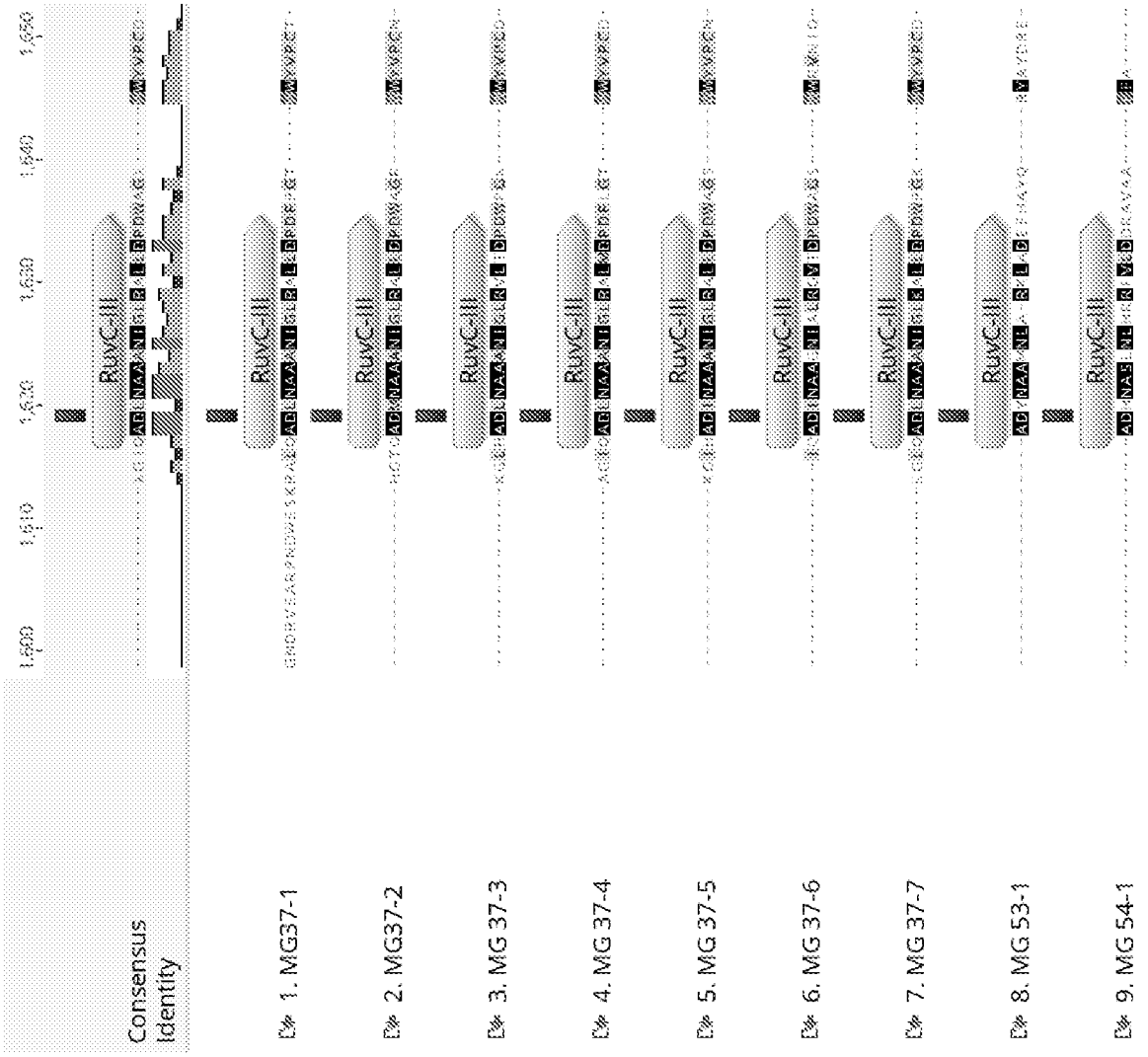

1,600
1,610
1,620
1,630
1,640
1,650
Consensus
Identity
RuvC-III
1. MG37-1
2. MG37-2
3. MG 37-3
4. MG 37-4
5. MG 37-5
6. MG 37-6
7. MG 37-7
8. MG 53-1
9. MG 54-1

FIG. 36

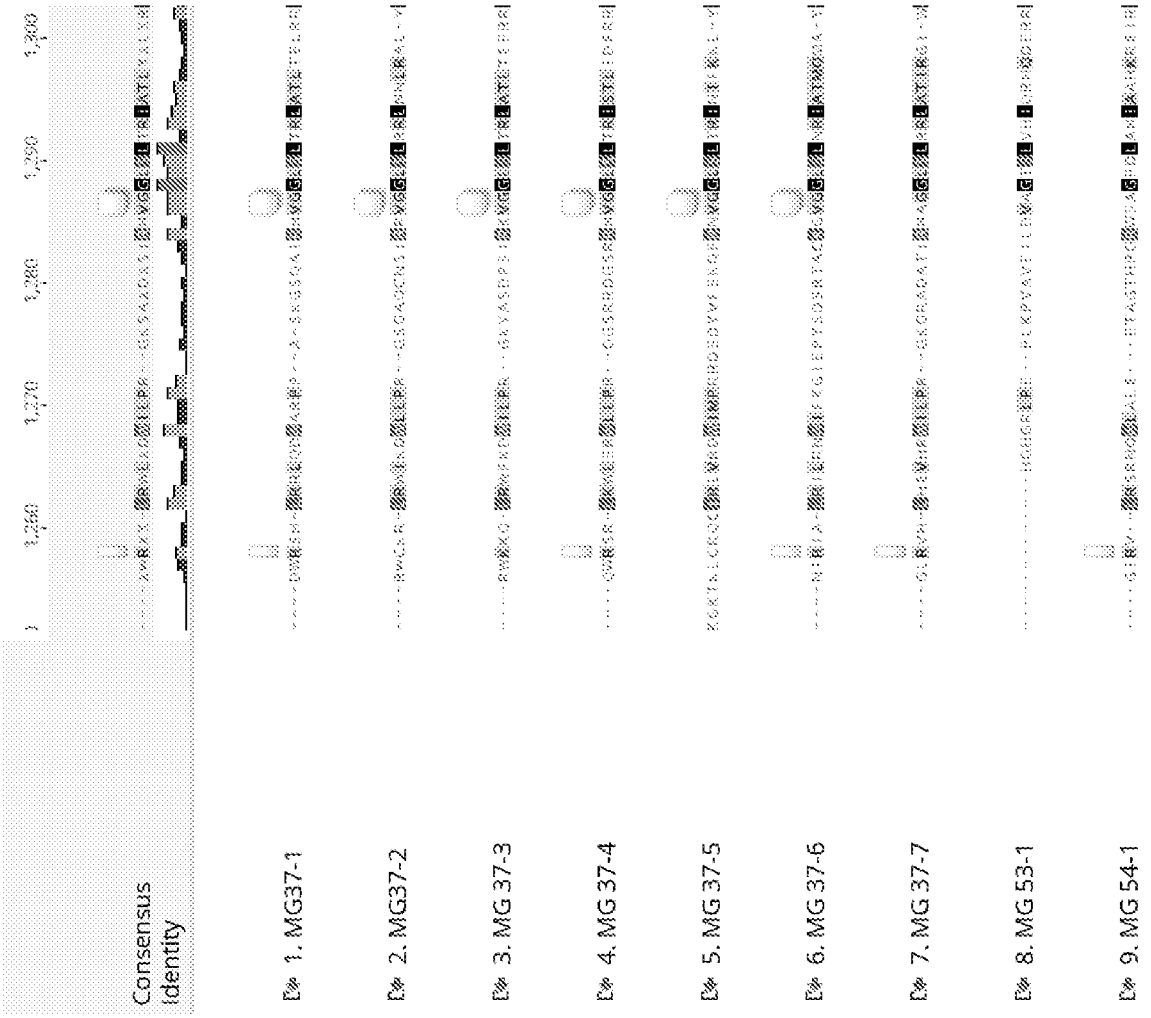

1
1,260
1,270
1,280
1,290
1,300
Consensus identity
1. MG37-1
2. MG37-2
3. MG 37-3
4. MG 37-4
5. MG 37-5
6. MG 37-6
7. MG 37-7
8. MG 53-1
9. MG 54-1

FIG. 37

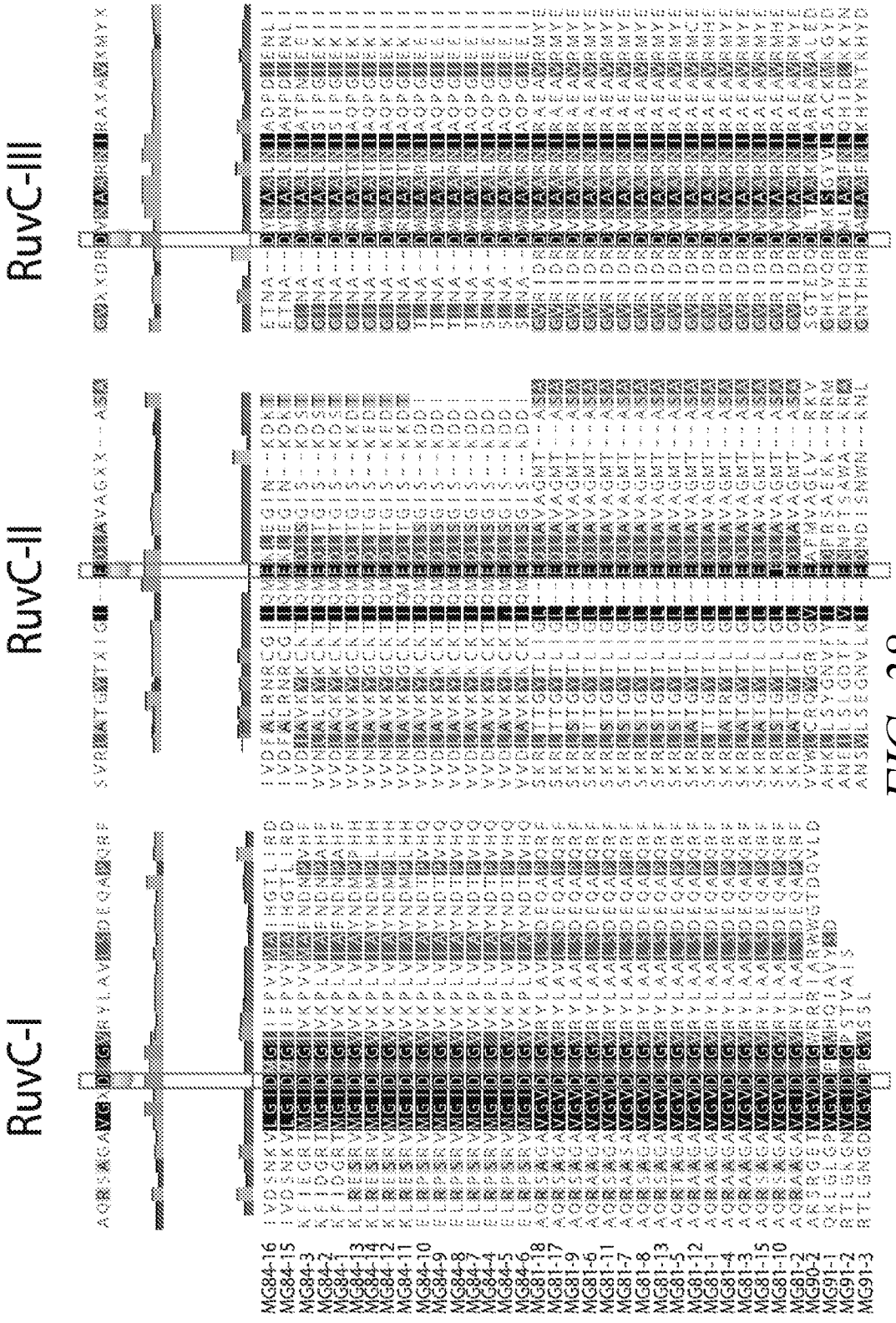

RuvC-I
RuvC-II
RuvC-III
MG84-16
MG84-15
MG84-3
MG84-2
MG84-1
MG84-13
MG84-14
MG84-12
MG84-11
MG84-10
MG84-9
MG84-8
MG84-7
MG84-4
MG84-5
MG84-6
MG81-18
MG81-17
MG81-9
MG81-6
MG81-11
MG81-7
MG81-8
MG81-13
MG81-5
MG81-12
MG81-1
MG81-4
MG81-3
MG81-15
MG81-10
MG81-2
MG90-2
MG91-1
MG91-2
MG91-3

FIG. 38

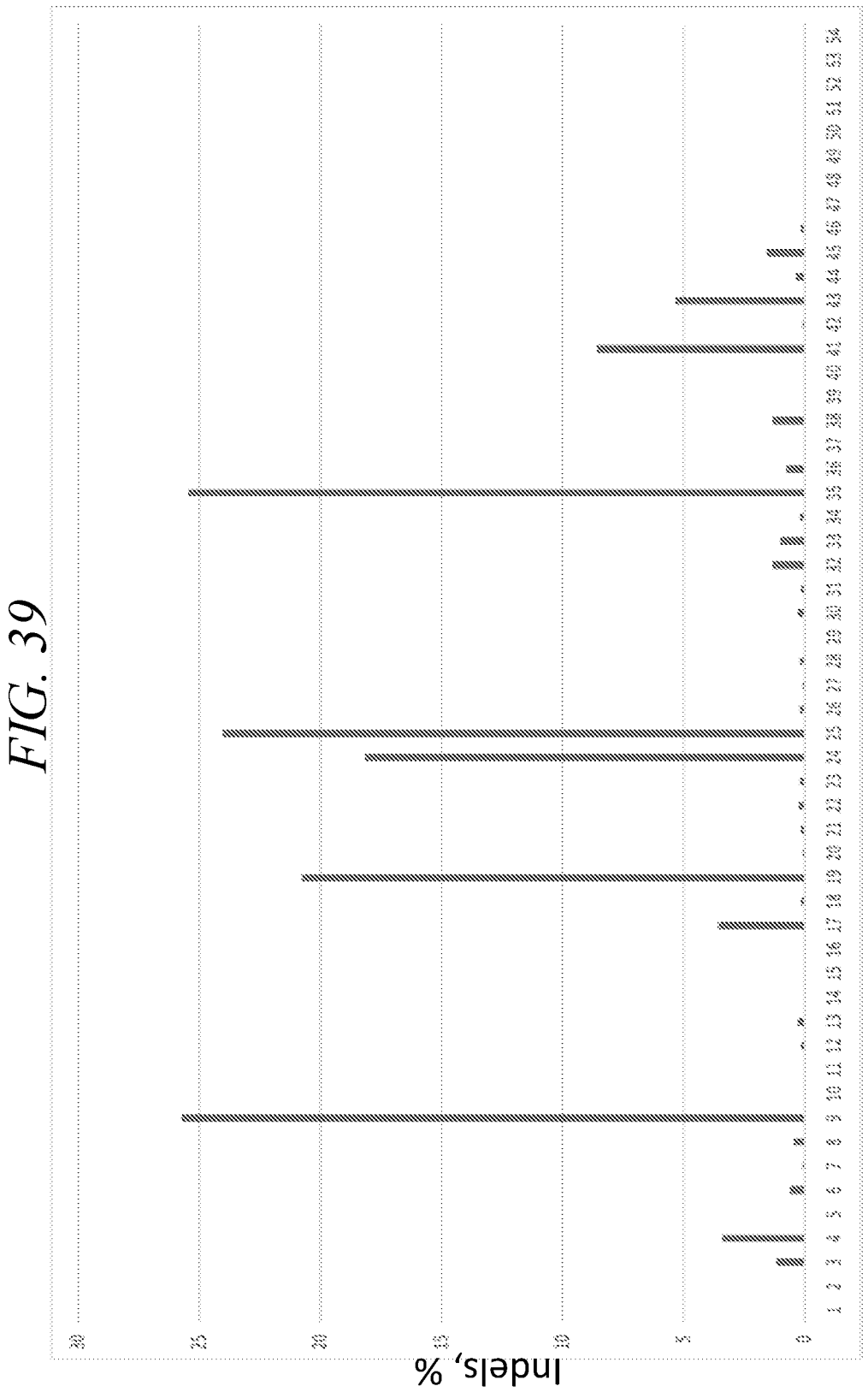


FIG. 39

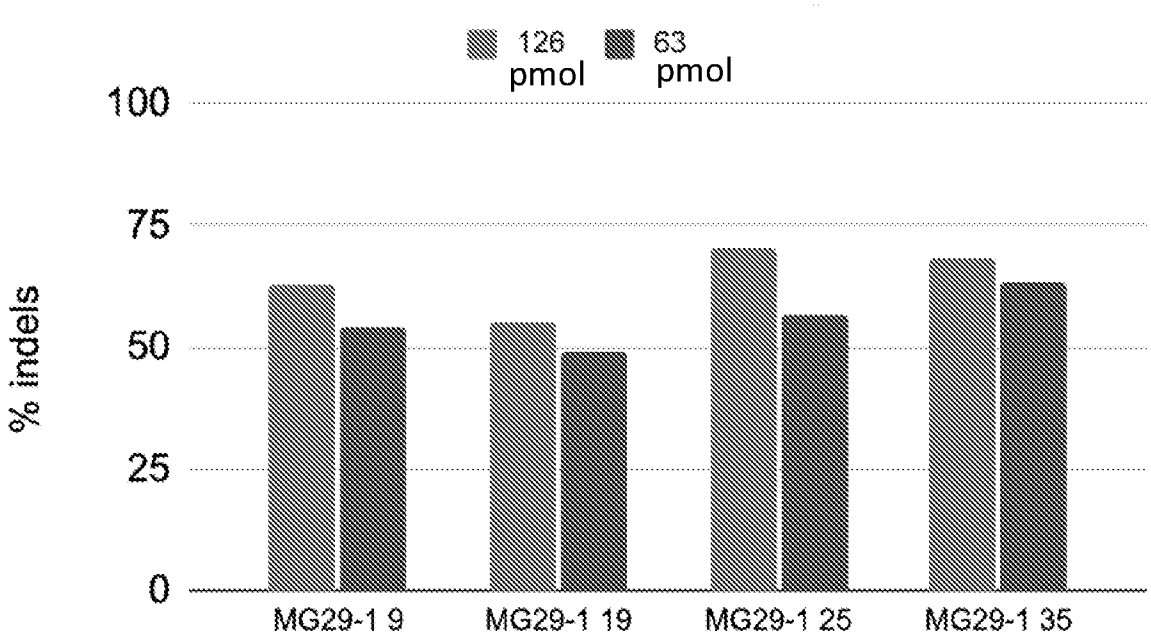

126 pmol
63 pmol
100
75
50
25
0
% indels
MG29-1 9
MG29-1 19
MG29-1 25
MG29-1 35

FIG. 40

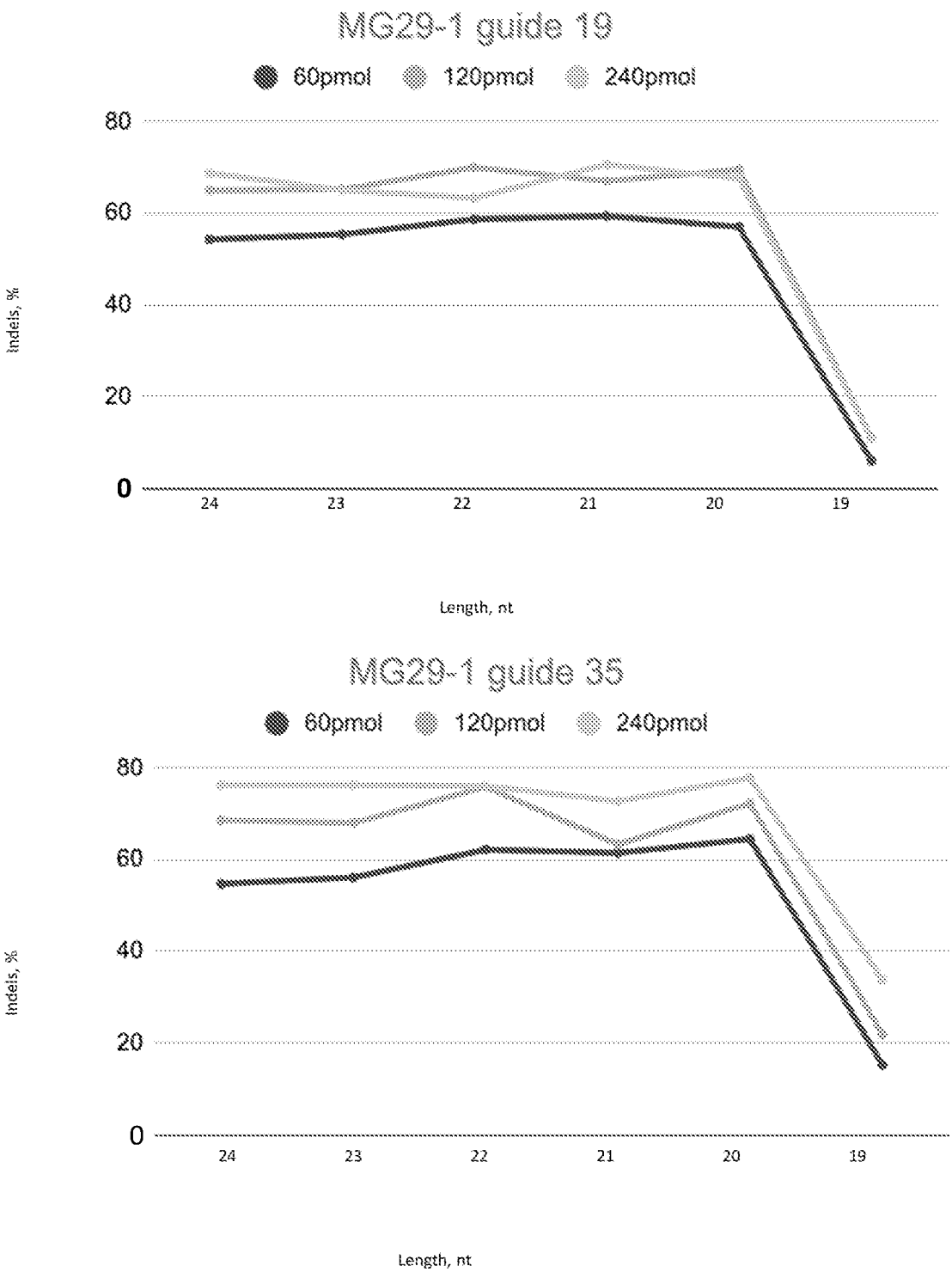

MG29-1 guide 19
60pmol
120pmol
240pmol
80
60
40
20
0
Indels, %
24
23
22
21
20
19
Length, nt
MG29-1 guide 35
60pmol
120pmol
240pmol
80
60
40
20
0
Indels, %
24
23
22
21
20
19
Length, nt

FIG. 41

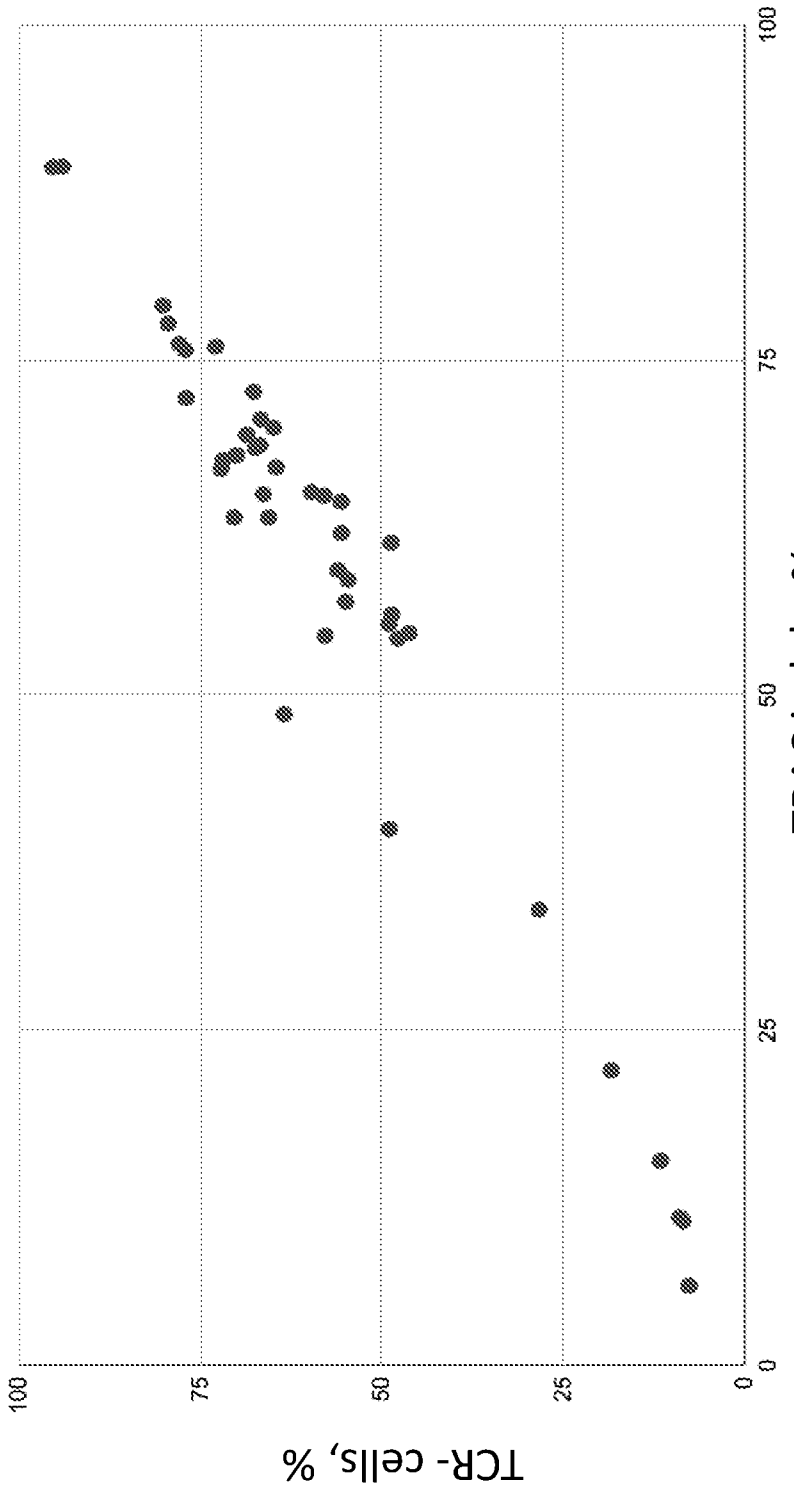


FIG. 42

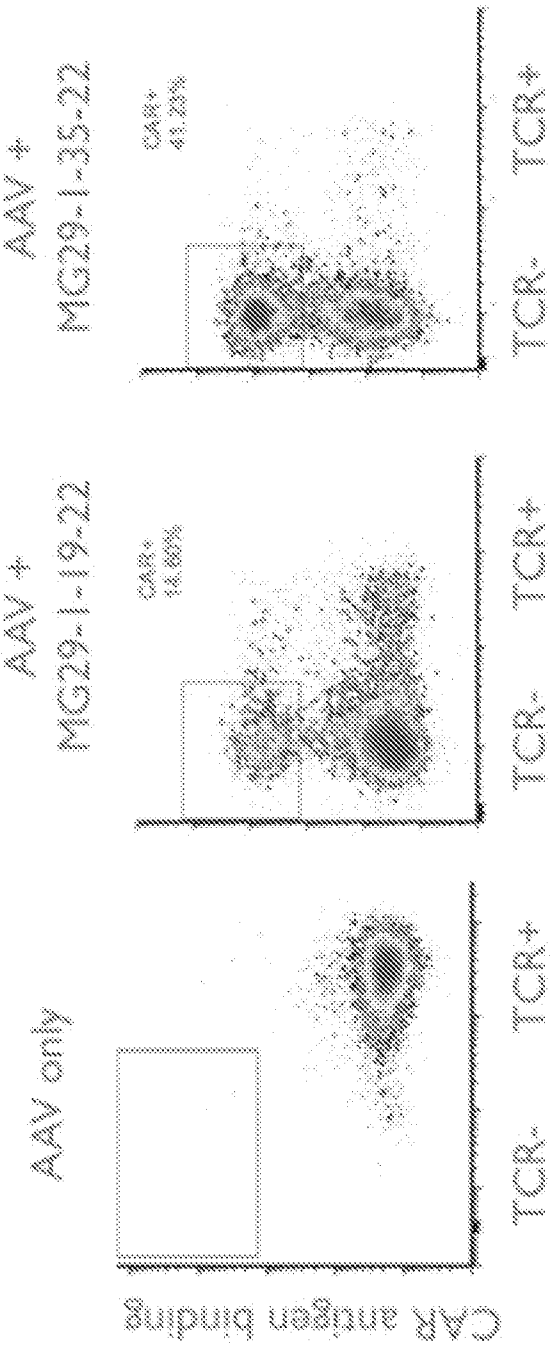

AAV only
AAV +
MG29-1-19-22
AAV +
MG29-1-35-22
CAR antigen binding
TCR-
TCR+
CAR+
44.60%
CAR+
41.23%

FIG. 43

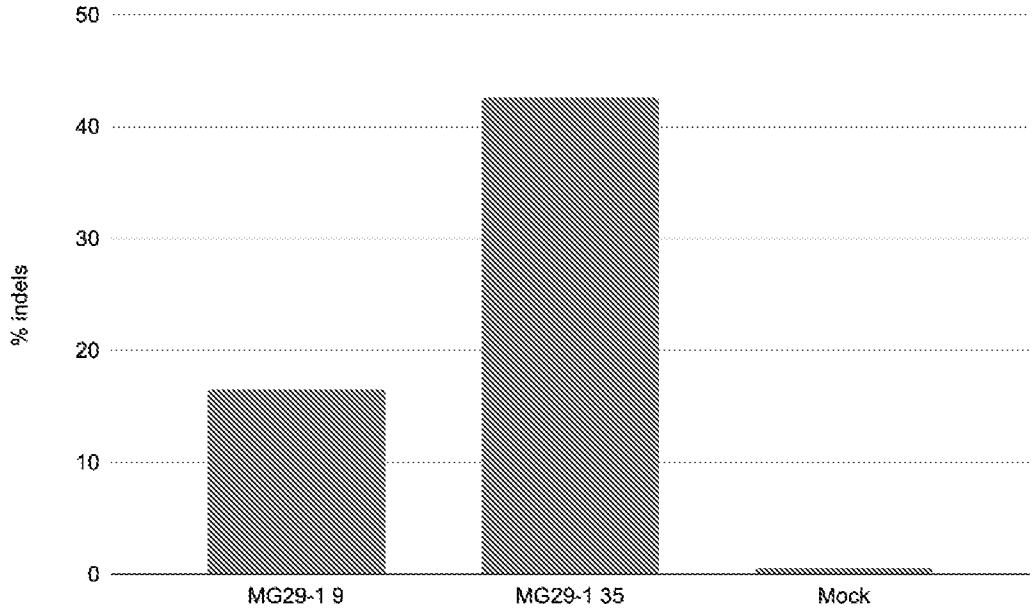

50
40
30
20
10
0
% indels
MG29-1 9
MG29-1 35
Mock

FIG. 44

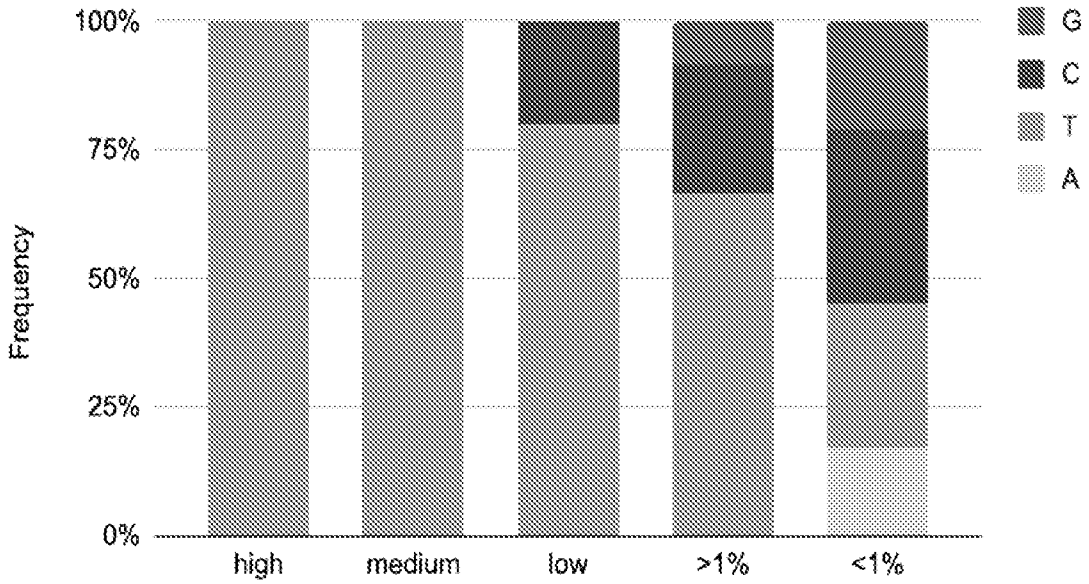

100%
75%
50%
25%
0%
Frequency
high
medium
low
>1%
<1%
G
C
T
A

FIG. 45

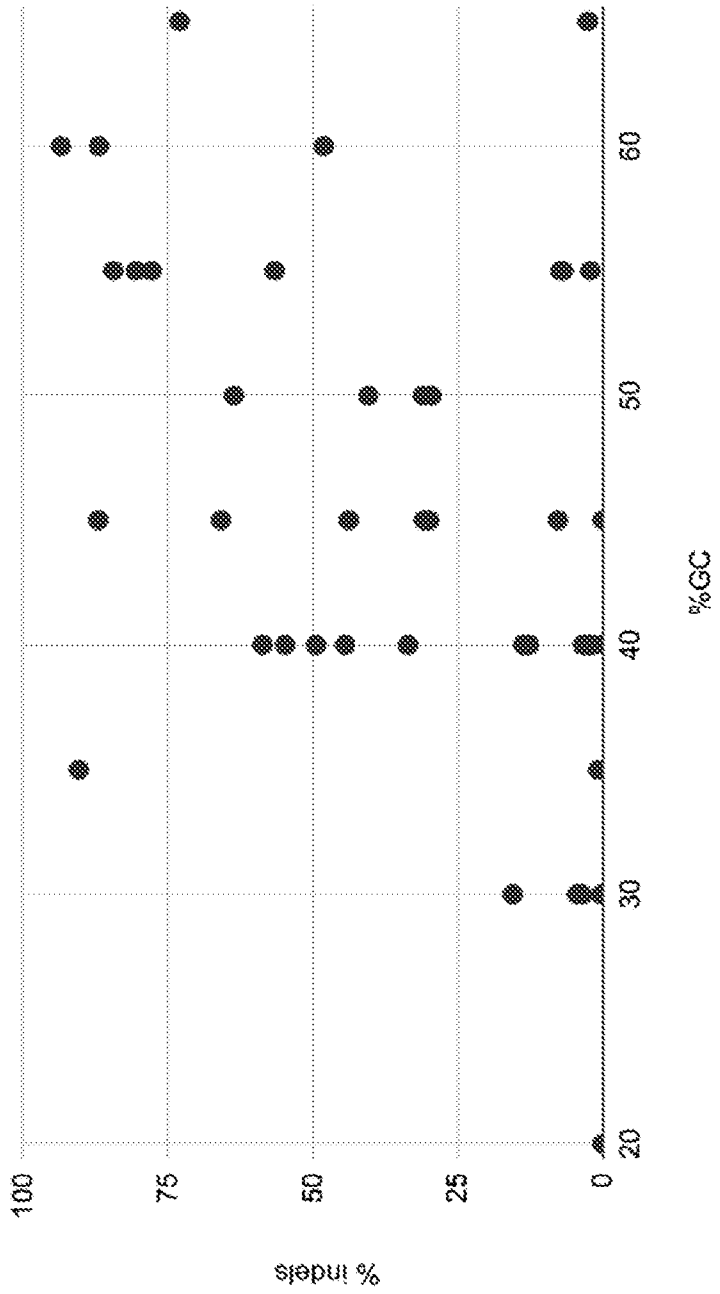


FIG. 46

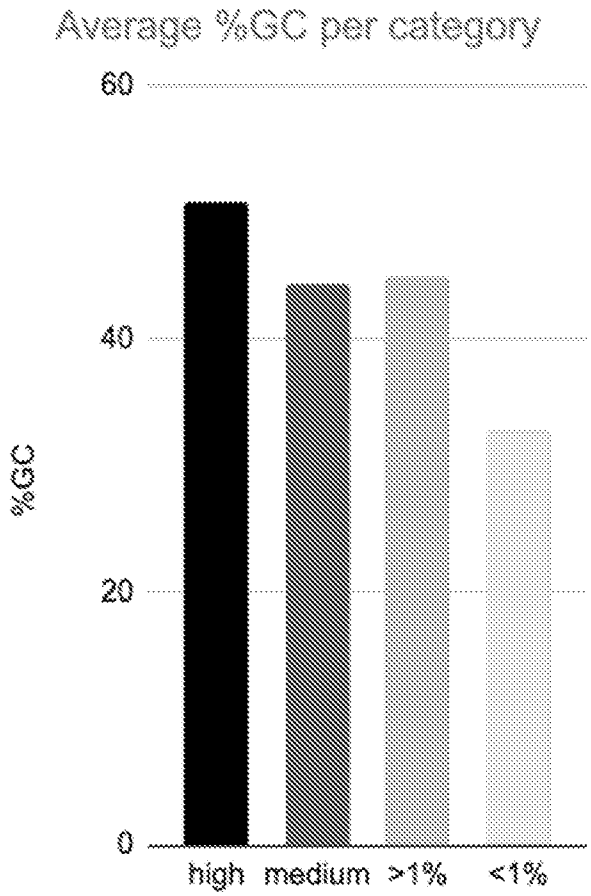


FIG. 46 (cont.)

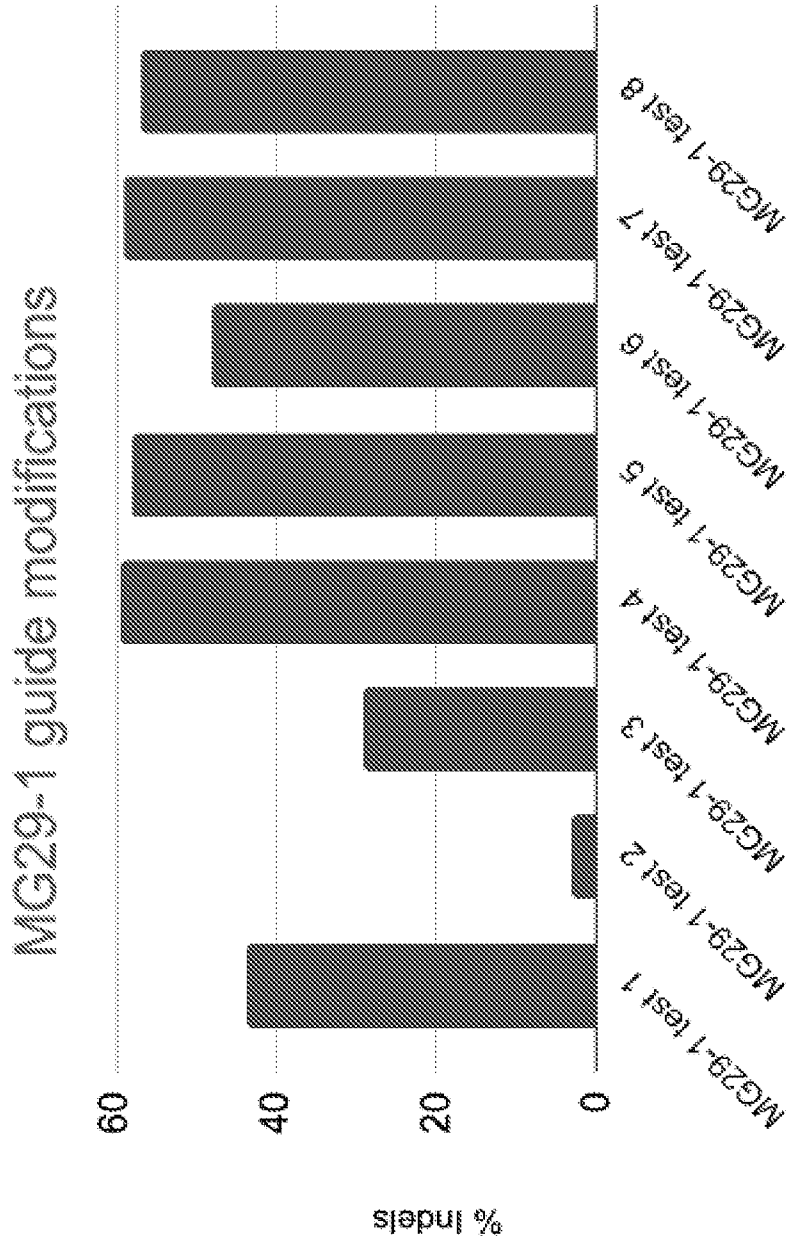

MG29-1 guide modifications
% Indels
0
20
40
60
MG29-1 test 1
MG29-1 test 2
MG29-1 test 3
MG29-1 test 4
MG29-1 test 5
MG29-1 test 6
MG29-1 test 7
MG29-1 test 8

FIG. 47

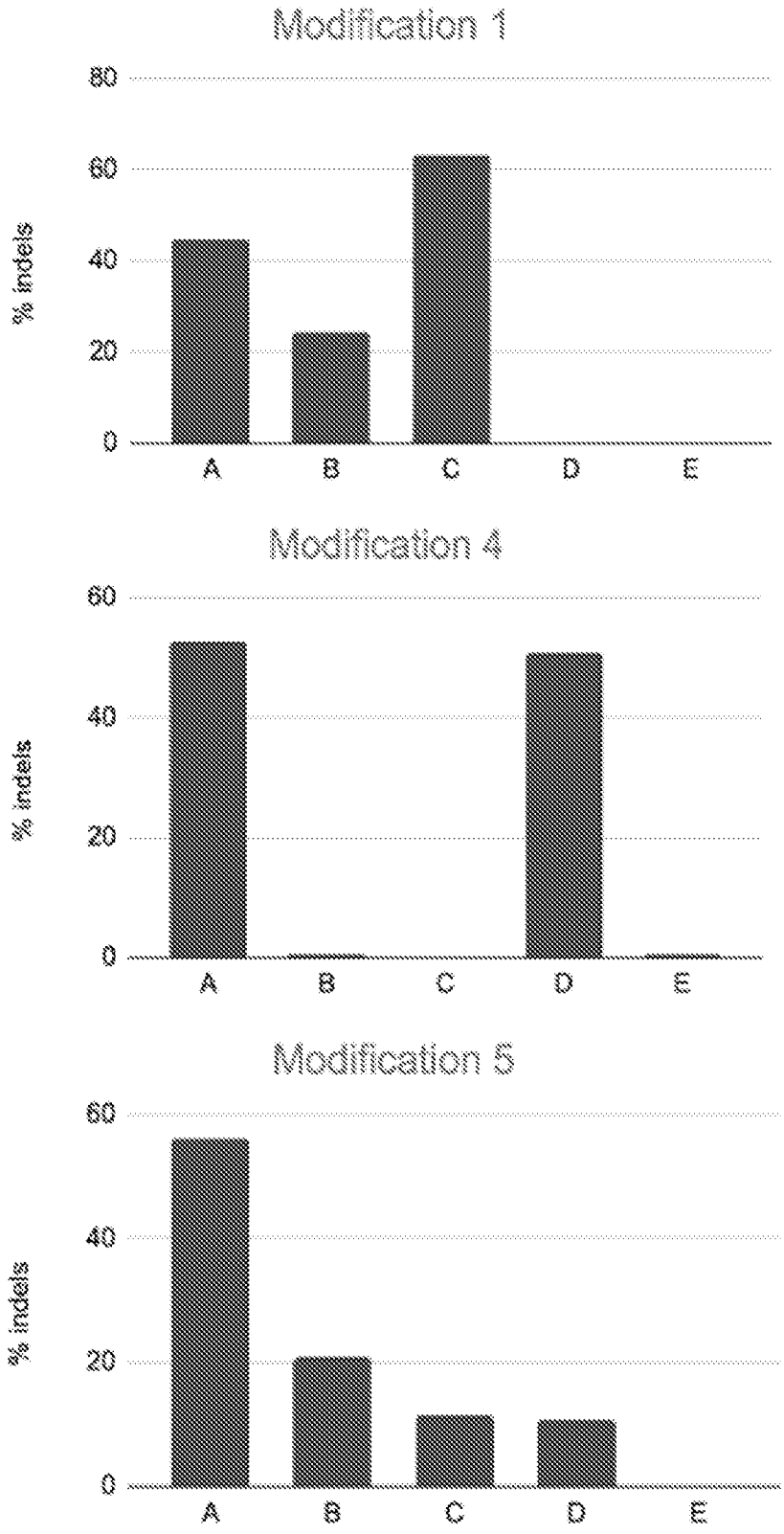

Modification 1
% indels
80
60
40
20
0
A
B
C
D
E
Modification 4
% indels
60
40
20
0
A
B
C
D
E
Modification 5
% indels
60
40
20
0
A
B
C
D
E

FIG. 48

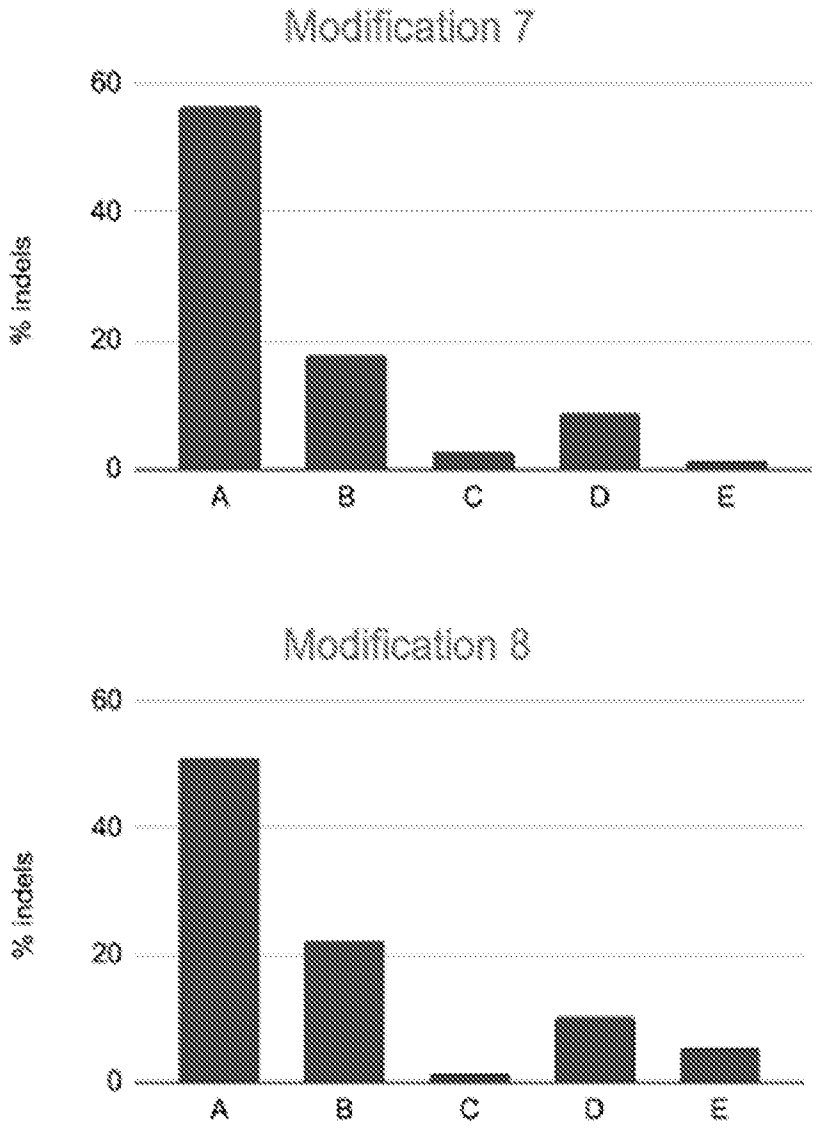

Modification 7
% indels
60
40
20
0
A
B
C
D
E
Modification 8
% indels
60
40
20
0
A
B
C
D
E

FIG. 48 (cont.)

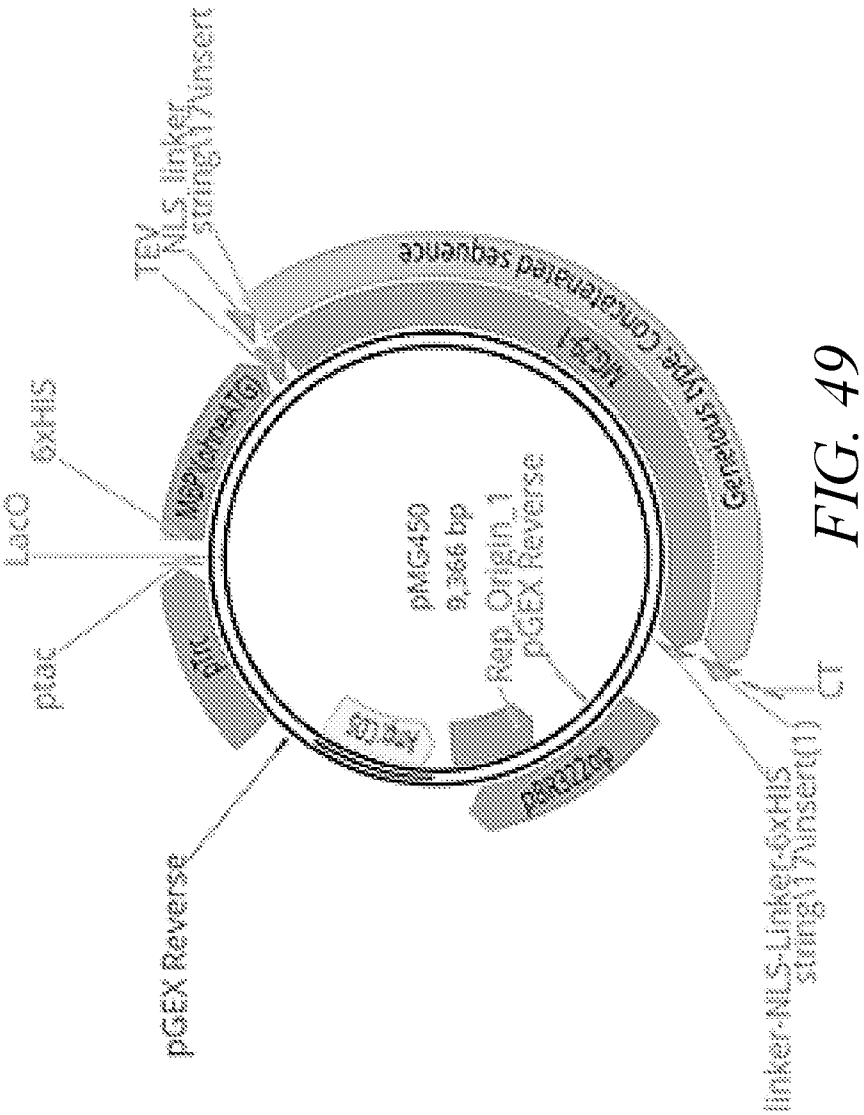

TEV
NLS linker
string\\7Ainsert
Cerevisiae type Consensus sequence
LacO
6xHIS
MBP (internal)
ptac
pMG450
9,346 bp
Rep Origin_1
pGEX Reverse
pGEX Reverse
linker-NLS-Linker-6xHIS
string\\1Ainsert(1)
CT

FIG. 49

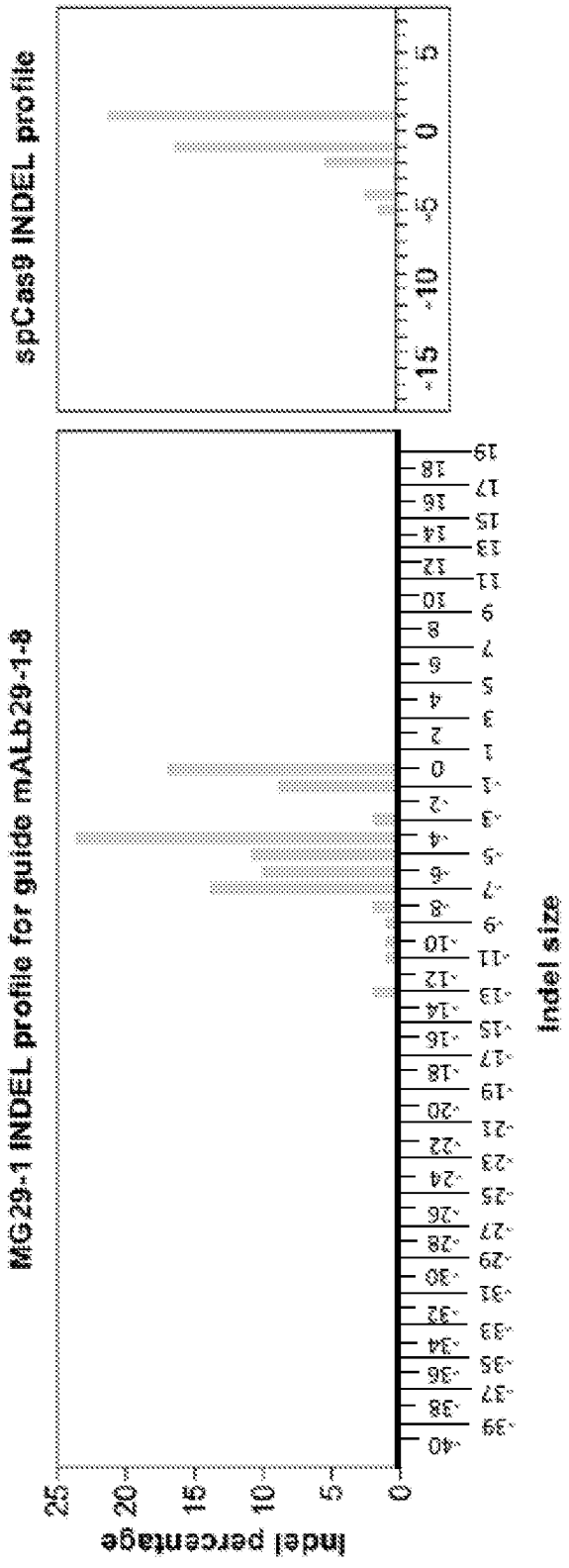

spCas9 INDEL profile
5
0
-5
-10
-15
MG29-1 INDEL profile for guide mALb29-1-8
indel percentage
25
20
15
10
5
0
indel size
-40
-39
-38
-37
-36
-35
-34
-33
-32
-31
-30
-29
-28
-27
-26
-25
-24
-23
-22
-21
-20
-19
-18
-17
-16
-15
-14
-13
-12
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19

FIG. 50

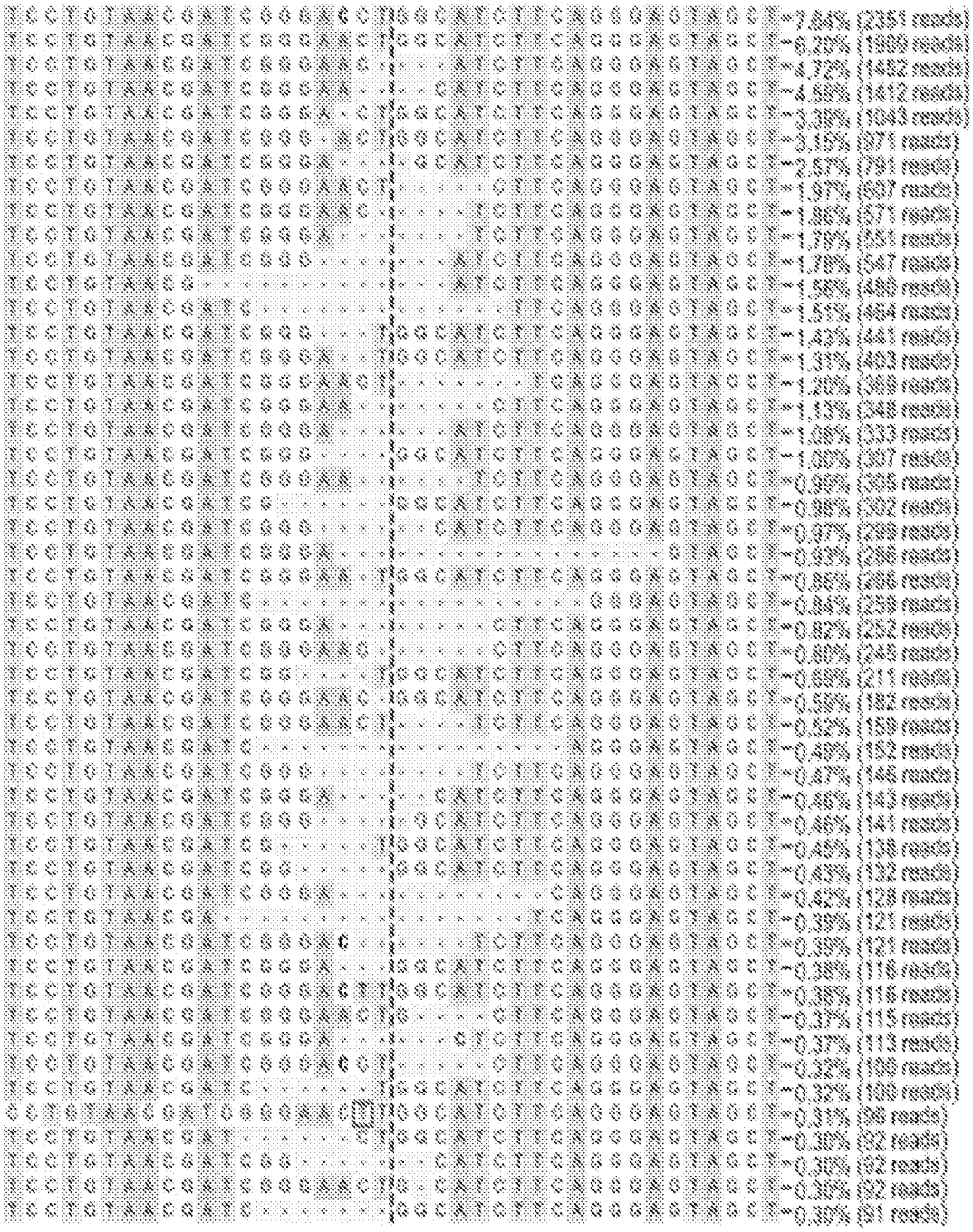

~7.84% (2351 reads)
~6.20% (1909 reads)
~4.72% (1452 reads)
~4.58% (1412 reads)
~3.38% (1043 reads)
~3.15% (971 reads)
~2.57% (791 reads)
~1.97% (607 reads)
~1.86% (571 reads)
~1.79% (551 reads)
~1.78% (547 reads)
~1.56% (480 reads)
~1.51% (464 reads)
~1.43% (441 reads)
~1.31% (403 reads)
~1.20% (369 reads)
~1.13% (348 reads)
~1.08% (333 reads)
~1.00% (307 reads)
~0.99% (305 reads)
~0.98% (302 reads)
~0.97% (299 reads)
~0.93% (286 reads)
~0.86% (266 reads)
~0.84% (259 reads)
~0.82% (252 reads)
~0.80% (245 reads)
~0.68% (211 reads)
~0.59% (182 reads)
~0.52% (159 reads)
~0.49% (152 reads)
~0.47% (146 reads)
~0.46% (143 reads)
~0.46% (141 reads)
~0.45% (138 reads)
~0.43% (132 reads)
~0.42% (128 reads)
~0.39% (121 reads)
~0.39% (121 reads)
~0.38% (116 reads)
~0.38% (116 reads)
~0.37% (115 reads)
~0.37% (113 reads)
~0.32% (100 reads)
~0.32% (100 reads)
~0.31% (96 reads)
~0.30% (92 reads)
~0.30% (92 reads)
~0.30% (92 reads)
~0.30% (91 reads)

FIG. 51

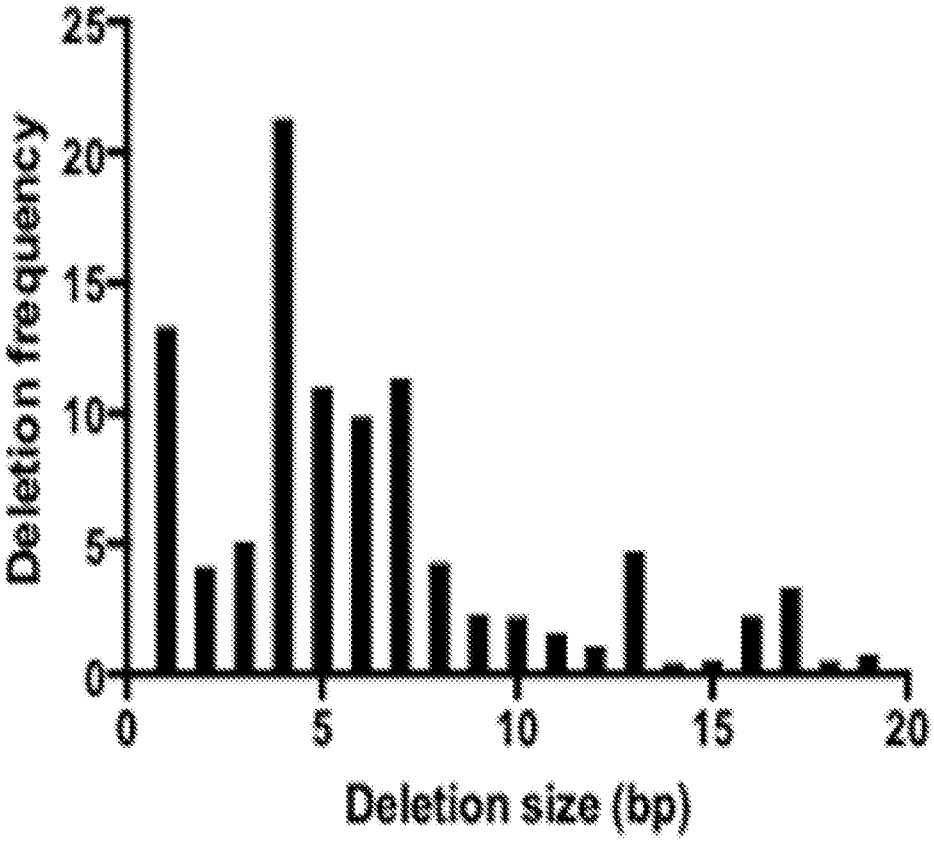


FIG. 51 (cont.)

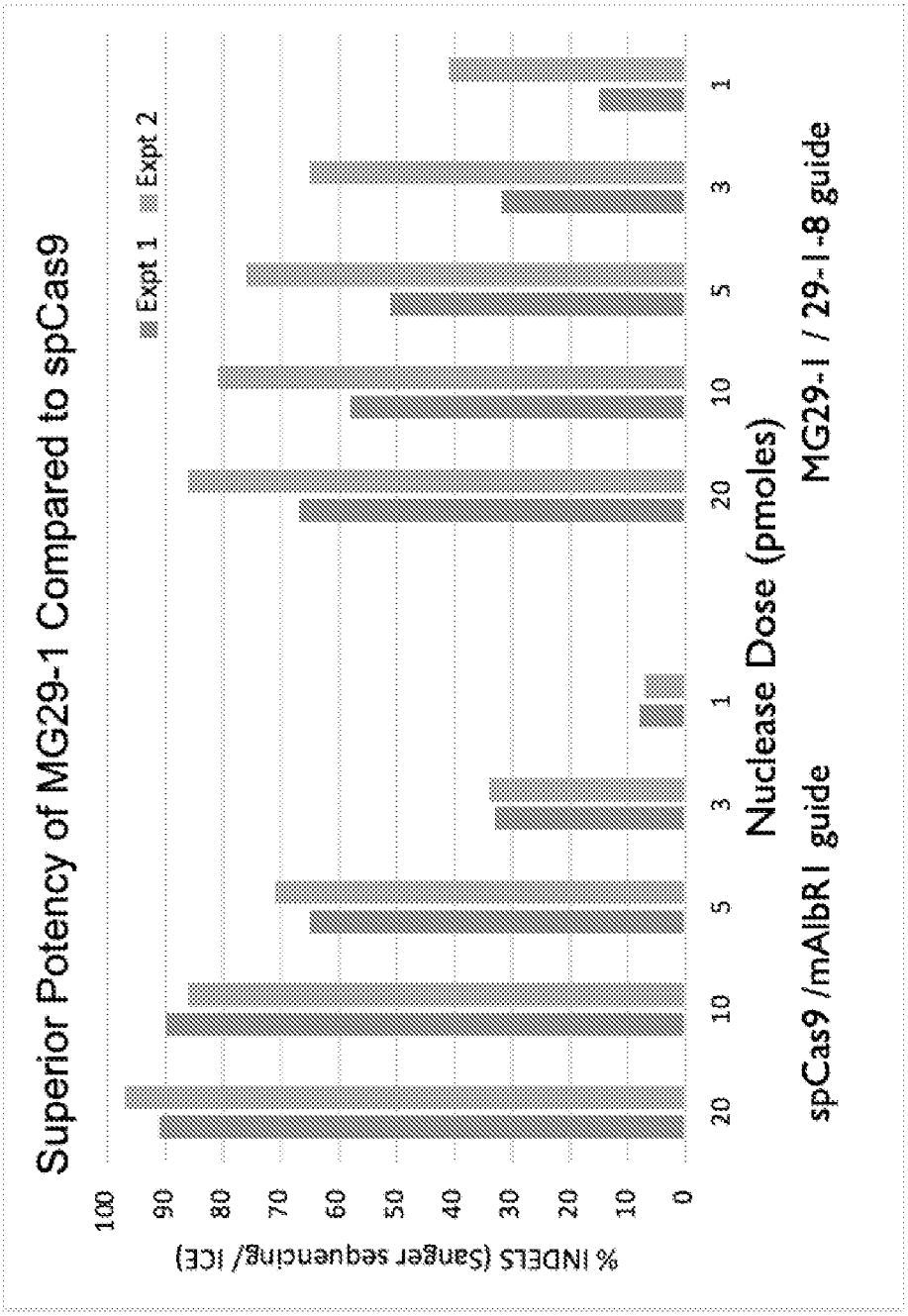

Superior Potency of MG29-1 Compared to spCas9
Expt 1
Expt 2
% INDELS (Sanger sequencing/ ICE)
100
90
80
70
60
50
40
30
20
10
0
20
10
5
3
1
20
10
5
3
1
Nuclease Dose (pmoles)
spCas9 /mAlbR1 guide
MG29-1 / 29-1-8 guide

FIG. 52

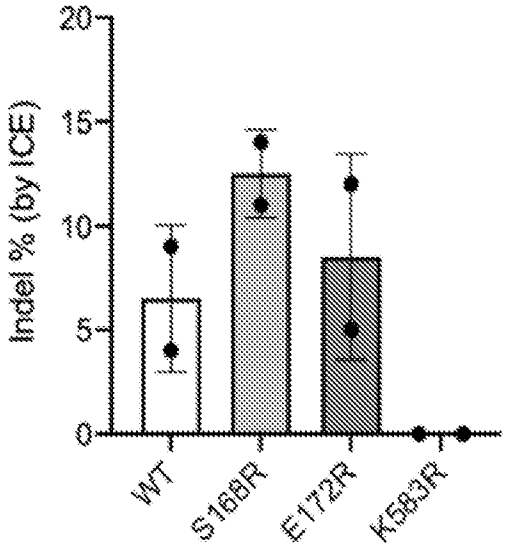


FIG. 53A

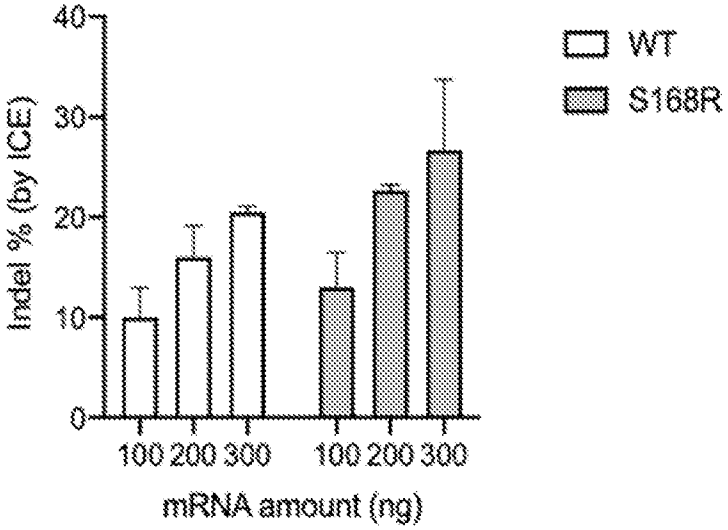


FIG. 53B

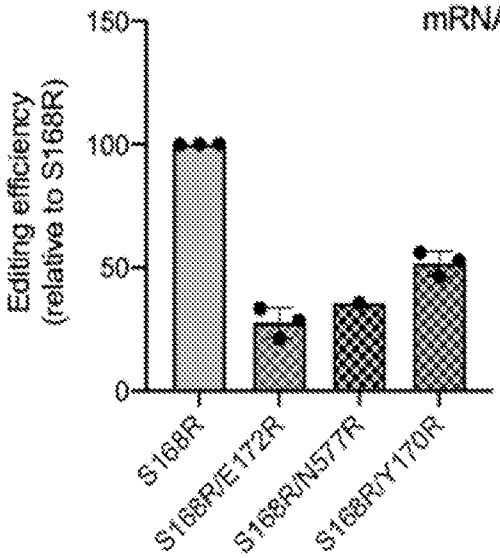


FIG. 53C

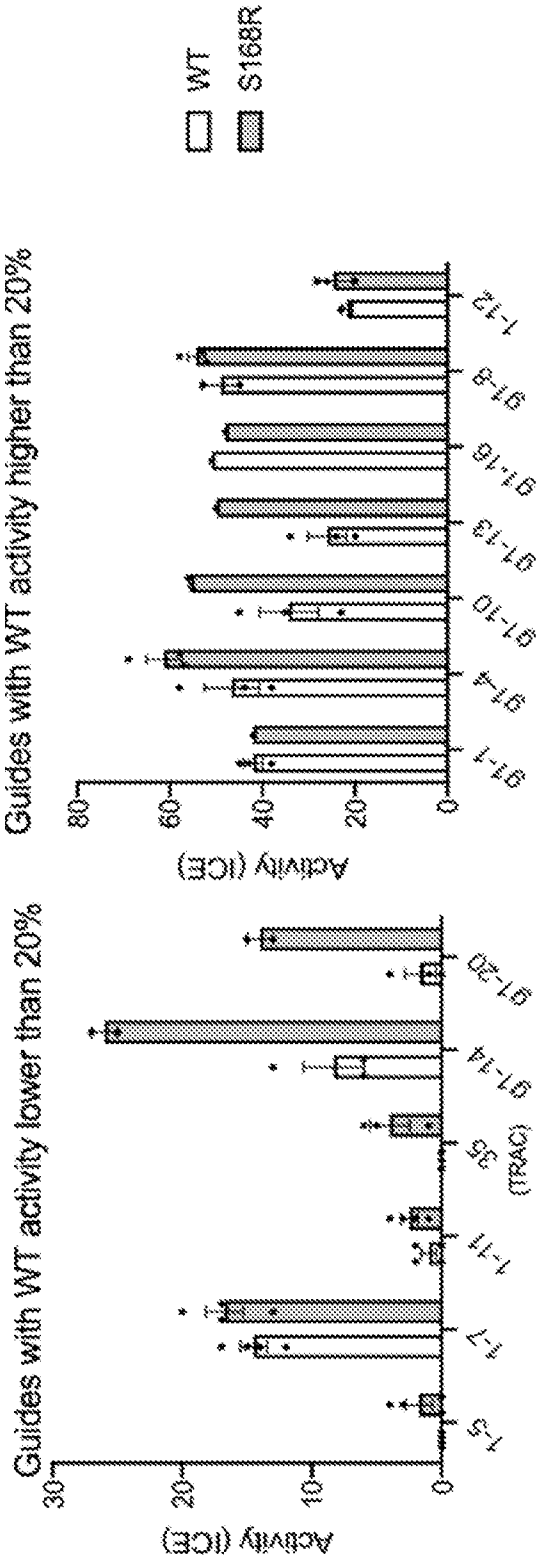

WT
S168R
Guides with WT activity higher than 20%
Activity (ICE)
80
60
40
20
0
g1-1
g1-4
g1-10
g1-13
g1-16
g1-9
1-12
Guides with WT activity lower than 20%
Activity (ICE)
30
20
10
0
1-5
1-7
1-11
35
(TRAC)
g1-14
g1-20

FIG. 53D

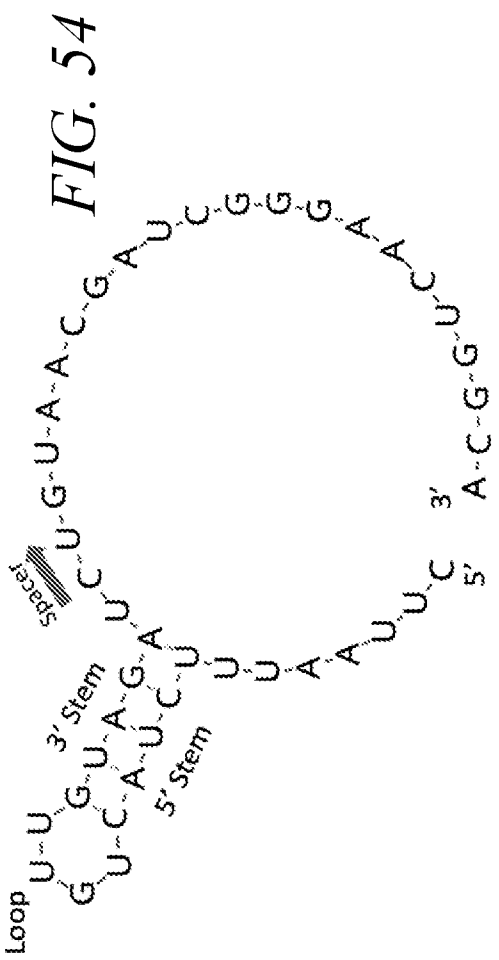

Loop
3' Stem
5' Stem
spacer
5'
3'

FIG. 54

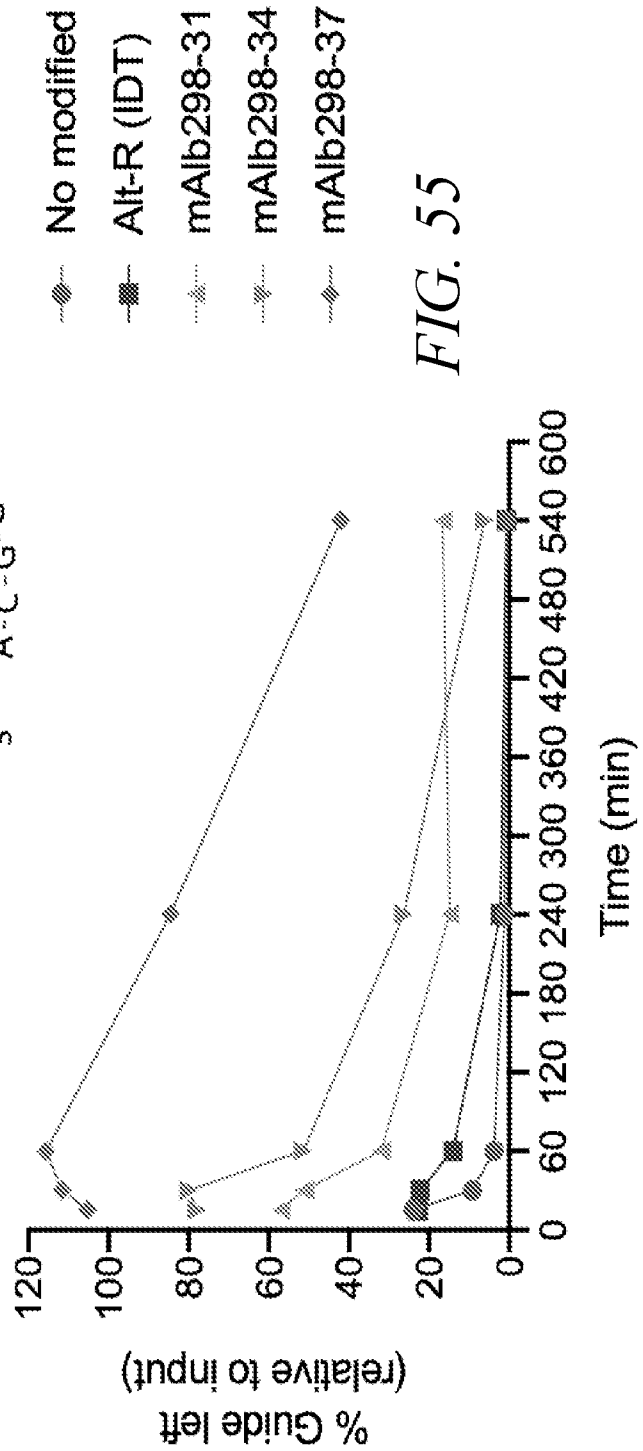

No modified
Alt-R (IDT)
mAlb298-31
mAlb298-34
mAlb298-37
% Guide left
(relative to input)
120
100
80
60
40
20
0
0
60
120
180
240
300
360
420
480
540
600
Time (min)

FIG. 55

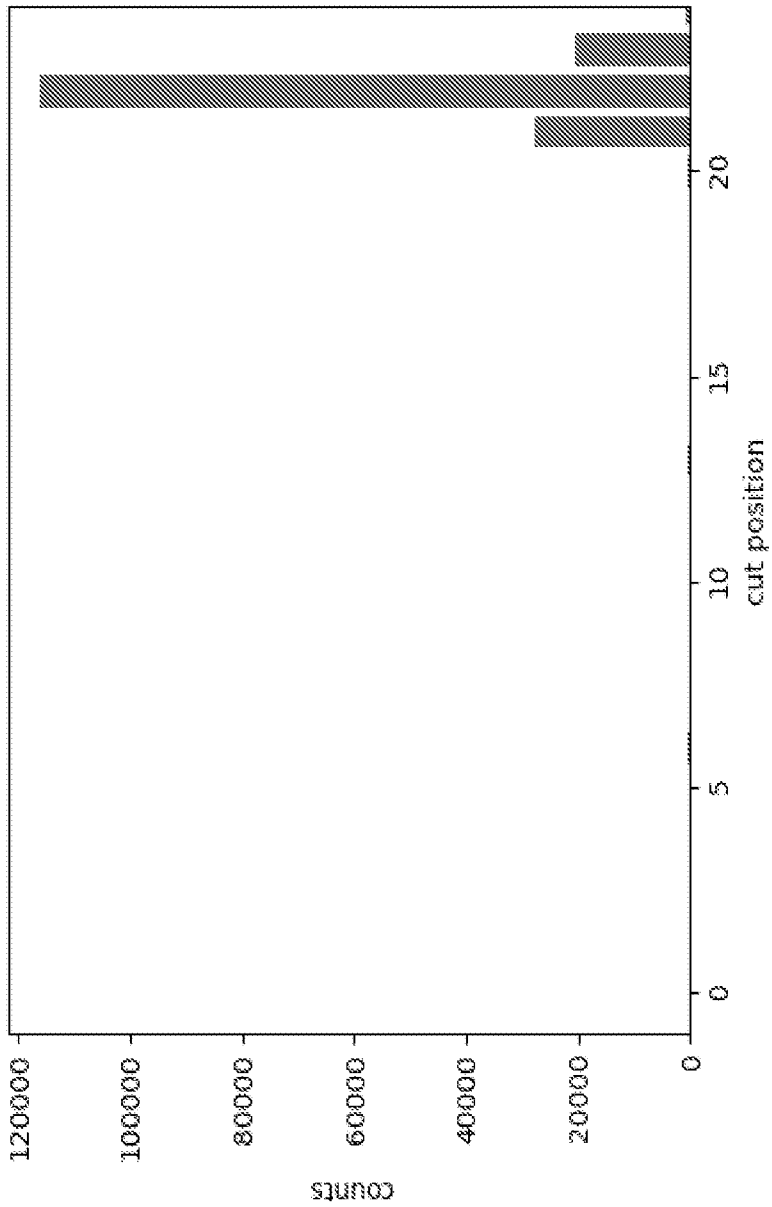

counts
120000
100000
80000
60000
40000
20000
0
0
5
10
15
20
cut position

FIG. 56A

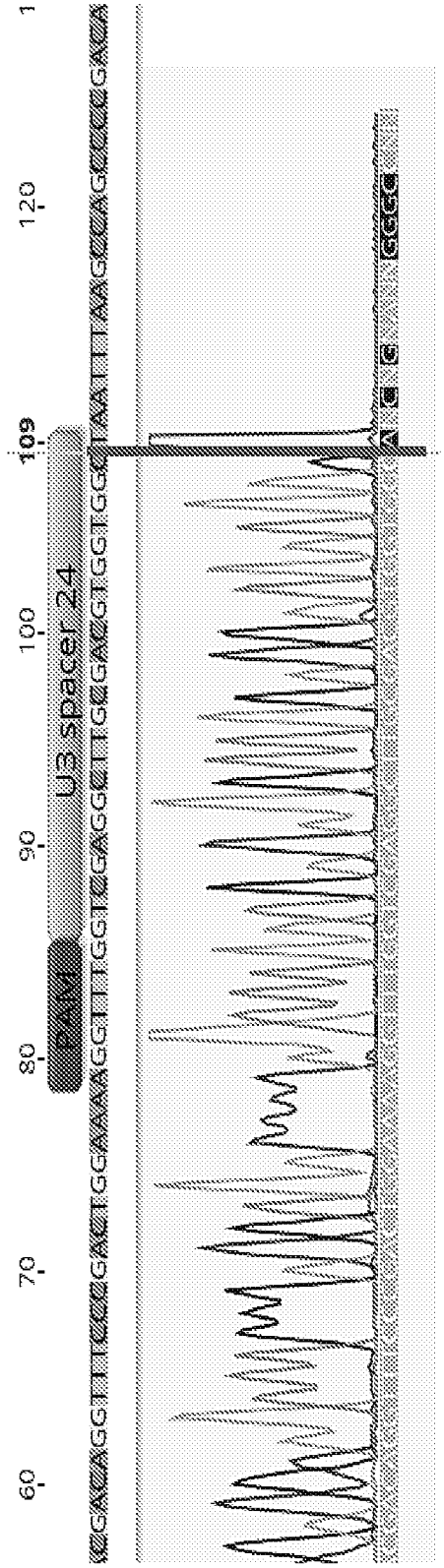


FIG. 56B

FIG. 56C

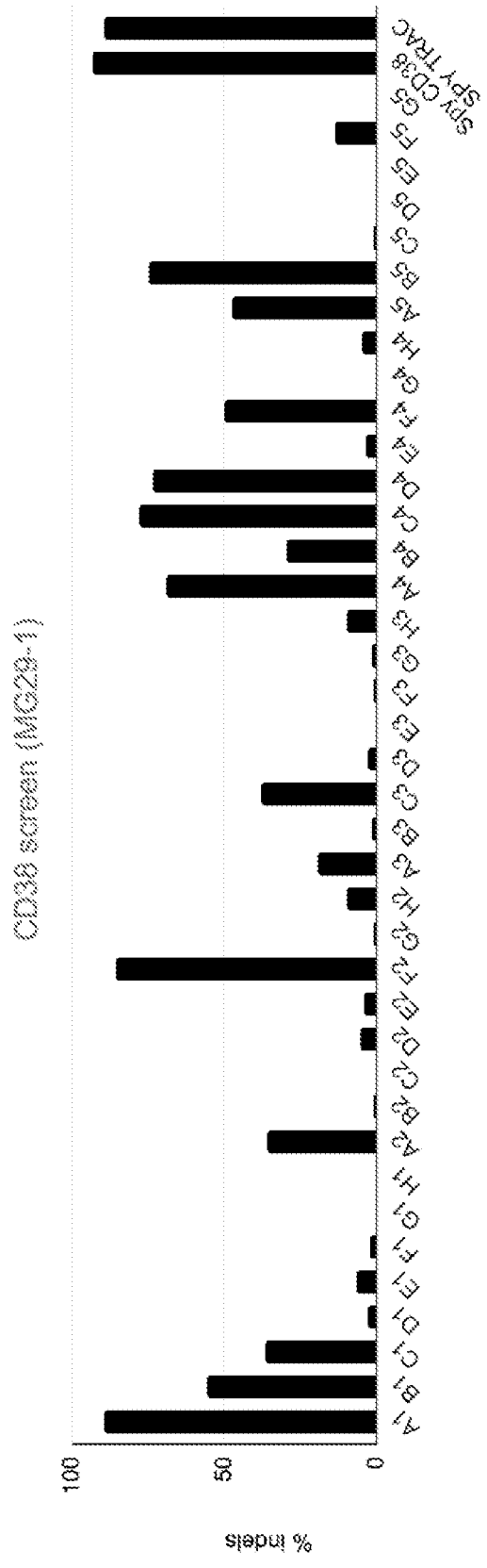

CD38 screen (MG28-1)
% Indels
100
50
0
A1
B1
C1
D1
E1
F1
G1
H1
A2
B2
C2
D2
E2
F2
G2
H2
A3
B3
C3
D3
E3
F3
G3
H3
A4
B4
C4
D4
E4
F4
G4
H4
A5
B5
C5
D5
E5
F5
G5
SpyCD38 SpyTRAC

FIG. 57

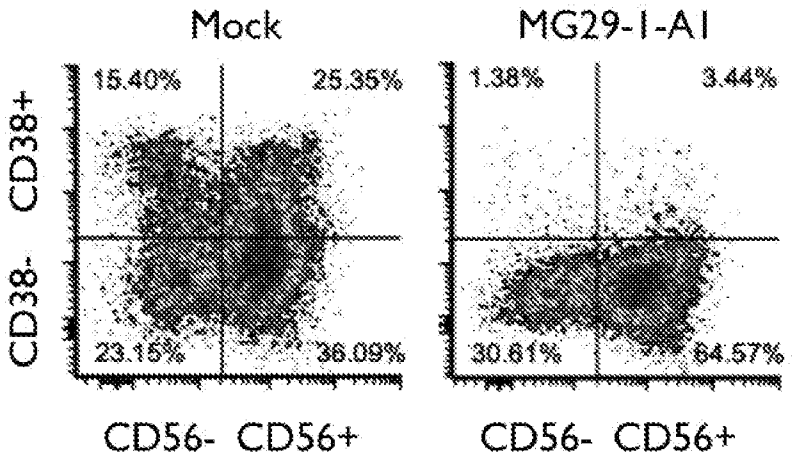

Mock
MG29-1-A1
15.40%
25.35%
23.15%
36.09%
1.38%
3.44%
30.61%
64.57%
CD38+
CD38-
CD56-
CD56+
CD56-
CD56+

FIG. 58

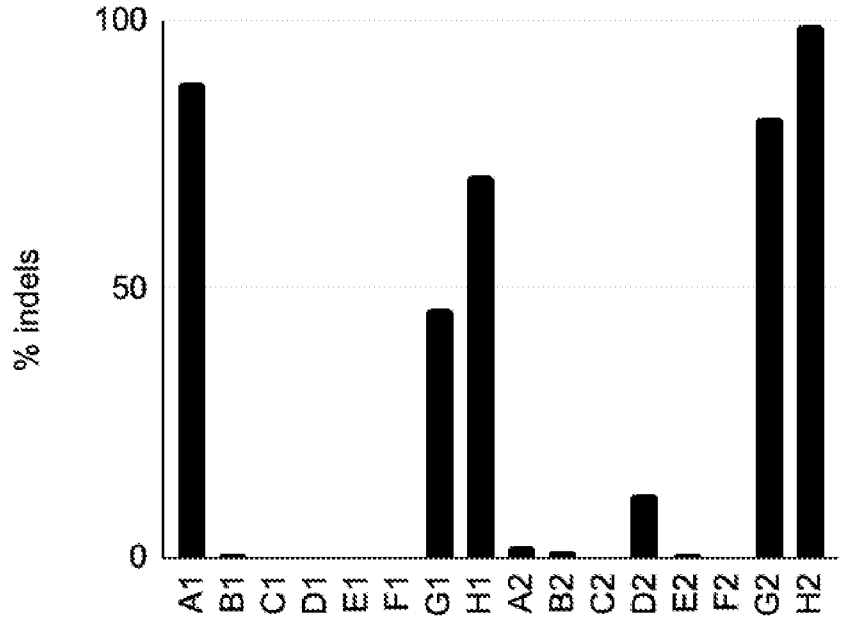

100
50
0
% indels
A1
B1
C1
D1
E1
F1
G1
H1
A2
B2
C2
D2
E2
F2
G2
H2

FIG. 59

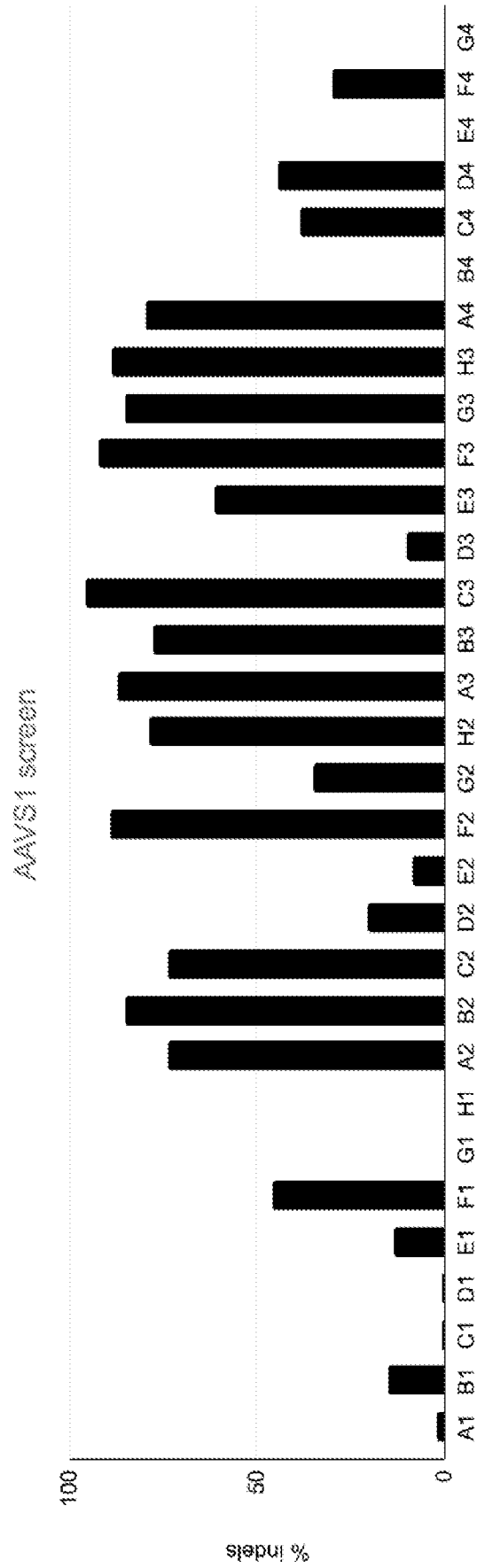

AAVS1 screen
% indels
100
50
0
A1 B1 C1 D1 E1 F1 G1 H1 A2 B2 C2 D2 E2 F2 G2 H2 A3 B3 C3 D3 E3 F3 G3 H3 A4 B4 C4 D4 E4 F4 G4

FIG. 60

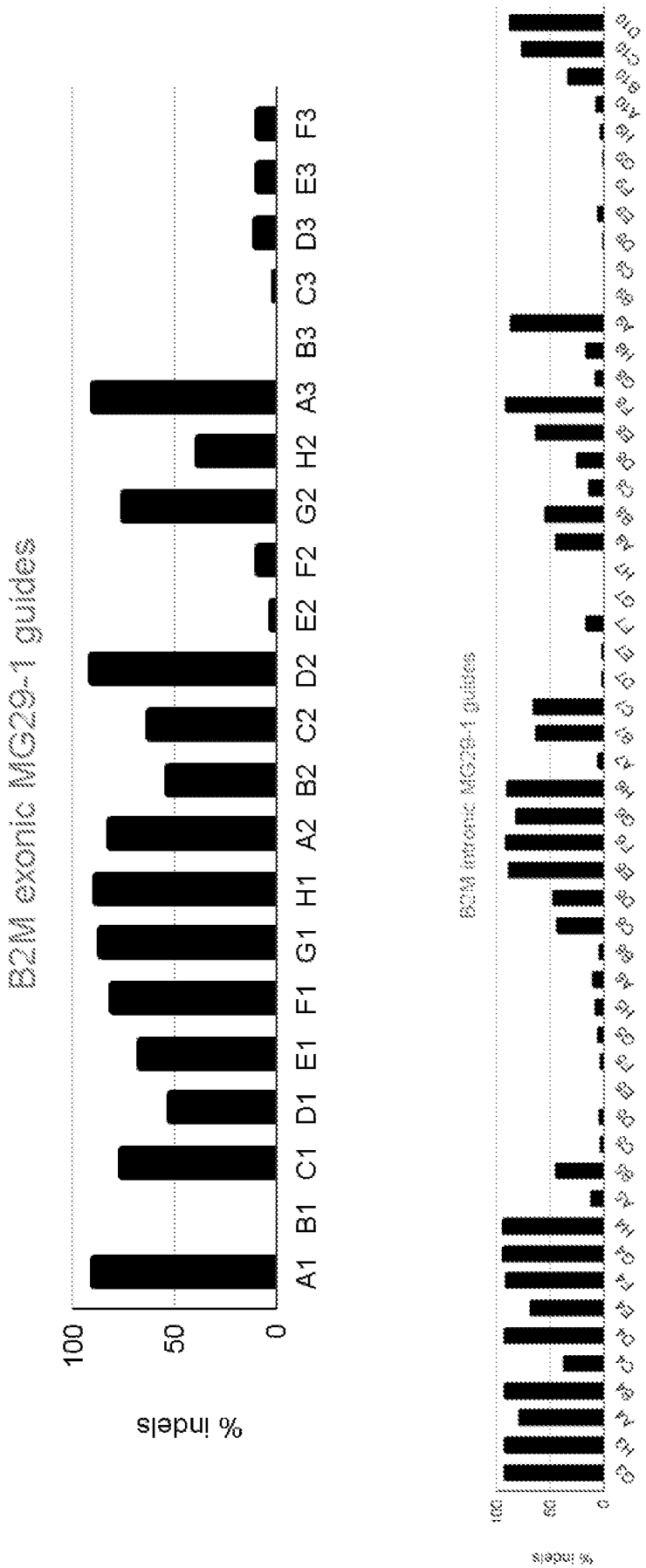

B2M exonic MG29-1 guides
% Indels
100
50
0
A1 B1 C1 D1 E1 F1 G1 H1 A2 B2 C2 D2 E2 F2 G2 H2 A3 B3 C3 D3 E3 F3

FIG. 61

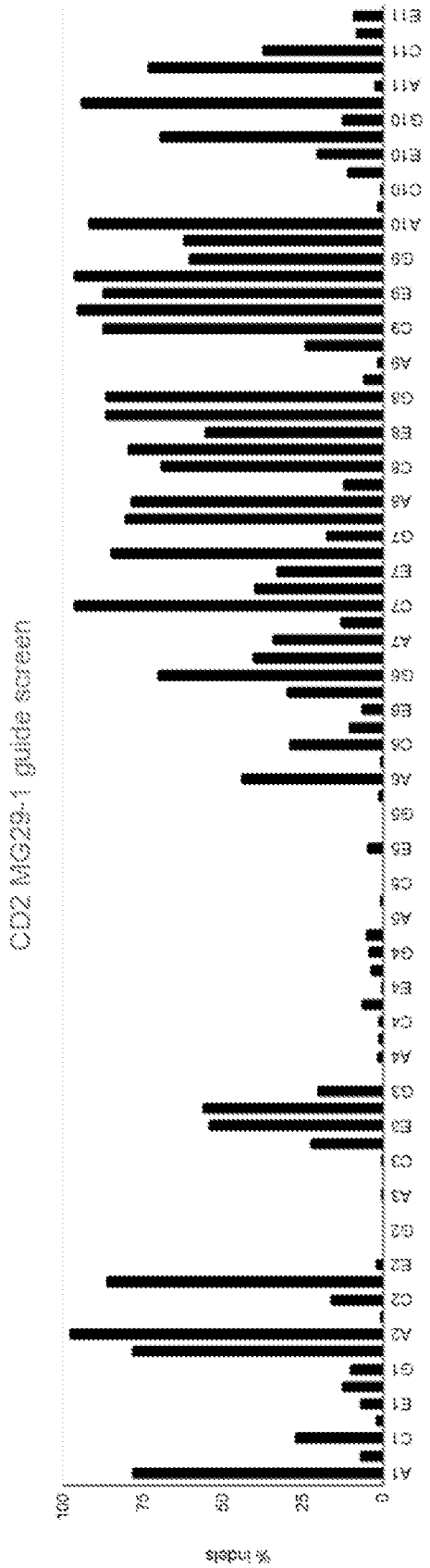

CD2 MG29-1 guide screen
% Indels
100
75
50
25
0
A1
C1
E1
G1
A2
C2
E2
G2
A3
C3
E3
G3
A4
C4
E4
G4
A5
C5
E5
G5
A6
C6
E6
G6
A7
C7
E7
G7
A8
C8
E8
G8
A9
C9
E9
G9
A10
C10
E10
G10
A11
C11
E11
G11

FIG. 62

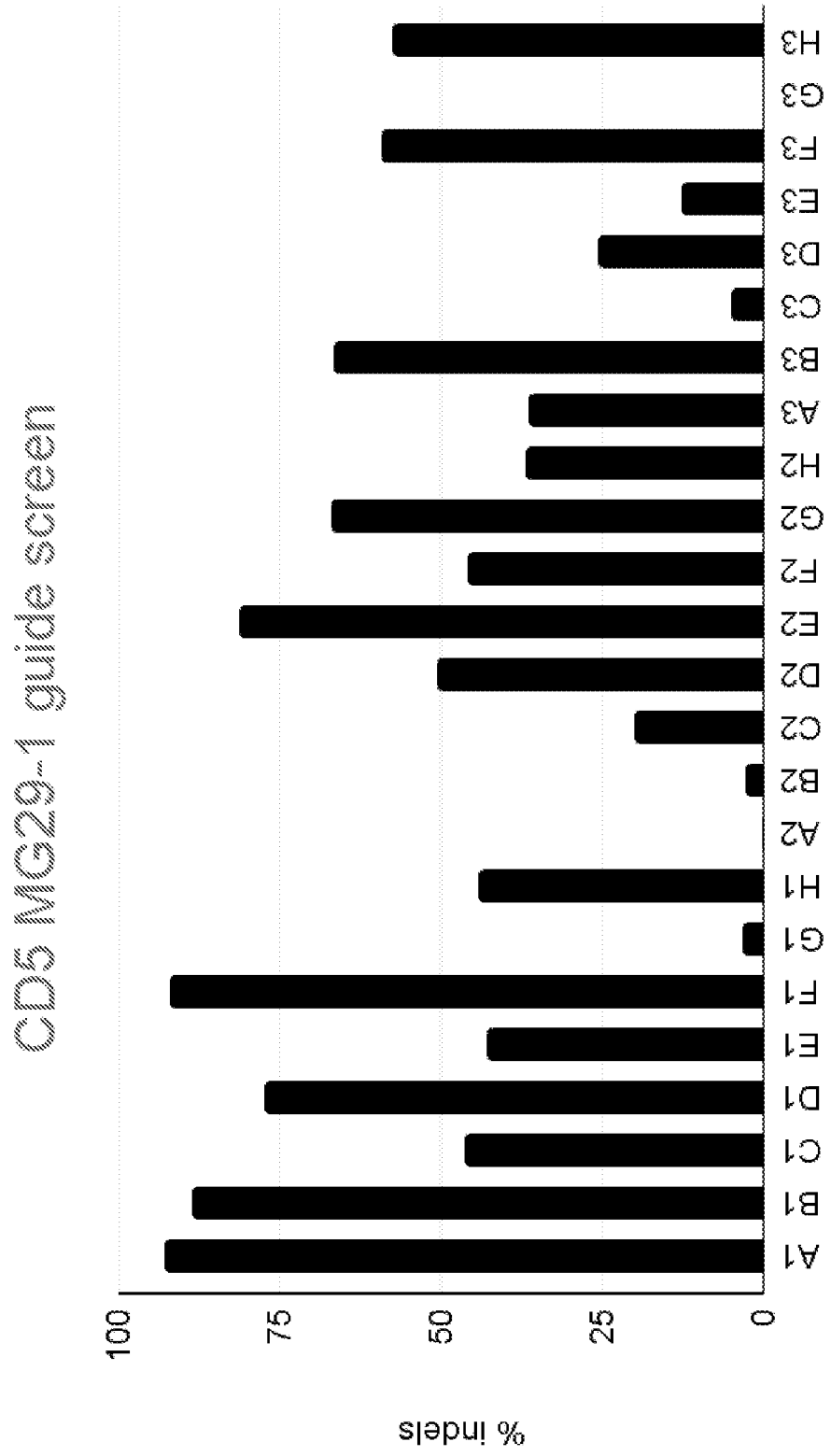

CD5 MG29-1 guide screen
% indels
100
75
50
25
0
A1
B1
C1
D1
E1
F1
G1
H1
A2
B2
C2
D2
E2
F2
G2
H2
A3
B3
C3
D3
E3
F3
G3
H3

FIG. 63

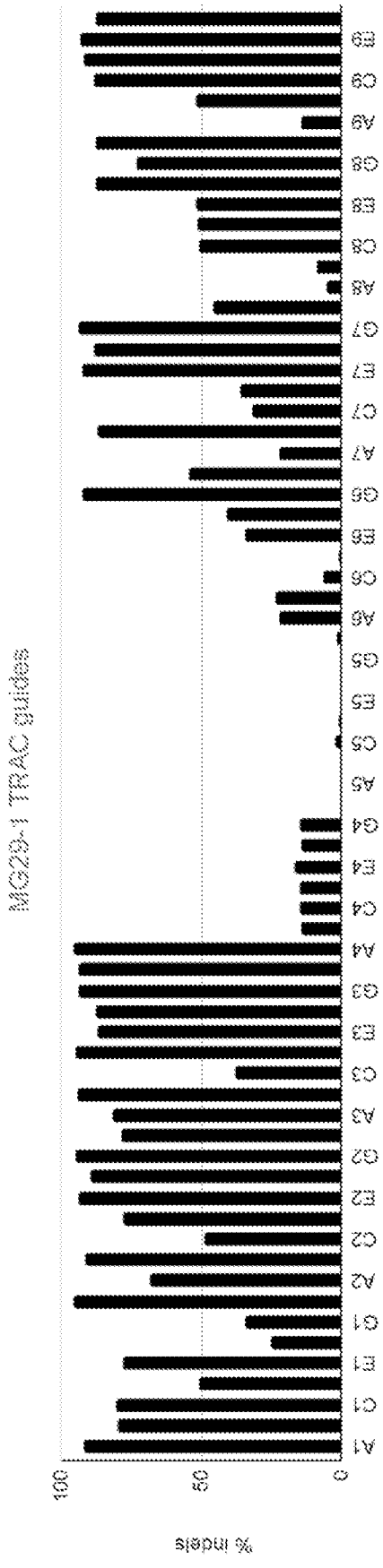


FIG. 64A

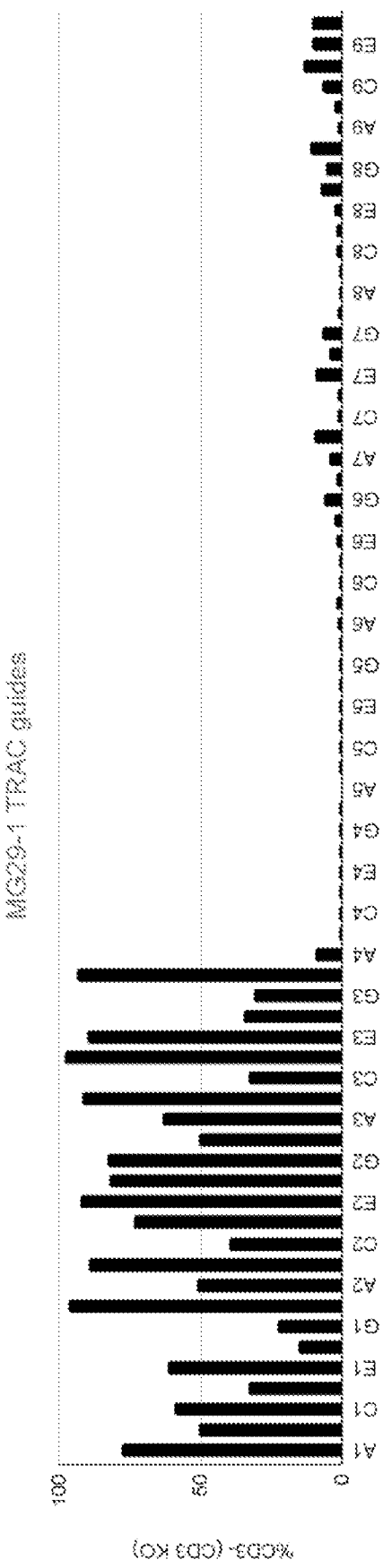


FIG. 64B

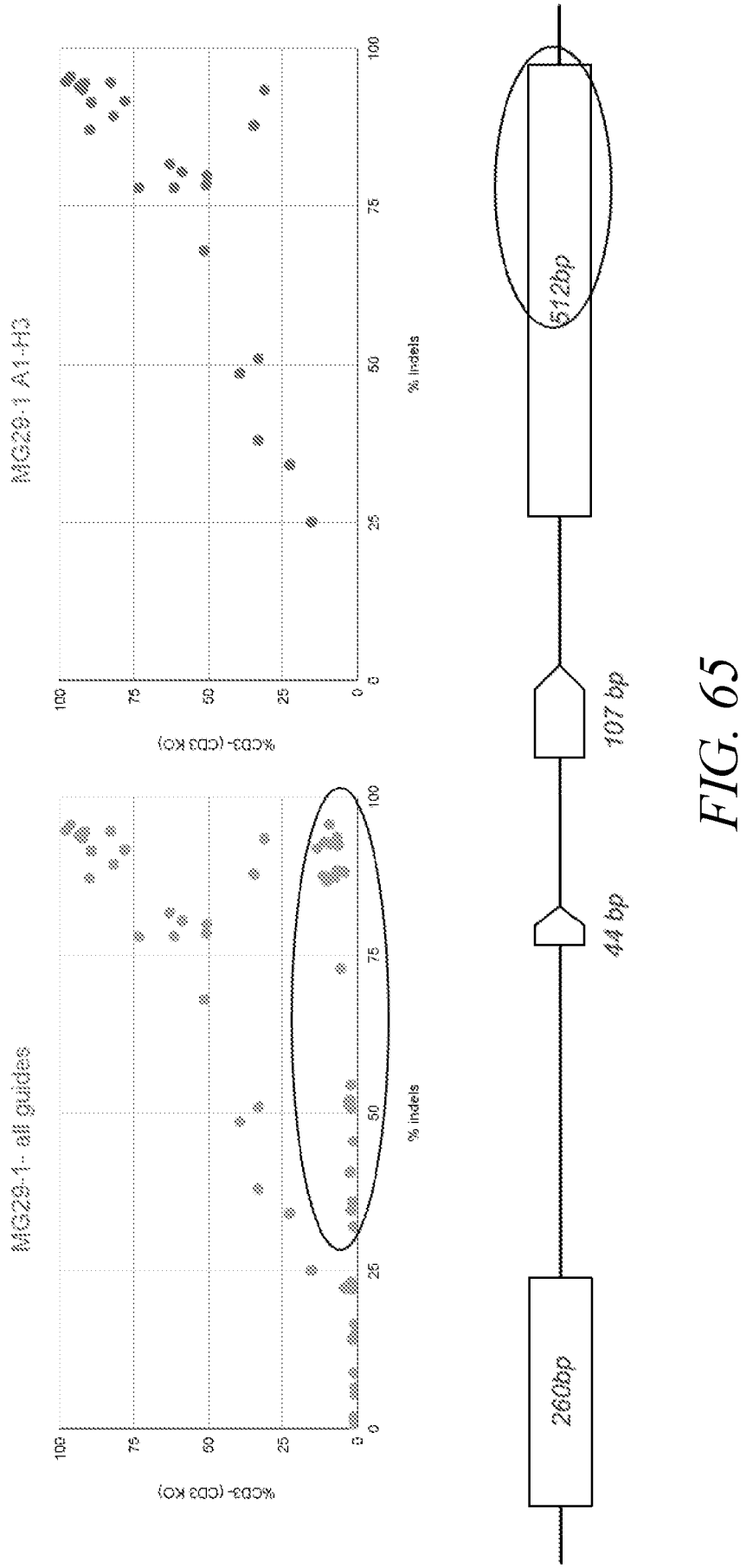

MG29-1 A1-H3
%CD3- (CD3 KO)
% Indels
0
25
50
75
100
MG29-1- all guides
%CD3- (CD3 KO)
% Indels
0
25
50
75
100
512bp
107 bp
44 bp
260bp

FIG. 65

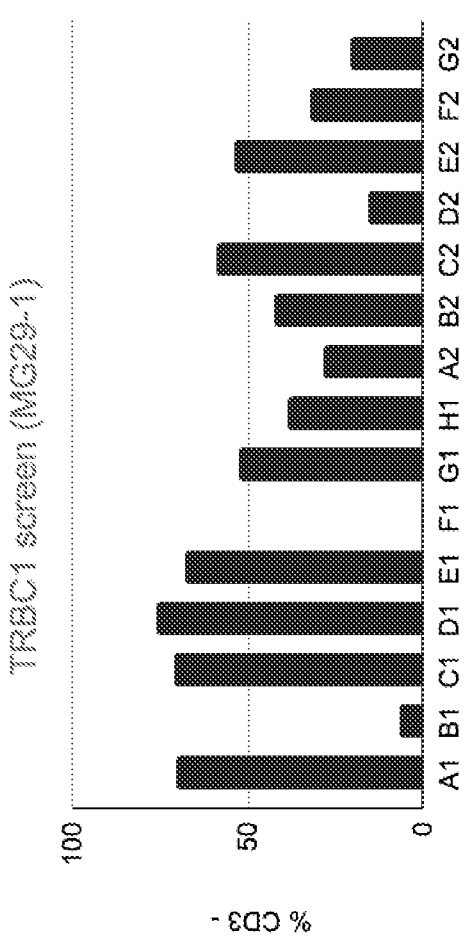


FIG. 66A

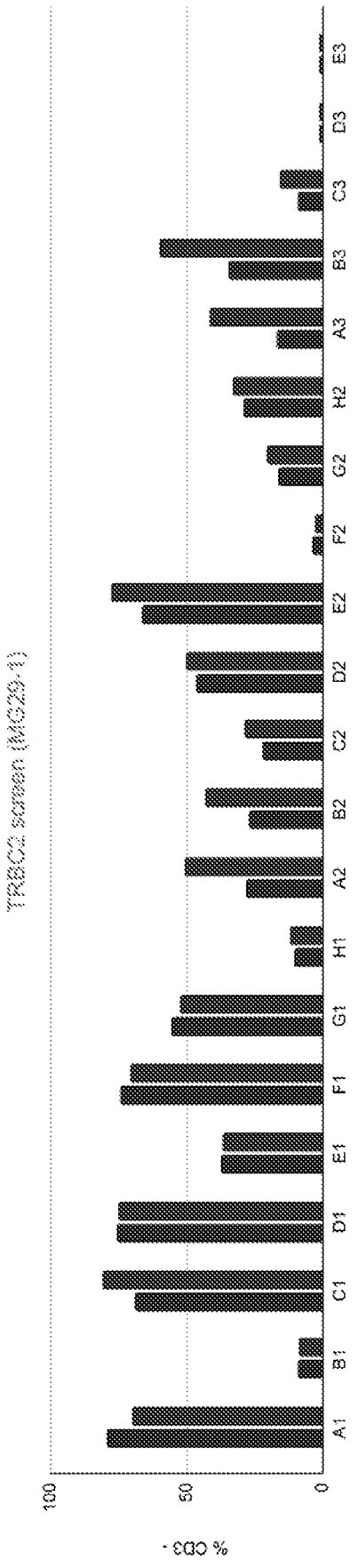


FIG. 66B

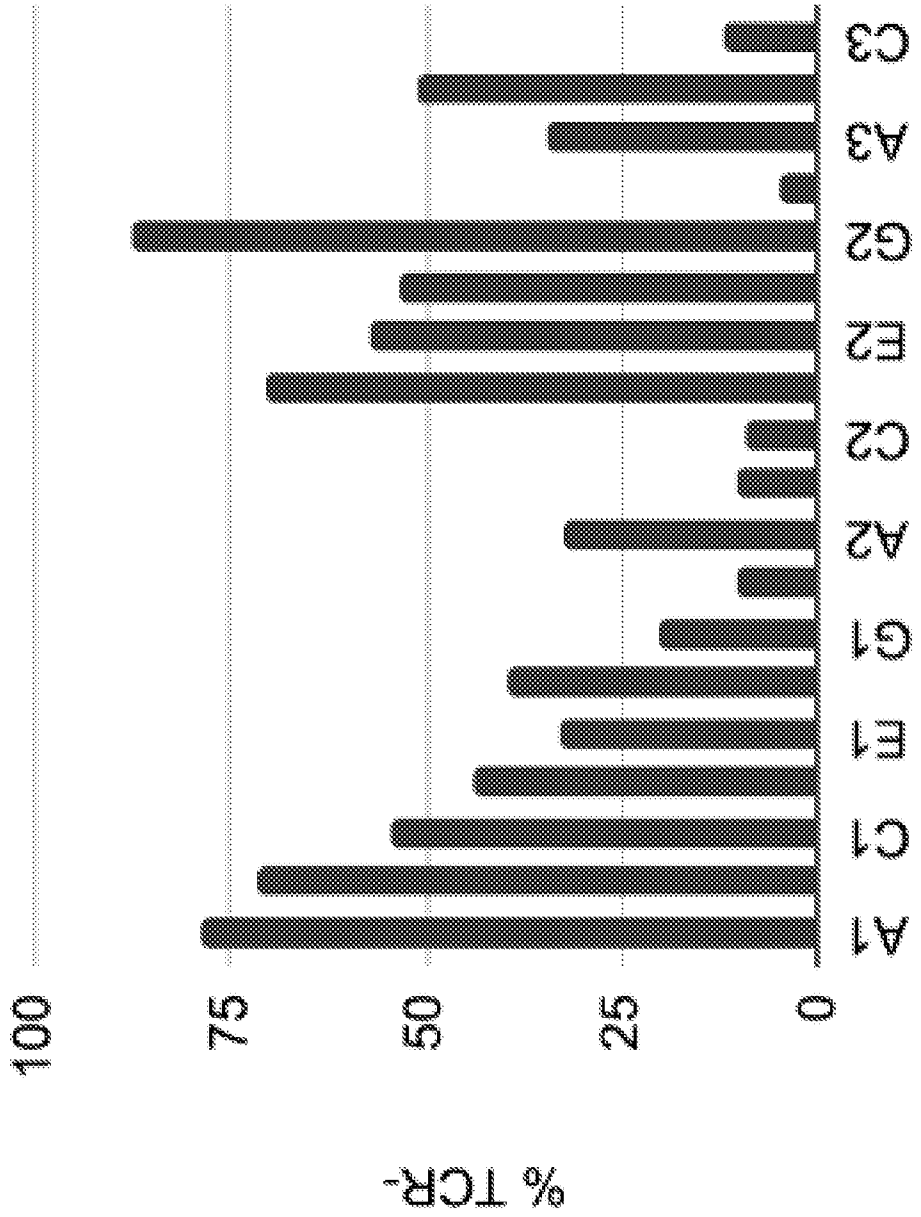


FIG. 66C

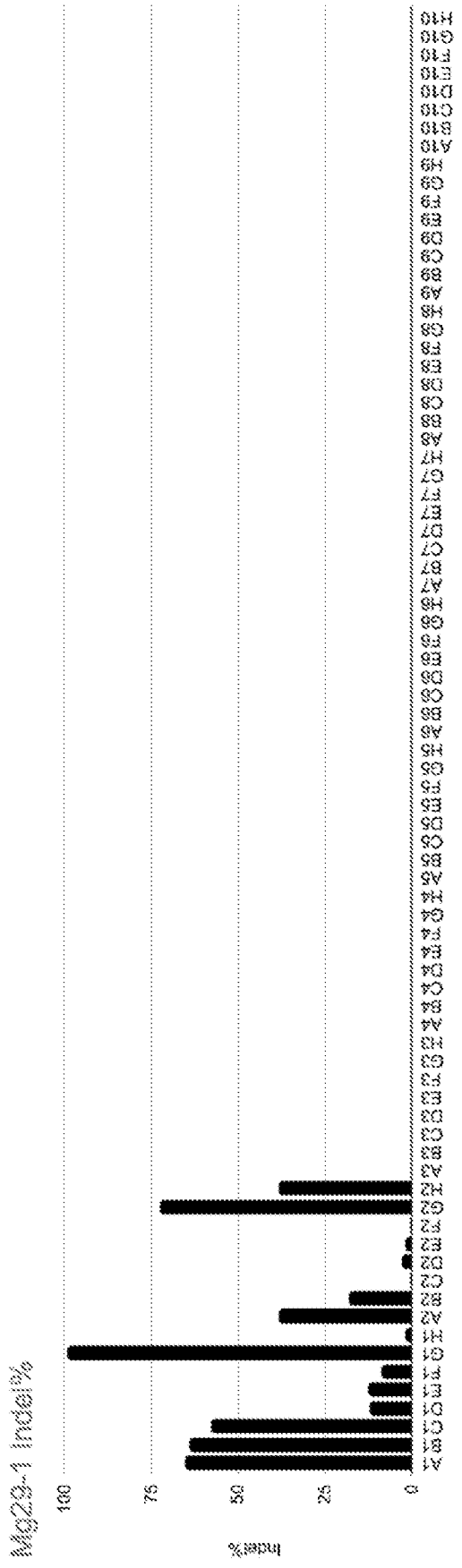


FIG. 67

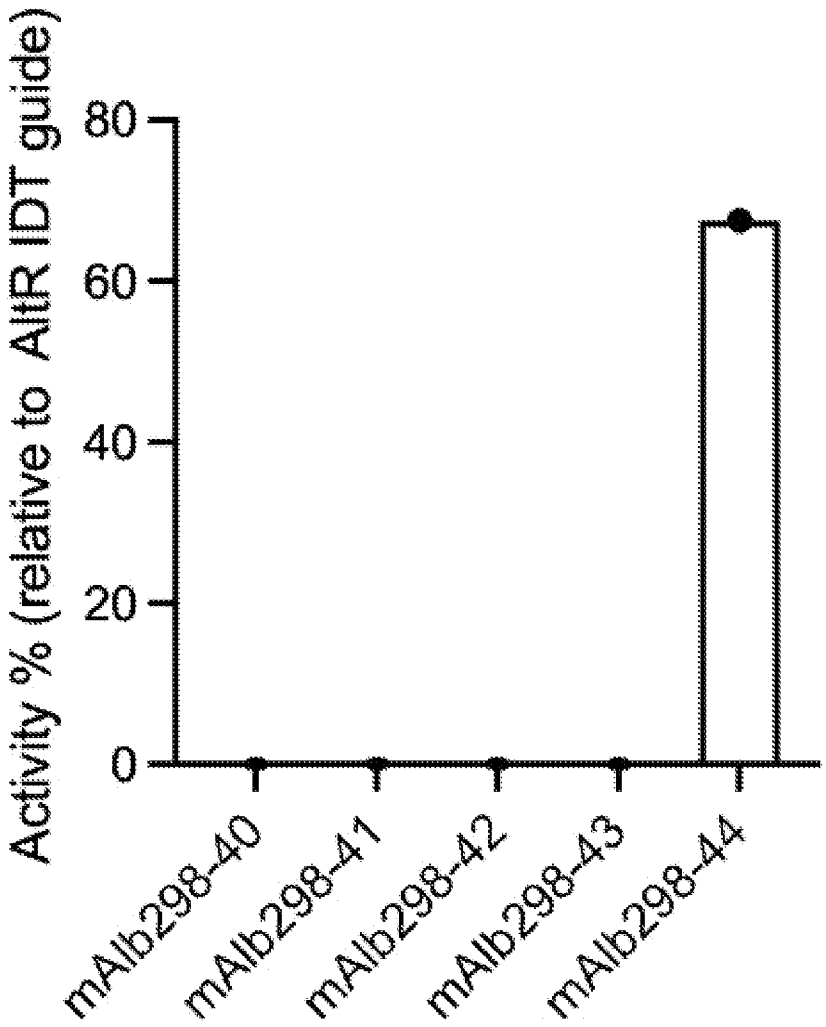

Activity % (relative to AltR IDT guide)
80
60
40
20
0
mAlb298-40
mAlb298-41
mAlb298-42
mAlb298-43
mAlb298-44

FIG. 68

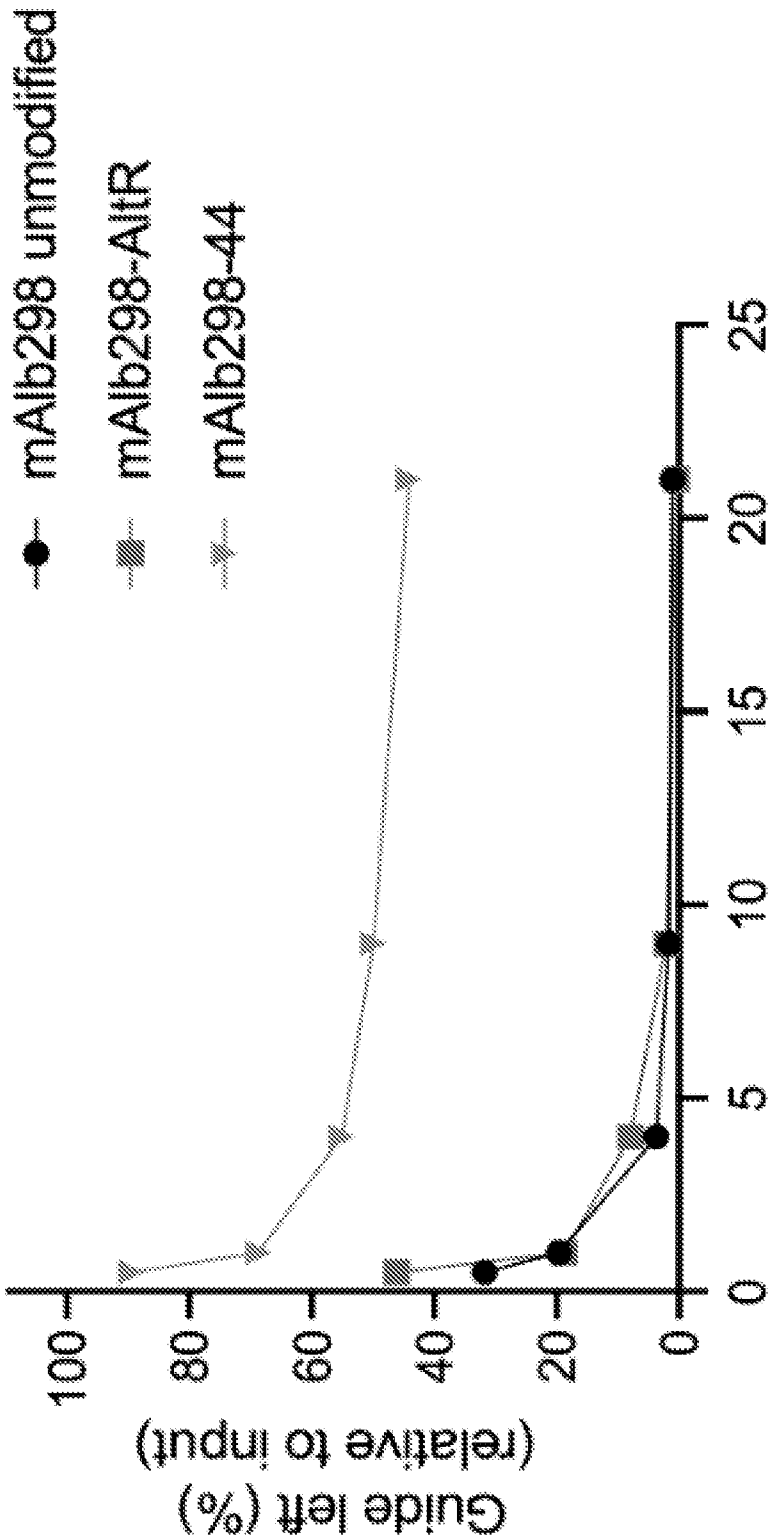

mAlb298 unmodified
mAlb298-AltR
mAlb298-44
Guide left (%)
(relative to input)
100
80
60
40
20
0
0
5
10
15
20
25
Time (H)

FIG. 69

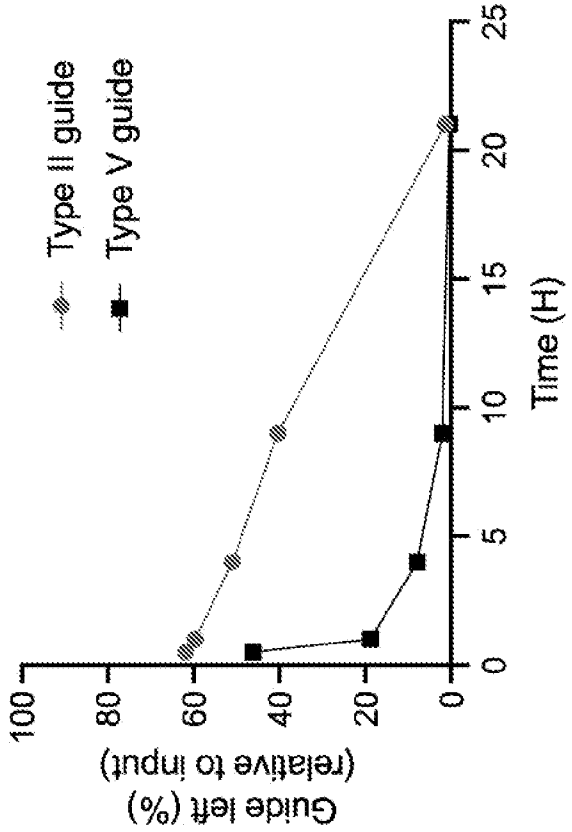


FIG. 70B

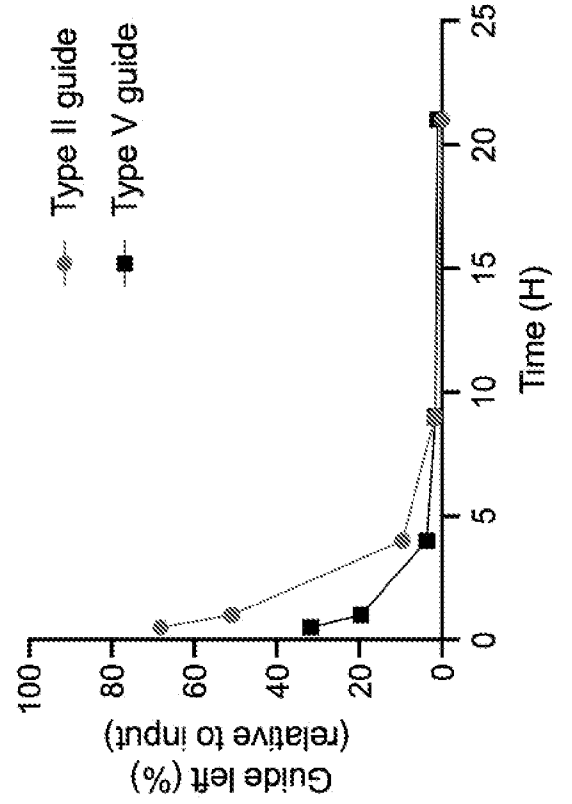
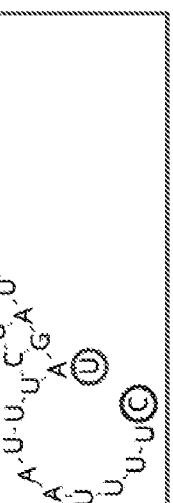
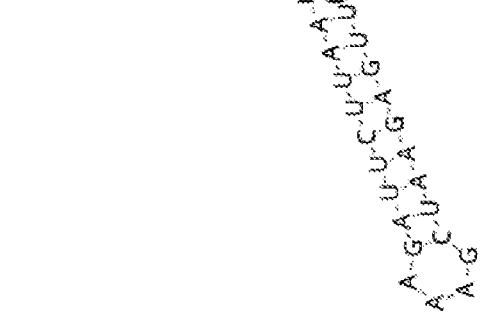


FIG. 70A

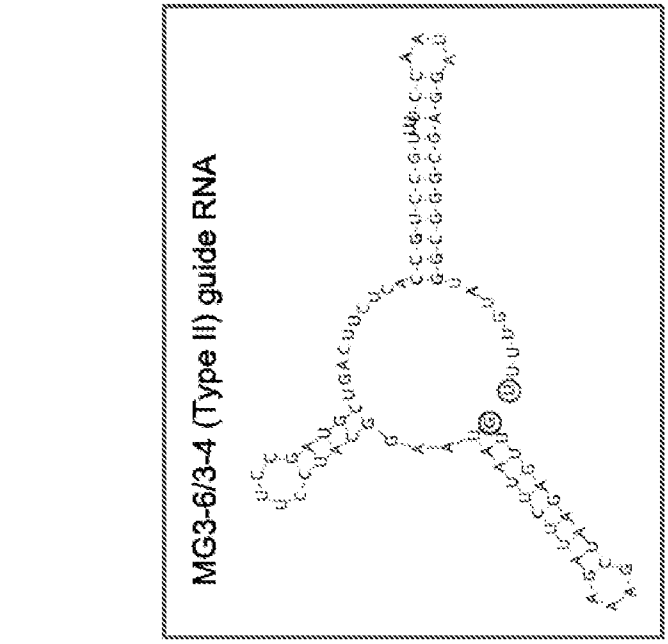
MG29-1 (Type V) guide RNA



MG29-1 (Type V) guide RNA with stem-loop 1 and stem-loop 2 of MG3-6/3-4 guide added at the 5' end



MG3-6/3-4 (Type II) guide RNA



MG29-1 (Type V) guide RNA with stem-loop 1 of MG3-6/3-4 guide added at the 5' end

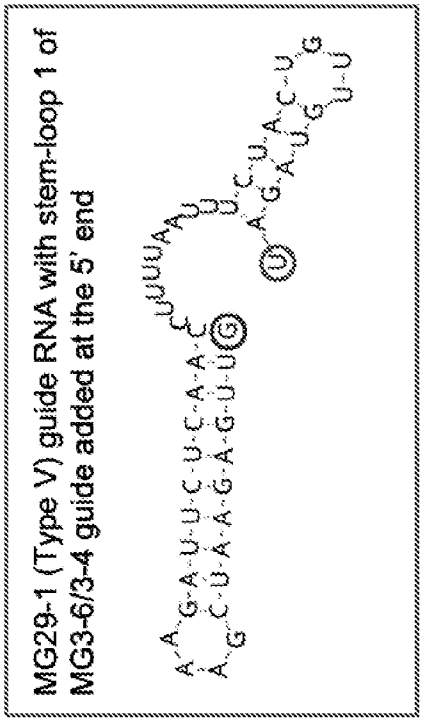


FIG. 71

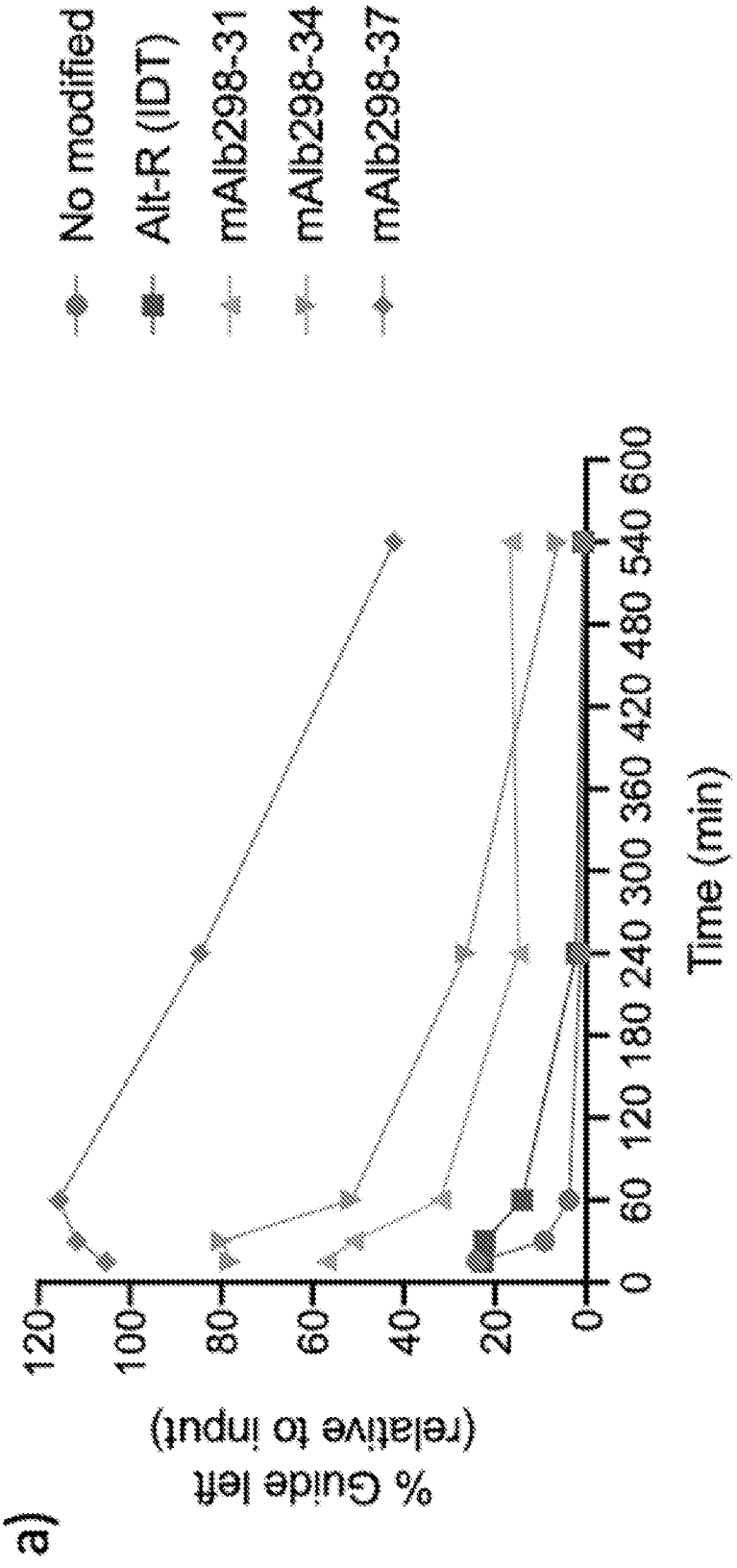

a)
% Guide left
(relative to input)
120
100
80
60
40
20
0
0
60
120
180
240
300
360
420
480
540
600
Time (min)
No modified
Alt-R (IDT)
mAlb298-31
mAlb298-34
mAlb298-37

FIG. 72

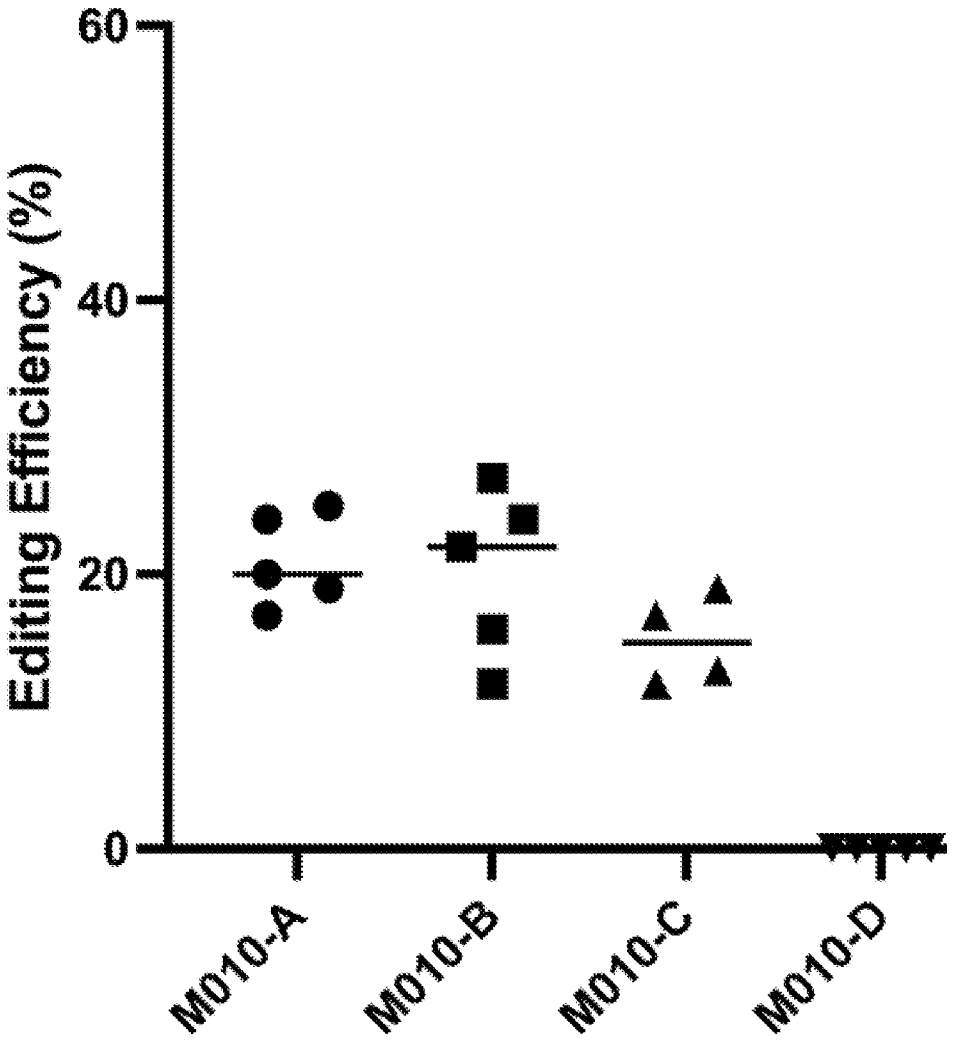


FIG. 73

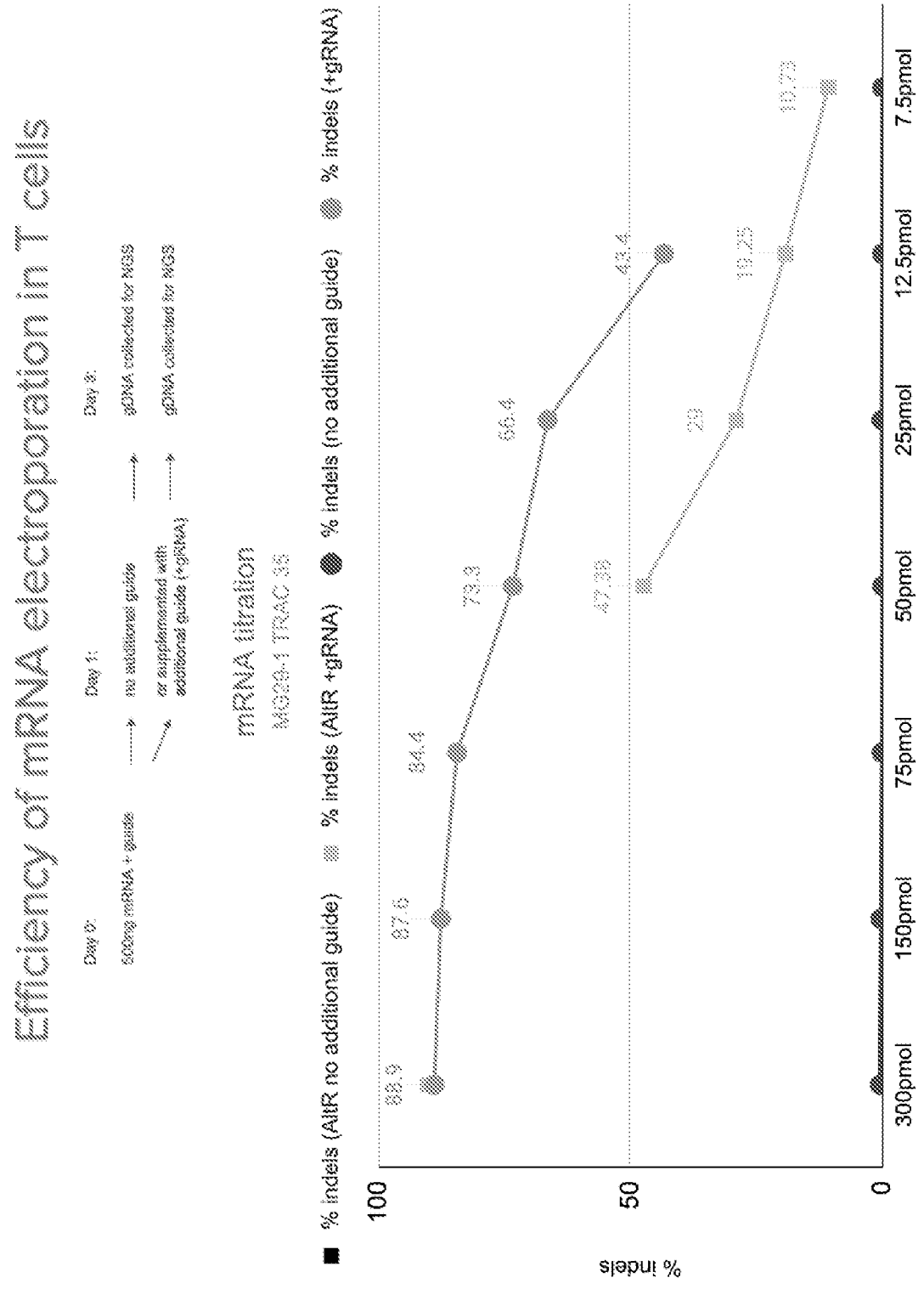

Efficiency of mRNA electroporation in T cells
Day 0:
500ng mRNA + guide
Day 1:
no additional guide
or supplemented with additional guide (+gRNA)
Day 8:
gDNA collected for NGS
gDNA collected for NGS
mRNA titration
MG29-1 TRAC 3S
% indels (AltR no additional guide)
% indels (AltR +gRNA)
% indels (no additional guide)
% indels (+gRNA)
% indels
100
50
0
300pmol
150pmol
75pmol
50pmol
25pmol
12.5pmol
7.5pmol
88.9
87.6
84.4
73.3
66.4
47.38
43.4
29
19.23
10.73

FIG. 74

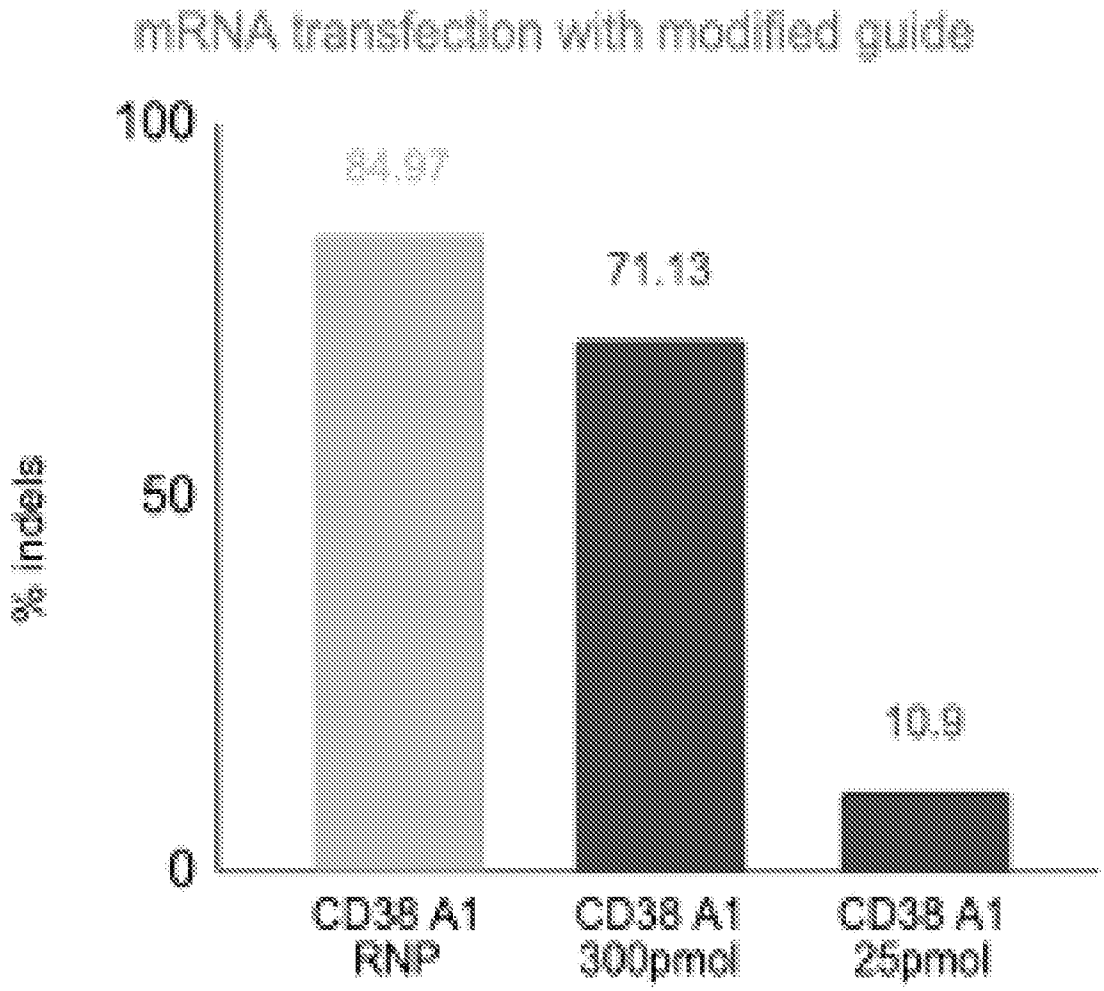

mRNA transfection with modified guide
100
50
0
% indels
84.97
71.13
10.9
CD38 A1 RNP
CD38 A1 300pmol
CD38 A1 25pmol

FIG. 75

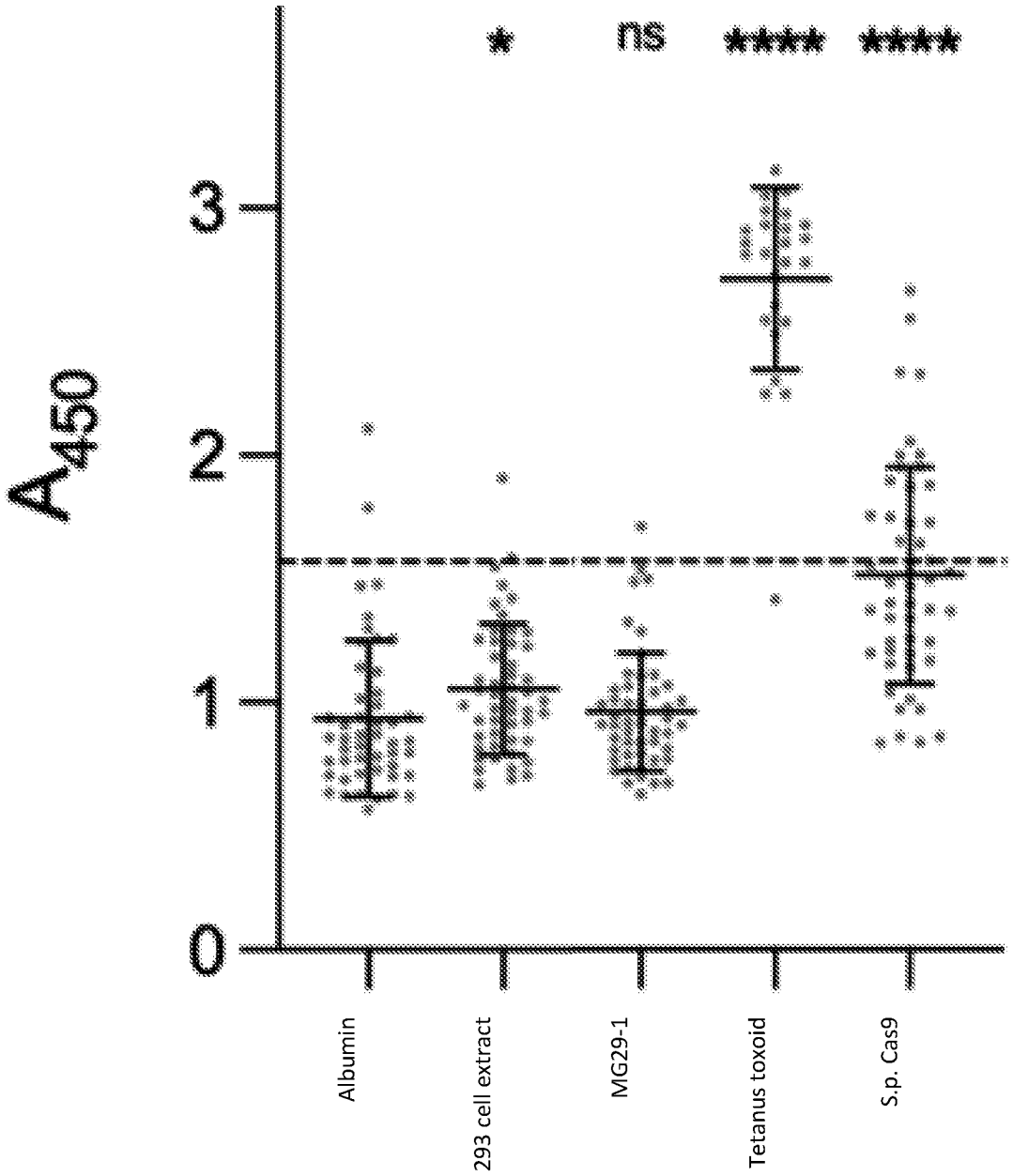

*
ns
****
****
A450
0
1
2
3
Albumin
293 cell extract
MG29-1
Tetanus toxoid
S.p. Cas9

FIG. 76

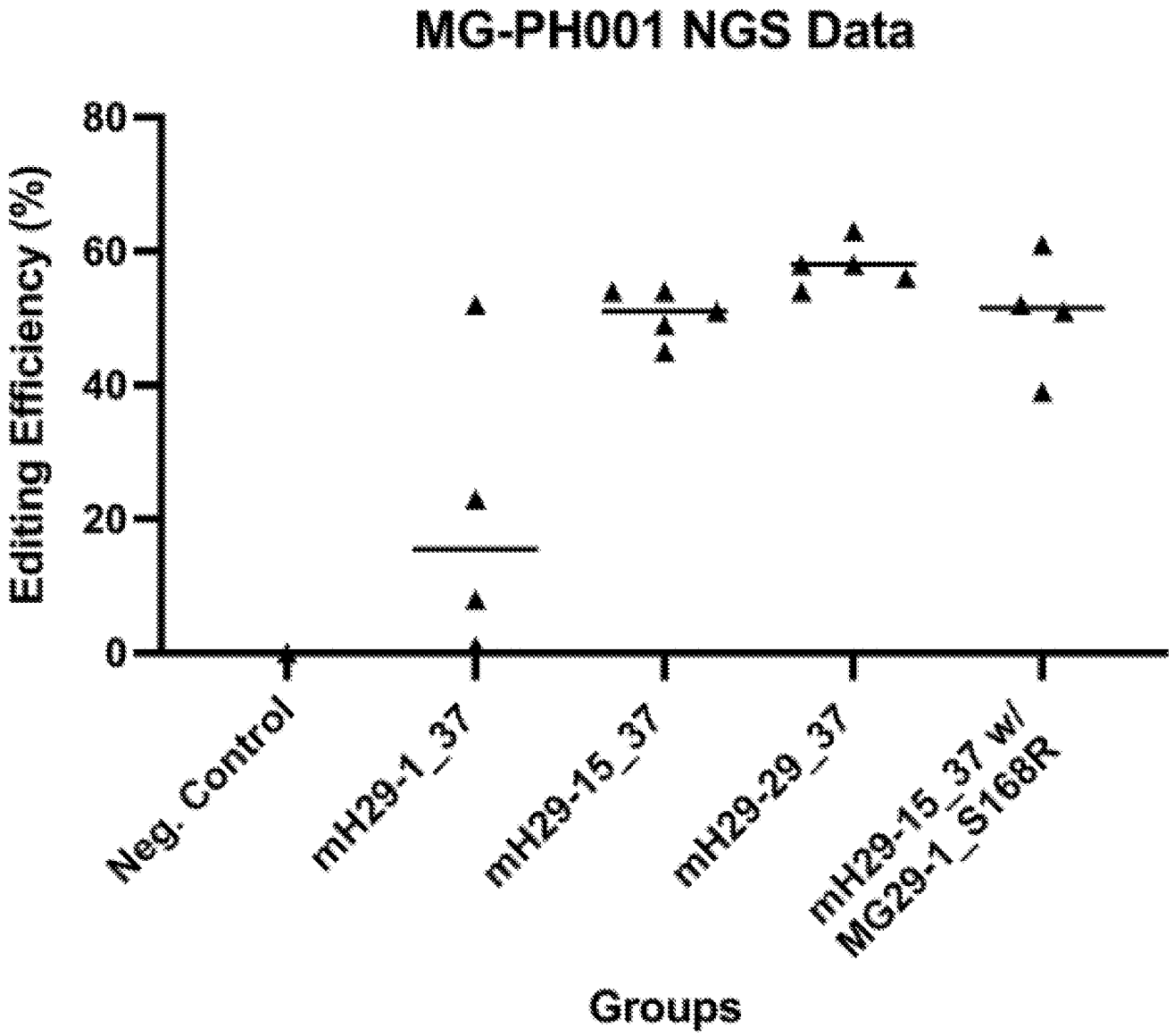

MG-PH001 NGS Data
Editing Efficiency (%)
80
60
40
20
0
Neg. Control
mH29-1_37
mH29-15_37
mH29-29_37
mH29-15_37 w/ MG29-1_S168R
Groups

FIG. 77

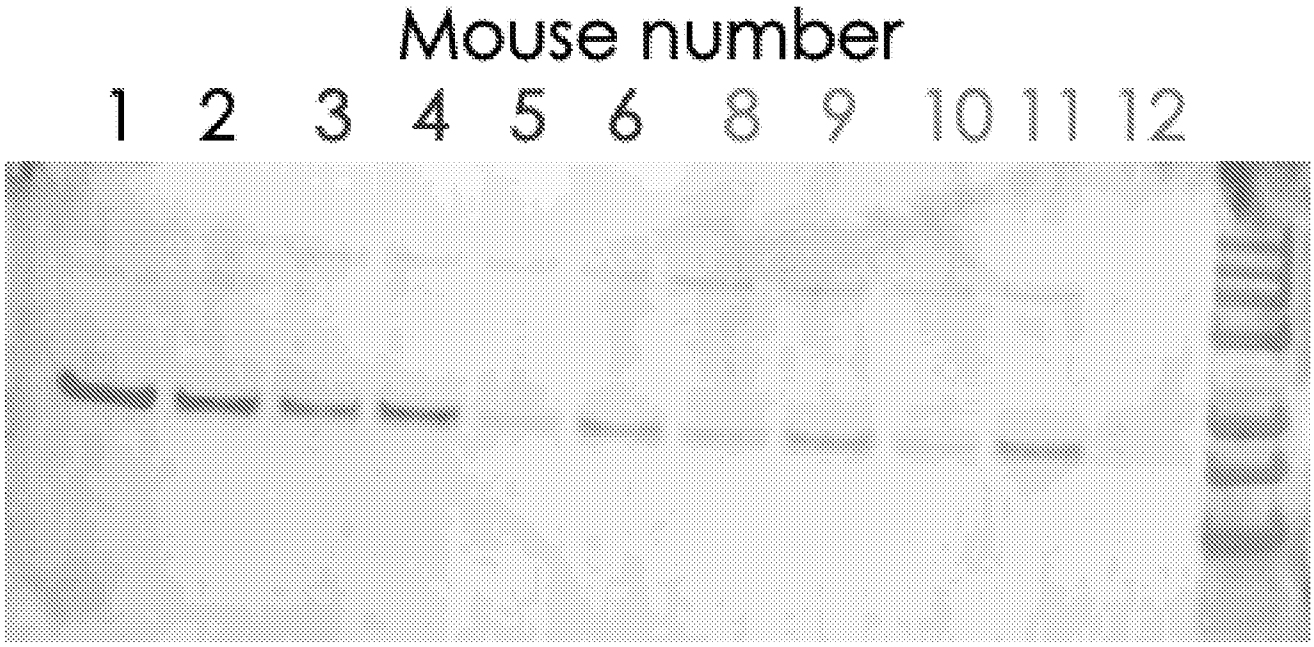


FIG. 78A

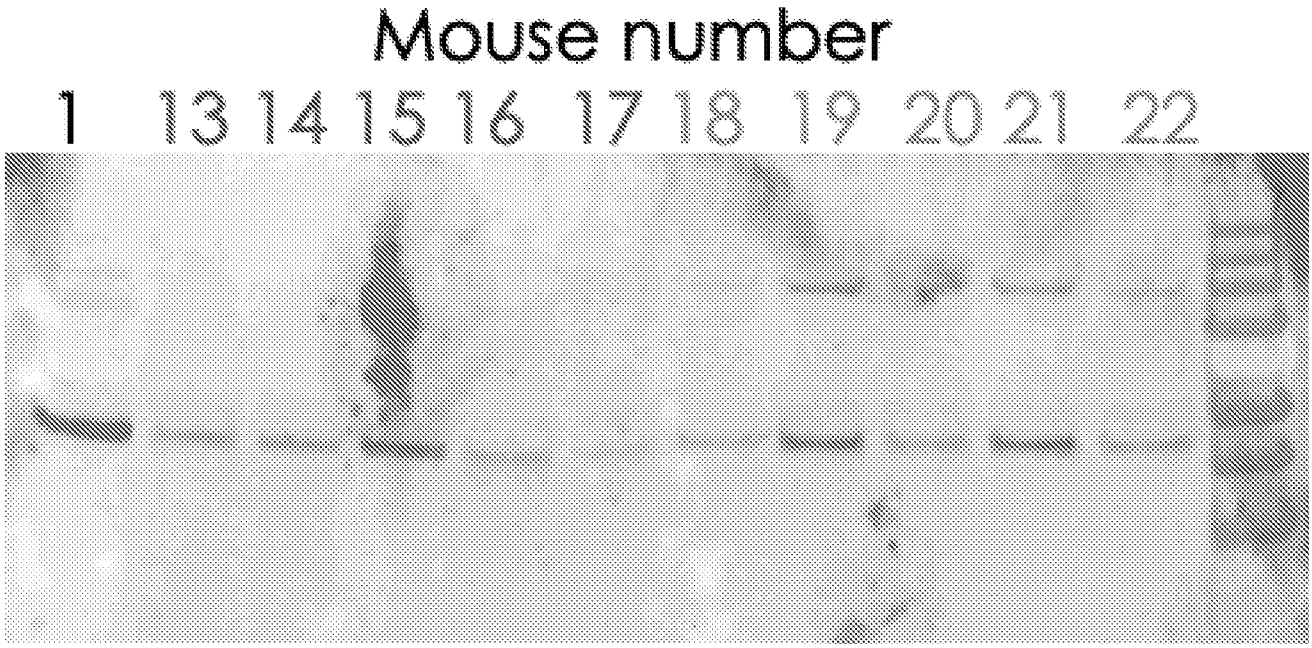


FIG. 78B

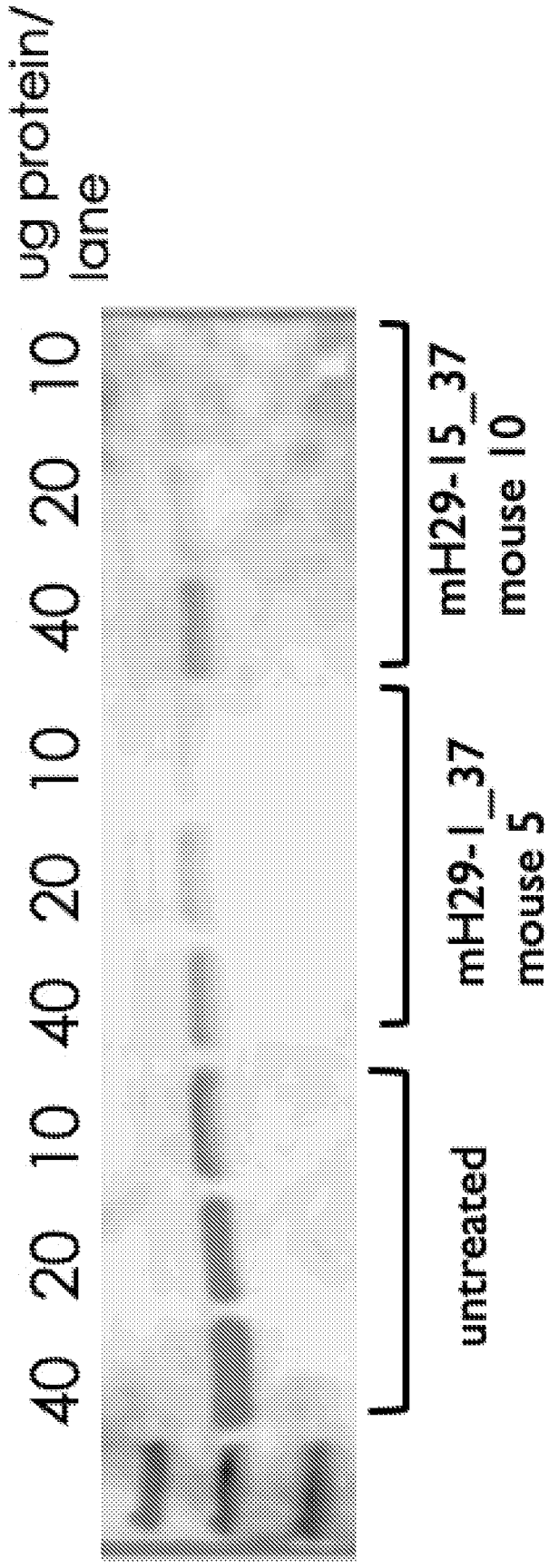

ug protein/ lane
40 20 10 40 20 10 40 20 10
untreated
mH29-I_37 mouse 5
mH29-I5_37 mouse 10

FIG. 79

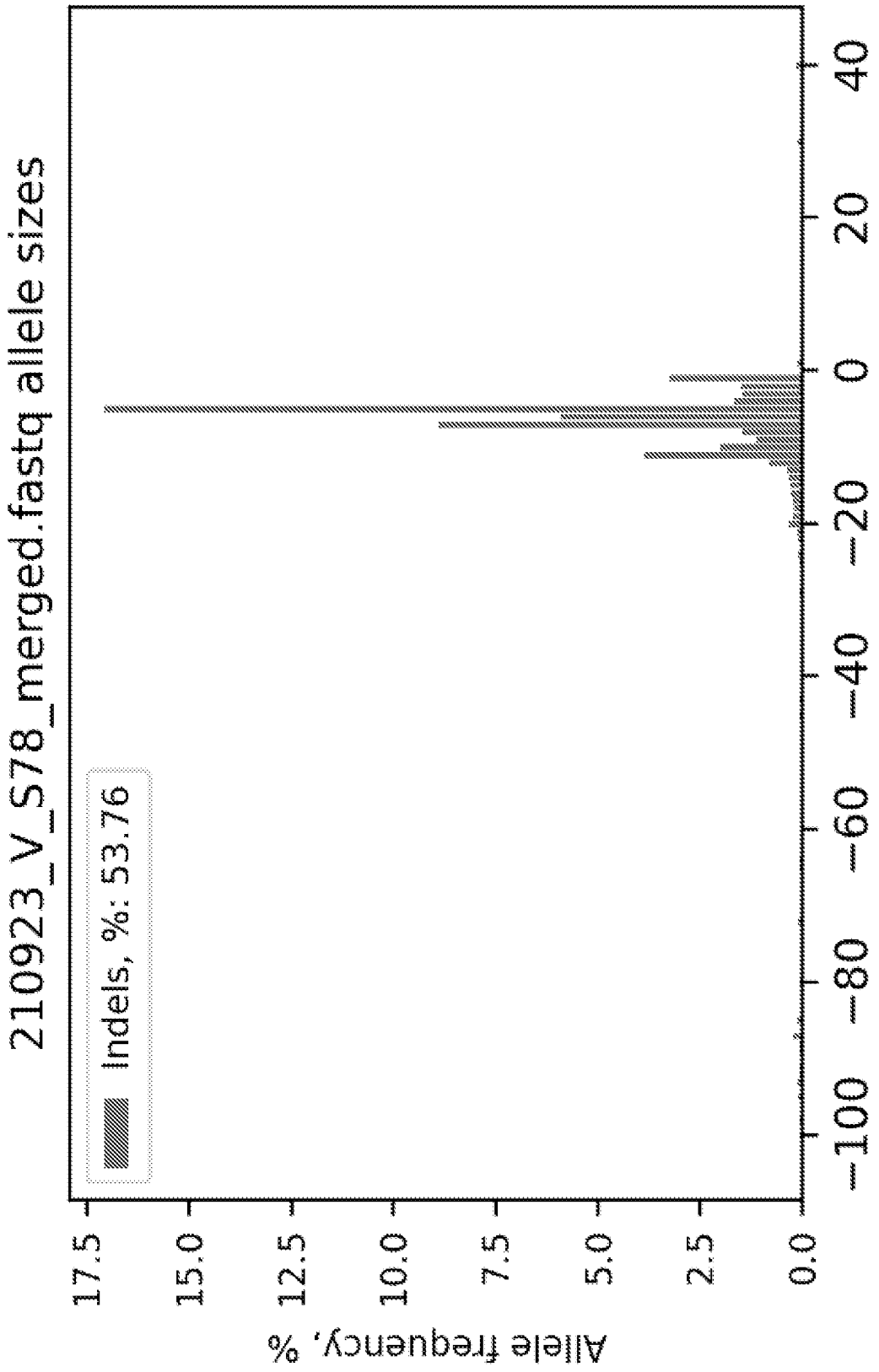

210923_V_S78_merged.fastq allele sizes
Indels, %: 53.76
Allele frequency, %
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
−100
−80
−60
−40
−20
0
20
40

FIG. 80

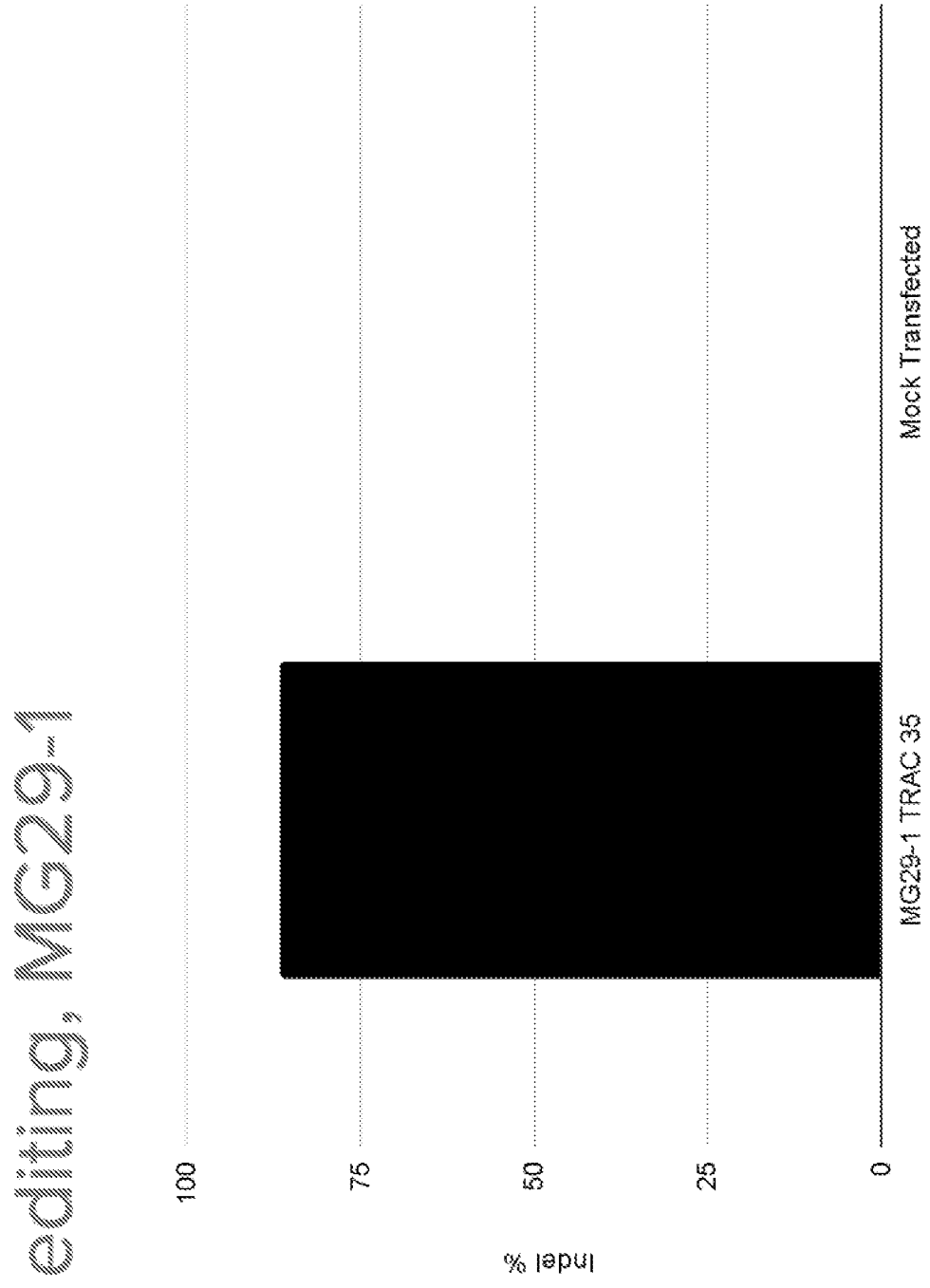

B cell editing, MG29-1
Indel %
100
75
50
25
0
MG29-1 TRAC 35
Mock Transfected

FIG. 81

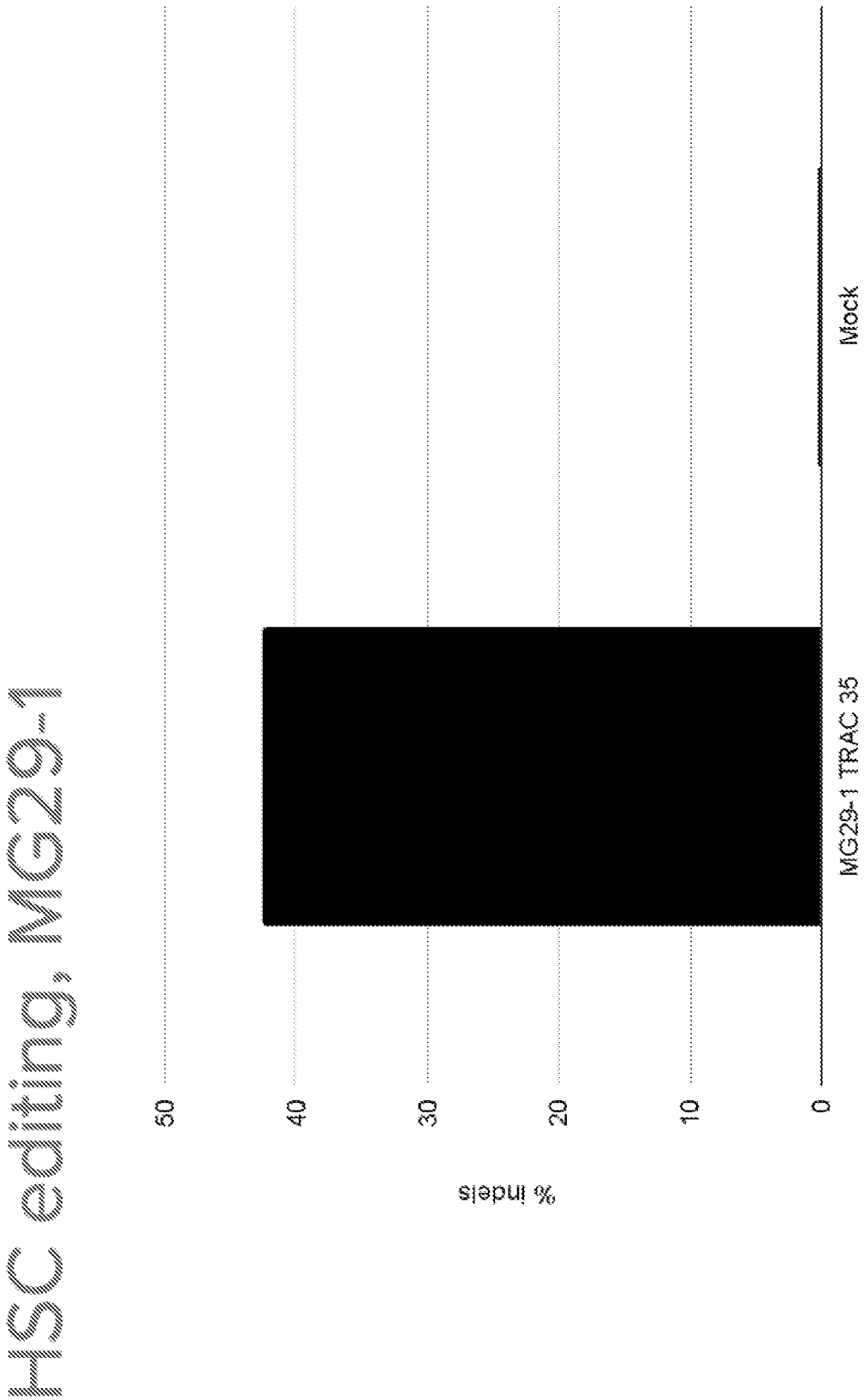


FIG. 82

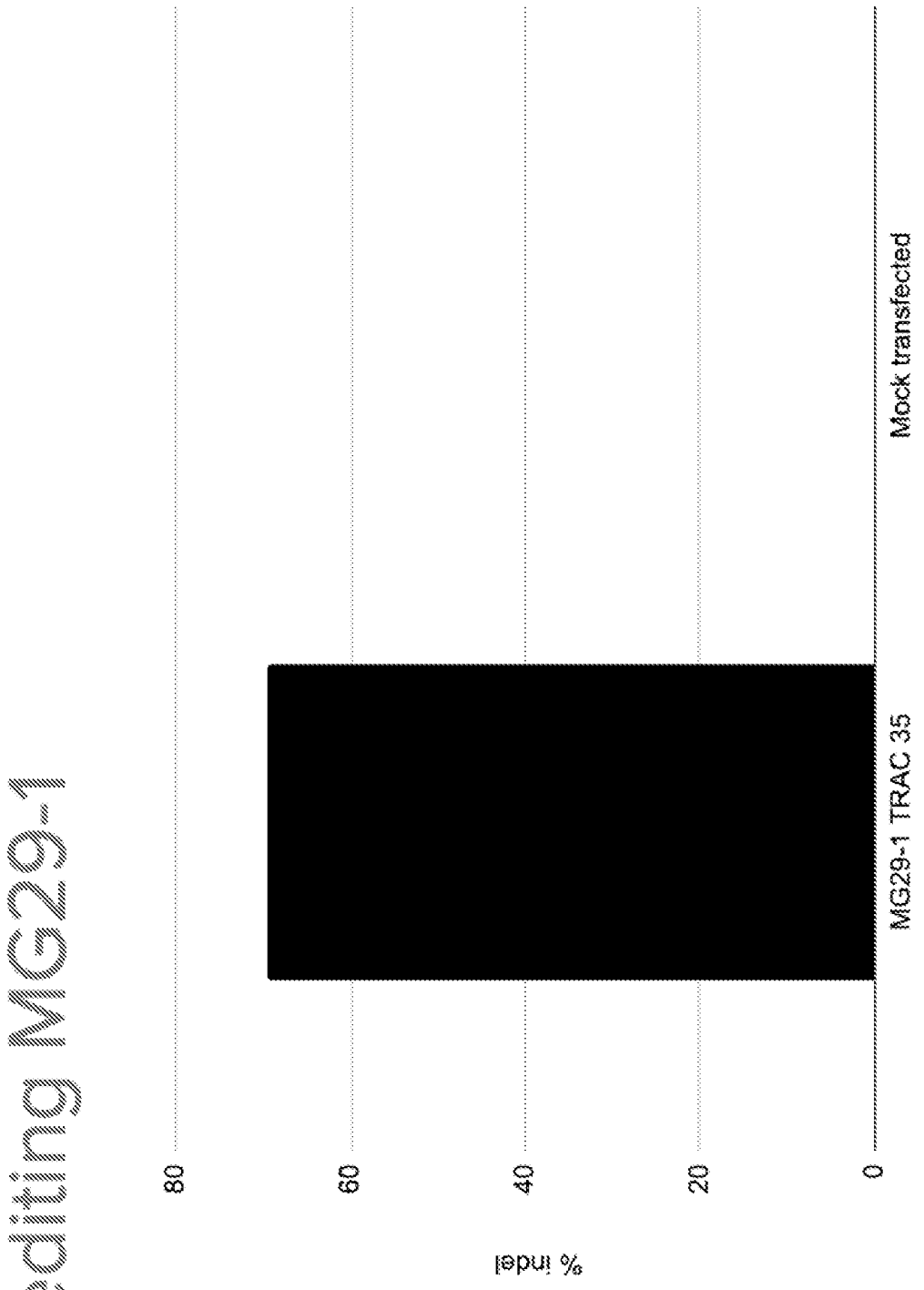

iPSC editing MG29-1
% indel
80
60
40
20
0
MG29-1 TRAC 35
Mock transfected

FIG. 83

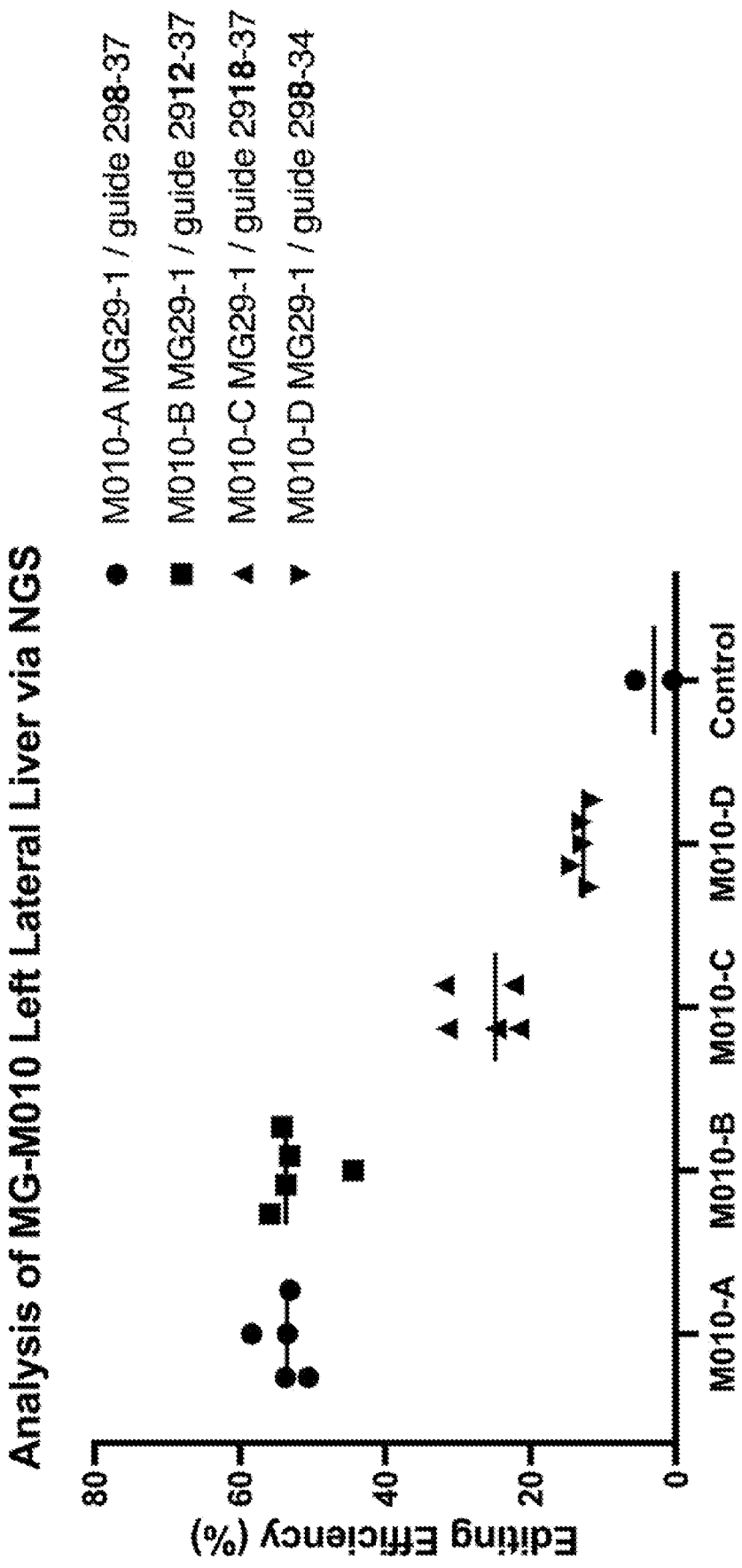

Analysis of MG-M010 Left Lateral Liver via NGS
M010-A MG29-1 / guide 298-37
M010-B MG29-1 / guide 2912-37
M010-C MG29-1 / guide 2918-37
M010-D MG29-1 / guide 298-34
Editing Efficiency (%)
80
60
40
20
0
M010-A
M010-B
M010-C
M010-D
Control

FIG. 84

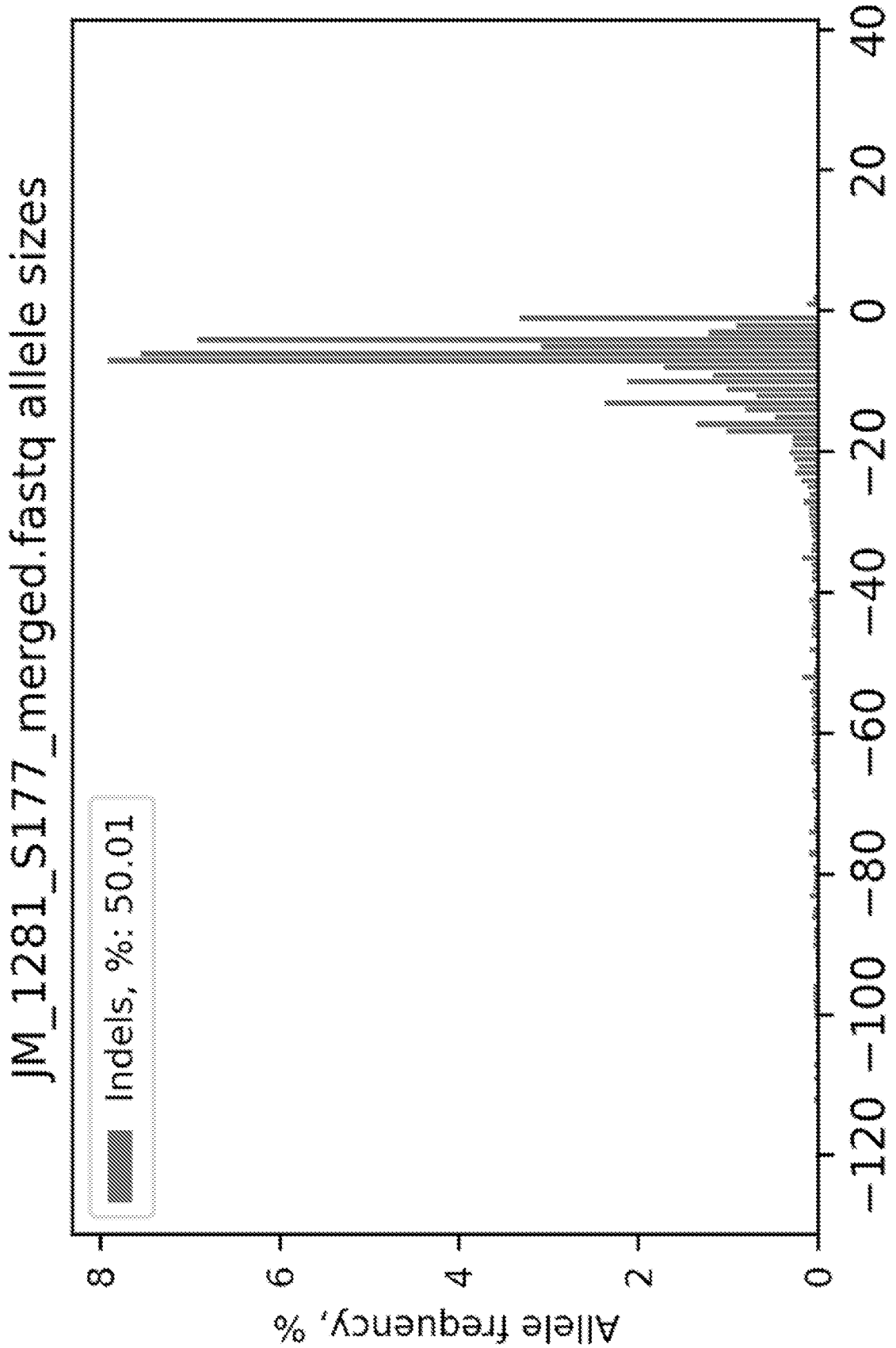
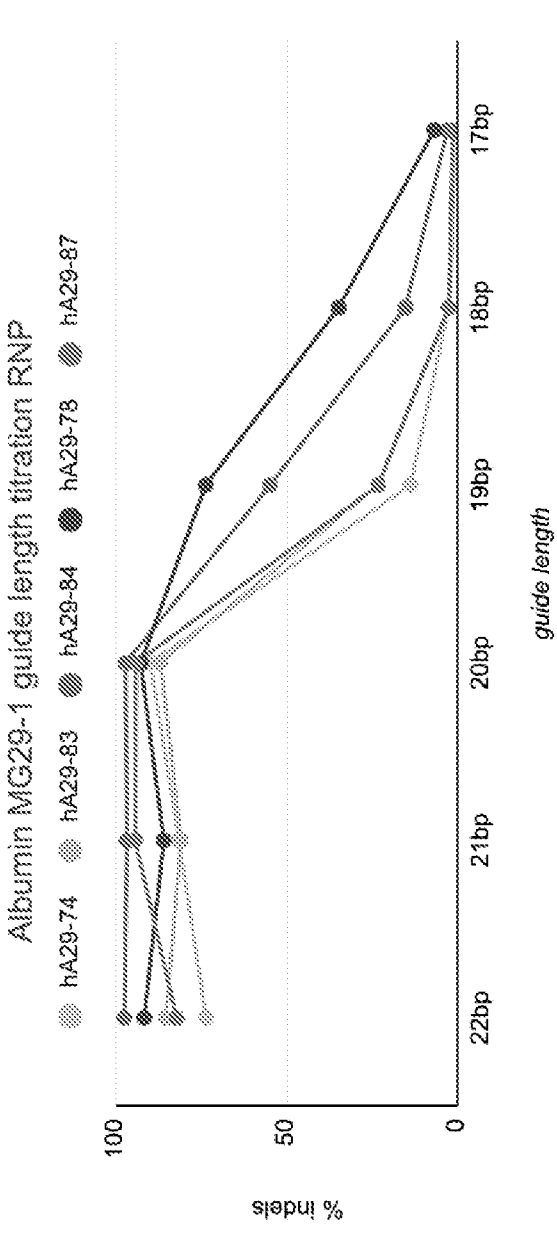

JM_1281_S177_merged.fastq allele sizes
Indels, %: 50.01
Allele frequency, %
0
2
4
6
8
−120
−100
−80
−60
−40
−20
0
20
40

FIG. 85


Albumin MG29-1 guide length titration RNP
hA29-74
hA29-83
hA29-84
hA29-78
hA29-87
% Indels
100
50
0
22bp
21bp
20bp
19bp
18bp
17bp
guide length

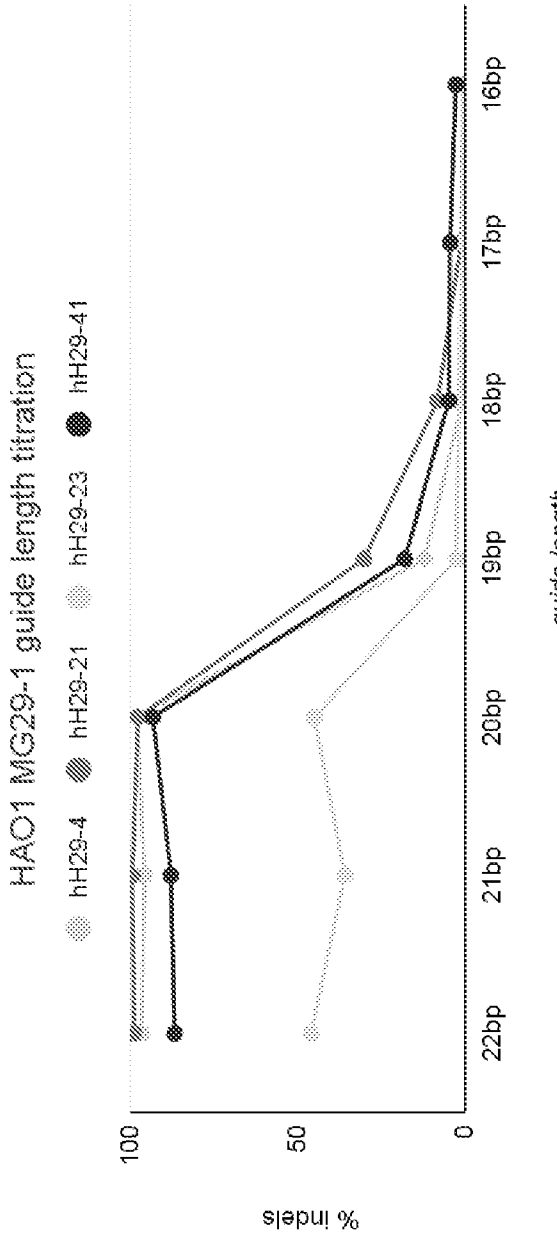

HAO1 MG29-1 guide length titration
hH29-4
hH29-21
hH29-23
hH29-41
% Indels
100
50
0
22bp
21bp
20bp
19bp
18bp
17bp
16bp
guide length

FIG. 86

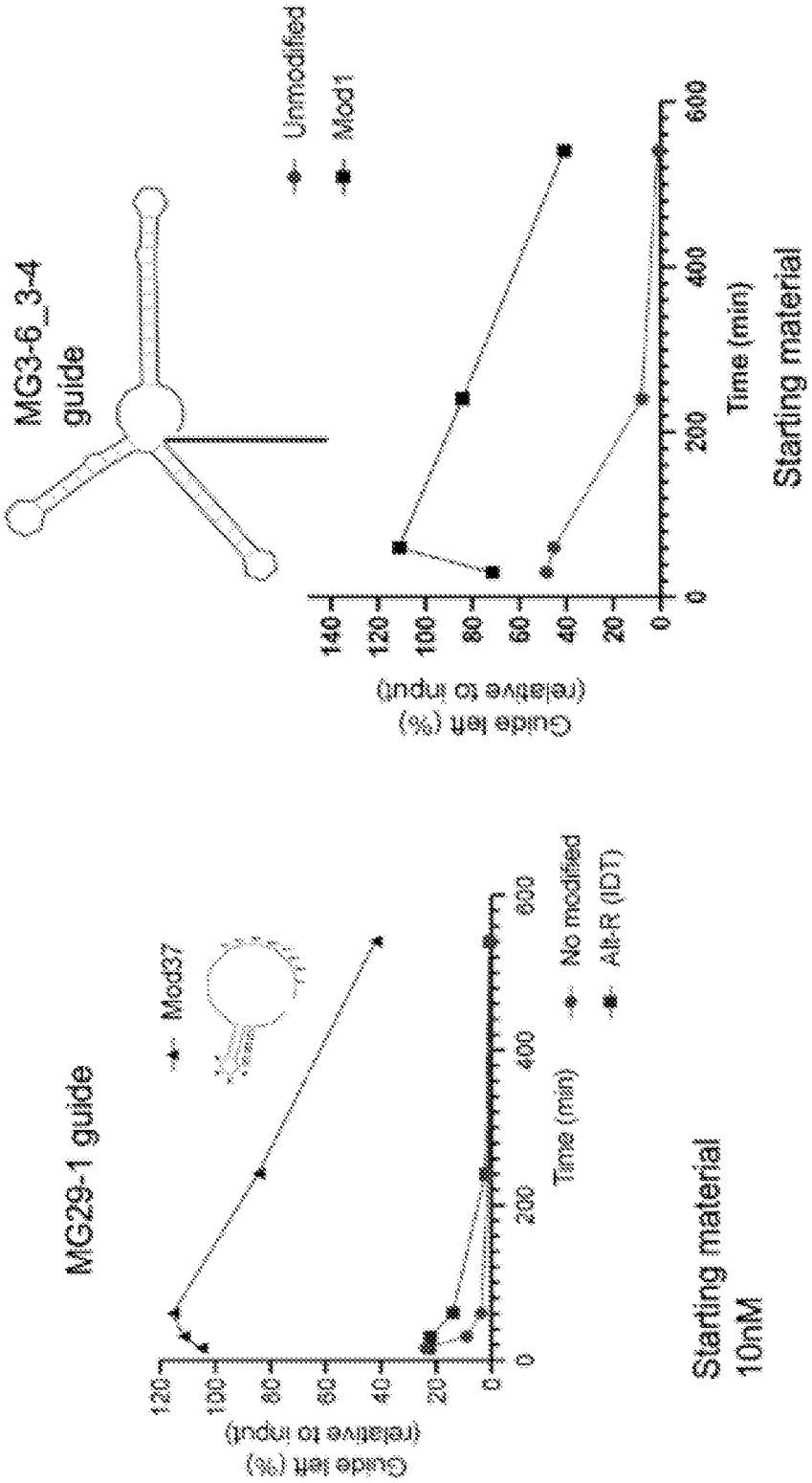

MG3-6_3-4 guide
Unmodified
Mod1
Guide left (%) (relative to input)
Time (min)
Starting material
MG29-1 guide
Mod37
No modified
Alt-R (IDT)
Starting material 10nM

FIG. 87

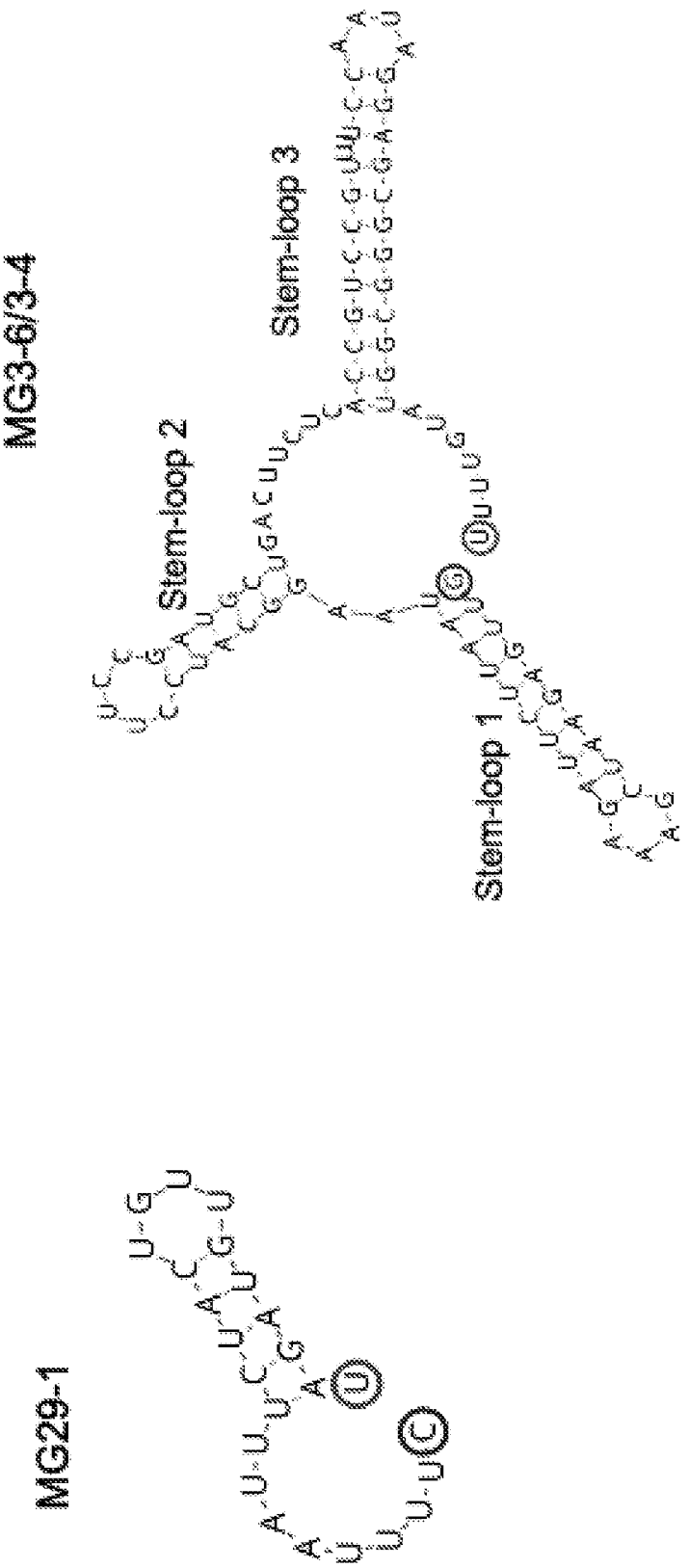


FIG. 88

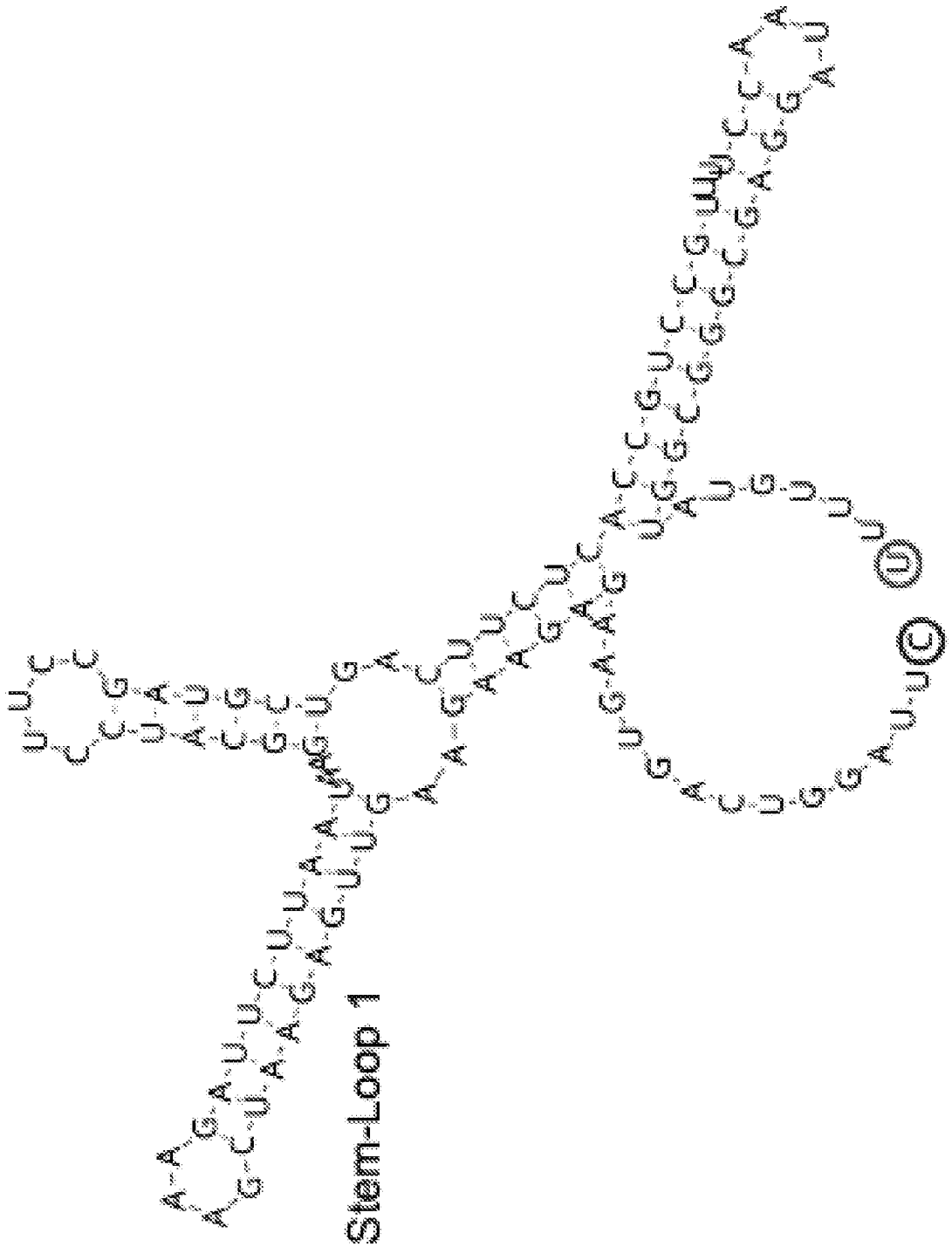


FIG. 89

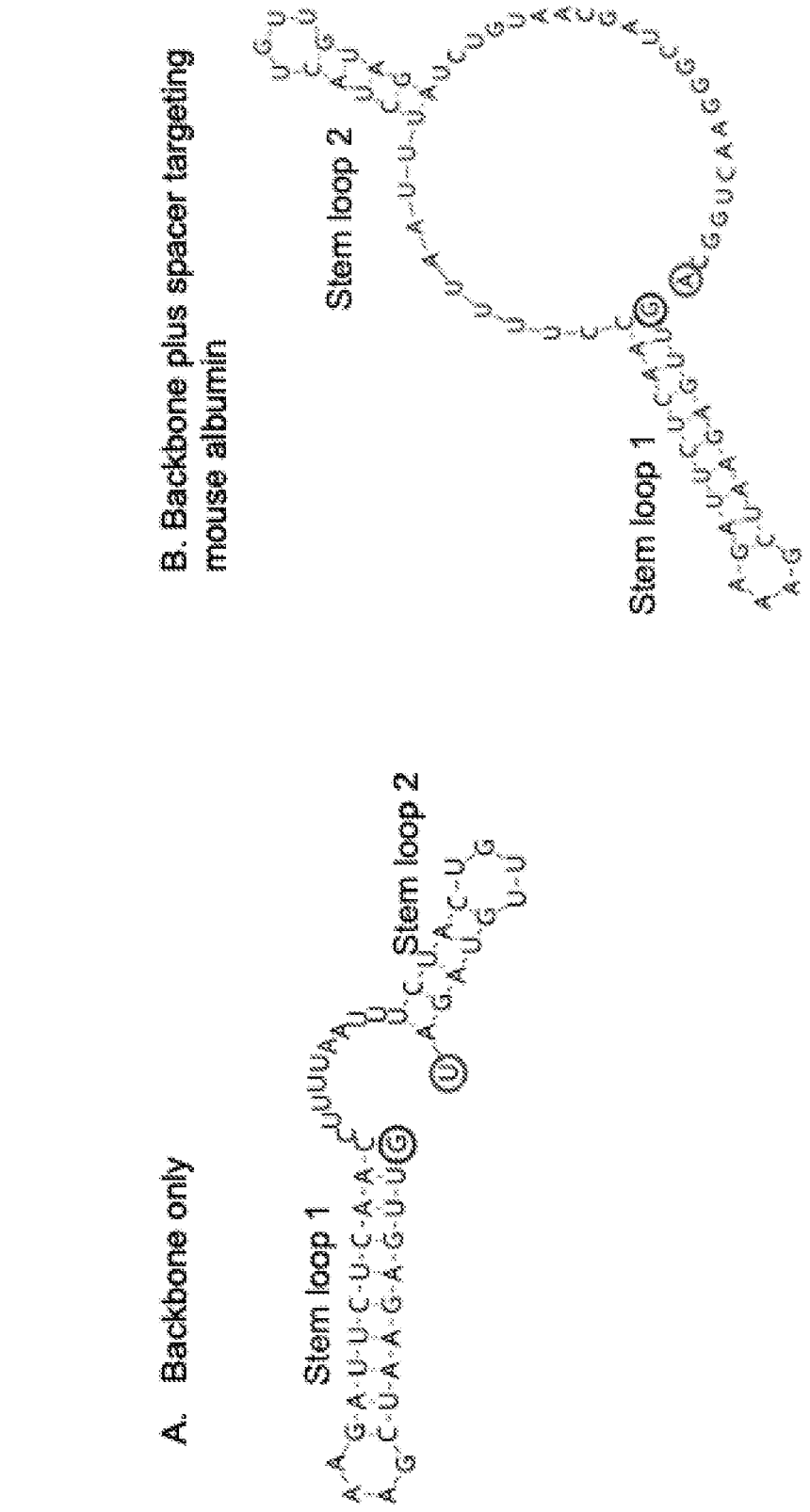

A. Backbone only
Stem loop 1
Stem loop 2
B. Backbone plus spacer targeting mouse albumin
Stem loop 2
Stem loop 1

FIG. 90

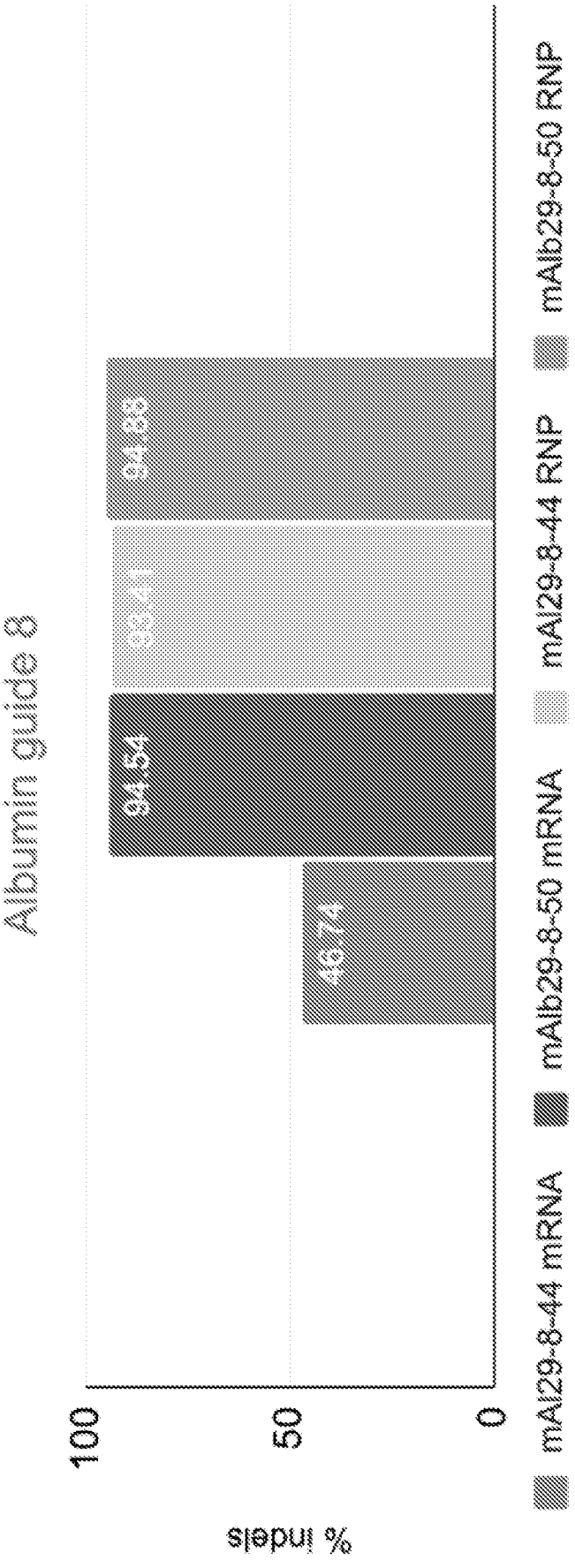

Albumin guide 8
% indels
100
50
0
94.54
48.74
93.41
94.88
mAl29-8-44 mRNA
mAlb29-8-50 mRNA
mAl29-8-44 RNP
mAlb29-8-50 RNP

FIG. 91

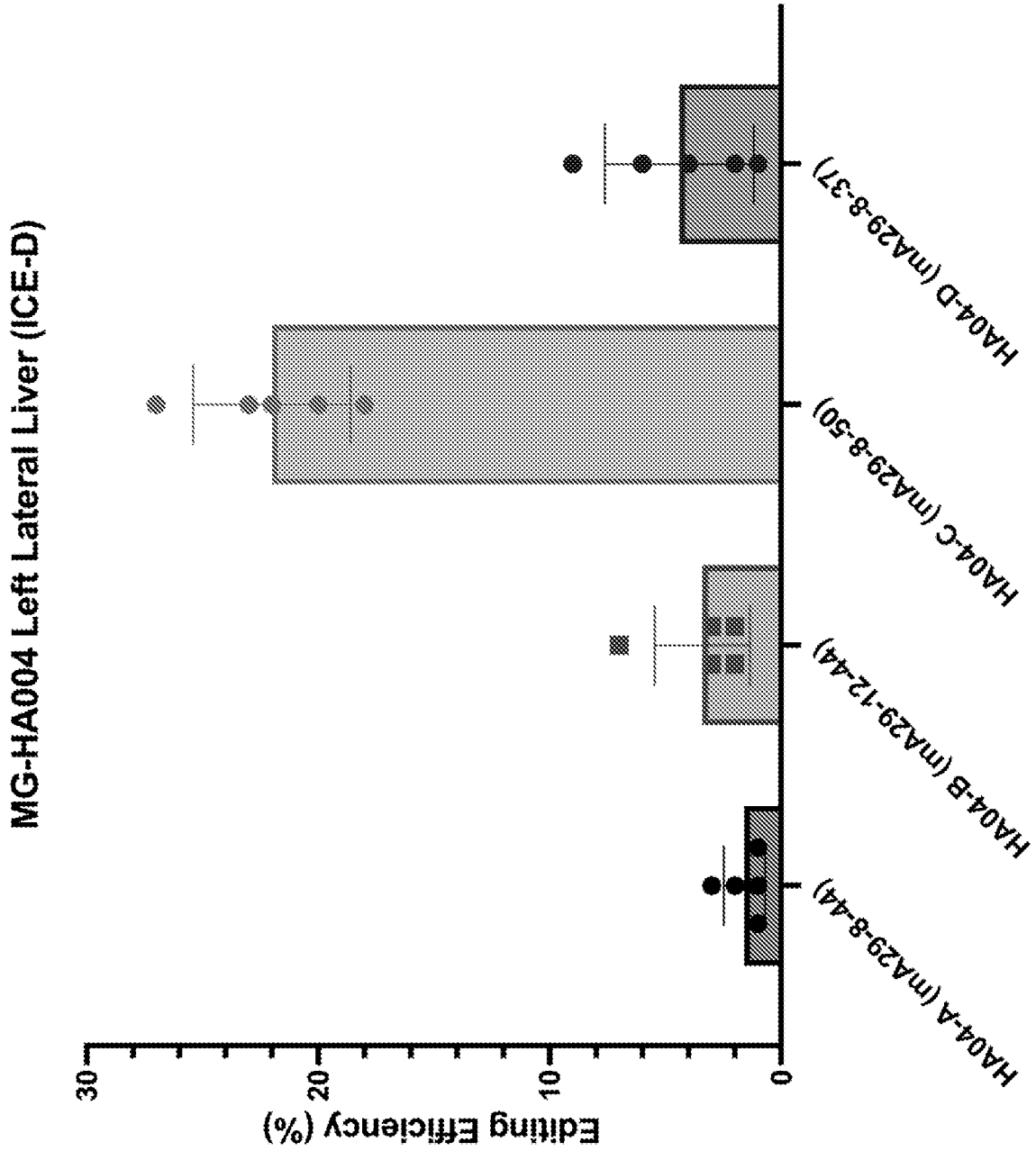

MG-HA004 Left Lateral Liver (ICE-D)
Editing Efficiency (%)
0
10
20
30
HA04-A (mA29-8-44)
HA04-B (mA29-12-44)
HA04-C (mA29-8-50)
HA04-D (mA29-8-37)

FIG. 92

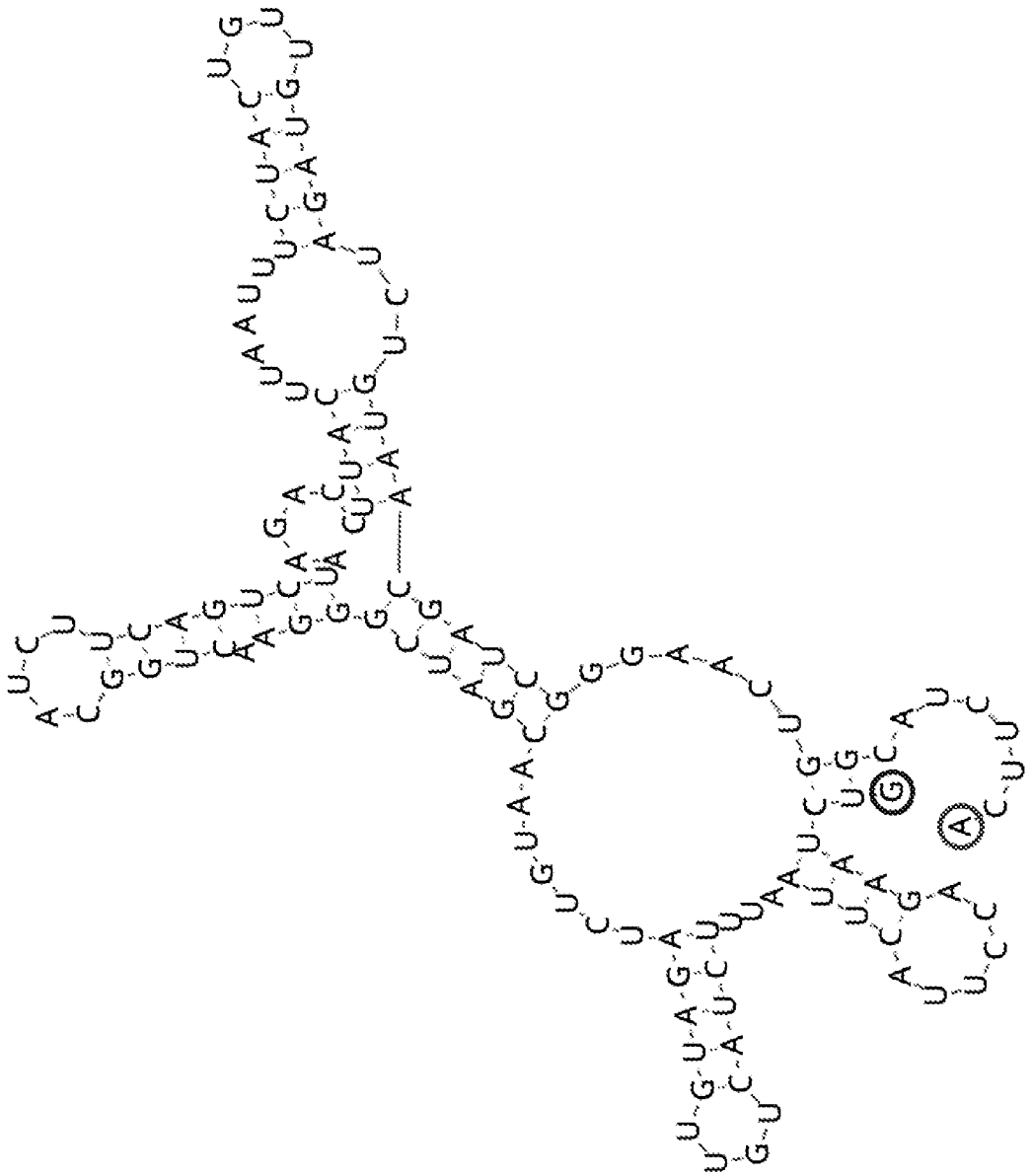


FIG. 93

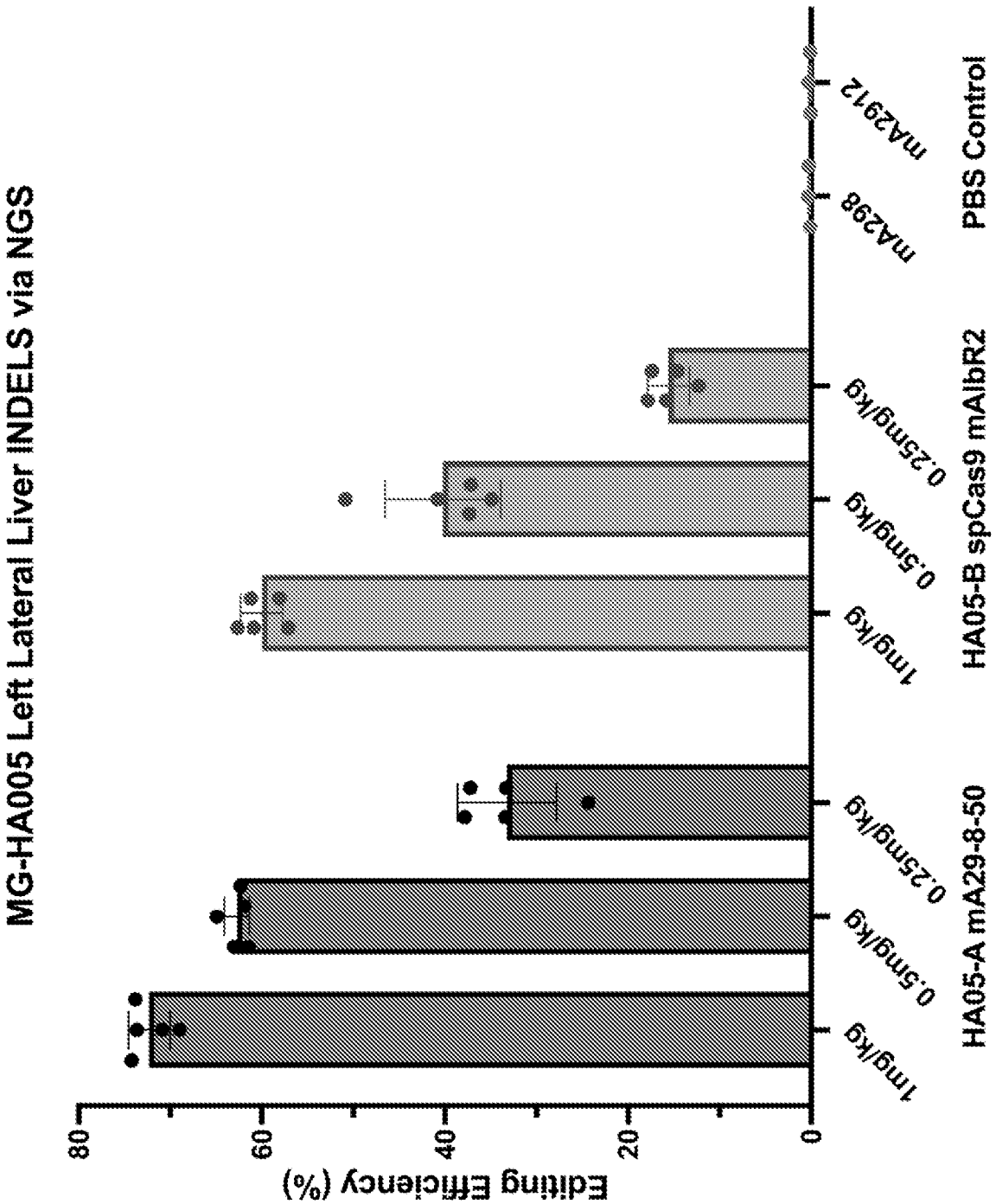

MG-HA005 Left Lateral Liver INDELS via NGS
Editing Efficiency (%)
0
20
40
60
80
1mg/kg
0.5mg/kg
0.25mg/kg
1mg/kg
0.5mg/kg
0.25mg/kg
mA298
mA2912
HA05-A mA29-8-50
HA05-B spCas9 mAlbR2
PBS Control

FIG. 94

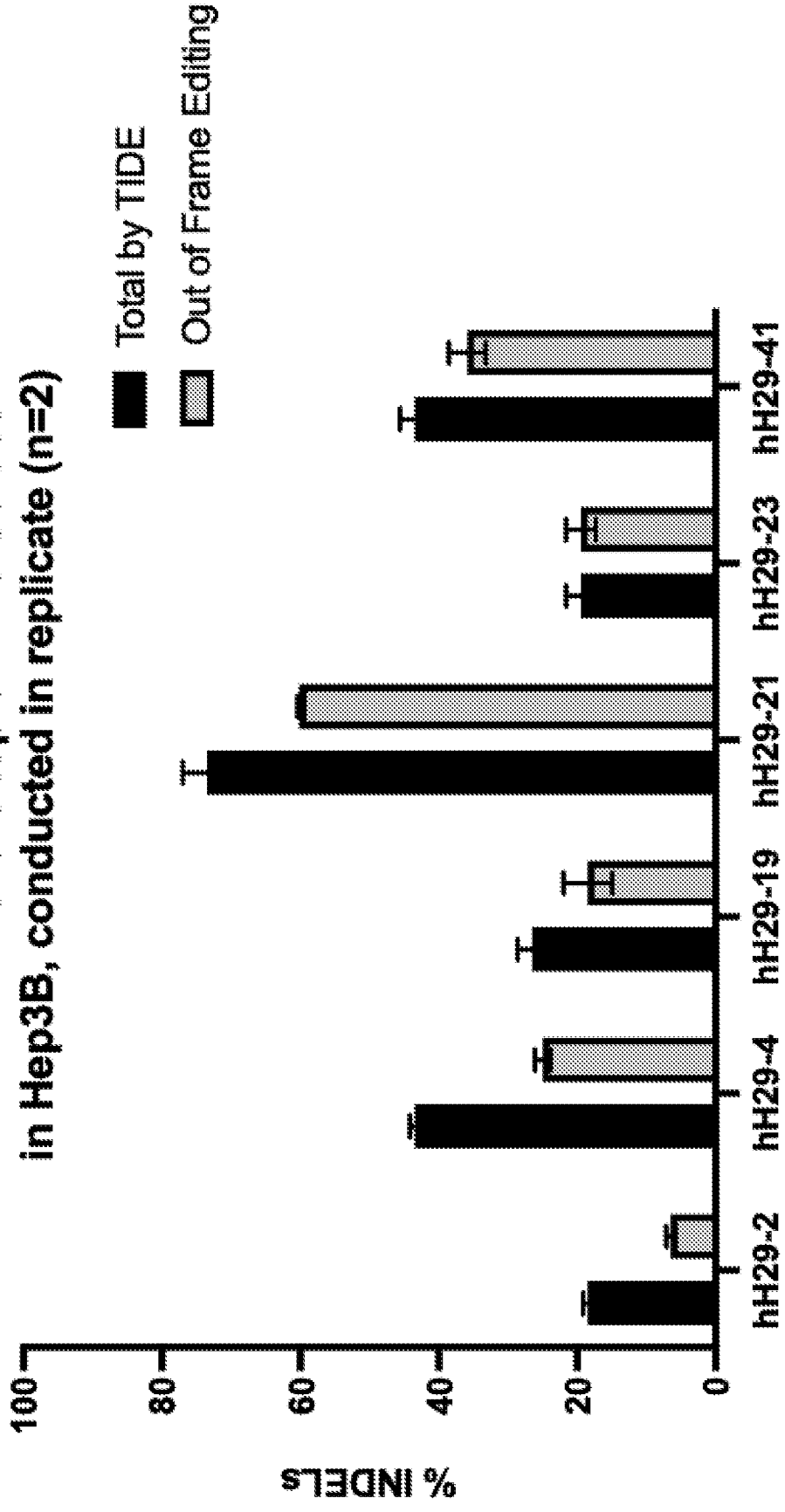

21-HAO-15 Results
mRNA TF of the top 6 hH29 Guides
in Hep3B, conducted in replicate (n=2)
Total by TIDE
Out of Frame Editing
% INDELS
100
80
60
40
20
0
hH29-2
hH29-4
hH29-19
hH29-21
hH29-23
hH29-41

FIG. 95

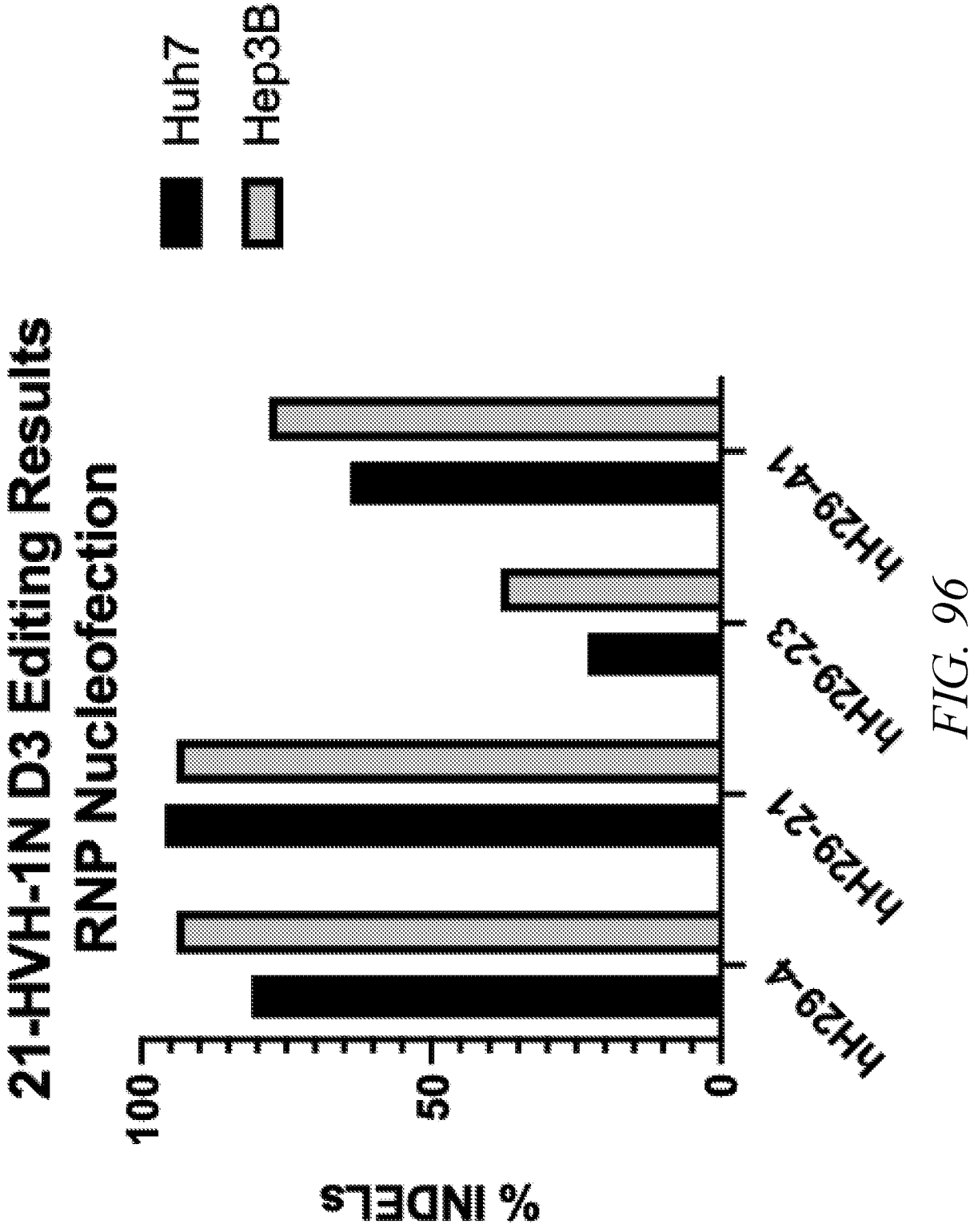

21-HVH-1N D3 Editing Results
RNP Nucleofection
Huh7
Hep3B
% INDELS
100
50
0
hH29-A
hH29-21
hH29-23
hH29-A1

FIG. 96

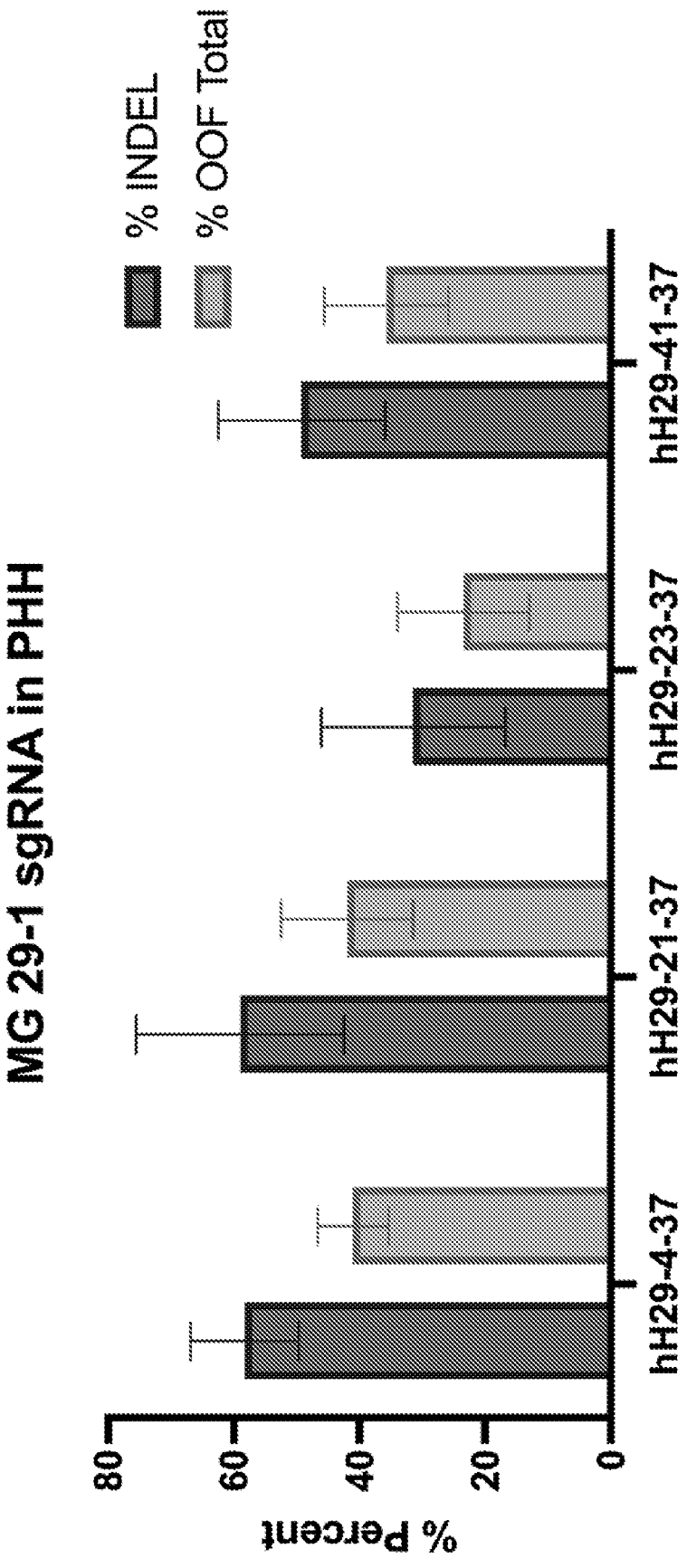

MG 29-1 sgRNA in PHH
% INDEL
% OOF Total
% Percent
0
20
40
60
80
hH29-4-37
hH29-21-37
hH29-23-37
hH29-41-37
Guide

FIG. 97

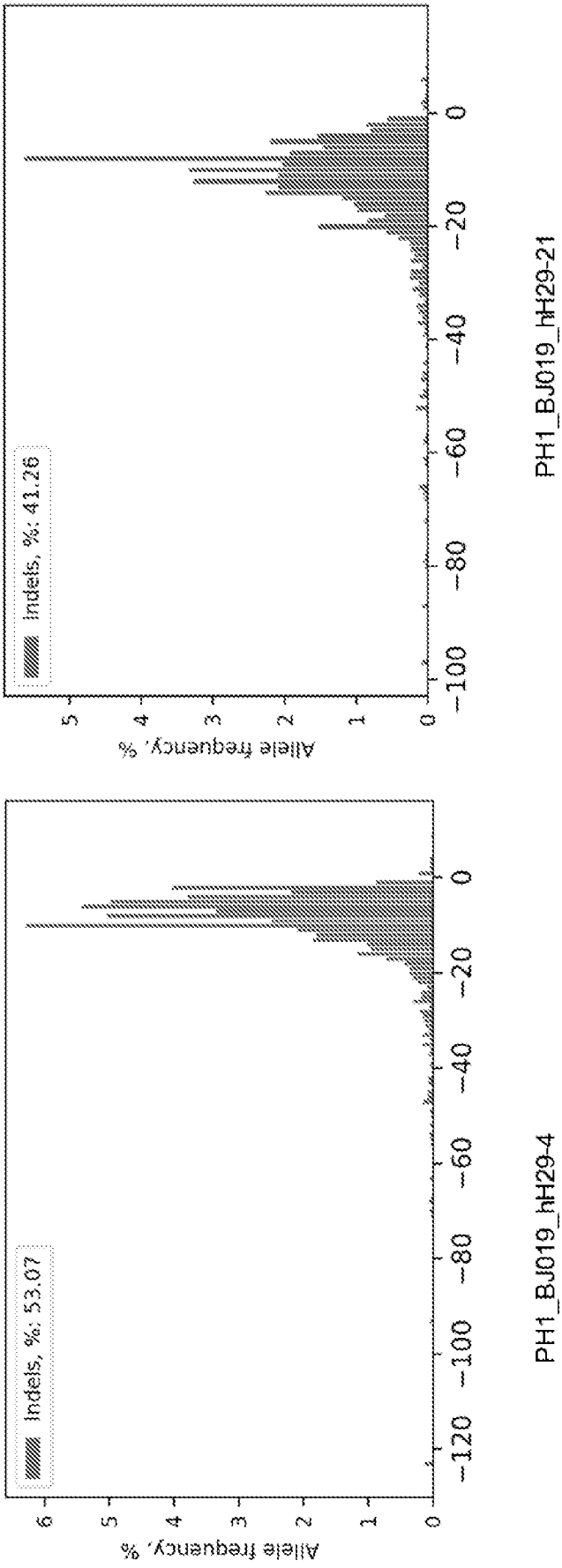

Indels, %: 53.07
Allele frequency, %
0
1
2
3
4
5
6
0
−20
−40
−60
−80
−100
−120
PH1_BJ019_hH29-4
Indels, %: 41.26
Allele frequency, %
0
1
2
3
4
5
0
−20
−40
−60
−80
−100
PH1_BJ019_hH29-21

FIG. 98A

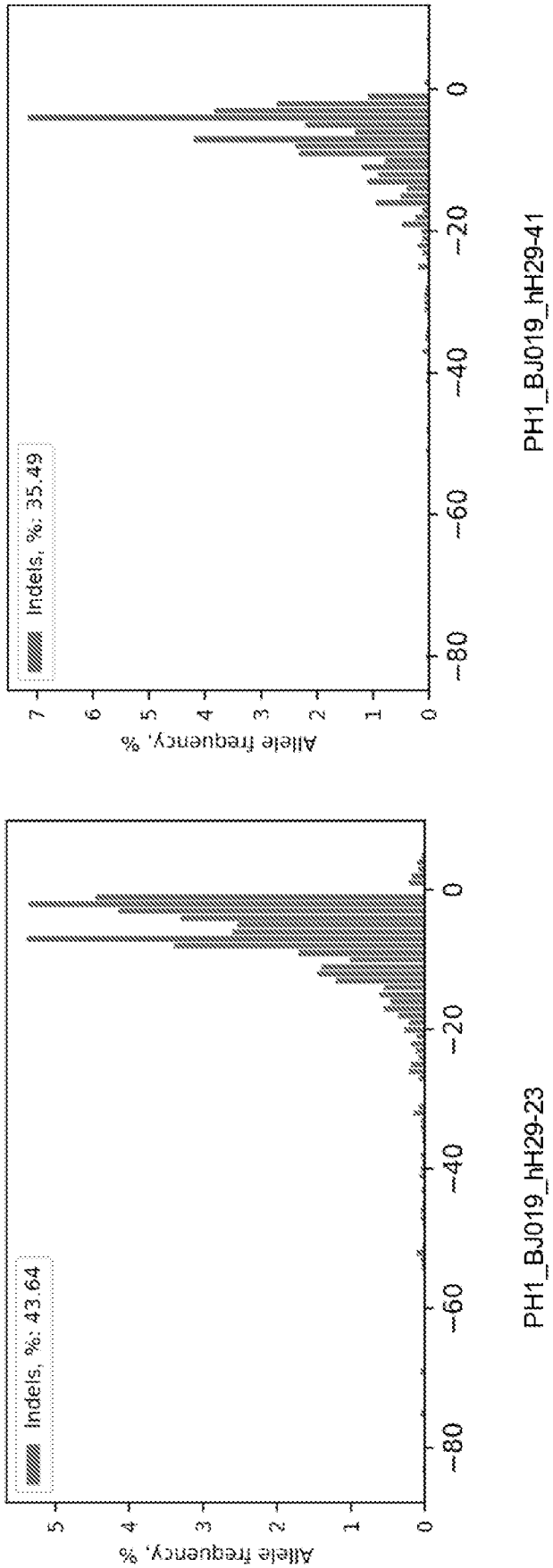

Allele frequency, %
Indels, %: 43.64
PH1_BJ019_hH29-23
Allele frequency, %
Indels, %: 35.49
PH1_BJ019_hH29-41

FIG. 98B

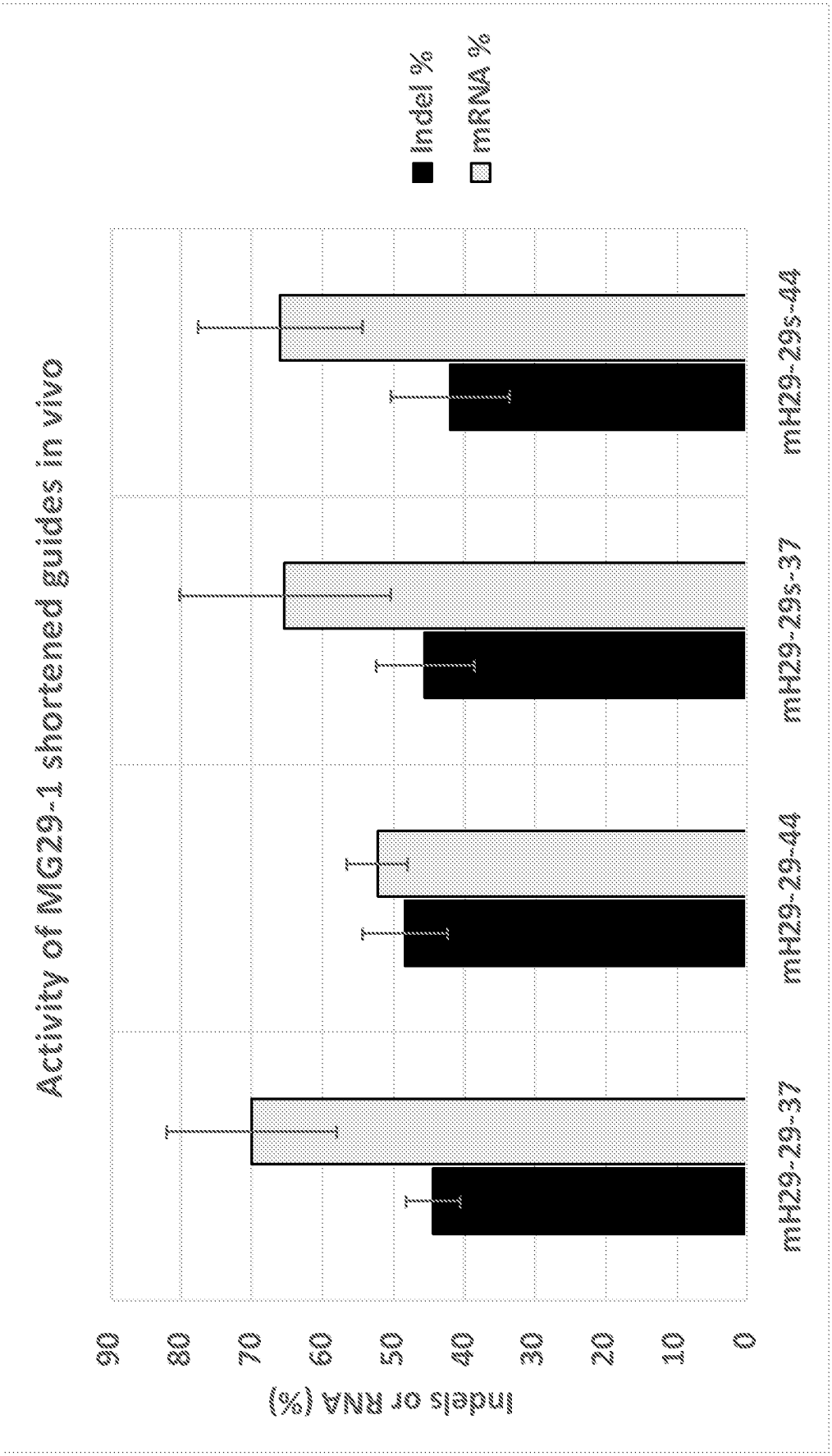

Activity of MG29-1 shortened guides in vivo
Indels or RNA (%)
0
10
20
30
40
50
60
70
80
90
Indel %
mRNA %
mH29-29-37
mH29-29-44
mH29-29s-37
mH29-29s-44

FIG. 99

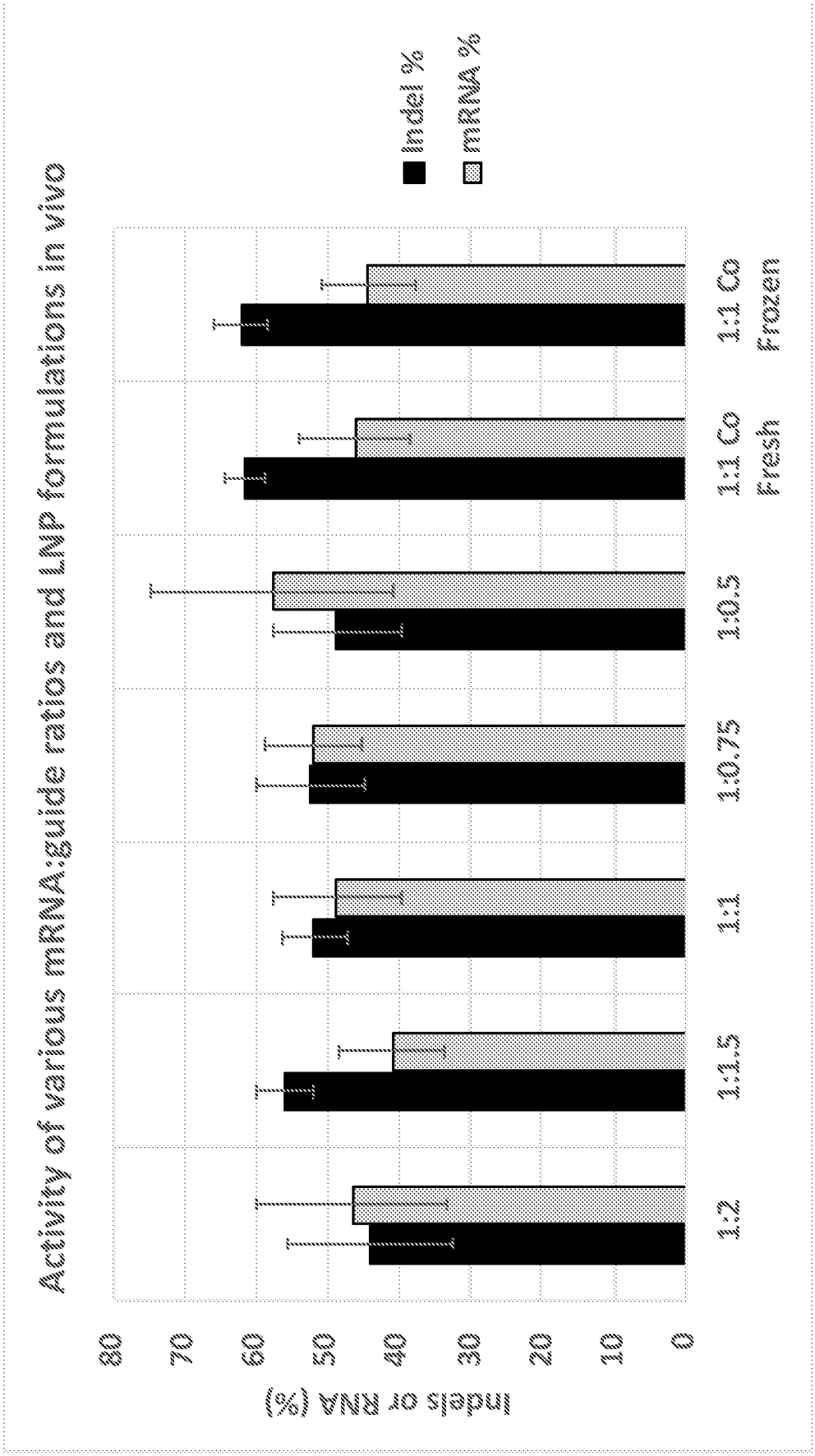

Activity of various mRNA:guide ratios and LNP formulations in vivo
Indel %
mRNA %
Indels or RNA (%)
0
10
20
30
40
50
60
70
80
1:2
1:1.5
1:1
1:0.75
1:0.5
1:1 Co Fresh
1:1 Co Frozen

FIG. 100

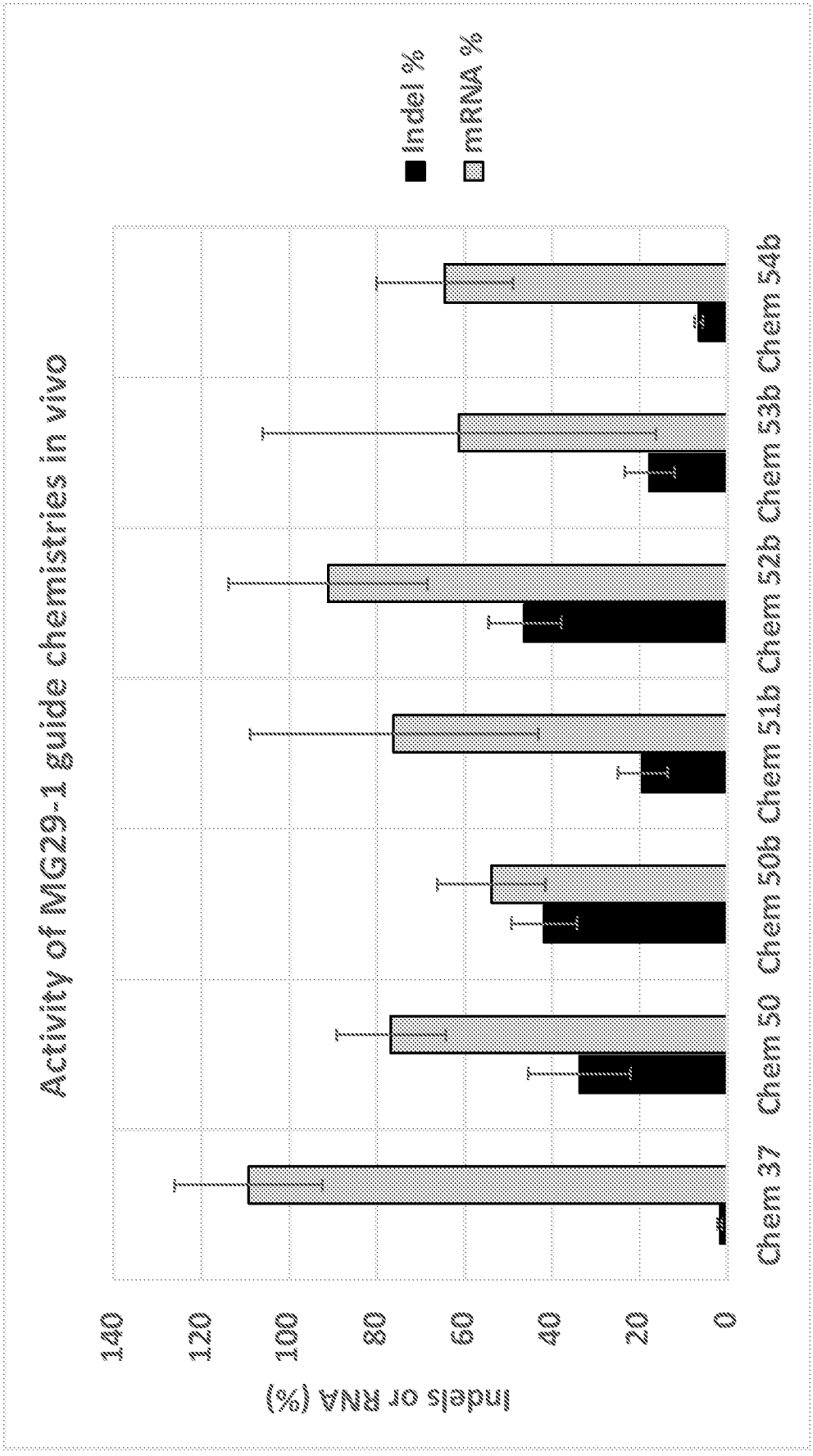

Activity of MG29-1 guide chemistries in vivo
Indel %
mRNA %
Indels or RNA (%)
0
20
40
60
80
100
120
140
Chem 37
Chem 50
Chem 50b
Chem 51b
Chem 52b
Chem 53b
Chem 54b

FIG. 101

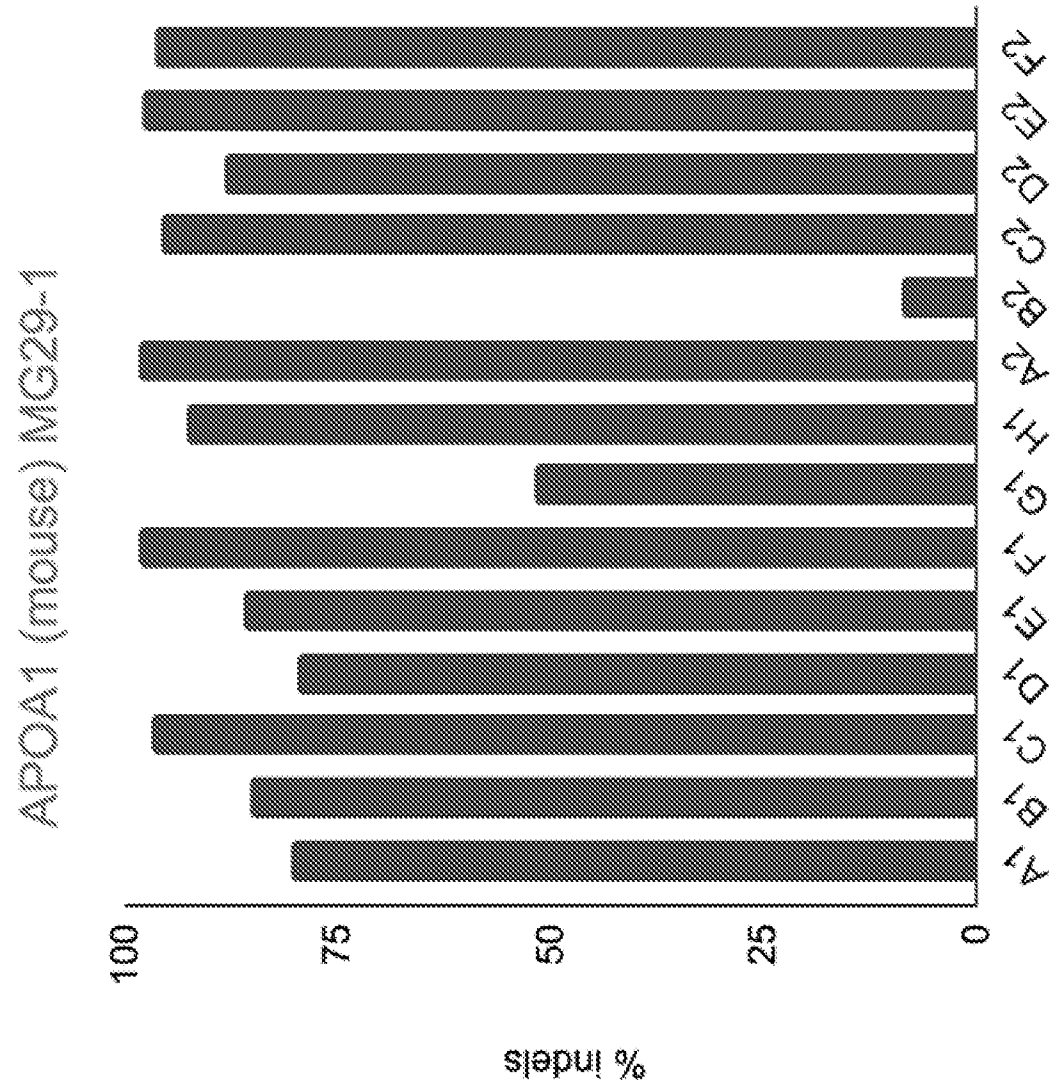

APOA1 (mouse) MG29-1
% indels
100
75
50
25
0
A1
B1
C1
D1
E1
F1
G1
H1
A2
B2
C2
D2
E2
F2
Well

FIG. 102

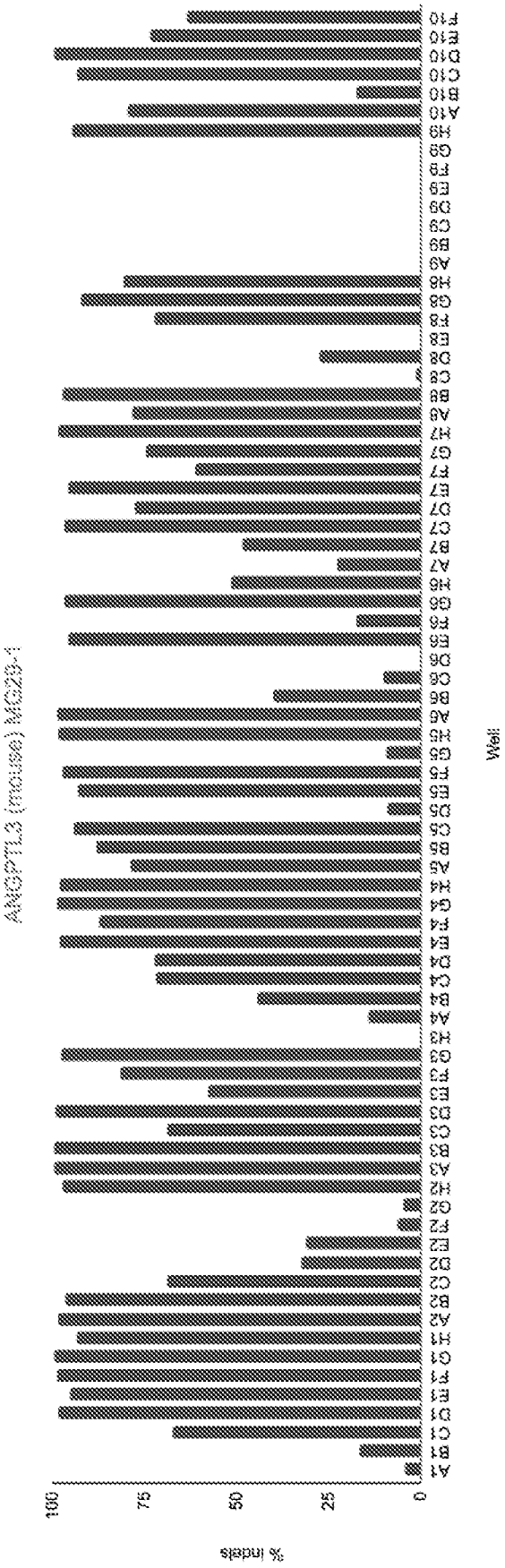

ANGPTL3 (mouse) MG29-1
% Indels
100
75
50
25
0
Well
A1 B1 C1 D1 E1 F1 G1 H1 A2 B2 C2 D2 E2 F2 G2 H2 A3 B3 C3 D3 E3 F3 G3 H3 A4 B4 C4 D4 E4 F4 G4 H4 A5 B5 C5 D5 E5 F5 G5 H5 A6 B6 C6 D6 E6 F6 G6 H6 A7 B7 C7 D7 E7 F7 G7 H7 A8 B8 C8 D8 E8 F8 G8 H8 A9 B9 C9 D9 E9 F9 G9 H9 A10 B10 C10 D10 E10 F10

FIG. 103

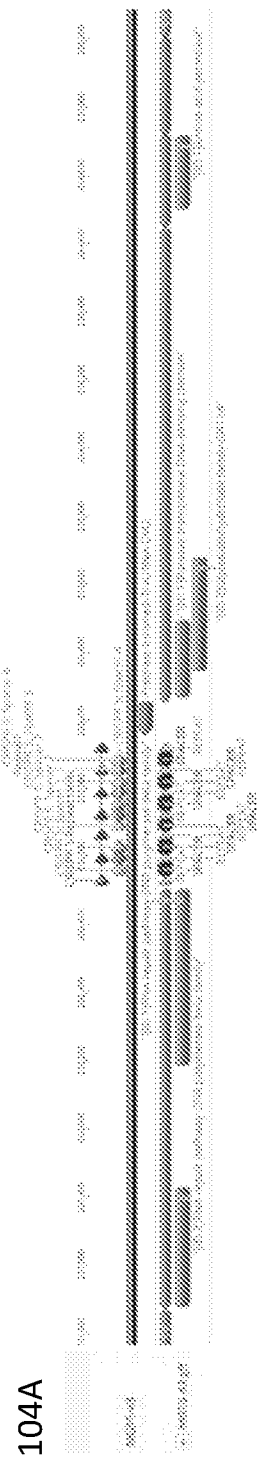


FIG. 104A

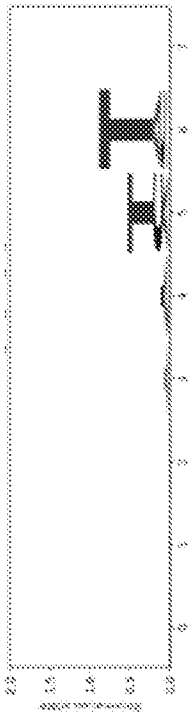


FIG. 104D

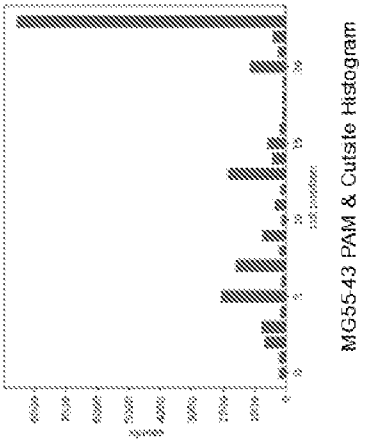


FIG. 104E

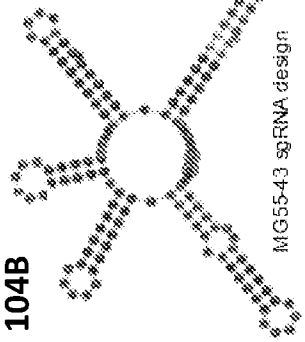


FIG. 104B

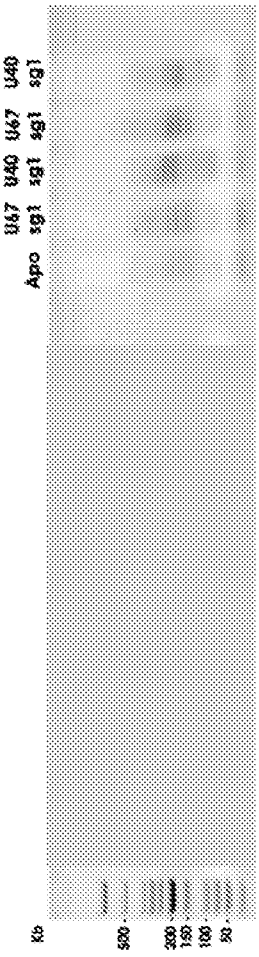


FIG. 104C

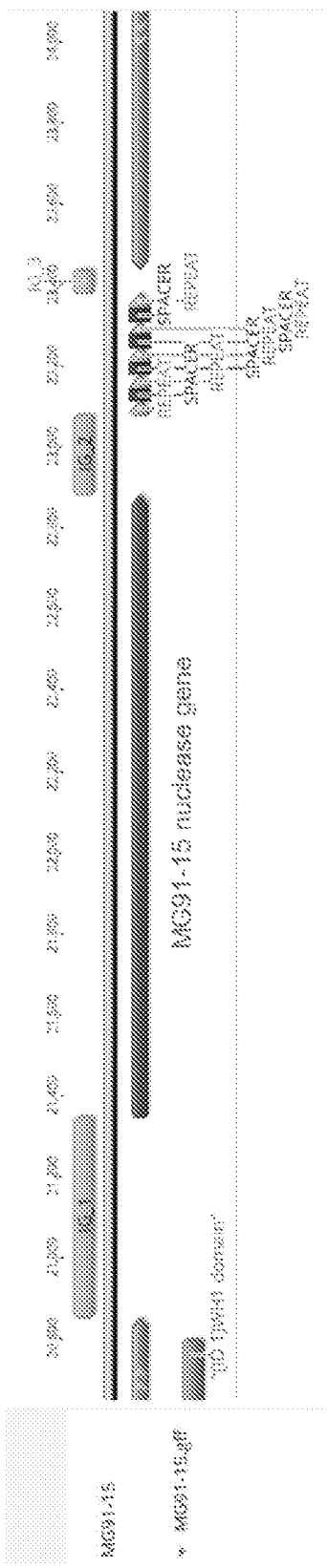


FIG. 105A

FIG. 105B

FIG. 105C

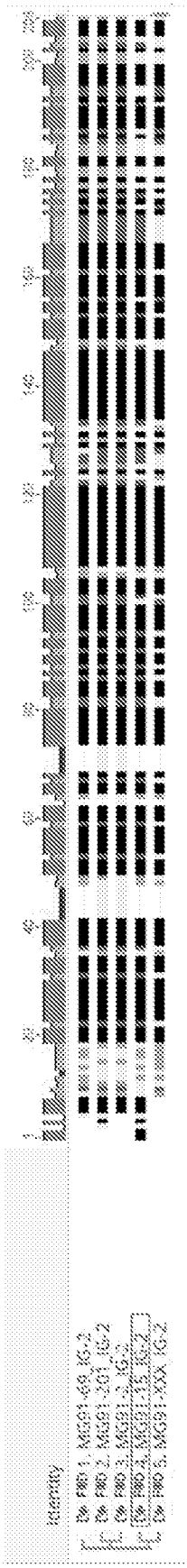


FIG. 106A

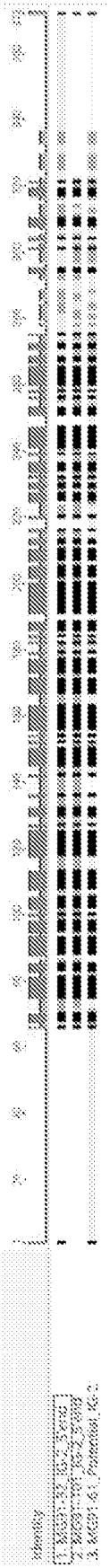
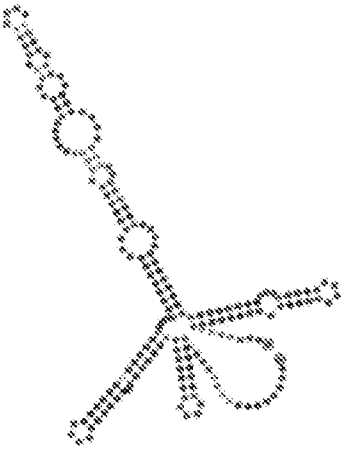


FIG. 106B



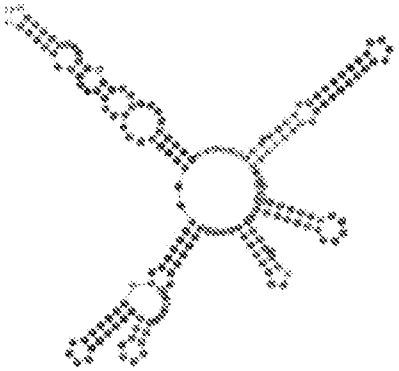
FIG. 106C

FIG. 107A



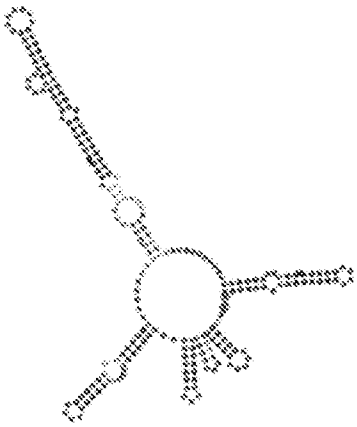
MG91-15 sgRNA without spacer

FIG. 107B

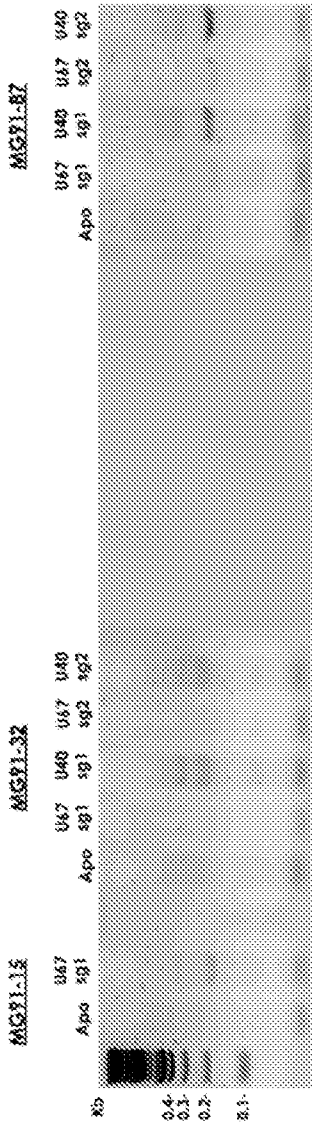


MG91-32 sgRNA without spacer

FIG. 107C

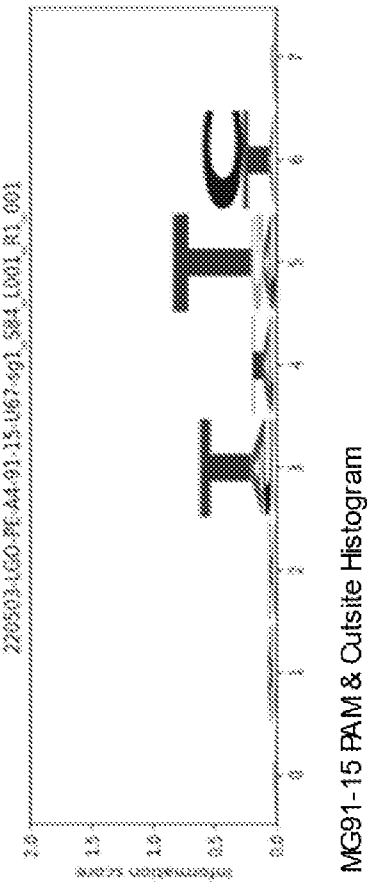


MG91-87 sgRNA without spacer



In vitro cleavage assay

FIG. 107D



MG91-15 PAM & Cutsite Histogram

FIG. 108A

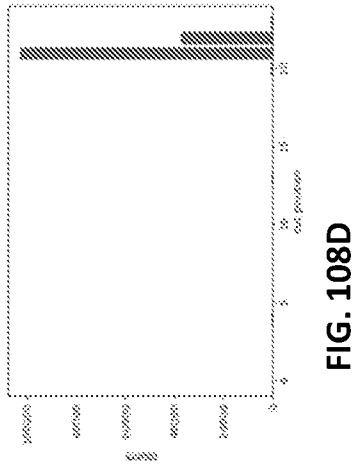
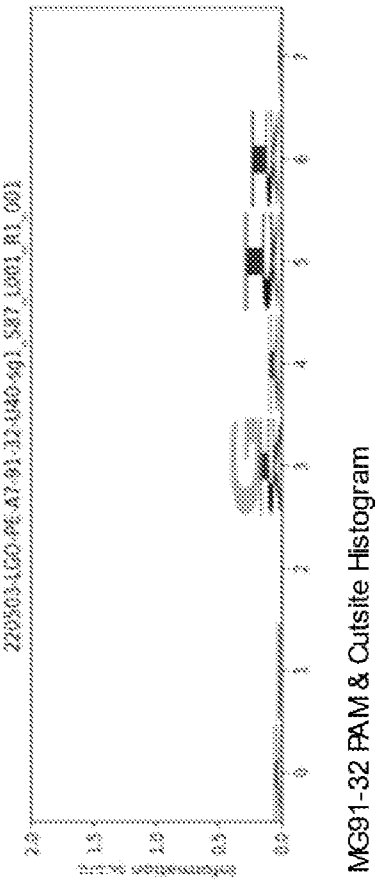


FIG. 108D



MG91-32 PAM & Cutsite Histogram

FIG. 108B

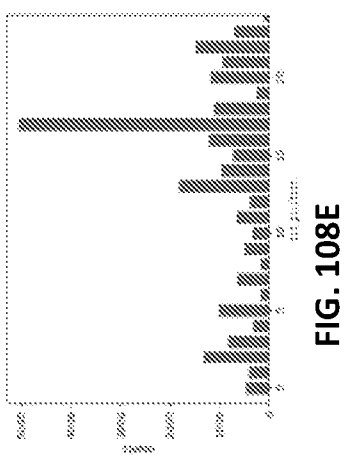
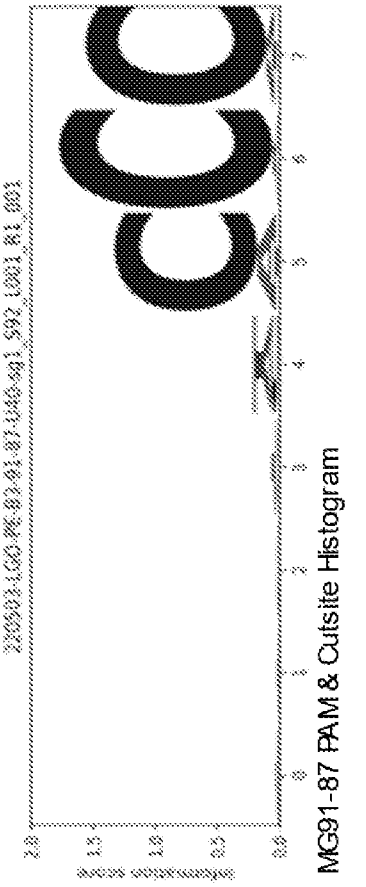


FIG. 108E



MG91-87 PAM & Cutsite Histogram

FIG. 108C

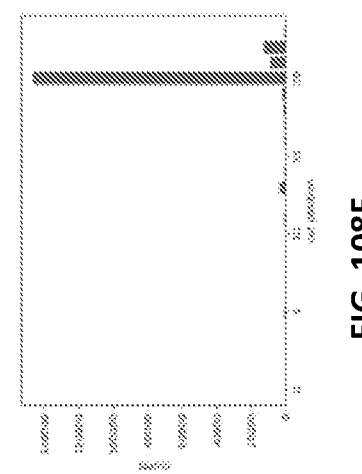
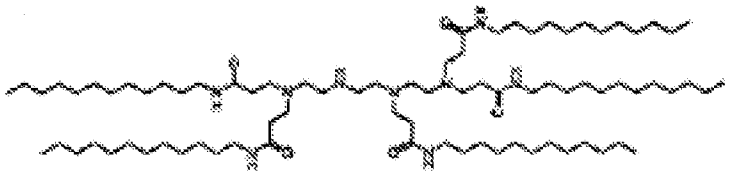


FIG. 108F

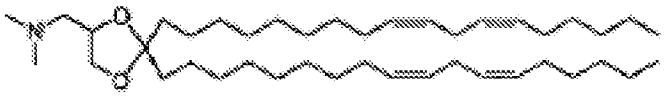
98N12-5 (TETA5-LAP)



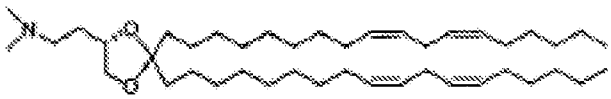
DLin DMA



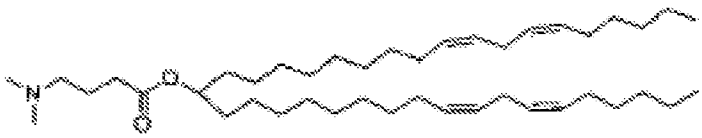
DLin-K-DMA (2,2-Dilinoleyl-4-dimethylaminomethyl-[1,3]-dioxolane)



DLin-KC2-DMA



DLin-MC3-DMA



C12-200

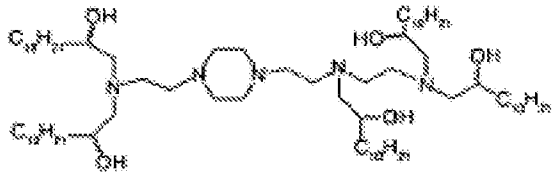


FIG. 109