



(51) International Patent Classification:

C12N 9/22 (2006.01) C12N 15/52 (2006.01)

C12N 15/10 (2006.01) C12N 15/86 (2006.01)

C12N 15/113 (2010.01) C12N 15/90 (2006.01)

(21) International Application Number:

PCT/US2024/056299

(22) International Filing Date:

15 November 2024 (15.11.2024)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/599,932 16 November 2023 (16.11.2023) US

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

(72) Inventors: THOMAS, Brian C.; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, California 94608 (US). ALEXANDER, Lisa; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, California 94608 (US). BROWN, Christopher; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, California 94608 (US). GOLTSMAN, Daniela S.A.; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, California 94608 (US). LAMOTHE, Rebecca; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, California 94608 (US). LAPERRIERE, Sarah; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, California 94608 (US). NOCEDAL,

Isabel; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, California 94608 (US). RALLAPALLI, Kartik; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, California 94608 (US). THOMAS, Nicole; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, California 94608 (US). D'ESPAUX, Leopold; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, California 94608 (US). BROTHERS, Molly; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, California 94608 (US). BELUR, Shweta; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, California 94608 (US). HACKNEY, Margaret; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, California 94608 (US). LUONG, Thao; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, California 94608 (US). WANG, Junkai; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, CA 94608 (US). ANGGITO, Bryan; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, CA 94608 (US). LE, Dinh; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, CA 94608 (US). PEKAREK, Brandon; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, CA 94608 (US). SENGER, Kate; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, CA 94608 (US). SZYMANSKI, Paul; c/o METAGENOMI, INC., 1545 Park Avenue, Emeryville, CA 94608 (US).

(74) Agent: DUSABAN GONZALES, Stephanie S. et al.; Goodwin Procter LLP, IP DOCKETING DEPT./7TH FL, 100 Northern Avenue, Boston, Massachusetts 02210 (US).

(54) Title: ENZYMES WITH RUV C DOMAINS

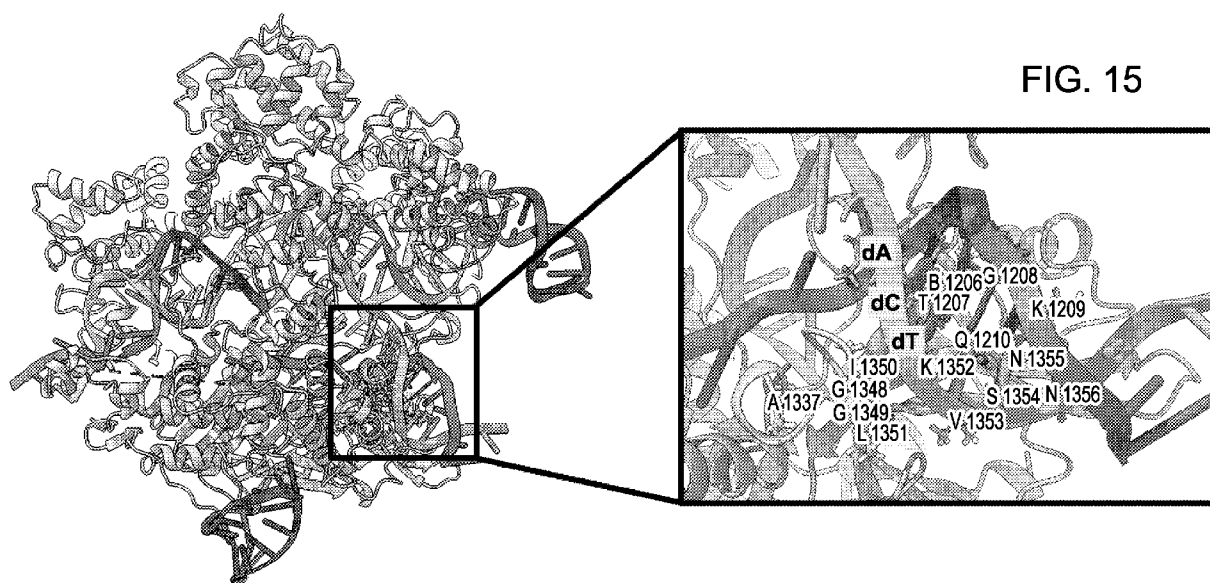


FIG. 15

(57) Abstract: The present disclosure provides for endonuclease enzymes having distinguishing domain features, as well as methods of using such enzymes or variants thereof.

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

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

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:

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

ENZYMES WITH RUVB DOMAINS

CROSS-REFERENCE

[0001] This application claims the benefit of and priority to U.S. Provisional Patent Application No. 63/599,932 filed November 16, 2023, which is incorporated by reference in its entirety herein.

SUMMARY

[0002] Described herein, in certain embodiments, are engineered nuclease systems comprising: an endonuclease comprising a sequence having at least 70% sequence identity to any one of SEQ ID NOs: 47-52, 659-698, 983-1010, 1263-1282, 1158-1159, and 975-981; and an engineered guide polynucleotide configured to form a complex with the endonuclease and to hybridize to a target nucleic acid sequence.

[0003] In some embodiments, the endonuclease comprises a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 47-52, 659-698, 983-1010, 1263-1282, 1158-1159, and 975-981.

[0004] In some embodiments, the endonuclease comprises a sequence having 90% sequence identity to any one of SEQ ID NOs: 47-52, 659-698, 983-1010, 1263-1282, 1158-1159, and 975-981.

[0005] In some embodiments, the engineered guide polynucleotide comprises a crRNA and a tracrRNA.

[0006] In some embodiments, the tracrRNA comprises a sequence having at least 90% sequence identity to any one of SEQ ID NOs: 710-722, 726-744, 745-767, 699, 700-702, and 703-709.

[0007] In some embodiments, the tracrRNA comprises a sequence having 100% sequence identity to any one of SEQ ID NOs: 710-722, 726-744, 745-767, 699, 700-702, and 703-709.

[0008] In some embodiments, the engineered guide polynucleotide is a single guide nucleic acid.

[0009] In some embodiments, the engineered guide polynucleotide is a dual guide nucleic acid.

[0010] In some embodiments, the engineered guide polynucleotide is RNA.

[0011] In some embodiments, the endonuclease is not a Cas9 endonuclease.

[0012] In some embodiments, the endonuclease has less than 80% identity to a Cas9 endonuclease.

[0013] In some embodiments, the endonuclease binds non-covalently to the engineered guide polynucleotide.

[0014] In some embodiments, the endonuclease is covalently linked to the engineered guide polynucleotide.

[0015] Described herein, in certain embodiments, are engineered nuclease systems comprising an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NO: 47; and an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 1137-1144, 1283-1392, 1502-1509, and 1786-2045.

[0016] Described herein, in certain embodiments, are engineered nuclease systems comprising an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NO: 48; and an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to SEQ ID NOs: 1153-1156, 1393-1493, 1494-1501, and 1510-1525.

[0017] Described herein, in certain embodiments, are engineered nuclease systems comprising an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263; and an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 57-58, 77-88, 101-119, 139-150, 163-181, 201-212, 225-243, 343-374, 407-413, 421-433, 447-453, 461-472, 485-491, 499-511, 525-531, 539-551, 565-578, 593-625, 710-722, 927-942, 945-950, 961-962, 969-970, 1906-1931, 1083-1094, 1096-1102, 1113-1122, 1123-1137, and 2433-2434.

[0018] Described herein, in certain embodiments, are engineered nuclease systems comprising an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NO: 51 or SEQ ID NO: 1264; and an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to SEQ ID NO: 61.

[0019] Described herein, in certain embodiments, are engineered nuclease systems comprising an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NOs: 659, 660, 1158, 1159, and 1267-1277; and an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 726-744, 843-880, 1011-1055, and 1145-1152.

[0020] Described herein, in certain embodiments, are engineered nuclease systems comprising an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 661-678 and 1278-1282; and an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 745-767 and 881-926.

[0021] Described herein, in certain embodiments, are engineered nuclease systems comprising an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NO: 659; and an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 699 and 973-974.

[0022] Described herein, in certain embodiments, are engineered nuclease systems comprising an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 696-698; and an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 700-702 and 957-960.

[0023] Described herein, in certain embodiments, are engineered nuclease systems comprising an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 975-981; and an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 703-709, 943-944, 951-956, 965-968, and 971-972.

[0024] In some embodiments, the engineered guide polynucleotide is a single guide nucleic acid.

[0025] In some embodiments, the engineered guide polynucleotide is a dual guide nucleic acid.

[0026] In some embodiments, the engineered guide polynucleotide is RNA.

[0027] In some embodiments, the endonuclease is not a Cas9 endonuclease.

[0028] In some embodiments, the endonuclease has less than 80% identity to a Cas9 endonuclease.

[0029] In some embodiments, the endonuclease binds non-covalently to the engineered guide polynucleotide.

[0030] In some embodiments, the endonuclease is covalently linked to the engineered guide polynucleotide.

[0031] In some embodiments, the endonuclease is fused to the engineered guide polynucleotide.

[0032] Described herein, in certain embodiments, are methods for modifying a target nucleic acid sequence comprising contacting the target nucleic acid sequence using the engineered nuclease system disclosed herein.

[0033] In some embodiments, modifying the target nucleic acid sequence comprises binding, nicking, or cleaving, the target nucleic acid sequence.

[0034] In some embodiments, the target nucleic acid sequence comprises genomic DNA, viral DNA, viral RNA, or bacterial DNA.

[0035] In some embodiments, the modification is in vitro.

[0036] In some embodiments, the modification is in vivo.

[0037] In some embodiments, the modification is ex vivo.

[0038] In some embodiments, the target nucleic acid sequence comprises a sequence of any one of SEQ ID NOs: 1161-1262, 1526-1785, and 2046-2050.

[0039] Described herein, in certain embodiments, are methods of modifying a target nucleic acid sequence in a mammalian cell comprising contacting the mammalian cell using the engineered nuclease systems disclosed herein.

[0040] In some embodiments, the method further comprises selecting cells comprising the modification.

[0041] Provided here are methods of modifying a hydroxyacid oxidase 1 (HAO1) gene comprising contacting the HAO1 gene using an engineered nuclease system comprising an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NO: 47; and an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 1283-1392 and 1502-1509.

[0042] In some embodiments, the engineered guide polynucleotide comprises a sequence of any one of SEQ ID NOs: 1283-1392 and 1502-1509.

[0043] Described herein, in certain embodiments, are methods of modifying ATPase copper transporting beta (ATP7B) comprising contacting ATP7B using an engineered nuclease system comprising an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263; and an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence.

[0044] In some embodiments, the engineered guide polynucleotide comprises a sequence of any one of SEQ ID NOs: 1056-1081.

[0045] In some embodiments, the target nucleic acid sequence comprises a sequence of any one of SEQ ID NOs: 1206-1231.

[0046] Described herein, in certain embodiments, are methods of modifying adeno-associated virus integration site 1 (AAVS1) comprising contacting AAVS1 using an engineered nuclease system comprising an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1010, 1263, 47, 659, 660, 1158, 1159, and 1267-1277; and an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 710-722.

[0047] In some embodiments, the engineered guide polynucleotide comprises a sequence of any one of SEQ ID NOs: 57-58, 101-119, 163-181, 225-243, 343-374, 21-433, 461-472, 499-511, 539-551, 1082, 1098-1102, 1393-1493, 1020-1049, and 1051-1055.

[0048] In some embodiments, the target nucleic acid sequence comprises a sequence of any one of SEQ ID NOs: 59-60, 120-138, 182-200, 243-262, 375-406, 434-446, 473-484, 512-524, 552-564, 1170-1205, 1232, and 1248-1252.

[0049] Described herein, in certain embodiments, are methods of modifying T cell receptor alpha constant (TRAC) comprising contacting TRAC using an engineered nuclease system comprising an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1010, 1263, 48, 659, 660, 1158, 1159, and 1267-1277; and an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 710-722.

[0050] In some embodiments, the engineered guide polynucleotide comprises a sequence of any one of SEQ ID NOs: 61, 77-88, 139-150, 201-212, 407-413, 447-453, 485-491, 525-531, 1096-1097, 1011-1019, 1050, 1494-1501, and 1510-1525.

[0051] In some embodiments, the target nucleic acid sequence comprises a sequence of any one of SEQ ID NOs: 62, 89-100, 151-162, 213-224, 414-420, 454-460, 492-489, 532-538, 1161-1168, and 1200.

[0052] Described herein, in certain embodiments, are methods of modifying an albumin gene comprising contacting the albumin gene using an engineered nuclease system comprising an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NO: 49-51, 679-694, 983-1002, 1003-1010, and 1263; and an engineered guide polynucleotide

configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence.

[0053] In some embodiments, the engineered guide polynucleotide comprises a sequence having any SEQ ID NO: 1083-1094.

[0054] In some embodiments, the target nucleic acid sequence comprises a sequence having SEQ ID NO: 1233-1244.

[0055] Described herein, in certain embodiments, are methods of modifying a beta-2-microglobulin (B2M) gene comprising contacting the B2M gene using an engineered nuclease system comprising an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263; and an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 710-722.

[0056] In some embodiments, the engineered guide polynucleotide comprises a sequence of any one of SEQ ID NOs: 565-578 and 1095.

[0057] In some embodiments, the target nucleic acid sequence comprises a sequence of any one of SEQ ID NOs: 579-592 and 1245.

[0058] Described herein, in certain embodiments, are methods of modifying a hemoglobin subunit beta (HBB) gene comprising contacting the HBB gene using an engineered nuclease system comprising an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263; and an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 710-722.

[0059] In some embodiments, the engineered guide polynucleotide comprises a sequence of any one of SEQ ID NOs: 593-625.

[0060] In some embodiments, the target nucleic acid sequence comprises a sequence of any one of SEQ ID NOs: 626-658.

[0061] Described herein, in certain embodiments, are methods of modifying a phenylalanine hydroxylase (PAH) gene comprising contacting the PAH gene using an engineered nuclease system comprising an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263; and an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence.

[0062] In some embodiments, the engineered guide polynucleotide comprises a sequence of any one of SEQ ID NOs: 1113-1122.

[0063] In some embodiments, the target nucleic acid sequence comprises a sequence of any one of SEQ ID NOs: 1253-1262.

[0064] Described herein, in certain embodiments, are methods of modifying an ataxin 2 (ATXN2) gene comprising contacting the ATXN2 gene using an engineered nuclease system comprising an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NO: 47; and an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence.

[0065] In some embodiments, the engineered guide polynucleotide comprises a sequence of any one of SEQ ID NOs: 1786-2045, and 2051-2055.

[0066] In some embodiments, the target nucleic acid sequence comprises a sequence of any one of SEQ ID NOs: 1526-1785, and 2046-2050.

[0067] Described herein, in certain embodiments, are cells comprising the engineered nuclease systems disclosed herein.

[0068] In some embodiments, the cell is a eukaryotic cell. In some embodiments, the cell is a mammalian cell. In some embodiments, the cell is an immortalized cell. In some embodiments, the cell is an insect cell. In some embodiments, the cell is a yeast cell. In some embodiments, the cell is a plant cell. In some embodiments, the cell is a fungal cell. In some embodiments, the cell is a prokaryotic cell.

[0069] In some embodiments, the cell is an A549, HEK-293, HEK-293T, BHK, CHO, HeLa, MRC5, Sf9, Cos-1, Cos-7, Vero, BSC 1, BSC 40, BMT 10, WI38, HeLa, Saos, C2C12, L cell, HT1080, HepG2, Huh7, K562, primary cell, or a derivative thereof.

[0070] In some embodiments, the cell is an engineered cell.

[0071] In some embodiments, the cell is a stable cell.

[0072] 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.

BRIEF DESCRIPTION OF THE DRAWINGS

[0073] The novel features of the disclosure are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present disclosure will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the disclosure are utilized, and the accompanying drawings (also “Figure” and “FIG.” herein), of which:

[0074] **FIG. 1** depicts the gene editing outcomes at the DNA level for TRAC (A1-D2) and AAVS1 (E2-G4) in K562 cells as described in Example 4.

[0075] **FIG. 2** depicts the gene editing outcomes at the DNA level for TRAC in K562 cells as described in Example 5.

[0076] **FIG. 3A** depicts analysis of the gene editing outcomes of first engineering round of the MG71-2 scaffold with guide hAAVS1 C3 at the DNA level.

[0077] **FIG. 3B** depicts analysis at the DNA level of second round of engineering for MG71-2 scaffold with guide hAAVS1 C3.

[0078] **FIG. 4** depicts analysis of the gene editing outcomes with original vs optimized MG71-2 scaffold at the DNA levels. Spacers indicated on the Y axis were tested with either a 22nt or 24nt spacer and either the original or shortened MG71-2 scaffold 11.

[0079] **FIGs. 5A-5B** depict analysis of the gene editing outcomes with original vs optimized MG71-2 scaffold at the DNA levels. **FIG. 5A** depicts editing outcomes with MG71-2 engineered guides (scaffold 21) at exon 2 of the human B2M gene. **FIG. 5B** depicts editing outcomes with MG71-2 engineered guides (scaffold 21) at exon 3 of the human HBB gene.

[0080] **FIG. 6** depicts a phylogenetic tree of the reconstructed ancestral variants of the MG71 family based on multiple sequence alignment. Ancestral sequences were generated for nodes of interest (shown with black circles).

[0081] **FIG. 7** depicts a SeqLogo of the protospacer adjacent motif (PAM) recognized by MG71-28 obtained from the in vitro cleavage assay and NGS sequencing. The MG71 family display a range of PAM specificities, with preference for an A in the 4th position.

[0082] **FIG. 8** depicts a SeqLogo of the protospacer adjacent motif (PAM) recognized by MG73-6 obtained from the in vitro cleavage assay and NGS sequencing. The nuclease MG73-6 was active in vitro with an nnRNCTW PAM.

[0083] **FIG. 9** depicts a SeqLogo of the protospacer adjacent motif (PAM) recognized by MG74-8 obtained from the in vitro cleavage assay and NGS sequencing. The MG74 family has A/C rich PAMs.

[0084] **FIG. 10** depicts a SeqLogo of the protospacer adjacent motif (PAM) recognized by MG87-102 obtained from the in vitro cleavage assay and NGS sequencing. The MG87 family nucleases have the strongest preference in the 5th and 6th bases from the spacer.

[0085] **FIG. 11** depicts a SeqLogo of the protospacer adjacent motif (PAM) recognized by MG88-11 obtained from the in vitro cleavage assay and NGS sequencing. The MG88 family tend to have purine-rich PAMs in the second through fourth positions.

[0086] **FIG. 12A-12B** depict SeqLogos of the protospacer adjacent motifs (PAMs) recognized by MG71 (**FIG. 12A: MG71-42; FIG. 12B: MG71-43**) ancestors obtained from the in vitro cleavage assay and NGS sequencing. The MG71 ASR proteins are active and have similar or more relaxed PAMs as the metagenomic protein MG71-2.

[0087] **FIG. 13** depicts the gene editing outcomes at the DNA level (shown in % indel) for AAVS1 in K562 cells.

[0088] **FIG. 14** depicts a phylogenetic tree of the reconstructed ancestral variants of the MG71 family. The phylogenetic tree was inferred with FastTree from a MAFFT multiple sequence alignment. Ancestral sequences were generated for nodes of interest (shown with white circles).

[0089] **FIG. 15** depicts alphafold2 predicted structure of MG71-2 with the gRNA and R-loop from crystal structure of SpCas9 bound to substrate DNA and guide (PDB ID: 4UN3). FIG. 15 shows predicted structure of MG71-2 after PID mutations. Residues predicted to be within the first interaction shell (5Å) of the PAM sequence are depicted as sticks and labelled.

[0090] **FIG. 16** depicts in vitro PAMs of MG71-2 and MG71-43 single point mutants. The nucleases were expressed in an in vitro transcription/translation reaction, then incubated with guide a 5N plasmid library. Cut plasmids were sequenced by NGS and aligned to generate the PAM. The first nucleotide “T” in the 0 position is masked as it represents the constant flanking base adjacent to the spacer in the 5N PAM enrichment library (TNNNNN).

[0091] **FIG. 17** depicts in vitro PAMs of MG71 ancestral reconstructions. The nucleases were expressed in an in vitro transcription/translation reaction, then incubated with guide a 5N plasmid library. Cut plasmids were sequenced by NGS and aligned to generate the PAM. The first nucleotide “T” in the 0 position is masked as it represents the constant flanking base adjacent to the spacer in the 5N PAM enrichment library (TNNNNN).

[0092] **FIG. 18** depicts MG87-70 tested for gene editing activity in human cells (K562 cells) with two doses of guide via nucleofection. 72 hrs later cells were harvested and gDNA was prepped to evaluate editing via NGS.

[0093] **FIG. 19** depicts MG87-21 tested for gene editing activity in human cells (K562 cells) with 650 pmol of gRNA and 500 ng of mRNA per well via nucleofection. 72 hrs later cells were harvested and gDNA was prepped to evaluate editing via NGS.

[0094] **FIG. 20** depicts MG71-2 and MG71-43 at ATP7B in PHH cells screened for active guides to the intronic region of ATP7B, between exon 5 and exon 6. Of 94 guides screened, 24 guides were identified as active >1% in PHH cells via NGS.

[0095] **FIG. 21** depicts MG71-2 and MG71-43 at ATP7B in Hep3B cells screened for active guides to the intronic region of ATP7B, between exon 5 and exon 6. Of 94 guides screened, 26 guides were identified as active >1% in Hep3B cells via NGS.

[0096] **FIG. 22** depicts MG71-2 and MG71-43 at Albumin in PHH cells screened for active guides to the intron 1 region of albumin. Of 19 guides screened, 10 guides were identified as active >1% in PHH cells for one or both nucleases. Additionally, 11 guides were designed only for MG71-43; of these guides, 2 were active in PHH.

[0097] **FIGs. 23A-23C** depict specificity of MG71-2 and MG71-43 showing MG71 family displaying high specificity in mammalian cells. Cells were assessed for double strand break (DSB) formation by co-nucleofection mRNA, gRNA, and annealed dsODN. The dsODN could be incorporated into DSBs and used as a priming site for NGS library prep, allowing for unbiased sampling of all DSBs created in the cell. The percent of dsODN reads coming from the desired target site was mapped for up to three replicates for nuclease MG71-2 (**FIG. 23A**). For all 8 guides, all the reads passing analysis criteria come from the on-target site. The percent of dsODN reads coming from the desired target site was mapped for up to 3 replicates for nuclease MG71-43. For 5/8 guides, all the reads passing analysis criteria come from the on-target site (**FIG. 23B**). The identified off-target sites for guide g6-MG71-2_AAVS1_H2 are shown relative to the on-target sequence (top, NNNRCY represents the PAM) (**FIG. 23C**). Dots denote matching position. The boxed bases represent changes from the on target. Two off targets are seen across the replicates. **FIG. 23C** discloses SEQ ID NOs: 2160-2163, 2160-2163 and 2160-2163, respectively, in order of appearance.

[0098] **FIGs. 24A-24B** depict specificity of MG87-70 displaying high specificity in mammalian cells. Cells were assessed for double strand break (DSB) formation by co-nucleofection mRNA, gRNA, and annealed dsODN. The dsODN could be incorporated into DSBs and used as a priming site for NGS library prep, allowing for unbiased sampling of all DSBs created in the cell. The percent of dsODN reads coming from the desired target site was mapped for up to 3 replicates for nuclease MG87-70. For 4/7 guides, all the reads passing analysis criteria come from the on-target site (**FIG. 24A**). The identified off-target sites for guide 87-70_sgRNA_P1B3 were shown relative to the on-target sequence (top, NNNNGT represents the PAM). Dots denote matching position (**FIG. 24B**). The boxed bases represent changes from the on target. Two off targets are seen across the replicates. **FIG. 24B** discloses SEQ ID NOs: 2164-2167, 2164-2167 and 2164-2167, respectively, in order of appearance.

[0099] **FIG. 25** depicts pooled screening of MG71 chimeras at PAH p.R408W in engineered immortalized K562 cells. Each nuclease was delivered as codon-optimized mRNA along with 10 guide RNAs that target the SNV of interest. Guides were offset from each other by 1 nt. Each guide's cut site lied within 10 nt of the disease mutation. Indel efficiency was evaluated by NGS.

[0100] **FIG. 26** depicts human HAO1 MG21-1 engineered guide RNA activity.

[0101] **FIG. 27** depicts human AAVS1 MG23-1 engineered guide RNA activity.

[0102] **FIG. 28** depicts effect of spacer length on gRNA activity (% indel). Analysis of gene-editing outcomes at the DNA level for guides of different spacer lengths on 1-2 different chemically modified backbones from three different Type II editing systems (MG23-1, MG21-1, and MG71-2). hAAVS1-23-1-G1-1 (SEQ ID NOs: 1486-1493), hTRAC-23-1-H7-1 (SEQ ID NOs: 1494-1500), hHAO1-21-1-F1-78 (SEQ ID NOs: 1502-1505), hHAO1-21-1-F1-157 (SEQ ID NOs: 1506-1509), hTRAC-71-2-C2-12 (SEQ ID NOs: 1510-1517), hTRAC-71-2-C2-255 (SEQ ID NOs: 1518-1525).

[0103] **FIG. 29** depicts guide RNA targeting the human *Atxn2* gene with MG21-1 in K562 cells. **FIG. 29** further shows analysis of *Atxn2* guides with MG21-1 mRNA in K562 cells. Nucleofection of MG21-1 along with the matching *Atxn2* guide RNA (500 ng mRNA/150 pmol guide) was performed into K562 (100,000 cells). Cells were harvested and genomic DNA prepared three 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. Graphs illustrate guides tested, targeting all exons of *Atxn2* region of interest.

[0104] **FIG. 30** depicts validation of a subset of MG21-1 synthetic guides on the *Atxn2* gene in human neural progenitor cells. **FIG. 30** further shows analysis of *Atxn2* guides with MG21-1 mRNA in human neural progenitor cells. NPC cells were co-transfected with 200 ng of mRNA and chemically-synthesized sgRNA targeting *Atxn2*. Cells were harvested after three days of culture, and genomic DNA (gDNA) was extracted. The target genomic regions were amplified from the extracted gDNA and the amplicons were sequenced on an Illumina MiSeq platform. Sequences were analyzed using a proprietary Python script to quantify gene editing efficiency. % OOF denotes out of frame mutations.

[0105] **FIG. 31** depicts mouse versions of active human synthetic guides showing *Atxn2* gene editing activity in mouse neuro-2A cells. The neuro-2a cells were transfected using a serial dilution of MG21-1 mRNA and mouse surrogate *Atxn2* guides via nucleofection. The highest concentration in the serial dilution was 500 ng of mRNA and 150 pmol of guide, followed by a 1:3 serial dilution series. Cells were harvested after three days of culture, and genomic DNA

(gDNA) was extracted. The target genomic regions were amplified from the extracted gDNA with primers designed for NGS-based sequencing. The amplicons were sequenced on an Illumina MiSeq platform and analyzed using a proprietary Python script to quantify gene editing efficiency.

[0106] FIG. 32 depicts design of AAV constructs bearing *Atxn2* mouse surrogate guide RNAs. The AAV cargo constructs were designed to express MG21-1 and a guide RNA from a single AAV. The payload was flanked by two, 141 bp inverted terminal repeats (ITRs). A U6 polymerase III promoter was placed upstream of the guide RNA sequence in the forward orientation at the 5' end of the cargo. This was followed by a CMV promoter driving the MG21-1 open reading frame. MG21-1 was N-terminally tagged with a V5 epitope. A synthetic polyadenylation sequence (SpA) follows MG21-1.

[0107] FIG. 33 depicts AAV constructs bearing MG21-1 and guides targeting mouse *Atxn2* show gene editing activity in mouse neuro-2A cells. The neuro-2a cells were transfected using a serial dilution of AAV cargo plasmids via lipofection. The highest plasmid concentration in the dilution series was 2 ug, followed by a 1:2 serial dilution. Cells were harvested after three days of culture, and genomic DNA (gDNA) was extracted. The target genomic regions were amplified from the extracted gDNA with primers designed for NGS-based sequencing. PCR products were purified and the amplicons were sequenced. Sequences were analyzed to quantify gene editing efficiency.

[0108] FIG. 34 depicts AAV constructs bearing MG21-1 and guides targeting mouse *Atxn2* show gene editing activity in mouse Neuro-2A cells. The neuro-2a cells were transfected using a serial dilution of AAV cargo plasmids via lipofection. The highest plasmid concentration in the dilution series was 500 ng, followed by a 1:2 serial dilution. Cells were harvested after three days of culture, and genomic DNA (gDNA) was extracted. The target genomic regions were amplified from the extracted gDNA with primers designed for NGS-based sequencing. PCR products were purified and the amplicons were sequenced. Sequences were analyzed using a proprietary Python script to quantify gene editing efficiency.

[0109] FIG. 35 depicts AAV constructs bearing MG21-1 and guides targeting mouse *Atxn2* showing protein knockdown in mouse neuro-2A cells. The neuro-2a cells were transfected with each of five AAV plasmids (1 µg of DNA) containing different mouse surrogate *Atxn2* guides using Lipofectamine 2000. Three days after transfection, cells were lysed, and 10 µg was run on a 4-20% Tris-Glycine gel. Proteins were transferred to a nitrocellulose membrane, blocked with 5% milk, then probed with the antibodies against the V5 epitope and *Atxn2*. Blots were reprobed with an actin antibody as a reference control.

[0110] **FIG. 36** depicts AAV constructs bearing MG21-1 and guides targeting mouse *Atxn2* show full-length genome packaging into AAV9 by alkaline gel electrophoresis. The AAVs were produced using a transient triple co-transfection method in HEK293 cells involving a plasmid encoding adenoviral helper proteins, a plasmid encoding AAV replication protein and AAV9 capsid protein, and a cargo plasmid encoding guide and MG21-1 flanked between two AAV2 ITRs. After transfection, the cells were incubated at 37 °C for 72 hrs prior to harvest. A crude viral lysate was created, which was then treated prior to purification via cesium chloride density gradient ultracentrifugation. The AAVs were then subjected to 3 rounds of dialysis prior to loading on an alkaline agarose gel to visualize the size and integrity of the vector genomes.

[0111] **FIG. 37** depicts in vivo study design, experimental procedure, and sample preparation. 6–8-week-old C57BL/6J mice were injected with AAV9 into both the olfactory bulb and M1 region of the motor cortex using a custom built stereotactic delivery system based on the RWD automated stereotaxic instrument. M1 injections consisted of a bilateral infusion of 750 nL of AAV9 infused at a rate of 5 nL/second at coordinates +1.00 AP, +/-1.50 ML, -1.50 DV relative to Bregma. Olfactory bulb injections consisted of a bilateral infusion of 250 nL of AAV9 infused at a rate of 5 nL/second and three target depths, coordinates +4.00 AP, +/- 0.70 ML, -2.5/-2.0/-1.5 DV relative to Bregma. A retention time of 5 minutes post injection was added to enhance absorption of the AAV into the tissue. Animals were sutured and allowed to recover until transferred back to their home cages. Two weeks post injection, animals were euthanized via deep inhalation of isoflurane. Whole brains were dissected fresh, the olfactory bulb and the M1 regions were further dissected and processed for downstream nucleic acid extraction. Genomic DNA and mRNA was extracted from fresh tissue using DNA and RNA isolation kits.

[0112] **FIG. 38** depicts tissue titer measurements demonstrate MG21-1 AAVs effectively transduce mouse olfactory bulb and motor cortex. 6–8-week-old C57BL/6J mice were injected with AAV9 prepared in-house into both the olfactory bulb and M1 region of the motor cortex using a custom built stereotactic delivery system. Two weeks post injection, animals were euthanized via deep inhalation of isoflurane. Whole brains were dissected fresh, the olfactory bulb and the M1 regions were further dissected and processed for downstream nucleic acid extraction. To determine the tissue titer of injected AAVs, 5 ng of gDNA was added to a ddPCR reaction containing custom probes targeting the central region of the MG21-1 sequence multiplexed with a premade assay targeting the *CyC1* gene. MG21-1 copy numbers were normalized by *CyC1* reference copies to produce a normalized MG21-1 copies/cell value. All individual data points are shown with the error bars denoting the average +/- standard error of the mean.

[0113] **FIG. 39** depicts MG21-1 transcripts are detected in the olfactory bulb and motor cortex after AAV9 injection. 6–8-week-old C57BL/6J mice were injected with AAV9 prepared in-house into both the olfactory bulb and M1 region of the motor cortex using a custom built stereotactic delivery system. Two weeks post injection, animals were euthanized via deep inhalation of isoflurane. Whole brains were dissected fresh, the olfactory bulb and the M1 regions were further dissected and processed for downstream nucleic acid extraction. mRNA was extracted from fresh tissue and then used to produce cDNA by reverse transcription. To quantify the expression of MG21-1, 5ng of cDNA mRNA equivalent was loaded into a ddPCR reaction containing custom primers and probes targeting the central region of the MG21-1 sequence. This assay was additionally multiplexed with an assay targeting the *Cyc1* gene. Data were graphed in GraphPad Prism 10. MG21-1 copies were normalized by *Cyc1* reference copies to produce a normalized MG21-1 copies/*Cyc1* copies value. All individual data points are shown with the error bars denoting the average +/- standard error of the mean.

[0114] **FIG. 40** depicts MG21-1 packaged into AAV9 effectively knocks down *Atxn2* mRNA in the olfactory bulb and motor cortex. 6–8-week-old C57BL/6J mice were injected with AAV9 prepared in-house into both the olfactory bulb and M1 region of the motor cortex using a custom built stereotactic delivery system. Two weeks post injection, animals were euthanized via deep inhalation of isoflurane. Whole brains were dissected fresh, the olfactory bulb and the M1 regions were further dissected and processed for downstream nucleic acid extraction. mRNA was extracted from fresh tissue and then used to produce cDNA by reverse transcription. To quantify the expression of *Atxn2*, we loaded 5 ng of cDNA mRNA equivalent to a ddPCR reaction containing an *Atxn2* premade expression quantification assay. This assay was additionally multiplexed with a premade assay targeting the *Cyc1* gene. Data were graphed in GraphPad Prism 10. *Atxn2* mRNA copies were normalized to *Cyc1* reference copies to produce a normalized *Atxn2* copies/*Cyc1* copies value. All individual data points are shown with the error bars denoting the average +/- standard error of the mean.

[0115] **FIG. 41** depicts a graph of indels within the mouse *Atxn2* gene are detectable within mouse olfactory bulb and motor cortex. All individual data points are shown with the error bars denoting the average +/- standard error of the mean.

[0116] **FIG. 42** depicts a graph of the quantification of *Atxn2* protein expression in olfactory bulb samples injected with AAV9. All individual data points are shown with the error bars denoting the average +/- standard error of the mean.

[0117] **FIG. 43** depicts data from stereotactic injection of an AAV encoding MG21-1 targeting *Atxn2* showing target protein knockdown by imaging. Scale bars represent 70 μm .

[0118] **FIG. 44** depicts data from stereotactic injection of an AAV encoding MG21-1_P1F12 guide shows *Atn2* protein knockdown by imaging. Scale bars represent 70 μ m.

[0119] **FIG. 45** depicts a schematic overview of the workflow. Single mismatches across every position of spacer and PAM were generated (1MM Target) and used to create the dual-target library members. The library was ordered and cloned into a lentiviral plasmid to create the plasmid library from which the lentiviral library was produced. K562 cells were infected with the dual-target lentibrary and following antibiotic selection the cells were edited via nucleofection and the gDNA harvested. After NGS sequencing the data was analyzed.

[0120] **FIGs. 46A** and **46B** depict bar graphs showing GC content and Percent Indels for library sgRNAs. The range of GC contents (**FIG. 46A**) and percent indels (**FIG. 46B**) for the guides from which the library was designed are plotted by nuclease.

[0121] **FIG. 47** depicts a schematic overview of the reference structure of dual-target oligo library member.

[0122] **FIG. 48** depicts graphs showing editing controls. Control oligonucleotides were generated to simulate four possible editing events: cleavage at the on-target, cleavage at the off-target, cleavage at both targets, and no cleavage. These control dual-target sequences possessed identical on- and off-target protospacers, with cleavage events regulated by their active or inactive PAM sequences. To generate inactive PAMs, bases were substituted in the active PAM consensus sequence with alternatives that are not recognized by the Cas system. Editing at the left or right end of the dual-target controls, following nucleofection with guides across MG71-2 and MG21-1, was plotted.

[0123] **FIG. 49** depicts the generation of a nuclease-specific single mismatch tolerance profile. Raw indel percentages at mismatched targets and their corresponding on-target were calculated for each library member. The raw indel percentages were normalized across each position by taking the ratio of observed off:on target editing, termed mismatch tolerance, for each guide. A nuclease-specific single mismatch tolerance profile was generated by plotting the mean and 95% confidence intervals of these single mismatch tolerances, by position, of all guides for a given nuclease.

[0124] **FIG. 50** depicts a graph showing single mismatch tolerance of 22 nucleotide spacers for MG71-2. 25 guides targeting human B2M, TRAC, AAVS1 loci were used to design dual-target library members in which the off-target sequence differs by 1 mismatch between the off and on-target sequence. The raw indel percentages at mismatched targets compared to on-target library members were normalized by taking the ratio of off:on target editing, termed mismatch tolerance, observed by position for each guide and a nuclease specific profile was generated by plotting the mean and 95% confidence intervals of single mismatch tolerance by position.

[0125] **FIG. 51** depicts a graph showing single mismatch tolerance of 24 nucleotide spacers for MG71-2. 7 guides targeting human HBB locus were used to design dual-target library members in which the off-target sequence differs by 1 mismatch between the off and on-target sequence. The raw indel percentages at mismatched targets compared to on-target library members were normalized by taking the ratio of off:on target editing, termed mismatch tolerance, observed by position for each guide and a nuclease specific profile was generated by plotting the mean and 95% confidence intervals of single mismatch tolerance by position.

[0126] **FIG. 52** depicts a graph showing single mismatch tolerance of MG71-2 with 22 and 24 nucleotide spacers overlaid. 25 guides of 22nt length targeting human B2M, TRAC, AAVS1 loci, and 7 guides of 24nt length targeting the human HBB locus were used to design dual-target library members in which the off-target sequence differs by 1 mismatch between the off and on-target sequence. The raw indel percentages at mismatched targets compared to on-target library members were normalized by taking the ratio of off:on target editing, termed mismatch tolerance, observed by position for each guide and a nuclease specific profile was generated by plotting the mean and 95% confidence intervals of single mismatch tolerance by position. The two plots were overlaid to visualize.

[0127] **FIG. 53** depicts a graph showing single mismatch tolerance of MG21-1. 21 guides targeting human PDCD1, HAO1, IL17, and TRAC loci were used to design dual-target library members in which the off-target sequence differs by 1 mismatch between the off and on-target sequence. The raw indel percentages at mismatched targets compared to on-target library members were normalized by taking the ratio of off:on target editing, termed mismatch tolerance, observed by position for each guide and a nuclease specific profile was generated by plotting the mean and 95% confidence intervals of single mismatch tolerance by position.

[0128] **FIG. 54** depicts a graph showing MG71-2 PAM preference. PAM preference scores were calculated for MG71-2 by converting the mismatch tolerance scores and their respective off-target PAMs into information-based representation matrices using logomaker.

[0129] **FIG. 55** depicts a graph showing MG21-1 PAM preference. PAM preference scores were calculated for MG21-1 by converting the mismatch tolerance scores and their respective off-target PAMs into information-based representation matrices using logomaker.

[0130] **FIG. 56** depicts an MG71-2 protospacer base preference heatmap. Heatmaps were created to illustrate the protospacer base preference at each mismatch position for different nucleases. The heatmaps were generated by grouping the data points by nuclease, mismatch location, and mismatched protospacer base, and then calculating the mean off-on ratio and 95% confidence interval for each group.

[0131] FIG. 57 depicts an MG21-1 protospacer base preference heatmap. Heatmaps were created to illustrate the protospacer base preference at each mismatch position for different nucleases. The heatmaps were generated by grouping the data points by nuclease, mismatch location, and mismatched protospacer base, and then calculating the mean off-on ratio and 95% confidence interval for each group.

BRIEF DESCRIPTION OF THE SEQUENCE LISTING

[0132] 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.

[0133] MG21

[0134] SEQ ID NO: 47 shows the full-length peptide sequence of an MG21 nuclease.

[0135] SEQ ID NOs: 1137-1144 show the nucleotide sequences of sgRNAs engineered to function with an MG21 nuclease.

[0136] MG23

[0137] SEQ ID NO: 48 shows the full-length peptide sequence of an MG23 nuclease.

[0138] SEQ ID NOs: 1153-1156 show the nucleotide sequences of sgRNAs engineered to function with an MG23 nuclease.

[0139] MG71

[0140] SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263 show the full-length peptide sequences of MG71 nucleases.

[0141] SEQ ID NOs: 53-54 show the peptide sequences of PAM-interacting domains of MG71 nucleases.

[0142] nnRMYnn, nnnACTnn nNNRMT, nRNaCT, nNNaCT, nRNRCT, nNNRH_Y, nNNRMT, nNNRC_Y, nNNR_YY, nNNRMT, nNNRC_Y, nNWRAT, nNNRMT, nNNRH_T, nNNRH_Y, nNNRH_Y, nNNR_YY, nNRNCN, nNNR_YY, nNRN_T, nNNAWT, nNRAMW, nNARMY, nNNRCT, nNNRY_T, nNAR_YY, nNNRC_Y, and nNNR_YY are PAM sequences compatible with MG71 nucleases.

[0143] SEQ ID NOs: 710-722 show the nucleotide sequences of MG71 tracrRNAs.

[0144] SEQ ID NOs: 779-791 show the nucleotide sequences of MG71 CRISPR repeats.

[0145] SEQ ID NOs: 927-942, 945-950, 961-962, 969-970, 1123-1137, and 2433-2434 show the nucleotide sequences of sgRNAs engineered to function with an MG71 nuclease. In SEQ ID NOs: 2433-2434, “n” can be any nucleotide.

[0146] MG73

[0147] SEQ ID NOs: 51 and 1264 show the full-length peptide sequence of a MG73 nuclease.

- [0148] SEQ ID NO: 55 shows the peptide sequence of a PAM-interacting domain of a MG73 nuclease.
- [0149] nnRnTTnn is a PAM sequence compatible with a MG73 nuclease.
- [0150] **MG89**
- [0151] SEQ ID NO: 52 shows the full-length peptide sequence of a MG89 nuclease.
- [0152] SEQ ID NO: 56 shows the peptide sequence of a PAM-interacting domain of a MG89 nuclease.
- [0153] nnnnCC is a PAM sequence compatible with a MG89 nuclease.
- [0154] **MG87**
- [0155] SEQ ID NOs: 659-660, 1158-1159, and 1267-1277 show the full-length peptide sequences of MG87 nucleases.
- [0156] SEQ ID NOs: 726-744 show the nucleotide sequences of MG87 tracrRNAs.
- [0157] SEQ ID NOs: 795-813 show the nucleotide sequences of MG87 CRISPR repeats.
- [0158] SEQ ID NOs: 843-880 and 1145-1152 show the nucleotide sequences of sgRNAs engineered to function with an MG87 nuclease.
- [0159] **MG88**
- [0160] SEQ ID NOs: 661-678 and 1278-1282 show the full-length peptide sequences of MG88 nucleases.
- [0161] SEQ ID NOs: 745-767 show the nucleotide sequences of MG88 tracrRNAs.
- [0162] SEQ ID NOs: 814-836 show the nucleotide sequences of MG88 CRISPR repeats.
- [0163] SEQ ID NOs: 881-926 show the nucleotide sequences of sgRNAs engineered to function with an MG88 nuclease.
- [0164] **MG17**
- [0165] SEQ ID NO: 695 shows the full-length peptide sequence of a MG17 nuclease.
- [0166] SEQ ID NO: 699 shows the nucleotide sequence of a MG17 tracrRNA.
- [0167] SEQ ID NO: 768 shows the nucleotide sequence of a MG17 CRISPR repeat.
- [0168] SEQ ID NOs: 973-974 show the nucleotide sequences of sgRNAs engineered to function with an MG17 nuclease.
- [0169] **MG18**
- [0170] SEQ ID NOs: 696-698 show the full-length peptide sequences of MG18 nucleases.
- [0171] SEQ ID NOs: 700-702 show the nucleotide sequences of MG18 tracrRNAs.
- [0172] SEQ ID NOs: 769-771 show the nucleotide sequences of MG18 CRISPR repeats.
- [0173] SEQ ID NOs: 957-960 show the nucleotide sequences of sgRNAs engineered to function with an MG18 nuclease.
- [0174] **MG46**

- [0175] SEQ ID NOs: 975-981 show the full-length peptide sequences of MG46 nucleases.
- [0176] SEQ ID NOs: 703-709 show the nucleotide sequences of MG46 tracrRNAs.
- [0177] SEQ ID NOs: 772-778 show the nucleotide sequences of MG46 CRISPR repeats.
- [0178] SEQ ID NOs: 943-944, 951-956, 965-968, and 971-972 show the nucleotide sequences of sgRNAs engineered to function with an MG46 nuclease.
- [0179] **MG74**
- [0180] SEQ ID NOs: 1265-1266 show the full-length peptide sequences of MG74 nucleases.
- [0181] **HAO1 Targeting with MG21-1**
- [0182] SEQ ID NOs: 1283-1392 and 1502-1509 show the nucleotide sequences of sgRNAs engineered to function with an MG21-1 nuclease in order to target the human HAO1 gene.
- [0183] **MG71-2 AAVS1 Targeting**
- [0184] SEQ ID NOs: 57-58, 101-119, 163-181, 225-243, 343-374, 421-433, 461-472, 499-511, 539-551, and 1098-1102 show the nucleotide sequences of sgRNAs engineered to function with an MG71-2 nuclease in order to target the AAVS1 gene.
- [0185] SEQ ID NOs: 59-60, 120-138, 182-200, 244-262, 375-406, 434-446, 473-484, 512-524, 552-564, and 1248-1252 show the DNA sequences of AAVS1 target sites.
- [0186] **MG71-2 TRAC Targeting**
- [0187] SEQ ID NOs: 77-88, 139-150, 201-212, 407-413, 447-453, 485-491, 525-531, and 1096-1097 show the nucleotide sequences of sgRNAs engineered to function with an MG71-2 nuclease in order to target TRAC.
- [0188] SEQ ID NOs: 89-100, 151-162, 213-224, 414-420, 454-460, 492-498, 532-538, and 1246-1247 show the DNA sequences of TRAC target sites.
- [0189] **MG71-2 B2M Targeting**
- [0190] SEQ ID NOs: 565-578 and 1095 show the nucleotide sequences of sgRNAs engineered to function with an MG71-2 nuclease in order to target the human B2M gene.
- [0191] SEQ ID NOs: 579-592 and 1245 show the DNA sequences of human B2M target sites.
- [0192] **MG71-2 HBB Targeting**
- [0193] SEQ ID NOs: 593-625 show the nucleotide sequences of sgRNAs engineered to function with an MG71-2 nuclease in order to target the human HBB gene.
- [0194] SEQ ID NOs: 626-658 show the DNA sequences of human HBB target sites.
- [0195] **MG71-2 Albumin Targeting**
- [0196] SEQ ID NOs: 1083-1094 show the nucleotide sequences of sgRNAs engineered to function with an MG71-2 nuclease in order to target the human Albumin gene.
- [0197] SEQ ID NOs: 1233-1244 show the DNA sequences of human Albumin target sites.
- [0198] **MG71-2 ATP7B Targeting**

- [0199] SEQ ID NOs: 1056-1081 show the nucleotide sequences of sgRNAs engineered to function with an MG71-2 nuclease in order to target human ATP7B.
- [0200] SEQ ID NOs: 1206-1231 show the DNA sequences of human ATP7B target sites.
- [0201] **MG71-2 PAH Targeting**
- [0202] SEQ ID NOs: 1113-1122 show the nucleotide sequences of sgRNAs engineered to function with an MG71-2 nuclease in order to target the human PAH gene.
- [0203] SEQ ID NOs: 1253-1262 show the DNA sequences of human PAH target sites.
- [0204] **MG73-1 TRAC Targeting**
- [0205] SEQ ID NO: 61 shows the nucleotide sequence of an sgRNA engineered to function with an MG73-1 nuclease in order to target TRAC.
- [0206] SEQ ID NO: 62 shows the DNA sequence of a TRAC target site.
- [0207] **MG89-2 TRAC Targeting**
- [0208] SEQ ID NOs: 63-69 and 263-302 show the nucleotide sequences of sgRNAs engineered to function with an MG89-2 nuclease in order to target TRAC.
- [0209] SEQ ID NOs: 70-76 and 303-342 show the DNA sequences of TRAC target sites.
- [0210] **MG87 TRAC Targeting**
- [0211] SEQ ID NOs: 1011-1019, and 1050 show the nucleotide sequences of sgRNAs engineered to function with an MG87 nuclease in order to target TRAC.
- [0212] SEQ ID NOs: 1161-1168 and 1200 show the DNA sequences of TRAC target sites.
- [0213] **MG87 AAVS1 Targeting**
- [0214] SEQ ID NOs: 1020-1049 and 1051-1055 show the nucleotide sequences of sgRNAs engineered to function with an MG87 nuclease in order to target the AAVS1 gene.
- [0215] SEQ ID NOs: 1169-1205 show the DNA sequences of AAVS1 target sites.
- [0216] **MG21-1 ATXN2 Targeting**
- [0217] SEQ ID NOs: 1786-2045 show the nucleotide sequences of sgRNAs engineered to function with an MG21-1 nuclease in order to target the human ATXN2 gene.
- [0218] SEQ ID NOs: 2051-2055 show the nucleotide sequences of sgRNAs engineered to function with an MG21-1 nuclease in order to target the mouse ATXN2 gene.
- [0219] SEQ ID NOs: 1526-1785 show the DNA sequences of human ATXN2 target sites.
- [0220] SEQ ID NOs: 2046-2050 show the DNA sequences of mouse ATXN2 target sites.
- [0221] **MG23-1 AAVS1 Targeting**
- [0222] SEQ ID NOs: 1393-1493 show the nucleotide sequences of sgRNAs engineered to function with an MG23-1 nuclease in order to target the AAVS1 gene.
- [0223] **MG23-1 TRAC Targeting**

[0224] SEQ ID NOs: 1494-1501 and 1510-1525 show the nucleotide sequences of sgRNAs engineered to function with an MG23-1 nuclease in order to target TRAC.

[0225] **AAV Cargo Plasmid**

[0226] SEQ ID NOs: 2056-2060 show the nucleic acid sequences of AAV cargo plasmids.

DETAILED DESCRIPTION

[0227] While various embodiments of the disclosure 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 disclosure. It should be understood that various alternatives to the embodiments of the disclosure described herein may be employed.

[0228] Whenever the term “at least,” “greater than,” or “greater than or equal to” precedes the first numerical value in a series of two or more numerical values, the term “at least,” “greater than” or “greater than or equal to” applies to each of the numerical values in that series of numerical values. For example, greater than or equal to 1, 2, or 3 is equivalent to greater than or equal to 1, greater than or equal to 2, or greater than or equal to 3.

[0229] Whenever the term “no more than,” “less than,” or “less than or equal to” precedes the first numerical value in a series of two or more numerical values, the term “no more than,” “less than,” or “less than or equal to” applies to each of the numerical values in that series of numerical values. For example, less than or equal to 3, 2, or 1 is equivalent to less than or equal to 3, less than or equal to 2, or less than or equal to 1.

[0230] 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.

[0231] 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.”

[0232] 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.

[0233] The term “nucleotide,” as used herein, refers to a base-sugar-phosphate combination. Contemplated nucleotides include naturally occurring nucleotides and synthetic nucleotides. Nucleotides are monomeric units of a nucleic acid sequence (*e.g.*, deoxyribonucleic acid (DNA) and ribonucleic acid (RNA)). The term nucleotide includes 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 include, for example, [α S]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 encompasses dideoxyribonucleoside triphosphates (ddNTPs) and their derivatives. Illustrative examples of ddNTPs 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) or quantum dots. Detectable labels include, for example, radioactive isotopes, fluorescent labels, chemiluminescent labels, bioluminescent labels, and enzyme labels. Fluorescent labels of nucleotides 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 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-rhodamine-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. The term nucleotide encompasses chemically modified nucleotides. An exemplary chemically-modified nucleotide is biotin-dNTP. Non-

limiting examples of biotinylated dNTPs 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).

[0234] The terms “polynucleotide,” “oligonucleotide,” and “nucleic acid” are used interchangeably to 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. Contemplated polynucleotides include a gene or fragment thereof. Exemplary polynucleotides include, but are not limited to, DNA, RNA, 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, cell-free polynucleotides including cell-free DNA (cfDNA) and cell-free RNA (cfRNA), nucleic acid probes, and primers. In a polynucleotide when referring to a T, a T means U (Uracil) in RNA and T (Thymine) in DNA. A polynucleotide can be exogenous or endogenous to a cell and/or exist in a cell-free environment. The term polynucleotide encompasses modified polynucleotides (*e.g.*, altered backbone, sugar, or nucleobase). If present, modifications to the nucleotide structure are imparted before or after assembly of the polymer. Non-limiting examples of modifications 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. The sequence of nucleotides may be interrupted by non-nucleotide components.

[0235] The terms “peptide,” “polypeptide,” and “protein” are used interchangeably herein to 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 is interrupted by non-amino acids. The terms include amino acid chains of any length, including full length proteins, and proteins with or without secondary 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, refer to natural and non-natural amino acids, including, but not limited to, modified amino acids. Modified amino acids include amino acids that have been chemically modified to include a group or a chemical moiety not naturally present on the amino acid. The term “amino acid” includes both D-amino acids and L-amino acids.

[0236] As used herein, “operably linked”, “operable linkage”, “operatively linked”, or grammatical equivalents thereof refer to an arrangement of genetic elements, *e.g.*, a promoter, an enhancer, a polyadenylation sequence, etc., wherein an operation (*e.g.*, movement or activation) of a first genetic element has some effect on the second genetic element. The effect on the second genetic element can be, but need not be, of the same type as operation of the first genetic element. For example, two genetic elements are operably linked if movement of the first element causes an activation of the second element. 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.

[0237] A “functional fragment” of a DNA or protein sequence 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 includes its ability to influence expression in a manner attributed to the full-length sequence.

[0238] The terms “engineered,” “synthetic,” and “artificial” are used interchangeably herein to refer to an object that has been modified by human intervention. For example, the terms refer to a polynucleotide or polypeptide that is non-naturally occurring. An engineered peptide has, but does not require, 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. Non-limiting examples include the following: a nucleic acid modified by changing its sequence to a sequence that does not occur in nature; a nucleic acid 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 synthesized *in vitro* with a sequence that does not exist in nature; a protein modified by changing its amino acid sequence to a sequence that does not exist in nature; an engineered protein acquiring a new function or property. An “engineered” system comprises at least one engineered component.

[0239] The term “tracrRNA” or “tracr sequence” means trans-activating CRISPR RNA. tracrRNA interacts with the CRISPR (cr) RNA to form a guide nucleic acid (*e.g.*, guide RNA or

gRNA) that may hybridize to a target nucleic acid and thereby directs an associated nuclease to the target nucleic acid.

[0240] As used herein, a “guide nucleic acid” or “guide polynucleotide” refers to a nucleic acid that may hybridize to a target nucleic acid and thereby directs an associated nuclease to the target nucleic acid. A guide nucleic acid is, but is not limited to, RNA (guide RNA or gRNA), DNA, or a mixture of RNA and DNA. A guide nucleic acid can include a crRNA or a tracrRNA or a combination of both. The term guide nucleic acid encompasses an engineered guide nucleic acid and a programmable guide nucleic acid to specifically bind to the target nucleic acid. 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 is the complementary strand. The strand of the double-stranded target polynucleotide that is complementary to the complementary strand, and therefore is not complementary to the guide nucleic acid is called noncomplementary strand. A guide nucleic acid having a polynucleotide chain is a “single guide nucleic acid.” A guide nucleic acid having two polynucleotide chains is a “double guide nucleic acid.” If not otherwise specified, the term “guide nucleic acid” is inclusive, referring to both single guide nucleic acids and double guide nucleic acids. A guide nucleic acid may comprise a segment referred to as a “nucleic acid-targeting segment” or a “nucleic acid-targeting sequence,” or a “spacer.” A nucleic acid-targeting segment can include a sub-segment referred to as a “protein binding segment” or “protein binding sequence” or “Cas protein binding segment.”

[0241] As used herein, the term “complex” refers to a joining of at least two components. The two components may each retain the properties/activities they had prior to forming the complex or gain properties as a result of forming the complex. The joining includes, but is not limited to, covalent bonding, non-covalent bonding (i.e., hydrogen bonding, ionic interactions, Van der Waals interactions, and hydrophobic bond), use of a linker, fusion, or any other suitable method. Contemplated components of the complex include polynucleotides, polypeptides, or combinations thereof. For example, a complex comprises an endonuclease and a guide polynucleotide.

[0242] The term “sequence identity” or “percent identity” in the context of two or more nucleic acids or polypeptide sequences, 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 parameters of the Smith-Waterman homology search algorithm with parameters of a match of 2, a mismatch of -1, and a gap of -1; MUSCLE with default parameters; MAFFT with parameters retree of 2 and maxiterations of 1000; Novafold with default parameters.

[0243] 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% identity to any one of the endonuclease protein sequences described herein (e.g. MG17, MG18, MG21, MG23, MG46, MG71, MG73, MG74, MG87, MG88, or MG89 family endonucleases 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 critical active site 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 residues predicted as essential. In some embodiments, a functional variant of any of the proteins described herein lacks substitution of all of the residues predicted as essential.

[0244] 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

[0245] 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 that represent large numbers of microbial species may offer the potential to drastically increase the number of new CRISPR/Cas systems known 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.

[0246] 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 (30-40 bp) 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.

[0247] Class 1 CRISPR-Cas systems have large, multisubunit effector complexes, and comprise Types I, III, and IV.

[0248] Type I CRISPR-Cas systems are considered of moderate complexity in terms of components. In Type I CRISPR-Cas systems, the array of RNA-targeting elements is transcribed as a long precursor crRNA (pre-crRNA) that is processed at repeat elements to liberate short, mature crRNAs that direct the nuclease complex to nucleic acid targets when they are followed by a suitable short consensus sequence called a protospacer-adjacent motif (PAM). This processing occurs via an endoribonuclease subunit (Cas6) of a large endonuclease complex called Cascade, which also comprises a nuclease (Cas3) that protein component of the crRNA-directed nuclease complex. Cas 1 nucleases function primarily as DNA nucleases.

[0249] Type III CRISPR systems may be characterized by the presence of a central nuclease, known as Cas10, alongside a repeat-associated mysterious protein (RAMP) that comprises Csm or Cmr protein subunits. Like in Type I systems, the mature crRNA is processed from a pre-crRNA using a Cas6-like enzyme. Unlike type I and II systems, type III systems appear to target and cleave DNA-RNA duplexes (such as DNA strands being used as templates for an RNA polymerase).

[0250] Type IV CRISPR-Cas systems possess an effector complex that consists of a highly reduced large subunit nuclease (csf1), two genes for RAMP proteins of the Cas5 (csf3) and Cas7 (csf2) groups, and, in some cases, a gene for a predicted small subunit; such systems are commonly found on endogenous plasmids.

[0251] Class II CRISPR-Cas systems generally have single-polypeptide multidomain nuclease effectors, and comprise Types II, V and VI.

[0252] 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 (Cas9) and the repeat sequence to form a precursor dsRNA structure, which is cleaved by endogenous RNase III to generate a mature Cas9 enzyme loaded with both tracrRNA and crRNA. Cas II nucleases are known as DNA nucleases. The Cas9 effector has a characteristic structure, consisting of 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.

[0253] Type V CRISPR-Cas systems are characterized by a nuclease effector (Cas12) structure similar to that of Type II/Cas9, 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 (Cas12) to cleave pre-crRNAs. Like Type-II CRISPR-Cas systems, Type V CRISPR-Cas systems are again known 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.

[0254] Type VI CRISPR-Cas systems are unique in that they appear to be the only class so far known as RNA-guided RNA endonucleases. Instead of RuvC-like domains, the single polypeptide effector of Type VI systems (Cas13) comprises two HEPN ribonuclease domains. Differing from both Type II and V systems, Type VI systems also appear to not need a tracrRNA for processing of pre-crRNA into crRNA. Similar to type V systems, however, some Type VI systems (e.g., C2C2) appear to possess robust single-stranded nonspecific nuclease (ribonuclease) activity activated by the first crRNA directed cleavage of a target RNA.

[0255] Because of their simpler architecture, Class 2 CRISPR-Cas have been most widely adopted for engineering and development as designer nuclease/genome editing applications.

[0256] One of the early adaptations of such a system for *in vitro* use involved (i) recombinantly-expressed, purified full-length Cas9 (e.g., a Class 2, Type II Cas enzyme) isolated from *S. pyogenes* SF370, (ii) purified mature ~42 nt crRNA bearing a ~20 nt 5' sequence complementary to the target DNA sequence desired to be cleaved followed by a 3' tracr-binding sequence (the whole crRNA being *in vitro* transcribed from a synthetic DNA template carrying a T7 promoter sequence); (iii) purified tracrRNA *in vitro* transcribed from a synthetic DNA template carrying a T7 promoter sequence, and (iv) Mg²⁺. A later improved, engineered system involved the crRNA of (ii) joined to the 5' end of (iii) by a linker (e.g., GAAA) to form a single fused synthetic guide RNA (sgRNA) capable of directing Cas9 to a target by itself.

[0257] Such engineered systems can be adapted for use in mammalian cells by providing DNA vectors encoding (i) an ORF encoding codon-optimized Cas9 (e.g., a Class 2, Type II Cas enzyme) under a suitable mammalian promoter with a C-terminal nuclear localization sequence (e.g., SV40 NLS) and a suitable polyadenylation signal (e.g., TK pA signal); and (ii) an ORF encoding an sgRNA (having a 5' sequence beginning with G followed by 20 nt of a

complementary targeting nucleic acid sequence joined to a 3' tracr-binding sequence, a linker, and the tracrRNA sequence) under a suitable Polymerase III promoter (e.g., the U6 promoter).

MG Enzymes

[0258] Described herein, in certain embodiments, are engineered nuclease systems comprising an endonuclease and an engineered guide polynucleotide. In some embodiments, the endonuclease is a Class 2, Type II endonuclease. In some embodiments, the endonuclease comprises a RuvC_III domain. In some embodiments, the endonuclease comprises an HNH domain. In some embodiments, the endonuclease is a double strand nuclease. In some embodiments, the endonuclease is catalytically dead. In some embodiments, the endonuclease is a double strand nuclease. In some embodiments, the endonuclease is modified. In some embodiments, the endonuclease is modified resulting in an endonuclease with nickase activity. In some embodiments, the modified endonuclease is a site-directed nickase

[0259] In some embodiments, the endonuclease is a MG21 endonuclease. In some embodiments, the endonuclease is a MG23 endonuclease. In some embodiments, the endonuclease is a MG71 endonuclease. In some embodiments, the endonuclease is a MG73 endonuclease. In some embodiments, the endonuclease is a MG74 endonuclease. In some embodiments, the endonuclease is a MG89 endonuclease. In some embodiments, the endonuclease is a MG87 endonuclease. In some embodiments, the endonuclease is a MG88 endonuclease. In some embodiments, the endonuclease is a MG17 endonuclease. In some embodiments, the endonuclease is a MG18 endonuclease. In some embodiments, the endonuclease is a MG46 endonuclease.

[0260] In some embodiments, the engineered nuclease system is discovered through metagenomic sequencing. In some embodiments, the metagenomic sequencing is conducted on samples collected from various environments. In some embodiments, the environment is a human microbiome, an animal microbiome, an environment with high temperatures, an environment with low temperatures, or sediment.

[0261] In some embodiments, the endonuclease is a MG21 endonuclease (i.e., SEQ ID NO: 47). In some embodiments, the endonuclease comprises a sequence having 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 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 47. In some embodiments, the endonuclease comprises a sequence having at least about 70% identity to SEQ ID NO: 47. In some embodiments, the endonuclease comprises a sequence having at least about 75% identity to SEQ ID NO: 47. In some embodiments, the endonuclease comprises a sequence having at least

about 80% identity to SEQ ID NO: 47. In some embodiments, the endonuclease comprises a sequence having at least about 85% identity to SEQ ID NO: 47. In some embodiments, the endonuclease comprises a sequence having at least about 90% identity to SEQ ID NO: 47. In some embodiments, the endonuclease comprises a sequence having at least about 95% identity to SEQ ID NO: 47. In some embodiments, the endonuclease comprises a sequence having at least about 96% identity to SEQ ID NO: 47. In some embodiments, the endonuclease comprises a sequence having at least about 97% identity to SEQ ID NO: 47. In some embodiments, the endonuclease comprises a sequence having at least about 98% identity to SEQ ID NO: 47. In some embodiments, the endonuclease comprises a sequence having at least about 99% identity to SEQ ID NO: 47. In some embodiments, the endonuclease comprises a sequence having 100% identity to SEQ ID NO: 47.

[0262] In some embodiments, the endonuclease is a MG23 endonuclease (i.e., SEQ ID NO: 48). In some embodiments, the endonuclease comprises a sequence having 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 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 48. In some embodiments, the endonuclease comprises a sequence having at least about 70% identity to SEQ ID NO: 48. In some embodiments, the endonuclease comprises a sequence having at least about 75% identity to SEQ ID NO: 48. In some embodiments, the endonuclease comprises a sequence having at least about 80% identity to SEQ ID NO: 48. In some embodiments, the endonuclease comprises a sequence having at least about 85% identity to SEQ ID NO: 48. In some embodiments, the endonuclease comprises a sequence having at least about 90% identity to SEQ ID NO: 48. In some embodiments, the endonuclease comprises a sequence having at least about 95% identity to SEQ ID NO: 48. In some embodiments, the endonuclease comprises a sequence having at least about 96% identity to SEQ ID NO: 48. In some embodiments, the endonuclease comprises a sequence having at least about 97% identity to SEQ ID NO: 48. In some embodiments, the endonuclease comprises a sequence having at least about 98% identity to SEQ ID NO: 48. In some embodiments, the endonuclease comprises a sequence having at least about 99% identity to SEQ ID NO: 48. In some embodiments, the endonuclease comprises a sequence having 100% identity to SEQ ID NO: 48.

[0263] In some embodiments, the endonuclease is a MG71 endonuclease (i.e., SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263). In some embodiments, the endonuclease comprises a sequence having 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 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: 49-51, 679-694, 983-1002, 1003-1010, and 1263. In some embodiments, the endonuclease comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263. In some embodiments, the endonuclease comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263. In some embodiments, the endonuclease comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263. In some embodiments, the endonuclease comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263. In some embodiments, the endonuclease comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263. In some embodiments, the endonuclease comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263. In some embodiments, the endonuclease comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263. In some embodiments, the endonuclease comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263. In some embodiments, the endonuclease comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263. In some embodiments, the endonuclease comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263. In some embodiments, the endonuclease comprises a sequence having 100% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263.

[0264] In some embodiments, the endonuclease is a MG73 endonuclease (i.e., SEQ ID NO: 51 and SEQ ID NO: 1264). In some embodiments, the endonuclease comprises a sequence having 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 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 51 or SEQ ID NO: 1264. In some embodiments, the endonuclease comprises a sequence having at least about 70% identity to SEQ ID NO: 51 or SEQ ID NO: 1264. In some embodiments, the endonuclease comprises a sequence having at least about 75% identity to SEQ ID NO: 51 or SEQ ID NO: 1264. In some embodiments, the endonuclease comprises a sequence having at least about 80% identity to SEQ ID NO: 51 or SEQ ID NO: 1264. In some embodiments, the endonuclease

comprises a sequence having at least about 85% identity to SEQ ID NO: 51 or SEQ ID NO: 1264. In some embodiments, the endonuclease comprises a sequence having at least about 90% identity to SEQ ID NO: 51 or SEQ ID NO: 1264. In some embodiments, the endonuclease comprises a sequence having at least about 95% identity to SEQ ID NO: 51 or SEQ ID NO: 1264. In some embodiments, the endonuclease comprises a sequence having at least about 96% identity to SEQ ID NO: 51 or SEQ ID NO: 1264. In some embodiments, the endonuclease comprises a sequence having at least about 97% identity to SEQ ID NO: 51 or SEQ ID NO: 1264. In some embodiments, the endonuclease comprises a sequence having at least about 98% identity to SEQ ID NO: 51 or SEQ ID NO: 1264. In some embodiments, the endonuclease comprises a sequence having at least about 99% identity to SEQ ID NO: 51 or SEQ ID NO: 1264. In some embodiments, the endonuclease comprises a sequence having 100% identity to SEQ ID NO: 51 or SEQ ID NO: 1264.

[0265] In some embodiments, the endonuclease is a MG89 endonuclease (i.e., SEQ ID NO: 52). In some embodiments, the endonuclease comprises a sequence having 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 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 52. In some embodiments, the endonuclease comprises a sequence having at least about 70% identity to SEQ ID NO: 52. In some embodiments, the endonuclease comprises a sequence having at least about 75% identity to SEQ ID NO: 52. In some embodiments, the endonuclease comprises a sequence having at least about 80% identity to SEQ ID NO: 52. In some embodiments, the endonuclease comprises a sequence having at least about 85% identity to SEQ ID NO: 52. In some embodiments, the endonuclease comprises a sequence having at least about 90% identity to SEQ ID NO: 52. In some embodiments, the endonuclease comprises a sequence having at least about 95% identity to SEQ ID NO: 52. In some embodiments, the endonuclease comprises a sequence having at least about 96% identity to SEQ ID NO: 52. In some embodiments, the endonuclease comprises a sequence having at least about 97% identity to SEQ ID NO: 52. In some embodiments, the endonuclease comprises a sequence having at least about 98% identity to SEQ ID NO: 52. In some embodiments, the endonuclease comprises a sequence having at least about 99% identity to SEQ ID NO: 52. In some embodiments, the endonuclease comprises a sequence having 100% identity to SEQ ID NO: 52.

[0266] In some embodiments, the endonuclease is a MG87 endonuclease (i.e., SEQ ID NOs: 659-660, 1158-1159, and 1267-1277). In some embodiments, the endonuclease comprises a sequence having 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 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: 659-660, 1158-1159, and 1267-1277. In some embodiments, the endonuclease comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277. In some embodiments, the endonuclease comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277. In some embodiments, the endonuclease comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277. In some embodiments, the endonuclease comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277. In some embodiments, the endonuclease comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277. In some embodiments, the endonuclease comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277. In some embodiments, the endonuclease comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277. In some embodiments, the endonuclease comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277. In some embodiments, the endonuclease comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277. In some embodiments, the endonuclease comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277. In some embodiments, the endonuclease comprises a sequence having 100% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277.

[0267] In some embodiments, the endonuclease is a MG88 endonuclease (i.e., SEQ ID NOs: 661-678 and 1278-1282). In some embodiments, the endonuclease comprises a sequence having 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 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: 661-678 and 1278-1282. In some embodiments, the endonuclease comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 661-678 and 1278-1282. In some embodiments, the endonuclease comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 661-678 and 1278-1282. In some embodiments, the endonuclease comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 661-678 and 1278-1282. In some embodiments, the endonuclease comprises a sequence having at least

about 85% identity to any one of SEQ ID NOs: 661-678 and 1278-1282. In some embodiments, the endonuclease comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 661-678 and 1278-1282. In some embodiments, the endonuclease comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 661-678 and 1278-1282. In some embodiments, the endonuclease comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 661-678 and 1278-1282. In some embodiments, the endonuclease comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 661-678 and 1278-1282. In some embodiments, the endonuclease comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 661-678 and 1278-1282. In some embodiments, the endonuclease comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 661-678 and 1278-1282. In some embodiments, the endonuclease comprises a sequence having 100% identity to any one of SEQ ID NOs: 661-678 and 1278-1282.

[0268] In some embodiments, the endonuclease is a MG17 endonuclease (i.e., SEQ ID NO: 659). In some embodiments, the endonuclease comprises a sequence having 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 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 659. In some embodiments, the endonuclease comprises a sequence having at least about 70% identity to SEQ ID NO: 659. In some embodiments, the endonuclease comprises a sequence having at least about 75% identity to SEQ ID NO: 659. In some embodiments, the endonuclease comprises a sequence having at least about 80% identity to SEQ ID NO: 659. In some embodiments, the endonuclease comprises a sequence having at least about 85% identity to SEQ ID NO: 659. In some embodiments, the endonuclease comprises a sequence having at least about 90% identity to SEQ ID NO: 659. In some embodiments, the endonuclease comprises a sequence having at least about 95% identity to SEQ ID NO: 659. In some embodiments, the endonuclease comprises a sequence having at least about 96% identity to SEQ ID NO: 659. In some embodiments, the endonuclease comprises a sequence having at least about 97% identity to SEQ ID NO: 659. In some embodiments, the endonuclease comprises a sequence having at least about 98% identity to SEQ ID NO: 659. In some embodiments, the endonuclease comprises a sequence having at least about 99% identity to SEQ ID NO: 659. In some embodiments, the endonuclease comprises a sequence having 100% identity to SEQ ID NO: 659.

[0269] In some embodiments, the endonuclease is a MG18 endonuclease (i.e., SEQ ID NOs: 696-698). In some embodiments, the endonuclease comprises a sequence having 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 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: 696-698. In some embodiments, the endonuclease comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 696-698. In some embodiments, the endonuclease comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 696-698. In some embodiments, the endonuclease comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 696-698. In some embodiments, the endonuclease comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 696-698. In some embodiments, the endonuclease comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 696-698. In some embodiments, the endonuclease comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 696-698. In some embodiments, the endonuclease comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 696-698. In some embodiments, the endonuclease comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 696-698. In some embodiments, the endonuclease comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 696-698. In some embodiments, the endonuclease comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 696-698. In some embodiments, the endonuclease comprises a sequence having 100% identity to any one of SEQ ID NOs: 696-698.

[0270] In some embodiments, the endonuclease is a MG46 endonuclease (i.e., SEQ ID NOs: 975-981). In some embodiments, the endonuclease comprises a sequence having 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 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: 975-981. In some embodiments, the endonuclease comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 975-981. In some embodiments, the endonuclease comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 975-981. In some embodiments, the endonuclease comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 975-981. In some embodiments, the endonuclease comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 975-981. In some embodiments, the endonuclease comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 975-981. In some embodiments, the endonuclease comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 975-981. In some embodiments, the endonuclease comprises a sequence having at least about 96% identity to any one of SEQ ID

NOs: 975-981. In some embodiments, the endonuclease comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 975-981. In some embodiments, the endonuclease comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 975-981. In some embodiments, the endonuclease comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 975-981. In some embodiments, the endonuclease comprises a sequence having 100% identity to any one of SEQ ID NOs: 975-981.

[0271] In some embodiments, the endonuclease is a MG74 endonuclease (i.e., SEQ ID NO: 1265 or SEQ ID NO: 1266). In some embodiments, the endonuclease comprises a sequence having 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 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 NO: 1265 or SEQ ID NO: 1266. In some embodiments, the endonuclease comprises a sequence having at least about 70% identity to any one of SEQ ID NO: 1265 or SEQ ID NO: 1266. In some embodiments, the endonuclease comprises a sequence having at least about 75% identity to any one of SEQ ID NO: 1265 or SEQ ID NO: 1266. In some embodiments, the endonuclease comprises a sequence having at least about 80% identity to any one of SEQ ID NO: 1265 or SEQ ID NO: 1266. In some embodiments, the endonuclease comprises a sequence having at least about 85% identity to any one of SEQ ID NO: 1265 or SEQ ID NO: 1266. In some embodiments, the endonuclease comprises a sequence having at least about 90% identity to any one of SEQ ID NO: 1265 or SEQ ID NO: 1266. In some embodiments, the endonuclease comprises a sequence having at least about 95% identity to any one of SEQ ID NO: 1265 or SEQ ID NO: 1266. In some embodiments, the endonuclease comprises a sequence having at least about 96% identity to any one of SEQ ID NO: 1265 or SEQ ID NO: 1266. In some embodiments, the endonuclease comprises a sequence having at least about 97% identity to any one of SEQ ID NO: 1265 or SEQ ID NO: 1266. In some embodiments, the endonuclease comprises a sequence having at least about 98% identity to any one of SEQ ID NO: 1265 or SEQ ID NO: 1266. In some embodiments, the endonuclease comprises a sequence having at least about 99% identity to any one of SEQ ID NO: 1265 or SEQ ID NO: 1266. In some embodiments, the endonuclease comprises a sequence having 100% identity to any one of SEQ ID NO: 1265 or SEQ ID NO: 1266.

[0272] In some embodiments, the endonuclease comprises a nuclear localization sequence (NLS). In some embodiments, the NLS is at an N-terminus of the endonuclease. In some embodiments, the NLS is at a C-terminus of the endonuclease. In some embodiments, the NLS is at an N-terminus and a C-terminus of the endonuclease.

[0273] In some embodiments, the NLS comprises a sequence of any one of SEQ ID NOs: 1-46, or 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 any one of SEQ ID NOs: 1-46. In some cases, the NLS comprises a sequence having at least about 80% identity to SEQ ID NOs: 1-46. In some cases, the NLS comprises a sequence having at least about 85% identity to SEQ ID NOs: 1-46. In some cases, the NLS comprises a sequence having at least about 90% identity to SEQ ID NOs: 1-46. In some cases, the NLS comprises a sequence having at least about 91% identity to SEQ ID NOs: 1-46. In some cases, the NLS comprises a sequence having at least about 92% identity to SEQ ID NOs: 1-46. In some cases, the NLS comprises a sequence having at least about 93% identity to SEQ ID NOs: 1-46. In some cases, the NLS comprises a sequence having at least about 94% identity to SEQ ID NOs: 1-46. In some cases, the NLS comprises a sequence having at least about 95% identity to SEQ ID NOs: 1-46. In some cases, the NLS comprises a sequence having at least about 96% identity to SEQ ID NOs: 1-46. In some cases, the NLS comprises a sequence having at least about 97% identity to SEQ ID NOs: 1-46. In some cases, the NLS comprises a sequence having at least about 98% identity to SEQ ID NOs: 1-46. In some cases, the NLS comprises a sequence having at least about 99% identity to SEQ ID NOs: 1-46. In some cases, the NLS comprises a sequence having 100% identity to SEQ ID NOs: 1-46.

Table 1: Example NLS Sequences that can be used with Cas Effectors According to the Disclosure

Source	NLS amino acid sequence	SEQ ID NO:
SV40	PKKKRKV	1
nucleoplasmin bipartite NLS	KRPAATKKAGQAKKKK	2
c-myc NLS	PAAKRVKLD	3
c-myc NLS	RQRRNELKRSP	4
hRNPA1 M9 NLS	NQSSNFGPMKGGNFGGRSSGPYGGGGQYFA KPRNQGGY	5
Importin-alpha IBB domain	RMRIZFKNKGKDTAELRRRRVEVSVELRKAK KDEQILKRRNV	6
Myoma T protein	VSRKRPRP	7

Source	NLS amino acid sequence	SEQ ID NO:
Myoma T protein	PPKKARED	8
p53	PQPKKKPL	9
mouse c-abl IV	SALIKKKKKMAP	10
influenza virus NS1	DRLRR	11
influenza virus NS1	PKQKKRK	12
Hepatitis virus delta antigen	RKLKKKIKKL	13
mouse Mx1 protein	REKKKFLKRR	14
human poly(ADP-ribose) polymerase	KRKGDEV DGVDEVAKKKSKK	15
steroid hormone receptor (human) glucocorticoid	RKCLQAGMNLEARKTKK	16
5' SV40 NLS	MAPKKKRKVGGGGS	17
3' SV40 1 NLS	SGGAPKKKRKV	18
3' SV40 2 NLS	SGGAPKKKRKV	19
3' SV40 3 NLS	SGGAPKKKRKVSGGAPKKKRKVSGGAPKKK RKV	20
5' Nucleoplasmin NLS	MKRPAATKKAGQAKKKKGGGGS	21
3' Nucleoplasmin 1 NLS	SGGKRPAATKKAGQAKKKK	22
3' Nucleoplasmin 2 NLS	SGGKRPAATKKAGQAKKKKSGGKRPAATKK AGQAKKKK	23
3' Nucleoplasmin 3 NLS	SGGKRPAATKKAGQAKKKKSGGKRPAATKK AGQAKKKKSGGKRPAATKKAGQAKKKK	24
5' c-Myc NLS	MPAAKRVKLDGGGGS	25
3' c-Myc 1 NLS	SGGPAAKRVKLD	26
3' c-Myc 2 NLS	SGGPAAKRVKLDSGGPAAKRVKLD	27
3' c-Myc 3 NLS	SGGPAAKRVKLDSGGPAAKRVKLDSGGPAA KRVKLD	28
Synthetic 5' Class 2-1 NLS	MQAAKRPRTTGGGGS	29
Synthetic 3' Class 2-1 1 NLS	SGGQAAKRPRTT	30

Source	NLS amino acid sequence	SEQ ID NO:
Synthetic 3' Class 2-1 2 NLS	SGGQAAKRPRTTSGGQAAKRPRTT	31
Synthetic 5' Class 2-2 NLS	MRAAKRPRTTGGGGS	32
Synthetic 3' Class 2-2 1 NLS	SGGRAAKRPRTT	33
Synthetic 3' Class 2-2 2 NLS	SGGRAAKRPRTTSGGRAAKRPRTT	34
Synthetic 5' Class 2-2 NLS	MRAAKRPRTTGGGGS	35
Synthetic 3' Class 2-2 1 NLS	SGGRAAKRPRTT	36
Synthetic 3' Class 2-2 2 NLS	SGGRAAKRPRTTSGGRAAKRPRTT	37
Synthetic 5' Class 3 NLS	MAAAKRSWSMAFGGGGS	38
Synthetic 3' Class 3 1 NLS	MSGGAAAKRSWSMAF	39
Synthetic 3' Class 3 2 NLS	MSGGAAAKRSWSMAFMSGGAAAKRSWSMAF	40
Synthetic 5' Class 4 NLS	MRAAKRKYFAVGGGGS	41
Synthetic 3' Class 4 1 NLS	MSGGRAAKRKYFAV	42
Synthetic 3' Class 4 2 NLS	MSGGRAAKRKYFAVMSGGRAAKRKYFAV	43
5' HCV NLS	MPKPQRKTKRGGGGS	44
3' HCV NLS	SGGPPRKKRTVV	45
3' hRNP A1 NLS	SGGFGNYNNQSSNFGPMKGGNFGGRSSGPY	46

Guide Polynucleotides

[0274] In some embodiments, the engineered nuclease system disclosed herein comprises an engineered guide polynucleotide, e.g., a guide ribonucleic acid (gRNA), a single gRNA, or a dual guide RNA.

[0275] In some embodiments, the guide RNAs comprise various structural elements including but not limited to: a spacer sequence which binds to the protospacer sequence (target sequence), a crRNA, and an optional tracrRNA. In some embodiments, the guide RNA comprises a crRNA comprising a spacer sequence. In some embodiments, the guide RNA comprises a tracrRNA or a modified tracrRNA.

[0276] In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides 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: 710-722. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 70% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 75% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 80% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 85% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 90% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 91% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 92% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 93% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 94% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 95% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 96% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 97% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 98% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 99% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having 100% identity to any one of SEQ ID NOs: 710-722.

[0277] In some cases, the tracrRNA comprises 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 any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 710-722. In some cases, the tracrRNA comprises a sequence having 100% identity to any one of SEQ ID NOs: 710-722.

[0278] In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides 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: 726-744. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 70% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 75% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence

comprising at least about 60-100 consecutive nucleotides having at least about 80% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 85% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 90% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 91% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 92% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 93% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 94% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 95% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 96% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 97% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 98% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 99% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having 100% identity to any one of SEQ ID NOs: 726-744.

[0279] In some cases, the tracrRNA comprises 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 any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises

a sequence having at least about 80% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 726-744. In some cases, the tracrRNA comprises a sequence having 100% identity to any one of SEQ ID NOs: 726-744.

[0280] In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides 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: 745-767. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 70% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 75% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 80% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 85% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 90% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 91% identity to

any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 92% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 93% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 94% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 95% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 96% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 97% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 98% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 99% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having 100% identity to any one of SEQ ID NOs: 745-767.

[0281] In some cases, the tracrRNA comprises 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 any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence having at least about 93% identity

to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 745-767. In some cases, the tracrRNA comprises a sequence having 100% identity to any one of SEQ ID NOs: 745-767.

[0282] In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides 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 SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 70% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 75% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 80% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 85% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 90% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 91% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 92% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 93% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 94% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 95% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100

consecutive nucleotides having at least about 96% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 97% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 98% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 99% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having 100% identity to SEQ ID NO: 699.

[0283] In some cases, the tracrRNA comprises 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 SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence having at least about 70% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence having at least about 75% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence having at least about 80% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence having at least about 85% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence having at least about 90% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence having at least about 91% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence having at least about 92% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence having at least about 93% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence having at least about 94% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence having at least about 95% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence having at least about 96% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence having at least about 97% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence having at least about 98% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence having at least about 99% identity to SEQ ID NO: 699. In some cases, the tracrRNA comprises a sequence having 100% identity to SEQ ID NO: 699.

[0284] In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides 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: 700-702. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 70% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 75% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 80% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 85% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 90% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 91% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 92% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 93% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 94% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 95% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 96% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 97% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 98% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 99% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having 100% identity to any one of SEQ ID NOs: 700-702.

[0285] In some cases, the tracrRNA comprises 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 any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 700-702. In some cases, the tracrRNA comprises a sequence having 100% identity to any one of SEQ ID NOs: 700-702.

[0286] In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides 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: 703-709. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 70% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 75% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence

comprising at least about 60-100 consecutive nucleotides having at least about 80% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 85% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 90% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 91% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 92% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 93% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 94% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 95% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 96% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 97% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 98% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having at least about 99% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence comprising at least about 60-100 consecutive nucleotides having 100% identity to any one of SEQ ID NOs: 703-709.

[0287] In some cases, the tracrRNA comprises 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 any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises

a sequence having at least about 80% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 703-709. In some cases, the tracrRNA comprises a sequence having 100% identity to any one of SEQ ID NOs: 703-709.

[0288] In some embodiments, the engineered nuclease system disclosed herein comprises an engineered guide polynucleotide, e.g., a guide ribonucleic acid (gRNA), a single gRNA, or a dual guide RNA.

[0289] In some embodiments, the engineered guide polynucleotide comprises 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 any one of SEQ ID NOs: 1137-1144. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 1137-1144. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 1137-1144. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 1137-1144. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 1137-1144. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 1137-1144. In some

embodiments, the engineered guide polynucleotide comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 1137-1144. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 1137-1144. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 1137-1144. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 1137-1144. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 1137-1144. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 1137-1144. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 1137-1144. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 1137-1144. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 1137-1144. In some embodiments, the engineered guide polynucleotide comprises a sequence having 100% identity to any one of SEQ ID NOs: 1137-1144.

[0290] In some embodiments, the engineered guide polynucleotide comprises 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 any one of SEQ ID NOs: 1153-1156. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 1153-1156. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 1153-1156. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 1153-1156. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 1153-1156. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 1153-1156. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 1153-1156. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 92% identity to any one of SEQ

ID NOs: 1153-1156. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 1153-1156. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 1153-1156. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 1153-1156. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 1153-1156. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 1153-1156. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 1153-1156. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 1153-1156. In some embodiments, the engineered guide polynucleotide comprises a sequence having 100% identity to any one of SEQ ID NOs: 1153-1156.

[0291] In some embodiments, the engineered guide polynucleotide comprises 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 any one of SEQ ID NOs: 927-942, 945-950, 961-962, 969-970, 1123-1137, and 2433-2434. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 927-942, 945-950, 961-962, 969-970, 1123-1137, and 2433-2434. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 927-942, 945-950, 961-962, 969-970, 1123-1137, and 2433-2434. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 927-942, 945-950, 961-962, 969-970, 1123-1137, and 2433-2434. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 927-942, 945-950, 961-962, 969-970, 1123-1137, and 2433-2434. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 927-942, 945-950, 961-962, 969-970, 1123-1137, and 2433-2434. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 927-942, 945-950, 961-962, 969-970, 1123-1137, and 2433-2434. In some embodiments, the engineered

guide polynucleotide comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 927-942, 945-950, 961-962, 969-970, 1123-1137, and 2433-2434. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 927-942, 945-950, 961-962, 969-970, 1123-1137, and 2433-2434. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 927-942, 945-950, 961-962, 969-970, 1123-1137, and 2433-2434. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 927-942, 945-950, 961-962, 969-970, 1123-1137, and 2433-2434. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 927-942, 945-950, 961-962, 969-970, 1123-1137, and 2433-2434. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 927-942, 945-950, 961-962, 969-970, 1123-1137, and 2433-2434. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 927-942, 945-950, 961-962, 969-970, 1123-1137, and 2433-2434. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 927-942, 945-950, 961-962, 969-970, 1123-1137, and 2433-2434. In some embodiments, the engineered guide polynucleotide comprises a sequence having 100% identity to any one of SEQ ID NOs: 927-942, 945-950, 961-962, 969-970, 1123-1137, and 2433-2434.

[0292] In some embodiments, the engineered guide polynucleotide comprises 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 any one of SEQ ID NOs: 843-880 and 1145-1152. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 843-880 and 1145-1152. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 843-880 and 1145-1152. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 843-880 and 1145-1152. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 843-880 and 1145-1152. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least

about 90% identity to any one of SEQ ID NOs: 843-880 and 1145-1152. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 843-880 and 1145-1152. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 843-880 and 1145-1152. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 843-880 and 1145-1152. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 843-880 and 1145-1152. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 843-880 and 1145-1152. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 843-880 and 1145-1152. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 843-880 and 1145-1152. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 843-880 and 1145-1152. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 843-880 and 1145-1152. In some embodiments, the engineered guide polynucleotide comprises a sequence having 100% identity to any one of SEQ ID NOs: 843-880 and 1145-1152.

[0293] In some embodiments, the engineered guide polynucleotide comprises 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 any one of SEQ ID NOs: 881-926. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 881-926. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 881-926. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 881-926. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 881-926. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 881-926. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 91% identity to any one of SEQ

ID NOs: 881-926. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 881-926. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 881-926. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 881-926. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 881-926. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 881-926. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 881-926. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 881-926. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 881-926. In some embodiments, the engineered guide polynucleotide comprises a sequence having 100% identity to any one of SEQ ID NOs: 881-926. In some embodiments, the engineered guide polynucleotide comprises 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 SEQ ID NO: 973 or SEQ ID NO: 974. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 70% identity to SEQ ID NO: 973 or SEQ ID NO: 974. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 75% identity to SEQ ID NO: 973 or SEQ ID NO: 974. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 80% identity to SEQ ID NO: 973 or SEQ ID NO: 974. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 85% identity to SEQ ID NO: 973 or SEQ ID NO: 974. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 90% identity to SEQ ID NO: 973 or SEQ ID NO: 974. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 91% identity to SEQ ID NO: 973 or SEQ ID NO: 974. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 92% identity to SEQ ID NO: 973 or SEQ ID NO: 974. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 93% identity to SEQ ID NO: 973 or SEQ ID NO: 974. In some embodiments, the engineered guide polynucleotide

comprises a sequence having at least about 94% identity to SEQ ID NO: 973 or SEQ ID NO: 974. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 95% identity to SEQ ID NO: 973 or SEQ ID NO: 974. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 96% identity to SEQ ID NO: 973 or SEQ ID NO: 974. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 97% identity to SEQ ID NO: 973 or SEQ ID NO: 974. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 98% identity to SEQ ID NO: 973 or SEQ ID NO: 974. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 99% identity to SEQ ID NO: 973 or SEQ ID NO: 974. In some embodiments, the engineered guide polynucleotide comprises a sequence having 100% identity to SEQ ID NO: 973 or SEQ ID NO: 974.

[0294] In some embodiments, the engineered guide polynucleotide comprises 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 any one of SEQ ID NOs: 957-960. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 957-960. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 957-960. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 957-960. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 957-960. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 957-960. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 957-960. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 957-960. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 957-960. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 957-960. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 957-960. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 96% identity to

any one of SEQ ID NOs: 957-960. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 957-960. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 957-960. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 957-960. In some embodiments, the engineered guide polynucleotide comprises a sequence having 100% identity to any one of SEQ ID NOs: 957-960.

[0295] In some embodiments, the engineered guide polynucleotide comprises 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 any one of SEQ ID NOs: 943-944, 951-956, 965-968, and 971-972. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 943-944, 951-956, 965-968, and 971-972. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 943-944, 951-956, 965-968, and 971-972. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 943-944, 951-956, 965-968, and 971-972. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 943-944, 951-956, 965-968, and 971-972. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 943-944, 951-956, 965-968, and 971-972. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 943-944, 951-956, 965-968, and 971-972. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 943-944, 951-956, 965-968, and 971-972. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 943-944, 951-956, 965-968, and 971-972. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 943-944, 951-956, 965-968, and 971-972. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 943-944, 951-956, 965-968, and 971-972. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about

96% identity to any one of SEQ ID NOs: 943-944, 951-956, 965-968, and 971-972. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 943-944, 951-956, 965-968, and 971-972. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 943-944, 951-956, 965-968, and 971-972. In some embodiments, the engineered guide polynucleotide comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 943-944, 951-956, 965-968, and 971-972. In some embodiments, the engineered guide polynucleotide comprises a sequence having 100% identity to any one of SEQ ID NOs: 943-944, 951-956, 965-968, and 971-972.

[0296] In some embodiments, the engineered guide polynucleotide comprises synthetic nucleotides or modified nucleotides. In some embodiments, the engineered guide polynucleotide comprises one or more inter-nucleoside linkers modified from the natural phosphodiester. In some embodiments, all of the inter-nucleoside linkers of the engineered guide polynucleotide, or contiguous nucleotide sequence thereof, are modified. For example, in some embodiments, the inter nucleoside linkage comprises Sulphur (S), such as a phosphorothioate inter-nucleoside linkage.

[0297] In some embodiments, the engineered guide polynucleotide comprises modifications to a ribose sugar or nucleobase. In some embodiments, the engineered guide polynucleotide comprises one or more nucleosides comprising a modified sugar moiety, wherein the modified sugar moiety is a modification of the sugar moiety when compared to the ribose sugar moiety found in deoxyribose nucleic acid (DNA) and RNA. In some embodiments, the modification is within the ribose ring structure. Exemplary modifications include, but are not limited to, replacement with a hexose ring (HNA), a bicyclic ring having a biradical bridge between the C2 and C4 carbons on the ribose ring (e.g., locked nucleic acids (LNA)), or an unlinked ribose ring which typically lacks a bond between the C2 and C3 carbons (e.g., UNA). In some embodiments, the sugar-modified nucleosides comprise bicyclohexose nucleic acids or tricyclic nucleic acids. In some embodiments, the modified nucleosides comprise nucleosides where the sugar moiety is replaced with a non-sugar moiety, for example peptide nucleic acids (PNA) or morpholino nucleic acids.

[0298] In some embodiments, the engineered guide polynucleotide comprises one or more modified sugars. In some embodiments, the sugar modifications comprise modifications made by altering the substituent groups on the ribose ring to groups other than hydrogen, or the 2'-OH group naturally found in DNA and RNA nucleosides. In some embodiments, substituents are introduced at the 2', 3', 4', or 5' positions, or combinations thereof. In some embodiments, nucleosides with modified sugar moieties comprise 2' modified nucleosides, e.g., 2' substituted

nucleosides. A 2' sugar modified nucleoside, in some embodiments, is a nucleoside that has a substituent other than -H or -OH at the 2' position (2' substituted nucleoside) or comprises a 2' linked biradical, and comprises 2' substituted nucleosides and LNA (2'-4' biradical bridged) nucleosides. Examples of 2'-substituted modified nucleosides comprise, but are not limited to, 2'-O-alkyl-RNA, 2'-O-methyl-RNA, 2'-alkoxy-RNA, 2'-O-methoxyethyl-RNA (MOE), 2'-amino-DNA, 2'-Fluoro-RNA, and 2'-F-ANA nucleosides. In some embodiments, the modification in the ribose group comprises a modification at the 2' position of the ribose group. In some embodiments, the modification at the 2' position of the ribose group is selected from the group consisting of 2'-O-methyl, 2'-fluoro, 2'-deoxy, and 2'-O-(2-methoxyethyl).

[0299] In some embodiments, the engineered guide polynucleotide comprises one or more modified sugars. In some embodiments, the engineered guide polynucleotide comprises only modified sugars. In certain embodiments, the engineered guide polynucleotide comprises greater than about 10%, 25%, 50%, 75%, or 90% modified sugars. In some embodiments, the modified sugar is a bicyclic sugar. In some embodiments, the modified sugar comprises a 2'-O-methoxyethyl group. In some embodiments, the engineered guide polynucleotide comprises both inter-nucleoside linker modifications and nucleoside modifications.

[0300] In some embodiments, the engineered guide polynucleotide comprises a sequence complementary to a eukaryotic, fungal, plant, mammalian, or human genomic polynucleotide sequence. In some embodiments, the engineered guide polynucleotide comprises a sequence complementary to a eukaryotic genomic polynucleotide sequence. In some embodiments, the engineered guide polynucleotide comprises a sequence complementary to a fungal genomic polynucleotide sequence. In some embodiments, the engineered guide polynucleotide comprises a sequence complementary to a plant genomic polynucleotide sequence. In some embodiments, the engineered guide polynucleotide comprises a sequence complementary to a mammalian genomic polynucleotide sequence. In some embodiments, the engineered guide polynucleotide comprises a sequence complementary to a human genomic polynucleotide sequence.

[0301] In some embodiments, the engineered guide polynucleotide is 30-250 nucleotides in length. In some embodiments, the engineered guide polynucleotide is more than 90 nucleotides in length. In some embodiments, the engineered guide polynucleotide is less than 245 nucleotides in length. In some embodiments, the engineered guide polynucleotide is 30, 40, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 220, 240, or more than 240 nucleotides in length. In some embodiments, the engineered guide polynucleotide is about 30 to about 40, about 30 to about 50, about 30 to about 60, about 30 to about 70, about 30 to about 80, about 30 to about 90, about 30 to about 100, about 30 to about 120, about 30 to about 140, about 30 to about 160, about 30 to about 180, about 30 to about 200, about 30 to about 220, about 30 to about 240, about 50 to about

60, about 50 to about 70, about 50 to about 80, about 50 to about 90, about 50 to about 100, about 50 to about 120, about 50 to about 140, about 50 to about 160, about 50 to about 180, about 50 to about 200, about 50 to about 220, about 50 to about 240, about 100 to about 120, about 100 to about 140, about 100 to about 160, about 100 to about 180, about 100 to about 200, about 100 to about 220, about 100 to about 240, about 160 to about 180, about 160 to about 200, about 160 to about 220, or about 160 to about 240 nucleotides in length.

MG Systems

[0302] Described herein, in certain embodiments, are engineered nuclease systems comprising an endonuclease comprising and an engineered guide polynucleotide. In some embodiments, the engineered guide polynucleotide comprises a tracrRNA. In some embodiments, the engineered guide polynucleotide comprises a guide nucleic acid (e.g., gRNA). In a polynucleotide when referring to a T, a T means U (Uracil) in RNA and T (Thymine) in DNA.

[0303] In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about 70% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263 and an engineered polynucleotide comprising a sequence having at least about 70% identity to any one of SEQ ID NOs: 710-722. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 75% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263 and an engineered polynucleotide comprising a sequence having at least about 75% identity to any one of SEQ ID NOs: 710-722. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 80% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263 and an engineered polynucleotide comprising a sequence having at least about 80% identity to any one of SEQ ID NOs: 710-722. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 85% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263 and an engineered polynucleotide comprising a sequence having at least about 85% identity to any one of SEQ ID NOs: 710-722. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 90% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263 and an engineered polynucleotide comprising a sequence having at least about 90% identity to any one of SEQ ID NOs: 710-722. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 95% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263 and an engineered polynucleotide comprising a sequence having at least about 95% identity to any one of SEQ ID NOs: 710-722. In some

embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 96% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263 and an engineered polynucleotide comprising a sequence having at least about 96% identity to any one of SEQ ID NOs: 710-722. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 97% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263 and an engineered polynucleotide comprising a sequence having at least about 97% identity to any one of SEQ ID NOs: 710-722. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 98% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263 and an engineered polynucleotide comprising a sequence having at least about 98% identity to any one of SEQ ID NOs: 710-722. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 99% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263 and an engineered polynucleotide comprising a sequence having at least about 99% identity to any one of SEQ ID NOs: 710-722. In some embodiments, the engineered nuclease system comprises an endonuclease comprising 100% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263 and an engineered polynucleotide comprising 100% identity to any one of SEQ ID NOs: 710-722.

[0304] In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about 70% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277 and an engineered polynucleotide comprising a sequence having at least about 70% identity to any one of SEQ ID NOs: 726-744. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 75% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277 and an engineered polynucleotide comprising a sequence having at least about 75% identity to any one of SEQ ID NOs: 726-744. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 80% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277 and an engineered polynucleotide comprising a sequence having at least about 80% identity to any one of SEQ ID NOs: 726-744. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 85% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277 and an engineered polynucleotide comprising a sequence having at least about 85% identity to any one of SEQ ID NOs: 726-744. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence

having at least about at least about 90% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277 and an engineered polynucleotide comprising a sequence having at least about 90% identity to any one of SEQ ID NOs: 726-744. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 95% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277 and an engineered polynucleotide comprising a sequence having at least about 95% identity to any one of SEQ ID NOs: 726-744. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 96% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277 and an engineered polynucleotide comprising a sequence having at least about 96% identity to any one of SEQ ID NOs: 726-744. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 97% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277 and an engineered polynucleotide comprising a sequence having at least about 97% identity to any one of SEQ ID NOs: 726-744. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 98% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277 and an engineered polynucleotide comprising a sequence having at least about 98% identity to any one of SEQ ID NOs: 726-744. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 99% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277 and an engineered polynucleotide comprising a sequence having at least about 99% identity to any one of SEQ ID NOs: 726-744. In some embodiments, the engineered nuclease system comprises an endonuclease comprising 100% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277 and an engineered polynucleotide comprising 100% identity to any one of SEQ ID NOs: 726-744.

[0305] In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about 70% identity to any one of SEQ ID NOs: 661-678 and 1278-1282 and an engineered polynucleotide comprising a sequence having at least about 70% identity to any one of SEQ ID NOs: 745-767. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 75% identity to any one of SEQ ID NOs: 661-678 and 1278-1282 and an engineered polynucleotide comprising a sequence having at least about 75% identity to any one of SEQ ID NOs: 745-767. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 80% identity to any one of SEQ ID NOs: 661-678 and 1278-1282 and an engineered polynucleotide comprising a sequence having at least about 80% identity to any one of SEQ ID NOs: 745-767. In some embodiments, the engineered

nuclease system comprises an endonuclease comprising sequence having at least about at least about 85% identity to any one of SEQ ID NOs: 661-678 and 1278-1282 and an engineered polynucleotide comprising a sequence having at least about 85% identity to any one of SEQ ID NOs: 745-767. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 90% identity to any one of SEQ ID NOs: 661-678 and 1278-1282 and an engineered polynucleotide comprising a sequence having at least about 90% identity to any one of SEQ ID NOs: 745-767. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 95% identity to any one of SEQ ID NOs: 661-678 and 1278-1282 and an engineered polynucleotide comprising a sequence having at least about 95% identity to any one of SEQ ID NOs: 745-767. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 96% identity to any one of SEQ ID NOs: 661-678 and 1278-1282 and an engineered polynucleotide comprising a sequence having at least about 96% identity to any one of SEQ ID NOs: 745-767. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 97% identity to any one of SEQ ID NOs: 661-678 and 1278-1282 and an engineered polynucleotide comprising a sequence having at least about 97% identity to any one of SEQ ID NOs: 745-767. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 98% identity to any one of SEQ ID NOs: 661-678 and 1278-1282 and an engineered polynucleotide comprising a sequence having at least about 98% identity to any one of SEQ ID NOs: 745-767. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 99% identity to any one of SEQ ID NOs: 661-678 and 1278-1282 and an engineered polynucleotide comprising a sequence having at least about 99% identity to any one of SEQ ID NOs: 745-767. In some embodiments, the engineered nuclease system comprises an endonuclease comprising 100% identity to any one of SEQ ID NOs: 661-678 and 1278-1282 and an engineered polynucleotide comprising 100% identity to any one of SEQ ID NOs: 745-767.

[0306] In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about 70% identity to SEQ ID NO: 659 and an engineered polynucleotide comprising a sequence having at least about 70% identity to SEQ ID NO: 699. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 75% identity to SEQ ID NO: 659 and an engineered polynucleotide comprising a sequence having at least about 75% identity to SEQ ID NO: 699. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 80% identity to SEQ ID NO: 659 and an engineered

polynucleotide comprising a sequence having at least about 80% identity to SEQ ID NO: 699. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 85% identity to SEQ ID NO: 659 and an engineered polynucleotide comprising a sequence having at least about 85% identity to SEQ ID NO: 699. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 90% identity to SEQ ID NO: 659 and an engineered polynucleotide comprising a sequence having at least about 90% identity to SEQ ID NO: 699. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 95% identity to SEQ ID NO: 659 and an engineered polynucleotide comprising a sequence having at least about 95% identity to SEQ ID NO: 699. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 96% identity to SEQ ID NO: 659 and an engineered polynucleotide comprising a sequence having at least about 96% identity to SEQ ID NO: 699. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 97% identity to SEQ ID NO: 659 and an engineered polynucleotide comprising a sequence having at least about 97% identity to SEQ ID NO: 699. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 98% identity to SEQ ID NO: 659 and an engineered polynucleotide comprising a sequence having at least about 98% identity to SEQ ID NO: 699. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 99% identity to SEQ ID NO: 659 and an engineered polynucleotide comprising a sequence having at least about 99% identity to SEQ ID NO: 699. In some embodiments, the engineered nuclease system comprises an endonuclease comprising 100% identity to SEQ ID NO: 659 and an engineered polynucleotide comprising 100% identity to SEQ ID NO: 699.

[0307] In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about 70% identity to any one of SEQ ID NOs: 696-698 and an engineered polynucleotide comprising a sequence having at least about 70% identity to any one of SEQ ID NOs: 700-702. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 75% identity to any one of SEQ ID NOs: 696-698 and an engineered polynucleotide comprising a sequence having at least about 75% identity to any one of SEQ ID NOs: 700-702. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 80% identity to any one of SEQ ID NOs: 696-698 and an engineered polynucleotide comprising a sequence having at least about 80% identity to any one of SEQ ID

NOs: 700-702. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 85% identity to any one of SEQ ID NOs: 696-698 and an engineered polynucleotide comprising a sequence having at least about 85% identity to any one of SEQ ID NOs: 700-702. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 90% identity to any one of SEQ ID NOs: 696-698 and an engineered polynucleotide comprising a sequence having at least about 90% identity to any one of SEQ ID NOs: 700-702. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 95% identity to any one of SEQ ID NOs: 696-698 and an engineered polynucleotide comprising a sequence having at least about 95% identity to any one of SEQ ID NOs: 700-702. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 96% identity to any one of SEQ ID NOs: 696-698 and an engineered polynucleotide comprising a sequence having at least about 96% identity to any one of SEQ ID NOs: 700-702. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 97% identity to any one of SEQ ID NOs: 696-698 and an engineered polynucleotide comprising a sequence having at least about 97% identity to any one of SEQ ID NOs: 700-702. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 98% identity to any one of SEQ ID NOs: 696-698 and an engineered polynucleotide comprising a sequence having at least about 98% identity to any one of SEQ ID NOs: 700-702. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 99% identity to any one of SEQ ID NOs: 696-698 and an engineered polynucleotide comprising a sequence having at least about 99% identity to any one of SEQ ID NOs: 700-702. In some embodiments, the engineered nuclease system comprises an endonuclease comprising 100% identity to any one of SEQ ID NOs: 696-698 and an engineered polynucleotide comprising 100% identity to any one of SEQ ID NOs: 700-702.

[0308] In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about 70% identity to any one of SEQ ID NOs: 975-981 and an engineered polynucleotide comprising a sequence having at least about 70% identity to any one of SEQ ID NOs: 703-709. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 75% identity to any one of SEQ ID NOs: 975-981 and an engineered polynucleotide comprising a sequence having at least about 75% identity to any one of SEQ ID NOs: 703-709. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least

about at least about 80% identity to any one of SEQ ID NOs: 975-981 and an engineered polynucleotide comprising a sequence having at least about 80% identity to any one of SEQ ID NOs: 703-709. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 85% identity to any one of SEQ ID NOs: 975-981 and an engineered polynucleotide comprising a sequence having at least about 85% identity to any one of SEQ ID NOs: 703-709. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 90% identity to any one of SEQ ID NOs: 975-981 and an engineered polynucleotide comprising a sequence having at least about 90% identity to any one of SEQ ID NOs: 703-709. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 95% identity to any one of SEQ ID NOs: 975-981 and an engineered polynucleotide comprising a sequence having at least about 95% identity to any one of SEQ ID NOs: 703-709. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 96% identity to any one of SEQ ID NOs: 975-981 and an engineered polynucleotide comprising a sequence having at least about 96% identity to any one of SEQ ID NOs: 703-709. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 97% identity to any one of SEQ ID NOs: 975-981 and an engineered polynucleotide comprising a sequence having at least about 97% identity to any one of SEQ ID NOs: 703-709. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 98% identity to any one of SEQ ID NOs: 975-981 and an engineered polynucleotide comprising a sequence having at least about 98% identity to any one of SEQ ID NOs: 703-709. In some embodiments, the engineered nuclease system comprises an endonuclease comprising sequence having at least about at least about 99% identity to any one of SEQ ID NOs: 975-981 and an engineered polynucleotide comprising a sequence having at least about 99% identity to any one of SEQ ID NOs: 703-709. In some embodiments, the engineered nuclease system comprises an endonuclease comprising 100% identity to any one of SEQ ID NOs: 975-981 and an engineered polynucleotide comprising 100% identity to any one of SEQ ID NOs: 703-709.

Cells

[0309] Described herein, in certain embodiments, is a cell comprising the systems described herein.

[0310] In some embodiments, the cell is a eukaryotic cell (e.g., a plant cell, an animal cell, a protist cell, or a fungi cell), a mammalian cell (a Chinese hamster ovary (CHO) cell, baby hamster kidney (BHK), human embryo kidney (HEK), mouse myeloma (NS0), or human retinal

cells), an immortalized cell (e.g., a HeLa cell, a COS cell, a HEK-293T cell, a MDCK cell, a 3T3 cell, a PC12 cell, a Huh7 cell, a HepG2 cell, a K562 cell, a N2a cell, or a SY5Y cell), an insect cell (e.g., a *Spodoptera frugiperda* cell, a *Trichoplusia ni* cell, a *Drosophila melanogaster* cell, a S2 cell, or a *Heliothis virescens* cell), a yeast cell (e.g., a *Saccharomyces cerevisiae* cell, a Cryptococcus cell, or a Candida cell), a plant cell (e.g., a parenchyma cell, a collenchyma cell, or a sclerenchyma cell), a fungal cell (e.g., a *Saccharomyces cerevisiae* cell, a Cryptococcus cell, or a Candida cell), or a prokaryotic cell (e.g., a *E. coli* cell, a streptococcus bacterium cell, a streptomyces soil bacteria cell, or an archaea cell). In some embodiments, the cell is a eukaryotic cell. In some embodiments, the cell is a mammalian cell. In some embodiments, the cell is an immortalized cell. In some embodiments, the cell is an insect cell. In some embodiments, the cell is a yeast cell. In some embodiments, the cell is a plant cell. In some embodiments, the cell is a fungal cell. In some embodiments, the cell is a prokaryotic cell.

[0311] In some embodiments, the cell is an A549, HEK-293, HEK-293T, BHK, CHO, HeLa, MRC5, Sf9, Cos-1, Cos-7, Vero, BSC 1, BSC 40, BMT 10, WI38, HeLa, Saos, C2C12, L cell, HT1080, HepG2, Huh7, K562, a primary cell, or derivative thereof.

Delivery and Vectors

[0312] Disclosed herein, in some embodiments, are nucleic acid sequences encoding an engineered nuclease system disclosed herein.

[0313] In some embodiments, the nucleic acid encoding the engineered nuclease system is a DNA, for example a linear DNA, a plasmid DNA, or a minicircle DNA. In some embodiments, the nucleic acid encoding the engineered nuclease system is an RNA, for example a mRNA.

[0314] In some embodiments, the nucleic acid encoding the engineered nuclease system is delivered by a nucleic acid-based vector. In some embodiments, the nucleic acid-based vector is a plasmid (e.g., circular DNA molecules that can autonomously replicate inside a cell), cosmid (e.g., pWE or sCos vectors), artificial chromosome, human artificial chromosome (HAC), yeast artificial chromosomes (YAC), bacterial artificial chromosome (BAC), P1-derived artificial chromosomes (PAC), phagemid, phage derivative, bacmid, or virus. In some embodiments, the nucleic acid-based vector is selected from the list consisting of: pSF-CMV-NEO-NH2-PPT-3XFLAG, pSF-CMV-NEO-COOH-3XFLAG, pSF-CMV-PURO-NH2-GST-TEV, pSF-OXB20-COOH-TEV-FLAG(R)-6His, pCEP4 pDEST27, pSF-CMV-Ub-KrYFP, pSF-CMV-FMDV-daGFP, pEF1a-mCherry-N1 vector, pEF1a-tdTomato vector, pSF-CMV-FMDV-Hygro, pSF-CMV-PGK-Puro, pMCP-tag(m), pSF-CMV-PURO-NH2-CMYC, pSF-OXB20-BetaGal, pSF-OXB20-Fluc, pSF-OXB20, pSF-Tac, pRI 101-AN DNA, pCambia2301, pTYB21, pKLAC2, pAc5.1/V5-His A, and pDEST8.

[0315] In some embodiments, the nucleic acid-based vector comprises a promoter. In some embodiments, the promoter is selected from the group consisting of a mini promoter, an inducible promoter, a constitutive promoter, and derivatives thereof. In some embodiments, the promoter is selected from the group consisting of CMV, CBA, EF1a, CAG, PGK, TRE, U6, UAS, T7, Sp6, lac, araBad, trp, Ptac, p5, p19, p40, Synapsin, CaMKII, GRK1, and derivatives thereof. In some embodiments the promoter is a U6 promoter. In some embodiments, the promoter is a CAG promoter. In some embodiments, the promoter is encoded by a sequence of any one of SEQ ID NOs: 190-191, or 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 of any one of SEQ ID NOs: 190-191.

[0316] In some embodiments, the nucleic acid-based vector is a virus. In some embodiments, the virus is an alphavirus, a parvovirus, an adenovirus, an AAV, a baculovirus, a Dengue virus, a lentivirus, a herpesvirus, a poxvirus, an anellovirus, a bocavirus, a vaccinia virus, or a retrovirus. In some embodiments, the virus is an alphavirus. In some embodiments, the virus is a parvovirus. In some embodiments, the virus is an adenovirus. In some embodiments, the virus is an AAV. In some embodiments, the virus is a baculovirus. In some embodiments, the virus is a Dengue virus. In some embodiments, the virus is a lentivirus. In some embodiments, the virus is a herpesvirus. In some embodiments, the virus is a poxvirus. In some embodiments, the virus is an anellovirus. In some embodiments, the virus is a bocavirus. In some embodiments, the virus is a vaccinia virus. In some embodiments, the virus is or a retrovirus.

[0317] In some embodiments, the AAV is AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAV13, AAV14, AAV15, AAV16, AAV-rh8, AAV-rh10, AAV-rh20, AAV-rh39, AAV-rh74, AAV-rhM4-1, AAV-hu37, AAV-Anc80, AAV-Anc80L65, AAV-7m8, AAV-PHP-B, AAV-PHP-EB, AAV-2.5, AAV-2tYF, AAV-3B, AAV-LK03, AAV-HSC1, AAV-HSC2, AAV-HSC3, AAV-HSC4, AAV-HSC5, AAV-HSC6, AAV-HSC7, AAV-HSC8, AAV-HSC9, AAV-HSC10, AAV-HSC11, AAV-HSC12, AAV-HSC13, AAV-HSC14, AAV-HSC15, AAV-TT, AAV-DJ/8, AAV-Myo, AAV-NP40, AAV-NP59, AAV-NP22, AAV-NP66, AAV-HSC16, or a derivative thereof. In some embodiments, the herpesvirus is HSV type 1, HSV-2, VZV, EBV, CMV, HHV-6, HHV-7, or HHV-8.

[0318] In some embodiments, the virus is AAV1 or a derivative thereof. In some embodiments, the virus is AAV2 or a derivative thereof. In some embodiments, the virus is AAV3 or a derivative thereof. In some embodiments, the virus is AAV4 or a derivative thereof. In some

embodiments, the virus is AAV5 or a derivative thereof. In some embodiments, the virus is AAV6 or a derivative thereof. In some embodiments, the virus is AAV7 or a derivative thereof. In some embodiments, the virus is AAV8 or a derivative thereof. In some embodiments, the virus is AAV9 or a derivative thereof. In some embodiments, the virus is AAV10 or a derivative thereof. In some embodiments, the virus is AAV11 or a derivative thereof. In some embodiments, the virus is AAV12 or a derivative thereof. In some embodiments, the virus is AAV13 or a derivative thereof. In some embodiments, the virus is AAV14 or a derivative thereof. In some embodiments, the virus is AAV15 or a derivative thereof. In some embodiments, the virus is AAV16 or a derivative thereof. In some embodiments, the virus is AAV-rh8 or a derivative thereof. In some embodiments, the virus is AAV-rh10 or a derivative thereof. In some embodiments, the virus is AAV-rh20 or a derivative thereof. In some embodiments, the virus is AAV-rh39 or a derivative thereof. In some embodiments, the virus is AAV-rh74 or a derivative thereof. In some embodiments, the virus is AAV-rhM4-1 or a derivative thereof. In some embodiments, the virus is AAV-hu37 or a derivative thereof. In some embodiments, the virus is AAV-Anc80 or a derivative thereof. In some embodiments, the virus is AAV-Anc80L65 or a derivative thereof. In some embodiments, the virus is AAV-7m8 or a derivative thereof. In some embodiments, the virus is AAV-PHP-B or a derivative thereof. In some embodiments, the virus is AAV-PHP-EB or a derivative thereof. In some embodiments, the virus is AAV-2.5 or a derivative thereof. In some embodiments, the virus is AAV-2tYF or a derivative thereof. In some embodiments, the virus is AAV-3B or a derivative thereof. In some embodiments, the virus is AAV-LK03 or a derivative thereof. In some embodiments, the virus is AAV-HSC1 or a derivative thereof. In some embodiments, the virus is AAV-HSC2 or a derivative thereof. In some embodiments, the virus is AAV-HSC3 or a derivative thereof. In some embodiments, the virus is AAV-HSC4 or a derivative thereof. In some embodiments, the virus is AAV-HSC5 or a derivative thereof. In some embodiments, the virus is AAV-HSC6 or a derivative thereof. In some embodiments, the virus is AAV-HSC7 or a derivative thereof. In some embodiments, the virus is AAV-HSC8 or a derivative thereof. In some embodiments, the virus is AAV-HSC9 or a derivative thereof. In some embodiments, the virus is AAV-HSC10 or a derivative thereof. In some embodiments, the virus is AAV-HSC11 or a derivative thereof. In some embodiments, the virus is AAV-HSC12 or a derivative thereof. In some embodiments, the virus is AAV-HSC13 or a derivative thereof. In some embodiments, the virus is AAV-HSC14 or a derivative thereof. In some embodiments, the virus is AAV-HSC15 or a derivative thereof. In some embodiments, the virus is AAV-TT or a derivative thereof. In some embodiments, the virus is AAV-DJ/8 or a derivative thereof. In some embodiments, the virus is AAV-Myo or a derivative thereof. In some embodiments, the virus is AAV-NP40 or a derivative thereof. In

some embodiments, the virus is AAV-NP59 or a derivative thereof. In some embodiments, the virus is AAV-NP22 or a derivative thereof. In some embodiments, the virus is AAV-NP66 or a derivative thereof. In some embodiments, the virus is AAV-HSC16 or a derivative thereof.

[0319] In some embodiments, the virus is HSV-1 or a derivative thereof. In some embodiments, the virus is HSV-2 or a derivative thereof. In some embodiments, the virus is VZV or a derivative thereof. In some embodiments, the virus is EBV or a derivative thereof. In some embodiments, the virus is CMV or a derivative thereof. In some embodiments, the virus is HHV-6 or a derivative thereof. In some embodiments, the virus is HHV-7 or a derivative thereof. In some embodiments, the virus is HHV-8 or a derivative thereof.

[0320] In some embodiments, the nucleic acid encoding the engineered nuclease system is delivered by a non-nucleic acid-based delivery system (e.g., a non-viral delivery system). In some embodiments, the non-viral delivery system is a liposome. In some embodiments, the nucleic acid is associated with a lipid. The nucleic acid associated with a lipid, in some embodiments, is encapsulated in the aqueous interior of a liposome, interspersed within the lipid bilayer of a liposome, attached to a liposome via a linking molecule that is associated with both the liposome and the nucleic acid, entrapped in a liposome, complexed with a liposome, dispersed in a solution containing a lipid, mixed with a lipid, combined with a lipid, contained as a suspension in a lipid, contained or complexed with a micelle, or otherwise associated with a lipid. In some embodiments, the nucleic acid is comprised in a lipid nanoparticle (LNP).

[0321] In some embodiments, the fusion protein or genome editing system is introduced into the cell in any suitable way, either stably or transiently. In some embodiments, a fusion protein or genome editing system is transfected into the cell. In some embodiments, the cell is transduced or transfected with a nucleic acid construct that encodes a fusion protein or genome editing system. For example, a cell is transduced (e.g., with a virus encoding a fusion protein or genome editing system), or transfected (e.g., with a plasmid encoding a fusion protein or genome editing system) with a nucleic acid that encodes a fusion protein or genome editing system, or the translated fusion protein or genome editing system. In some embodiments, the transduction is a stable or transient transduction. In some embodiments, cells expressing a fusion protein or genome editing system or containing a fusion protein or genome editing system are transduced or transfected with one or more gRNA molecules, for example, when the fusion protein or genome editing system comprises a CRISPR nuclease. In some embodiments, a plasmid expressing a fusion protein or genome editing system is introduced into cells through electroporation, transient (e.g., lipofection) and stable genome integration (e.g., piggybac) and viral transduction (for example lentivirus or AAV) or other methods known to those of skill in the art. In some embodiments, the gene editing system is introduced into the cell as one or more polypeptides. In some

embodiments, delivery is achieved through the use of RNP complexes. Delivery methods to cells for polypeptides and/or RNPs are known in the art, for example by electroporation or by cell squeezing.

[0322] Exemplary methods of delivery of nucleic acids include lipofection, nucleofection, electroporation, stable genome integration (e.g., piggybac), microinjection, biolistics, virosomes, liposomes, immunoliposomes, polycation or lipid nucleic acid conjugates, naked DNA, artificial virions, and agent-enhanced uptake of DNA. Lipofection is described in e.g., U.S. Pat. Nos. 5,049,386; 4,946,787; and 4,897,355) and lipofection reagents are sold commercially (e.g., Transfectam™, Lipofectin™ and SF Cell Line 4D-Nucleofector X Kit™ (Lonza)). Cationic and neutral lipids that are suitable for efficient receptor-recognition lipofection of polynucleotides include those of WO 91/17424 and WO 91/16024. In some embodiments, the delivery is to cells (e.g., *in vitro* or *ex vivo* administration) or target tissues (e.g., *in vivo* administration). In some embodiments, the nucleic acid is comprised in a liposome or a nanoparticle that specifically targets a host cell.

[0323] Additional methods for the delivery of nucleic acids to cells are known to those skilled in the art. See, for example, US 2003/0087817.

Methods of Use

[0324] Described herein, in certain embodiments, are methods for modifying a target nucleic acid comprising providing an engineered nuclease system disclosed herein. In some embodiments, the engineered nuclease system comprises an endonuclease and an engineered guide polynucleotide. In some embodiments, the target nucleic acid is double stranded. In some embodiments, the target nucleic acid is double stranded DNA. In some embodiments, the target nucleic acid is single stranded.

[0325] In some embodiments, the methods are used to introduce a modification in the genome of a cell. In some embodiments, the modification is an insertion, deletion, or mutation. In some embodiments, the methods are used to introduce site-directed insertions, deletions, and/or mutations in the genome of a cell (for example an insertion and a mutation). In some embodiments, the methods are used in combination with a nucleic acid template to facilitate site-directed insertions into the genome of a cell.

[0326] In some embodiments, the cell is a human cell. In some embodiments, the cell genome or a vector comprised in the cell is modified. In some embodiments, the cell genome is modified *ex vivo*. In some embodiments, the cell genome is modified *in vivo*.

[0327] In some embodiments, the engineered guide polynucleotide targets a gene in a cell. In some embodiments, the engineered guide polynucleotide targets a gene in a mammalian cell. In

some embodiments, the mammalian cell is a pig, a cow, a goat, a sheep, a rodent, a rat, a mouse, a non-human primate, or a human cell.

[0328] In some embodiments, the target gene is the HAO1 gene. In some embodiments, the gRNA comprises a sequence having at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 47. In some embodiments, the gRNA comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 1283-1392 and 1502-1509. In some embodiments, the gRNA comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 1283-1392 and 1502-1509. In some embodiments, the gRNA comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 1283-1392 and 1502-1509. In some embodiments, the gRNA comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 1283-1392 and 1502-1509. In some embodiments, the gRNA comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 1283-1392 and 1502-1509. In some embodiments, the gRNA comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 1283-1392 and 1502-1509. In some embodiments, the gRNA comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 1283-1392 and 1502-1509. In some embodiments, the gRNA comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 1283-1392 and 1502-1509. In some embodiments, the gRNA comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 1283-1392 and 1502-1509. In some embodiments, the gRNA comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 1283-1392 and 1502-1509. In some embodiments, the gRNA comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 1283-1392 and 1502-1509. In some embodiments, the gRNA comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 1283-1392 and 1502-1509. In some embodiments, the gRNA comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 1283-1392 and 1502-1509. In some embodiments, the gRNA comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 1283-1392 and 1502-1509. In some embodiments, the gRNA comprises a sequence having 100% identity to any one of SEQ ID NOs: 1283-1392 and 1502-1509.

[0329] In some embodiments, the target gene is the AAVS1 gene. In some embodiments, the gRNA is engineered to function with an MG71 endonuclease (e.g., MG71-2; SEQ ID NO: 50) comprising a sequence having at least at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 49-51, 679-694,

983-1002, 1003-1010, and 1263. In some embodiments, the gRNA comprises 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 any one of SEQ ID NOs: 57-58, 101-119, 163-181, 225-243, 343-374, 421-433, 461-472, 499-511, 539-551, and 1098-1102. In some embodiments, the gRNA comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 57-58, 101-119, 163-181, 225-243, 343-374, 421-433, 461-472, 499-511, 539-551, and 1098-1102. In some embodiments, the gRNA comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 57-58, 101-119, 163-181, 225-243, 343-374, 421-433, 461-472, 499-511, 539-551, and 1098-1102. In some embodiments, the gRNA comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 57-58, 101-119, 163-181, 225-243, 343-374, 421-433, 461-472, 499-511, 539-551, and 1098-1102. In some embodiments, the gRNA comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 57-58, 101-119, 163-181, 225-243, 343-374, 421-433, 461-472, 499-511, 539-551, and 1098-1102. In some embodiments, the gRNA comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 57-58, 101-119, 163-181, 225-243, 343-374, 421-433, 461-472, 499-511, 539-551, and 1098-1102. In some embodiments, the gRNA comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 57-58, 101-119, 163-181, 225-243, 343-374, 421-433, 461-472, 499-511, 539-551, and 1098-1102. In some embodiments, the gRNA comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 57-58, 101-119, 163-181, 225-243, 343-374, 421-433, 461-472, 499-511, 539-551, and 1098-1102. In some embodiments, the gRNA comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 57-58, 101-119, 163-181, 225-243, 343-374, 421-433, 461-472, 499-511, 539-551, and 1098-1102. In some embodiments, the gRNA comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 57-58, 101-119, 163-181, 225-243, 343-374, 421-433, 461-472, 499-511, 539-551, and 1098-1102. In some embodiments, the gRNA comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 57-58, 101-119, 163-181, 225-243, 343-374, 421-433, 461-472, 499-511, 539-551, and 1098-1102. In some embodiments, the gRNA comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 57-58, 101-119, 163-181, 225-243, 343-374, 421-433, 461-472, 499-511, 539-551, and 1098-1102. In some embodiments, the gRNA comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 57-58, 101-119, 163-181, 225-243, 343-374, 421-433, 461-472, 499-511, 539-551,

and 1098-1102. In some embodiments, the gRNA comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 57-58, 101-119, 163-181, 225-243, 343-374, 421-433, 461-472, 499-511, 539-551, and 1098-1102. In some embodiments, the gRNA comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 57-58, 101-119, 163-181, 225-243, 343-374, 421-433, 461-472, 499-511, 539-551, and 1098-1102. In some embodiments, the gRNA comprises a sequence having 100% identity to any one of SEQ ID NOs: 57-58, 101-119, 163-181, 225-243, 343-374, 421-433, 461-472, 499-511, 539-551, and 1098-1102.

[0330] In some embodiments, the gRNA (e.g., that functions with an MG71 nuclease) hybridizes to the AAVS1 gene. In some embodiments, the gRNA hybridizes to 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 any one of SEQ ID NOs: 59-60, 120-138, 182-200, 244-262, 375-406, 434-446, 473-484, 512-524, 552-564, and 1248-1252. In some embodiments, the gRNA hybridizes to a sequence having at least about 70% identity to any one of SEQ ID NOs: 59-60, 120-138, 182-200, 244-262, 375-406, 434-446, 473-484, 512-524, 552-564, and 1248-1252. In some embodiments, the gRNA hybridizes to a sequence having at least about 75% identity to any one of SEQ ID NOs: 59-60, 120-138, 182-200, 244-262, 375-406, 434-446, 473-484, 512-524, 552-564, and 1248-1252. In some embodiments, the gRNA hybridizes to a sequence having at least about 80% identity to any one of SEQ ID NOs: 59-60, 120-138, 182-200, 244-262, 375-406, 434-446, 473-484, 512-524, 552-564, and 1248-1252. In some embodiments, the gRNA hybridizes to a sequence having at least about 85% identity to any one of SEQ ID NOs: 59-60, 120-138, 182-200, 244-262, 375-406, 434-446, 473-484, 512-524, 552-564, and 1248-1252. In some embodiments, the gRNA hybridizes to a sequence having at least about 90% identity to any one of SEQ ID NOs: 59-60, 120-138, 182-200, 244-262, 375-406, 434-446, 473-484, 512-524, 552-564, and 1248-1252. In some embodiments, the gRNA hybridizes to a sequence having at least about 91% identity to any one of SEQ ID NOs: 59-60, 120-138, 182-200, 244-262, 375-406, 434-446, 473-484, 512-524, 552-564, and 1248-1252. In some embodiments, the gRNA hybridizes to a sequence having at least about 92% identity to any one of SEQ ID NOs: 59-60, 120-138, 182-200, 244-262, 375-406, 434-446, 473-484, 512-524, 552-564, and 1248-1252. In some embodiments, the gRNA hybridizes to a sequence having at least about 93% identity to any one of SEQ ID NOs: 59-60, 120-138, 182-200, 244-262, 375-406, 434-446, 473-484, 512-524,

552-564, and 1248-1252. In some embodiments, the gRNA hybridizes to a sequence having at least about 94% identity to any one of SEQ ID NOs: 59-60, 120-138, 182-200, 244-262, 375-406, 434-446, 473-484, 512-524, 552-564, and 1248-1252. In some embodiments, the gRNA hybridizes to a sequence having at least about 95% identity to any one of SEQ ID NOs: 59-60, 120-138, 182-200, 244-262, 375-406, 434-446, 473-484, 512-524, 552-564, and 1248-1252. In some embodiments, the gRNA hybridizes to a sequence having at least about 96% identity to any one of SEQ ID NOs: 59-60, 120-138, 182-200, 244-262, 375-406, 434-446, 473-484, 512-524, 552-564, and 1248-1252. In some embodiments, the gRNA hybridizes to a sequence having at least about 97% identity to any one of SEQ ID NOs: 59-60, 120-138, 182-200, 244-262, 375-406, 434-446, 473-484, 512-524, 552-564, and 1248-1252. In some embodiments, the gRNA hybridizes to a sequence having at least about 98% identity to any one of SEQ ID NOs: 59-60, 120-138, 182-200, 244-262, 375-406, 434-446, 473-484, 512-524, 552-564, and 1248-1252. In some embodiments, the gRNA hybridizes to a sequence having at least about 99% identity to any one of SEQ ID NOs: 59-60, 120-138, 182-200, 244-262, 375-406, 434-446, 473-484, 512-524, 552-564, and 1248-1252. In some embodiments, the gRNA hybridizes to a sequence having 100% identity to any one of SEQ ID NOs: 59-60, 120-138, 182-200, 244-262, 375-406, 434-446, 473-484, 512-524, 552-564, and 1248-1252.

[0331] In some embodiments, the target gene is TRAC. In some embodiments, the gRNA is engineered to function with an MG71 endonuclease (e.g., MG71-2; SEQ ID NO: 50) comprising a sequence having at least at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263. In some embodiments, the gRNA comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 77-88, 139-150, 201-212, 407-413, 447-453, 485-491, 525-531, and 1096-1097. In some embodiments, the gRNA comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 77-88, 139-150, 201-212, 407-413, 447-453, 485-491, 525-531, and 1096-1097. In some embodiments, the gRNA comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 77-88, 139-150, 201-212, 407-413, 447-453, 485-491, 525-531, and 1096-1097. In some embodiments, the gRNA comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 77-88, 139-150, 201-212, 407-413, 447-453, 485-491, 525-531, and 1096-1097. In some embodiments, the gRNA comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 77-88, 139-150, 201-212, 407-413, 447-453, 485-491, 525-531, and 1096-1097. In some embodiments, the gRNA comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 77-88, 139-150, 201-212, 407-413, 447-453, 485-491, 525-531, and 1096-1097. In

some embodiments, the gRNA comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 77-88, 139-150, 201-212, 407-413, 447-453, 485-491, 525-531, and 1096-1097. In some embodiments, the gRNA comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 77-88, 139-150, 201-212, 407-413, 447-453, 485-491, 525-531, and 1096-1097. In some embodiments, the gRNA comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 77-88, 139-150, 201-212, 407-413, 447-453, 485-491, 525-531, and 1096-1097. In some embodiments, the gRNA comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 77-88, 139-150, 201-212, 407-413, 447-453, 485-491, 525-531, and 1096-1097. In some embodiments, the gRNA comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 77-88, 139-150, 201-212, 407-413, 447-453, 485-491, 525-531, and 1096-1097. In some embodiments, the gRNA comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 77-88, 139-150, 201-212, 407-413, 447-453, 485-491, 525-531, and 1096-1097. In some embodiments, the gRNA comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 77-88, 139-150, 201-212, 407-413, 447-453, 485-491, 525-531, and 1096-1097. In some embodiments, the gRNA comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 77-88, 139-150, 201-212, 407-413, 447-453, 485-491, 525-531, and 1096-1097. In some embodiments, the gRNA comprises a sequence having 100% identity to any one of SEQ ID NOs: 77-88, 139-150, 201-212, 407-413, 447-453, 485-491, 525-531, and 1096-1097.

[0332] In some embodiments, the gRNA (e.g., that functions with an MG71 nuclease) hybridizes to a TRAC 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 any one of SEQ ID NOs: 89-100, 151-162, 213-224, 414-420, 454-460, 492-498, 532-538, and 1246-1247. In some embodiments, the gRNA hybridizes to a TRAC sequence having at least about 70% identity to any one of SEQ ID NOs: 89-100, 151-162, 213-224, 414-420, 454-460, 492-498, 532-538, and 1246-1247. In some embodiments, the gRNA hybridizes to a TRAC sequence having at least about 75% identity to any one of SEQ ID NOs: 89-100, 151-162, 213-224, 414-420, 454-460, 492-498, 532-538, and 1246-1247. In some embodiments, the gRNA hybridizes to a TRAC sequence having at least about 80% identity to any one of SEQ ID NOs: 89-100, 151-162, 213-224, 414-420, 454-460, 492-498, 532-538, and 1246-1247. In some embodiments, the gRNA hybridizes to a TRAC sequence having at least about 85% identity to any one of SEQ ID NOs: 89-100, 151-162, 213-224, 414-420, 454-460,

492-498, 532-538, and 1246-1247. In some embodiments, the gRNA hybridizes to a TRAC sequence having at least about 90% identity to any one of SEQ ID NOs: 89-100, 151-162, 213-224, 414-420, 454-460, 492-498, 532-538, and 1246-1247. In some embodiments, the gRNA hybridizes to a TRAC sequence having at least about 91% identity to any one of SEQ ID NOs: 89-100, 151-162, 213-224, 414-420, 454-460, 492-498, 532-538, and 1246-1247. In some embodiments, the gRNA hybridizes to a TRAC sequence having at least about 92% identity to any one of SEQ ID NOs: 89-100, 151-162, 213-224, 414-420, 454-460, 492-498, 532-538, and 1246-1247. In some embodiments, the gRNA hybridizes to a TRAC sequence having at least about 93% identity to any one of SEQ ID NOs: 89-100, 151-162, 213-224, 414-420, 454-460, 492-498, 532-538, and 1246-1247. In some embodiments, the gRNA hybridizes to a TRAC sequence having at least about 94% identity to any one of SEQ ID NOs: 89-100, 151-162, 213-224, 414-420, 454-460, 492-498, 532-538, and 1246-1247. In some embodiments, the gRNA hybridizes to a TRAC sequence having at least about 95% identity to any one of SEQ ID NOs: 89-100, 151-162, 213-224, 414-420, 454-460, 492-498, 532-538, and 1246-1247. In some embodiments, the gRNA hybridizes to a TRAC sequence having at least about 96% identity to any one of SEQ ID NOs: 89-100, 151-162, 213-224, 414-420, 454-460, 492-498, 532-538, and 1246-1247. In some embodiments, the gRNA hybridizes to a TRAC sequence having at least about 97% identity to any one of SEQ ID NOs: 89-100, 151-162, 213-224, 414-420, 454-460, 492-498, 532-538, and 1246-1247. In some embodiments, the gRNA hybridizes to a TRAC sequence having at least about 98% identity to any one of SEQ ID NOs: 89-100, 151-162, 213-224, 414-420, 454-460, 492-498, 532-538, and 1246-1247. In some embodiments, the gRNA hybridizes to a TRAC sequence having at least about 99% identity to any one of SEQ ID NOs: 89-100, 151-162, 213-224, 414-420, 454-460, 492-498, 532-538, and 1246-1247. In some embodiments, the gRNA hybridizes to a TRAC sequence having 100% identity to any one of SEQ ID NOs: 89-100, 151-162, 213-224, 414-420, 454-460, 492-498, 532-538, and 1246-1247.

[0333] In some embodiments, the target gene is the B2M gene. In some embodiments, the gRNA is engineered to function with an MG71 endonuclease (e.g., MG71-2; SEQ ID NO: 50) comprising a sequence having at least at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263. In some embodiments, the gRNA comprises 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 any one of SEQ ID NOs: 565-578 and 1095. In some embodiments, the gRNA comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 565-578 and 1095. In some embodiments, the gRNA comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 565-578 and 1095. In some embodiments, the gRNA comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 565-578 and 1095. In some embodiments, the gRNA comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 565-578 and 1095. In some embodiments, the gRNA comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 565-578 and 1095. In some embodiments, the gRNA comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 565-578 and 1095. In some embodiments, the gRNA comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 565-578 and 1095. In some embodiments, the gRNA comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 565-578 and 1095. In some embodiments, the gRNA comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 565-578 and 1095. In some embodiments, the gRNA comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 565-578 and 1095. In some embodiments, the gRNA comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 565-578 and 1095. In some embodiments, the gRNA comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 565-578 and 1095. In some embodiments, the gRNA comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 565-578 and 1095. In some embodiments, the gRNA comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 565-578 and 1095. In some embodiments, the gRNA comprises a sequence having 100% identity to any one of SEQ ID NOs: 565-578 and 1095.

[0334] In some embodiments, the gRNA (e.g., that functions with an MG71 nuclease) hybridizes to a B2M gene 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 any one of SEQ ID NOs: 579-592 and 1245. In some embodiments, the gRNA hybridizes to a B2M sequence having at least about 70% identity to any one of SEQ ID NOs: 579-592 and 1245. In some embodiments, the gRNA hybridizes to a B2M sequence having at least about 75% identity to any one of SEQ ID NOs: 579-592 and 1245. In some embodiments, the gRNA hybridizes to a B2M gene sequence having at least about 80%

identity to any one of SEQ ID NOs: 579-592 and 1245. In some embodiments, the gRNA hybridizes to a B2M gene sequence having at least about 85% identity to any one of SEQ ID NOs: 579-592 and 1245. In some embodiments, the gRNA hybridizes to a B2M gene sequence having at least about 90% identity to any one of SEQ ID NOs: 579-592 and 1245. In some embodiments, the gRNA hybridizes to a B2M sequence having at least about 91% identity to any one of SEQ ID NOs: 579-592 and 1245. In some embodiments, the gRNA hybridizes to a B2M gene sequence having at least about 92% identity to any one of SEQ ID NOs: 579-592 and 1245. In some embodiments, the gRNA hybridizes to a B2M gene sequence having at least about 93% identity to any one of SEQ ID NOs: 579-592 and 1245. In some embodiments, the gRNA hybridizes to a B2M gene sequence having at least about 94% identity to any one of SEQ ID NOs: 579-592 and 1245. In some embodiments, the gRNA hybridizes to a B2M gene sequence having at least about 95% identity to any one of SEQ ID NOs: 579-592 and 1245. In some embodiments, the gRNA hybridizes to a B2M gene sequence having at least about 96% identity to any one of SEQ ID NOs: 579-592 and 1245. In some embodiments, the gRNA hybridizes to a B2M sequence having at least about 97% identity to any one of SEQ ID NOs: 579-592 and 1245. In some embodiments, the gRNA hybridizes to a B2M sequence having at least about 98% identity to any one of SEQ ID NOs: 579-592 and 1245. In some embodiments, the gRNA hybridizes to a B2M sequence having at least about 99% identity to any one of SEQ ID NOs: 579-592 and 1245. In some embodiments, the gRNA hybridizes to a B2M sequence having 100% identity to any one of SEQ ID NOs: 579-592 and 1245.

[0335] In some embodiments, the target gene is the HBB gene. In some embodiments, the gRNA is engineered to function with an MG71 endonuclease (e.g., MG71-2; SEQ ID NO: 50) comprising a sequence having at least at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263. In some embodiments, the gRNA comprises 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 any one of SEQ ID NOs: 593-625. In some embodiments, the gRNA comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 593-625. In some embodiments, the gRNA comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 593-625. In some embodiments, the gRNA comprises a sequence having at least about 80% identity to any

one of SEQ ID NOs: 593-625. In some embodiments, the gRNA comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 593-625. In some embodiments, the gRNA comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 593-625. In some embodiments, the gRNA comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 593-625. In some embodiments, the gRNA comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 593-625. In some embodiments, the gRNA comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 593-625. In some embodiments, the gRNA comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 593-625. In some embodiments, the gRNA comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 593-625. In some embodiments, the gRNA comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 593-625. In some embodiments, the gRNA comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 593-625. In some embodiments, the gRNA comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 593-625. In some embodiments, the gRNA comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 593-625. In some embodiments, the gRNA comprises a sequence having 100% identity to any one of SEQ ID NOs: 593-625.

[0336] In some embodiments, the gRNA (e.g., that functions with an MG71 nuclease) hybridizes to the HBB gene. In some embodiments, the gRNA hybridizes to 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 any one of SEQ ID NOs: 626-658. In some embodiments, the gRNA hybridizes to a sequence having at least about 70% identity to any one of SEQ ID NOs: 626-658. In some embodiments, the gRNA hybridizes to a sequence having at least about 75% identity to any one of SEQ ID NOs: 626-658. In some embodiments, the gRNA hybridizes to a sequence having at least about 80% identity to any one of SEQ ID NOs: 626-658. In some embodiments, the gRNA hybridizes to a sequence having at least about 85% identity to any one of SEQ ID NOs: 626-658. In some embodiments, the gRNA hybridizes to a sequence having at least about 90% identity to any one of SEQ ID NOs: 626-658. In some embodiments, the gRNA hybridizes to a sequence having at least about 91% identity to any one of SEQ ID NOs: 626-658. In some embodiments, the gRNA hybridizes to a sequence having at least about 92% identity to any one of SEQ ID NOs: 626-658. In some embodiments, the gRNA hybridizes to a sequence having at least about 93% identity to any one

of SEQ ID NOs: 626-658. In some embodiments, the gRNA hybridizes to a sequence having at least about 94% identity to any one of SEQ ID NOs: 626-658. In some embodiments, the gRNA hybridizes to a sequence having at least about 95% identity to any one of SEQ ID NOs: 626-658. In some embodiments, the gRNA hybridizes to a sequence having at least about 96% identity to any one of SEQ ID NOs: 626-658. In some embodiments, the gRNA hybridizes to a sequence having at least about 97% identity to any one of SEQ ID NOs: 626-658. In some embodiments, the gRNA hybridizes to a sequence having at least about 98% identity to any one of SEQ ID NOs: 626-658. In some embodiments, the gRNA hybridizes to a sequence having at least about 99% identity to any one of SEQ ID NOs: 626-658. In some embodiments, the gRNA hybridizes to a sequence having 100% identity to any one of SEQ ID NOs: 626-658.

[0337] In some embodiments, the target gene is the albumin gene. In some embodiments, the gRNA is engineered to function with an MG71 endonuclease (e.g., MG71-2; SEQ ID NO: 50) comprising a sequence having at least at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263. In some embodiments, the gRNA comprises 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 any one of SEQ ID NOs: 1083-1094. In some embodiments, the gRNA comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 1083-1094. In some embodiments, the gRNA comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 1083-1094. In some embodiments, the gRNA comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 1083-1094. In some embodiments, the gRNA comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 1083-1094. In some embodiments, the gRNA comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 1083-1094. In some embodiments, the gRNA comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 1083-1094. In some embodiments, the gRNA comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 1083-1094. In some embodiments, the gRNA comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 1083-1094. In some embodiments, the gRNA comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 1083-1094. In some embodiments, the gRNA comprises a sequence having at least about 95% identity to any one of SEQ ID NOs:

1083-1094. In some embodiments, the gRNA comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 1083-1094. In some embodiments, the gRNA comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 1083-1094. In some embodiments, the gRNA comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 1083-1094. In some embodiments, the gRNA comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 1083-1094. In some embodiments, the gRNA comprises a sequence having 100% identity to any one of SEQ ID NOs: 1083-1094.

[0338] In some embodiments, the gRNA (e.g., that functions with an MG71 nuclease) hybridizes to the albumin gene. In some embodiments, the gRNA hybridizes to 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 any one of SEQ ID NOs: 1233-1244. In some embodiments, the gRNA hybridizes to a sequence having at least about 70% identity to any one of SEQ ID NOs: 1233-1244. In some embodiments, the gRNA hybridizes to a sequence having at least about 75% identity to any one of SEQ ID NOs: 1233-1244. In some embodiments, the gRNA hybridizes to a sequence having at least about 80% identity to any one of SEQ ID NOs: 1233-1244. In some embodiments, the gRNA hybridizes to a sequence having at least about 85% identity to any one of SEQ ID NOs: 1233-1244. In some embodiments, the gRNA hybridizes to a sequence having at least about 90% identity to any one of SEQ ID NOs: 1233-1244. In some embodiments, the gRNA hybridizes to a sequence having at least about 91% identity to any one of SEQ ID NOs: 1233-1244. In some embodiments, the gRNA hybridizes to a sequence having at least about 92% identity to any one of SEQ ID NOs: 1233-1244. In some embodiments, the gRNA hybridizes to a sequence having at least about 93% identity to any one of SEQ ID NOs: 1233-1244. In some embodiments, the gRNA hybridizes to a sequence having at least about 94% identity to any one of SEQ ID NOs: 1233-1244. In some embodiments, the gRNA hybridizes to a sequence having at least about 95% identity to any one of SEQ ID NOs: 1233-1244. In some embodiments, the gRNA hybridizes to a sequence having at least about 96% identity to any one of SEQ ID NOs: 1233-1244. In some embodiments, the gRNA hybridizes to a sequence having at least about 97% identity to any one of SEQ ID NOs: 1233-1244. In some embodiments, the gRNA hybridizes to a sequence having at least about 98% identity to any one of SEQ ID NOs: 1233-1244. In some embodiments, the gRNA hybridizes to a sequence having at least about 99% identity to any one of SEQ ID NOs:

1233-1244. In some embodiments, the gRNA hybridizes to a sequence having 100% identity to any one of SEQ ID NOs: 1233-1244.

[0339] In some embodiments, the target gene is ATP7B. In some embodiments, the gRNA is engineered to function with an MG71 endonuclease (e.g., MG71-2; SEQ ID NO: 50) comprising a sequence having at least at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263. In some embodiments, the gRNA comprises 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 any one of SEQ ID NOs: 1906-1931. In some embodiments, the gRNA comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 1906-1931. In some embodiments, the gRNA comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 1906-1931. In some embodiments, the gRNA comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 1906-1931. In some embodiments, the gRNA comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 1906-1931. In some embodiments, the gRNA comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 1906-1931. In some embodiments, the gRNA comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 1906-1931. In some embodiments, the gRNA comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 1906-1931. In some embodiments, the gRNA comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 1906-1931. In some embodiments, the gRNA comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 1906-1931. In some embodiments, the gRNA comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 1906-1931. In some embodiments, the gRNA comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 1906-1931. In some embodiments, the gRNA comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 1906-1931. In some embodiments, the gRNA comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 1906-1931. In some embodiments, the gRNA comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 1906-1931. In some embodiments, the gRNA comprises a sequence having 100% identity to any one of SEQ ID NOs: 1906-1931.

[0340] In some embodiments, the gRNA (e.g., that functions with an MG71 nuclease) hybridizes to ATP7B. In some embodiments, the gRNA hybridizes to 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 any one of SEQ ID NOs: 1206-1231. In some embodiments, the gRNA hybridizes to a sequence having at least about 70% identity to any one of SEQ ID NOs: 1206-1231. In some embodiments, the gRNA hybridizes to a sequence having at least about 75% identity to any one of SEQ ID NOs: 1206-1231. In some embodiments, the gRNA hybridizes to a sequence having at least about 80% identity to any one of SEQ ID NOs: 1206-1231. In some embodiments, the gRNA hybridizes to a sequence having at least about 85% identity to any one of SEQ ID NOs: 1206-1231. In some embodiments, the gRNA hybridizes to a sequence having at least about 90% identity to any one of SEQ ID NOs: 1206-1231. In some embodiments, the gRNA hybridizes to a sequence having at least about 91% identity to any one of SEQ ID NOs: 1206-1231. In some embodiments, the gRNA hybridizes to a sequence having at least about 92% identity to any one of SEQ ID NOs: 1206-1231. In some embodiments, the gRNA hybridizes to a sequence having at least about 93% identity to any one of SEQ ID NOs: 1206-1231. In some embodiments, the gRNA hybridizes to a sequence having at least about 94% identity to any one of SEQ ID NOs: 1206-1231. In some embodiments, the gRNA hybridizes to a sequence having at least about 95% identity to any one of SEQ ID NOs: 1206-1231. In some embodiments, the gRNA hybridizes to a sequence having at least about 96% identity to any one of SEQ ID NOs: 1206-1231. In some embodiments, the gRNA hybridizes to a sequence having at least about 97% identity to any one of SEQ ID NOs: 1206-1231. In some embodiments, the gRNA hybridizes to a sequence having at least about 98% identity to any one of SEQ ID NOs: 1206-1231. In some embodiments, the gRNA hybridizes to a sequence having at least about 99% identity to any one of SEQ ID NOs: 1206-1231. In some embodiments, the gRNA hybridizes to a sequence having 100% identity to any one of SEQ ID NOs: 1206-1231.

[0341] In some embodiments, the target gene is the PAH gene. In some embodiments, the gRNA is engineered to function with an MG71 endonuclease (e.g., MG71-2; SEQ ID NO: 50) comprising a sequence having at least at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263. In some embodiments, the gRNA comprises 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 any one of SEQ ID NOs: 1113-1122. In some embodiments, the gRNA comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 1113-1122. In some embodiments, the gRNA comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 1113-1122. In some embodiments, the gRNA comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 1113-1122. In some embodiments, the gRNA comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 1113-1122. In some embodiments, the gRNA comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 1113-1122. In some embodiments, the gRNA comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 1113-1122. In some embodiments, the gRNA comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 1113-1122. In some embodiments, the gRNA comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 1113-1122. In some embodiments, the gRNA comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 1113-1122. In some embodiments, the gRNA comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 1113-1122. In some embodiments, the gRNA comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 1113-1122. In some embodiments, the gRNA comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 1113-1122. In some embodiments, the gRNA comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 1113-1122. In some embodiments, the gRNA comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 1113-1122. In some embodiments, the gRNA comprises a sequence having 100% identity to any one of SEQ ID NOs: 1113-1122.

[0342] In some embodiments, the gRNA (e.g., that functions with an MG71 nuclease) hybridizes to the PAH gene. In some embodiments, the gRNA hybridizes to 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 any one of SEQ ID NOs: 1253-1262. In some embodiments, the gRNA hybridizes to a sequence having at least about 70% identity to any one of SEQ ID NOs: 1253-1262. In some embodiments, the

gRNA hybridizes to a sequence having at least about 75% identity to any one of SEQ ID NOs: 1253-1262. In some embodiments, the gRNA hybridizes to a sequence having at least about 80% identity to any one of SEQ ID NOs: 1253-1262. In some embodiments, the gRNA hybridizes to a sequence having at least about 85% identity to any one of SEQ ID NOs: 1253-1262. In some embodiments, the gRNA hybridizes to a sequence having at least about 90% identity to any one of SEQ ID NOs: 1253-1262. In some embodiments, the gRNA hybridizes to a sequence having at least about 91% identity to any one of SEQ ID NOs: 1253-1262. In some embodiments, the gRNA hybridizes to a sequence having at least about 92% identity to any one of SEQ ID NOs: 1253-1262. In some embodiments, the gRNA hybridizes to a sequence having at least about 93% identity to any one of SEQ ID NOs: 1253-1262. In some embodiments, the gRNA hybridizes to a sequence having at least about 94% identity to any one of SEQ ID NOs: 1253-1262. In some embodiments, the gRNA hybridizes to a sequence having at least about 95% identity to any one of SEQ ID NOs: 1253-1262. In some embodiments, the gRNA hybridizes to a sequence having at least about 96% identity to any one of SEQ ID NOs: 1253-1262. In some embodiments, the gRNA hybridizes to a sequence having at least about 97% identity to any one of SEQ ID NOs: 1253-1262. In some embodiments, the gRNA hybridizes to a sequence having at least about 98% identity to any one of SEQ ID NOs: 1253-1262. In some embodiments, the gRNA hybridizes to a sequence having at least about 99% identity to any one of SEQ ID NOs: 1253-1262. In some embodiments, the gRNA hybridizes to a sequence having 100% identity to any one of SEQ ID NOs: 1253-1262.

[0343] In some embodiments, the target gene is TRAC. In some embodiments, the gRNA is engineered to function with an MG73 endonuclease (e.g., MG73-1; SEQ ID NO: 51) comprising a sequence having at least at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 51 or SEQ ID NO: 1264. In some embodiments, the gRNA comprises 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 SEQ ID NO: 61. In some embodiments, the gRNA comprises a sequence having at least about 70% identity to SEQ ID NO: 61. In some embodiments, the gRNA comprises a sequence having at least about 75% identity to SEQ ID NO: 61. In some embodiments, the gRNA comprises a sequence having at least about 80% identity to SEQ ID NO: 61. In some embodiments, the gRNA comprises a sequence having at

least about 85% identity to SEQ ID NO: 61. In some embodiments, the gRNA comprises a sequence having at least about 90% identity to SEQ ID NO: 61. In some embodiments, the gRNA comprises a sequence having at least about 91% identity to SEQ ID NO: 61. In some embodiments, the gRNA comprises a sequence having at least about 92% identity to SEQ ID NO: 61. In some embodiments, the gRNA comprises a sequence having at least about 93% identity to SEQ ID NO: 61. In some embodiments, the gRNA comprises a sequence having at least about 94% identity to SEQ ID NO: 61. In some embodiments, the gRNA comprises a sequence having at least about 95% identity to SEQ ID NO: 61. In some embodiments, the gRNA comprises a sequence having at least about 96% identity to SEQ ID NO: 61. In some embodiments, the gRNA comprises a sequence having at least about 97% identity to SEQ ID NO: 61. In some embodiments, the gRNA comprises a sequence having at least about 98% identity to SEQ ID NO: 61. In some embodiments, the gRNA comprises a sequence having at least about 99% identity to SEQ ID NO: 61. In some embodiments, the gRNA comprises a sequence having 100% identity to SEQ ID NO: 61.

[0344] In some embodiments, the gRNA (e.g., that functions with an MG73 nuclease) hybridizes to TRAC. In some embodiments, the gRNA hybridizes to 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 SEQ ID NO: 62. In some embodiments, the gRNA hybridizes to a sequence having at least about 70% identity to SEQ ID NO: 62. In some embodiments, the gRNA hybridizes to a sequence having at least about 75% identity to SEQ ID NO: 62. In some embodiments, the gRNA hybridizes to a sequence having at least about 80% identity to SEQ ID NO: 62. In some embodiments, the gRNA hybridizes to a sequence having at least about 85% identity to SEQ ID NO: 62. In some embodiments, the gRNA hybridizes to a sequence having at least about 90% identity to SEQ ID NO: 62. In some embodiments, the gRNA hybridizes to a sequence having at least about 91% identity to SEQ ID NO: 62. In some embodiments, the gRNA hybridizes to a sequence having at least about 92% identity to SEQ ID NO: 62. In some embodiments, the gRNA hybridizes to a sequence having at least about 93% identity to SEQ ID NO: 62. In some embodiments, the gRNA hybridizes to a sequence having at least about 94% identity to SEQ ID NO: 62. In some embodiments, the gRNA hybridizes to a sequence having at least about 95% identity to SEQ ID NO: 62. In some embodiments, the gRNA hybridizes to a sequence having at least about 96% identity to SEQ ID NO: 62. In some embodiments, the gRNA hybridizes to a sequence having at

least about 97% identity to SEQ ID NO: 62. In some embodiments, the gRNA hybridizes to a sequence having at least about 98% identity to SEQ ID NO: 62. In some embodiments, the gRNA hybridizes to a sequence having at least about 99% identity to SEQ ID NO: 62. In some embodiments, the gRNA hybridizes to a sequence having 100% identity to SEQ ID NO: 62.

[0345] In some embodiments, the target gene is TRAC. In some embodiments, the gRNA is engineered to function with an MG89 endonuclease comprising a sequence having at least at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 52. In some embodiments, the gRNA comprises 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 any one of SEQ ID NOs: 63-69 and 263-302. In some embodiments, the gRNA comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 63-69 and 263-302. In some embodiments, the gRNA comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 63-69 and 263-302. In some embodiments, the gRNA comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 63-69 and 263-302. In some embodiments, the gRNA comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 63-69 and 263-302. In some embodiments, the gRNA comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 63-69 and 263-302. In some embodiments, the gRNA comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 63-69 and 263-302. In some embodiments, the gRNA comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 63-69 and 263-302. In some embodiments, the gRNA comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 63-69 and 263-302. In some embodiments, the gRNA comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 63-69 and 263-302. In some embodiments, the gRNA comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 63-69 and 263-302. In some embodiments, the gRNA comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 63-69 and 263-302. In some embodiments, the gRNA comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 63-69 and 263-302. In some embodiments, the gRNA comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 63-69 and 263-302. In some embodiments, the gRNA comprises a sequence having at least about 99% identity to any one of

SEQ ID NOS: 63-69 and 263-302. In some embodiments, the gRNA comprises a sequence having 100% identity to any one of SEQ ID NOS: 63-69 and 263-302.

[0346] In some embodiments, the gRNA (e.g., that functions with an MG89 nuclease) hybridizes to TRAC. In some embodiments, the gRNA hybridizes to 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 any one of SEQ ID NOS: 70-76 and 303-342. In some embodiments, the gRNA hybridizes to a sequence having at least about 70% identity to any one of SEQ ID NOS: 70-76 and 303-342. In some embodiments, the gRNA hybridizes to a sequence having at least about 75% identity to any one of SEQ ID NOS: 70-76 and 303-342. In some embodiments, the gRNA hybridizes to a sequence having at least about 80% identity to any one of SEQ ID NOS: 70-76 and 303-342. In some embodiments, the gRNA hybridizes to a sequence having at least about 85% identity to any one of SEQ ID NOS: 70-76 and 303-342. In some embodiments, the gRNA hybridizes to a sequence having at least about 90% identity to any one of SEQ ID NOS: 70-76 and 303-342. In some embodiments, the gRNA hybridizes to a sequence having at least about 91% identity to any one of SEQ ID NOS: 70-76 and 303-342. In some embodiments, the gRNA hybridizes to a sequence having at least about 92% identity to any one of SEQ ID NOS: 70-76 and 303-342. In some embodiments, the gRNA hybridizes to a sequence having at least about 93% identity to any one of SEQ ID NOS: 70-76 and 303-342. In some embodiments, the gRNA hybridizes to a sequence having at least about 94% identity to any one of SEQ ID NOS: 70-76 and 303-342. In some embodiments, the gRNA hybridizes to a sequence having at least about 95% identity to any one of SEQ ID NOS: 70-76 and 303-342. In some embodiments, the gRNA hybridizes to a sequence having at least about 96% identity to any one of SEQ ID NOS: 70-76 and 303-342. In some embodiments, the gRNA hybridizes to a sequence having at least about 97% identity to any one of SEQ ID NOS: 70-76 and 303-342. In some embodiments, the gRNA hybridizes to a sequence having at least about 98% identity to any one of SEQ ID NOS: 70-76 and 303-342. In some embodiments, the gRNA hybridizes to a sequence having at least about 99% identity to any one of SEQ ID NOS: 70-76 and 303-342. In some embodiments, the gRNA hybridizes to a sequence having 100% identity to any one of SEQ ID NOS: 70-76 and 303-342.

[0347] In some embodiments, the target gene is TRAC. In some embodiments, the gRNA is engineered to function with an MG87 endonuclease comprising a sequence having at least at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at

least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277. In some embodiments, the gRNA comprises 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 any one of SEQ ID NOs: 1011-1019 and 1050. In some embodiments, the gRNA comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 1011-1019 and 1050. In some embodiments, the gRNA comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 1011-1019 and 1050. In some embodiments, the gRNA comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 1011-1019 and 1050. In some embodiments, the gRNA comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 1011-1019 and 1050. In some embodiments, the gRNA comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 1011-1019 and 1050. In some embodiments, the gRNA comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 1011-1019 and 1050. In some embodiments, the gRNA comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 1011-1019 and 1050. In some embodiments, the gRNA comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 1011-1019 and 1050. In some embodiments, the gRNA comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 1011-1019 and 1050. In some embodiments, the gRNA comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 1011-1019 and 1050. In some embodiments, the gRNA comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 1011-1019 and 1050. In some embodiments, the gRNA comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 1011-1019 and 1050. In some embodiments, the gRNA comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 1011-1019 and 1050. In some embodiments, the gRNA comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 1011-1019 and 1050. In some embodiments, the gRNA comprises a sequence having 100% identity to any one of SEQ ID NOs: 1011-1019 and 1050.

[0348] In some embodiments, the gRNA (e.g., that functions with an MG87 nuclease) hybridizes to TRAC. In some embodiments, the gRNA hybridizes to 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 any one of SEQ ID NOs: 1161-1169 and 1200. In some embodiments, the gRNA hybridizes to a sequence having at least about 70% identity to any one of SEQ ID NOs: 1161-1169 and 1200. In some embodiments, the gRNA hybridizes to a sequence having at least about 75% identity to any one of SEQ ID NOs: 1161-1169 and 1200. In some embodiments, the gRNA hybridizes to a sequence having at least about 80% identity to any one of SEQ ID NOs: 1161-1169 and 1200. In some embodiments, the gRNA hybridizes to a sequence having at least about 85% identity to any one of SEQ ID NOs: 1161-1169 and 1200. In some embodiments, the gRNA hybridizes to a sequence having at least about 90% identity to any one of SEQ ID NOs: 1161-1169 and 1200. In some embodiments, the gRNA hybridizes to a sequence having at least about 91% identity to any one of SEQ ID NOs: 1161-1169 and 1200. In some embodiments, the gRNA hybridizes to a sequence having at least about 92% identity to any one of SEQ ID NOs: 1161-1169 and 1200. In some embodiments, the gRNA hybridizes to a sequence having at least about 93% identity to any one of SEQ ID NOs: 1161-1169 and 1200. In some embodiments, the gRNA hybridizes to a sequence having at least about 94% identity to any one of SEQ ID NOs: 1161-1169 and 1200. In some embodiments, the gRNA hybridizes to a sequence having at least about 95% identity to any one of SEQ ID NOs: 1161-1169 and 1200. In some embodiments, the gRNA hybridizes to a sequence having at least about 96% identity to any one of SEQ ID NOs: 1161-1169 and 1200. In some embodiments, the gRNA hybridizes to a sequence having at least about 97% identity to any one of SEQ ID NOs: 1161-1169 and 1200. In some embodiments, the gRNA hybridizes to a sequence having at least about 98% identity to any one of SEQ ID NOs: 1161-1169 and 1200. In some embodiments, the gRNA hybridizes to a sequence having at least about 99% identity to any one of SEQ ID NOs: 1161-1169 and 1200. In some embodiments, the gRNA hybridizes to a sequence having 100% identity to any one of SEQ ID NOs: 1161-1169 and 1200.

[0349] In some embodiments, the target gene is the AAVS1 gene. In some embodiments, the gRNA is engineered to function with an MG87 endonuclease comprising a sequence having at least at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 659-660, 1158-1159, and 1267-1277. In some embodiments, the gRNA comprises 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 any one of SEQ ID NOs: 1020-1049 and 1051-1055. In some embodiments, the gRNA comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 1020-1049 and 1051-1055. In some embodiments, the gRNA comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 1020-1049 and 1051-1055. In some embodiments, the gRNA comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 1020-1049 and 1051-1055. In some embodiments, the gRNA comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 1020-1049 and 1051-1055. In some embodiments, the gRNA comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 1020-1049 and 1051-1055. In some embodiments, the gRNA comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 1020-1049 and 1051-1055. In some embodiments, the gRNA comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 1020-1049 and 1051-1055. In some embodiments, the gRNA comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 1020-1049 and 1051-1055. In some embodiments, the gRNA comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 1020-1049 and 1051-1055. In some embodiments, the gRNA comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 1020-1049 and 1051-1055. In some embodiments, the gRNA comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 1020-1049 and 1051-1055. In some embodiments, the gRNA comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 1020-1049 and 1051-1055. In some embodiments, the gRNA comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 1020-1049 and 1051-1055. In some embodiments, the gRNA comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 1020-1049 and 1051-1055. In some embodiments, the gRNA comprises a sequence having 100% identity to any one of SEQ ID NOs: 1020-1049 and 1051-1055.

[0350] In some embodiments, the gRNA (e.g., that functions with an MG87 nuclease) hybridizes to the AAVS1 gene. In some embodiments, the gRNA hybridizes to 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 any one of SEQ ID NOs: 1170-1205. In some embodiments, the gRNA hybridizes to a sequence having at least about 70% identity to any one of SEQ ID NOs: 1170-1205. In some

embodiments, the gRNA hybridizes to a sequence having at least about 75% identity to any one of SEQ ID NOs: 1170-1205. In some embodiments, the gRNA hybridizes to a sequence having at least about 80% identity to any one of SEQ ID NOs: 1170-1205. In some embodiments, the gRNA hybridizes to a sequence having at least about 85% identity to any one of SEQ ID NOs: 1170-1205. In some embodiments, the gRNA hybridizes to a sequence having at least about 90% identity to any one of SEQ ID NOs: 1170-1205. In some embodiments, the gRNA hybridizes to a sequence having at least about 91% identity to any one of SEQ ID NOs: 1170-1205. In some embodiments, the gRNA hybridizes to a sequence having at least about 92% identity to any one of SEQ ID NOs: 1170-1205. In some embodiments, the gRNA hybridizes to a sequence having at least about 93% identity to any one of SEQ ID NOs: 1170-1205. In some embodiments, the gRNA hybridizes to a sequence having at least about 94% identity to any one of SEQ ID NOs: 1170-1205. In some embodiments, the gRNA hybridizes to a sequence having at least about 95% identity to any one of SEQ ID NOs: 1170-1205. In some embodiments, the gRNA hybridizes to a sequence having at least about 96% identity to any one of SEQ ID NOs: 1170-1205. In some embodiments, the gRNA hybridizes to a sequence having at least about 97% identity to any one of SEQ ID NOs: 1170-1205. In some embodiments, the gRNA hybridizes to a sequence having at least about 98% identity to any one of SEQ ID NOs: 1170-1205. In some embodiments, the gRNA hybridizes to a sequence having at least about 99% identity to any one of SEQ ID NOs: 1170-1205. In some embodiments, the gRNA hybridizes to a sequence having 100% identity to any one of SEQ ID NOs: 1170-1205.

[0351] In some embodiments, the target gene is the ATXN2 gene. In some embodiments, the gRNA is engineered to function with an MG21 endonuclease comprising a sequence having at least at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 47. In some embodiments, the gRNA comprises 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 any one of SEQ ID NOs: 1786-2045. In some embodiments, the gRNA comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 1786-2045. In some embodiments, the gRNA comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 1786-2045. In some embodiments, the gRNA comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 1786-2045. In some embodiments, the gRNA

comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 1786-2045. In some embodiments, the gRNA comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 1786-2045. In some embodiments, the gRNA comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 1786-2045. In some embodiments, the gRNA comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 1786-2045. In some embodiments, the gRNA comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 1786-2045. In some embodiments, the gRNA comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 1786-2045. In some embodiments, the gRNA comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 1786-2045. In some embodiments, the gRNA comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 1786-2045. In some embodiments, the gRNA comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 1786-2045. In some embodiments, the gRNA comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 1786-2045. In some embodiments, the gRNA comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 1786-2045. In some embodiments, the gRNA comprises a sequence having 100% identity to any one of SEQ ID NOs: 1786-2045.

[0352] In some embodiments, the gRNA (e.g., that functions with an MG87 nuclease) hybridizes to the ATXN2 gene. In some embodiments, the gRNA hybridizes to 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 any one of SEQ ID NOs: 1526-1785. In some embodiments, the gRNA hybridizes to a sequence having at least about 70% identity to any one of SEQ ID NOs: 1526-1785. In some embodiments, the gRNA hybridizes to a sequence having at least about 75% identity to any one of SEQ ID NOs: 1526-1785. In some embodiments, the gRNA hybridizes to a sequence having at least about 80% identity to any one of SEQ ID NOs: 1526-1785. In some embodiments, the gRNA hybridizes to a sequence having at least about 85% identity to any one of SEQ ID NOs: 1526-1785. In some embodiments, the gRNA hybridizes to a sequence having at least about 90% identity to any one of SEQ ID NOs: 1526-1785. In some embodiments, the gRNA hybridizes to a sequence having at least about 91% identity to any one of SEQ ID NOs: 1526-1785. In some embodiments, the gRNA hybridizes to a sequence having at least about 92% identity to any one of SEQ ID NOs: 1526-1785. In some embodiments, the gRNA hybridizes to a sequence having at least about 93% identity to any one of SEQ ID NOs: 1526-1785. In some embodiments, the

gRNA hybridizes to a sequence having at least about 94% identity to any one of SEQ ID NOs: 1526-1785. In some embodiments, the gRNA hybridizes to a sequence having at least about 95% identity to any one of SEQ ID NOs: 1526-1785. In some embodiments, the gRNA hybridizes to a sequence having at least about 96% identity to any one of SEQ ID NOs: 1526-1785. In some embodiments, the gRNA hybridizes to a sequence having at least about 97% identity to any one of SEQ ID NOs: 1526-1785. In some embodiments, the gRNA hybridizes to a sequence having at least about 98% identity to any one of SEQ ID NOs: 1526-1785. In some embodiments, the gRNA hybridizes to a sequence having at least about 99% identity to any one of SEQ ID NOs: 1526-1785. In some embodiments, the gRNA hybridizes to a sequence having 100% identity to any one of SEQ ID NOs: 1526-1785.

[0353] In some embodiments, the target gene is the AAVS1 gene. In some embodiments, the gRNA is engineered to function with an MG23 endonuclease comprising a sequence having at least at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 48. In some embodiments, the gRNA comprises 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 any one of SEQ ID NOs: 1393-1493. In some embodiments, the gRNA comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 1393-1493. In some embodiments, the gRNA comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 1393-1493. In some embodiments, the gRNA comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 1393-1493. In some embodiments, the gRNA comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 1393-1493. In some embodiments, the gRNA comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 1393-1493. In some embodiments, the gRNA comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 1393-1493. In some embodiments, the gRNA comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 1393-1493. In some embodiments, the gRNA comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 1393-1493. In some embodiments, the gRNA comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 1393-1493. In some embodiments, the gRNA comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 1393-1493. In some embodiments, the gRNA comprises a sequence having at least

about 96% identity to any one of SEQ ID NOs: 1393-1493. In some embodiments, the gRNA comprises a sequence having at least about 97% identity to any one of SEQ ID NOs: 1393-1493. In some embodiments, the gRNA comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 1393-1493. In some embodiments, the gRNA comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 1393-1493. In some embodiments, the gRNA comprises a sequence having 100% identity to any one of SEQ ID NOs: 1393-1493.

[0354] In some embodiments, the target gene is the AAVS1 gene. In some embodiments, the gRNA is engineered to function with an MG23 endonuclease comprising a sequence having at least at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 48. In some embodiments, the gRNA comprises 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 any one of SEQ ID NOs: 1494-1501 and 1510-1525. In some embodiments, the gRNA comprises a sequence having at least about 70% identity to any one of SEQ ID NOs: 1494-1501 and 1510-1525. In some embodiments, the gRNA comprises a sequence having at least about 75% identity to any one of SEQ ID NOs: 1494-1501 and 1510-1525. In some embodiments, the gRNA comprises a sequence having at least about 80% identity to any one of SEQ ID NOs: 1494-1501 and 1510-1525. In some embodiments, the gRNA comprises a sequence having at least about 85% identity to any one of SEQ ID NOs: 1494-1501 and 1510-1525. In some embodiments, the gRNA comprises a sequence having at least about 90% identity to any one of SEQ ID NOs: 1494-1501 and 1510-1525. In some embodiments, the gRNA comprises a sequence having at least about 91% identity to any one of SEQ ID NOs: 1494-1501 and 1510-1525. In some embodiments, the gRNA comprises a sequence having at least about 92% identity to any one of SEQ ID NOs: 1494-1501 and 1510-1525. In some embodiments, the gRNA comprises a sequence having at least about 93% identity to any one of SEQ ID NOs: 1494-1501 and 1510-1525. In some embodiments, the gRNA comprises a sequence having at least about 94% identity to any one of SEQ ID NOs: 1494-1501 and 1510-1525. In some embodiments, the gRNA comprises a sequence having at least about 95% identity to any one of SEQ ID NOs: 1494-1501 and 1510-1525. In some embodiments, the gRNA comprises a sequence having at least about 96% identity to any one of SEQ ID NOs: 1494-1501 and 1510-1525. In some embodiments, the gRNA comprises a sequence having at least about 97% identity

to any one of SEQ ID NOs: 1494-1501 and 1510-1525. In some embodiments, the gRNA comprises a sequence having at least about 98% identity to any one of SEQ ID NOs: 1494-1501 and 1510-1525. In some embodiments, the gRNA comprises a sequence having at least about 99% identity to any one of SEQ ID NOs: 1494-1501 and 1510-1525. In some embodiments, the gRNA comprises a sequence having 100% identity to any one of SEQ ID NOs: 1494-1501 and 1510-1525.

EXAMPLES

Example 1 - Gene editing outcomes at the DNA level for TRAC and AAVS1 in K562 cells

[0355] Nucleofection of MG71-2 mRNA along with the matching guide RNA (500 ng mRNA/150 pmol guide) was performed into K562 cells (200,000). Cells were harvested and genomic DNA prepared three 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 and analyzed to measure gene editing (**FIG. 1**).

Example 2 – Gene editing outcomes at the DNA level for TRAC in K562 cells

[0356] Nucleofection of MG89-2 mRNA along with the matching guide RNA (500 ng mRNA/150 pmol guide) was performed into K562 cells (200,000). Cells were harvested and genomic DNA prepared three 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 and analyzed to measure gene editing (**FIG. 2**).

Example 3 – Analysis of gene editing outcomes of the MG71-2 scaffold with guide hAAVS1 C3 at the DNA level

[0357] 1.2×10^5 K562 cells were nucleofected with 500 ng mRNA/150 pmol guide, 250 ng mRNA/75 pmol guide, or 125 ng mRNA/37.5 pmol guide. Cells were harvested and genomic DNA prepared three 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. Amplicons were sequenced and analyzed to measure gene editing (**FIG. 3A**). For analysis at the DNA level of second round of engineering for MG71-2 scaffold with guide hAAVS1 C3, 1.2×10^5 K562 cells were nucleofected with 500 ng mRNA/150 pmol guide and processed for downstream steps as described above (**FIG. 3B**).

Example 4 – Analysis of gene editing outcomes with original vs optimized MG71-2 scaffold at the DNA levels

[0358] 1.2×10^5 K562 cells were nucleofected with 500 ng mRNA/150 pmol guide and processed for downstream steps. Spacers indicated on the Y axis were tested with either a 22nt or 24nt spacer and either the original or shortened MG71-2 scaffold 11 (**FIG. 4**).

Example 5 – Analysis of gene editing outcomes with original vs optimized MG71-2 scaffold at the DNA levels

[0359] 1.2×10^5 K562 cells were nucleofected with 500 ng mRNA/150 pmol guide. Cells were harvested and genomic DNA prepared three 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. Amplicons were sequenced and analyzed to measure gene editing. Editing outcomes with MG71-2 engineered guides (scaffold 21) at exon 2 of the human B2M gene are shown in **FIG. 5A**. Editing outcomes with MG71-2 engineered guides (scaffold 21) at exon 3 of the human HBB gene are shown in **FIG. 5B**.

Example 6 – In silico identification of Type II CRISPR effectors**[0360] *In silico identification of Type II CRISPR effectors***

[0361] Putative Type II CRISPR effectors were identified by searching an extensive database. The resulting homologs were filtered to include those with an e-value $\leq 1e^{-5}$, a length ≥ 500 aa, and an associated CRISPR loci predicted using minCED. The effectors were dereplicated at 99% amino acid identity (AAI), globally aligned, and a phylogenetic tree was constructed.

[0362] *Ancestral reconstruction of MG71 nucleases*

[0363] In order to generate further diversity of the MG71 family of Type II nucleases, ancestral sequence reconstruction (ASR) algorithms were used. ASR is a computational technique that uses existing protein sequences and the relationships inferred between them to reconstruct potential ancestral sequences from ancient organisms. This technique was used to reconstruct sequences from the MG71 family. For this analysis, 432 Type II nuclease sequences were aligned and a phylogenetic tree was built. The trees were rooted using two Type II-B sequences as outgroups. Sequence reconstruction was done. Insertions and deletions were identified manually for each reconstructed node. Four ancestral sequences were reconstructed with high confidence: MG71-40, MG71-41, MG71-42, MG71-43 (**FIG. 6**, SEQ ID NOs: 691-694).

[0364] *Results*

[0365] Type II effectors were identified from the MG71, MG87, and MG88 families (SEQ ID NOs: 659-690). All the Type II effectors have the catalytic residues required for activity and range in length between 1,031 and 1,438 aa.

Example 7 – sgRNA design and in vitro activity of MG71, MG73, MG74, MG87, and MG88 nucleases

[0366] *sgRNA design*

[0367] TracrRNAs were predicted. TracrRNAs (SEQ ID NOs: 699-767) and repeats (SEQ ID NOs: 768-836) were folded, trimmed, and connected with a tetraloop sequence GAAA, or TTCG if GAAA altered the fold. The effectors were screened with multiple sgRNA designs (SEQ ID NOs: 837-974) consisting of a short and long scaffold with multiple spacers varying in length (20, 22, and 24 nt) and GC content (40% and 67%).

[0368] *In vitro activity assay*

[0369] Nucleases were expressed using 5-10 nM of a PCR-generated template. After expression, the nucleases were diluted 10-fold and incubated for 1 hour in a mixture containing 5 nM of an 8N protospacer adjacent motif (PAM) library plasmid and 50 nM sgRNA in 10 mM Tris pH 7.5, 100 mM NaCl, and 10 mM MgCl₂. The plasmid digest was cleaned using SPRI beads and eluted in TE buffer. The digested PAM plasmids (15 nM) were blunt-end ligated to double-stranded adapter oligos (150 nM) with T4 ligase in 1X T4 ligase buffer (NEB). The ligated product was sequenced using 150 bp single read amplicon sequencing. The resulting reads were filtered by a quality score >20. PAMs were identified by mapping the reads to the PAM plasmid backbones requiring a perfect match. SeqLogos of the PAMs were generated using and PAM sequences were determined by the height of each nucleotide. The cut site was identified by calculating the distance between the PAM and the ligated adapter.

[0370] *Results*

[0371] Multiple candidates from each family were active, demonstrating site-directed nuclease activity with their respective sgRNA designs. The cut sites and PAMs identified from the NGS data are summarized in **Tables 2-6** and example PAM SeqLogos are in **FIGs. 7-11**. Each family has a variety of PAMs, demonstrating the versatility of these systems as potential gene editing agents. The nucleases derived from metagenomic data from the MG71 family, which includes the MG18 and MG46 families, display a range of PAM specificities, with preference for an A in the 4th position (**FIG. 7**). The MG74, including MG17, family has A/C rich PAMs (**FIG. 9**). The MG87 family of nucleases have the strongest preference in the 5th and 6th bases from the spacer (**FIG. 10**). The MG88 family tend to have purine-rich PAMs in the second through fourth positions (**FIG. 11**).

Table 2: Summary of activity for MG71 family

Nuclease	PAM	Cut site
MG71-36	nRnRMY	3
MG71-37	nnnRCC	3
MG71-38	nRWRYY	3
MG71-39	nRRRY	3
MG71-27	nnCA	3
MG71-28	nnnAHT	3
MG46-69	nRWRYY	3
MG71-29	nRMRYYY	3
MG71-30	nRTR	3
MG71-31	nnnAYT	3
MG46-88	nnnRMY	3
MG46-90	nnYMA	3
MG46-93	nnnAM	3
MG18-72	nnnATT	3
MG18-90	nnRAMW	3
MG71-32	nRWAMY	3
MG18-103	nnnACT	3
MG46-147	nnTWT	3
MG46-153	nnTAYY	3
MG71-33	nnTRRW	3
MG46-163	nRTAnY	3

Table 3: Summary of activity for MG73 family

Nuclease	PAM	Cut site
MG73-6	nnRnCTW	3

Table 4: Summary of activity for MG74 family

Nuclease	PAM	Cut site
MG74-14	NNNNMNAN	3
MG74-8	NNNNMNAN	3
MG17-17	NNNNCMWA	3

Table 5: Summary of activity for MG87 family

Nuclease	PAM	Cut site
MG87-102	NNNRGTYN	3
MG87-103	NNNTGTYN	3
MG87-21	NNNNCCNN	3
MG87-37	NNNRRTYN	3
MG87-6	NNNNCCNN	3
MG87-63	NNNNMYNN	3
MG87-67	NNNNGTNN	3
MG87-69	NNNNGTNN	3

Nuclease	PAM	Cut site
MG87-70	NNNNGTNN	3
MG87-71	NNNWGTWN	3
MG87-74	NNTTCCYN	3
MG87-75	NNNNCYYN	3
MG87-83	NNACRTWN	3
MG87-88	NNARGTTN	3
MG87-94	NNATYYYN	3

Table 6: Summary of activity for MG88 family

Nuclease	PAM	Cut site
MG88-11	NRRRWNNN	3
MG88-16	NRRRNGNN	3
MG88-22	NRRRAKNN	3
MG88-23	NRRRANNN	3
MG88-32	NRRWYTNN	3
MG88-33	NRRRWANN	3
MG88-34	NRRANNNN	3
MG88-35	NAMRNYTN	3
MG88-36	NRRNATNN	3
MG88-37	NRRRRNNN	3
MG88-38	NRRARYNN	3
MG88-39	NRRNAYNN	3
MG88-40	NRRRNYNN	3
MG88-41	NRRNRANN	3
MG88-42	NRRRRGNA	3
MG88-43	NRRRANNN	3
MG88-44	NRRRRNNN	3
MG88-45	NRRRRWNN	3
MG88-46	NNANGNNN	3
MG88-47	NRRNRNNN	3
MG88-48	WNRRNAAN	2
MG88-49	NRRRGNNN	3
MG88-5	NRRNNYNN	3

[0372] Additional candidates were constructed using ancestral sequence reconstruction. Ancestors MG71-42 and MG71-43 share 91.0% and 81.1% AAI with MG71-2, respectively. Two out of three tested candidates showed robust activity with MG71-2, MG71-1, or MG18-1 sgRNA (**FIGs. 12A-12B, Table 7**). The resulting PAM of MG71-42 is similar to the natural metagenomic nuclease MG71-2. The PAM of MG71-43 is more relaxed in vitro than that of MG71-2, but the MG71-2 PAM NNNACT is a subset of the observed PAM.

Table 7: Summary of activity for MG71 ASRs

Nuclease	PAM	Cut site
----------	-----	----------

MG71-42	NNNACT	3
MG71-43	NNNRYY	3

Example 8 – Activity of MG71-2 gRNAs in mammalian cells

[0373] To further test the activity of MG71-2, guides targeting a region of the human AAVS1 locus were designed with the MG71-2 in vitro PAM. 50,000 K562 cells were nucleofected with 500 ng MG71-2 mRNA and 150 pmol chemically-synthesized guide RNA in a 96-well plate format and a cell-type specific program recommended by the manufacturer. gDNA was extracted at 72 h post-nucleofection. PCR primers appropriate for use in NGS-based DNA sequencing were generated, optimized, and used to amplify the target sequence. The amplicons were sequenced and analyzed to measure gene editing.

[0374] *Results*

[0375] Several guides showed > 50% activity, further supporting the robust activities of the nuclease (FIG. 13).

Example 9 – In vitro activity of Type II CRISPR systems with rational point mutations

[0376] *Ancestral Sequence Reconstruction (ASR) to generate MG71 variants*

[0377] To generate additional diversity of the MG71 family of Type II nucleases, ancestral sequence reconstruction (ASR) algorithms were used, as described in Example 6 above. For this analysis, 432 Type II nuclease sequences were aligned using MAFFT with parameters, a phylogenetic tree was built, and ancestral nodes were reconstructed. Eight ancestral nodes were reconstructed: MG71-88, MG71-89, MG71-90, MG71-91, MG71-92, MG71-95, MG71-96, MG71-97 (FIG. 14, SEQ ID NOs. 1003-1010).

[0378] *In vitro activity assay*

[0379] As described in Example 7 above, nucleases were expressed using 5-15 nM of a PCR-generated template with a T7 promoter. After expression, the nucleases were diluted 10-fold and digested for 1 hr in a mixture containing 5 nM protospacer adjacent motif (PAM) library plasmid and 50 nM sgRNA in 10 mM Tris pH 7.5, 100 mM NaCl, and 10 mM MgCl₂. The digest was cleaned using the SPRI beads and eluted in the TE buffer. The digested PAM plasmids (15 nM) were blunt-end ligated to double-stranded adapter oligos (150 nM) with T4 ligase in 1X T4 ligase buffer (NEB). The ligated product was amplified with NGS adapters and sequenced using 300 bp single-read amplicon sequencing. The resulting reads were filtered by a quality score >20. The PAMs were identified by mapping the reads to the PAM plasmid backbones requiring a perfect match. SeqLogos of the PAMs were generated and the PAM sequences were determined by the

height of each nucleotide. The cut site was identified by calculating the distance between the PAM and the ligated adapter.

[0380] Results

[0381] Based on the Alphafold2 predicted structure of MG71-2 overlaid with the structure of SpCas9 (PDB ID: 4UN3), 16 residues within the first interaction shell of the PAM base pair (FIG. 15) were selected. To change the PAM sequence of MG71-2, the 16 residues were mutated so that positively charged residues were either installed or removed. Analogous residues were identified in MG71-43 (SEQ ID NO: 694) to generate four MG71-43 mutants. These were screened in vitro with the MG71-2 optimized sgRNA 71-2_C2D2_U40_22 (SEQ ID NO: 982). The target sequence of the guide RNA matched the sequence adjacent to the PAM library with a PAM design of TNNNNNA (5N PAM). Because the PAM region has a designed T in the first position, the exact nucleotide preference was not known, based on the WT MG71-2 protein's PAM, it is assumed to be N. This is reported in the table below as a lower-case n since it is not empirically tested. These variants were all active in vitro (Table 8, FIG. 16):

Table 8: PAMs of single-point mutant variants of MG71 proteins

Protein	Point mutation	Protein Seq ID	PAM
prt1556	MG71-2(G1208R)	983	nNNRMT
prt1557	MG71-2(A1337R)	984	nRNACT
prt1558	MG71-2(R1206A)	985	nNNACT
prt1559	MG71-2(T1207R)	986	nRNRCT
prt1560	MG71-2(K1209A)	987	nNNRH Y
prt1561	MG71-2(C1348R)	988	nNNRMT
prt1562	MG71-2(G1349K)	989	nNNRCY
prt1568	MG71-2(Q1210R)	990	nNNRYY
prt1569	MG71-2(I1350R)	991	nNNRMT
prt1570	MG71-2(L1351R)	992	nNNRCY
prt1571	MG71-2(K1352A)	993	nNWRAT
prt1572	MG71-2(V1353R)	994	nNNRMT
prt1573	MG71-2(S1354R)	995	nNNRHT
prt1574	MG71-2(N1355R)	996	nNNRH Y
prt1575	MG71-2(N1356R)	997	nNNRH Y
prt1576	MG71-2(I1357R)	998	nNNRH Y
prt22902	MG71-43(Q1206R)	999	nRNRYY
prt22903	MG71-43(Q1206A)	1000	nRNRCN
prt22904	MG71-43(I1347R)	1001	nNNRYY
prt22905	MG71-43(K1348A)	1002	nRNRNT

[0382] From these results, we see multiple residues can affect the PAM, with the strongest effect to the PAM specificity resulting from the mutations in residues K1352, V1353, and Q1210 in MG71-2 (protein SEQ ID NOs: 990, 993, and 994; PAM's nNWRAT, nNNRMT, and nNNRYY) and the homologous residues Q1206 and K1348 in MG71-43 (protein SEQ ID NOs: 1000 and 1002; PAM's nRNRCN and nRNRNT). In contrast other mutations in MG71-2, especially R1206A, L1351R, N1355R and N1356R variants (SEQ ID NOs: 985, 992, 996, and 997) does not lead to changes to the native PAM motif of the wild-type enzyme.

[0383] Additionally, eight ancestral sequence reconstructions (ASRs) were screened for activity in vitro with the same MG71-2 sgRNA. The PAMs for these variants are shown in **Table 9** and **FIG. 17**. As has been observed with other nucleases of the MG71 family, we see a tendency for A or R in the fourth position and T or Y in the 6th position of the PAM.

Table 9: PAMs of ASR variants of MG71 proteins

Protein	Protein Seq ID	PAM
MG71-88	1003	nNNAWT
MG71-89	1004	nNRAMW
MG71-90	1005	nNARMY
MG71-91	1006	nNNRCT
MG71-92	1007	nNNRYT
MG71-95	1008	nNARYY
MG71-96	1009	nNNRCY
MG71-97	1010	nNNRYY

Example 10 – In cell editing activity of MG nucleases at endogenous loci

[0384] mRNA synthesis

[0385] The CDS codifying for nucleases MG87-21 (SEQ ID NO: 1158), MG87-70 (SEQ ID NO: 1159), MG71-2 (SEQ ID NO: 50), MG71-43 (SEQ ID NO: 694), MG21-1 (SEQ ID NO: 47), and MG23-1 (SEQ ID NO: 48) were cloned into a pUC19 plasmid including a RNA-pol T7 promoter, 5' and 3' UTRs, appropriate NLS sequences, and a 100 nt polyA tail (**Table 10**). To linearize the plasmid, 100 µg of plasmid was digested with SapI. The plasmid was purified with phenol/chloroform and precipitated with 70% ethanol. The DNA pellet was resuspended in 20 µl of nuclease-free water. For *in vitro* transcription, 1 µg of linearized plasmid DNA was added to a 20 µl reaction containing 1X reaction buffer (40 mM Tris-HCl pH 7.5, 16.5 mM MgCl₂, 50 mM NaCl, 2.5 mM Spermidine, 1 mM DTT) and 750 units of Hi-T7 RNA Polymerase. The reaction was incubated at 50°C for 1 hr. Transcribed mRNA was purified. Alternatively, the *in vitro* transcription template could be generated via the PCR instead of plasmid digestion. In this case,

the ORF+UTRs was cloned into a plasmid without the promoter or a polyA tail and these were added via overhangs on the amplification primers. After clean-up via SPRI beads or gel extraction, this PCR product was used in the same *in vitro* transcription conditions. The optimized RNA sequences are referenced in **Table 10**.

Table 10: mRNA sequence information for different nucleases

Nuclease system	mRNA sequence name	Seq ID of corresponding mRNA
MG71-2	mr557	1103
MG71-43	mr759	1104
MG87-21	MG87-21_mRNA	1105
MG71-88	mr1346	1106
MG87-70	mr328	1107
MG23-1	mr600	1108
MG21-1	mr006	1109
prt22905	mr1511	1110
prt22903	mr1505	1111
prt22904	mr1510	1112
prt22902	mr1504	1157

[0386] Mammalian cell activity assays

[0387] K562 cells were cultured in IMDM + 10% FBS at 37 °C and 5% CO₂ per 100,000-150,000 cells were nucleofected with the nuclease mRNA along with the indicated chemically-synthesized guide RNA (500 ng mRNA/150 pmol guide per well). 72 hr post-nucleofection, the cells were harvested and genomic DNA was extracted. The PCR primers appropriate for use in the NGS-based DNA sequencing were generated, optimized, and used to amplify the individual target sequences for each guide RNA using DNA polymerase. Amplicons were sequenced and analyzed.

[0388] Primary Human Hepatocytes were thawed and resuspended in Hepatocyte Basal Media (HBM) supplemented with HCM SingleQuots and 5% Fetal Bovine Serum (FBS). 50,000 cells were nucleofected with nuclease mRNA along with the indicated chemically-synthesized guide RNA (500 ng mRNA/150 pmol guide per well). 72 hr post-nucleofection, the cells were harvested and the genomic DNA was extracted. PCR primers appropriate for use in the NGS-based DNA sequencing were generated, optimized, and used to amplify the individual target sequences for each guide RNA using DNA polymerase. Amplicons were sequenced and analyzed.

[0389] Hep3B cells were cultured in EMEM + 10% FBS + 1% penstrep at 37°C and 5% CO₂. Once an adequate amount of cells had been cultured, 0.25% Trypsin-EDTA was added to the flasks to harvest the cells. Harvested Hep3B cells were resuspended in buffer at a concentration

of 50,000 cells/20uL. 50,000 cells/well were nucleofected with nuclease mRNA along with the indicated chemically-synthesized guide RNA (500 ng mRNA/150 pmol guide per well). 72 hr post-nucleofection, the cells were harvested and the genomic DNA extracted. The PCR primers appropriate for use in the NGS-based DNA sequencing were generated, optimized, and used to amplify the individual target sequences for each guide RNA using DNA polymerase. Amplicons were sequenced and analyzed.

[0390] **Results**

[0391] MG87-70 (SEQ ID NO: 1107) was screened in K562 cells with 47 guides targeting sites with NNNNGT PAMs in AAVS1 or TRAC at two doses of guide, 150 pmol or 650 pmol. 39 guides showed activity >2% in at least one of the screen conditions (SEQ ID NOs: 1011-1049, **FIG. 18**). The highest observed editing with the 150 pmol dose was 93.8%. Indel rates only increase slightly at the higher dose, suggesting 150 pmol is near but slightly under-saturating editing activity. MG87-21 (SEQ ID NO: 1105) was screened in the K562 cells with 47 guides targeting sites with NNNNCC PAMs in AAVS1 or TRAC with high dose guide conditions only (650 pmol per well instead of 150 pmol per well), showing 6 guides with activity >2% (SEQ ID NOs: 1900-1905, **FIG. 19**). These results demonstrate that the MG87 family of nucleases is active in human cells for targeted genome editing, with MG87-70 in particular being a high efficiency nuclease.

[0392] MG71-2 and MG71-43 (SEQ ID NO: 1103, 1104) were screened for active guides to the intronic region of ATP7B, between exon 5 and exon 6 using an NNNACT PAM for guide design. Of 94 guides screened, 24 guides were identified as active >1% in PHH cells and 26 guides were identified as active >1% in Hep3B cells (SEQ ID NOs: 1056-1082, **FIGs. 20-21**). The two nucleases show similar levels of editing activity at these conditions.

[0393] MG71-2 and MG71-43 (SEQ ID NOs: 1103, 1104) were screened for active guides to the intron 1 region of albumin, using an NNNACT PAM for guide design. Of 19 guides screened, 10 guides were identified as active >1% in PHH cells for one or both nucleases (SEQ ID NOs: 1083-1092, **FIG. 22**). Additionally, 11 guides were designed only for MG71-43 using a less restrictive PAM of NNNRMY. Of these guides, 2 were active in PHH (SEQ ID NOs: 1093-1094, **FIG. 22**).

Example 11 – Metagenomic and ASR nucleases demonstrate high specificity

[0394] ***In cell specificity assay***

[0395] The study utilized an off-target discovery method to identify potential off-targets of the nuclease. Wild type-K562 cells were subcultured in a T75 flask using modified Improved Minimum Essential Medium (IMEM) with 10% fetal bovine serum (FBS) until confluent. On the

day of the experiment, cells were harvested, washed in PBS, and 120K cells per 20uL were resuspended in the Nucleofector Solution (SF buffer). Cells were nucleofected with 500 ng of mRNA, 150pmol of guide, and 50 pmol of dsODN with K562 specific program and allowed to recover for 72 hrs. dsODN was freshly prepared on the day of experiment by annealing (heat and slow cool) equimolar ratios of modified oligos in TE 8 buffer to make a 50 μ M stock solution. Controls include wells that are dsODN only, as well as a condition with only mRNA and guide (minus dsODN well). After 3 days, gDNA was extracted. Target amplicons were amplified, and sent to NGS for sequencing to verify editing efficiency and dsODN incorporation. For detection of off-target editing events, the gDNA was then quantified and 400 ng was used as input for enzymatic fragmentation, end repair, and ligation of adapters and final sequencing library preparation. Fragmented adapter ligated DNA was size selected using magnetic beads and used as input material for the PCR reaction that enriches for fragments containing the dsODN sequence. This PCR step was carried out in the form of two different reactions: one targeting the plus strand and the second targeting the minus strand, utilizing the primers that bind to the dsODN and end ligated adapters, thus enriching for fragments containing the dsODN. The resulting DNA was quantified, normalized, and used for input in an indexing PCR which added Illumina sequencing adapters and as well as sample barcodes to each library. The final libraries were pooled and purified before being quantified and sequenced. Reads were filtered to a max edit distance of 7 and read count of 10 to call off-targets.

[0396] *Results*

[0397] MG71-2 and MG71-43 (SEQ ID NOs: 1103, 1104) were screened with 8 guides that showed editing >80% (SEQ ID NOs: 1095-1102). For MG71-2, no detectable off targets were found for any of the guides. For MG71-43, five guides showed no detectable off-targets in 3 replicates. For two guides, one off target was detected in one out of three replicates. The remaining guide had two reproducible off-target sites in all three replicates, but still >98% of reads mapped to the on-target site (**FIGs. 23A-23C**). These results showed that the MG71 family of proteins is highly specific, both as the WT protein (MG71-2) and after ancestral reconstruction (MG71-43).

[0398] MG87-70 (SEQ ID NO: 1107) was screened with 7 guides that showed editing >80% (SEQ ID NOs: 1023, 1025, 1029, 1030, 1042, 1045, 1049). Four guides showed no detectable off-targets in 3 replicates, showing the high specificity of this nuclease for certain sites. For two guides, one off target was detected in a subset of the replicates, but still had >99% on-target reads. The remaining guide g1_AAVS1_P2B3 had two reproducible off-target sites in all three replicates with overall on target reads <80% of total reads (**FIGs. 24A-24B**).

Example 12 – Improved nucleases demonstrate high editing on therapeutically relevant loci in engineered mammalian cells**[0399] *In-cell testing of nucleases in engineered K562s***

[0400] To test the activity of nucleases in mammalian cells with therapeutic contexts the study used custom-engineered K562 cells. A stable cell line was generated in K562 cells using a lentiviral vector encoding a combination of therapeutically relevant targets (~4.5kb) (SEQ ID NO: 2061) at a low MOI. This 4.5kb sequence contained about 500 nt of various disease targets (PAH, Wilson's disease, and others) with about 250 nt on either side of the therapeutically relevant edit. The mutation that would correspond to the gene of interest would then be changed to the diseased state and allow for the correction of a therapeutically relevant edit. In order to amplify the engineered target of interest over the endogenous target, 20 nt of non-native sequences replaced endogenous gene sequence about 125 nt on the 5' and 3' end of the target of interest provide primer binding sites that would result in a 250 nt amplicon for NGS processing. Transduced cells were selected with puromycin from 3 days to 10 days post-transduction. These stable cell lines are here on referred to as engineered K562 cells.

[0401] Initial testing used a pooled guide approach to streamline experimental design where each MG71 variant received the same set of 10 guides regardless of PAM (guides with a cut site within 10 nt of the target mutation). For pooled testing, engineered K562 cells containing the disease mutation were nucleofected with 500 ng of mRNA and 150 pmol total guide (15 pmol each guide) into 50,000 cells and recovered for three (3) days at 37 °C in a 96-well format. gDNA was extracted, target regions were amplified, and processed for NGS, and percent editing was analyzed. Guides were also tested in individual transfections whereby each well contained 1ug of IVT mRNA of each nuclease and 150 pmol of guide RNA with the same nucleofection protocol as above. gDNA was extracted, target regions were amplified, and processed for NGS. Percent editing was analyzed.

[0402] *Results*

[0403] A subset of nucleases were tested in a pooled format to reduce the number of nucleofections required for screening. For the pooled screens, 10 guides were pooled equimolar and used in transfection (15 pmol each, 150 pmol total). The PAH pool was composed of SEQ ID NOs: 1113-1122. Protein MG71-88 showed activity at the engineered PAH site (**FIGs. 25A**).

Example 13 – Human HAO1 MG21-1 engineered guide RNA activity

[0404] The study analyzed gene-editing outcomes at the DNA level for guide hHAO1-211-F1 (**FIG. 26**). Hep3B cells were nucleofected with 500 ng of MG21-1 mRNA and 3 pmol of each engineered guide (SEQ ID NOs: 1283-1392) (50,000 cells per well in a 96-well plate). Cells

were harvested and genomic DNA extracted 72 hrs post-transfection. PCR primers appropriate for use in NGS-based DNA sequencing were generated, optimized, and used to amplify the target sequence for hHAO1-211-F1. The amplicons were sequenced and analyzed to measure gene editing. Compared to end-modified guide (3x 2'OMe and 3x phosphorothioate bonds at both ends of the gRNA, SEQ ID NO: 1283), changes to the sequence and chemical modifications on the guide produced up to 2-fold higher editing, with SEQ ID NOs: 1290, 1319, 1336, 1360, and 1375 producing the greatest improvement in activity.

Example 14 – Human AAVS1 MG23-1 engineered guide RNA activity

[0405] The study analyzed gene-editing outcomes at the DNA level for guide hAAVS1-231-G1 (**FIG. 27**). Hep3B cells were nucleofected with 500 ng of MG23-1 mRNA and 150 pmol of each engineered guide (SEQ ID NOs: 1393-1485) (50,000 cells per well in a 96-well plate). Cells were harvested and genomic DNA extracted 72 hrs post-transfection. PCR primers appropriate for use in NGS-based DNA sequencing were generated, optimized, and used to amplify the target sequence for hAAVS1-231-G1. The amplicons were sequenced and analyzed to measure gene editing. Compared to end-modified guide (3x 2'OMe and 3x phosphorothioate bonds at both ends of the gRNA, SEQ ID NO: 1393), changes to the sequence and chemical modifications on the guide produced up to 1.7-fold higher editing, with SEQ ID NOs: 1414 and 1441 producing the greatest improvement in activity.

Example 15 – Effect of spacer length on gRNA activity

[0406] The study analyzed gene-editing outcomes at the DNA level for guides of different spacer lengths on 1-2 different chemically modified backbones from four different type II editing systems (MG23-1, MG21-1, and MG71-2) (**FIG. 28**).

[0407] Type II system 23-1 was tested with two spacer sequences (hAAVS1-231-G1 and hTRAC-23-1-H7) across eight spacer lengths (18-25 nt) on one chemically modified backbone pattern. - Type II system 21-1 was tested with one spacer sequence (hHAO1-21-1-F1) across four spacer lengths (22-25 nt) and two different chemically modified backbones (78 and 157). Type II system 71-2 was tested with one spacer (hTRAC-71-2-C2) across eight spacer lengths (18-25 nt) and two different chemically modified backbones (12 and 255).

[0408] Hep3B cells were nucleofected with 500 ng of nuclease mRNA and 65 pmol (231 nuclease), 30 pmol (368 nuclease), 2.5 pmol (211 nuclease), or 7.5 pmol (71-2 nuclease) of each engineered guide (SEQ ID NOs: 1486-1525) (50,000 cells per well in a 96-well plate). Cells were harvested and genomic DNA extracted 72 hrs post- transfection. PCR primers appropriate for use in NGS-based DNA sequencing were generated, optimized, and used to amplify the target sequence for each spacer. The amplicons were sequenced and analyzed to measure gene editing.

Across all four type II systems, 22-23 nt spacers show optimal activity. For systems where shorter spacers were tested (23-1 and 71-2), spacer lengths of 20 nt and below completely lost activity. For spacers longer than 23 nt whether activity was maintained or decreased depended on the type II system or spacer sequence tested.

Example 16 – Guide screen against the human ATXN2 locus in K562 cells

[0409] In this example, guides targeting all exons of the human ATXN2 were tested in an immortalized lymphoblast-derived cell line K562.

[0410] *Methods*

[0411] *Cell culture, transfections, next generation sequencing, and editing analysis*

[0412] K562 cells were grown and passaged in Iscove's Modified Dulbecco's Medium (IMDM) supplemented with 10% (v/v) fetal bovine serum at 37 °C with 5% CO₂. Sequences for MG21-1 mRNA were codon optimized for human expression, then synthesized and cloned into a high copy kanamycin plasmid. Synthesized constructs encoding T7 promoter, UTRs, base editor ORF, and NLS sequences were digested with HindII and BamHI, and ligated into a pUC19 plasmid backbone with T4 DNA ligase and 1x reaction buffer. The complete MG21-1 mRNA plasmid consists of an origin of replication, kanamycin resistance cassette, the synthesized construct, and an encoded polyA tail. MG21-1 mRNA was synthesized via in vitro transcription (IVT) using the linearized mRNA plasmid. This plasmid was linearized by incubation at 37 °C for 4 hours with SpeI enzyme. A standard linearization reaction consists of a 50 uL reaction containing 10 ug pDNA, 40 units SpeI, and 1x reaction buffer. Linearized plasmid was purified and precipitated and resuspended in nuclease free water at an adjusted concentration of 500 ng/uL. The IVT reaction to generate base editor mRNA was performed at 37°C for 4 hr under the following conditions: 1.5 ug linearized plasmid; 4 mM ATP, CTP, GTP, N1-methyl pseudo-UTP, and CleanCap AG 3'OME; HiScribe Master Mix and 1x transcription buffer. After 4 hr, IVT was stopped, and plasmid DNA was digested with the addition of 250 U/mL DNaseI and incubated for 15 min at 37 °C. Purification of base editor mRNA was performed. Transcript concentration was determined by UV and further analyzed by capillary gel electrophoresis.

[0413] Following mRNA production, 1.0E+05 K562 cells were co-nucleofected with 500 ng of the above mRNAs and 150 pmol chemically- synthesized sgRNA targeting all exons of ATXN2. Guide RNA spacer sequence SEQ ID NOs: 1526-1785 and full guide sequence SEQ ID NOs: 1786-2045 were evaluated.

[0414] Cells were grown for 3 days, harvested, and gDNA was extracted. Targeted regions for editing were amplified using DNA polymerase with primers appropriate for use with NGS- based

DNA sequencing, and extracted DNA as the templates. PCR products were purified. Amplicons were sequenced and analyzed to measure gene editing.

[0415] Results

[0416] The study provided that up to 97.34% editing across 260 guides, with 29 guides showed >50% editing (**FIG. 29**). Based on these high editing levels and recent results demonstrating a high correlation between indel formation and functional knockout it is anticipated that several guides and their mouse equivalents will have robust decreases in ATXN2 protein levels in vitro and in vivo. Results can be found in **FIGs. 30-31, 33-35, and 42-44**.

Example 17 – Validation of synthetic human ATXN2 guides in human neural progenitor cells

[0417] In this example, chemically-modified, synthetic guides targeting the human ATXN2 gene were tested in human iPSC-derived neural progenitor cells (NPC). The guides and MG21-1 nuclease mRNA were co-transfected into the NPC cells via lipofection, and indels were measured.

[0418] Methods

[0419] Cell culture, transfections, next generation sequencing, and editing analysis

[0420] NPC cells were cultured and passaged at 37 °C with 5% CO₂. A total of 1.5 x 10⁴ NPC cells were co-transfected with 200 ng of mRNA and chemically-synthesized sgRNA targeting ATXN2 at a molar ratio of 1:5 (mRNA:sgRNA). Guide spacer sequence SEQ ID NOs: 1558, 1579, 1619, 1622, 1628, 1639, 1667, 1733, 1757, and 1757 and full guide sequence SEQ ID NOs: 1818, 1839, 1879, 1882, 1888, 1899, 1927, 1993, 2017, and 2038 were evaluated.

[0421] Cells were harvested after 3 days of culture, and genomic DNA (gDNA) was extracted. The target genomic regions were amplified from the extracted gDNA using DNA polymerase with primers designed for NGS-based sequencing. PCR products were purified. The amplicons were sequenced and analyzed to quantify gene editing efficiency.

[0422] Results

[0423] The study showed that the gene editing efficiencies ranged from 39% to 76% across the 10 MG21-1 guides tested, with the highest efficiency achieved by the P3E8 guide (**FIG. 30**).

Example 18 – Validation of synthetic mouse surrogate ATXN2 guides in mouse Neuro-2a cells

[0424] In this example, human ATXN2 guides from Examples 16 and 17 with up to two mismatches to the mouse genome were converted to the corresponding mouse sequences. These mouse surrogate guides were tested in Neuro-2a cells as synthetic guides. The guides were

transfected into the cells along with the corresponding editing system mRNA using nucleofection, and indels at the *Atxn2* locus were measured.

[0425] Methods

[0426] Cell culture, transfections, next generation sequencing, and editing analysis

[0427] Neuro-2a cells were cultured and passaged in the Dulbecco's Modified Eagle's Medium (supplemented with 10% FBS at 37 °C with 5% CO₂). A total of 1.5 x 10⁵ Neuro-2a cells were transfected using a serial dilution of MG21-1 mRNA and mouse surrogate *Atxn2* guides via nucleofection. The highest concentration in the serial dilution was 500 ng of mRNA and 150 pmol of guide, followed by a 1:3 serial dilution series. Guide spacer sequence SEQ ID NOs: 2046-2050 and full guide sequence SEQ ID NOs: 2051-2055 were evaluated.

[0428] Cells were harvested after three days of culture, and genomic DNA (gDNA) was extracted. The target genomic regions were amplified from the extracted gDNA using DNA polymerase with primers designed for the NGS-based sequencing. The PCR products were purified. The amplicons were sequenced and analyzed to quantify the gene editing efficiency.

[0429] Results

[0430] The study showed that up to 97% gene editing efficiency with the P2G1 guide at the highest concentration (500 ng of mRNA and 150 pmol of guide, **FIG. 31**). Additionally, a dose-dependent increase in the editing was noted across all guides tested.

Example 19 – Validation of editing by mouse surrogate ATXN2 guides expressed from AAV cargo plasmids in mouse Neuro-2a cells

[0431] In this example, the AAV cargo plasmids were synthesized bearing a U6 promoter driving expression of the mouse surrogate guide RNAs created in Example 18, upstream of a CMV promoter driving V5 epitope-tagged MG21-1 bearing a synthetic poly-A (SpA). These plasmids were transfected into the mouse neuro-2a cells, and editing at the *Atxn2* locus was measured as a function of plasmid amount.

[0432] Methods

[0433] All-in-one AAV cargo design

[0434] AAV cargo constructs were designed to express MG21-1 and a guide RNA from a single AAV. The payload was flanked by two, 141 bp inverted terminal repeats (ITRs). A U6 polymerase III promoter was placed upstream of the guide RNA sequence in the forward orientation at the 5' end of the cargo. This was followed by a CMV promoter driving the MG21-1 open reading frame. MG21-1 was N-terminally tagged with a V5 epitope. A synthetic poly-adenylation sequence (SpA) followed MG21-1.

[0435] Cell culture, transfections, next generation sequencing, and editing analysis

[0436] Neuro-2a cells were cultured and passaged in Dulbecco's Modified Eagle's Medium supplemented with 10% FBS at 37 °C with 5% CO₂. A total of 5 x 10⁴ neuro-2a cells were transfected using a serial dilution of the AAV cargo plasmids via lipofection. SEQ ID NOs: 2056-2060 were evaluated. The highest plasmid concentration in the dilution series was 2 ug, followed by a 1:2 serial dilution.

[0437] Cells were harvested after three days of culture, and genomic DNA (gDNA) was extracted. The target genomic regions were amplified from the extracted gDNA using DNA polymerase with the primers designed for the NGS-based sequencing. The PCR products were purified. The amplicons were sequenced and analyzed to quantify gene editing efficiency.

[0438] ***Results***

[0439] Using the AAV cargo plasmids (**FIG. 32**), in this study were observed up to 21% gene editing efficiency with the P1F12 guide at a plasmid concentration of 250 ng (**FIG. 33**). The P2F6 guide exhibited the lowest editing efficiency (2-3%), while the other three guides demonstrated intermediate levels of the gene editing. Notably, none of the guides reached an editing plateau.

Example 20 – Validation of editing by mouse surrogate ATXN2 guides expressed from AAV cargo plasmids in mouse Neuro-2a cells

[0440] In this example, the same AAV cargo plasmids used in Example 19 were examined in an additional experimental repeat, employing the lower plasmid concentrations to maximize editing efficiency. These plasmids were transfected into the mouse neuro-2a cells, and editing at the *Atxn2* locus was measured as a function of plasmid amount.

[0441] **All-in-one AAV cargo design**

[0442] AAV cargo constructs were designed to express MG21-1 and a guide RNA from a single AAV. The payload was flanked by two, 141 bp inverted terminal repeats (ITRs). A U6 polymerase III promoter was placed upstream of the guide RNA sequence in the forward orientation at the 5' end of the cargo. This was followed by a CMV promoter driving the MG21-1 open reading frame. MG21-1 was N-terminally tagged with a V5 epitope. A synthetic poly-adenylation sequence (SpA) followed MG21-1.

[0443] **Cell culture, transfections, next generation sequencing, and editing analysis**

[0444] Neuro-2a cells were cultured and passaged in the Dulbecco's Modified Eagle's Medium supplemented with 10% FBS at 37 °C with 5% CO₂. A total of 5 x 10⁴ Neuro-2a cells were transfected using a serial dilution of AAV cargo plasmids via lipofection. SEQ ID NOs: 2056-2060 were evaluated. The highest plasmid concentration in the dilution series was 2 ug, followed by a 1:2 serial dilution.

[0445] Cells were harvested after 3 days of culture, and genomic DNA (gDNA) was extracted. The target genomic regions were amplified from the extracted gDNA using DNA polymerase with primers designed for the NGS-based sequencing. The PCR products were purified. The amplicons were sequenced and analyzed to quantify the gene editing efficiency.

[0446] **Results**

[0447] The study showed that up to 32% gene editing efficiency with the P1F12 guide at a plasmid concentration of 250 ng (**FIG. 34**). The P2F6 guide exhibited the lowest editing efficiency (1-5%), while the other three guides demonstrated intermediate levels of gene editing.

Example 21– Knockdown of mouse ATXN2 protein in Neuro-2a cells transfected with AAV cargo plasmids

[0448] In this example, the AAV cargo plasmids bearing V5-tagged MG21-1 and mouse *Atnx2* guide RNAs under the control of promoters were transfected into mouse Neuro-2a cells. The ability of these plasmids to knockdown ATXN2 protein levels was measured by semi-quantitative Western blot.

[0449] **Methods**

[0450] Neuro-2a cells were cultured and passaged in the Dulbecco's Modified Eagle's Medium supplemented with 10% FBS at 37 °C with 5% CO₂. For transfection, a total of 2 x 10⁵ Neuro-2a cells were transfected with each of five AAV plasmids (1 µg of DNA) containing different mouse surrogate *Atnx2* guides. SEQ ID NOs: 2056-2060 were evaluated. Three days after transfection, the cells were lysed in the RIPA buffer containing protease and phosphatase inhibitor cocktail. Of the cell lysate, 10 µg was loaded onto a 4-20% Tris-Glycine gel and transferred to the nitrocellulose membrane. Blots were blocked with 5% milk, then probed with the antibodies against the V5 epitope (1:1000) and Ataxin2 (1:1000), detected using the fluorescent secondary antibody AF647-anti-mouse IgG (1:10000) and anti-rabbit IgG-IR800 (1:10,000) respectively. The blots were reprobed with actin antibody (1:2000) as a reference control.

[0451] **Results**

[0452] The highest knockdown level of mouse ATXN2 protein was 62% achieved by an AAV plasmid expressing guide P1F12 (**FIG. 35**). Protein knockdown of 44%, 41%, 38%, and 20% were achieved with AAV plasmids expressing guides P1A5, P2A1, P2G1, and P2F6, respectively.

Example 22– Packaging of full-length AAV cargo sequences into AAV9 as assessed by alkaline DNA gel electrophoresis

[0453] In this example, the packaging of full-length AAV cargos into AAV9 virions is demonstrated using alkaline DNA gel electrophoresis.

[0454] *Methods*

[0455] *AAV production*

[0456] AAVs were produced using a transient triple co-transfection method involving a plasmid encoding adenoviral helper proteins, a plasmid encoding AAV rep protein and capsid proteins, and a plasmid encoding guide and MG21-1 flanked between two AAV2 ITRs. SEQ ID NOs: 2056-2060 were evaluated. One day prior to transfection, Viral Production Cells 2.0 were seeded at 1.5e6 cells/mL in BalanCD HEK293 medium supplemented with glutamine at 1:100 dilution. On the day of transfection, the three plasmids were mixed at a 1:1:1 mass ratio with a total plasmid mass of 2 µg per mL of cell culture volume. The plasmid mixture was diluted in 100 µL of 1X PBS per mL of cell culture volume. To make the final transfection mixture, 1 µL of RevIT AAV enhancer and 2 µL of TransIT Transfection Reagent per mL of cell culture volume were mixed into the diluted plasmid mixture, and incubated at room temperature for 30 minutes. The entire transfection mixture was added into the cells and incubated at 37 °C for 72 hrs prior to harvest. The cells were harvested by centrifuging at 1500 g for 15 minutes at 4 °C and resuspended in a TMS buffer containing 50 mM of Tris-HCl, 2 mM MgCl₂, and 150 mM NaCl. The cell pellets in the TMS buffer were subjected to physical lysis through 3 rapid freeze and thaw cycles, resulting in a crude viral lysate.

[0457] *AAV Purification*

[0458] The crude viral lysate was treated with 100 U/mL of DENARASE® for 1 hr to digest exogenous nucleic acids. The treated viral lysate was clarified by centrifugation at 12000 g for 60 minutes at 4 °C in two 30-minute steps, with the supernatant retained after each centrifugation step. The clarified viral lysate was further purified via cesium chloride density gradient ultracentrifugation. To prepare the gradient, 5 mL of cesium chloride in 1X DPBS at a density of 1.3 g/cm³ was added to the bottom of a 17 mL open-top ultracentrifuge tube 5 mL of cesium chloride in 1X DPBS at a density of 1.5 g/cm³ was then slowly underlaid below the cesium chloride at 1.3 g/cm³ density, and 7 mL of the clarified viral lysate was slowly overlaid on top of the cesium chloride at 1.3 g/cm³ density. The ultracentrifuge tubes were loaded and centrifuged at 32000 rpm for 6 hrs at room temperature. To retrieve AAV particles from the gradients, an 18G needle in a 5 mL syringe was punctured into the side of the ultracentrifuge tube at the interface of the two different cesium chloride densities, and 2.5 mL of the sedimented AAV was pulled. The AAV underwent 3 rounds of dialysis in a solution of 0.001% Pluronic F-68 in 1X DPBS with calcium and magnesium at pH 7.10. The AAV suspension was finally sterilized with a 0.45 µm syringe filter.

[0459] Alkaline gel electrophoresis

[0460] The alkaline gel was made by dissolving 1% agarose (w/v) in a solution of 30 mM NaCl and 2 mM EDTA in water, microwaved until a homogenous solution was formed, and then casted onto a gel casting tray. The solidified gel was equilibrated in a running buffer containing 30 mM NaOH and 2 mM EDTA for 30 minutes. The AAV samples were prepared by loading 2e10 vector genomes in 25 μ L of 0.001% Pluronic F-68 in 1X DPBS with calcium and magnesium at pH 7.10 and 5 μ L of 6X alkaline loading dye. The solution was incubated at 95 °C for 5 minutes then cooled to 4 °C for 3 minutes. The prepared samples were loaded onto the gel and ran at 50V for 210 minutes. Once the run had completed, the gel was placed in a neutralization solution consisting of 0.5 M Tris-HCl and 1 mM EDTA and incubated at room temperature for 1 hr on a rocking shaker. The neutralization solution was removed and replaced with a fresh neutralization solution and SYBR gold stain, and the gel was incubated for 1 hr at room temperature on a rocking shaker. The gel was imaged.

[0461] Results

[0462] The size of band detected by alkaline gel electrophoresis for each AAV corresponds to the expected cargo size (**FIG. 36**). This suggests that packaging of the full-length cargo construct occurred.

Example 23– AAV9 virus particles bearing the MG21-1 editing system show effective transduction when injected into mouse olfactory bulb and cortex

[0463] In this example, the AAV9 virus preparations shown in Example 22 were injected into the mouse olfactory bulb and motor cortex. The tissues were analyzed for the presence of viral genome copies to confirm transduction.

[0464] Methods**[0465] Stereotactic injections of AAV9 and preparation of CNS tissue**

[0466] 6–8-week-old C57BL/6J mice were injected with AAV9 prepared in-house into both the olfactory bulb and M1 region of the motor cortex using a custom built stereotactic delivery system based on the RWD automated stereotaxic instrument (**FIG. 37**). M1 injections consisted of a bilateral infusion of 750 nL of AAV9 infused at a rate of 5nl/second at coordinates +1.00 AP, +/-1.50 ML, -1.50 DV relative to Bregma. Olfactory bulb injections consisted of a bilateral infusion of 250 nL of AAV9 infused at a rate of 5 nL/second and three target depths, coordinates +4.00 AP, +/- 0.70 ML, -2.5/-2.0/-1.5 DV relative to Bregma. A retention time of 5 minutes post injection was added to enhance absorption of the AAV into the tissue. Animals were sutured and allowed to recover until transferred back to their home cages. 2 weeks post injection, animals were euthanized via deep inhalation of isoflurane. Whole brains were dissected fresh, the

olfactory bulb and the M1 regions were further dissected and processed for downstream nucleic acid extraction. Genomic DNA and mRNA was extracted from fresh tissue.

[0467] Genomic DNA ddPCR and data analysis

[0468] Purified genomic DNA (gDNA) was quantified. To determine the tissue titer of injected AAVs, 5 ng of gDNA was added to a ddPCR reaction containing custom probes targeting the central region of the MG21-1 sequence multiplexed with a premade assay targeting the *CycI* gene. Data were graphed. MG21-1 copies were normalized by *CycI* reference copies to produce a normalized MG21-1 copies/cell value (**FIG. 38**). All individual data points are shown with the error bars denoting the average +/- standard error of the mean.

[0469] *Results*

[0470] Viral genome copies per cell varied across AAVs and injection sites, ranging from 2 to 63 copies per cell with an average of 8 copies per cell for olfactory bulb and 9 copies per cell for M1 motor cortex injections. The highest copy numbers were seen for P1F12 AAVs in the olfactory bulb, and P2G1 AAVs in the cortex.

Example 24— MG21-1 transcripts are detected in tissues injected with virus

[0471] In this example, the AAV9 virus preparations shown in Example 22 were injected into the mouse olfactory bulb and motor cortex. The tissues were analyzed for the presence of MG21-1 mRNA by ddPCR in order to confirm expression of the editing system.

[0472] *Methods*

[0473] Stereotactic injections of AAV9 and preparation of CNS tissue

[0474] The methods are similar to as provided in Example 23.

[0475] RNA ddPCR and data analysis

[0476] Purified mRNA was quantified. The mRNA was then used to produce cDNA by reverse transcription. To quantify the expression of AAVs in this study were loaded 5 ng of cDNA mRNA equivalent into a ddPCR reaction containing custom primers and probes targeting the central region of the MG21-1 sequence. This assay was additionally multiplexed targeting the *CycI* gene. Data were graphed. MG21-1 copies were normalized by *CycI* reference copies to produce a normalized MG21-1 copies/*CycI* copies value (**FIG. 39**). All individual data points are shown with the error bars denoting the average +/- standard error of the mean.

[0477] *Results*

[0478] MG21-1 mRNA copies varied across AAVs and injection sites. The highest mRNA expression was seen for P1F12 AAV in the olfactory bulb, and P2G1 AAV in the cortex, consistent with the viral copy numbers obtained in **FIG. 38**.

Example 25– Evaluation of ATXN2 mRNA knockdown in tissues injected with MG21-1 AAVs

[0479] In this example, the AAV9 virus preparations shown in Example 22 were injected into the mouse olfactory bulb and motor cortex. The tissues were analyzed for the level of *Atxn2* mRNA by ddPCR in order to gauge knockdown of target gene expression.

[0480] **Methods**

[0481] Stereotactic injections of AAV9 and preparation of CNS tissue

[0482] The methods are similar to as provided in Example 23.

[0483] RNA ddPCR and data analysis

[0484] Purified mRNA was quantified. The mRNA was then used to produce cDNA by reverse transcription. To quantify the expression of *Atxn2* in this study were loaded 5 ng of cDNA mRNA equivalent to a ddPCR reaction containing an *Atxn2* premade expression quantification assay. This assay was additionally multiplexed with a premade assay targeting the *Cyc1* gene. Data were graphed in GraphPad Prism 10 (**FIG. 40**). *Atxn2* mRNA copies were normalized to *Cyc1* reference copies to produce a normalized *Atxn2* copies/*Cyc1* copies value. All individual data points are shown with the error bars denoting the average +/- standard error of the mean.

[0485] **Results**

[0486] *Atxn2* mRNA copies showed the greatest reduction in average mRNA expression in olfactory bulb samples injected with P1F12 AAV, and in cortex samples receiving the P2G1 AAV, consistent with the viral copy numbers and MG21-1 mRNA expression data obtained in **FIG. 38** and **FIG. 39**, respectively.

Example 26– Indels within the ATXN2 gene are detectable in response to MG21-1 AAV delivery

[0487] In this example, tissues harvested from the stereotactic injection of MG21-1 AAVs were analyzed for indels within the mouse *Atxn2* gene.

[0488] **Methods**

[0489] Stereotactic injections of AAV9 and preparation of CNS tissue

[0490] The methods are similar to as provided in Example 23.

[0491] Indel sequencing and data analysis

[0492] Purified genomic DNA (gDNA) was quantified. The gDNA was then used as a template for a PCR reaction to produce amplicons for next-generation sequencing (NGS). The reaction used primers to specifically amplify a region surrounding the targeted cutsite. Amplicons were then sent for NGS using 300x1bp reads. Raw reads were filtered and aligned to a reference genome using a custom script. Insertions and deletions (indels) were quantified using a custom

script and graphed as a percentage of all filtered reads. Data were graphed. All individual data points are shown with the error bars denoting the average +/- standard error of the mean.

[0493] Results

[0494] Indels within the mouse *Atxn2* gene were highest in olfactory bulb samples injected with P1F12 AAV, and in cortex samples receiving the P2G1 AAV, consistent with the viral copy numbers, MG21-1 mRNA expression data, and *Atxn2* transcript data obtained in **FIGs. 38-41**.

Example 27– Qualification of ATXN2 protein expression in the mouse olfactory bulb after injection of AAV9 viruses expressing the MG21-1 system

[0495] In this example, tissues harvested from the stereotactic injection of MG21-1 AAVs were analyzed for indels within the mouse *Atxn2* gene.

[0496] Methods

[0497] Stereotactic injections of AAV9 and preparation of CNS tissue

[0498] The methods are similar to as provided in Example 23. For protein extraction, Whole brains were dissected fresh and immediately frozen on dry ice, the olfactory bulb and the M1 regions were further dissected and processed for downstream protein extraction. Protein was extracted by homogenizing frozen tissue in 2.5 mL tubes with ceramic beads in lysis buffer (1X HALT + 1x PMSF in RIPA buffer) on a tissue homogenizer to create a homogenate. The homogenate was then clarified by centrifugation at max speed at 4 °C for 10 minutes. The clarified homogenate was then used in a Pierce BCA protein estimation assay.

[0499] Western blot and data analysis

[0500] 20 µg of clarified protein was loaded onto a tris-glycine reducing gel, transferred to a nitrocellulose membrane and blotted using antibodies against *Atxn2* and Actinb. Membranes were then blotted with secondary fluorescent antibodies and imaged on a fluorescent reader. *Atxn2* signal was quantified using background subtraction and then normalized to ACTB background subtracted signal. Data were graphed. *Atxn2* protein was normalized to ACTB protein to produce a normalized *Atxn2*/ACTB copies value (**FIG. 42**). All individual data points are shown with the error bars denoting the average +/- standard error of the mean.

[0501] Results

[0502] Average *Atxn2* protein levels appeared to be reduced in olfactory bulb samples injected with the P1F12 guide-expressing virus.

Example 28– Imaging of an olfactory bulb injected with a mixture of MG21-1 P1F12 and eGFP AAVs shows localized knockdown of ATXN2 protein

[0503] In this example, an AAV expressing MG21-1 with the P1F12 guide and an AAV expressing eGFP were co-injected into the mouse olfactory bulb. The tissue was sectioned and

Atxn2 protein visualized by immunofluorescence to detect knockdown. V5 and eGFP expression were used to delineate the injected area.

[0504] *Methods*

[0505] Stereotactic injections of AAV9 and preparation of CNS tissue

[0506] 6–8-week-old C57BL/6J mice were injected with an AAV9 mixture consisting of MG21-1 P1F12 and U1a-EGFP prepared in house into the olfactory bulb using a custom built stereotactic delivery system based on the RWD automated stereotaxic instrument. Olfactory bulb injections consisted of a bilateral infusion of 250 nL of AAV9 infused at a rate of 5nL/second and three target depths, coordinates +4.00 AP, +/- 0.70 ML, -2.5/-2.0/-1.5 DV relative to Bregma. A retention time of 5 minutes post injection was added to enhance absorption of the AAV into the tissue. Animals were sutured and allowed to recover until transferred back to their home cages. 2 weeks post injection, animals were euthanized via deep inhalation of isoflurane. The animals were first perfused transcardially with 10 mL PBS followed by 10 mL 4% paraformaldehyde. Whole brains were dissected and then cryoprotected in a solution of 30% sucrose prepared in PBS.

[0507] Slicing, immunohistochemistry, and imaging

[0508] Cryoprotected whole brains were frozen in OCT, mounted and then sliced into 40 µm coronal sections using a Leica cryostat. Free-floating sections were blocked, permeabilized and then stained overnight at 4 °C with antibodies against *Atxn2* (1:100) V5 (1:200) and GFP (1:1000). Stained sections were washed 3 times with 0.01% PBST and then incubated at room temperature for 2 hrs for secondary staining using Alexa-Fluor secondary antibodies. Finally, sections were washed 3 more times with 0.01% PBST and mounted on glass slides sealed with Fluoromount-G. Mounted slides were imaged. 40 µm Z-stacks were acquired by manually setting laser power and sensor gain to maximize signal-to-noise ratio for each channel. Data was then reconstructed for visualization and exported (**FIG. 43**). Scale bars represent 70 µm.

[0509] *Results*

[0510] *Atxn2* was seen to be expressed in a prominent ring within the olfactory bulb. A loss in *Atxn2* expression appeared to correlate with the injected mixture of AAVs, as inferred by V5 and eGFP expression. This result suggested that the AAV expressing MG21-1 and guide P1F12 was capable of knocking down *Atxn2* protein expression *in vivo*.

Example 29– Stereotactic injection of an AAV encoding MG21-1_P1F12 guide shows ATXN2 protein knockdown by imaging

[0511] In this example, an AAV expressing MG21-1 with the P1F12 guide and an AAV expressing eGFP were co-injected into the mouse olfactory bulb. The tissue was sectioned and

Atxn2 protein visualized by immunofluorescence to detect knockdown. V5 and eGFP expression were used to delineate the injected area. This was a higher magnification of the images shown in Example 28.

[0512] *Methods*

[0513] *Stereotactic injections of AAV9 and preparation of CNS tissue*

[0514] 6-8 week old C57BL/6J mice were injected with an AAV9 mixture consisting of MG21-1 P1F12 and U1a-EGFP prepared in house into the olfactory bulb using a custom built stereotactic delivery system based on the RWD automated stereotaxic instrument. Olfactory bulb injections consisted of a bilateral infusion of 250 nL of AAV9 infused at a rate of 5 nL/second and three target depths, coordinates +4.00 AP, +/- 0.70 ML, -2.5/-2.0/-1.5 DV relative to Bregma. A retention time of 5 minutes post injection was added to enhance absorption of the AAV into the tissue. Animals were sutured and allowed to recover until transferred back to their home cages. 2 weeks post injection, animals were euthanized via deep inhalation of isoflurane. The animals were first perfused transcardially with 10 mL PBS followed by 10 mL 4% paraformaldehyde. Whole brains were dissected and then cryoprotected in a solution of 30% sucrose prepared in PBS.

[0515] *Slicing, immunohistochemistry, and imaging*

[0516] Cryoprotected whole brains were frozen in OCT, mounted and then sliced into 40 μ m coronal sections using a Leica cryostat. Free-floating sections were blocked, permeabilized and then stained overnight at 4 °C with antibodies against *Atxn2* (1:100) V5 (1:200) and GFP (1:1000). Stained sections were washed 3 times with 0.01% PBST and then incubated at room temperature for 2 hrs for secondary staining using Alexa-Fluor secondary antibodies. Finally, sections were washed 3 more times with 0.01% PBST and mounted on glass slides sealed with Fluoromount-G. Mounted slides were imaged on at 20X magnification. 40 μ m Z-stacks were acquired by manually setting laser power and sensor gain to maximize signal-to-noise ratio for each channel. Data was then reconstructed for visualization and exported (**FIG. 44**). Scale bars represent 70 μ m.

[0517] *Results*

[0518] This Example provides a higher magnification view of the samples presented in Example 28. *Atxn2* was seen to be expressed in a prominent ring within the olfactory bulb. A loss in *Atxn2* expression appeared to correlate with the injected mixture of AAVs, as inferred by V5 and eGFP expression. This result suggests that the AAV expressing MG21-1 and guide P1F12 was capable of knocking down *Atxn2* protein expression *in vivo*.

Example 30 - Mismatch tolerance evaluation of Type II nuclease

[0519] Summary

[0520] The specificity of CRISPR systems depends not only on the sequences of guide RNA (gRNA) and target DNA, but also the Cas enzyme that is being used. Lentivirus was used to deliver a library of dual-targets oligos, one containing a single mismatch in the spacer or PAM compared to the on-target, which are then edited in 96-well format by nucleofection with mRNA or RNP with chemically modified guides. The mismatch tolerance, obtained from the ratio of off:on target editing for each dual target oligo, was plotted by guide and highlights the guide-intrinsic mismatch tolerance observed. When the mean and 95% confidence intervals were plotted across all guides for a given nuclease, by position and by mismatch type, a nuclease-specific profile was observed. Here, data on MG71-2 and MG21-1 nucleases are shown.

[0521] Methods

[0522] A 96-well mismatch tolerance assay was developed (**FIG. 45**) in which cells were transduced with a library of dual-targets, one containing a single mismatch in the spacer or PAM compared to the on-target and edited by nucleofection to directly compare the editing efficiencies observed at both mismatched and on-target sequences using chemically modified guides.

[0523] Library Design

[0524] Single guide RNAs that represented a range of GC contents and percent indels were chosen for use in this assay (**FIGs. 46A** and **46B**). An oligonucleotide library was generated from 49 guides, 21 guides for MG21-1 (SEQ ID NOs: 2062-2082) and 28 guides for MG71-2 (SEQ ID NOs: 2083-2110) to investigate their on- and off-target effects. Each library member contained two targets: an on-target sequence and an off-target sequence with mismatches systematically introduced at every position within both the protospacer adjacent motif (PAM) and the spacer region.

[0525] In addition, control oligonucleotides were generated to simulate four possible editing events: cleavage at the on-target, cleavage at the off-target, cleavage at both targets, and no cleavage.

[0526] The library construct, reference structure shown in (**FIG. 47**), contains of the following components ordered left to right: a 20-bp left primer adapter sequence, an 18-bp barcode 1, a 24-32 bp off-target sequence, a 28 bp linker segregating the two target sequences, a 24-32 bp on-target sequence, an 18-bp barcode 2, and an 20-bp right primer adapter sequence.

[0527] The lentilibrary transduced cells (1×10^6) were edited via nucleofection of either mRNA and guide (2500 ng mRNA, 1000 pmol guide) or RNP (300 pmol of protein complexed with 500 pmol of sgRNA).

[0528] Cells were harvested and gDNA extracted 72 hours post-transfection. 5 ug of gDNA was amplified with primers containing adapters for sequencing for 20 cycles.

[0529] Editing efficiency was computed based on two quantification windows: one around the off-target sequence and one around the on-target sequence. Groups with fewer than 100 reads, less than 1% on-target editing, or no editing at either target were discarded from downstream analysis. A final single mismatch tolerance score was calculated for each group as the ratio of off-target to on-target editing. Single mismatch tolerance scores above 0.5 were considered to be highly tolerant, while those below 0.5 were considered to have low tolerance. PAM preference scores were also calculated by converting these mismatch tolerance scores and their respective off-target PAMs into information-based representation matrices using logomaker.

[0530] To further visualize the data, heatmaps were created to illustrate the protospacer base preference at each mismatch position for different nucleases. The heatmaps were generated by grouping the data points by nuclease, mismatch location, and mismatch base, and then calculating the mean off-on ratio and 95% confidence interval for each group. The resulting data was pivoted to create a matrix format suitable for heatmap visualization via seaborn, with color coding to indicate the mismatch tolerance at each base and position.

[0531] *Results*

[0532] Editing controls were generated to characterize the editing outcomes of a dual-target library. The editing at left or right end was turned “OFF” by modifying the native PAM sequence (“ON”) associated with each on-target sequence. Editing observed across the four types of editing controls, no editing (OFF/OFF), left-end editing (OFF/ON, right-end editing, and both-end editing (ON/ON) were plotted in (FIG. 48) and show editing can be controlled through PAM inactivation and indels correctly assigned to the left or right end target.

[0533] Raw indel percentages at mismatched targets compared to on-target library members were plotted by guide to compare the ranges of editing levels observed. The raw indel percentages were normalized by taking the ratio of off: on target editing, termed mismatch tolerance, observed by position for each guide. A nuclease-specific single mismatch tolerance profile was generated by plotting the mean and 95% confidence intervals of single mismatch tolerance, by position, of all guides for a given nuclease (FIG. 49). The guides included in the library for MG71-2 contained spacers of two different lengths. FIG. 50 shows the single mismatch tolerance profile observed with 25 MG71-2 guides of 22 nucleotide length targeting human B2M, TRAC, and AAVS1 loci. FIG. 51 shows the single mismatch tolerance profile observed with 7 MG71-2 guides of 24-nucleotide length targeting the HBB locus, and when the two MG71-2 single mismatch tolerance profiles are overlaid on top of one another they show a high degree of similarity (FIG. 52). The single mismatch tolerance profile observed from 21 guides of 22 nucleotide lengths targeting human PDCD1, HAO1, IL17, TRAC loci, is shown in (FIG. 53). Both Type II enzymes, MG71-2 and MG21-1, demonstrated lowest tolerance for

single nucleotide mismatches at PAM proximal sites, increasing in tolerance moving distally from PAM.

[0534] PAM preference scores were calculated by converting mismatch tolerance scores and their respective off-target PAMs into information-based representation matrices using logomaker. The PAM preferences for MG71-2 across both 22nt and 24nt spacers are represented as seqlogos in **FIG. 54** depicting a consensus PAM of nnnACTnn. The PAM preferences for MG21-1 across is represented as a seqlogo in **FIG. 55** depicting a consensus PAM of RnGMAR.

[0535] Heatmaps were created to illustrate the protospacer base preference at each mismatched position for different nucleases. Due to the fixed nucleotide nature of certain nuclease PAMs, not all mismatches are represented and are thus represented by white blank squares. For example, if all PAM sequences for a given position are a C, there would be no PAM mismatches to a C (and thus because all protospacers contain a G, only A, C, and T mismatches would be displayed).

FIG. 56 depicts the protospacer base preferences for MG71-2, and demonstrates slightly different protospacer base preferences for mismatches depending on the spacer length. **FIG. 57** depicts the protospacer base preferences for MG21-1.

References

Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).

Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE* **5**, e9490 (2010).

Yang, Z. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).

[0536] While preferred embodiments of the present disclosure 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 disclosure be limited by the specific examples provided within the specification. While the disclosure 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 disclosure. Furthermore, it shall be understood that all aspects of the disclosure 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 disclosure described herein may be employed in practicing the disclosure. It is therefore contemplated that the disclosure shall also cover any such alternatives, modifications, variations

or equivalents. It is intended that the following claims define the scope of the disclosure and that methods and structures within the scope of these claims and their equivalents be covered thereby.

Table 11. Spacer / Target Site Information of sgRNAs

Cat.	SEQ ID NO:	Description	Spacer/ Target Site Information
MG87-70 sgRNA targeting human TRAC	1011	87-70_sgRNA_P1B1	Target site SEQ ID NO: 1161
MG87-70 sgRNA targeting human TRAC	1012	87-70_sgRNA_P1C1	Target site SEQ ID NO: 1162
MG87-70 sgRNA targeting human TRAC	1013	87-70_sgRNA_P1D1	Target site SEQ ID NO: 1163
MG87-70 sgRNA targeting human TRAC	1014	87-70_sgRNA_P1G1	Target site SEQ ID NO: 1164
MG87-70 sgRNA targeting human TRAC	1015	87-70_sgRNA_P1H1	Target site SEQ ID NO: 1165
MG87-70 sgRNA targeting human TRAC	1016	87-70_sgRNA_P1A2	Target site SEQ ID NO: 1166
MG87-70 sgRNA targeting human TRAC	1017	87-70_sgRNA_P1C2	Target site SEQ ID NO: 1167
MG87-70 sgRNA targeting human TRAC	1018	87-70_sgRNA_P1D2	Target site SEQ ID NO: 1168
MG87-70 sgRNA targeting human TRAC	1019	87-70_sgRNA_P1F2	Target site SEQ ID NO: 1169
MG87-70 sgRNA targeting human AAVS1	1020	87-70_sgRNA_P1G2	Target site SEQ ID NO: 1170
MG87-70 sgRNA targeting human AAVS1	1021	87-70_sgRNA_P1H2	Target site SEQ ID NO: 1171
MG87-70 sgRNA targeting human AAVS1	1022	87-70_sgRNA_P1A3	Target site SEQ ID NO: 1172
MG87-70 sgRNA targeting human AAVS1	1023	87-70_sgRNA_P1B3	Target site SEQ ID NO: 1173
MG87-70 sgRNA targeting human AAVS1	1024	87-70_sgRNA_P1C3	Target site SEQ ID NO: 1174
MG87-70 sgRNA targeting human AAVS1	1025	87-70_sgRNA_P1D3	Target site SEQ ID NO: 1175
MG87-70 sgRNA targeting human AAVS1	1026	87-70_sgRNA_P1G3	Target site SEQ ID NO: 1176
MG87-70 sgRNA targeting human AAVS1	1027	87-70_sgRNA_P1H3	Target site SEQ ID NO: 1177
MG87-70 sgRNA targeting human AAVS1	1028	87-70_sgRNA_P1A4	Target site SEQ ID NO: 1178

Cat.	SEQ ID NO:	Description	Spacer/ Target Site Information
MG87-70 sgRNA targeting human AAVS1	1029	87-70_sgRNA_P1B4	Target site SEQ ID NO: 1179
MG87-70 sgRNA targeting human AAVS1	1030	87-70_sgRNA_P1C4	Target site SEQ ID NO: 1180
MG87-70 sgRNA targeting human AAVS1	1031	87-70_sgRNA_P1D4	Target site SEQ ID NO: 1181
MG87-70 sgRNA targeting human AAVS1	1032	87-70_sgRNA_P1E4	Target site SEQ ID NO: 1182
MG87-70 sgRNA targeting human AAVS1	1033	87-70_sgRNA_P1F4	Target site SEQ ID NO: 1183
MG87-70 sgRNA targeting human AAVS1	1034	87-70_sgRNA_P1G4	Target site SEQ ID NO: 1184
MG87-70 sgRNA targeting human AAVS1	1035	87-70_sgRNA_P1H4	Target site SEQ ID NO: 1185
MG87-70 sgRNA targeting human AAVS1	1036	87-70_sgRNA_P1B5	Target site SEQ ID NO: 1186
MG87-70 sgRNA targeting human AAVS1	1037	87-70_sgRNA_P1C5	Target site SEQ ID NO: 1187
MG87-70 sgRNA targeting human AAVS1	1038	87-70_sgRNA_P1D5	Target site SEQ ID NO: 1188
MG87-70 sgRNA targeting human AAVS1	1039	87-70_sgRNA_P1E5	Target site SEQ ID NO: 1189
MG87-70 sgRNA targeting human AAVS1	1040	87-70_sgRNA_P1F5	Target site SEQ ID NO: 1190
MG87-70 sgRNA targeting human AAVS1	1041	87-70_sgRNA_P1G5	Target site SEQ ID NO: 1191
MG87-70 sgRNA targeting human AAVS1	1042	87-70_sgRNA_P1H5	Target site SEQ ID NO: 1192
MG87-70 sgRNA targeting human AAVS1	1043	87-70_sgRNA_P1A6	Target site SEQ ID NO: 1193
MG87-70 sgRNA targeting human AAVS1	1044	87-70_sgRNA_P1B6	Target site SEQ ID NO: 1194
MG87-70 sgRNA targeting human AAVS1	1045	87-70_sgRNA_P1C6	Target site SEQ ID NO: 1195
MG87-70 sgRNA targeting human AAVS1	1046	87-70_sgRNA_P1D6	Target site SEQ ID NO: 1196
MG87-70 sgRNA targeting human AAVS1	1047	87-70_sgRNA_P1E6	Target site SEQ ID NO: 1197
MG87-70 sgRNA targeting human AAVS1	1048	87-70_sgRNA_P1F6	Target site SEQ ID NO: 1198

Cat.	SEQ ID NO:	Description	Spacer/ Target Site Information
MG87-70 sgRNA targeting human AAVS1	1049	87-70_sgRNA_P1G6	Target site SEQ ID NO: 1199
MG87-21 sgRNA targeting human TRAC	1050	87-21_sgRNA_P1B1	Target site SEQ ID NO: 1200
MG87-21 sgRNA targeting human AAVS1	1051	87-21_sgRNA_P1F4	Target site SEQ ID NO: 1201
MG87-21 sgRNA targeting human AAVS1	1052	87-21_sgRNA_P1G5	Target site SEQ ID NO: 1202
MG87-21 sgRNA targeting human AAVS1	1053	87-21_sgRNA_P1A6	Target site SEQ ID NO: 1203
MG87-21 sgRNA targeting human AAVS1	1054	87-21_sgRNA_P1C6	Target site SEQ ID NO: 1204
MG87-21 sgRNA targeting human AAVS1	1055	87-21_sgRNA_P1G6	Target site SEQ ID NO: 1205
MG71-2 sgRNA targeting ATP7B	1056	71-2_ATP7B_P1G1	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1206
MG71-2 sgRNA targeting ATP7B	1057	71-2_ATP7B_P1G2	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1207
MG71-2 sgRNA targeting ATP7B	1058	71-2_ATP7B_P1C3	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1208
MG71-2 sgRNA targeting ATP7B	1059	71-2_ATP7B_P1H3	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1209
MG71-2 sgRNA targeting ATP7B	1060	71-2_ATP7B_P1A4	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1210
MG71-2 sgRNA targeting ATP7B	1061	71-2_ATP7B_P1C4	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1211
MG71-2 sgRNA targeting ATP7B	1062	71-2_ATP7B_P1D4	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1212
MG71-2 sgRNA targeting ATP7B	1063	71-2_ATP7B_P1B5	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1213
MG71-2 sgRNA targeting ATP7B	1064	71-2_ATP7B_P1C5	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1214
MG71-2 sgRNA targeting ATP7B	1065	71-2_ATP7B_P1D5	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1215
MG71-2 sgRNA targeting ATP7B	1066	71-2_ATP7B_P1E5	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1216
MG71-2 sgRNA targeting ATP7B	1067	71-2_ATP7B_P1H5	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1217
MG71-2 sgRNA targeting ATP7B	1068	71-2_ATP7B_P1D6	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1218

Cat.	SEQ ID NO:	Description	Spacer/ Target Site Information
MG71-2 sgRNA targeting ATP7B	1069	71-2_ATP7B_P1E6	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1219
MG71-2 sgRNA targeting ATP7B	1070	71-2_ATP7B_P1F6	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1220
MG71-2 sgRNA targeting ATP7B	1071	71-2_ATP7B_P1G6	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1221
MG71-2 sgRNA targeting ATP7B	1072	71-2_ATP7B_P1C7	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1222
MG71-2 sgRNA targeting ATP7B	1073	71-2_ATP7B_P1G8	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1223
MG71-2 sgRNA targeting ATP7B	1074	71-2_ATP7B_P1D9	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1224
MG71-2 sgRNA targeting ATP7B	1075	71-2_ATP7B_P1H9	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1225
MG71-2 sgRNA targeting ATP7B	1076	71-2_ATP7B_P1A10	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1226
MG71-2 sgRNA targeting ATP7B	1077	71-2_ATP7B_P1B10	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1227
MG71-2 sgRNA targeting ATP7B	1078	71-2_ATP7B_P1E10	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1228
MG71-2 sgRNA targeting ATP7B	1079	71-2_ATP7B_P1G10	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1229
MG71-2 sgRNA targeting ATP7B	1080	71-2_ATP7B_P1H10	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1230
MG71-2 sgRNA targeting ATP7B	1081	71-2_ATP7B_P1A11	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1231
MG71-2 sgRNA targeting AAVS1	1082	71-2_AAVS1_P1G12	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1232
MG71-2 sgRNA targeting Albumin	1083	gr2169	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1233
MG71-2 sgRNA targeting Albumin	1084	gr2171	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1234
MG71-2 sgRNA targeting Albumin	1085	gr2172	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1235
MG71-2 sgRNA targeting Albumin	1086	gr2175	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1236
MG71-2 sgRNA targeting Albumin	1087	gr2177	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1237
MG71-2 sgRNA targeting Albumin	1088	gr2178	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1238

Cat.	SEQ ID NO:	Description	Spacer/ Target Site Information
MG71-2 sgRNA targeting Albumin	1089	gr2179	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1239
MG71-2 sgRNA targeting Albumin	1090	gr2180	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1240
MG71-2 sgRNA targeting Albumin	1091	gr2186	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1241
MG71-2 sgRNA targeting Albumin	1092	gr2189	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1242
MG71-2 sgRNA targeting Albumin	1093	gr2198	Used for MG71-43 only; Target site SEQ ID NO: 1243
MG71-2 sgRNA targeting Albumin	1094	gr2201	Used for MG71-43 only; Target site SEQ ID NO: 1244
MG71-2 sgRNA targeting B2M	1095	g1-MG71-2_B2M_C2	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1245
MG71-2 sgRNA targeting TRAC	1096	g2-MG71-2_TRAC_F1	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1246
MG71-2 sgRNA targeting TRAC	1097	g3-MG71-2_TRAC_G1	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1247
MG71-2 sgRNA targeting AAVS1	1098	g4-MG71-2_AAVS1_C2	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1248
MG71-2 sgRNA targeting AAVS1	1099	g5-MG71-2_AAVS1_D2	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1249
MG71-2 sgRNA targeting AAVS1	1100	g6-MG71-2_AAVS1_F2	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1250
MG71-2 sgRNA targeting AAVS1	1101	g7-MG71-2_AAVS1_H2	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1251
MG71-2 sgRNA targeting AAVS1	1102	g8-MG71-2_AAVS1_D3	Used for both 71-2 and 71-43; Target site SEQ ID NO: 1252
MG71-2 variant sgRNA targeting PAH R408W	1113	pool_gRNA_R408W_rv_1	Target site SEQ ID NO: 1253
MG71-2 variant sgRNA targeting PAH R408W	1114	pool_gRNA_R408W_rv_2	Target site SEQ ID NO: 1254
MG71-2 variant sgRNA targeting PAH R408W	1115	pool_gRNA_R408W_rv_3	Target site SEQ ID NO: 1255
MG71-2 variant sgRNA targeting PAH R408W	1116	pool_gRNA_R408W_rv_4	Target site SEQ ID NO: 1256
MG71-2 variant sgRNA targeting PAH R408W	1117	pool_gRNA_R408W_rv_5	Target site SEQ ID NO: 1257
MG71-2 variant sgRNA targeting PAH R408W	1118	pool_gRNA_R408W_rv_6	Target site SEQ ID NO: 1258

Cat.	SEQ ID NO:	Description	Spacer/ Target Site Information
MG71-2 variant sgRNA targeting PAH R408W	1119	pool_gRNA_R408W_rv_7	Target site SEQ ID NO: 1259
MG71-2 variant sgRNA targeting PAH R408W	1120	pool_gRNA_R408W_rv_8	Target site SEQ ID NO: 1260
MG71-2 variant sgRNA targeting PAH R408W	1121	pool_gRNA_R408W_rv_9	Target site SEQ ID NO: 1261
MG71-2 variant sgRNA targeting PAH R408W	1122	pool_gRNA_R408W_rv_10	Target site SEQ ID NO: 1262
MG23-1 guide targeting human AAVS1	1486	gr1825_hAAVS1-231-G1-224	spacer length 18
MG23-1 guide targeting human AAVS1	1487	gr1826_hAAVS1-231-G1-225	spacer length 19
MG23-1 guide targeting human AAVS1	1488	gr1827_hAAVS1-231-G1-226	spacer length 20
MG23-1 guide targeting human AAVS1	1489	gr1828_hAAVS1-231-G1-227	spacer length 21
MG23-1 guide targeting human AAVS1	1490	gr1829_hAAVS1-231-G1-1	spacer length 22
MG23-1 guide targeting human AAVS1	1491	gr1830_hAAVS1-231-G1-228	spacer length 23
MG23-1 guide targeting human AAVS1	1492	gr1831_hAAVS1-231-G1-229	spacer length 24
MG23-1 guide targeting human AAVS1	1493	gr1832_hAAVS1-231-G1-230	spacer length 25
MG23-1 guide targeting human TRAC	1494	gr1833_hTRAC-231-H7-224	spacer length 18
MG23-1 guide targeting human TRAC	1495	gr1834_hTRAC-231-H7-225	spacer length 19
MG23-1 guide targeting human TRAC	1496	gr1835_hTRAC-231-H7-226	spacer length 20
MG23-1 guide targeting human TRAC	1497	gr1836_hTRAC-231-H7-227	spacer length 21
MG23-1 guide targeting human TRAC	1498	gr1837_hTRAC-231-H7-1	spacer length 22
MG23-1 guide targeting human TRAC	1499	gr1838_hTRAC-231-H7-228	spacer length 23
MG23-1 guide targeting human TRAC	1500	gr1839_hTRAC-231-H7-229	spacer length 24
MG23-1 guide targeting human TRAC	1501	gr1840_hTRAC-231-H7-230	spacer length 25
MG21-1 guide targeting human HAO1	1502	gr1878_hHAO1-211-F1-78	spacer length 22
MG21-1 guide targeting human HAO1	1503	gr1879_hHAO1-211-F1-240	spacer length 23
MG21-1 guide targeting human HAO1	1504	gr1880_hHAO1-211-F1-239	spacer length 24
MG21-1 guide targeting human HAO1	1505	gr1881_hHAO1-211-F1-241	spacer length 25
MG21-1 guide targeting human HAO1	1506	gr1882_hHAO1-211-F1-157	spacer length 22
MG21-1 guide targeting human HAO1	1507	gr1883_hHAO1-211-F1-242	spacer length 23
MG21-1 guide targeting human HAO1	1508	gr1884_hHAO1-211-F1-243	spacer length 24

Cat.	SEQ ID NO:	Description	Spacer/ Target Site Information
MG21-1 guide targeting human HAO1	1509	gr1885_hHAO1-211-F1-244	spacer length 25
MG71-2 guide targeting human TRAC	1510	hTRAC-712-C2-519	spacer length 18
MG71-2 guide targeting human TRAC	1511	hTRAC-712-C2-520	spacer length 19
MG71-2 guide targeting human TRAC	1512	hTRAC-712-C2-521	spacer length 20
MG71-2 guide targeting human TRAC	1513	hTRAC-712-C2-12	spacer length 21
MG71-2 guide targeting human TRAC	1514	hTRAC-712-C2-522	spacer length 22
MG71-2 guide targeting human TRAC	1515	hTRAC-712-C2-523	spacer length 23
MG71-2 guide targeting human TRAC	1516	hTRAC-712-C2-524	spacer length 24
MG71-2 guide targeting human TRAC	1517	hTRAC-712-C2-525	spacer length 25
MG71-2 guide targeting human TRAC	1518	hTRAC-712-C2-526	spacer length 18
MG71-2 guide targeting human TRAC	1519	hTRAC-712-C2-527	spacer length 19
MG71-2 guide targeting human TRAC	1520	hTRAC-712-C2-528	spacer length 20
MG71-2 guide targeting human TRAC	1521	hTRAC-712-C2-255	spacer length 21
MG71-2 guide targeting human TRAC	1522	hTRAC-712-C2-529	spacer length 22
MG71-2 guide targeting human TRAC	1523	hTRAC-712-C2-530	spacer length 23
MG71-2 guide targeting human TRAC	1524	hTRAC-712-C2-531	spacer length 24
MG71-2 guide targeting human TRAC	1525	hTRAC-712-C2-532	spacer length 25
MG21-1 sgRNA Targeting Human ATXN2	1786	MG21-1 Human ATXN2 P1A1	CCGCCGGAGGTCGCCGCGA CCA (SEQID NO: 2168)
MG21-1 sgRNA Targeting Human ATXN2	1787	MG21-1 Human ATXN2 P1B1	CGGAGGTTCGCCGCGACCAC CGA (SEQID NO: 2169)
MG21-1 sgRNA Targeting Human ATXN2	1788	MG21-1 Human ATXN2 P1C1	GAGGAGGGAGCCGTGGCC GAGG (SEQID NO: 2170)
MG21-1 sgRNA Targeting Human ATXN2	1789	MG21-1 Human ATXN2 P1D1	CTCCGGCGGCGGGAGGCCC GGC (SEQID NO: 2171)
MG21-1 sgRNA Targeting Human ATXN2	1790	MG21-1 Human ATXN2 P1E1	ATTAAAGTTACTCACCGTA GAC (SEQID NO: 2172)
MG21-1 sgRNA Targeting Human ATXN2	1791	MG21-1 Human ATXN2 P1F1	TCCTTTGTTACTGTTTCGAC CT (SEQID NO: 2173)
MG21-1 sgRNA Targeting Human ATXN2	1792	MG21-1 Human ATXN2 P1G1	AATATTTAATCTTTTCAGA GGT (SEQID NO: 2174)
MG21-1 sgRNA Targeting Human ATXN2	1793	MG21-1 Human ATXN2 P1H1	CAAAGGACTGCCTCAGTCT ACG (SEQID NO: 2175)
MG21-1 sgRNA Targeting Human ATXN2	1794	MG21-1 Human ATXN2 P1A2	TGAGTAACTTTAATGTTAC TTA (SEQID NO: 2176)
MG21-1 sgRNA Targeting Human ATXN2	1795	MG21-1 Human ATXN2 P1B2	GAGTAACTTTAATGTTACT TAT (SEQID NO: 2177)
MG21-1 sgRNA Targeting Human ATXN2	1796	MG21-1 Human ATXN2 P1C2	AGTAACTTTAATGTTACTT ATT (SEQID NO: 2178)
MG21-1 sgRNA Targeting Human ATXN2	1797	MG21-1 Human ATXN2 P1D2	CTCATATTTGCATAGATTCC AT (SEQID NO: 2179)

Cat.	SEQ ID NO:	Description	Spacer/ Target Site Information
MG21-1 sgRNA Targeting Human ATXN2	1798	MG21-1 Human ATXN2 P1E2	TCATATTTGCATAGATTCC ATC (SEQID NO: 2180)
MG21-1 sgRNA Targeting Human ATXN2	1799	MG21-1 Human ATXN2 P1F2	CCATCAAAAAGAAATCTGGA ATA (SEQID NO: 2181)
MG21-1 sgRNA Targeting Human ATXN2	1800	MG21-1 Human ATXN2 P1G2	CATCAAAAAGAAATCTGGAA TAT (SEQID NO: 2182)
MG21-1 sgRNA Targeting Human ATXN2	1801	MG21-1 Human ATXN2 P1H2	ATCAAAAAGAAATCTGGAAT ATT (SEQID NO: 2183)
MG21-1 sgRNA Targeting Human ATXN2	1802	MG21-1 Human ATXN2 P1A3	TCAAAAAGAAATCTGGAATA TTA (SEQID NO: 2184)
MG21-1 sgRNA Targeting Human ATXN2	1803	MG21-1 Human ATXN2 P1B3	CAAAAAGAAATCTGGAATAT TAA (SEQID NO: 2185)
MG21-1 sgRNA Targeting Human ATXN2	1804	MG21-1 Human ATXN2 P1C3	AAAAGAAATCTGGAATATT AAA (SEQID NO: 2186)
MG21-1 sgRNA Targeting Human ATXN2	1805	MG21-1 Human ATXN2 P1D3	AAAGAAATCTGGAATATTA AAA (SEQID NO: 2187)
MG21-1 sgRNA Targeting Human ATXN2	1806	MG21-1 Human ATXN2 P1E3	AAGAAATCTGGAATATTAA AAA (SEQID NO: 2188)
MG21-1 sgRNA Targeting Human ATXN2	1807	MG21-1 Human ATXN2 P1F3	TGTTGTAAAGTTATTAGATT ATT (SEQID NO: 2189)
MG21-1 sgRNA Targeting Human ATXN2	1808	MG21-1 Human ATXN2 P1G3	TTTACTTTGTAAGAAAAAG AAA (SEQID NO: 2190)
MG21-1 sgRNA Targeting Human ATXN2	1809	MG21-1 Human ATXN2 P1H3	CTAAAAACATATAATTTAT TTA (SEQID NO: 2191)
MG21-1 sgRNA Targeting Human ATXN2	1810	MG21-1 Human ATXN2 P1A4	TAAAAACATATAATTTATT TAT (SEQID NO: 2192)
MG21-1 sgRNA Targeting Human ATXN2	1811	MG21-1 Human ATXN2 P1B4	GGGCTCCAAATGTGAAGTA CAA (SEQID NO: 2193)
MG21-1 sgRNA Targeting Human ATXN2	1812	MG21-1 Human ATXN2 P1C4	AAGTGAAAAATGGAGGTAT ATA (SEQID NO: 2194)
MG21-1 sgRNA Targeting Human ATXN2	1813	MG21-1 Human ATXN2 P1D4	GAGTTTTTAAACTTACAG TCC (SEQID NO: 2195)
MG21-1 sgRNA Targeting Human ATXN2	1814	MG21-1 Human ATXN2 P1E4	TACCACAACAAAGTCTGAA CAT (SEQID NO: 2196)
MG21-1 sgRNA Targeting Human ATXN2	1815	MG21-1 Human ATXN2 P1F4	TCAAGTACCAAATCACACT AAA (SEQID NO: 2197)
MG21-1 sgRNA Targeting Human ATXN2	1816	MG21-1 Human ATXN2 P1G4	AAATCACACTAAAATGAAA AAC (SEQID NO: 2198)
MG21-1 sgRNA Targeting Human ATXN2	1817	MG21-1 Human ATXN2 P1H4	GATTTGGTACTTGATGCCG CAC (SEQID NO: 2199)
MG21-1 sgRNA Targeting Human ATXN2	1818	MG21-1 Human ATXN2 P1A5	ATTTGGTACTTGATGCCGC ACA (SEQID NO: 2200)
MG21-1 sgRNA Targeting Human ATXN2	1819	MG21-1 Human ATXN2 P1B5	TTTGGTACTTGATGCCGCA CAT (SEQID NO: 2201)
MG21-1 sgRNA Targeting Human ATXN2	1820	MG21-1 Human ATXN2 P1C5	TAAAGATATGGACTCCAGT TAT (SEQID NO: 2202)
MG21-1 sgRNA Targeting Human ATXN2	1821	MG21-1 Human ATXN2 P1D5	AAGATATGGACTCCAGTTA TGC (SEQID NO: 2203)
MG21-1 sgRNA Targeting Human ATXN2	1822	MG21-1 Human ATXN2 P1E5	AGCAGAGTCAGTAAAAGC ATCT (SEQID NO: 2204)
MG21-1 sgRNA Targeting Human ATXN2	1823	MG21-1 Human ATXN2 P1F5	AGTGCTAAAGTGAATGGCG AAC (SEQID NO: 2205)
MG21-1 sgRNA Targeting Human ATXN2	1824	MG21-1 Human ATXN2 P1G5	TGCTAAAGTGAATGGCGAA CAC (SEQID NO: 2206)
MG21-1 sgRNA Targeting Human ATXN2	1825	MG21-1 Human ATXN2 P1H5	GCTAAAGTGAATGGCGAAC ACA (SEQID NO: 2207)
MG21-1 sgRNA Targeting Human ATXN2	1826	MG21-1 Human ATXN2 P1A6	GATGCAGGTGAACCTCACAG CCA (SEQID NO: 2208)

Cat.	SEQ ID NO:	Description	Spacer/ Target Site Information
MG21-1 sgRNA Targeting Human ATXN2	1827	MG21-1 Human ATXN2 P1B6	AGCCAATGAGGAACTTGAG GCT (SEQID NO: 2209)
MG21-1 sgRNA Targeting Human ATXN2	1828	MG21-1 Human ATXN2 P1C6	TGAGGCTTTGGAAAATGAC GTA (SEQID NO: 2210)
MG21-1 sgRNA Targeting Human ATXN2	1829	MG21-1 Human ATXN2 P1D6	AACTTTTCAAACCTTACGTA TAC (SEQID NO: 2211)
MG21-1 sgRNA Targeting Human ATXN2	1830	MG21-1 Human ATXN2 P1E6	ATTGGGATCCCATCCATTA GAC (SEQID NO: 2212)
MG21-1 sgRNA Targeting Human ATXN2	1831	MG21-1 Human ATXN2 P1F6	TTGGGATCCCATCCATTAG ACT (SEQID NO: 2213)
MG21-1 sgRNA Targeting Human ATXN2	1832	MG21-1 Human ATXN2 P1G6	TGGGATCCCATCCATTAGA CTA (SEQID NO: 2214)
MG21-1 sgRNA Targeting Human ATXN2	1833	MG21-1 Human ATXN2 P1H6	CCCAATGATATGTTTCGAT ATA (SEQID NO: 2215)
MG21-1 sgRNA Targeting Human ATXN2	1834	MG21-1 Human ATXN2 P1A7	CCAATGATATGTTTCGATA TAA (SEQID NO: 2216)
MG21-1 sgRNA Targeting Human ATXN2	1835	MG21-1 Human ATXN2 P1B7	CAATGATATGTTTCGATAT AAT (SEQID NO: 2217)
MG21-1 sgRNA Targeting Human ATXN2	1836	MG21-1 Human ATXN2 P1C7	GTTTATCTTCGTATACGTAA GT (SEQID NO: 2218)
MG21-1 sgRNA Targeting Human ATXN2	1837	MG21-1 Human ATXN2 P1D7	CTTTCTAAGGGCACTCTGA AAC (SEQID NO: 2219)
MG21-1 sgRNA Targeting Human ATXN2	1838	MG21-1 Human ATXN2 P1E7	TTTCTAAGGGCACTCTGAA ACA (SEQID NO: 2220)
MG21-1 sgRNA Targeting Human ATXN2	1839	MG21-1 Human ATXN2 P1F7	TTCTAAGGGCACTCTGAAA CAT (SEQID NO: 2221)
MG21-1 sgRNA Targeting Human ATXN2	1840	MG21-1 Human ATXN2 P1G7	TCTAAGGGCACTCTGAAAC ATG (SEQID NO: 2222)
MG21-1 sgRNA Targeting Human ATXN2	1841	MG21-1 Human ATXN2 P1H7	TAAGGGCACTCTGAAACAT GAG (SEQID NO: 2223)
MG21-1 sgRNA Targeting Human ATXN2	1842	MG21-1 Human ATXN2 P1A8	AAGGGCACTCTGAAACATG AGG (SEQID NO: 2224)
MG21-1 sgRNA Targeting Human ATXN2	1843	MG21-1 Human ATXN2 P1B8	AGGGCACTCTGAAACATGA GGA (SEQID NO: 2225)
MG21-1 sgRNA Targeting Human ATXN2	1844	MG21-1 Human ATXN2 P1C8	GGGCACTCTGAAACATGAG GAA (SEQID NO: 2226)
MG21-1 sgRNA Targeting Human ATXN2	1845	MG21-1 Human ATXN2 P1D8	CACTCTGAAACATGAGGAA AAG (SEQID NO: 2227)
MG21-1 sgRNA Targeting Human ATXN2	1846	MG21-1 Human ATXN2 P1E8	TTCTCATGTTTCAGAGTGCC CC (SEQID NO: 2228)
MG21-1 sgRNA Targeting Human ATXN2	1847	MG21-1 Human ATXN2 P1F8	CCTCATGTTTCAGAGTGCC CTT (SEQID NO: 2229)
MG21-1 sgRNA Targeting Human ATXN2	1848	MG21-1 Human ATXN2 P1G8	ATGTTTCAGAGTGCCCTTA GAA (SEQID NO: 2230)
MG21-1 sgRNA Targeting Human ATXN2	1849	MG21-1 Human ATXN2 P1H8	GTGCCCTTAGAAAGAGATA ACT (SEQID NO: 2231)
MG21-1 sgRNA Targeting Human ATXN2	1850	MG21-1 Human ATXN2 P1A9	ATAACTCAGAAGAATTTTT AAA (SEQID NO: 2232)
MG21-1 sgRNA Targeting Human ATXN2	1851	MG21-1 Human ATXN2 P1B9	ACTCAGAAGAATTTTTAAA ACG (SEQID NO: 2233)
MG21-1 sgRNA Targeting Human ATXN2	1852	MG21-1 Human ATXN2 P1C9	AAGAATTTTTTAAACGGGA AGC (SEQID NO: 2234)
MG21-1 sgRNA Targeting Human ATXN2	1853	MG21-1 Human ATXN2 P1D9	GAAGCAAGGGCAAACCAG TTAG (SEQID NO: 2235)
MG21-1 sgRNA Targeting Human ATXN2	1854	MG21-1 Human ATXN2 P1E9	AAGCAAGGGCAAACCAGTT AGC (SEQID NO: 2236)
MG21-1 sgRNA Targeting Human ATXN2	1855	MG21-1 Human ATXN2 P1F9	CCAGTACAAAGCTCGAGTG GCC (SEQID NO: 2237)

Cat.	SEQ ID NO:	Description	Spacer/ Target Site Information
MG21-1 sgRNA Targeting Human ATXN2	1856	MG21-1 Human ATXN2 P1G9	GCCCTGGAAAATGATGATA GGA (SEQID NO: 2238)
MG21-1 sgRNA Targeting Human ATXN2	1857	MG21-1 Human ATXN2 P1H9	CCCTGGAAAATGATGATAG GAG (SEQID NO: 2239)
MG21-1 sgRNA Targeting Human ATXN2	1858	MG21-1 Human ATXN2 P1A10	CTGGAAAATGATGATAGGA GTG (SEQID NO: 2240)
MG21-1 sgRNA Targeting Human ATXN2	1859	MG21-1 Human ATXN2 P1B10	TGGAAAATGATGATAGGAG TGA (SEQID NO: 2241)
MG21-1 sgRNA Targeting Human ATXN2	1860	MG21-1 Human ATXN2 P1C10	GGAAAATGATGATAGGAGT GAG (SEQID NO: 2242)
MG21-1 sgRNA Targeting Human ATXN2	1861	MG21-1 Human ATXN2 P1D10	GAAAATGATGATAGGAGTG AGG (SEQID NO: 2243)
MG21-1 sgRNA Targeting Human ATXN2	1862	MG21-1 Human ATXN2 P1E10	AGGAAGAAAAATACACAG CAGT (SEQID NO: 2244)
MG21-1 sgRNA Targeting Human ATXN2	1863	MG21-1 Human ATXN2 P1F10	GGAAGAAAAATACACAGC AGTT (SEQID NO: 2245)
MG21-1 sgRNA Targeting Human ATXN2	1864	MG21-1 Human ATXN2 P1G10	CACAGCATAAACACTAGGT ATT (SEQID NO: 2246)
MG21-1 sgRNA Targeting Human ATXN2	1865	MG21-1 Human ATXN2 P1H10	ACAGCATAAACACTAGGTA TTT (SEQID NO: 2247)
MG21-1 sgRNA Targeting Human ATXN2	1866	MG21-1 Human ATXN2 P1A11	AATTCGGGTTGAAATCTGA AGT (SEQID NO: 2248)
MG21-1 sgRNA Targeting Human ATXN2	1867	MG21-1 Human ATXN2 P1B11	TGACTTCTCTATTTCTTTGT CC (SEQID NO: 2249)
MG21-1 sgRNA Targeting Human ATXN2	1868	MG21-1 Human ATXN2 P1C11	TGATTTAGAGGATCTCTCA TTT (SEQID NO: 2250)
MG21-1 sgRNA Targeting Human ATXN2	1869	MG21-1 Human ATXN2 P1D11	GATTTAGAGGATCTCTCAT TTT (SEQID NO: 2251)
MG21-1 sgRNA Targeting Human ATXN2	1870	MG21-1 Human ATXN2 P1E11	ATTTAGAGGATCTCTCATTT TC (SEQID NO: 2252)
MG21-1 sgRNA Targeting Human ATXN2	1871	MG21-1 Human ATXN2 P1F11	TAGAGGATCTCTCATTTTC AGG (SEQID NO: 2253)
MG21-1 sgRNA Targeting Human ATXN2	1872	MG21-1 Human ATXN2 P1G11	AGGGAAAATAAATATATTC CTC (SEQID NO: 2254)
MG21-1 sgRNA Targeting Human ATXN2	1873	MG21-1 Human ATXN2 P1H11	AAAATAAATATATTCCTCC TGG (SEQID NO: 2255)
MG21-1 sgRNA Targeting Human ATXN2	1874	MG21-1 Human ATXN2 P1A12	AAATAAATATATTCCTCCT GGA (SEQID NO: 2256)
MG21-1 sgRNA Targeting Human ATXN2	1875	MG21-1 Human ATXN2 P1B12	TAAATATATTCCTCCTGGA CAA (SEQID NO: 2257)
MG21-1 sgRNA Targeting Human ATXN2	1876	MG21-1 Human ATXN2 P1C12	TATATTCCTCCTGGACAAA GAA (SEQID NO: 2258)
MG21-1 sgRNA Targeting Human ATXN2	1877	MG21-1 Human ATXN2 P1D12	ATATTCCTCCTGGACAAAG AAA (SEQID NO: 2259)
MG21-1 sgRNA Targeting Human ATXN2	1878	MG21-1 Human ATXN2 P1E12	AAAGAAATAGAGAAGTCA TATC (SEQID NO: 2260)
MG21-1 sgRNA Targeting Human ATXN2	1879	MG21-1 Human ATXN2 P1F12	AAGAAATAGAGAAGTCAT ATCC (SEQID NO: 2261)
MG21-1 sgRNA Targeting Human ATXN2	1880	MG21-1 Human ATXN2 P1G12	TCAACCCGAATTCTGGTTC AGA (SEQID NO: 2262)
MG21-1 sgRNA Targeting Human ATXN2	1881	MG21-1 Human ATXN2 P1H12	ACCCGAATTCTGGTTCAGA CCA (SEQID NO: 2263)
MG21-1 sgRNA Targeting Human ATXN2	1882	MG21-1 Human ATXN2 P2A1	TTCAGACCAAAGAGTAGTT AAT (SEQID NO: 2264)
MG21-1 sgRNA Targeting Human ATXN2	1883	MG21-1 Human ATXN2 P2B1	AAGACATGCGTTTAGGCAT AGT (SEQID NO: 2265)
MG21-1 sgRNA Targeting Human ATXN2	1884	MG21-1 Human ATXN2 P2C1	TTAGGCATAGTAGAGACAG GAG (SEQID NO: 2266)

Cat.	SEQ ID NO:	Description	Spacer/ Target Site Information
MG21-1 sgRNA Targeting Human ATXN2	1885	MG21-1 Human ATXN2 P2D1	TAGAGACAGGAGCTGGAG AACC (SEQID NO: 2267)
MG21-1 sgRNA Targeting Human ATXN2	1886	MG21-1 Human ATXN2 P2E1	GAGAGTTGGGACCTGACTG GTA (SEQID NO: 2268)
MG21-1 sgRNA Targeting Human ATXN2	1887	MG21-1 Human ATXN2 P2F1	GACTGGTAGCGAGAAGGTG GGC (SEQID NO: 2269)
MG21-1 sgRNA Targeting Human ATXN2	1888	MG21-1 Human ATXN2 P2G1	ACTGGTAGCGAGAAGGTGG GCG (SEQID NO: 2270)
MG21-1 sgRNA Targeting Human ATXN2	1889	MG21-1 Human ATXN2 P2H1	GGTAGCGAGAAGGTGGGC GAGA (SEQID NO: 2271)
MG21-1 sgRNA Targeting Human ATXN2	1890	MG21-1 Human ATXN2 P2A2	GAAGGTGGCGAGAGGAA GGAG (SEQID NO: 2272)
MG21-1 sgRNA Targeting Human ATXN2	1891	MG21-1 Human ATXN2 P2B2	AAGGAGATGGGCAAGGCG ATGG (SEQID NO: 2273)
MG21-1 sgRNA Targeting Human ATXN2	1892	MG21-1 Human ATXN2 P2C2	GCGATGGCCAGGGAACACC TGA (SEQID NO: 2274)
MG21-1 sgRNA Targeting Human ATXN2	1893	MG21-1 Human ATXN2 P2D2	ACACCTGACAGAACAAATG ATA (SEQID NO: 2275)
MG21-1 sgRNA Targeting Human ATXN2	1894	MG21-1 Human ATXN2 P2E2	CACCTGACAGAACAAATGA TAT (SEQID NO: 2276)
MG21-1 sgRNA Targeting Human ATXN2	1895	MG21-1 Human ATXN2 P2F2	ACCTGACAGAACAAATGAT ATG (SEQID NO: 2277)
MG21-1 sgRNA Targeting Human ATXN2	1896	MG21-1 Human ATXN2 P2G2	CAGGTCCCACTCTCTTCC ACC (SEQID NO: 2278)
MG21-1 sgRNA Targeting Human ATXN2	1897	MG21-1 Human ATXN2 P2H2	CCGAGGGACTGGTCTTGC TAC (SEQID NO: 2279)
MG21-1 sgRNA Targeting Human ATXN2	1898	MG21-1 Human ATXN2 P2A3	AGGGACTGGTCTTGCTAC TGG (SEQID NO: 2280)
MG21-1 sgRNA Targeting Human ATXN2	1899	MG21-1 Human ATXN2 P2B3	GGGACATCCTTGGAGGCC TAA (SEQID NO: 2281)
MG21-1 sgRNA Targeting Human ATXN2	1900	MG21-1 Human ATXN2 P2C3	ACATCCTTGGAGGCCCTAA GGA (SEQID NO: 2282)
MG21-1 sgRNA Targeting Human ATXN2	1901	MG21-1 Human ATXN2 P2D3	TTGTATCCCACAACCCACC CAG (SEQID NO: 2283)
MG21-1 sgRNA Targeting Human ATXN2	1902	MG21-1 Human ATXN2 P2E3	CAGTAGCAAGGACCAGTCC CTC (SEQID NO: 2284)
MG21-1 sgRNA Targeting Human ATXN2	1903	MG21-1 Human ATXN2 P2F3	AACGTGGTCATCAGTGGTC AGT (SEQID NO: 2285)
MG21-1 sgRNA Targeting Human ATXN2	1904	MG21-1 Human ATXN2 P2G3	TTTGTTACATACCCTCACTA GA (SEQID NO: 2286)
MG21-1 sgRNA Targeting Human ATXN2	1905	MG21-1 Human ATXN2 P2H3	TAGCAGGCGTAGGAGATGC AGC (SEQID NO: 2287)
MG21-1 sgRNA Targeting Human ATXN2	1906	MG21-1 Human ATXN2 P2A4	TAGGCATGGCAACAGCTTC AGT (SEQID NO: 2288)
MG21-1 sgRNA Targeting Human ATXN2	1907	MG21-1 Human ATXN2 P2B4	CTTCAGTTGGAATAATACC AGC (SEQID NO: 2289)
MG21-1 sgRNA Targeting Human ATXN2	1908	MG21-1 Human ATXN2 P2C4	TCAGTTGGAATAATACCAG CTT (SEQID NO: 2290)
MG21-1 sgRNA Targeting Human ATXN2	1909	MG21-1 Human ATXN2 P2D4	CAGTTGGAATAATACCAGC TTG (SEQID NO: 2291)
MG21-1 sgRNA Targeting Human ATXN2	1910	MG21-1 Human ATXN2 P2E4	AGACCTGGGTCTATGAGTT TTA (SEQID NO: 2292)
MG21-1 sgRNA Targeting Human ATXN2	1911	MG21-1 Human ATXN2 P2F4	GGGATAATCTTGGAAGTAG AAG (SEQID NO: 2293)
MG21-1 sgRNA Targeting Human ATXN2	1912	MG21-1 Human ATXN2 P2G4	GGATAATCTTGGAAGTAGA AGA (SEQID NO: 2294)
MG21-1 sgRNA Targeting Human ATXN2	1913	MG21-1 Human ATXN2 P2H4	GATAATCTTGGAAGTAGAA GAA (SEQID NO: 2295)

Cat.	SEQ ID NO:	Description	Spacer/ Target Site Information
MG21-1 sgRNA Targeting Human ATXN2	1914	MG21-1 Human ATXN2 P2A5	AAAACTCATAGACCCAGGT CTC (SEQID NO: 2296)
MG21-1 sgRNA Targeting Human ATXN2	1915	MG21-1 Human ATXN2 P2B5	TCATAGACCCAGGTCTCCC AGA (SEQID NO: 2297)
MG21-1 sgRNA Targeting Human ATXN2	1916	MG21-1 Human ATXN2 P2C5	TCCTACGCCTGCTAGTCCT GCA (SEQID NO: 2298)
MG21-1 sgRNA Targeting Human ATXN2	1917	MG21-1 Human ATXN2 P2D5	CTCTAACCTTTGTTTTTCAGC TT (SEQID NO: 2299)
MG21-1 sgRNA Targeting Human ATXN2	1918	MG21-1 Human ATXN2 P2E5	TTAATATTTTCTTTATTCCC TG (SEQID NO: 2300)
MG21-1 sgRNA Targeting Human ATXN2	1919	MG21-1 Human ATXN2 P2F5	TCTTGAAGCCTGGAATCTT TAG (SEQID NO: 2301)
MG21-1 sgRNA Targeting Human ATXN2	1920	MG21-1 Human ATXN2 P2G5	TGAAGCCTGGAATCTTTAG CTA (SEQID NO: 2302)
MG21-1 sgRNA Targeting Human ATXN2	1921	MG21-1 Human ATXN2 P2H5	TCTTTAGCTAGAAAACAAT TTT (SEQID NO: 2303)
MG21-1 sgRNA Targeting Human ATXN2	1922	MG21-1 Human ATXN2 P2A6	CTTTAGCTAGAAAACAATT TTT (SEQID NO: 2304)
MG21-1 sgRNA Targeting Human ATXN2	1923	MG21-1 Human ATXN2 P2B6	AAAGATTCCAGGCTTCAAG ATC (SEQID NO: 2305)
MG21-1 sgRNA Targeting Human ATXN2	1924	MG21-1 Human ATXN2 P2C6	AGATCAGAGGCAGAACTCT CCT (SEQID NO: 2306)
MG21-1 sgRNA Targeting Human ATXN2	1925	MG21-1 Human ATXN2 P2D6	TCAGAGGCAGAACTCTCCT GCA (SEQID NO: 2307)
MG21-1 sgRNA Targeting Human ATXN2	1926	MG21-1 Human ATXN2 P2E6	AGGCAGAACTCTCCTGCAG GGA (SEQID NO: 2308)
MG21-1 sgRNA Targeting Human ATXN2	1927	MG21-1 Human ATXN2 P2F6	GGCAGAACTCTCCTGCAGG GAA (SEQID NO: 2309)
MG21-1 sgRNA Targeting Human ATXN2	1928	MG21-1 Human ATXN2 P2G6	GCAGAACTCTCCTGCAGGG AAT (SEQID NO: 2310)
MG21-1 sgRNA Targeting Human ATXN2	1929	MG21-1 Human ATXN2 P2H6	ATCACCTAGCTTCTCAAAA GCT (SEQID NO: 2311)
MG21-1 sgRNA Targeting Human ATXN2	1930	MG21-1 Human ATXN2 P2A7	AGCTGAAAACAAAGGTTAG AGT (SEQID NO: 2312)
MG21-1 sgRNA Targeting Human ATXN2	1931	MG21-1 Human ATXN2 P2B7	CATCAATCTGTTTCTATGT TC (SEQID NO: 2313)
MG21-1 sgRNA Targeting Human ATXN2	1932	MG21-1 Human ATXN2 P2C7	ATATCACCAGTTGTTTCTG AAC (SEQID NO: 2314)
MG21-1 sgRNA Targeting Human ATXN2	1933	MG21-1 Human ATXN2 P2D7	TCACCAGTTGTTTCTGAAC ATA (SEQID NO: 2315)
MG21-1 sgRNA Targeting Human ATXN2	1934	MG21-1 Human ATXN2 P2E7	ACATAGAAAACAGATTGAT GAT (SEQID NO: 2316)
MG21-1 sgRNA Targeting Human ATXN2	1935	MG21-1 Human ATXN2 P2F7	CATAGAAAACAGATTGATG ATT (SEQID NO: 2317)
MG21-1 sgRNA Targeting Human ATXN2	1936	MG21-1 Human ATXN2 P2G7	TAAAGAAATTTAAGAATGA TTT (SEQID NO: 2318)
MG21-1 sgRNA Targeting Human ATXN2	1937	MG21-1 Human ATXN2 P2H7	CTGGGCTGGAAGTCTGAAC CCC (SEQID NO: 2319)
MG21-1 sgRNA Targeting Human ATXN2	1938	MG21-1 Human ATXN2 P2A8	TGCTCCGTGTTACTAAGTA TTG (SEQID NO: 2320)
MG21-1 sgRNA Targeting Human ATXN2	1939	MG21-1 Human ATXN2 P2B8	GCTCCGTGTTACTAAGTAT TGA (SEQID NO: 2321)
MG21-1 sgRNA Targeting Human ATXN2	1940	MG21-1 Human ATXN2 P2C8	CAGTTGCTGCTGCTATTTTC AA (SEQID NO: 2322)
MG21-1 sgRNA Targeting Human ATXN2	1941	MG21-1 Human ATXN2 P2D8	TTAGTAGTTGATCCATAGA TTC (SEQID NO: 2323)
MG21-1 sgRNA Targeting Human ATXN2	1942	MG21-1 Human ATXN2 P2E8	CAGAAGTAGAACTTGGCTG TAA (SEQID NO: 2324)

Cat.	SEQ ID NO:	Description	Spacer/ Target Site Information
MG21-1 sgRNA Targeting Human ATXN2	1943	MG21-1 Human ATXN2 P2F8	GTAGAACTTGGCTGTAAC AAA (SEQID NO: 2325)
MG21-1 sgRNA Targeting Human ATXN2	1944	MG21-1 Human ATXN2 P2G8	TAGAACTTGGCTGTAAC AAT (SEQID NO: 2326)
MG21-1 sgRNA Targeting Human ATXN2	1945	MG21-1 Human ATXN2 P2H8	AACTTGGCTGTAAC AAA (SEQID NO: 2327)
MG21-1 sgRNA Targeting Human ATXN2	1946	MG21-1 Human ATXN2 P2A9	TTCTGAATCTATGGATCA CTA (SEQID NO: 2328)
MG21-1 sgRNA Targeting Human ATXN2	1947	MG21-1 Human ATXN2 P2B9	ATCTATGGATCAACTACT AAC (SEQID NO: 2329)
MG21-1 sgRNA Targeting Human ATXN2	1948	MG21-1 Human ATXN2 P2C9	ATGGATCAACTACTAAAC AAA (SEQID NO: 2330)
MG21-1 sgRNA Targeting Human ATXN2	1949	MG21-1 Human ATXN2 P2D9	ATCAACTACTAAACAAAA TAG (SEQID NO: 2331)
MG21-1 sgRNA Targeting Human ATXN2	1950	MG21-1 Human ATXN2 P2E9	CAACTACTAAACAAAAAT GAG (SEQID NO: 2332)
MG21-1 sgRNA Targeting Human ATXN2	1951	MG21-1 Human ATXN2 P2F9	AACTACTAAACAAAAATAG AGA (SEQID NO: 2333)
MG21-1 sgRNA Targeting Human ATXN2	1952	MG21-1 Human ATXN2 P2G9	ACTACTAAACAAAAATAGA GAG (SEQID NO: 2334)
MG21-1 sgRNA Targeting Human ATXN2	1953	MG21-1 Human ATXN2 P2H9	CTACTAAACAAAAATAGAG AGG (SEQID NO: 2335)
MG21-1 sgRNA Targeting Human ATXN2	1954	MG21-1 Human ATXN2 P2A10	AGAAAAATCAAGAGATTG ATC (SEQID NO: 2336)
MG21-1 sgRNA Targeting Human ATXN2	1955	MG21-1 Human ATXN2 P2B10	AAGTGCTAAGGATTCTTC ATT (SEQID NO: 2337)
MG21-1 sgRNA Targeting Human ATXN2	1956	MG21-1 Human ATXN2 P2C10	CTGTACCAGTGGCAGCAGC AAG (SEQID NO: 2338)
MG21-1 sgRNA Targeting Human ATXN2	1957	MG21-1 Human ATXN2 P2D10	GTTCACTTCCAGCCCAG CAT (SEQID NO: 2339)
MG21-1 sgRNA Targeting Human ATXN2	1958	MG21-1 Human ATXN2 P2E10	GACTTCCAGCCCAGCATGT AAA (SEQID NO: 2340)
MG21-1 sgRNA Targeting Human ATXN2	1959	MG21-1 Human ATXN2 P2F10	ACTTCCAGCCCAGCATGTA AAC (SEQID NO: 2341)
MG21-1 sgRNA Targeting Human ATXN2	1960	MG21-1 Human ATXN2 P2G10	CTTCCAGCCCAGCATGTAA ACA (SEQID NO: 2342)
MG21-1 sgRNA Targeting Human ATXN2	1961	MG21-1 Human ATXN2 P2H10	GCATGTAAACAAGAGAAA GACG (SEQID NO: 2343)
MG21-1 sgRNA Targeting Human ATXN2	1962	MG21-1 Human ATXN2 P2A11	CATGTAAACAAGAGAAAG ACGA (SEQID NO: 2344)
MG21-1 sgRNA Targeting Human ATXN2	1963	MG21-1 Human ATXN2 P2B11	TGTAAACAAGAGAAAGAC GATA (SEQID NO: 2345)
MG21-1 sgRNA Targeting Human ATXN2	1964	MG21-1 Human ATXN2 P2C11	TAAACAAGAGAAAGACGA TAAG (SEQID NO: 2346)
MG21-1 sgRNA Targeting Human ATXN2	1965	MG21-1 Human ATXN2 P2D11	AAACAAGAGAAAGACGAT AAGG (SEQID NO: 2347)
MG21-1 sgRNA Targeting Human ATXN2	1966	MG21-1 Human ATXN2 P2E11	ACAAGAGAAAGACGATAA GGAA (SEQID NO: 2348)
MG21-1 sgRNA Targeting Human ATXN2	1967	MG21-1 Human ATXN2 P2F11	CAAGAGAAAGACGATAAG GAAG (SEQID NO: 2349)
MG21-1 sgRNA Targeting Human ATXN2	1968	MG21-1 Human ATXN2 P2G11	AAGAGAAAGACGATAAGG AAGA (SEQID NO: 2350)
MG21-1 sgRNA Targeting Human ATXN2	1969	MG21-1 Human ATXN2 P2H11	AAGAGAAAGAACGCAG CTGA (SEQID NO: 2351)
MG21-1 sgRNA Targeting Human ATXN2	1970	MG21-1 Human ATXN2 P2A12	CAAAGAAAGTAATAAACCT ACC (SEQID NO: 2352)
MG21-1 sgRNA Targeting Human ATXN2	1971	MG21-1 Human ATXN2 P2B12	AAAGAAAGTAATAAACCTA CCT (SEQID NO: 2353)

Cat.	SEQ ID NO:	Description	Spacer/ Target Site Information
MG21-1 sgRNA Targeting Human ATXN2	1972	MG21-1 Human ATXN2 P2C12	GAAAGTAATAAACCTACCT GAG (SEQID NO: 2354)
MG21-1 sgRNA Targeting Human ATXN2	1973	MG21-1 Human ATXN2 P2D12	CAATGTTGATTTCCTAACTT GC (SEQID NO: 2355)
MG21-1 sgRNA Targeting Human ATXN2	1974	MG21-1 Human ATXN2 P2E12	AATGTTGATTTCCTAACTTG CC (SEQID NO: 2356)
MG21-1 sgRNA Targeting Human ATXN2	1975	MG21-1 Human ATXN2 P2F12	ATGTTGATTTCCTAACTTGC CT (SEQID NO: 2357)
MG21-1 sgRNA Targeting Human ATXN2	1976	MG21-1 Human ATXN2 P2G12	AAATCAACATTGAATCCCA ATG (SEQID NO: 2358)
MG21-1 sgRNA Targeting Human ATXN2	1977	MG21-1 Human ATXN2 P2H12	AGTTCAACCCACGTTCCCTT CTC (SEQID NO: 2359)
MG21-1 sgRNA Targeting Human ATXN2	1978	MG21-1 Human ATXN2 P3A1	ACTGGATACATCATATTTG GTG (SEQID NO: 2360)
MG21-1 sgRNA Targeting Human ATXN2	1979	MG21-1 Human ATXN2 P3B1	GATACATCATATTTGGTGC AAA (SEQID NO: 2361)
MG21-1 sgRNA Targeting Human ATXN2	1980	MG21-1 Human ATXN2 P3C1	GTGCTTGAGGCCGAGGTGA AGT (SEQID NO: 2362)
MG21-1 sgRNA Targeting Human ATXN2	1981	MG21-1 Human ATXN2 P3D1	GGTAGTAGAAGGCTTTGGC TAC (SEQID NO: 2363)
MG21-1 sgRNA Targeting Human ATXN2	1982	MG21-1 Human ATXN2 P3E1	CAGTCCCAGTGAGCCCAGG CGT (SEQID NO: 2364)
MG21-1 sgRNA Targeting Human ATXN2	1983	MG21-1 Human ATXN2 P3F1	ATAGGTATTGGGTATAAAG GCT (SEQID NO: 2365)
MG21-1 sgRNA Targeting Human ATXN2	1984	MG21-1 Human ATXN2 P3G1	TGGGTATAAAGGCTTGAGA GAA (SEQID NO: 2366)
MG21-1 sgRNA Targeting Human ATXN2	1985	MG21-1 Human ATXN2 P3H1	GGGTATAAAGGCTTGAGAG AAT (SEQID NO: 2367)
MG21-1 sgRNA Targeting Human ATXN2	1986	MG21-1 Human ATXN2 P3A2	GGTATAAAGGCTTGAGAGA ATT (SEQID NO: 2368)
MG21-1 sgRNA Targeting Human ATXN2	1987	MG21-1 Human ATXN2 P3B2	GTATAAAGGCTTGAGAGAA TTA (SEQID NO: 2369)
MG21-1 sgRNA Targeting Human ATXN2	1988	MG21-1 Human ATXN2 P3C2	TATAAAGGCTTGAGAGAAT TAA (SEQID NO: 2370)
MG21-1 sgRNA Targeting Human ATXN2	1989	MG21-1 Human ATXN2 P3D2	CAGTGAATCAAGCCAAGAC ATA (SEQID NO: 2371)
MG21-1 sgRNA Targeting Human ATXN2	1990	MG21-1 Human ATXN2 P3E2	GAATCAAGCCAAGACATAT AGA (SEQID NO: 2372)
MG21-1 sgRNA Targeting Human ATXN2	1991	MG21-1 Human ATXN2 P3F2	GACATATAGAGCAGGTAAA GGT (SEQID NO: 2373)
MG21-1 sgRNA Targeting Human ATXN2	1992	MG21-1 Human ATXN2 P3G2	GAGACTGATAATGTGGCAC ATG (SEQID NO: 2374)
MG21-1 sgRNA Targeting Human ATXN2	1993	MG21-1 Human ATXN2 P3H2	GATTTGGGAAGCTGCTGAGG ACT (SEQID NO: 2375)
MG21-1 sgRNA Targeting Human ATXN2	1994	MG21-1 Human ATXN2 P3A3	GGACTGTAGGCAACATATT GCG (SEQID NO: 2376)
MG21-1 sgRNA Targeting Human ATXN2	1995	MG21-1 Human ATXN2 P3B3	GCATATTTGGTACTGCAGA AAA (SEQID NO: 2377)
MG21-1 sgRNA Targeting Human ATXN2	1996	MG21-1 Human ATXN2 P3C3	AGCATGTGCCACATTATCA GTC (SEQID NO: 2378)
MG21-1 sgRNA Targeting Human ATXN2	1997	MG21-1 Human ATXN2 P3D3	TCATGAGCCCCGTACTGAG TTG (SEQID NO: 2379)
MG21-1 sgRNA Targeting Human ATXN2	1998	MG21-1 Human ATXN2 P3E3	CATGAGCCCCGTACTGAGT TGC (SEQID NO: 2380)
MG21-1 sgRNA Targeting Human ATXN2	1999	MG21-1 Human ATXN2 P3F3	TATAGACATGAGGATGCTG TGT (SEQID NO: 2381)
MG21-1 sgRNA Targeting Human ATXN2	2000	MG21-1 Human ATXN2 P3G3	GACATGAGGATGCTGTGTT CAA (SEQID NO: 2382)

Cat.	SEQ ID NO:	Description	Spacer/ Target Site Information
MG21-1 sgRNA Targeting Human ATXN2	2001	MG21-1 Human ATXN2 P3H3	ATGAGGATGCTGTGTTCAA ACA (SEQID NO: 2383)
MG21-1 sgRNA Targeting Human ATXN2	2002	MG21-1 Human ATXN2 P3A4	GGATGCTGTGTTCAAACAA AAA (SEQID NO: 2384)
MG21-1 sgRNA Targeting Human ATXN2	2003	MG21-1 Human ATXN2 P3B4	TCATGTCTATAGTCCTGTA ATA (SEQID NO: 2385)
MG21-1 sgRNA Targeting Human ATXN2	2004	MG21-1 Human ATXN2 P3C4	GAGCAGACGCATGCGATGT ATG (SEQID NO: 2386)
MG21-1 sgRNA Targeting Human ATXN2	2005	MG21-1 Human ATXN2 P3D4	AAAAACCCAACTCACTGGC AAA (SEQID NO: 2387)
MG21-1 sgRNA Targeting Human ATXN2	2006	MG21-1 Human ATXN2 P3E4	AAACCCAACTCACTGGCAA AGT (SEQID NO: 2388)
MG21-1 sgRNA Targeting Human ATXN2	2007	MG21-1 Human ATXN2 P3F4	AACCCAACTCACTGGCAAA GTA (SEQID NO: 2389)
MG21-1 sgRNA Targeting Human ATXN2	2008	MG21-1 Human ATXN2 P3G4	GGTAATTTGGGACATGCTG AAT (SEQID NO: 2390)
MG21-1 sgRNA Targeting Human ATXN2	2009	MG21-1 Human ATXN2 P3H4	AATTTGGGACATGCTGAAT TTG (SEQID NO: 2391)
MG21-1 sgRNA Targeting Human ATXN2	2010	MG21-1 Human ATXN2 P3A5	TGGGACATGCTGAATTTGG GAA (SEQID NO: 2392)
MG21-1 sgRNA Targeting Human ATXN2	2011	MG21-1 Human ATXN2 P3B5	ACATGCTGAATTTGGGAAA TAG (SEQID NO: 2393)
MG21-1 sgRNA Targeting Human ATXN2	2012	MG21-1 Human ATXN2 P3C5	GCATGTCCCAAATTACCAT ACA (SEQID NO: 2394)
MG21-1 sgRNA Targeting Human ATXN2	2013	MG21-1 Human ATXN2 P3D5	GTCCCAAATTACCATACAA CAA (SEQID NO: 2395)
MG21-1 sgRNA Targeting Human ATXN2	2014	MG21-1 Human ATXN2 P3E5	CAGACCTCTGAGTTGCCCT TAC (SEQID NO: 2396)
MG21-1 sgRNA Targeting Human ATXN2	2015	MG21-1 Human ATXN2 P3F5	GTTGGCTTTGCTGCTGTCCA GT (SEQID NO: 2397)
MG21-1 sgRNA Targeting Human ATXN2	2016	MG21-1 Human ATXN2 P3G5	GAGGGTGTGGAGTATGTGG GTG (SEQID NO: 2398)
MG21-1 sgRNA Targeting Human ATXN2	2017	MG21-1 Human ATXN2 P3H5	TAGCGTTAGGGTGCGCATA CTG (SEQID NO: 2399)
MG21-1 sgRNA Targeting Human ATXN2	2018	MG21-1 Human ATXN2 P3A6	TTAGGGTGCGCATACTGCT GAG (SEQID NO: 2400)
MG21-1 sgRNA Targeting Human ATXN2	2019	MG21-1 Human ATXN2 P3B6	AGGGAGCCCGTGGAAACTA AAG (SEQID NO: 2401)
MG21-1 sgRNA Targeting Human ATXN2	2020	MG21-1 Human ATXN2 P3C6	GGGAGCCCGTGGAAACTAA AGT (SEQID NO: 2402)
MG21-1 sgRNA Targeting Human ATXN2	2021	MG21-1 Human ATXN2 P3D6	GGAGCCCGTGGAAACTAAA GTG (SEQID NO: 2403)
MG21-1 sgRNA Targeting Human ATXN2	2022	MG21-1 Human ATXN2 P3E6	GAGCCCGTGGAAACTAAAG TGA (SEQID NO: 2404)
MG21-1 sgRNA Targeting Human ATXN2	2023	MG21-1 Human ATXN2 P3F6	AGCCCGTGGAAACTAAAGT GAA (SEQID NO: 2405)
MG21-1 sgRNA Targeting Human ATXN2	2024	MG21-1 Human ATXN2 P3G6	CCGTGGAAACTAAAGTGAA AGA (SEQID NO: 2406)
MG21-1 sgRNA Targeting Human ATXN2	2025	MG21-1 Human ATXN2 P3H6	CCTCAGCCTTCAGCTACCC CCA (SEQID NO: 2407)
MG21-1 sgRNA Targeting Human ATXN2	2026	MG21-1 Human ATXN2 P3A7	GTCATCCTGCACCCAGTCC TGT (SEQID NO: 2408)
MG21-1 sgRNA Targeting Human ATXN2	2027	MG21-1 Human ATXN2 P3B7	CTGCACCCAGTCCTGTTCA GGT (SEQID NO: 2409)
MG21-1 sgRNA Targeting Human ATXN2	2028	MG21-1 Human ATXN2 P3C7	GGTTGGTATACGCCGGCTG AAC (SEQID NO: 2410)
MG21-1 sgRNA Targeting Human ATXN2	2029	MG21-1 Human ATXN2 P3D7	GAACGTGAGAAGGATGGA TCGT (SEQID NO: 2411)

Cat.	SEQ ID NO:	Description	Spacer/ Target Site Information
MG21-1 sgRNA Targeting Human ATXN2	2030	MG21-1 Human ATXN2 P3E7	GTAAAGACAGTCTGTTGTG CTG (SEQID NO: 2412)
MG21-1 sgRNA Targeting Human ATXN2	2031	MG21-1 Human ATXN2 P3F7	TATTCTGTGGCGACTGCGT GTT (SEQID NO: 2413)
MG21-1 sgRNA Targeting Human ATXN2	2032	MG21-1 Human ATXN2 P3G7	ACTGCTGCTGTGGACTGGC CAG (SEQID NO: 2414)
MG21-1 sgRNA Targeting Human ATXN2	2033	MG21-1 Human ATXN2 P3H7	GTGCTGATGGTGCTGCAAA GCG (SEQID NO: 2415)
MG21-1 sgRNA Targeting Human ATXN2	2034	MG21-1 Human ATXN2 P3A8	TGCTGATGGTGCTGCAAAG CGA (SEQID NO: 2416)
MG21-1 sgRNA Targeting Human ATXN2	2035	MG21-1 Human ATXN2 P3B8	CTGATGGTGCTGCAAAGCG ACA (SEQID NO: 2417)
MG21-1 sgRNA Targeting Human ATXN2	2036	MG21-1 Human ATXN2 P3C8	TGCAAAGCGACAGGAAAG AATT (SEQID NO: 2418)
MG21-1 sgRNA Targeting Human ATXN2	2037	MG21-1 Human ATXN2 P3D8	GCAAAGCGACAGGAAAGA ATTG (SEQID NO: 2419)
MG21-1 sgRNA Targeting Human ATXN2	2038	MG21-1 Human ATXN2 P3E8	CAAAGCGACAGGAAAGAA TTGA (SEQID NO: 2420)
MG21-1 sgRNA Targeting Human ATXN2	2039	MG21-1 Human ATXN2 P3F8	TGCCTCCAACACGCAGTCG CCA (SEQID NO: 2421)
MG21-1 sgRNA Targeting Human ATXN2	2040	MG21-1 Human ATXN2 P3G8	CACCCACATGGCCACGT ACC (SEQID NO: 2422)
MG21-1 sgRNA Targeting Human ATXN2	2041	MG21-1 Human ATXN2 P3H8	TCGTCATTAGCATCATTGG CGC (SEQID NO: 2423)
MG21-1 sgRNA Targeting Human ATXN2	2042	MG21-1 Human ATXN2 P3A9	TCATTGGCGCATGGGCAGT TGG (SEQID NO: 2424)
MG21-1 sgRNA Targeting Human ATXN2	2043	MG21-1 Human ATXN2 P3B9	TTGGCGCATGGGCAGTTGG ATG (SEQID NO: 2425)
MG21-1 sgRNA Targeting Human ATXN2	2044	MG21-1 Human ATXN2 P3C9	CATTCCTGACTGTACATGA GCC (SEQID NO: 2426)
MG21-1 sgRNA Targeting Human ATXN2	2045	MG21-1 Human ATXN2 P3D9	CCTGACTGTACATGAGCCT GAA (SEQID NO: 2427)
MG21-1 sgRNA Targeting Mouse ATXN2	2051	mATXN21-P1A5	ACTTGGTACTTGATGCTGC ACA (SEQID NO: 2428)
MG21-1 sgRNA Targeting Mouse ATXN2	2052	mATXN21-P1F12	AAGAAACAGAGAAGTCCT ATCC (SEQID NO: 2429)
MG21-1 sgRNA Targeting Mouse ATXN2	2053	mATXN21-P2A1	CTCAGACCAAAGAGTAGTT AAT (SEQID NO: 2430)
MG21-1 sgRNA Targeting Mouse ATXN2	2054	mATXN21-P2G1	ACTGGTAGCGAGAAGGTGG GCG (SEQID NO: 2431)
MG21-1 sgRNA Targeting Mouse ATXN2	2055	mATXN21-P2F6	GGCAGAACTCTCCTGCAGG GAG (SEQID NO: 2432)

CLAIMS

WHAT IS CLAIMED IS:

1. An engineered nuclease system comprising:
 - a) an endonuclease comprising a sequence having at least 70% sequence identity to any one of SEQ ID NOs: 47-52, 659-698, 983-1010, 1263-1282, 1158-1159, and 975-981; and
 - b) an engineered guide polynucleotide configured to form a complex with the endonuclease and to hybridize to a target nucleic acid sequence.
2. The engineered nuclease system of claim 1, wherein the endonuclease comprises a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 47-52, 659-698, 983-1010, 1263-1282, 1158-1159, and 975-981.
3. The engineered nuclease system of claim 1, wherein the endonuclease comprises a sequence having 90% sequence identity to any one of SEQ ID NOs: 47-52, 659-698, 983-1010, 1263-1282, 1158-1159, and 975-981.
4. The engineered nuclease system of any one of claims 1-3, wherein the engineered guide polynucleotide comprises a crRNA and a tracrRNA.
5. The engineered nuclease system of claim 4, wherein the tracrRNA comprises a sequence having at least 90% sequence identity to any one of SEQ ID NOs: 710-722, 726-744, 745-767, 699, 700-702, and 703-709.
6. The engineered nuclease system of claim 4, wherein the tracrRNA comprises a sequence having 100% sequence identity to any one of SEQ ID NOs: 710-722, 726-744, 745-767, 699, 700-702, and 703-709.
7. The engineered nuclease system of any one of claims 1-6, wherein the engineered guide polynucleotide is a single guide nucleic acid.
8. The engineered nuclease system of any one of claims 1-6, wherein the engineered guide polynucleotide is a dual guide nucleic acid.
9. The engineered nuclease system of any one of claims 1-8, wherein the engineered guide polynucleotide is RNA.
10. The engineered nuclease system of any one of claims 1-9, wherein the endonuclease is not a Cas9 endonuclease.
11. The engineered nuclease system of any one of claims 1-10, wherein the endonuclease has less than 80% identity to a Cas9 endonuclease.
12. The engineered nuclease system of any one of claims 1-11, wherein the endonuclease binds non-covalently to the engineered guide polynucleotide.
13. The engineered nuclease system of any one of claims 1-11, wherein the endonuclease is covalently linked to the engineered guide polynucleotide.

14. An engineered nuclease system comprising:
- a) an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NO: 47; and
 - b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 1137-1144, 1283-1392, 1502-1509, and 1786-2045.
15. An engineered nuclease system comprising:
- a) an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NO: 48; and
 - b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to SEQ ID NOs: 1153-1156, 1393-1493, 1494-1501, and 1510-1525.
16. An engineered nuclease system comprising:
- a) an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263; and
 - b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 57-58, 77-88, 101-119, 139-150, 163-181, 201-212, 225-243, 343-374, 407-413, 421-433, 447-453, 461-472, 485-491, 499-511, 525-531, 539-551, 565-578, 593-625, 710-722, 927-942, 945-950, 961-962, 969-970, 1906-1931, 1083-1094, 1096-1102, 1113-1122, 1123-1137, and 2433-2434.
17. An engineered nuclease system comprising:
- a) an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NO: 51 or SEQ ID NO: 1264; and
 - b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to SEQ ID NO: 61.
18. An engineered nuclease system comprising:
- a) an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NOs: 659, 660, 1158, 1159, and 1267-1277; and
 - b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide

polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 726-744, 843-880, 1011-1055, and 1145-1152.

19. An engineered nuclease system comprising:

a) an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 661-678 and 1278-1282; and

b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 745-767 and 881-926.

20. An engineered nuclease system comprising:

a) an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NO: 659; and

b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 699 and 973-974.

21. An engineered nuclease system comprising:

a) an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 696-698; and

b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 700-702 and 957-960.

22. An engineered nuclease system comprising:

a) an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 975-981; and

b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 703-709, 943-944, 951-956, 965-968, and 971-972.

23. The engineered nuclease system of any one of claims 14-22, wherein the engineered guide polynucleotide is a single guide nucleic acid.

24. The engineered nuclease system of any one of claims 14-22, wherein the engineered guide polynucleotide is a dual guide nucleic acid.

25. The engineered nuclease system of any one of claims 14-24, wherein the engineered guide polynucleotide is RNA.
26. The engineered nuclease system of any one of claims 1-25, wherein the endonuclease is not a Cas9 endonuclease.
27. The engineered nuclease system of any one of claims 14-26, wherein the endonuclease has less than 80% identity to a Cas9 endonuclease.
28. The engineered nuclease system of any one of claims 14-27, wherein the endonuclease binds non-covalently to the engineered guide polynucleotide.
29. The engineered nuclease system of any one of claims 14-27, wherein the endonuclease is covalently linked to the engineered guide polynucleotide.
30. The engineered nuclease system of any one of claims 14-27, wherein the endonuclease is fused to the engineered guide polynucleotide.
31. A method for modifying a target nucleic acid sequence comprising contacting the target nucleic acid sequence using the engineered nuclease system of any one of claims 1-30.
32. The method of claim 31, wherein modifying the target nucleic acid sequence comprises binding, nicking, or cleaving, the target nucleic acid sequence.
33. The method of any one of claims 31-32, wherein the target nucleic acid sequence comprises genomic DNA, viral DNA, viral RNA, or bacterial DNA.
34. The method of any one of claims 31-33, wherein the modification is *in vitro*.
35. The method of any one of claims 31-33, wherein the modification is *in vivo*.
36. The method of any one of claims 31-33, wherein the modification is *ex vivo*.
37. The method of any one of claims 31-36, wherein the target nucleic acid sequence comprises a sequence of any one of SEQ ID NOs: 1161-1262, 1526-1785, and 2046-2050.
38. A method of modifying a target nucleic acid sequence in a mammalian cell comprising contacting the mammalian cell using the engineered nuclease system of any one of claims 1-30.
39. The method of claim 38, further comprising selecting cells comprising the modification.
40. A method of modifying a hydroxyacid oxidase 1 (HAO1) gene comprising contacting the HAO1 gene using an engineered nuclease system comprising:
 - a) an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NO: 47; and
 - b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 1283-1392 and 1502-1509.

41. The method of claim 40, wherein the engineered guide polynucleotide comprises a sequence of any one of SEQ ID NOs: 1283-1392 and 1502-1509.
42. A method of modifying ATPase copper transporting beta (ATP7B) comprising contacting ATP7B using an engineered nuclease system comprising:
- a) an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263; and
 - b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence.
43. The method of claim 42, wherein the engineered guide polynucleotide comprises a sequence of any one of SEQ ID NOs: 1056-1081.
44. The method of claim 42, wherein the target nucleic acid sequence comprises a sequence of any one of SEQ ID NOs: 1206-1231.
45. A method of modifying adeno-associated virus integration site 1 (AAVS1) comprising contacting AAVS1 using an engineered nuclease system comprising:
- a) an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1010, 1263, 47, 659, 660, 1158, 1159, and 1267-1277; and
 - b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 710-722.
46. The method of claim 45, wherein the engineered guide polynucleotide comprises a sequence of any one of SEQ ID NOs: 57-58, 101-119, 163-181, 225-243, 343-374, 21-433, 461-472, 499-511, 539-551, 1082, 1098-1102, 1393-1493, 1020-1049, and 1051-1055.
47. The method of claim 45, wherein the target nucleic acid sequence comprises a sequence of any one of SEQ ID NOs: 59-60, 120-138, 182-200, 243-262, 375-406, 434-446, 473-484, 512-524, 552-564, 1170-1205, 1232, and 1248-1252.
48. A method of modifying T cell receptor alpha constant (TRAC) comprising contacting TRAC using an engineered nuclease system comprising:
- a) an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1010, 1263, 48, 659, 660, 1158, 1159, and 1267-1277; and
 - b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide

polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 710-722.

49. The method of claim 48, wherein the engineered guide polynucleotide comprises a sequence of any one of SEQ ID NOs: 61, 77-88, 139-150, 201-212, 407-413, 447-453, 485-491 525-531, 1096-1097, 1011-1019, 1050, 1494-1501, and 1510-1525.

50. The method of claim 48, wherein the target nucleic acid sequence comprises a sequence of any one of SEQ ID NOs: 62, 89-100, 151-162, 213-224, 414-420, 454-460, 492-489, 532-538, 1161-1168, and 1200.

51. A method of modifying an albumin gene comprising contacting the albumin gene using an engineered nuclease system comprising:

a) an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NO: 49-51, 679-694, 983-1002, 1003-1010, and 1263; and

b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence.

52. The method of claim 51, wherein the engineered guide polynucleotide comprises a sequence having any SEQ ID NO: 1083-1094.

53. The method of claim 48, wherein the target nucleic acid sequence comprises a sequence having SEQ ID NO: 1233-1244.

54. A method of modifying a beta-2-microglobulin (B2M) gene comprising contacting the B2M gene using an engineered nuclease system comprising:

a) an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263; and

b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 710-722.

55. The method of claim 54, wherein the engineered guide polynucleotide comprises a sequence of any one of SEQ ID NOs: 565-578 and 1095.

56. The method of claim 54, wherein the target nucleic acid sequence comprises a sequence of any one of SEQ ID NOs: 579-592 and 1245.

57. A method of modifying a hemoglobin subunit beta (HBB) gene comprising contacting the HBB gene using an engineered nuclease system comprising:

a) an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263; and

b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence, the engineered guide polynucleotide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 710-722.

58. The method of claim 57, wherein the engineered guide polynucleotide comprises a sequence of any one of SEQ ID NOs: 593-625.

59. The method of claim 57, wherein the target nucleic acid sequence comprises a sequence of any one of SEQ ID NOs: 626-658.

60. A method of modifying a phenylalanine hydroxylase (PAH) gene comprising contacting the PAH gene using an engineered nuclease system comprising:

a) an endonuclease comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 49-51, 679-694, 983-1002, 1003-1010, and 1263; and

b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence.

61. The method of claim 60, wherein the engineered guide polynucleotide comprises a sequence of any one of SEQ ID NOs: 1113-1122.

62. The method of claim 60, wherein the target nucleic acid sequence comprises a sequence of any one of SEQ ID NOs: 1253-1262.

63. A method of modifying an ataxin 2 (ATXN2) gene comprising contacting the ATXN2 gene using an engineered nuclease system comprising:

c) an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NO: 47; and

d) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence.

64. The method of claim 63, wherein the engineered guide polynucleotide comprises a sequence of any one of SEQ ID NOs: 1786-2045, and 2051-2055.

65. The method of claim 63, wherein the target nucleic acid sequence comprises a sequence of any one of SEQ ID NOs: 1526-1785, and 2046-2050.

66. A cell comprising the engineered nuclease system of any one of claims 1-30.

67. The cell of claim 66, wherein the cell is a eukaryotic cell.

68. The cell of claim 66, wherein the cell is a mammalian cell.

69. The cell of claim 66, wherein the cell is an immortalized cell.

70. The cell of claim 66, wherein the cell is an insect cell.

71. The cell of claim 66, wherein the cell is a yeast cell.

72. The cell of claim 66, wherein the cell is a plant cell.

- 73. The cell of claim 66, wherein the cell is a fungal cell.
- 74. The cell of claim 66, wherein the cell is a prokaryotic cell.
- 75. The cell of claim 66, wherein the cell is an A549, HEK-293, HEK-293T, BHK, CHO, HeLa, MRC5, Sf9, Cos-1, Cos-7, Vero, BSC 1, BSC 40, BMT 10, WI38, HeLa, Saos, C2C12, L cell, HT1080, HepG2, Huh7, K562, primary cell, or a derivative thereof.
- 76. The cell of claim 66, wherein the cell is an engineered cell.
- 77. The cell of claim 66, wherein the cell is a stable cell.

MG71-2 optimization

20nt 22nt 24nt

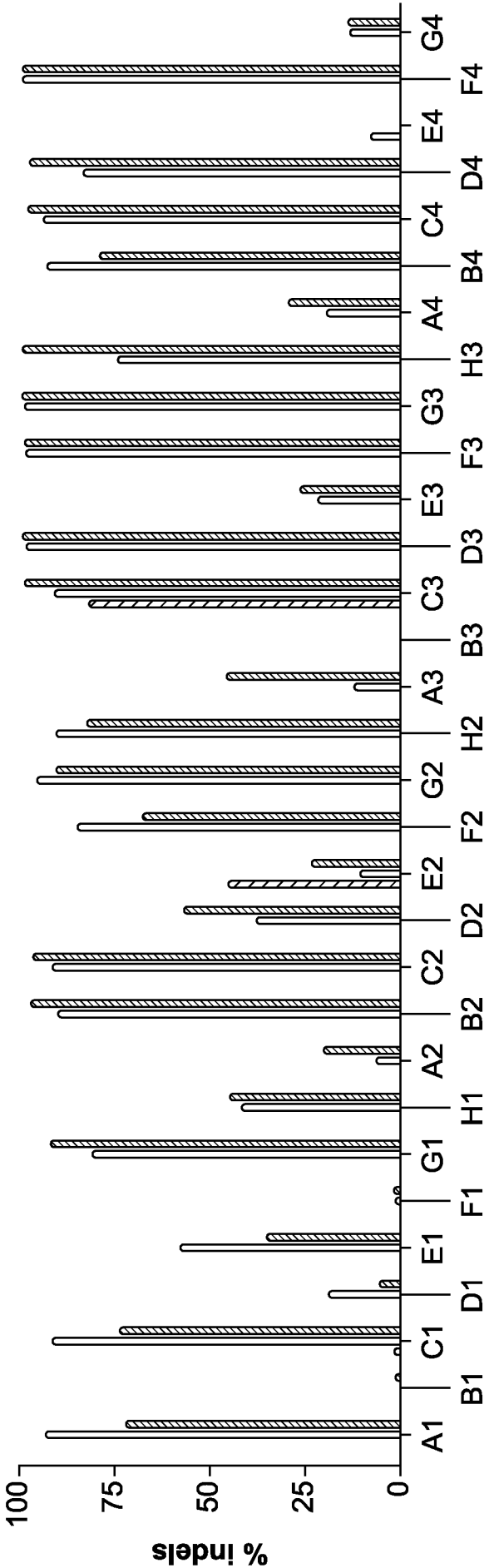


FIG. 1

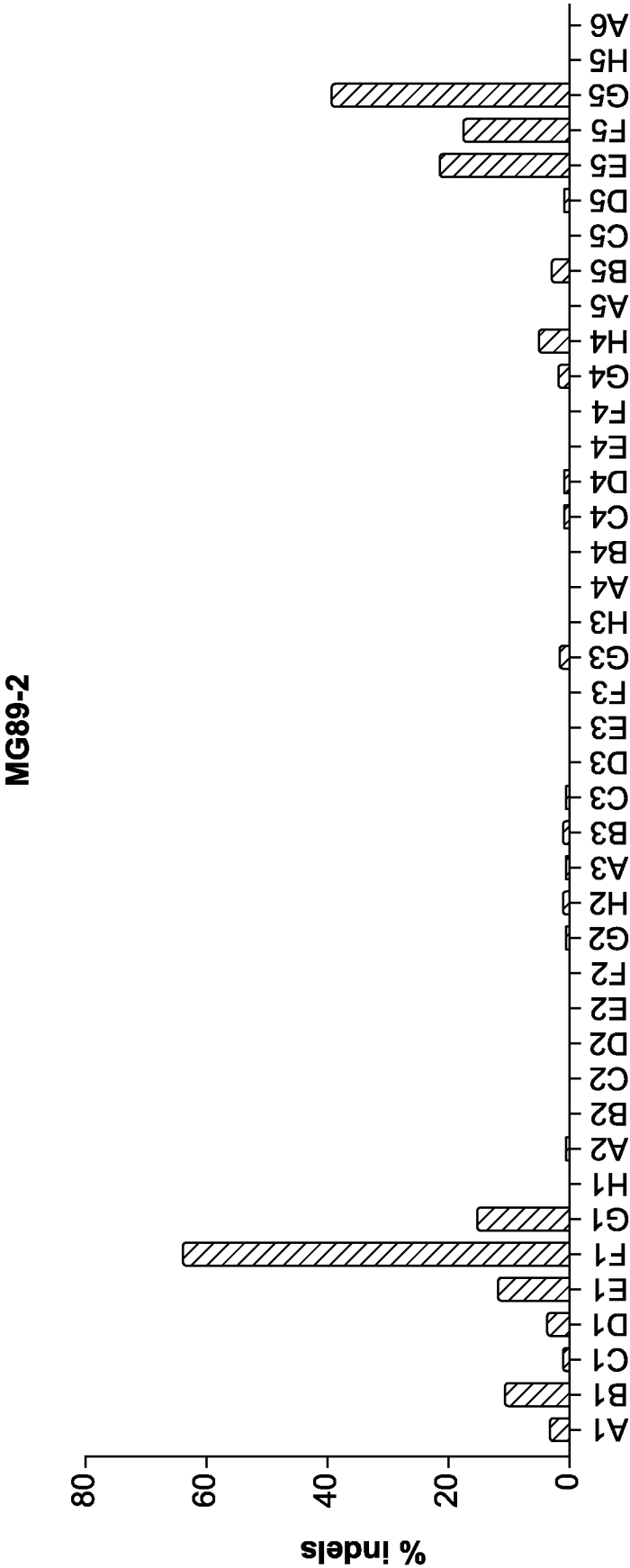


FIG. 2

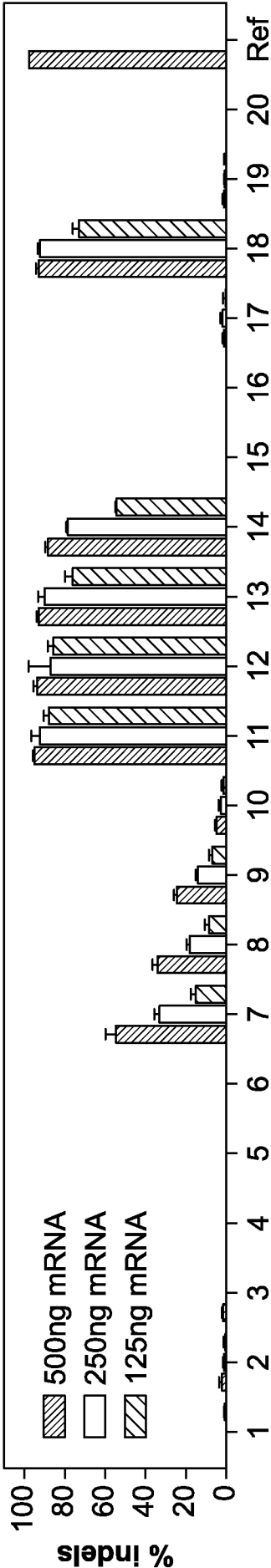


FIG. 3A

MG71-2 guide optimization round 2

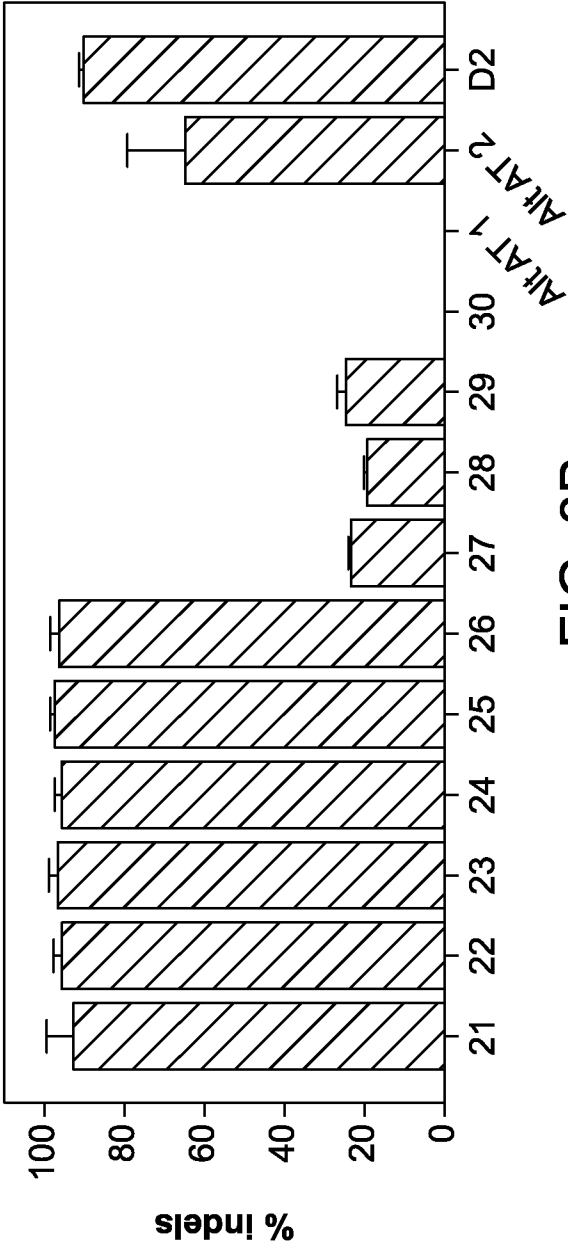


FIG. 3B

MG71-2 guide optimization at TRAC and AAVS1

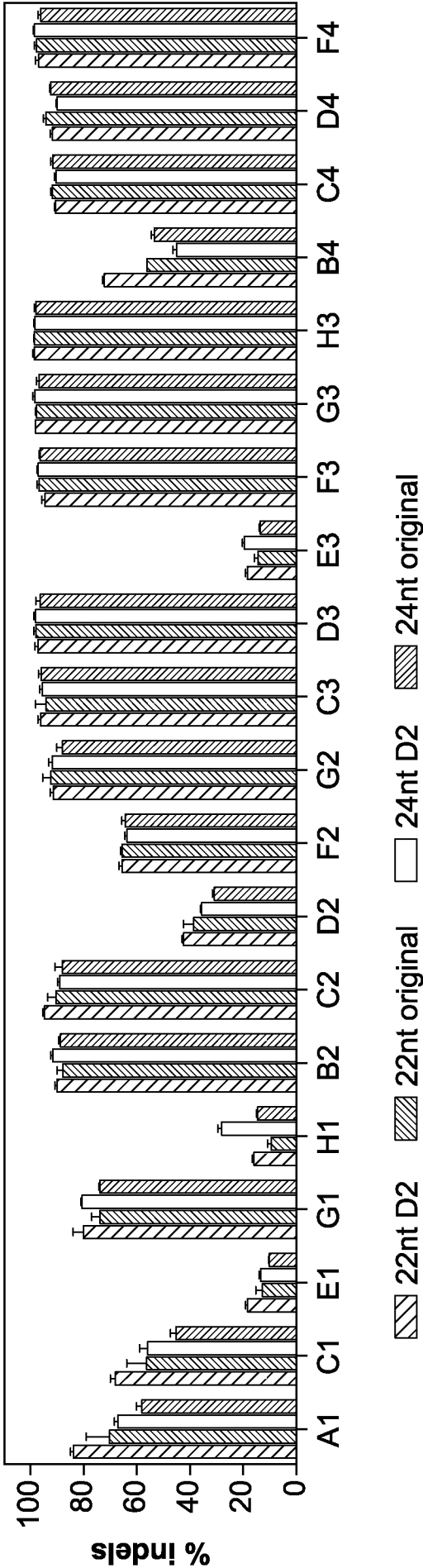


FIG. 4

MG71-2 hB2M exon 2

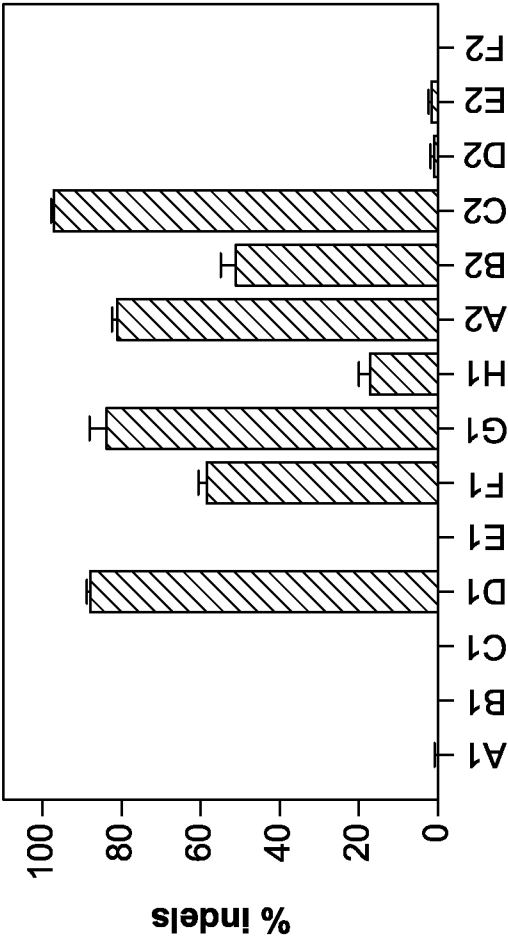


FIG. 5A

MG71-2 HBB exon 3 results

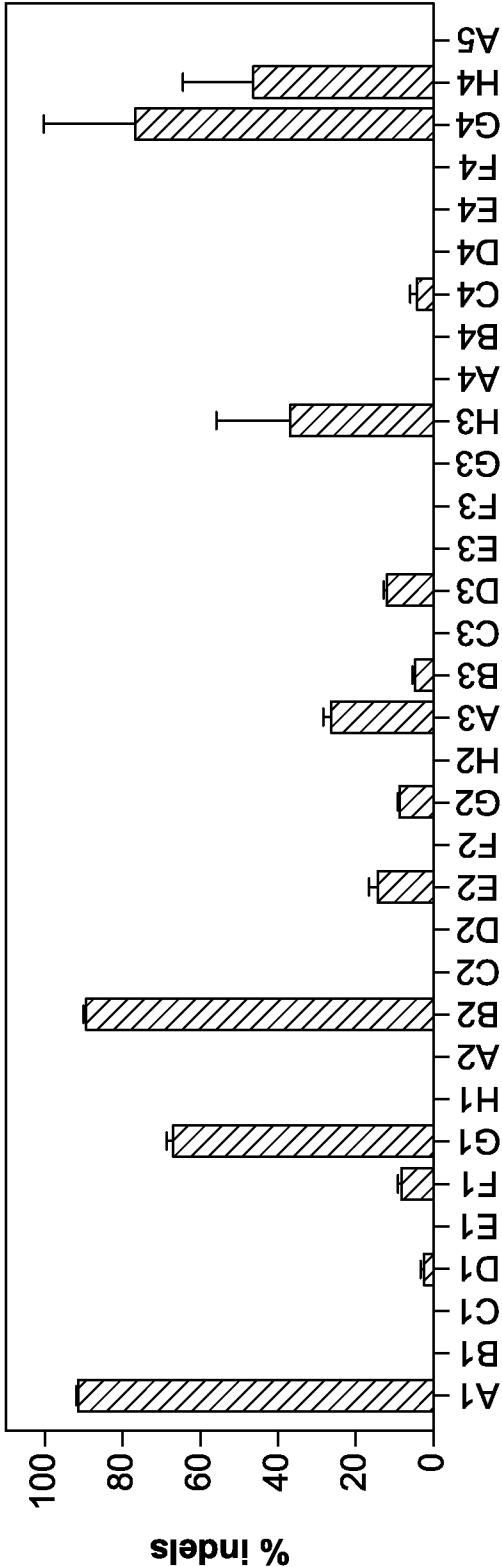


FIG. 5B

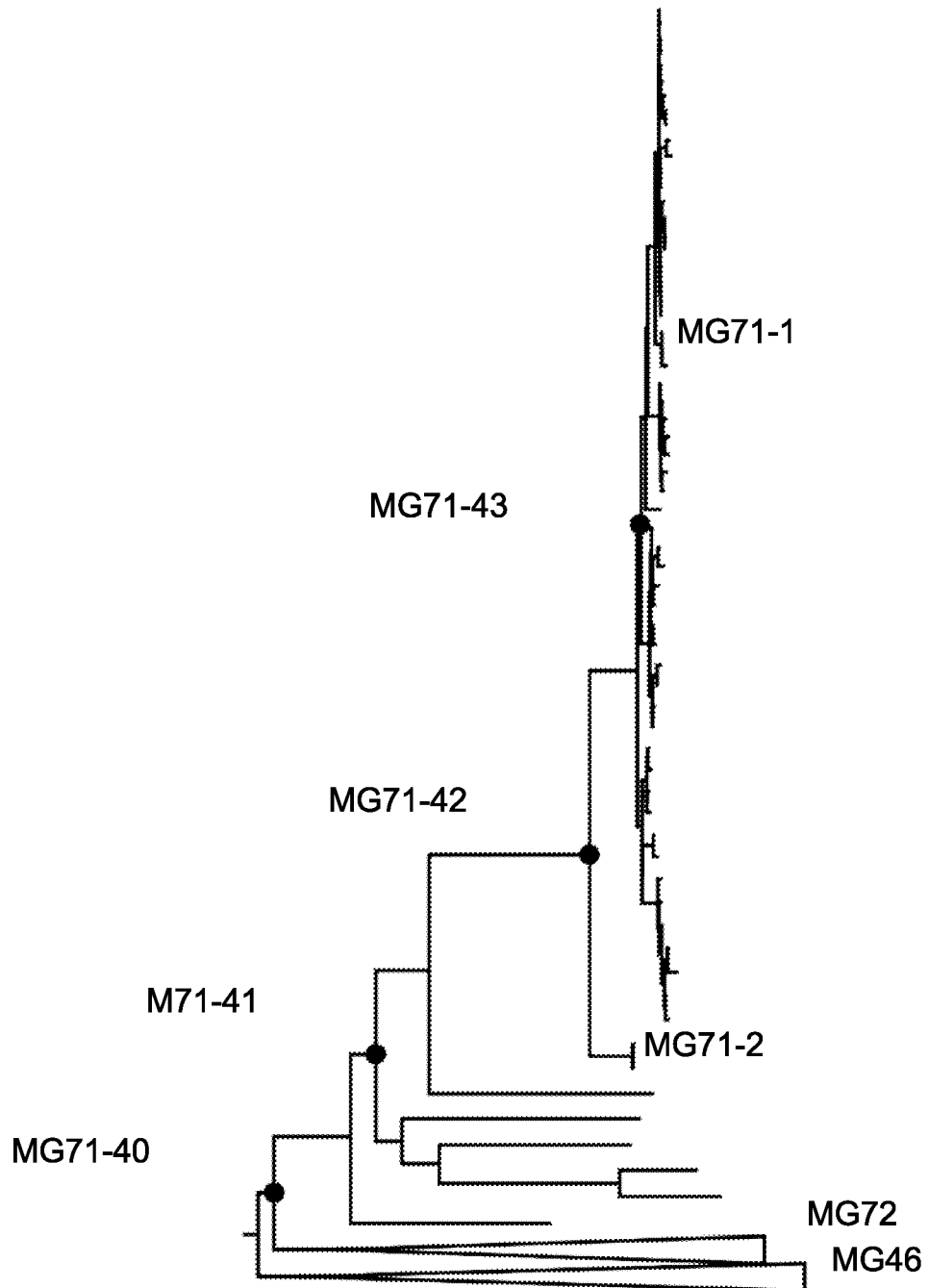


FIG. 6

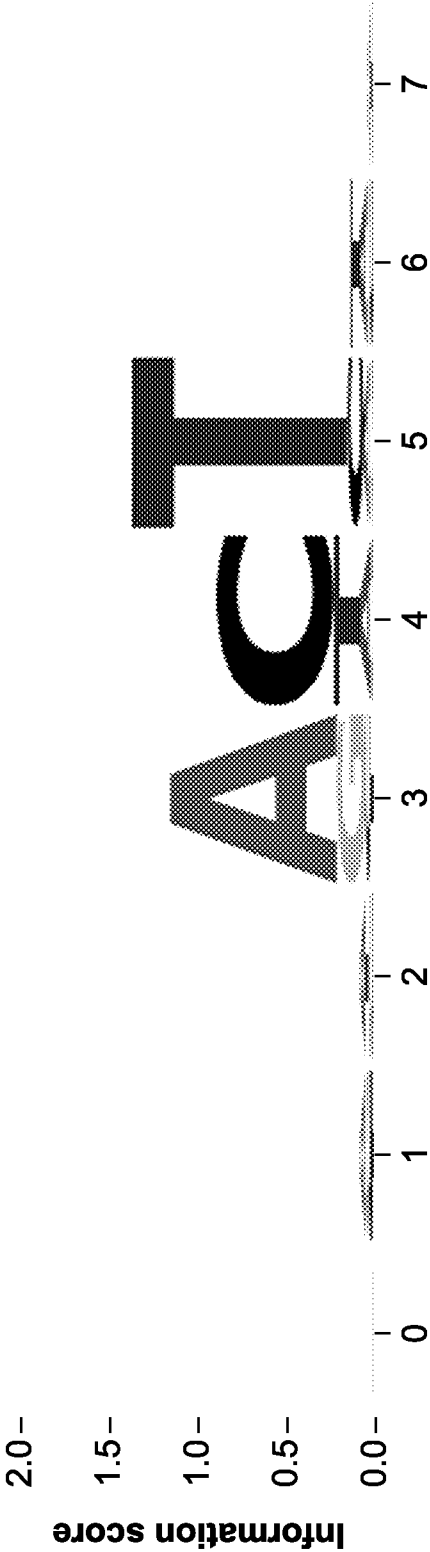


FIG. 7

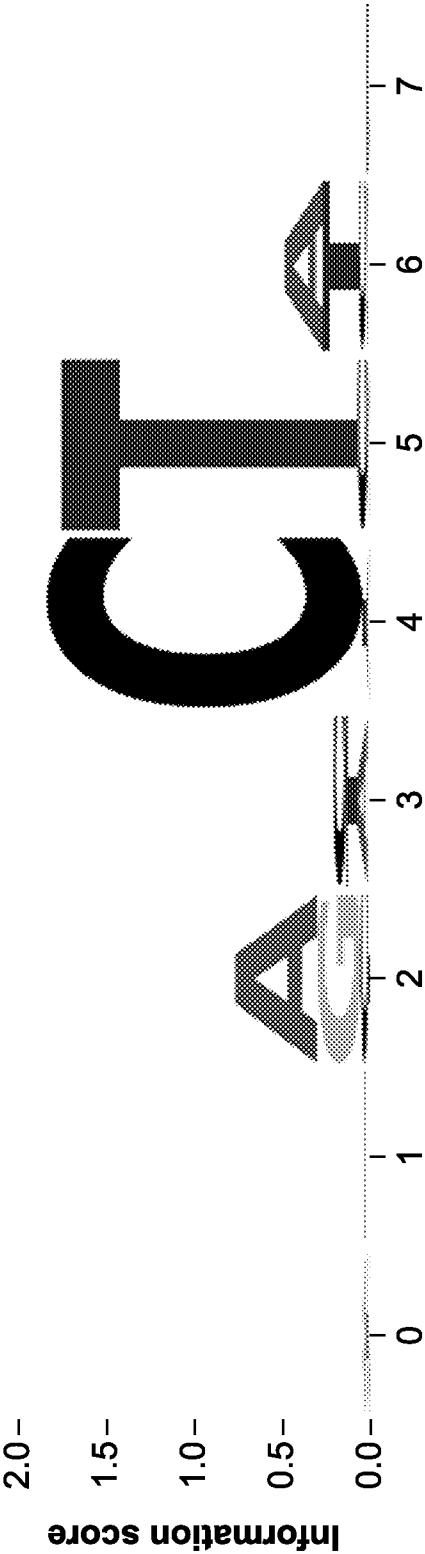


FIG. 8

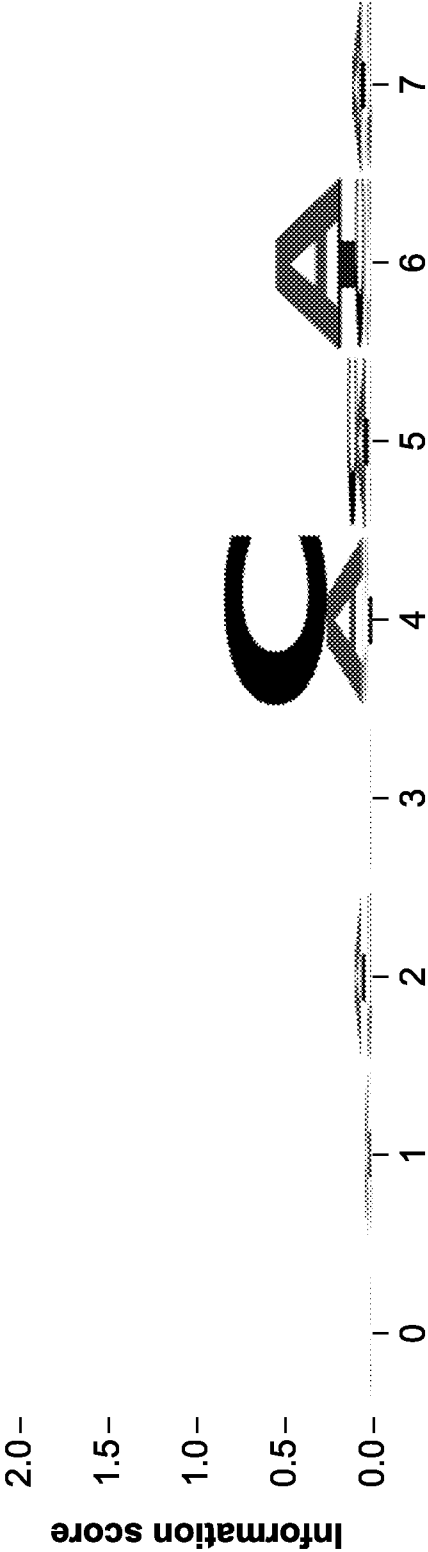


FIG. 9

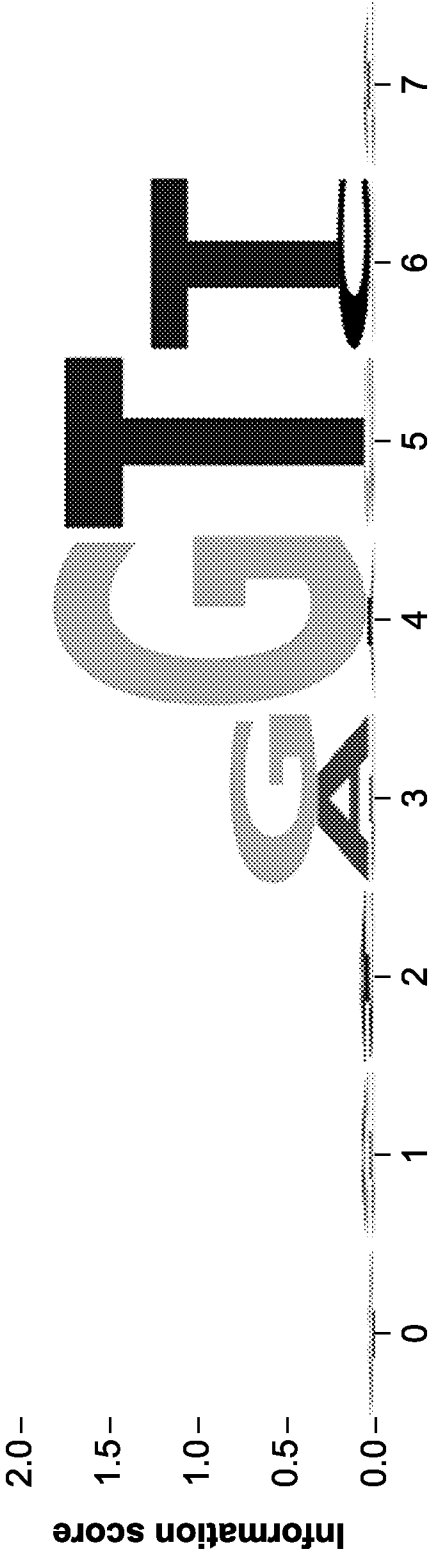


FIG. 10

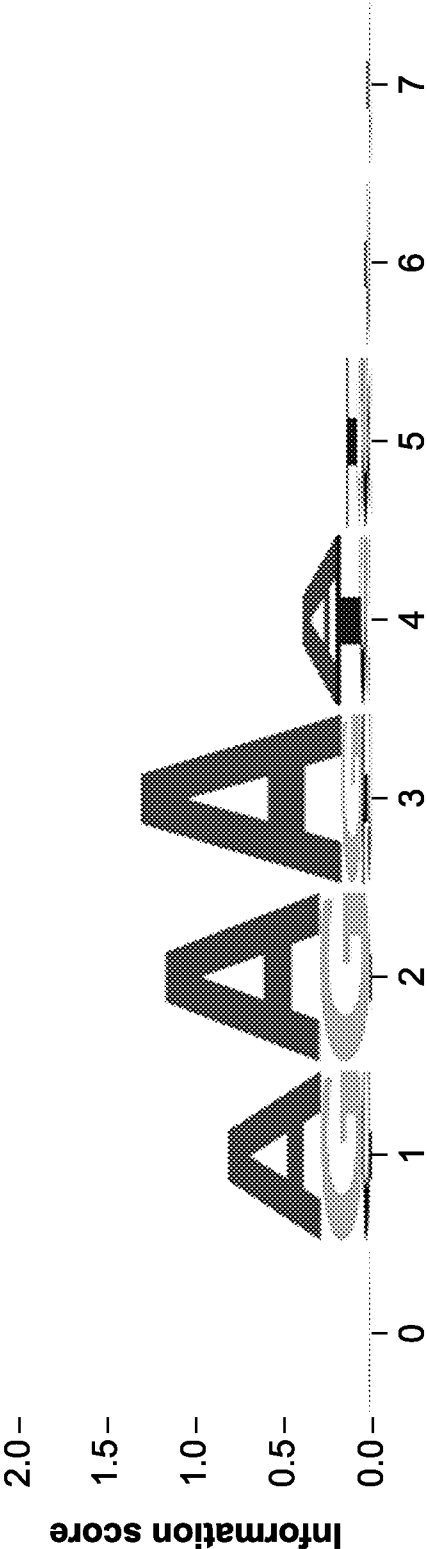


FIG. 11

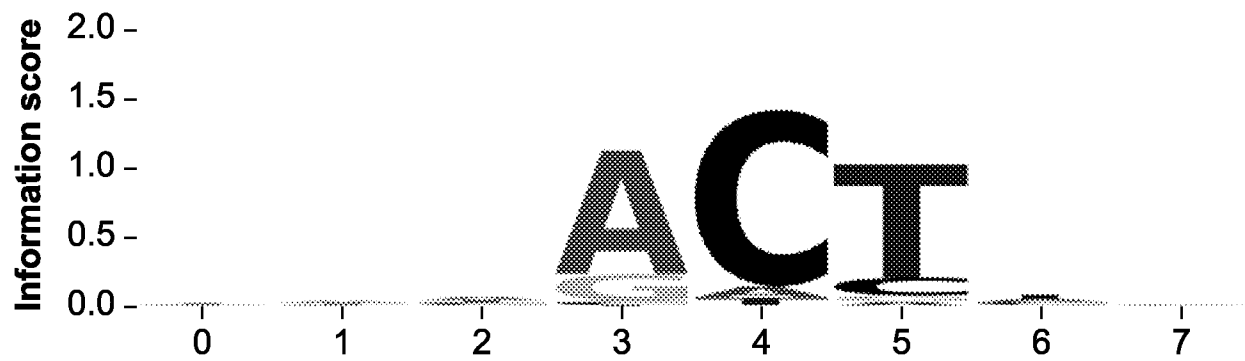


FIG. 12A

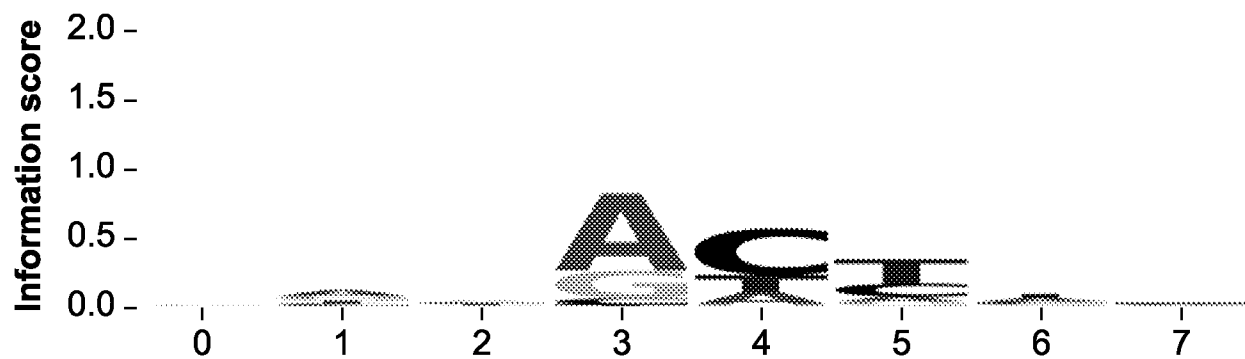


FIG. 12B

hAAVS1 editing with MG71-2

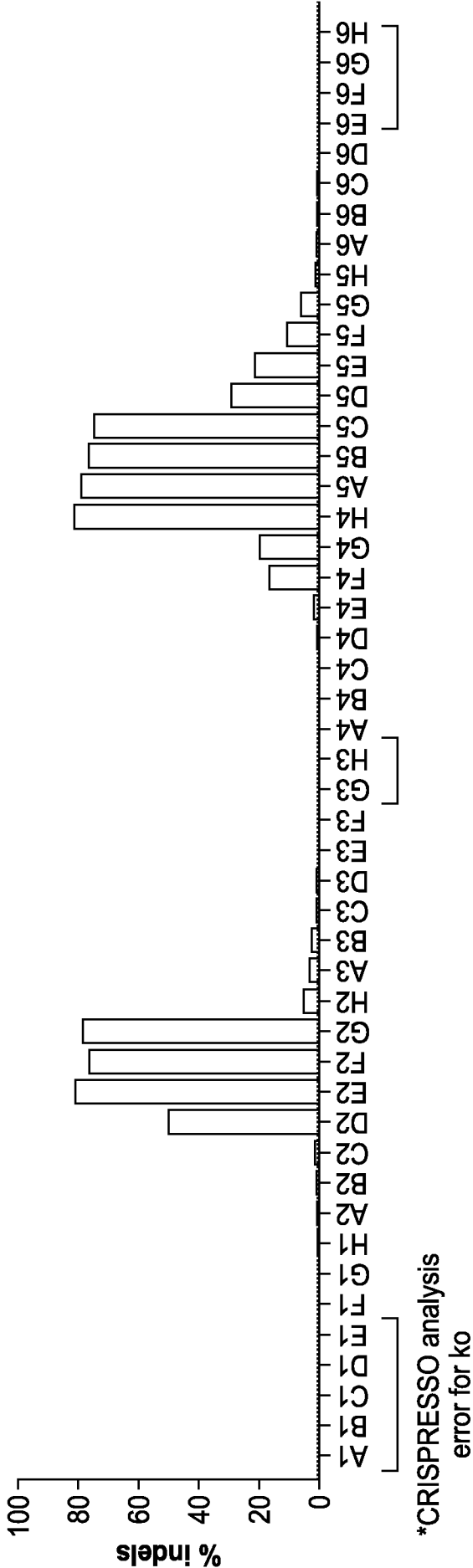
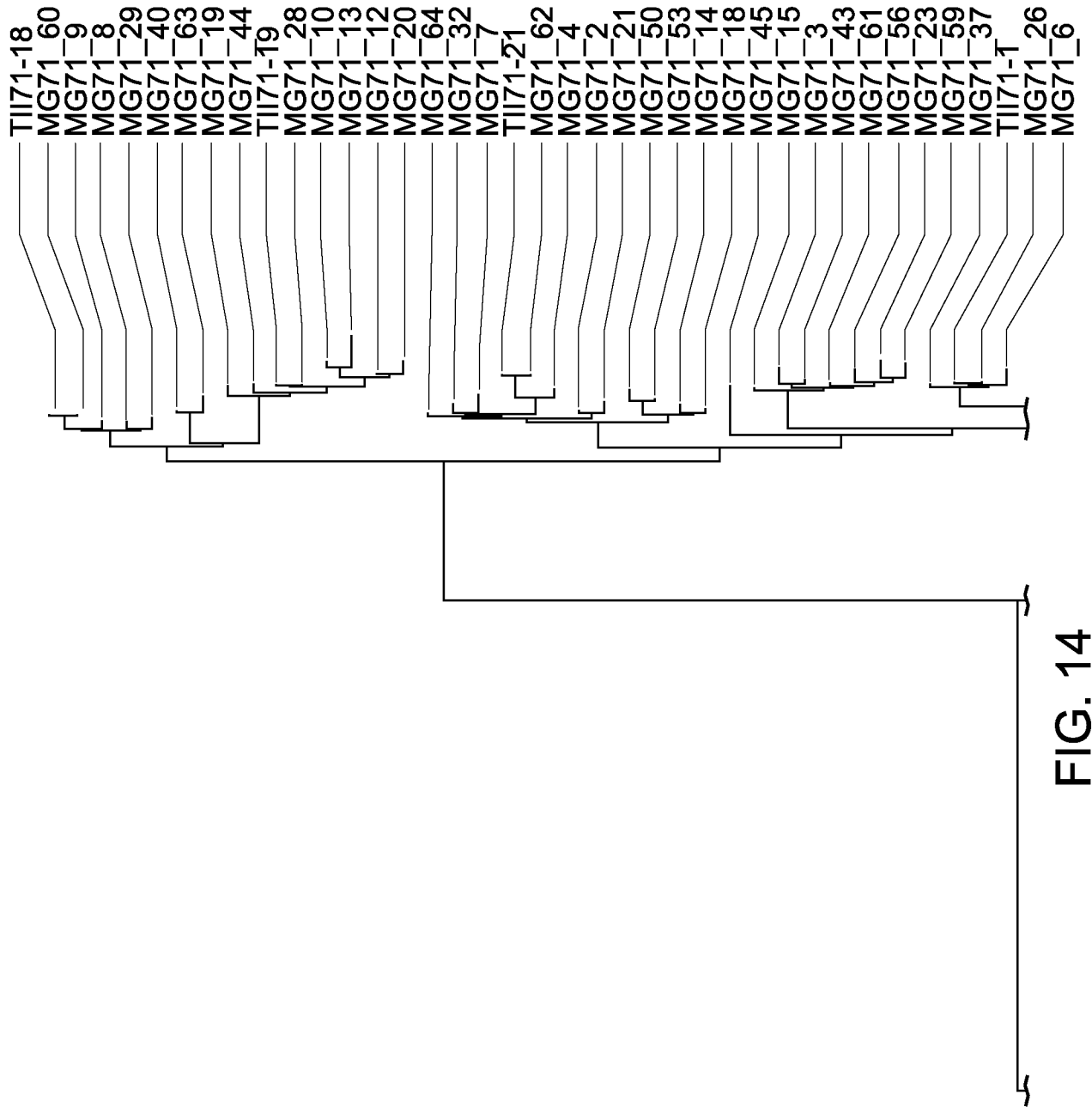


FIG. 13



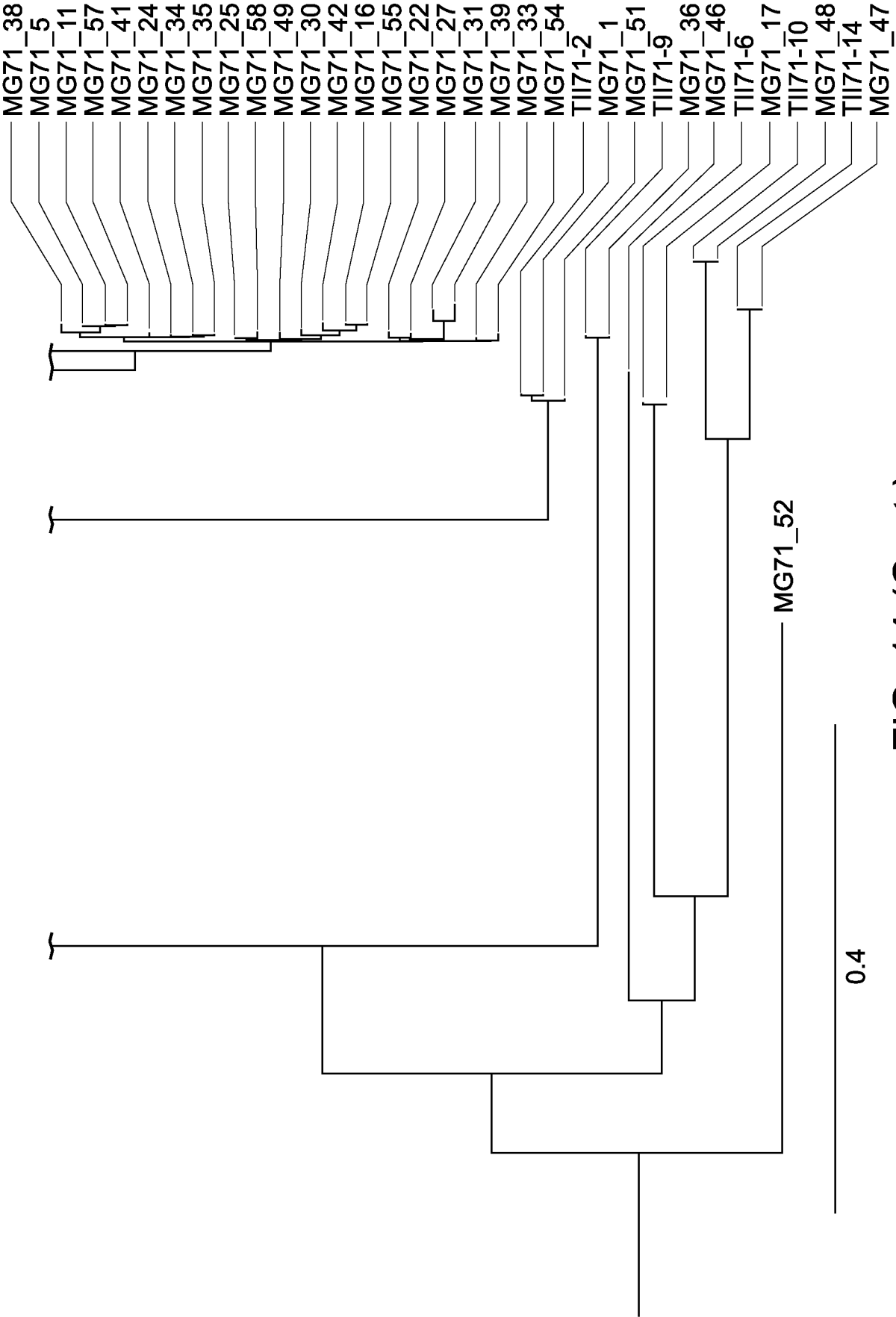


FIG. 14 (Cont.)

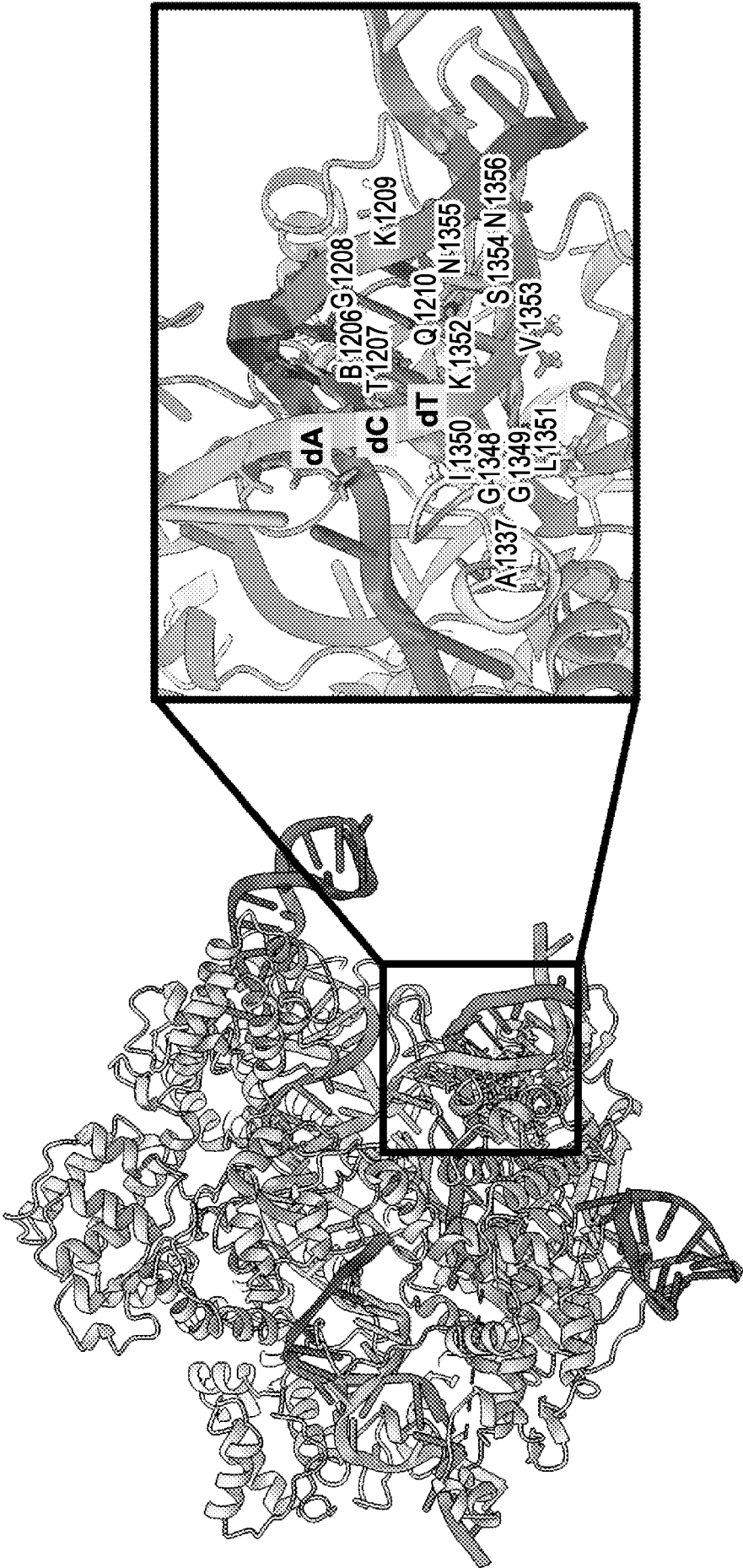


FIG. 15

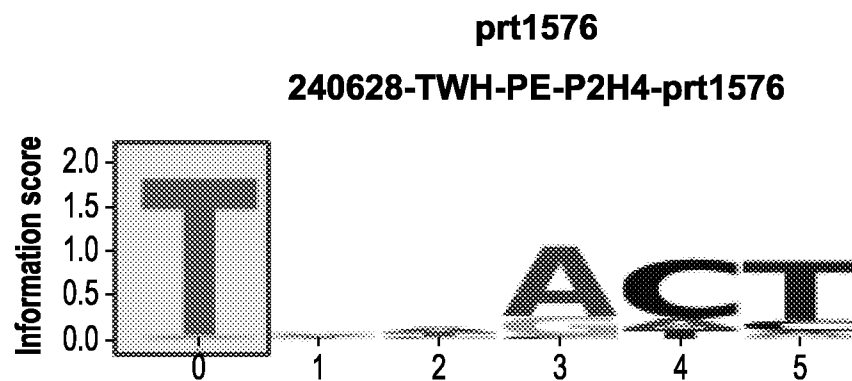
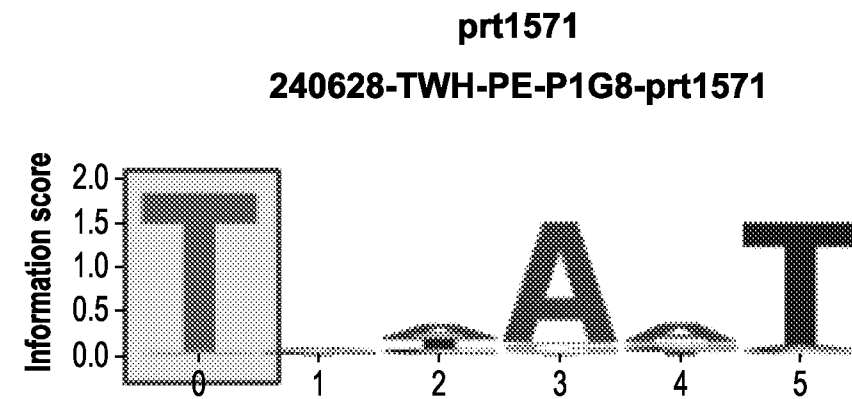
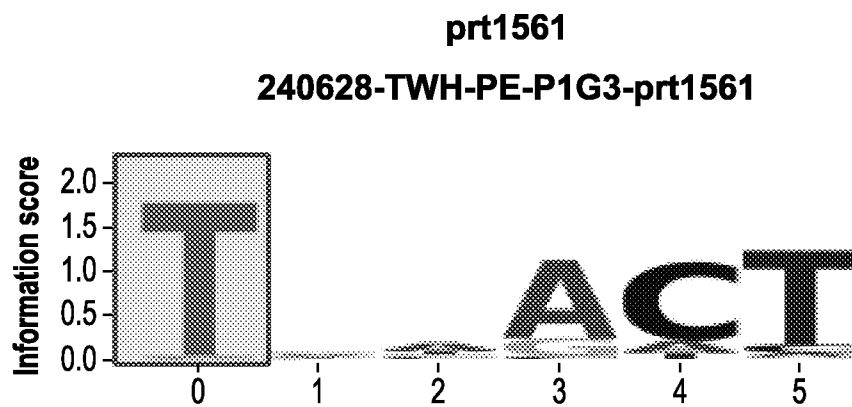
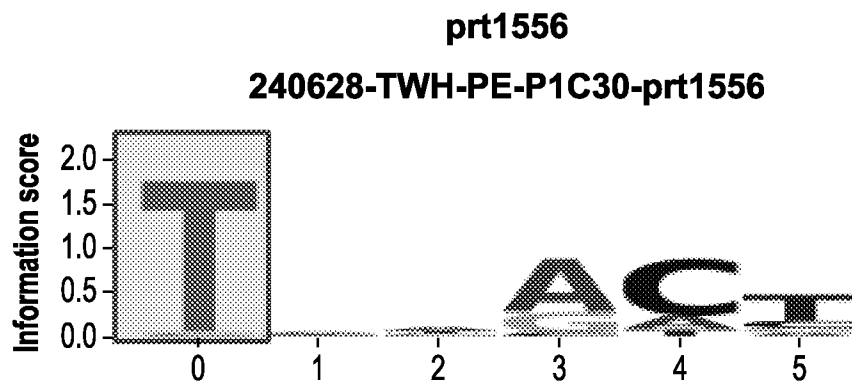
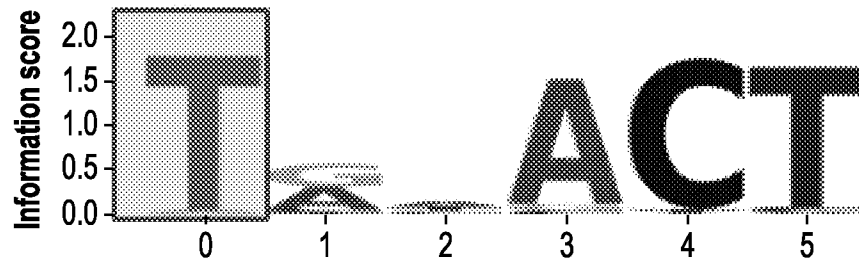


FIG. 16

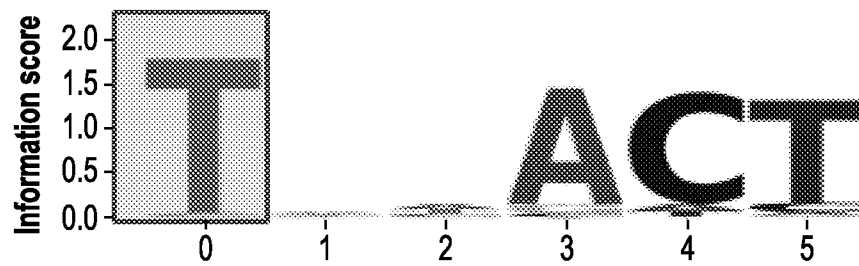
prt1557

240628-TWH-PE-P1C11-prt1557



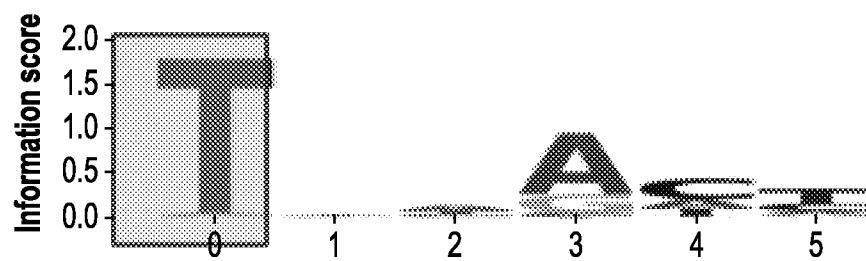
prt1562

240628-TWH-PE-P1G4-prt1562



prt1572

240628-TWH-PE-P1G9-prt1572



prt22902

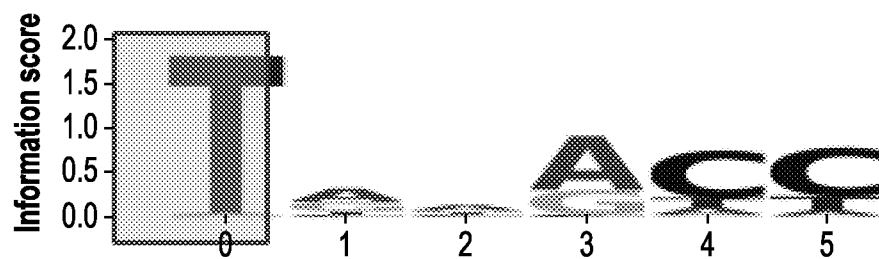
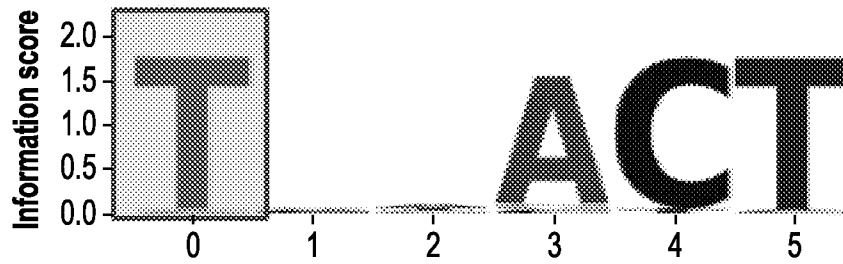


FIG. 16 (Cont. 1)

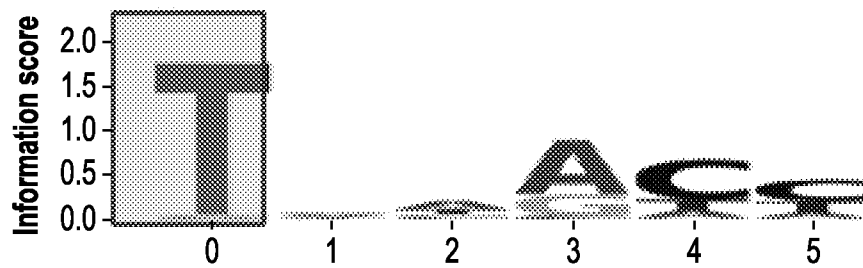
prt1558

240628-TWH-PE-P1C12-prt1558



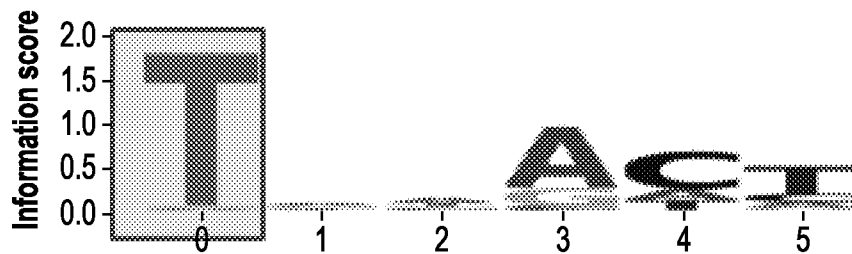
prt1568

240628-TWH-PE-P1G5-prt1568



prt1573

240628-TWH-PE-P1G10-prt1573



prt22903

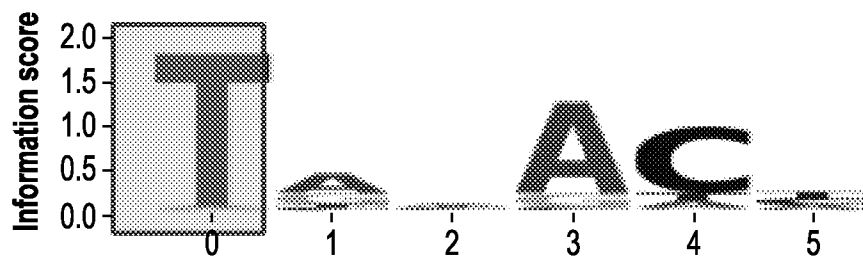
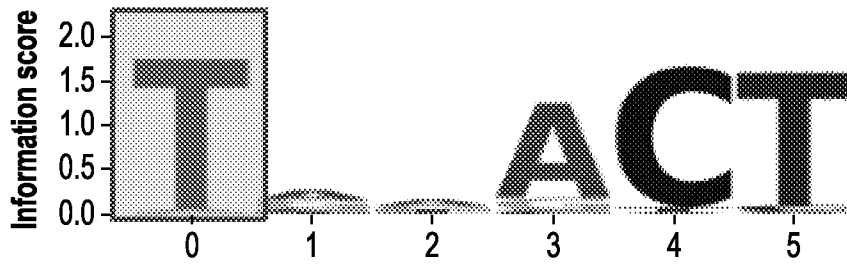
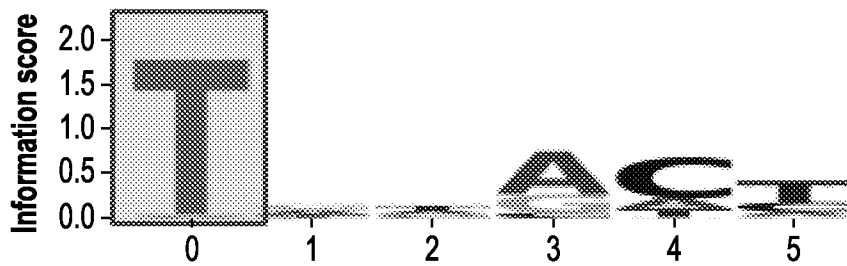
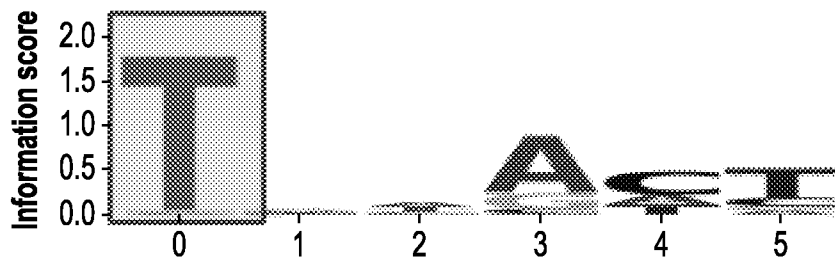
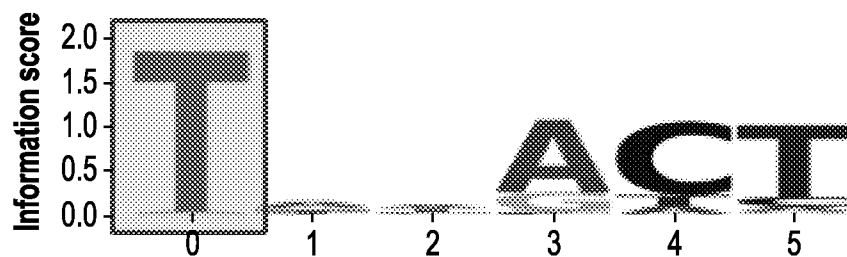
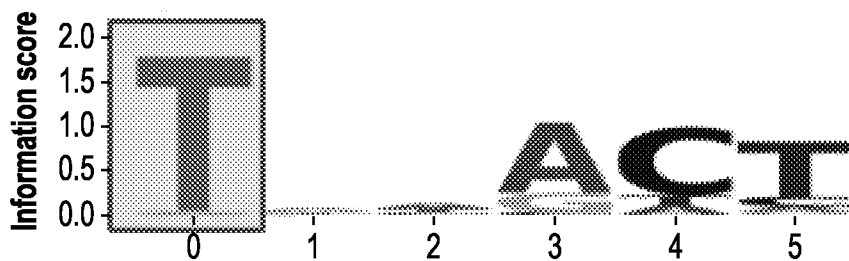
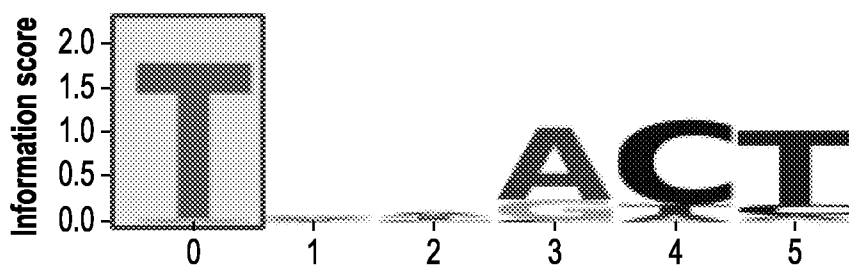
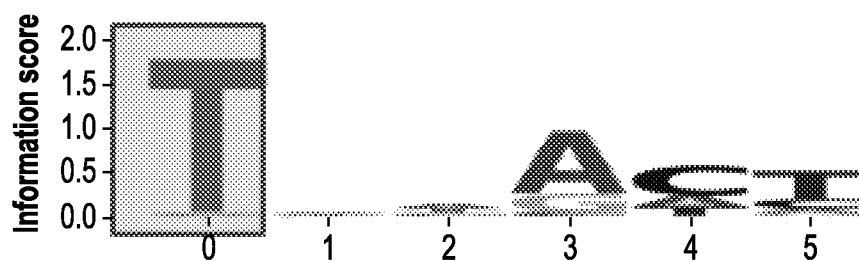
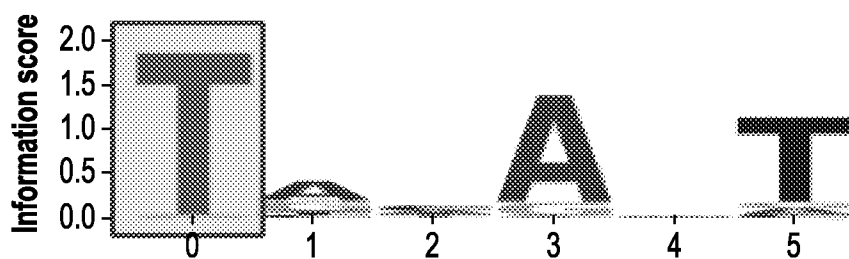


FIG. 16 (Cont. 2)

prt1559**240628-TWH-PE-P1G1-prt1559****prt1569****240628-TWH-PE-P1G6-prt1569****prt1574****240628-TWH-PE-P1G11-prt1574****prt22904****FIG. 16 (Cont. 3)**

prt1560**240628-TWH-PE-P1G2-prt1560****prt1570****240628-TWH-PE-P1G7-prt1570****prt1575****240628-TWH-PE-P2C1-prt1575****prt22905****FIG. 16 (Cont. 4)**

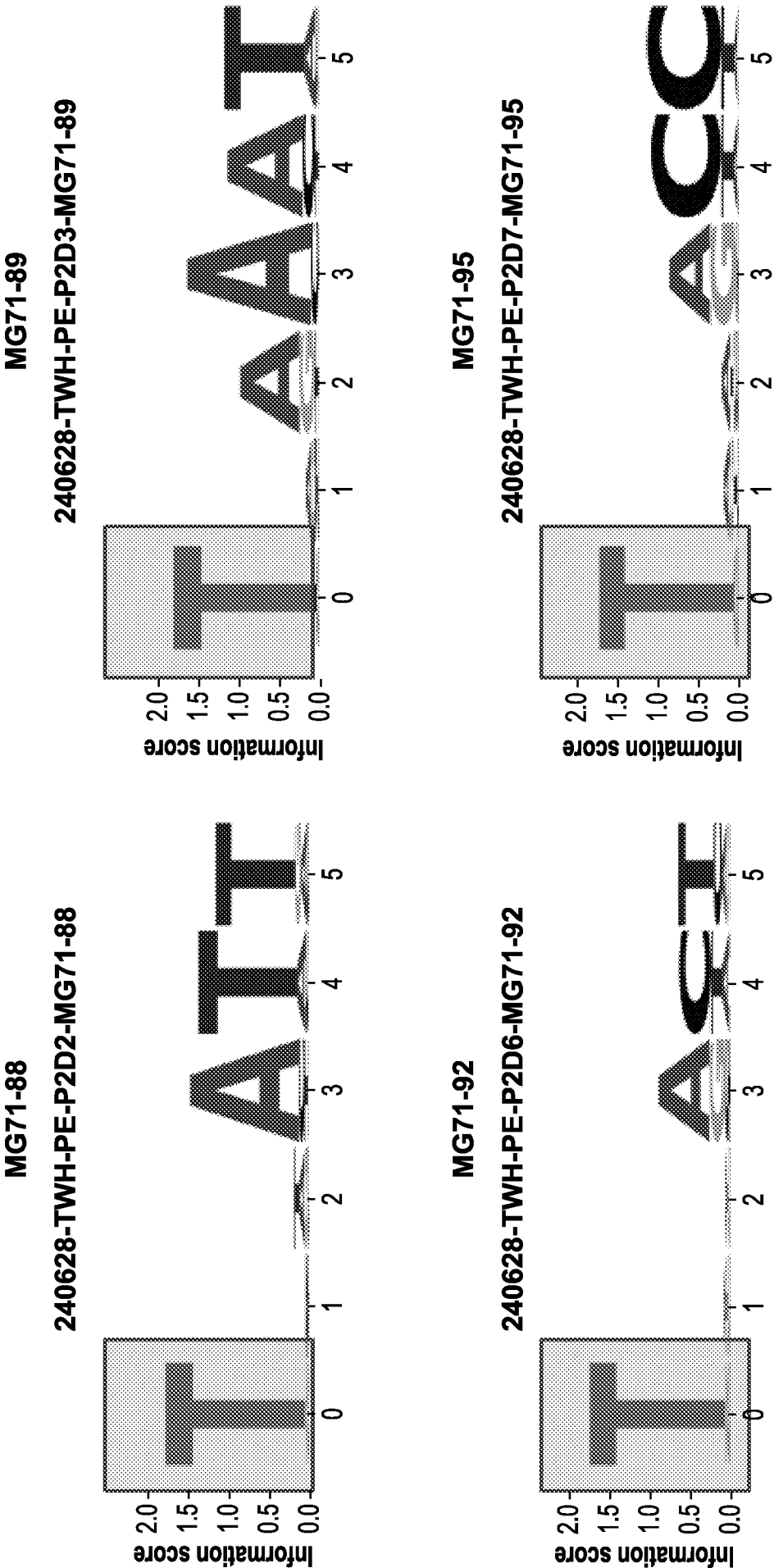


FIG. 17

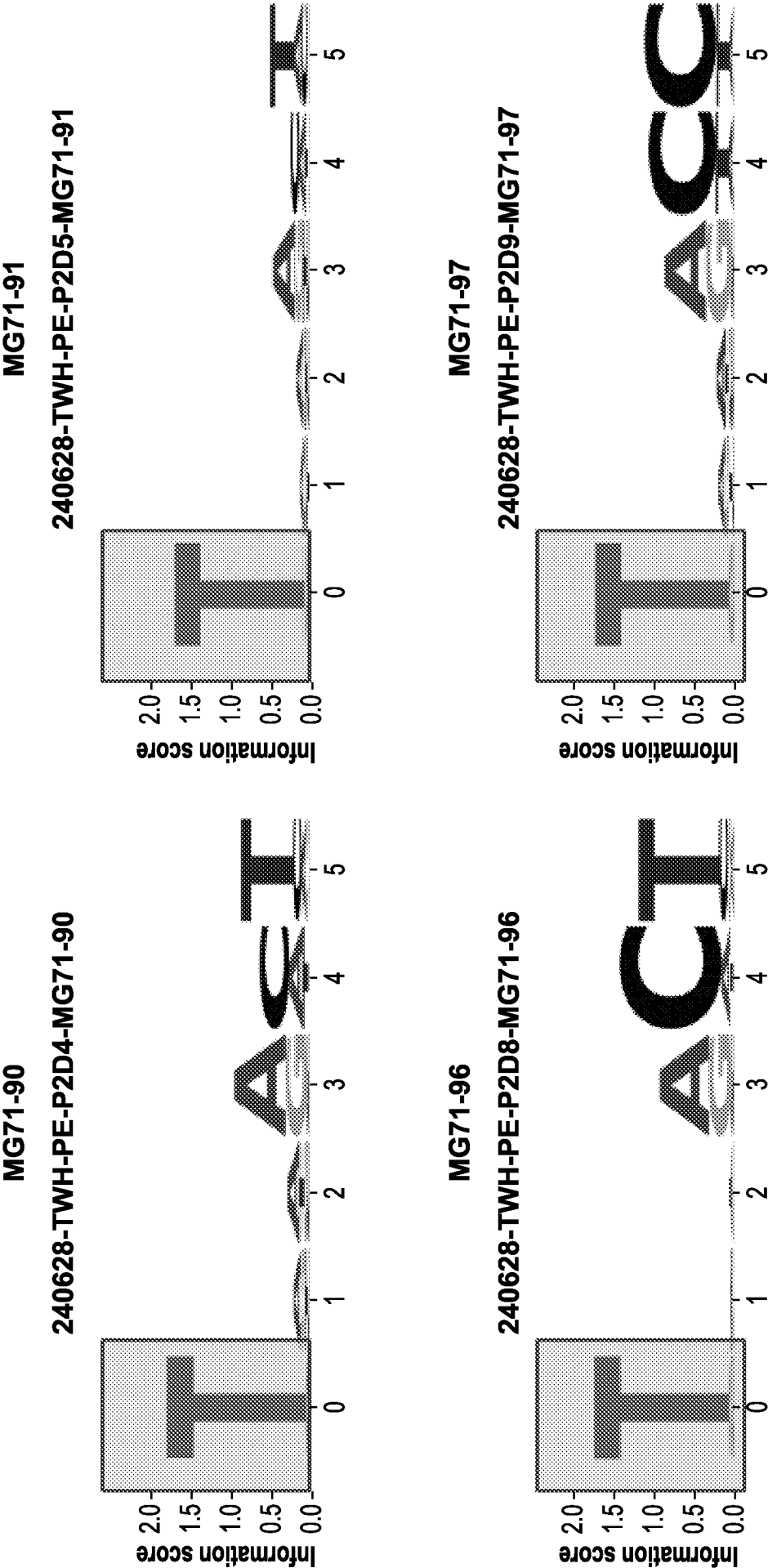


FIG. 17 (Cont.)

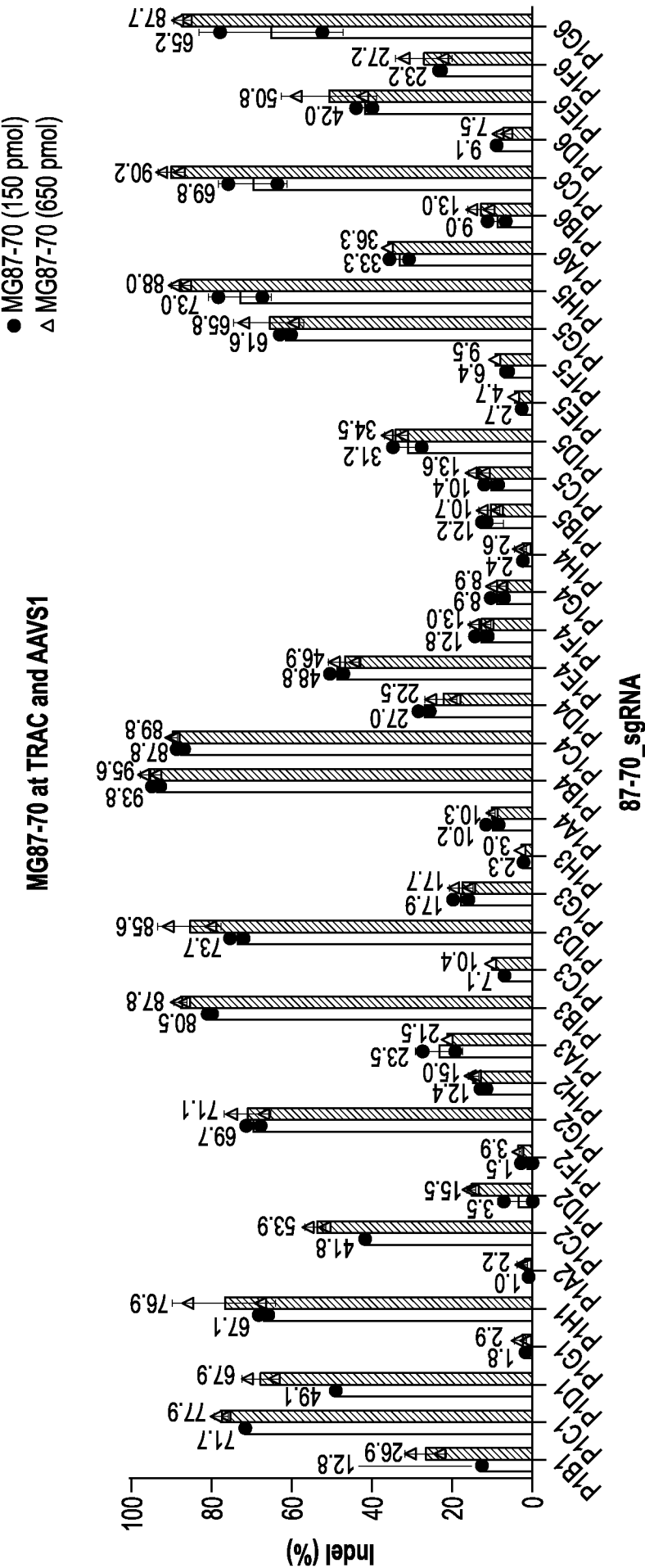


FIG. 18

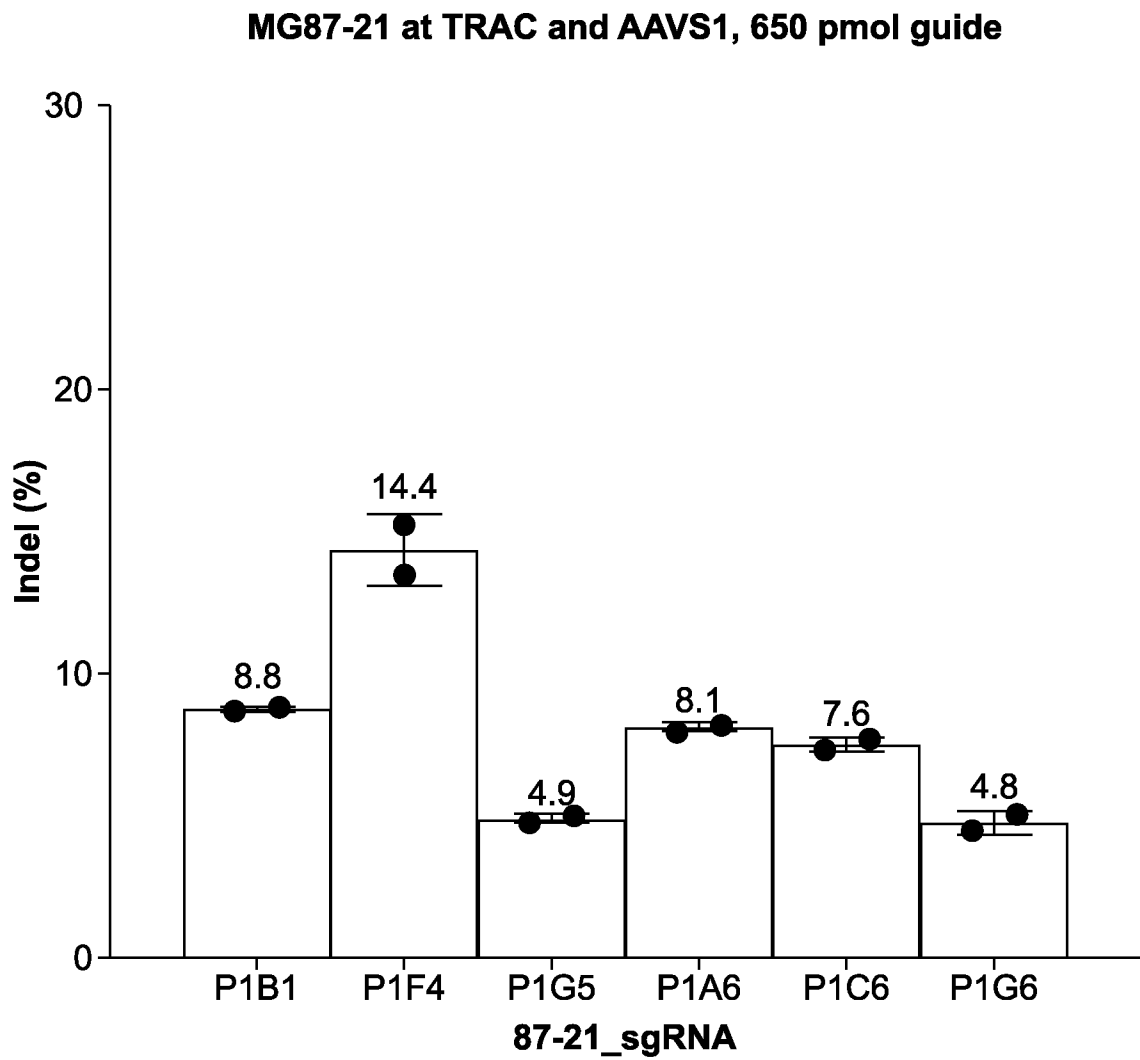


FIG. 19

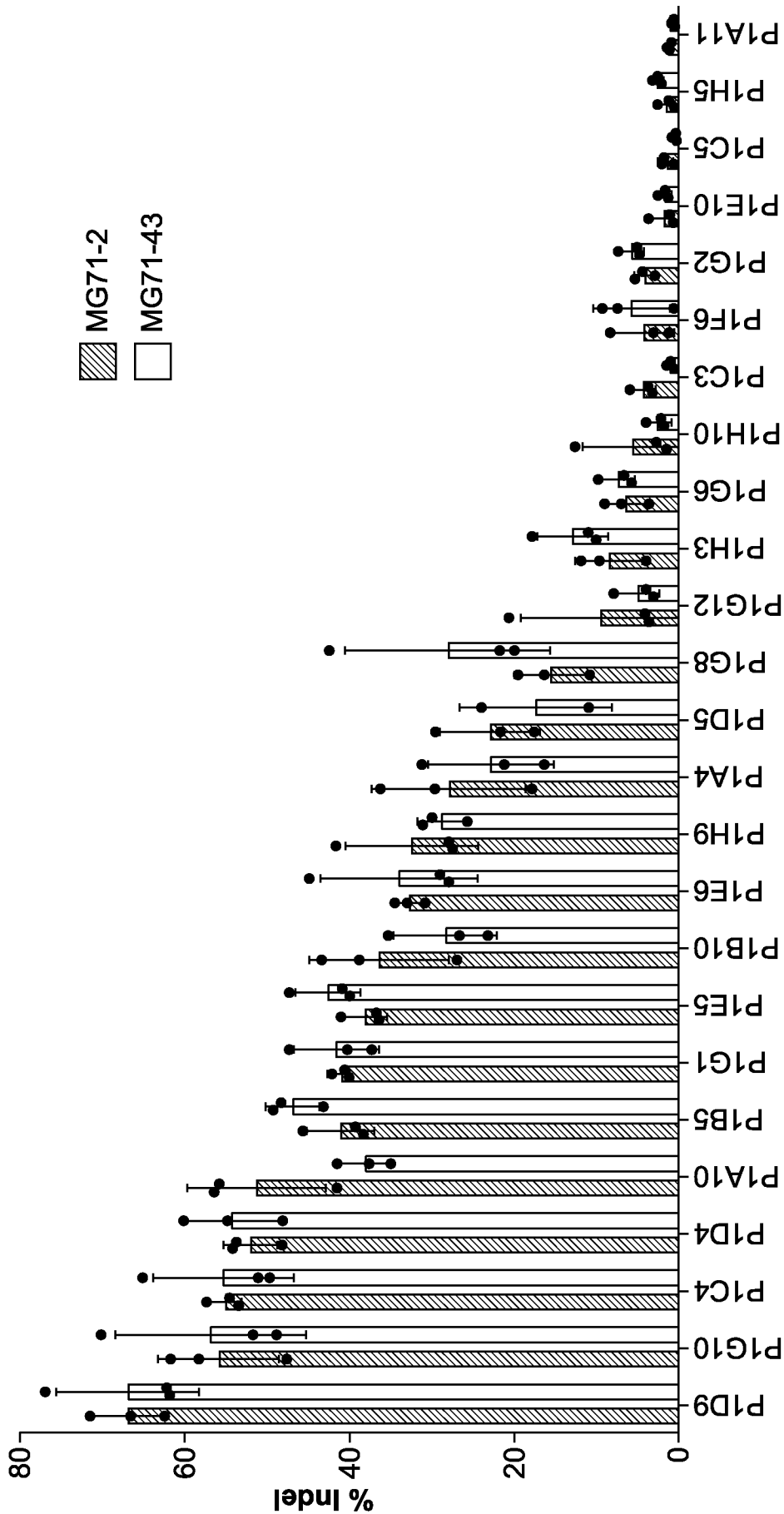


FIG. 20

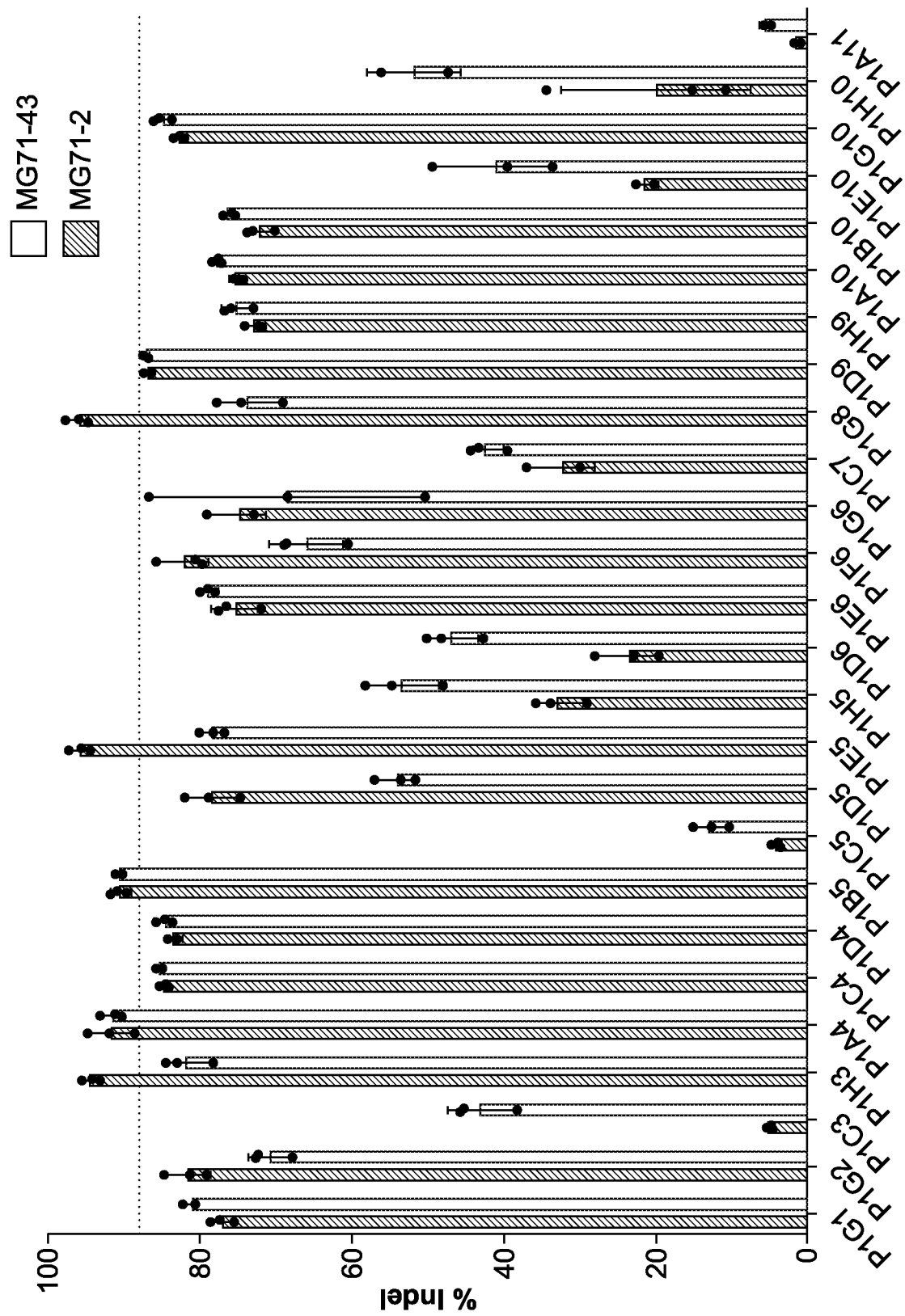


FIG. 21

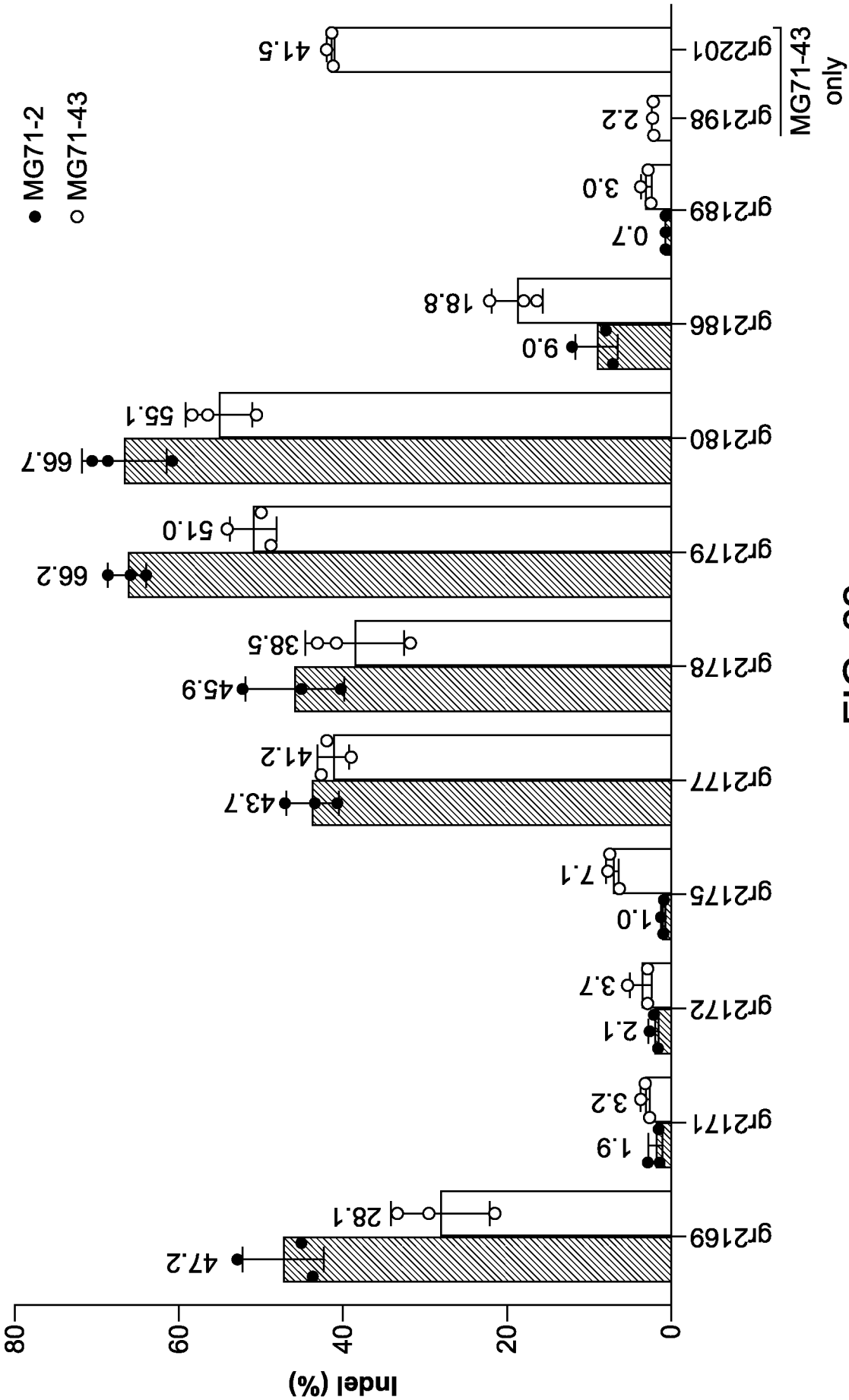


FIG. 22

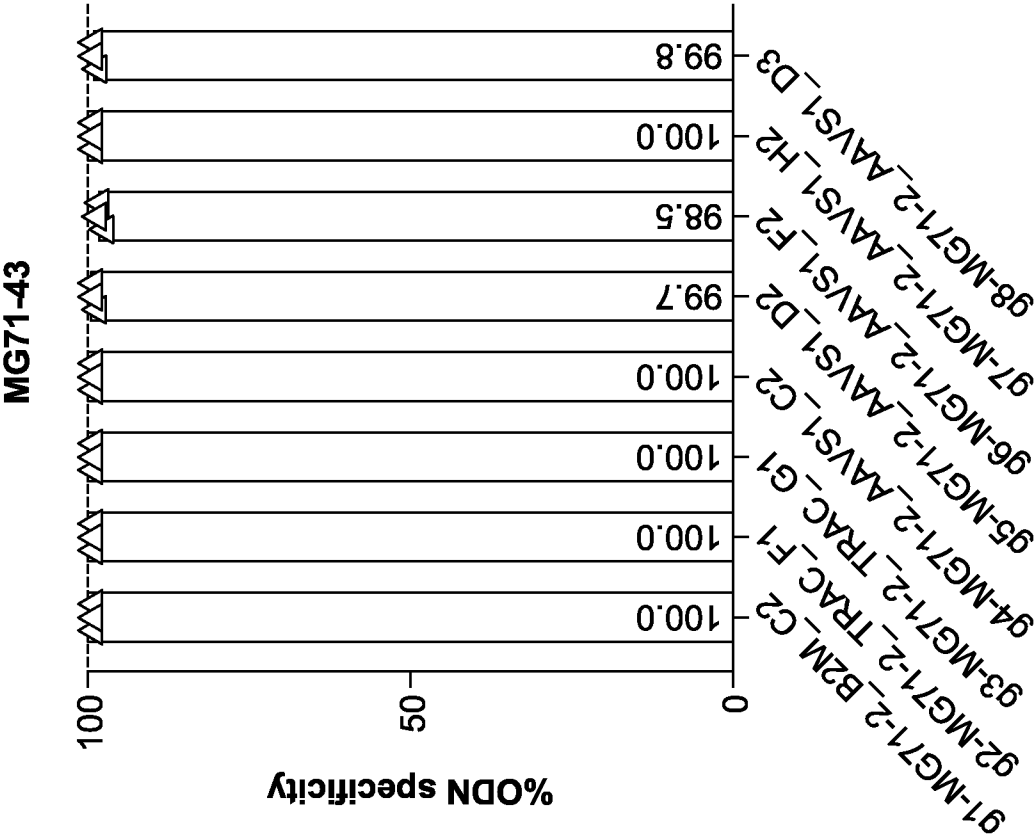


FIG. 23B

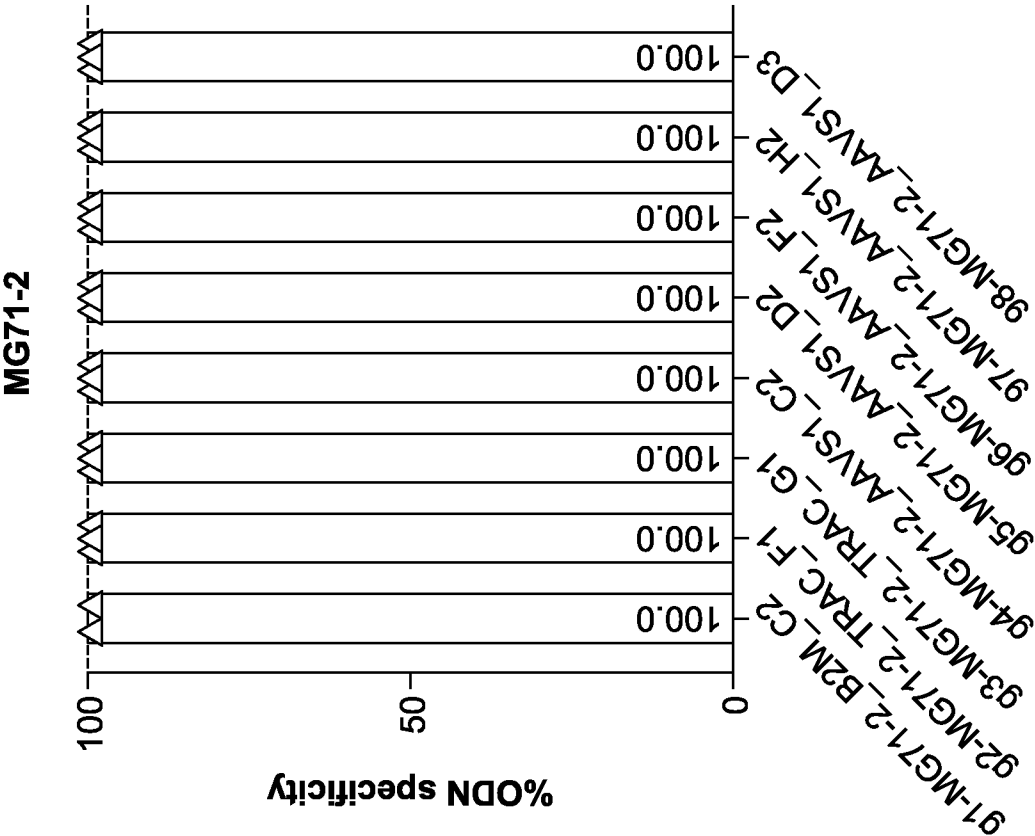


FIG. 23A

**MG71-43 protein with
guide g6-MG71-2_AAVS1_H2
3 replicates**

5	10	15	20	25	
T	A	G	G	G	G
C	C	C	C	G	G
G	G	T	T	G	G
A	G	G	A	G	G
A	A	G	N	N	N
R	C	Y			
Reads					
3000					
9					
29					
5	10	15	20	25	
T	A	G	G	G	G
C	C	C	C	G	G
T	T	G	G	A	G
A	G	G	A	A	G
N	N	N	N	R	C
Y					
Reads					
3448					
15					
35					
5	10	15	20	25	
T	A	G	G	G	G
C	C	C	C	G	G
T	T	G	G	A	G
A	G	G	A	A	G
N	N	N	N	R	C
Y					
Reads					
2602					
13					
31					

FIG. 23C

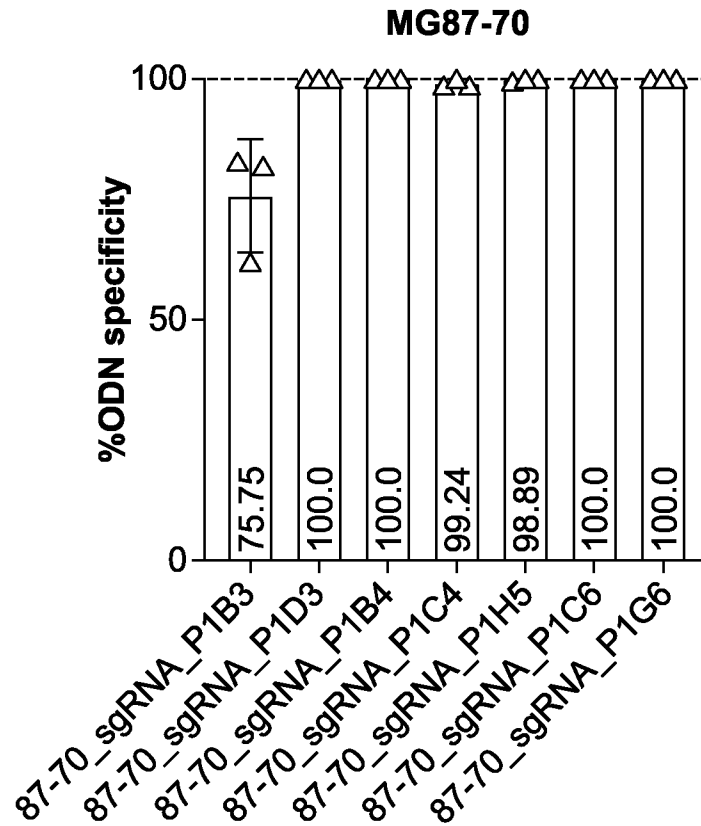


FIG. 24A

**MG87-70 protein with guide
87-70_sgRNA_P1B3 3 replicates**

5	10	15	20	25	
TAGCACA GACTAGAGAGGTAAGNNNGT					Reads
.....					1940
.....T.....A.....					19
AC A T A A					1158
5	10	15	20	25	
TAGCACA GACTAGAGAGGTAAGNNNGT					Reads
.....					3648
.....T.....A.....					13
AC A T A A					781
5	10	15	20	25	
TAGCACA GACTAGAGAGGTAAGNNNGT					Reads
.....					4596
.....T.....A.....					3
AC A T A A					950

FIG. 24B

**Activity of variants with guide pools in
engineered K562 cells**

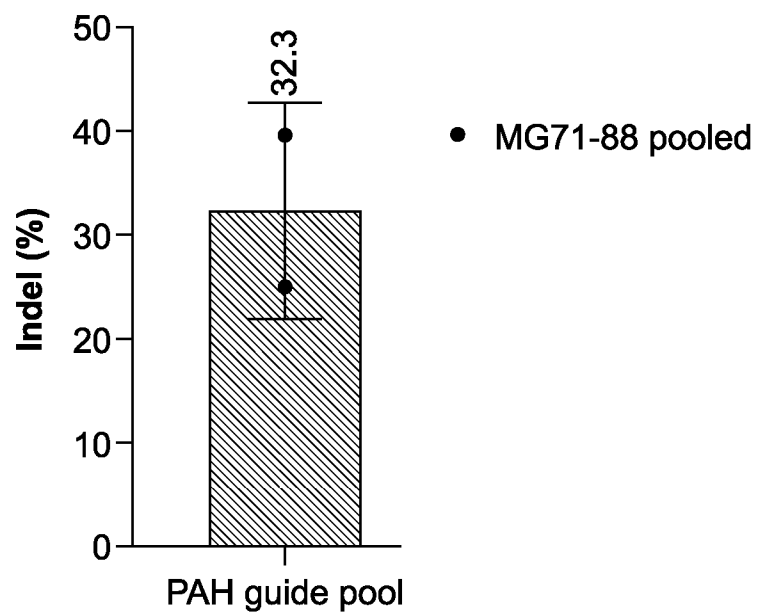


FIG. 25

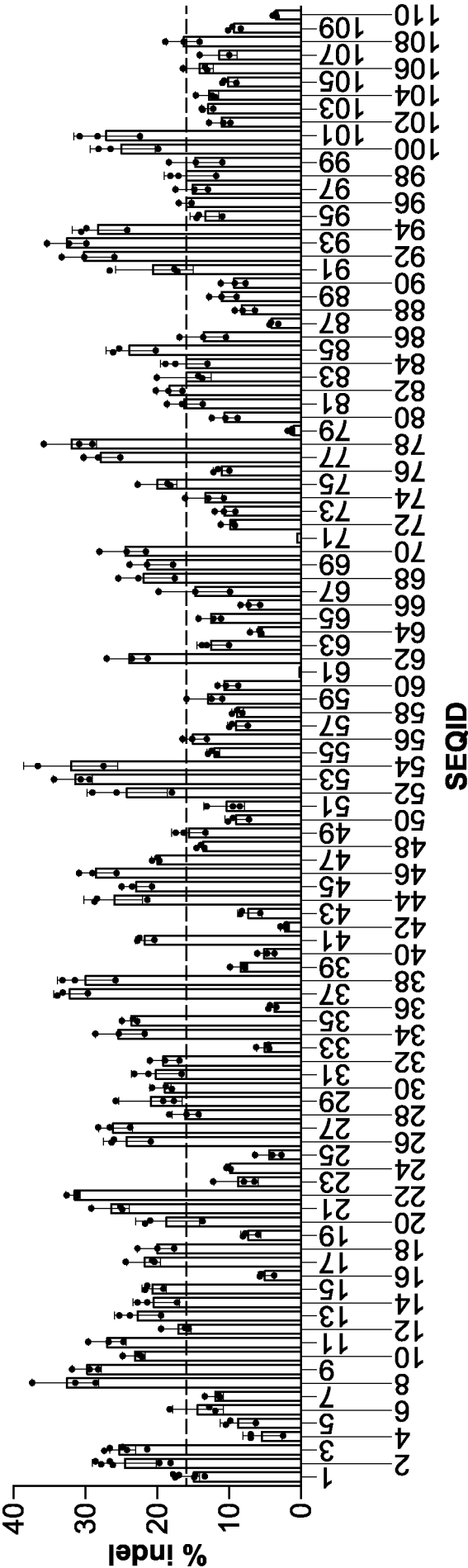


FIG. 26

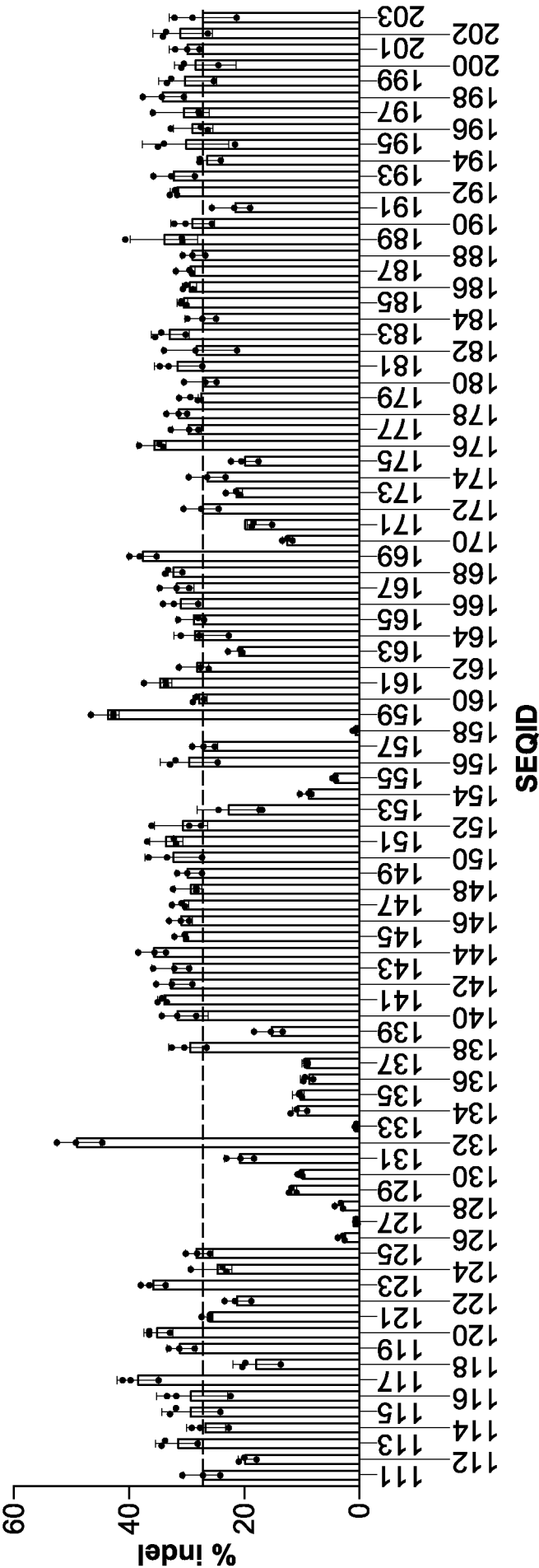


FIG. 27

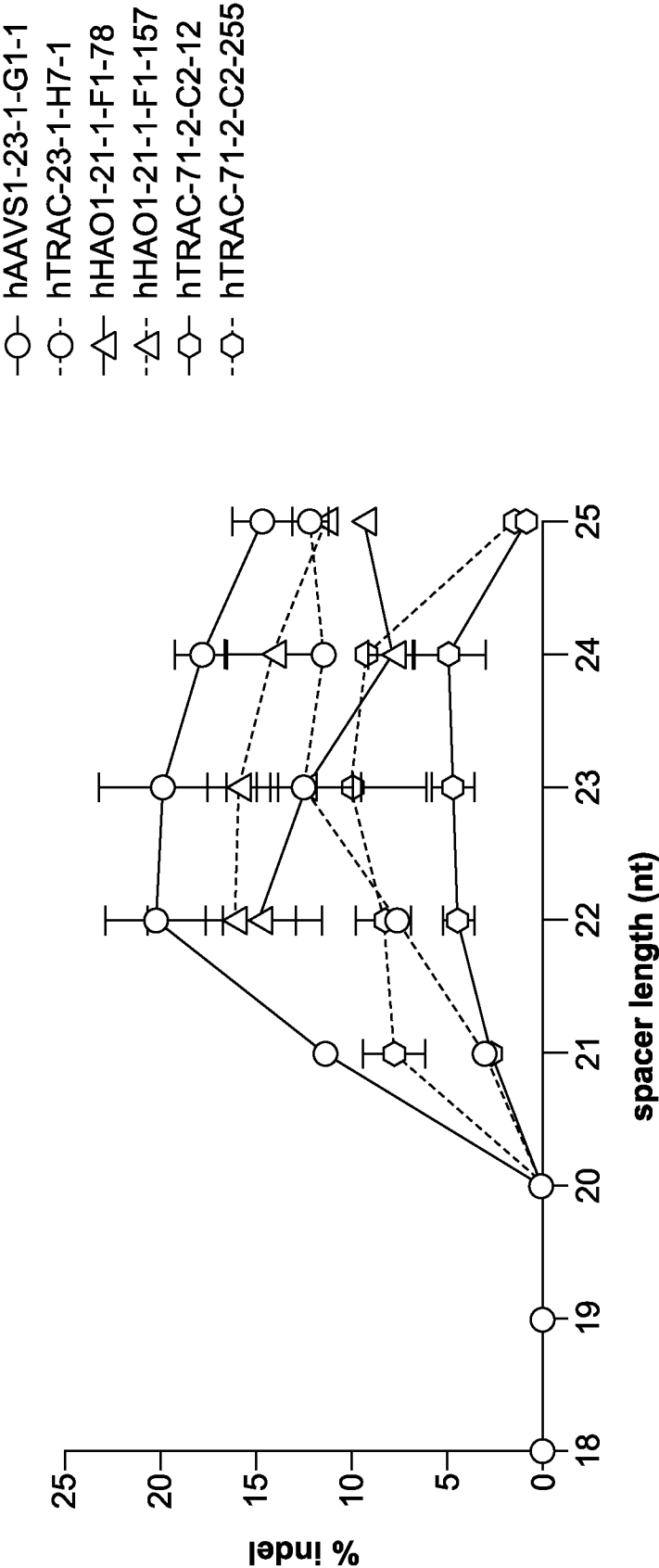


FIG. 28

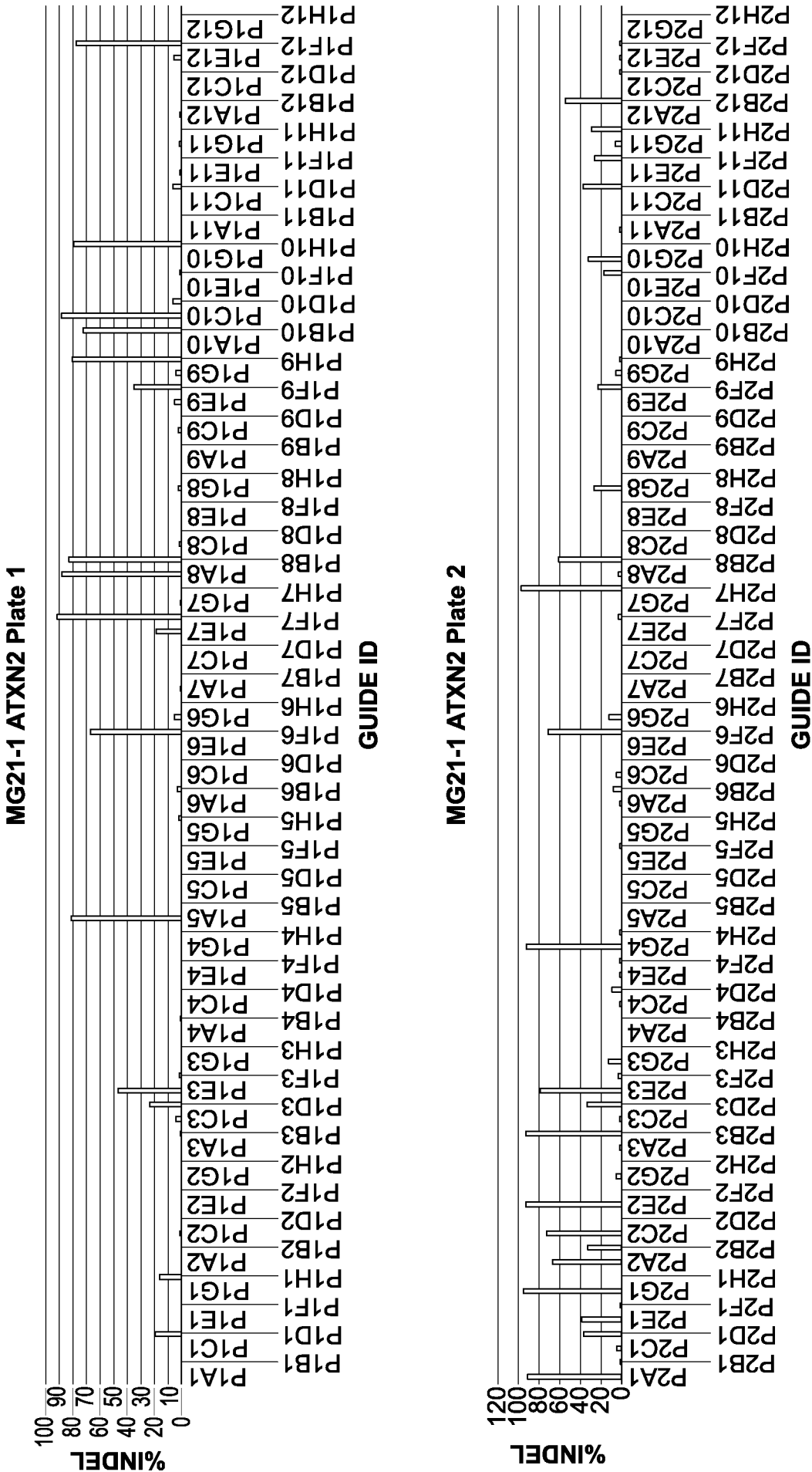


FIG. 29

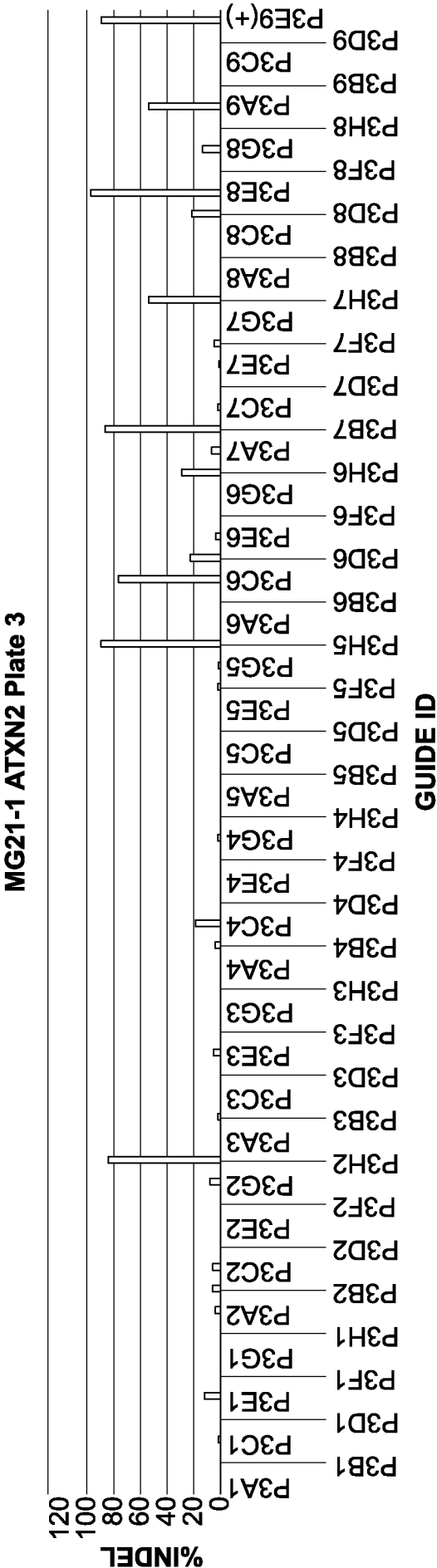


FIG. 29 (Cont.)

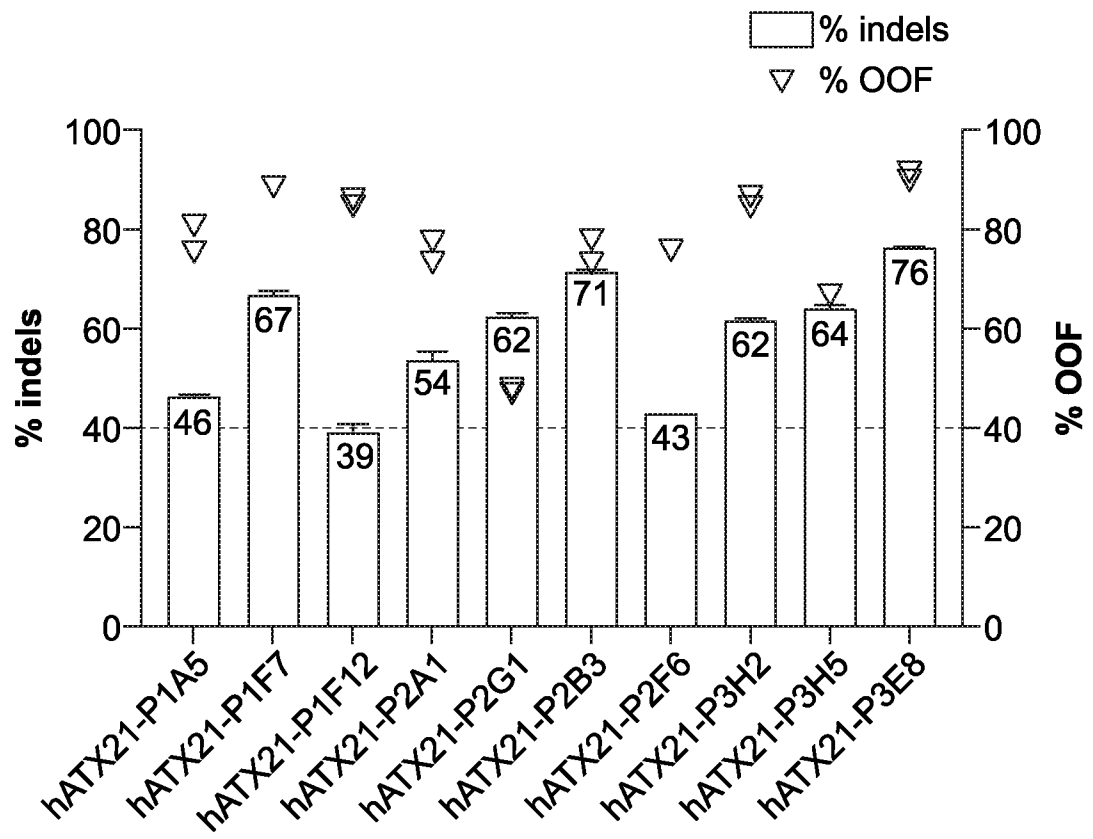


FIG. 30

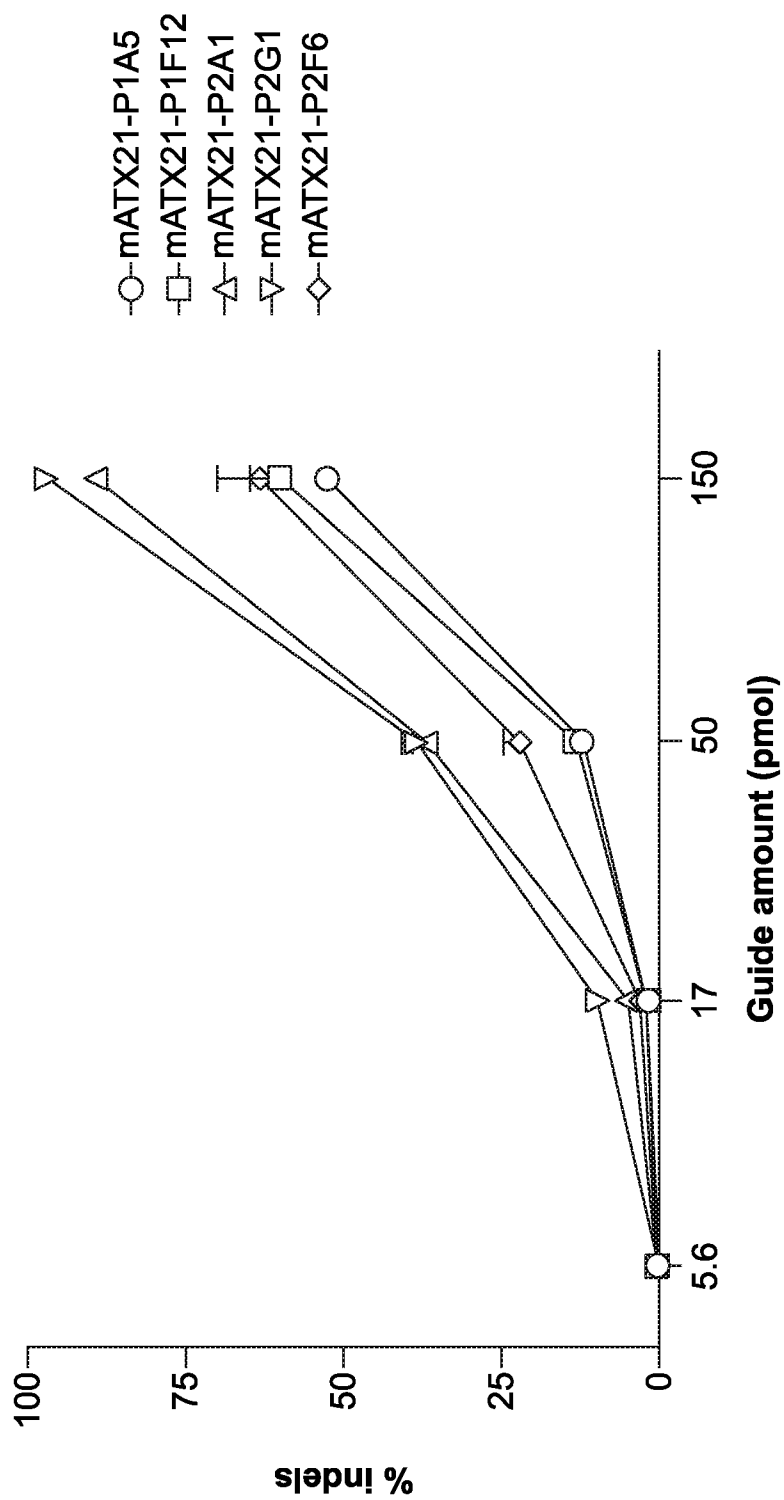


FIG. 31

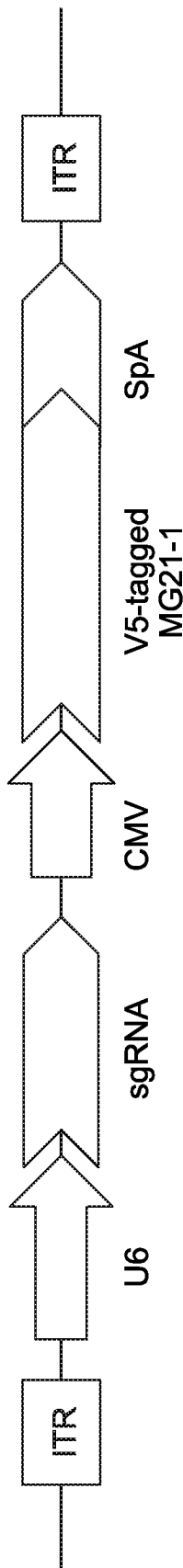


FIG. 32

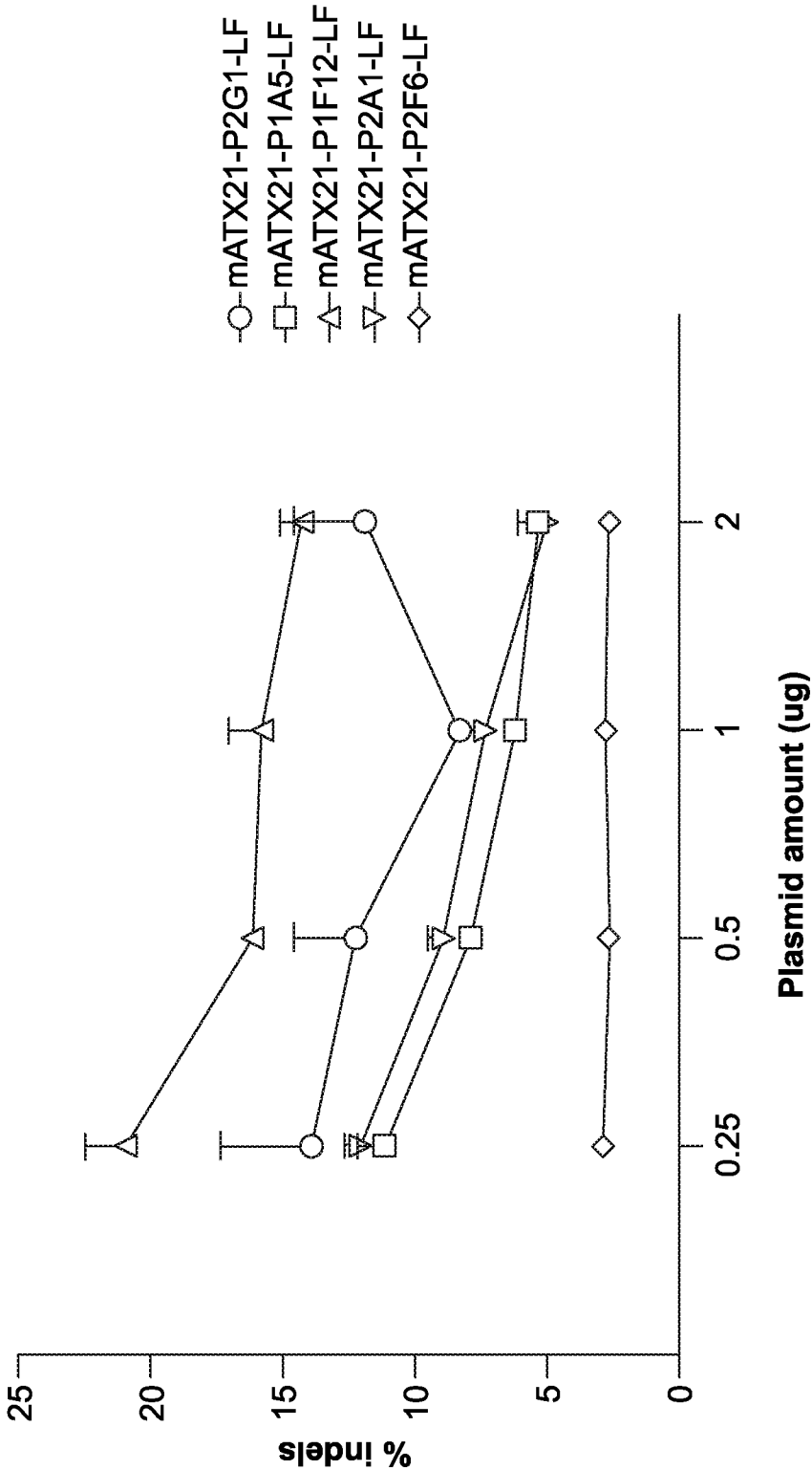


FIG. 33

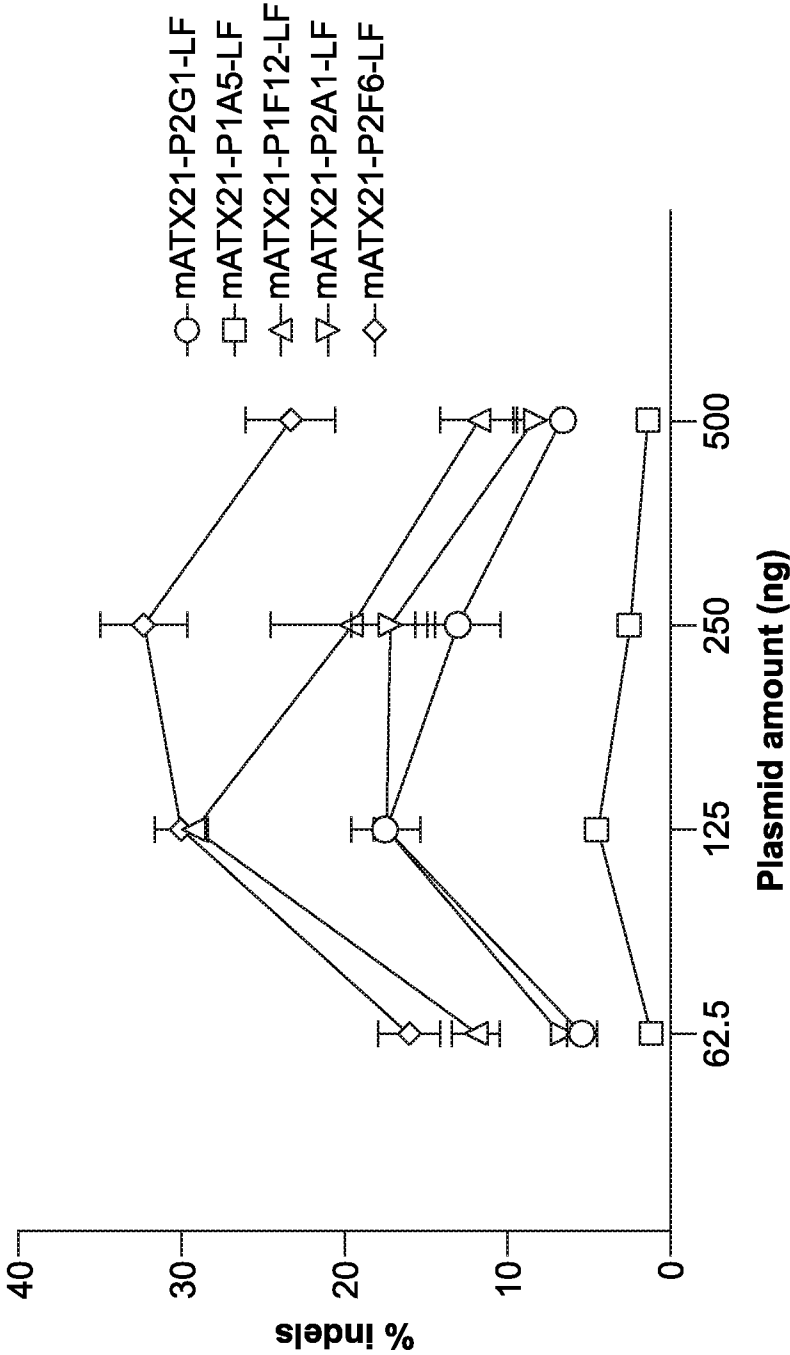


FIG. 34

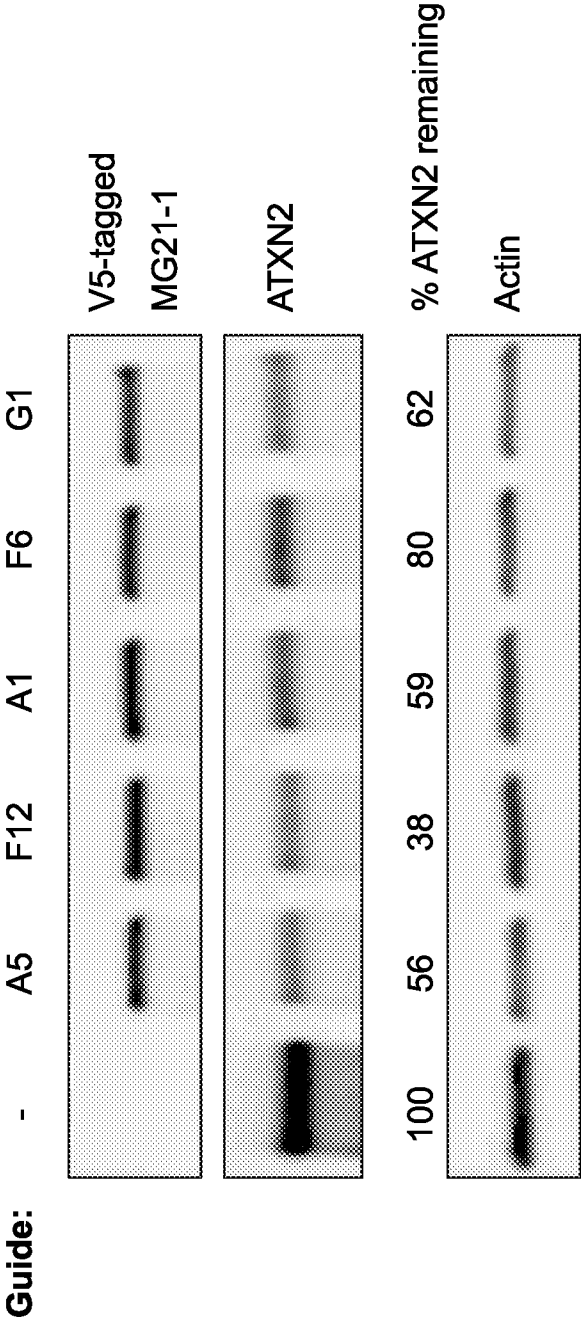


FIG. 35

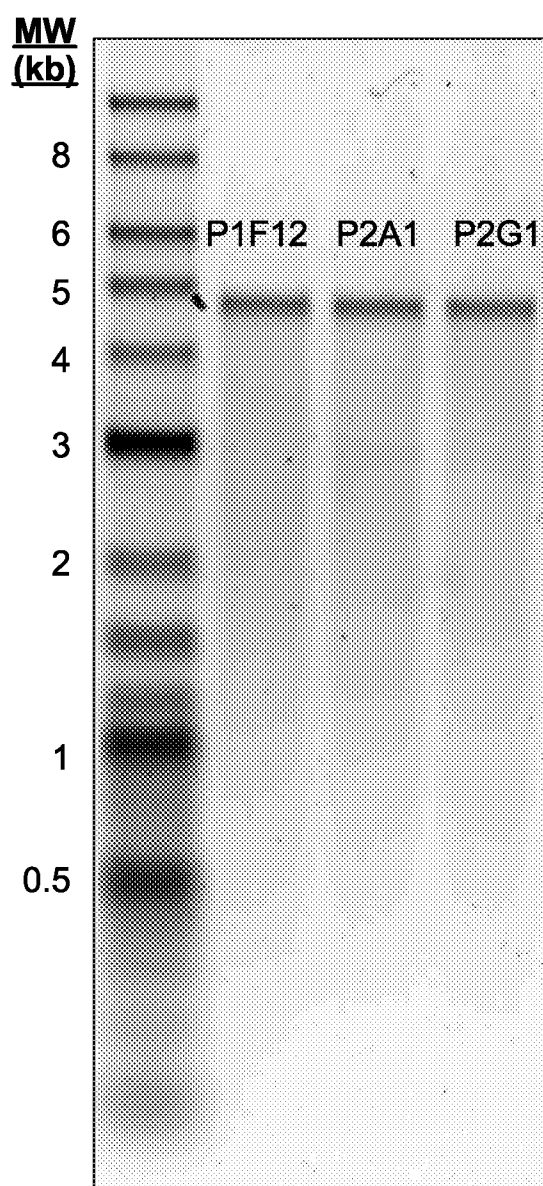


FIG. 36

Groups overview and description				
Group #	Serotype	Cargo full name	Purpose	# Mice
1	N/A	Untreated	Indels/Protein	4
2 (STX)	AAV9	CMV_V5_MG21-1_SpA_mATXN2-P1F12_FW	Indels/Protein	4
3 (STX)	AAV9	CMV_V5_MG21-1_SpA_mATXN2-P2A1_FW	Indels/Protein	4
4 (STX)	AAV9	CMV V5_MG21-1_SpA_mATXN2-P2G1_FW	Indels/Protein	4

Stereotactic targeting M1 motor cortex and olfactory bulb

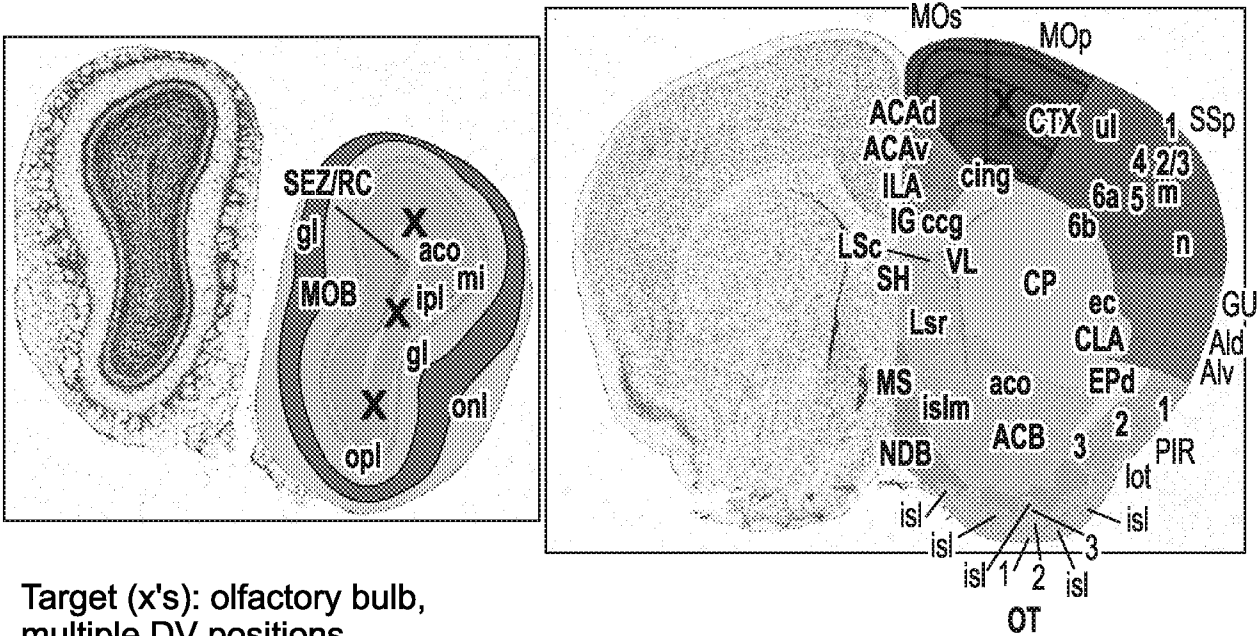
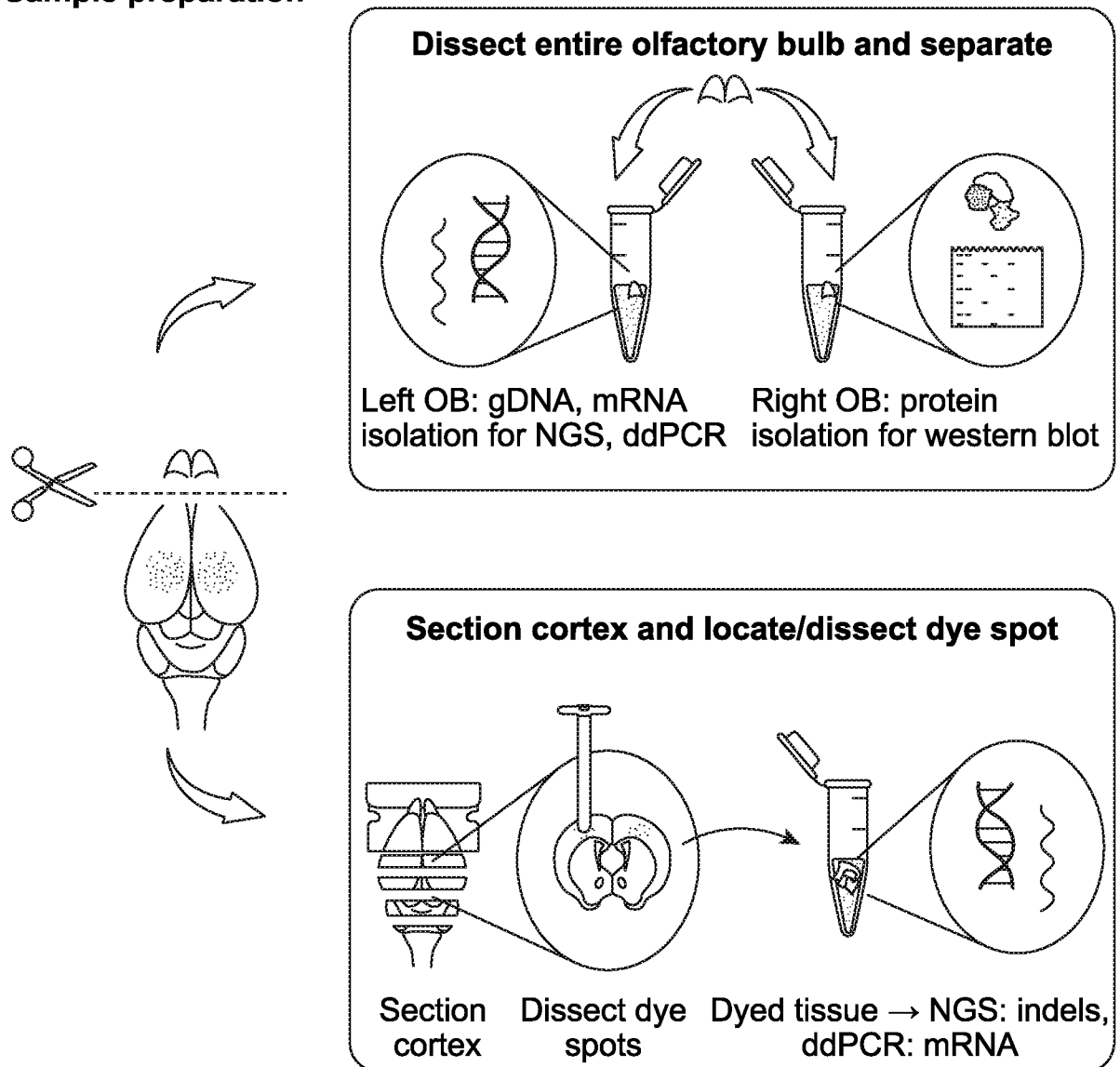


FIG. 37

Sample preparation**FIG. 37 (Cont.)**

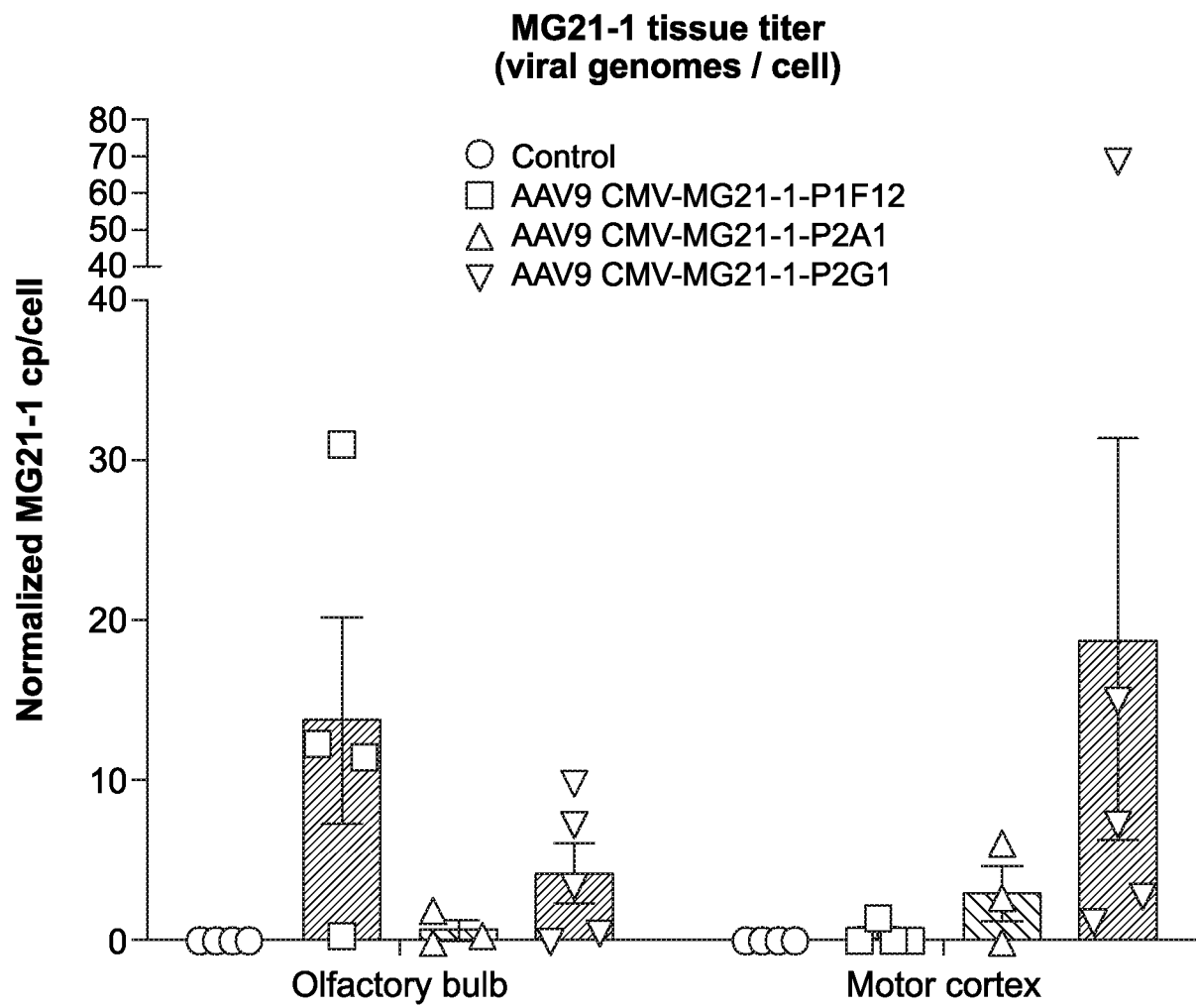


FIG. 38

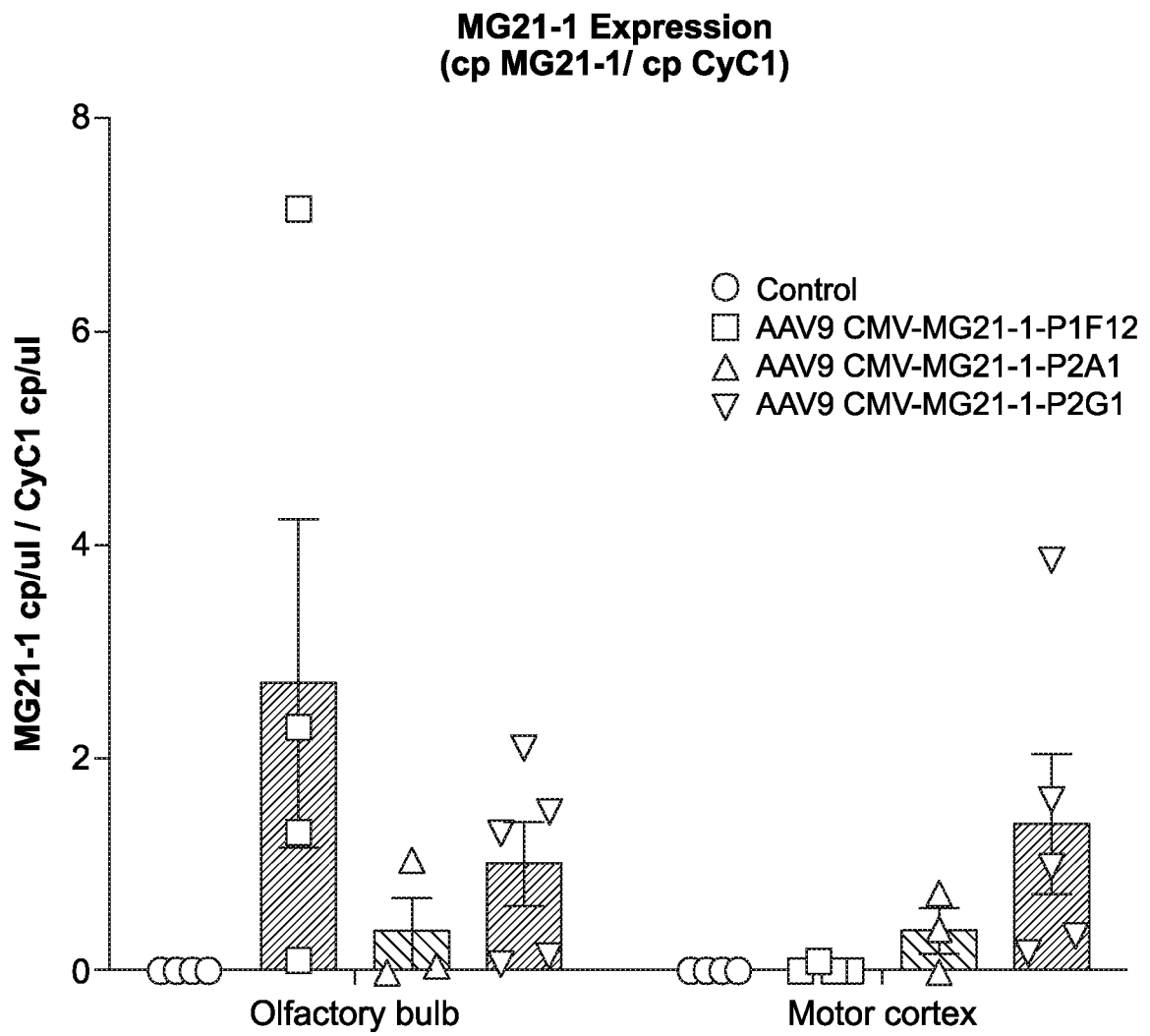


FIG. 39

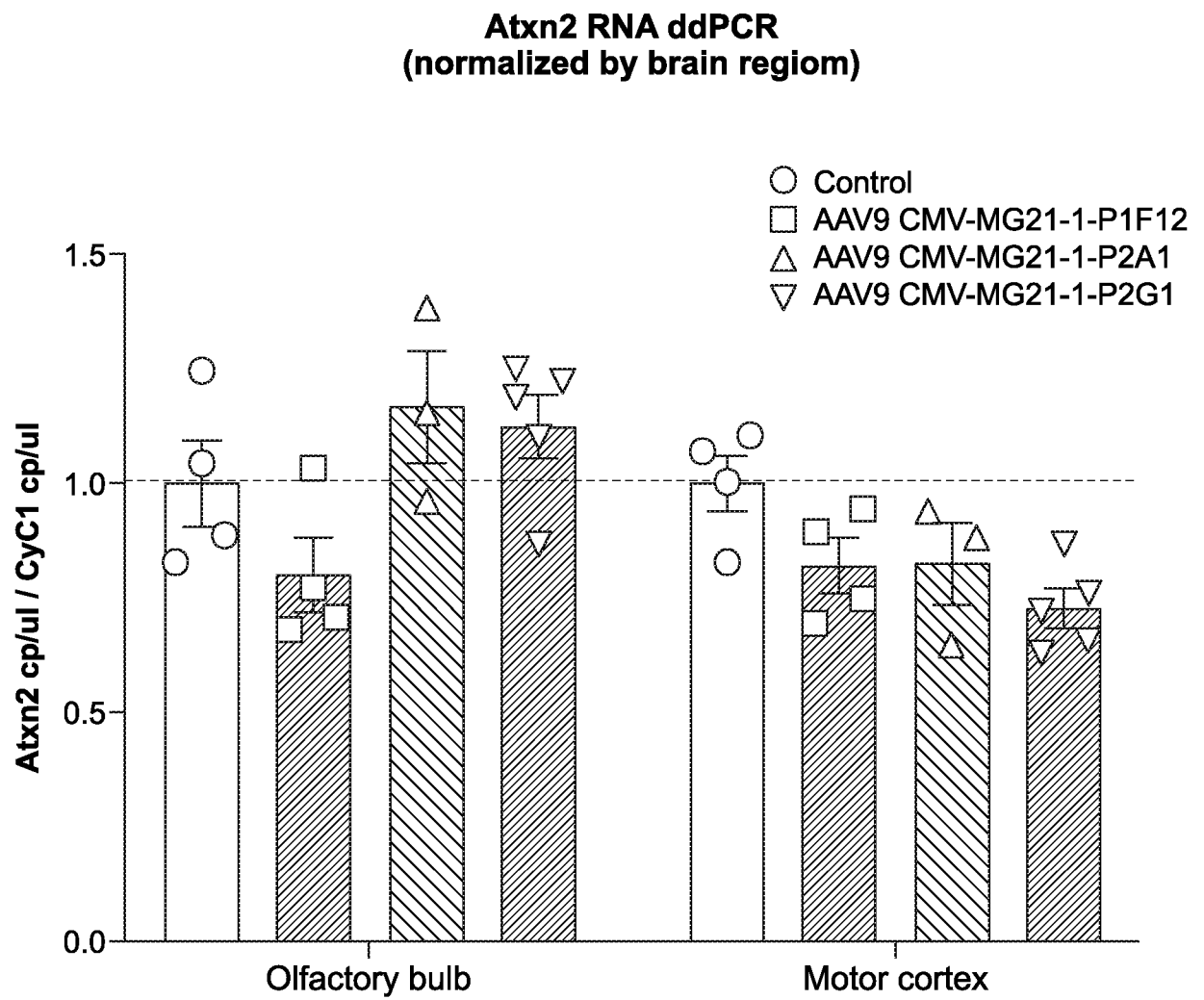
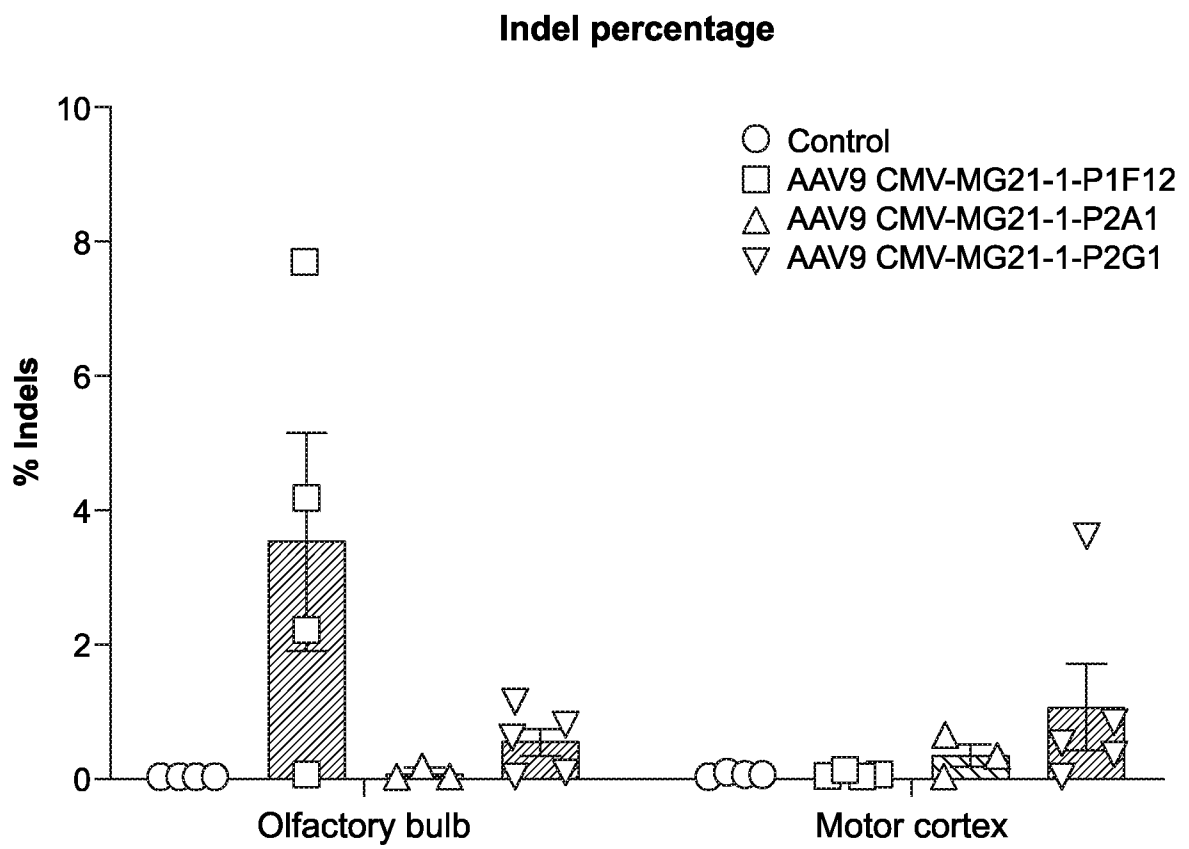


FIG. 40

**FIG. 41**

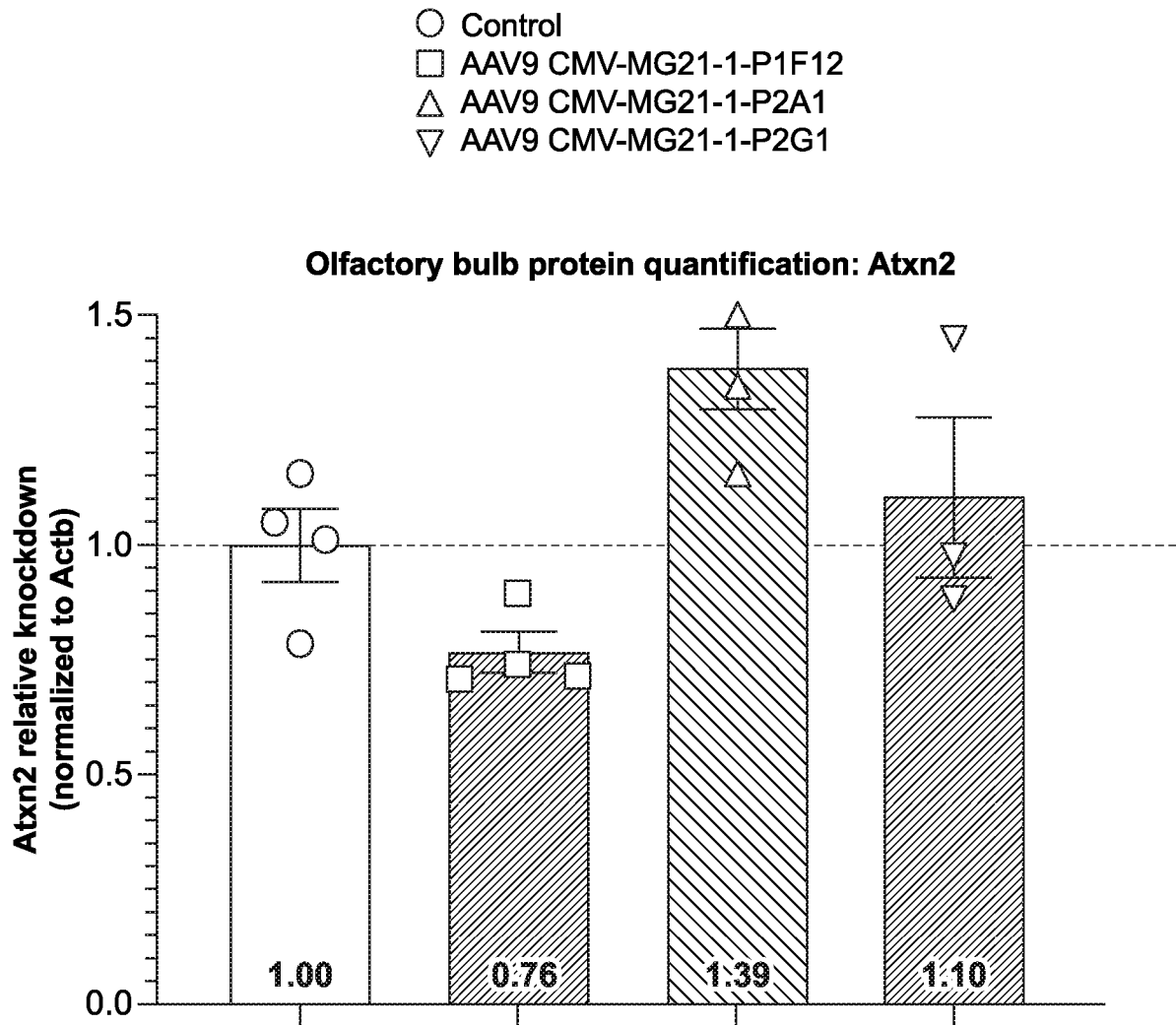


FIG. 42

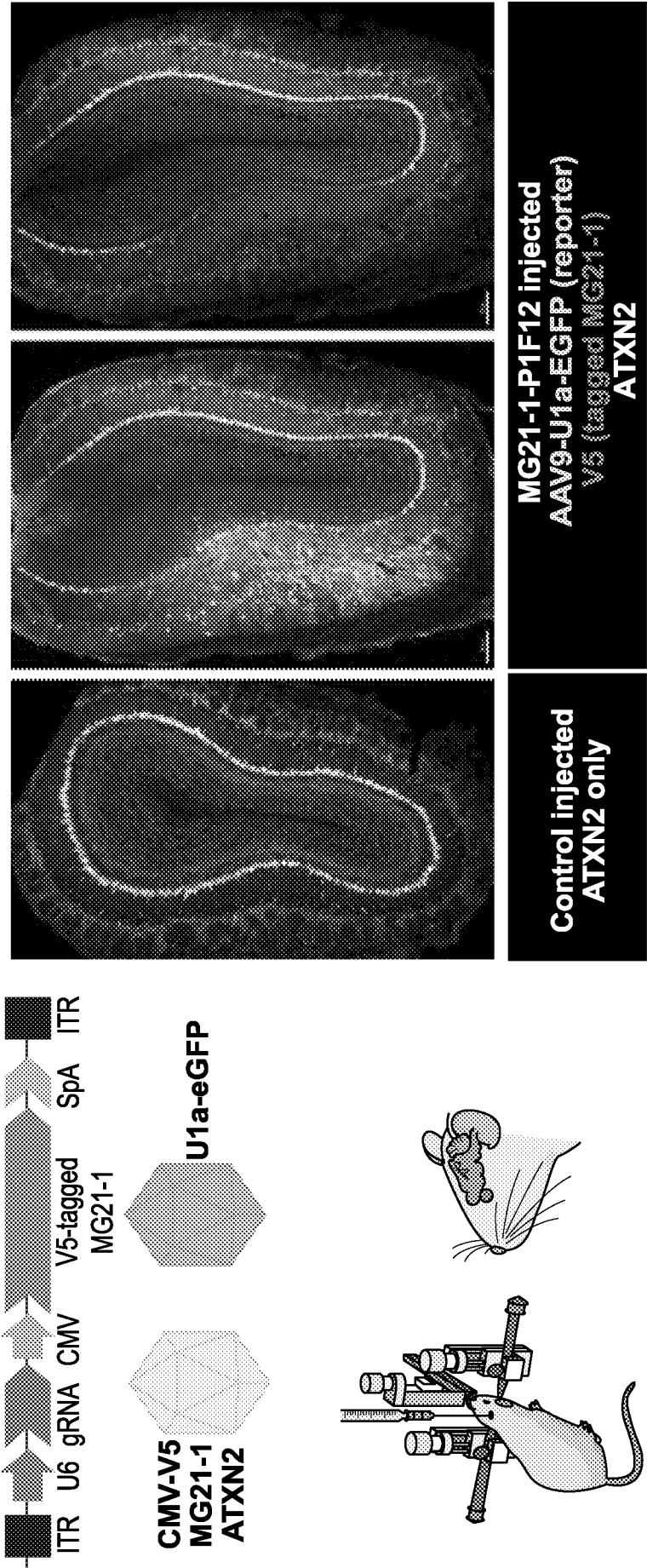


FIG. 43

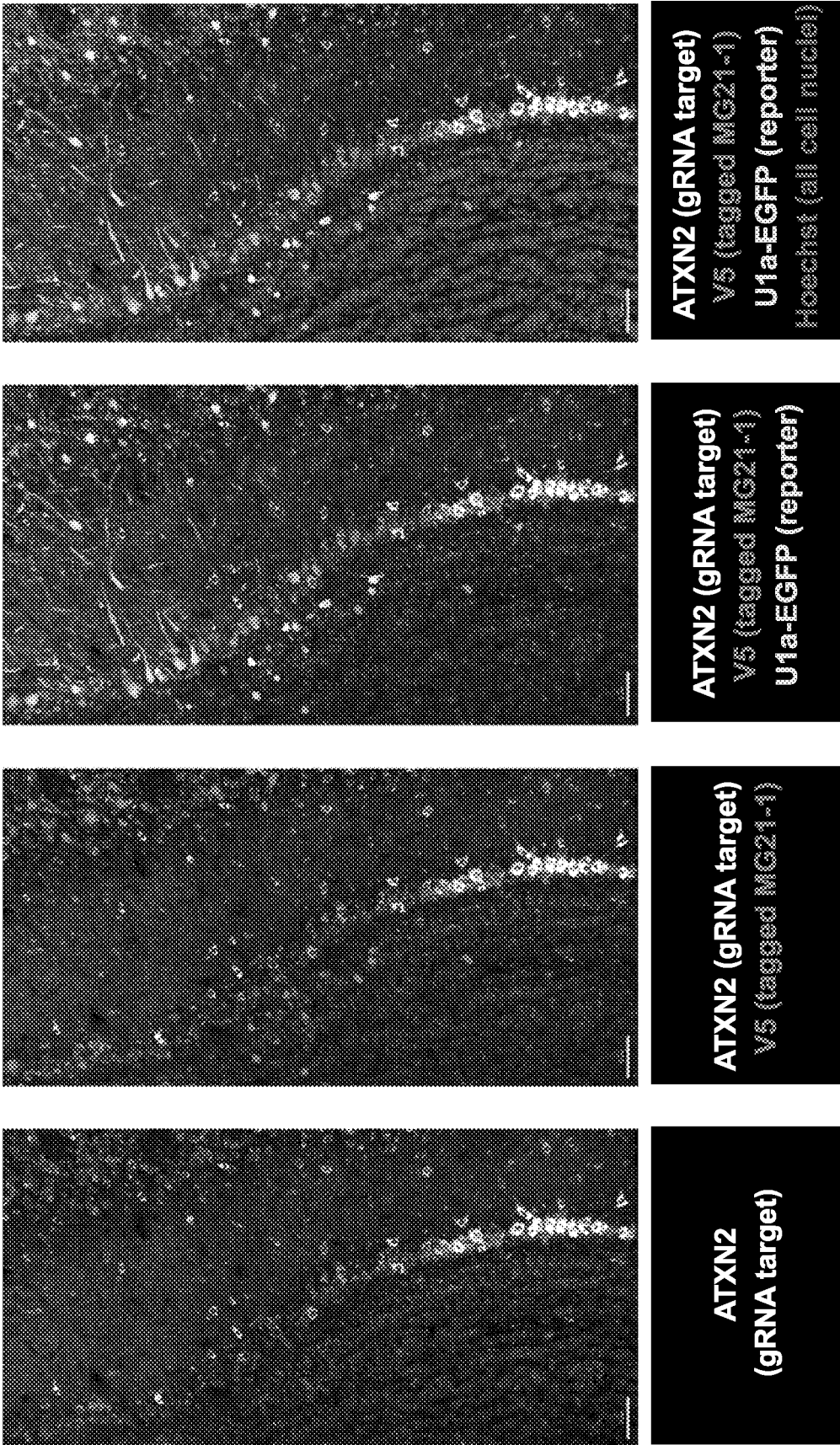


FIG. 44

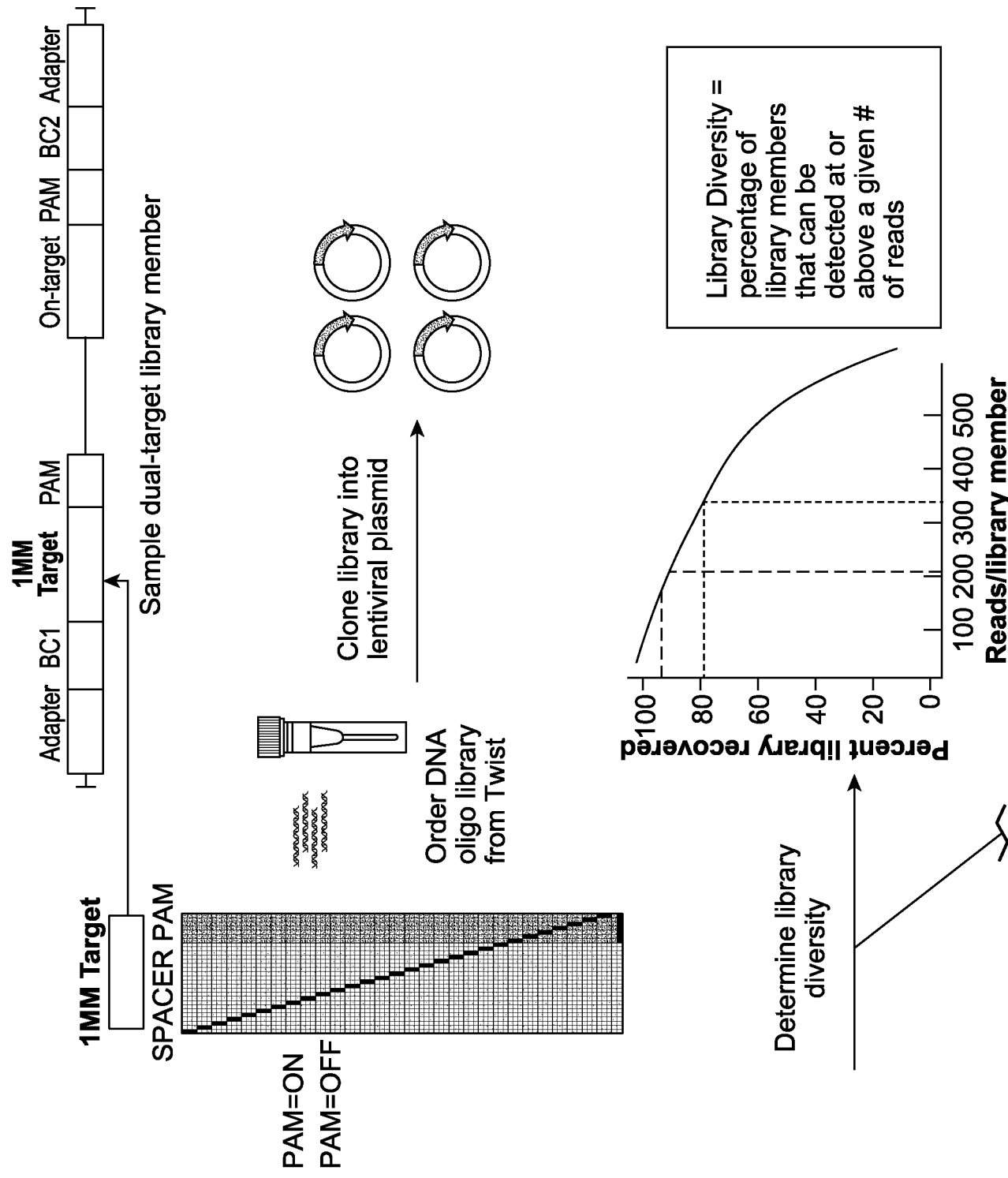


FIG. 45

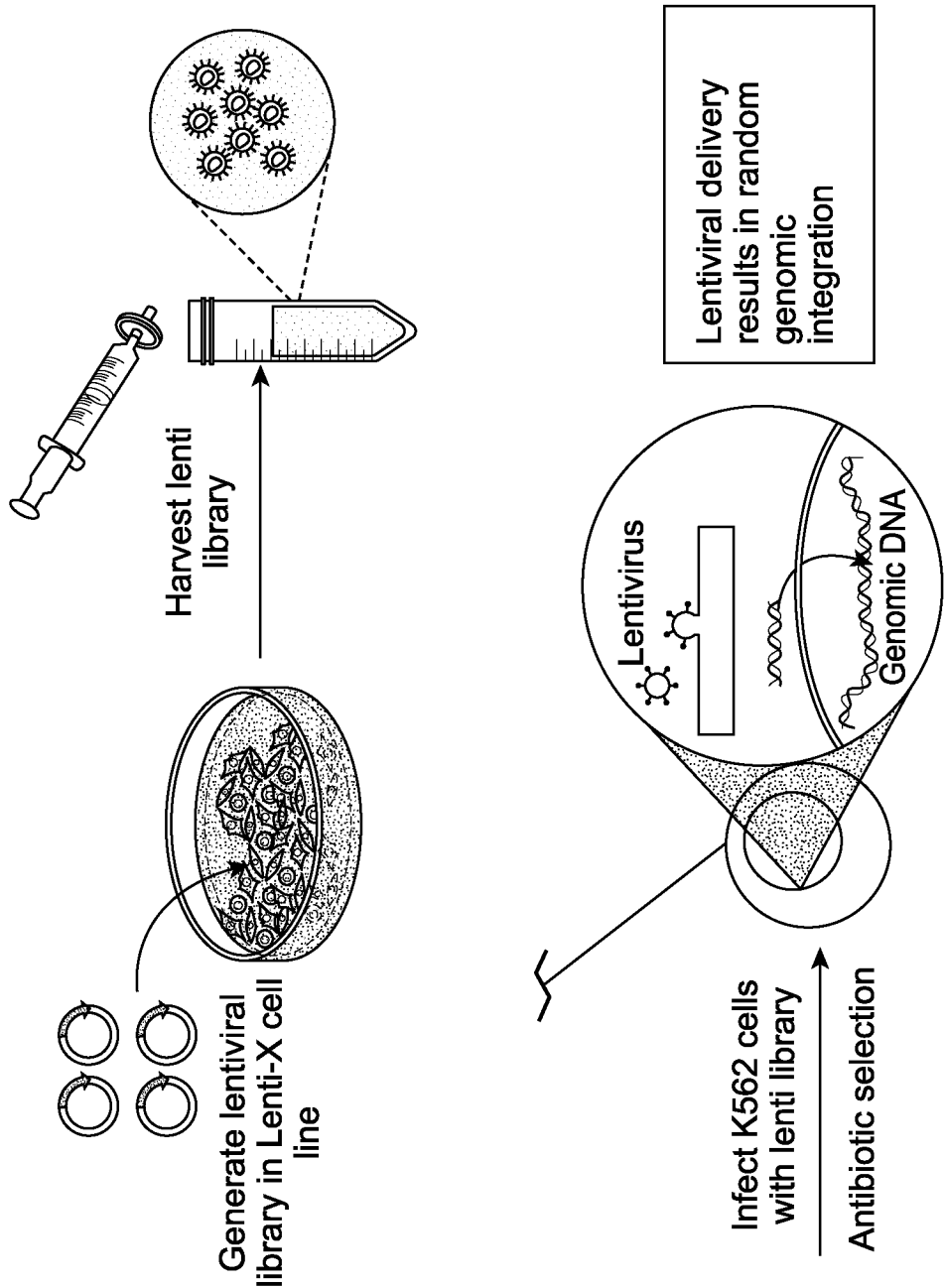


FIG. 45 (Cont. 1)

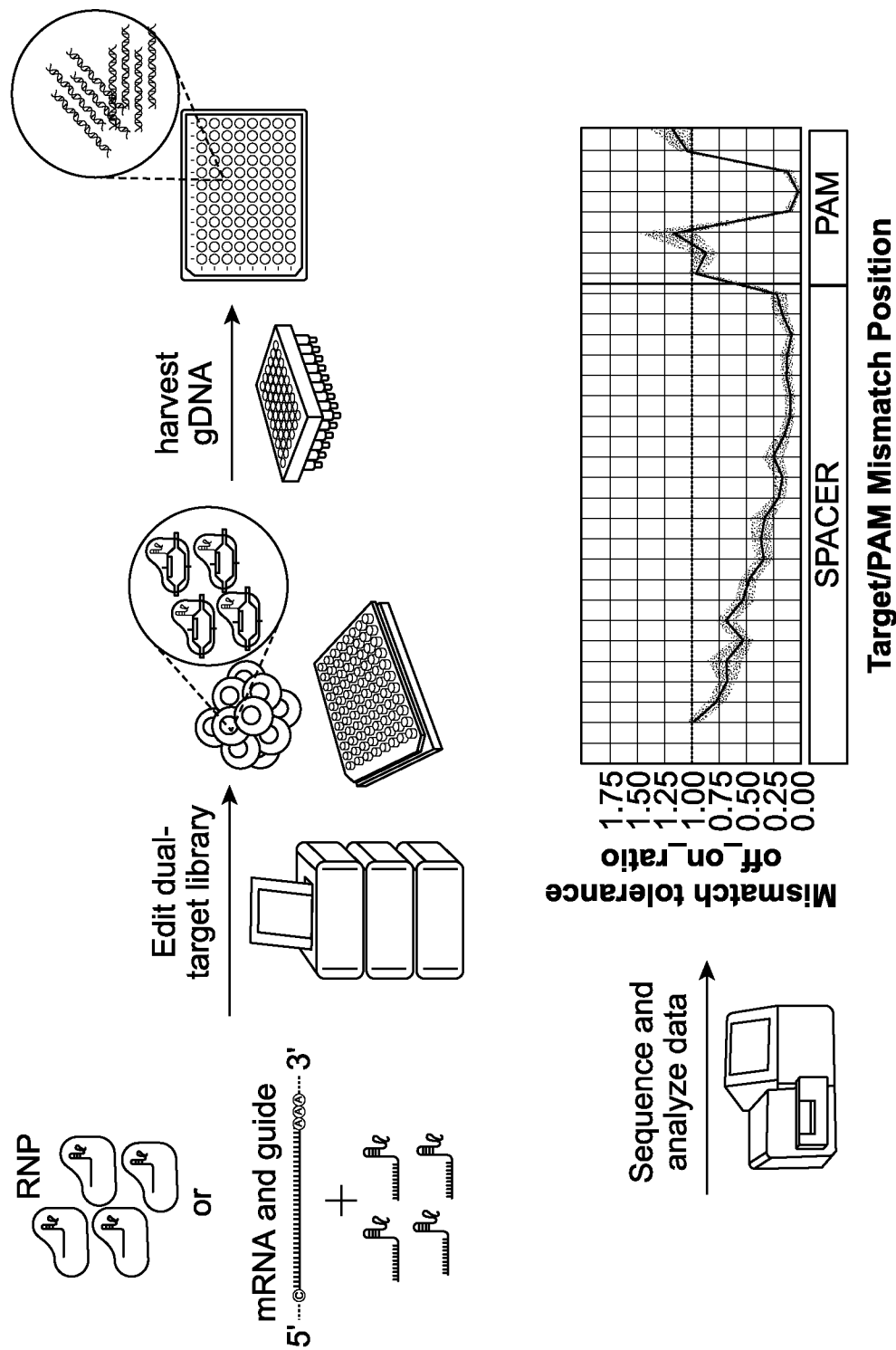


FIG. 45 (Cont. 2)

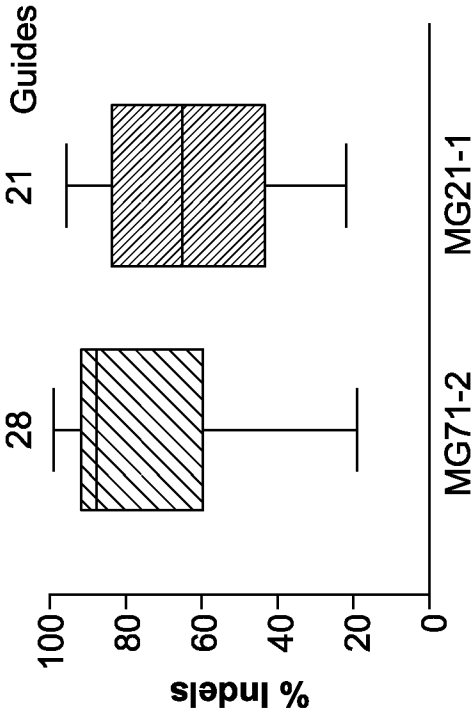


FIG. 46B

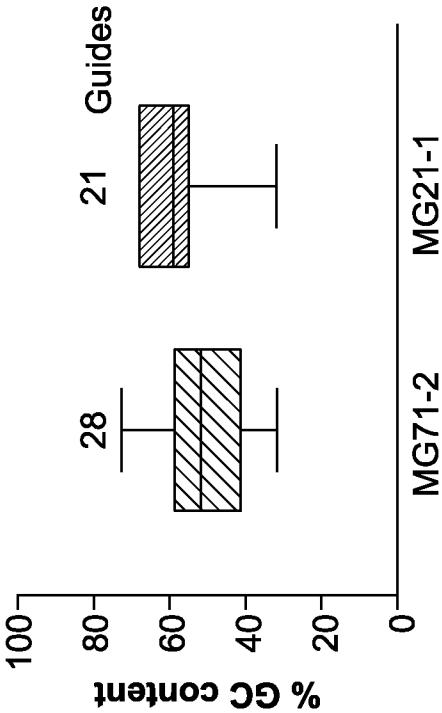


FIG. 46A

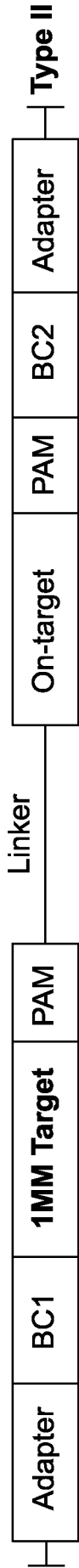


FIG. 47

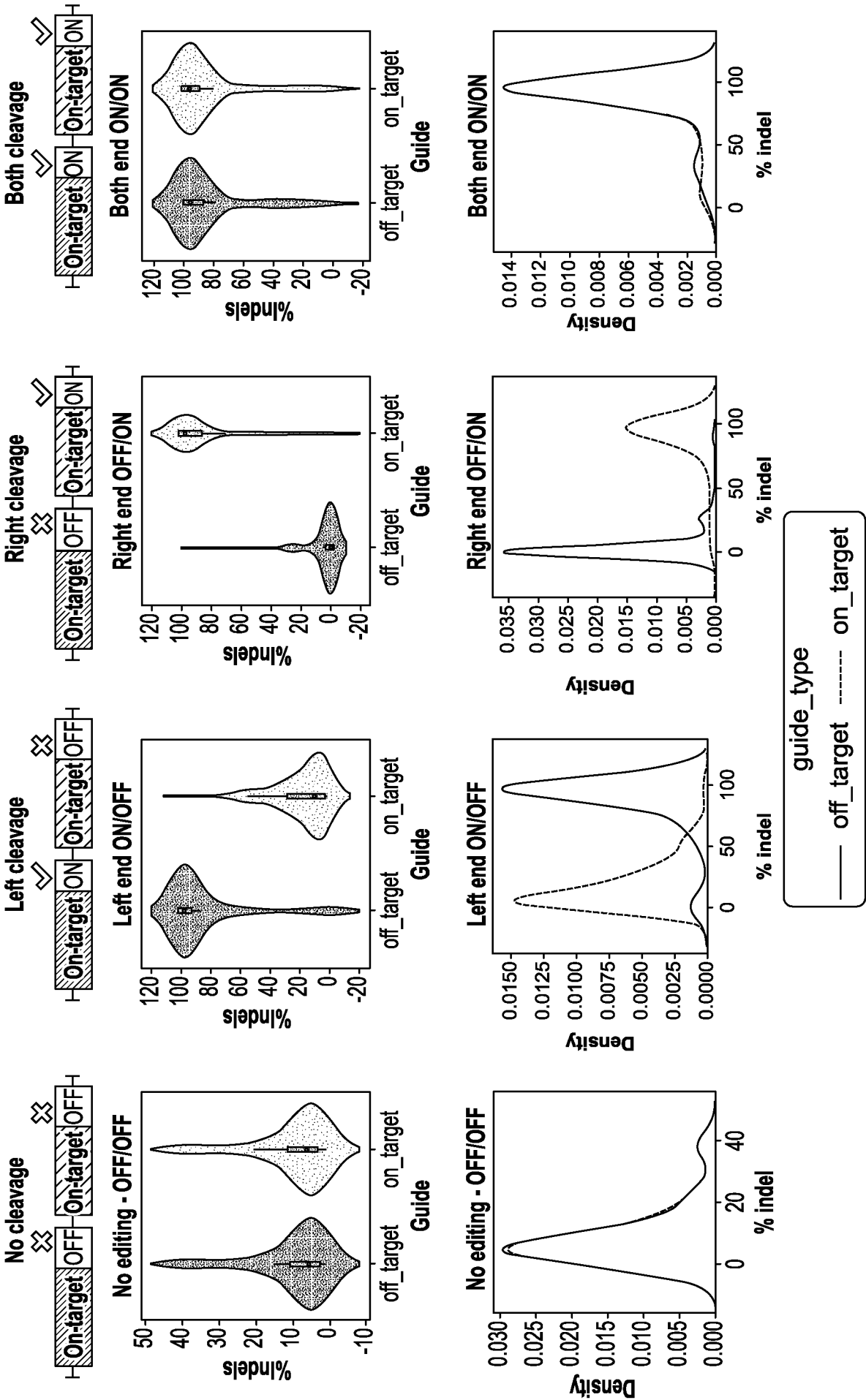


FIG. 48

MG71-2

Guide = TRAC-B2

Guide = HBB-B2

Guide = hB2M-A2

Guide = TRAC-H1

Guide = hB2M-G1

Guide = AAVS1-G2

Guide = AAVS1-F4

Guide = hB2M-D1

Guide = AAVS1-F2

Guide = HBB-H4

Guide = HBB-H3

Guide = HBB-A1

Guide = hB2M-B2

Guide = hB2M-C2

Guide = TRAC-C2

off_target on_target

off_target on_target

off_target on_target

off_target on_target

off_target on_target

guide_type

guide_type

guide_type

guide_type

guide_type

- Off target
- On target

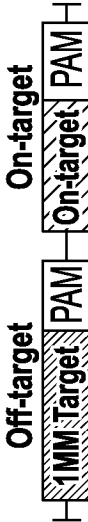


FIG. 49

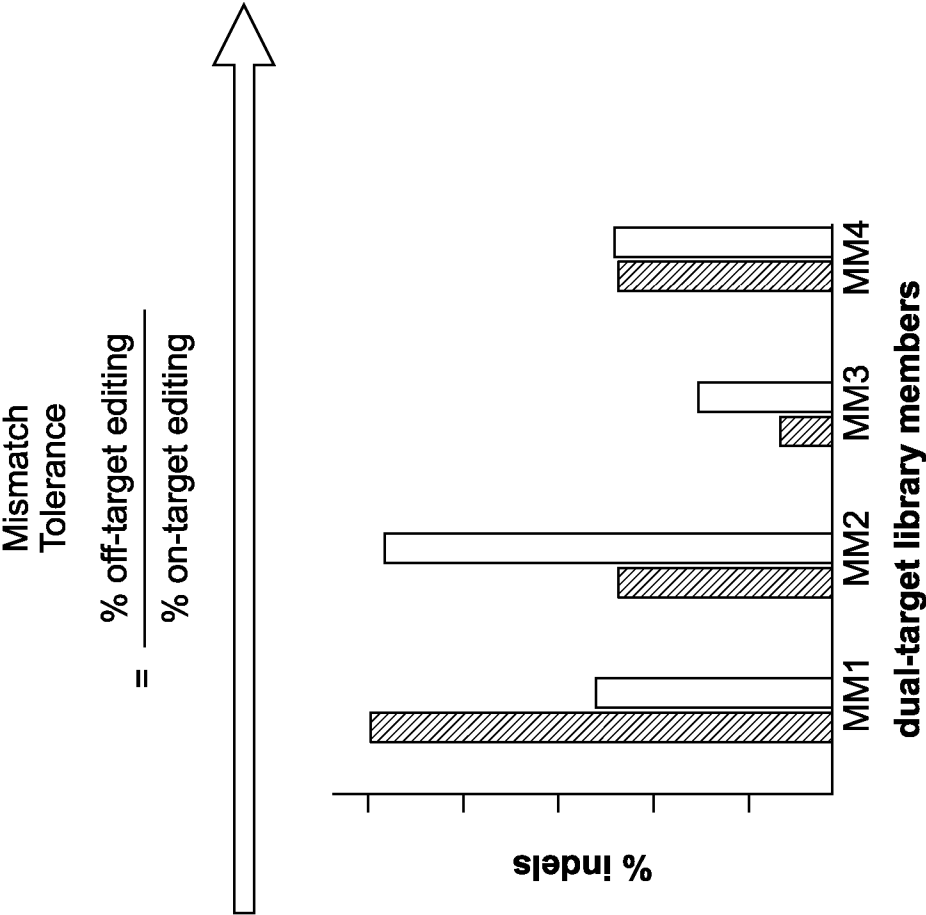


FIG. 49 (Cont. 1)

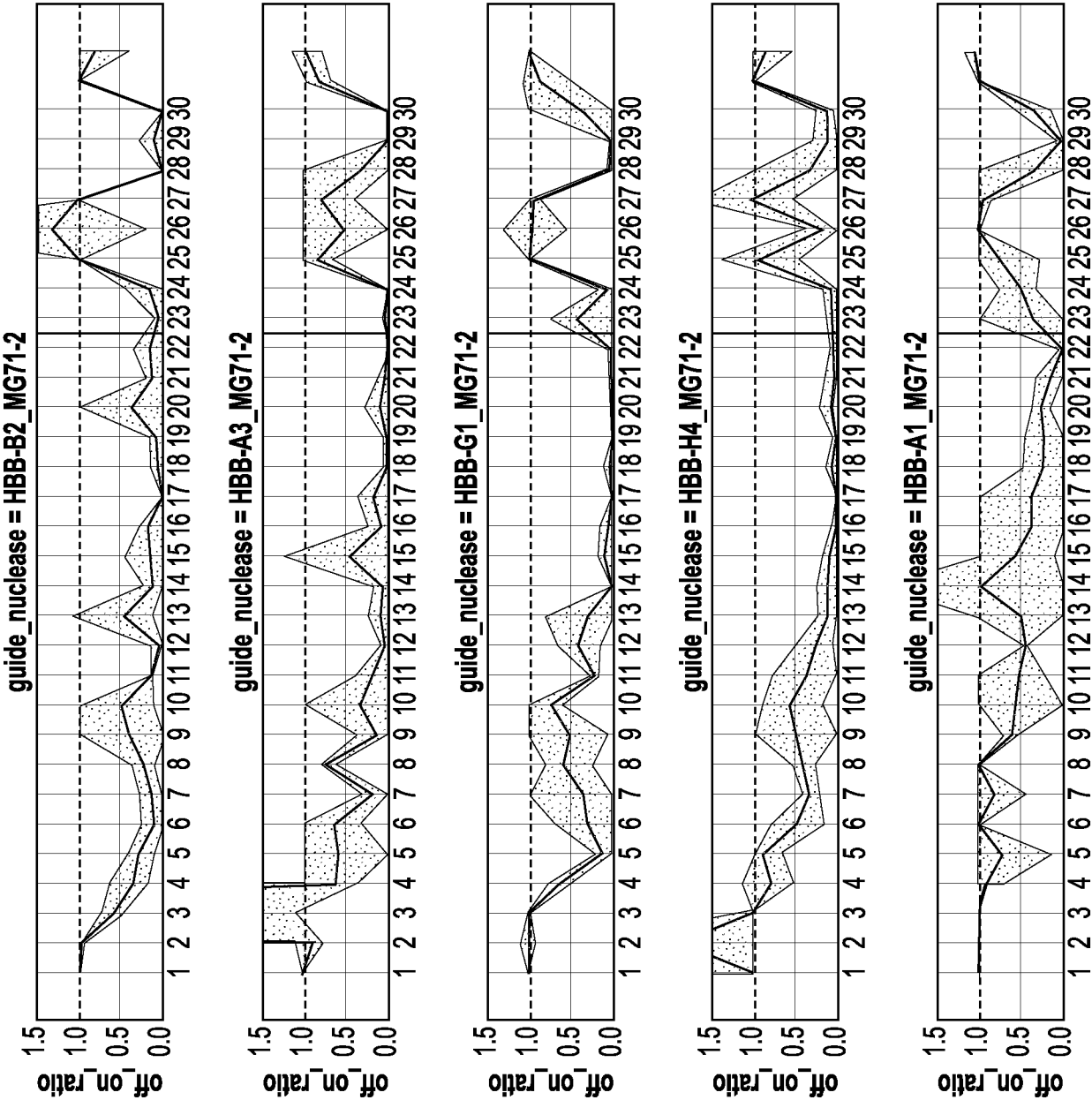


FIG. 49 (Cont. 2)

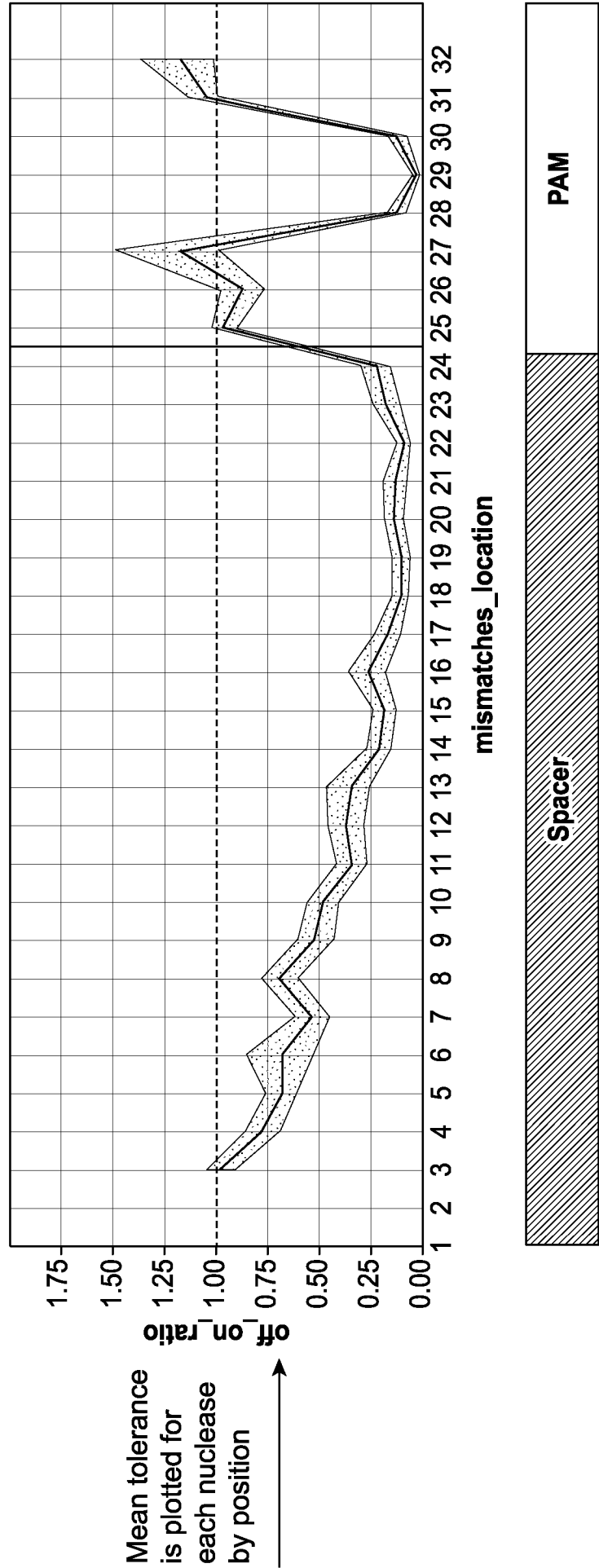


FIG. 49 (Cont. 3)

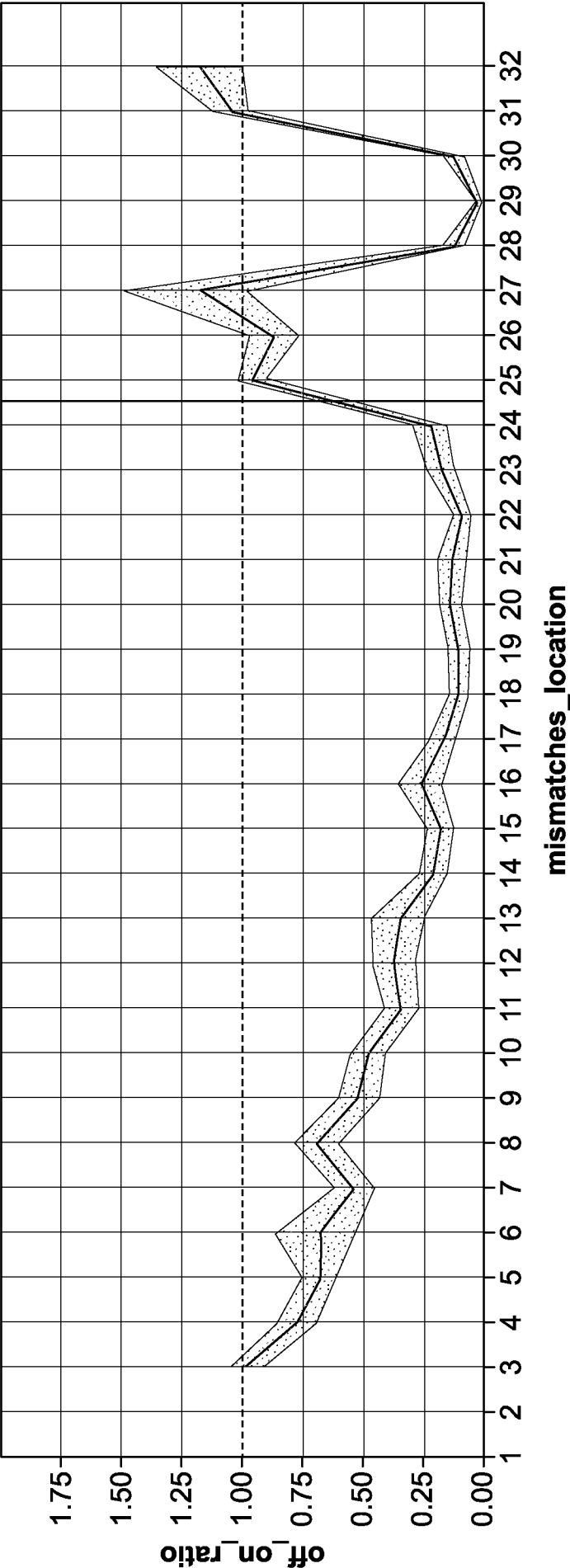


FIG. 50

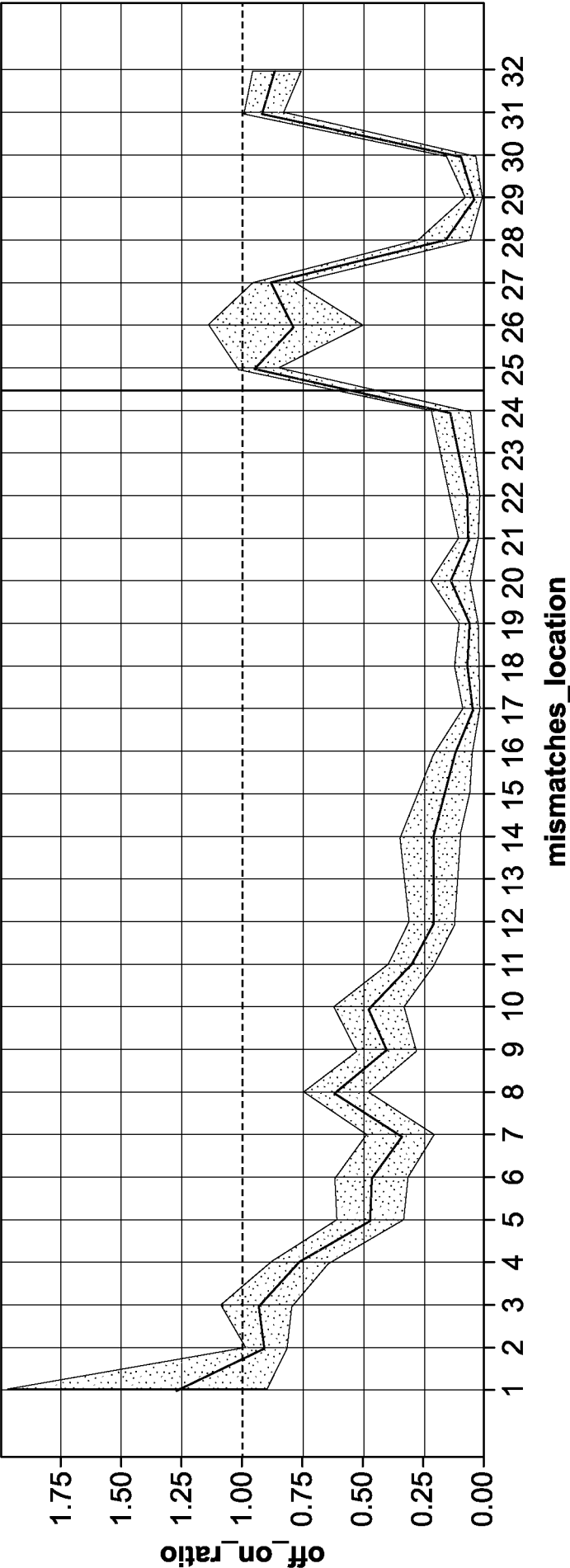


FIG. 51

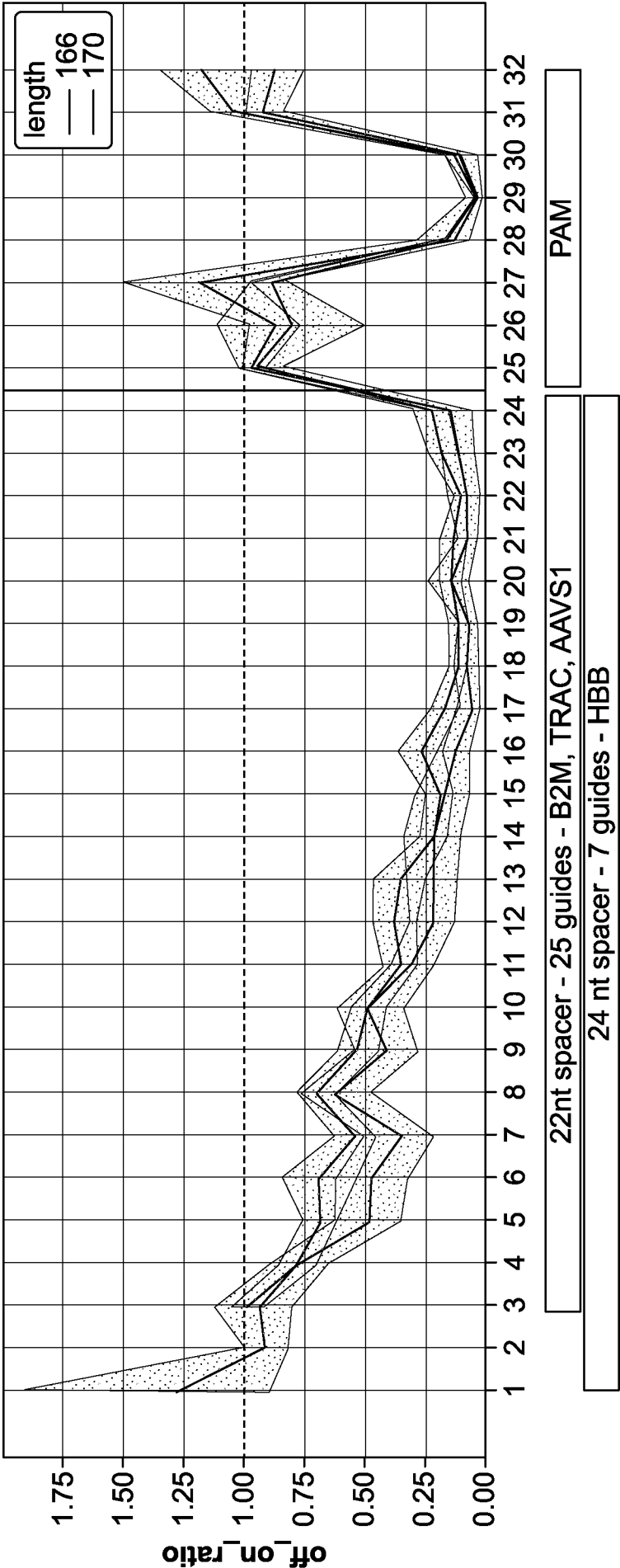


FIG. 52

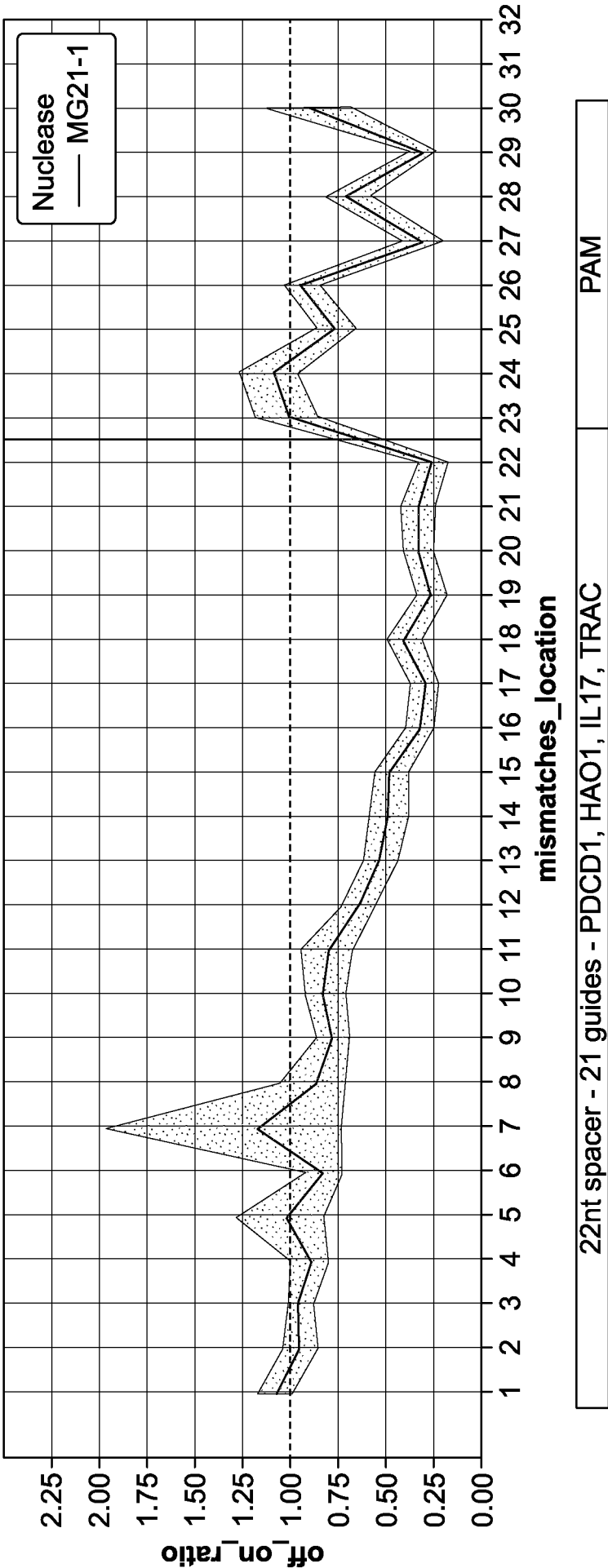


FIG. 53

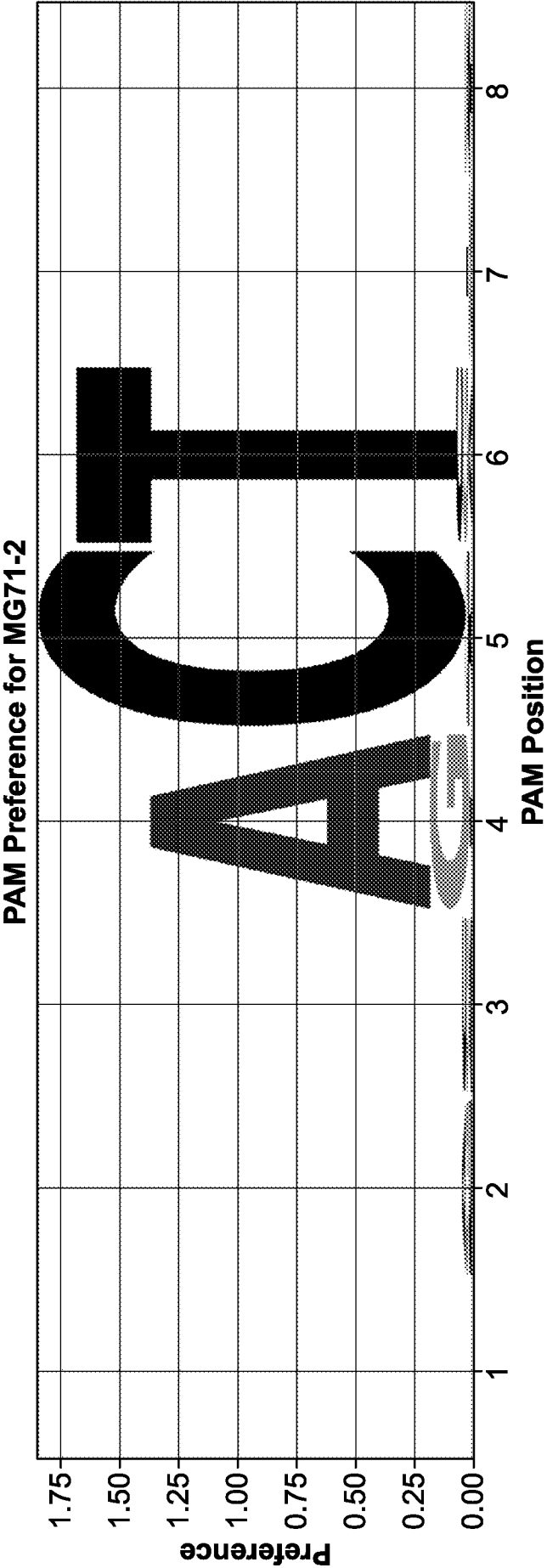


FIG. 54

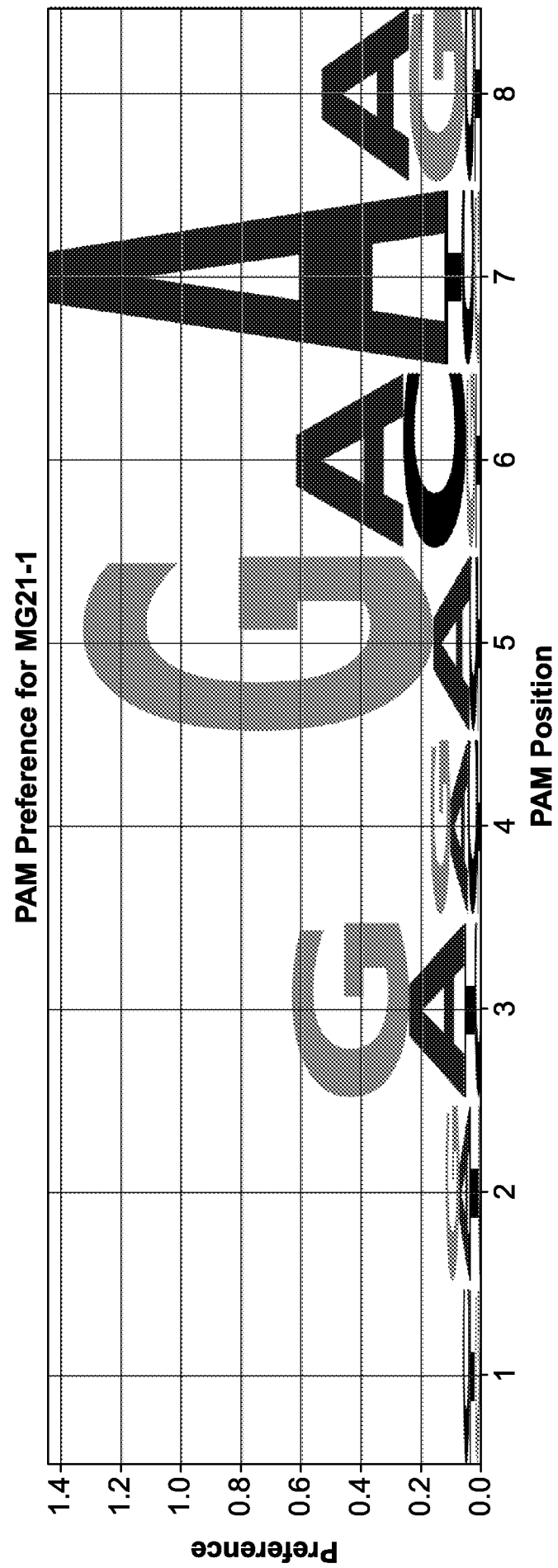


FIG. 55

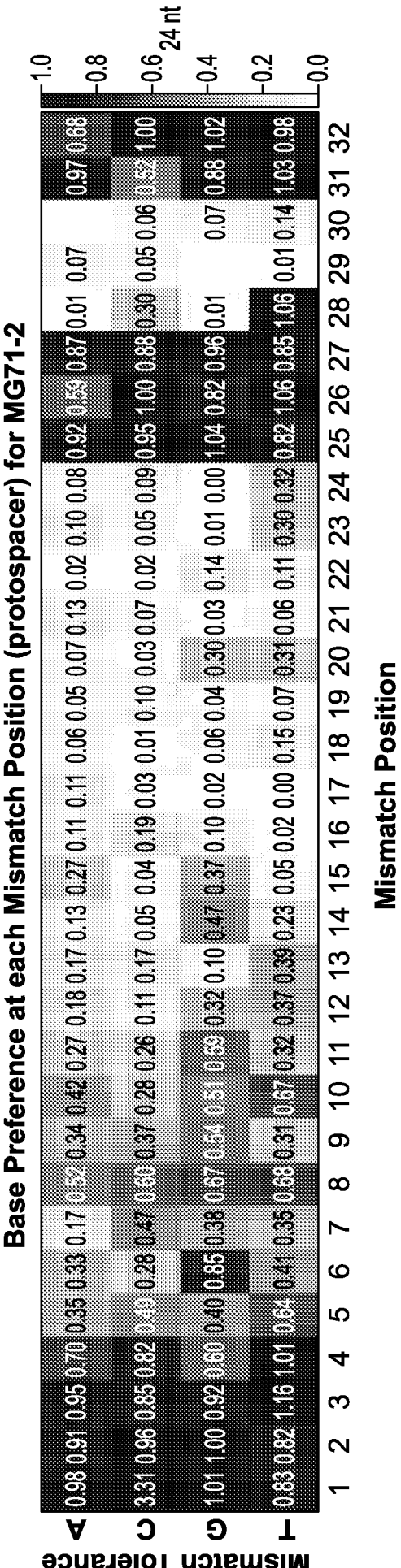
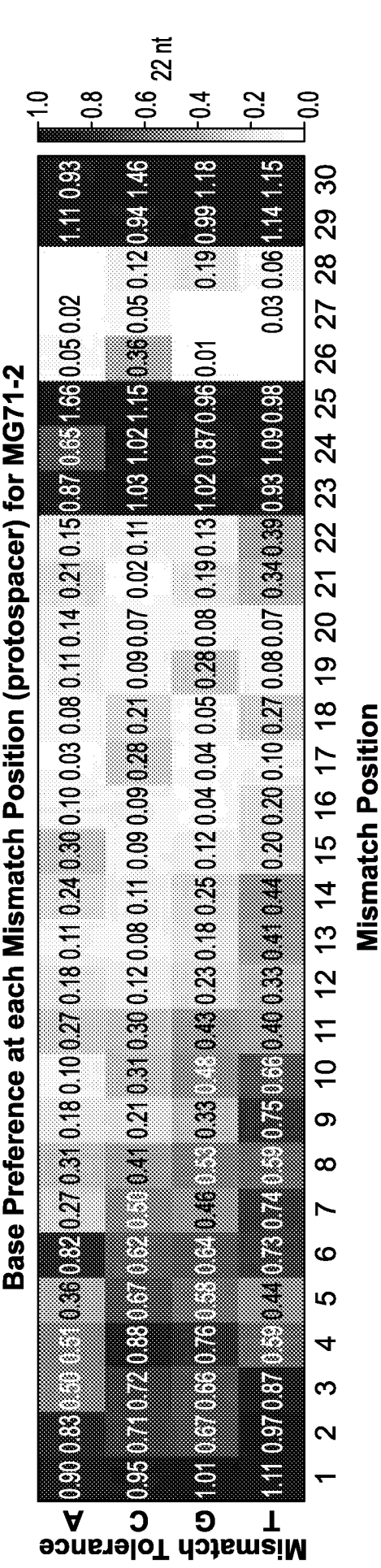


FIG. 56

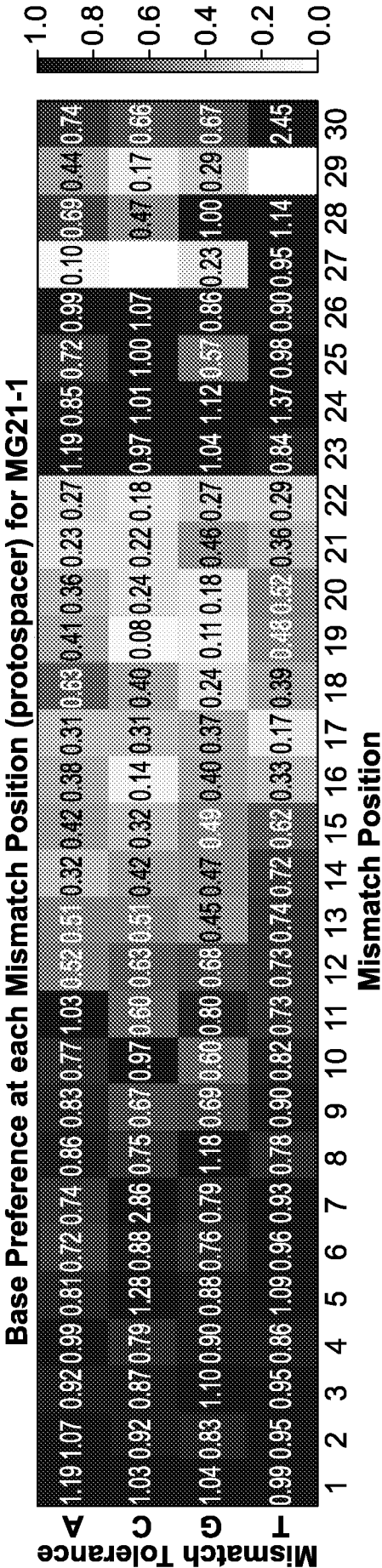


FIG. 57

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2024/056299

A. CLASSIFICATION OF SUBJECT MATTERIPC: **C12N 9/22** (2024.01); **C12N 15/10** (2024.01); **C12N 15/113** (2024.01); **C12N 15/52** (2024.01); **C12N 15/86** (2024.01); **C12N 15/90** (2024.01)CPC: **C12N 9/22**; **C12N 15/102**; **C12N 15/113**; **C12N 15/52**; **C12N 15/86**; **C12N 15/907**; **C12N 2310/20**; **C12N 2750/14143**

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History Document

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

See Search History Document

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

See Search History Document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2020/168291 A1 (METAGENOMI IP TECHNOLOGIES LLC) 20 August 2020 (20.08.2020) entire document	1-4, 14, 23
A	WO 2022/232638 A2 (METAGENOMI INC.) 03 November 2022 (03.11.2022) entire document	1-4, 14, 23
A	US 2022/0235340 A1 (THE BROAD INSTITUTE INC. et al.) 28 July 2022 (28.07.2022) entire document	1-4, 14, 23



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

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

“D” document cited by the applicant in the international application

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

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

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

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

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

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

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

“&” document member of the same patent family

Date of the actual completion of the international search

24 December 2024 (24.12.2024)

Date of mailing of the international search report

18 February 2025 (18.02.2025)

Name and mailing address of the ISA/US

COMMISSIONER FOR PATENTS
MAIL STOP PCT, ATTN: ISA/US
P.O. Box 1450
Alexandria, VA 22313-1450
UNITED STATES OF AMERICA

Authorized officer

TAINA MATOS

Facsimile No. 571-273-8300

Telephone No. 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2024/056299

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

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☒ forming part of the international application as filed.
 - b. ☐ furnished subsequent to the international filing date for the purposes of international search (Rule 13ter.1(a)),
☐ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.
2. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.
3. Additional comments:

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: **7-13, 25-39, 66-77**
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees need to be paid.

Group I+: claims 1-6 and 14-24 are drawn to engineered nuclease systems.

Group II+: claims 40 and 41 are drawn to methods of modifying a hydroxyacid oxidase 1 (HAO1) gene.

Group III+: claims 42-44 are drawn to methods of modifying ATPase copper transporting beta (ATP7B).

Group IV+: claims 45-47 are drawn to methods of modifying adeno-associated virus integration site 1 (AAVS1).

Group V+: claims 48-50 are drawn to methods of modifying T cell receptor alpha constant (TRAC).

Group VI+: claims 51-53 are drawn to methods of modifying an albumin gene.

Group VII+: claims 54-56 are drawn to methods of modifying a beta-2-microglobulin (B2M) gene.

Group VIII+: claims 57-59 are drawn to methods of modifying a hemoglobin subunit beta (HBB) gene.

Group IX+: claims 60-62 are drawn to methods of modifying a phenylalanine hydroxylase (PAH) gene.

Group X: claims 63-65 are drawn to methods of modifying an ataxin 2 (ATXN2) gene.

The first invention of Group I+ is restricted to an engineered nuclease system comprising: a) an endonuclease comprising SEQ ID NO: 47; and b) an engineered guide polynucleotide comprising SEQ ID NO: 1137. The first named invention has been selected based on the guidance set forth in section 10.54 of the PCT International Search and Preliminary Examination Guidelines. Specifically, the first named invention was selected based on the first listed compound species presented in the claims (see claims 1 and 14). It is believed that claims 1-4, 14, and 23 read on this first named invention and thus these claims will be searched without fee to the extent that they read on an engineered nuclease system comprising: a) an endonuclease comprising SEQ ID NO: 47; and b) an engineered guide polynucleotide comprising SEQ ID NO: 1137.

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

An exemplary election of Group II+ is a guide RNA comprising SEQ ID NO: 1283, and methods of modifying a hydroxyacid oxidase 1 (HAO1) gene comprising the same.

An exemplary election of Group III+ is an endonuclease comprising SEQ ID NO: 49, and methods of modifying ATPase copper transporting beta (ATP7B) comprising the same.

An exemplary election of Group IV+ is an endonuclease comprising SEQ ID NO: 49, a guide polynucleotide comprising SEQ ID NO: 710, and methods of modifying adeno-associated virus integration site 1 (AAVS1) comprising the same.

An exemplary election of Group V+ is an endonuclease comprising SEQ ID NO: 49, a guide polynucleotide comprising SEQ ID NO: 710, and methods of modifying T cell receptor alpha constant (TRAC) comprising the same.

An exemplary election of Group VI+ is an endonuclease comprising SEQ ID NO: 49, and methods of modifying an albumin gene comprising the same.

An exemplary election of Group VII+ is an endonuclease comprising SEQ ID NO: 49, a guide polynucleotide comprising SEQ ID NO: 710, and methods of modifying a beta-2-microglobulin (B2M) gene comprising the same.

An exemplary election of Group VIII+ is an endonuclease comprising SEQ ID NO: 49, a guide polynucleotide comprising SEQ ID NO: 710, and methods of modifying a hemoglobin subunit beta (HBB) gene comprising the same.

An exemplary election of Group IX+ is an endonuclease comprising SEQ ID NO: 49, and methods of modifying a phenylalanine hydroxylase (PAH) gene comprising the same.

Applicant is invited to elect nuclease systems comprising endonucleases and guide polynucleotides, and their respective, corresponding, SEQ ID NOs to be searched in a specific combination by paying an additional fee for each election. An exemplary election would be an engineered nuclease system comprising an endonuclease comprising SEQ ID NO: 48 and a guide polynucleotide comprising SEQ ID NO: 1153. Additional endonucleases and guide polynucleotides and their respective, corresponding, SEQ ID NOs will be searched upon the payment of additional fees. Applicants must specify the claims that read on any additional elected inventions. Applicants must further indicate, if applicable, the claims which read on the first named invention if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched/examined.

The inventions listed in Groups I+-X do not relate to a single general inventive concept under PCT Rule 13.1, because under PCT Rule 13.2 they lack the same or corresponding special technical features for the following reasons:

The Groups I+, II+, III+, IV+, V+, VI+, VII+, VIII+, and IX+ formulas do not share a significant structural element responsible for engineered nuclease systems requiring the selection of alternative endonucleases where "an endonuclease comprising a sequence having at least 70% sequence identity to any one of SEQ ID NOs: 47-52, 659-698, 983-1010, 1263-1282, 1158-1159, and 975-981."

The special technical features of Group I+, engineered nuclease systems, are not present in Groups II+-X; the special technical features of Group II+, methods of modifying a hydroxyacid oxidase 1 (HAO1) gene, are not present in Groups I+, or III+-X; the special technical features of Group III+, methods of modifying ATPase copper transporting beta (ATP7B), are not present in Groups I+, II+, or IV+-X; the special technical features of Group IV+, methods of modifying adeno-associated virus integration site 1 (AAVS1), are not present in Groups I+-III+, or V+-X; the special technical features of Group V+, methods of modifying T cell receptor alpha constant (TRAC), are not present in Groups I+-IV+ or VI+-X; the special technical features of Group VI+, methods of modifying an albumin gene, are not present in Groups I+-V+ or VII+-X; the special technical features of Group VII+, methods of modifying a beta-2-microglobulin (B2M) gene, are not present in Groups I+-VI+ or VIII+-X; the special technical features of Group VIII+, methods of modifying a hemoglobin subunit

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

beta (HBB) gene, are not present in Groups I+-VII+, IX+, or X; the special technical features of Group IX+, modifying a phenylalanine hydroxylase (PAH) gene, are not present in Groups I+-VIII+ or X; and the special technical features of Group X, methods of modifying an ataxin 2 (ATXN2) gene, are not present in Groups I+-IX+.

Additionally, even if Groups I+-X were considered to share the technical features of an engineered nuclease system comprising: a) an endonuclease; and b) an engineered guide polynucleotide configured to form a complex with the endonuclease and to hybridize to a target nucleic acid sequence; method of modifying a hydroxyacid oxidase 1 (HAO 1) gene comprising contacting the HAO1 gene using an engineered nuclease system comprising: a) an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NO: 47; and b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence; method of modifying ATPase copper transporting beta (ATP7B) comprising contacting ATP7B using an engineered nuclease system comprising: a) an endonuclease; and b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence; method of modifying adeno-associated virus integration site 1 (AAVS1) comprising contacting AAVS1 using an engineered nuclease system comprising: a) an endonuclease; and b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence; method of modifying T cell receptor alpha constant (TRAC) comprising contacting TRAC using an engineered nuclease system comprising: a) an endonuclease; and b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence; method of modifying an albumin gene comprising contacting the albumin gene using an engineered nuclease system comprising: a) an endonuclease; and b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence; method of modifying a beta-2-microglobulin (B2M) gene comprising contacting the B2M gene using an engineered nuclease system comprising: a) an endonuclease; and b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence; method of modifying a hemoglobin subunit beta (HBB) gene comprising contacting the HBB gene using an engineered nuclease system comprising: a) an endonuclease; and b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence; method of modifying a phenylalanine hydroxylase (PAH) gene comprising contacting the PAH gene using an engineered nuclease system comprising: a) an endonuclease; and b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence; and an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NO: 47, these shared technical features do not represent a contribution over the prior art.

Specifically, WO 2020/168291 A1 to Metagenomi IP Technologies LLC (hereinafter, "Metagenomi") teaches an engineered nuclease system comprising: a) an endonuclease; and b) an engineered guide polynucleotide configured to form a complex with the endonuclease and to hybridize to a target nucleic acid sequence (the present disclosure provides for an engineered nuclease system, comprising: (a) an endonuclease comprising a RuvC III domain and an HNH domain, wherein the endonuclease is derived from an uncultivated microorganism, wherein the endonuclease is a class 2, type II Cas endonuclease; and (b) an engineered guide ribonucleic acid structure configured to form a complex with the endonuclease comprising: (i) a guide ribonucleic acid sequence configured to hybridize to a target deoxyribonucleic acid sequence; and (ii) a tracr ribonucleic acid sequence configured to bind to the endonuclease, Para. [0004]); and an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NO: 47 (the endonuclease comprises a sequence at least 70%, 80%, or 90% identical to a sequence selected from the group consisting of SEQ ID NO: 1512, Para. [0028]; SEQ ID NO: 1512 of Metagenomi comprises 100% sequence identity to instant SEQ ID NO: 47).

Further, US 2022/0235340 A1 to The Broad Institute Inc. et al. (hereinafter, "Broad") teaches method of modifying a hydroxyacid oxidase 1 (HAO 1) gene comprising contacting the HAO1 gene using an engineered nuclease system comprising: a) an endonuclease; and b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence (the method can include delivering a composition, system, and/or component thereof to a subject or cell thereof, or to an infectious or symbiotic organism by a suitable delivery technique and/or composition, Para. [0454]; the disclosure provides non-naturally occurring or engineered DNA or RNA-targeting systems comprising a novel DNA or RNA-targeting CRISPR effector protein and at least one targeting nucleic acid component like a guide RNA and wherein the CRISPR effector protein is a Cas protein, Abstract; Exemplary chromosomal sequences can be

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

found in Table 2, Para. [0503]; Table 2 discloses HAO1); method of modifying ATPase copper transporting beta (ATP7B) comprising contacting ATP7B using an engineered nuclease system comprising: a) an endonuclease; and b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence (the method can include delivering a composition, system, and/or component thereof to a subject or cell thereof, or to an infectious or symbiotic organism by a suitable delivery technique and/or composition, Para. [0454]; the disclosure provides non-naturally occurring or engineered DNA or RNA-targeting systems comprising a novel DNA or RNA-targeting CRISPR effector protein and at least one targeting nucleic acid component like a guide RNA and wherein the CRISPR effector protein is a Cas protein, Abstract; Exemplary chromosomal sequences can be found in Table 2, Para. [0503]; Table 2 discloses ATP7B); method of modifying adeno-associated virus integration site 1 (AAVS1) comprising contacting AAVS1 using an engineered nuclease system comprising: a) an endonuclease; and b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence (the method can include delivering a composition, system, and/or component thereof to a subject or cell thereof, or to an infectious or symbiotic organism by a suitable delivery technique and/or composition, Para. [0454]; the disclosure provides non-naturally occurring or engineered DNA or RNA-targeting systems comprising a novel DNA or RNA-targeting CRISPR effector protein and at least one targeting nucleic acid component like a guide RNA and wherein the CRISPR effector protein is a Cas protein, Abstract; Directing of transgene(s) to a specific locus in a cell can minimize or avoid such risks and advantageously provide for uniform expression of the transgene(s) by the cells. Without limitation, suitable 'safe harbor' loci for directed transgene integration include... AAVS1, Para. [0567]); method of modifying T cell receptor alpha constant (TRAC) comprising contacting TRAC using an engineered nuclease system comprising: a) an endonuclease; and b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence (the method can include delivering a composition, system, and/or component thereof to a subject or cell thereof, or to an infectious or symbiotic organism by a suitable delivery technique and/or composition, Para. [0454]; the disclosure provides non-naturally occurring or engineered DNA or RNA-targeting systems comprising a novel DNA or RNA-targeting CRISPR effector protein and at least one targeting nucleic acid component like a guide RNA and wherein the CRISPR effector protein is a Cas protein, Abstract; Homology-directed repair (HDR) strategies are known and described elsewhere in this specification allowing to insert transgenes into desired loci (e.g., TRAC locus), Para. [0567]); method of modifying an albumin gene comprising contacting the albumin gene using an engineered nuclease system comprising: a) an endonuclease; and b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence (the method can include delivering a composition, system, and/or component thereof to a subject or cell thereof, or to an infectious or symbiotic organism by a suitable delivery technique and/or composition, Para. [0454]; the disclosure provides non-naturally occurring or engineered DNA or RNA-targeting systems comprising a novel DNA or RNA-targeting CRISPR effector protein and at least one targeting nucleic acid component like a guide RNA and wherein the CRISPR effector protein is a Cas protein, Abstract; A preferred target for the liver, whether in vitro or in vivo, is the albumin gene, Para. [0514]); method of modifying a beta-2-microglobulin (B2M) gene comprising contacting the B2M gene using an engineered nuclease system comprising: a) an endonuclease; and b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence (the method can include delivering a composition, system, and/or component thereof to a subject or cell thereof, or to an infectious or symbiotic organism by a suitable delivery technique and/or composition, Para. [0454]; the disclosure provides non-naturally occurring or engineered DNA or RNA-targeting systems comprising a novel DNA or RNA-targeting CRISPR effector protein and at least one targeting nucleic acid component like a guide RNA and wherein the CRISPR effector protein is a Cas protein, Abstract; In certain embodiments, editing of cells, particularly cells intended for adoptive cell therapies, more particularly immunoresponsive cells such as T cells, may be performed to knock-out or knock-down expression of one or more MHC constituent proteins, such as... beta-2 microglobulin (B2M), in a cell, Para. [0579]); method of modifying a hemoglobin subunit beta (HBB) gene comprising contacting the HBB gene using an engineered nuclease system comprising: a) an endonuclease; and b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence (the method can include delivering a composition, system, and/or component thereof to a subject or cell thereof, or to an infectious or symbiotic organism by a suitable delivery technique and/or composition, Para. [0454]; the disclosure provides non-naturally occurring or engineered DNA or RNA-targeting systems comprising a novel DNA or RNA-targeting CRISPR effector protein and at least one targeting nucleic acid component like a guide RNA and wherein the CRISPR effector protein is a Cas protein, Abstract; Exemplary chromosomal sequences can be found in Table 2, Para. [0503]; Table 2 discloses HBB); method of modifying a phenylalanine hydroxylase

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

(PAH) gene comprising contacting the PAH gene using an engineered nuclease system comprising: a) an endonuclease; and b) an engineered guide polynucleotide configured to form a complex with the endonuclease and hybridize to a target nucleic acid sequence (the method can include delivering a composition, system, and/or component thereof to a subject or cell thereof, or to an infectious or symbiotic organism by a suitable delivery technique and/or composition, Para. [0454]; the disclosure provides non-naturally occurring or engineered DNA or RNA-targeting systems comprising a novel DNA or RNA-targeting CRISPR effector protein and at least one targeting nucleic acid component like a guide RNA and wherein the CRISPR effector protein is a Cas protein, Abstract; Exemplary chromosomal sequences can be found in Table 2, Para. [0503]; Table 2 discloses PAH).

Broad fails to explicitly disclose an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NO: 47.

Metagenomi teaches an endonuclease comprising a sequence having at least 80% sequence identity to SEQ ID NO: 47 (the endonuclease comprises a sequence at least 70%, 80%, or 90% identical to a sequence selected from the group consisting of SEQ ID NO: 1512, Para. [0028]; SEQ ID NO: 1512 of Metagenomi comprises 100% sequence identity to instant SEQ ID NO: 47).

It would have been obvious to one of ordinary skill in the art before the priority date to modify the teachings of Broad by using an endonuclease as taught by Metagenomi. The motivation being to use an endonuclease having an amino acid sequence suitable for genetically modifying a target gene such as HAO1.

The inventions listed in Groups I+-X therefore lack unity under Rule 13 because they do not share a same or corresponding special technical features.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: **1-4, 14, 23**

- Remark on Protest**
- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.