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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 September 2015 (17.09.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/138510 A1

(51) International Patent Classification:
C12N 15/113 (2010.01) *C12N 15/86* (2006.01)
A61K 31/7088 (2006.01) *A61K 48/00* (2006.01)
C12N 9/22 (2006.01)

(21) International Application Number:
PCT/US2015/019790

(22) International Filing Date:
10 March 2015 (10.03.2015)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/950,733 10 March 2014 (10.03.2014) US
62/036,576 12 August 2014 (12.08.2014) US

(71) Applicant: **EDITAS MEDICINE., INC.** [US/US]; 300 Third Street, First Floor, Cambridge, MA 02142 (US).

(72) Inventors: **MAEDER, Morgan, L.**; 41 Spring Park Ave, Unit 1, Jamaica Pllain, MA 02130 (US). **BUMCROT, David, A.**; 30 Leicester Road, Belmont, MA 02478 (US). **SHEN, Shen**; 118 Hillside Road, Watertown, MA 02478 (US).

(74) Agent: **LU, Shihua**; Lando & Anastasi LLP, Riverfront Office Park, One Main Street, Suite 1100, Cambridge, MA 02142 (US).

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

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

Published:
— *with international search report (Art. 21(3))*

(54) Title: CRISPR/CAS-RELATED METHODS AND COMPOSITIONS FOR TREATING LEBER'S CONGENITAL AMAUROSIS 10 (LCA10)

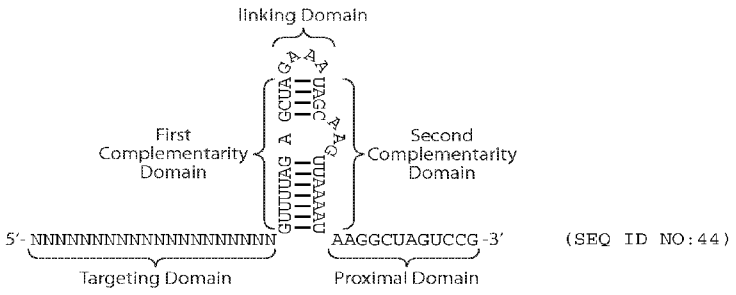


Fig. 1B

(57) Abstract: CRISPR/CAS-related compositions and methods for treatment of Leber's Congenital Amaurosis 10 (LCAIO) are disclosed.

WO 2015/138510 A1

CRISPR/CAS-RELATED METHODS AND COMPOSITIONS FOR TREATING LEBER'S CONGENITAL AMAUROSIS 10 (LCA10)

CROSS-REFERENCE TO RELATED APPLICATIONS

The present application claims the benefit of U.S. Provisional Application No. 61/950,733, filed March 10, 2014, and U.S. Provisional Application No. 62/036,576, filed August 12, 2014, the contents of which are hereby incorporated by reference in their entirety.

FIELD OF THE INVENTION

The invention relates to CRISPR/CAS-related methods and components for editing of a target nucleic acid sequence, and applications thereof in connection with Leber's Congenital Amaurosis 10 (LCA10).

BACKGROUND

Leber's congenital amaurosis (LCA) is the most severe form of inherited retinal dystrophy, with an onset of disease symptoms in the first years of life (Leber, T., Archiv für Ophthalmologie (in German). 15 (3): 1–25, 1869) and an estimated prevalence of approximately 1 in 50,000 worldwide (Koenekoop et al., Clin Experiment Ophthalmol. 35(5): 473-485, 2007; Stone, Am J Ophthalmol. 144(6): 791-811, 2007). Genetically, LCA is a heterogeneous disease. To date, fifteen genes have been identified with mutations that result in LCA (den Hollander et al., Prog Retin Eye Res. 27(4): 391-419, 2008; Estrada-Cuzcano et al., Invest Ophthalmol Vis Sci. 52(2): 834-9, 2011). The *CEP290* gene is the most frequently mutated LCA gene accounting for approximately 15% of all cases (Stone, Am J Ophthalmol. 144(6): 791-811, 2007; den Hollander et al., Prog Retin Eye Res. 27(4): 391-419, 2008; den Hollander et al., Am J Hum Genet. 79(3): 556-561, 2006; Perrault et al., Hum Mutat. 28(4):4 16, 2007). Severe mutations in *CEP290* have also been reported to cause systemic diseases that are characterized by brain defects, kidney malformations, polydactyly and/or obesity (Baal et al., Am J Hum Genet. 81, 170-179, 2007; den Hollander et al., Prog Retin Eye Res. 27(4): 391-419, 2008; Helou et al., J Med Genet. 44: 657-663, 2007; Valente et al., Nat Genet. 38: 623-625, 2006). Patients with LCA and early-onset retinal dystrophy often carry hypomorphic *CEP290* alleles (Stone, Am J Ophthalmol. 144(6): 791-811, 2007; den Hollander et al., Am J Hum Genet. 79(3): 556-561,

2006; Perrault et al., Hum Mutat. 28(4):4 16, 2007; Coppieters et al., Hum Mutat 31, E1709-E1766. 2010; Littink et al., Invest Ophthalmol Vis Sci 51, 3646-3652, 2010).

LCA, and other retinal dystrophies such as Retinitis Pigmentosa (RP), have long been considered incurable diseases. However, the first phase I/II clinical trials using gene augmentation therapy have led to promising results in a selected group of adult LCA/RP patients with mutations in the RPE65 gene (Bainbridge et al., N Engl J Med. 358, 2231-2239, 2008; Cideciyan et al., Proc Natl Acad Sci U S A. 105, 15112-15117, 2008; Hauswirth et al., N Engl J Med. 358, 2240-2248, 2008; Maguire et al., N Engl J Med. 358: 2240-2248, 2008). Unilateral subretinal injections of adeno-associated virus particles carrying constructs encoding the wild-type RPE65 cDNA were shown to be safe and moderately effective in some patients, without causing any adverse effects. In a follow-up study including adults and children, visual improvements were more sustained, especially in the children all of whom gained ambulatory vision (Maguire et al., Lancet. 374, 1597-1605, 2009). Although these studies demonstrated the potential to treat LCA using gene augmentation therapy and increased the development of therapeutic strategies for other genetic subtypes of retinal dystrophies (den Hollander et al., J Clin Invest 120: 3042-3053, 2010), it is hard to control the expression levels of the therapeutic genes when using gene augmentation therapy.

Leber's congenital amaurosis 10 (LCA10), one type of LCA, is is an inherited (autosomal recessive) retinal degenerative disease characterized by severe loss of vision at birth. All subjects having LCA10 have had at least one c.2991+1655A to G (adenine to guanine) mutation in the *CEP290* gene. Heterozygous nonsense, frameshift, and splice-site mutations have been identified on the remaining allele. A c.2991+1655A to G mutation in the *CEP290* gene give rise to a cryptic splice donor cite in intron 26 which results in the inclusion of an aberrant exon of 128 bp in the mutant *CEP290* mRNA, and inserts a premature stop codon (P.C998X). The sequence of the cryptic exon contains part of an *Alu* repeat. There are currently no approved therapeutics for LCA10.

Despite advances that have been made using gene therapy, there remains a need for therapeutics to treat retinal dystrophies, including LCA10.

SUMMARY OF THE INVENTION

Methods and compositions discussed herein, provide for treating or delaying the onset or progression of diseases of the eye, e.g., disorders that affect retinal cells, e.g., photoreceptor cells.

Methods and compositions discussed herein, provide for treating or delaying the onset or progression of Leber's Congenital Amaurosis 10 (LCA10), an inherited retinal degenerative disease characterized by severe loss of vision at birth. LCA10 is caused by a mutation in the *CEP290* gene, e.g., a c.2991+1655A to G (adenine to guanine) mutation in the *CEP290* gene which gives rise to a cryptic splice site in intron 26. This is a mutation at nucleotide 1655 of intron 26 of *CEP290*, e.g., an A to G mutation. *CEP290* is also known as: *CT87; MKS4; POC3; rd16; BBS14; JBTS5; LCA10; NPHP6; SLSN6;* and *3H11Ag.*

Methods and compositions discussed herein, provide for treating or delaying the onset or progression of LCA10 by gene editing, e.g., using CRISPR-Cas9 mediated methods to alter a LCA10 target position, as disclosed below.

"LCA10 target position", as used herein, refers to nucleotide 1655 of intron 26 of the *CEP290* gene, and the mutation at that site that gives rise to a cryptic splice donor site in intron 26 which results in the inclusion of an aberrant exon of 128bp (c.2991+1523 to c.2991+1650) in the mutant *CEP290* mRNA, and inserts a premature stop codon (p.C998X). The sequence of the cryptic exon contains part of an *Alu* repeat region. The *Alu* repeats span from c.2991+1162 to c.2991+1638. In an embodiment, the LCA10 target position is occupied by an adenine (A) to guanine (G) mutation (c.2991+1655A to G).

In one aspect, methods and compositions discussed herein, provide for altering a LCA10 target position in the *CEP290* gene. The methods and compositions described herein introduce one or more breaks near the site of the LCA target position (e.g., c.2991+1655A to G) in at least one allele of the *CEP290* gene. Altering the LCA10 target position refers to (1) break-induced introduction of an indel (also referred to herein as NHEJ-mediated introduction of an indel) in close proximity to or including a LCA10 target position (e.g., c.2991+1655A to G), or (2) break-induced deletion (also referred to herein as NHEJ-mediated deletion) of genomic sequence including the mutation at a LCA10 target position (e.g., c.2991+1655A to G). Both approaches give rise to the loss or destruction of the cryptic splice site resulting from the mutation at the LCA10 target position (e.g., c.2991+1655A to G).

In an embodiment, a single strand break is introduced in close proximity to or at the LCA10 target position (e.g., c.2991+1655A to G) in the *CEP290* gene. While not wishing to be bound by theory, it is believed that break-induced indels (e.g., indels created following NHEJ) destroy the cryptic splice site. In an embodiment, the single strand break will be accompanied by an additional single strand break, positioned by a second gRNA molecule.

In an embodiment, a double strand break is introduced in close proximity to or at the LCA10 target position (e.g., c.2991+1655A to G) in the *CEP290* gene. While not wishing to be bound by theory, it is believed that break-induced indels (e.g., indels created following NHEJ) destroy the cryptic splice site. In an embodiment, a double strand break will be accompanied by an additional single strand break may be positioned by a second gRNA molecule. In an embodiment, a double strand break will be accompanied by two additional single strand breaks positioned by a second gRNA molecule and a third gRNA molecule.

In an embodiment, a pair of single strand breaks is introduced in close proximity to or at the LCA10 target position (e.g., c.2991+1655A to G) in the *CEP290* gene. While not wishing to be bound by theory, it is believed that break-induced indels destroy the cryptic splice site. In an embodiment, the pair of single strand breaks will be accompanied by an additional double strand break, positioned by a third gRNA molecule. In an embodiment, the pair of single strand breaks will be accompanied by an additional pair of single strand breaks positioned by a third gRNA molecule and a fourth gRNA molecule.

In an embodiment, two double strand breaks are introduced to flank the LCA10 target position in the *CEP290* gene (one 5' and the other one 3' to the mutation at the LCA10 target position, e.g., c.2991+1655A to G) to remove (e.g., delete) the genomic sequence including the mutation at the LCA10 target position. It is contemplated herein that in an embodiment the break-induced deletion of the genomic sequence including the mutation at the LCA10 target position is mediated by NHEJ. In an embodiment, the breaks (i.e., the two double strand breaks) are positioned to avoid unwanted target chromosome elements, such as repeat elements, e.g., an *Alu* repeat. The breaks, i.e., two double strand breaks, can be positioned upstream and downstream of the LCA10 target position, as discussed herein.

In an embodiment, one double strand break (either 5' or 3' to the mutation at the LCA10 target position, e.g., c.2991+1655A to G) and two single strand breaks (on the other side of the mutation at the LCA10 target position from the double strand break) are introduced to flank the

LCA10 target position in the *CEP290* gene to remove (e.g., delete) the genomic sequence including the mutation at the LCA10 target position. It is contemplated herein that in an embodiment the break-induced deletion of the genomic sequence including the mutation at the LCA10 target position is mediated by NHEJ. In an embodiment, the breaks (i.e., the double strand break and the two single strand breaks) are positioned to avoid unwanted target chromosome elements, such as repeat elements, e.g., an *Alu* repeat. The breaks, e.g., one double strand break and two single strand breaks, can be positioned upstream and downstream of the LCA10 target position, as discussed herein.

In an embodiment, two pairs of single strand breaks (two 5' and the other two 3' to the mutation at the LCA10 target position, e.g., c.2991+1655A to G) are introduced to flank the LCA10 target position in the *CEP290* gene to remove (e.g., delete) the genomic sequence including the mutation at the LCA10 target position. It is contemplated herein that in an embodiment the break-induced deletion of the genomic sequence including the mutation at the LCA10 target position is mediated by NHEJ. In an embodiment, the breaks (e.g., two pairs of single strand breaks) are positioned to avoid unwanted target chromosome elements, such as repeat elements, e.g., an *Alu* repeat. The breaks, e.g., two pairs of single strand breaks, can be positioned upstream or downstream of the LCA10 target position, as discussed herein.

The LCA10 target position may be targeted by cleaving with either a single nuclease or dual nickases, e.g., to induce break-induced indel in close proximity to or including the LCA10 target position or break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene. The method can include acquiring knowledge of the mutation carried by the subject, e.g., by sequencing the appropriate portion of the *CEP290* gene.

In one aspect, disclosed herein is a gRNA molecule, e.g., an isolated or non-naturally occurring gRNA molecule, comprising a targeting domain which is complementary with a target domain from the *CEP290* gene.

When two or more gRNAs are used to position two or more cleavage events, e.g., double strand or single strand breaks, in a target nucleic acid, it is contemplated that in an embodiment the two or more cleavage events may be made by the same or different Cas9 proteins. For example, when two gRNAs are used to position two double strand breaks, a single Cas9 nuclease may be used to create both double strand breaks. When two or more gRNAs are used to position two or more single stranded breaks (single strand breaks), a single Cas9 nickase may be used to

create the two or more single strand breaks. When two or more gRNAs are used to position at least one double strand break and at least one single strand break, two Cas9 proteins may be used, e.g., one Cas9 nuclease and one Cas9 nickase. It is contemplated that in an embodiment when two or more Cas9 proteins are used that the two or more Cas9 proteins may be delivered sequentially to control specificity of a double strand versus a single strand break at the desired position in the target nucleic acid.

In some embodiments, the targeting domain of the first gRNA molecule and the targeting domain of the second gRNA molecule hybridize to the target domain from the target nucleic acid molecule (i.e., the *CEP290* gene) through complementary base pairing to opposite strands of the target nucleic acid molecule. In some embodiments, the first gRNA molecule and the second gRNA molecule are configured such that the PAMs are oriented outward.

In an embodiment, the targeting domain of a gRNA molecule is configured to avoid unwanted target chromosome elements, such as repeat elements, e.g., an Alu repeat, or the endogenous *CEP290* splice sites, in the target domain. The gRNA molecule may be a first, second, third and/or fourth gRNA molecule.

In an embodiment, the targeting domain of a gRNA molecule is configured to position a cleavage event sufficiently far from a preselected nucleotide, e.g., the nucleotide of a coding region, such that the nucleotide is not altered. In an embodiment, the targeting domain of a gRNA molecule is configured to position an intronic cleavage event sufficiently far from an intron/exon border, or naturally occurring splice signal, to avoid alteration of the exonic sequence or unwanted splicing events. The gRNA molecule may be a first, second, third and/or fourth gRNA molecule, as described herein.

In an embodiment, the LCA10 target position in the *CEP290* gene is targeted. In an embodiment, the targeting domain comprises a sequence that is the same as, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from, a targeting domain sequence from **Table 10**. In some embodiments, the targeting domain is selected from those in **Table 10**. For example, in certain embodiments, the targeting domain is:

GACACTGCCAATAGGGATAGGT;
GTCAAAAGCTACCGGTTACCTG;
GTTCTGTCCTCAGTAAAAGGTA;
GAATAGTTTGTTCTGGGTAC;
GAGAAAGGGATGGGCACTTA;
GATGCAGAACTAGTGTAGAC;

GTCACATGGGAGTCACAGGG; or
GAGTATCTCCTGTTTGGCA.

In an embodiment, when two or more gRNAs are used to position two or more breaks, e.g., two or more single strand breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Table 10**. In an embodiment, the two or more gRNAs or targeting domains are selected from one or more of the pairs of gRNAs or targeting domains described herein, e.g., as indicated in **Table 10**. In an embodiment, when two or more gRNAs are used to position four breaks, e.g., four single strand breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Table 10**.

In an embodiment, the LCA10 target position in the *CEP290* gene is targeted. In an embodiment, the targeting domain comprises a sequence that is the same as, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from, a targeting domain sequence from **Table 1A-1D**. In some embodiments, the targeting domain is selected from those in **Table 1A-1D**. For example, in certain embodiments, the targeting domain is:

GAGAUACUCACAAUUACAAC; or
GAUACUCACAAUUACAACUG.

In an embodiment, when two or more gRNAs are used to position two or more breaks, e.g., two or more single strand breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Tables 1A-1D**. In an embodiment, when two or more gRNAs are used to position four breaks, e.g., four single strand breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Tables 1A-1D**.

In an embodiment, the LCA10 target position in the *CEP290* gene is targeted. In an embodiment, the targeting domain comprises a sequence that is the same as, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from, a targeting domain sequence from **Tables 2A-2C**. In some embodiments, the targeting domain is selected from those in **Tables 2A-2C**. For example, in certain embodiments, the targeting domain is:

GAGAUACUCACAAUUACAAC; or
GAUACUCACAAUUACAA.

In an embodiment, when two or more gRNAs are used to position two or more breaks, e.g., two or more single stranded breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Tables 2A-2C**. In an embodiment, when two or more

gRNAs are used to position four breaks, e.g., four single strand breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Tables 2A-2C**.

In an embodiment, the LCA10 target position in the *CEP290* gene is targeted. In an embodiment, the targeting domain comprises a sequence that is the same as, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from, a targeting domain sequence from **Tables 6A-6D**. In some embodiments, the targeting domain is selected from those in **Tables 6A-6D**. For example, in certain embodiments, the targeting domain is:

GCACCUGGCCCCAGUUGUAAUU.

In an embodiment, when two or more gRNAs are used to position two or more breaks, e.g., two or more single stranded breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Tables 6A-6D**. In an embodiment, when two or more gRNAs are used to position four breaks, e.g., four single strand breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Tables 6A-6D**.

In an embodiment, the LCA10 target position in the *CEP290* gene is targeted. In an embodiment, the targeting domain comprises a sequence that is the same as, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from, a targeting domain sequence from **Tables 3A-3D**. In some embodiments, the targeting domain is selected from those in **Tables 3A-3D**. For example, in certain embodiments, the targeting domain is:

GCUACCGGUUACCUGAA;
GCAGAACUAGUGUAGAC;
GUUGAGUAUCUCCUGUU;
GAUGCAGAACUAGUGUAGAC; or
GCUUGAACUCUGUGCCAAAC.

In an embodiment, when two or more gRNAs are used to position two or more breaks, e.g., two or more single stranded breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Tables 3A-3D**. In an embodiment, when two or more gRNAs are used to position four breaks, e.g., four single strand breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Tables 3A-3D**.

In an embodiment, the LCA10 target position in the *CEP290* gene is targeted. In an embodiment, the targeting domain comprises a sequence that is the same as, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from, a targeting domain sequence from **Tables 7A-7D**. In

some embodiments, the targeting domain is selected from those in **Tables 7A-7D**. For example, in certain embodiments, the targeting domain is:

GAAAGAUGAAAAAUACUCUU;
GAAAUAGAUGUAGAUUG;
GAAAUAUUAAGGGCUCUUCC;
GAACAAAAGCCAGGGACCAU;
GAACUCUAUACCUUUUACUG;
GAAGAAUGGAAUAGAUAAUA;
GAAUAGUUUGUUCUGGGUAC;
GAAUGGAAUAGAUAAUA;
GAAUUUACAGAGUGCAUCCA;
GAGAAAAAGGAGCAUGAAAC;
GAGAGCCACAGUGCAUG;
GAGGUAGAAUCAAGAAG;
GAGUGCAUCCAUGGUCC;
GAUAACUACAAAGGGUC;
GAUAGAGACAGGAAUAA;
GAUGAAAAAUACUCUUU;
GAUGACAUGAGGUAAGU;
GAUGCAGAACUAGUGUAGAC;
GCAGAACUAGUGUAGAC;
GCAUGUGGUGUCAAAUA;
GCCUGAACAAGUUUUGAAAC;
GCUACCGGUUACCUGAA;
GCUCUUUUCUAUAUAUA;
GCUUGAACUCUGUGCCAAAC;
GCUUUUGACAGUUUUUAAGG;
GCUUUUGUUCCUUGGAA;
GGAACAAAAGCCAGGGACCA;
GGACUUGACUUUUACCCUUC;
GGAGAAUAGUUUGUUCU;
GGAGUCACAUGGGAGUCACA;
GGAUAGGACAGAGGACA;
GGCUGUAAGAUAACUACAAA;
GGGAGAAUAGUUUGUUC;
GGGAGUCACAUGGGAGUCAC;
GGGCUCUUCCUGGACCA;
GGGUACAGGGGUAAGAGAAA;
GGUCCCUGGCUUUUGUUCCU;
GUAAAGGUUCAUGAGACUAG;
GUAACAUAAUCACCUCUCUU;
GUAAGACUGGAGAUAGAGAC;
GUACAGGGGUAAGAGAA;

GUAGCUUUUGACAGUUUUUA;
GUCACAUGGGAGUCACA;
GUGGAGAGCCACAGUGCAUG;
GUUACAAUCUGUGAAUA;
GUUCUGUCCUCAGUAAA;
GUUGAGUAUCUCCUGUU;
GUUUAGAAUGAUCAUUCUUG;
GUUUGUUCUGGGUACAG;
UAAAAACUGUCAAAAGCUAC;
UAAAAGGUAUAGAGUUCAAG;
UAAAUCAUGCAAGUGACCUA; or
UAAGAUAACUACAAAGGGUC.

In an embodiment, when two or more gRNAs are used to position two or more breaks, e.g., two or more single stranded breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Tables 7A-7D**. In an embodiment, when two or more gRNAs are used to position four breaks, e.g., four single strand breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Tables 7A-7D**.

In an embodiment, the LCA10 target position in the *CEP290* gene is targeted. In an embodiment, the targeting domain comprises a sequence that is the same as, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from, a targeting domain sequence from **Table 4A-4D**. In some embodiments, the targeting domain is selected from those in **Table 4A-4D**. For example, in certain embodiments, the targeting domain is:

GAAUCCUGAAAGCUACU.

In an embodiment, when two or more gRNAs are used to position two or more breaks, e.g., two or more single stranded breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Tables 4A-4D**. In an embodiment, when two or more gRNAs are used to position four breaks, e.g., four single strand breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Tables 4A-4D**.

In an embodiment, the LCA10 target position in the *CEP290* gene is targeted. In an embodiment, the targeting domain comprises a sequence that is the same as, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from, a targeting domain sequence from **Tables 8A-8E**. In some embodiments, the targeting domain is selected from those in **Tables 8A-8E**. For example, in certain embodiments, the targeting domain is:

GCUAAAUCAUGCAAGUGACCUAAG;

GGUCACUUGCAUGAUUUAG;
GUCACUUGCAUGAUUUAG;
GCCUAGGACUUUCUAAUGCUGGA;
GGACUUUCUAAUGCUGGA;
GGGACCAUGGGAGAAUAGUUUGUU;
GGACCAUGGGAGAAUAGUUUGUU;
GACCAUGGGAGAAUAGUUUGUU;
GGUCCCUGGCUUUUGUUCCUUGGA;
GUCCCUGGCUUUUGUUCCUUGGA;
GAAAACGUUGUUCUGAGUAGCUUU;
GUUGUUCUGAGUAGCUUU;
GGUCCCUGGCUUUUGUUCCU;
GUCCCUGGCUUUUGUUCCU;
GACAUCUUGUGGAUAAUGUAUCA;
GUCCUAGGCAAGAGACAUCUU;
GCCAGCAAAAGCUUUUGAGCUAA;
GCAAAAGCUUUUGAGCUAA;
GAUCUUAUUCUACUCCUGUGA;
GCUUUCAGGAUUCCUACUAAAUU;
GUUCUGUCCUCAGUAAAAGGUA;
GAACAACGUUUUCAUUUA;
GUAGAAUAUCAUAAGUUACAAUCU;
GAAUAUCAUAAGUUACAAUCU;
GUGGCUGUAAGAUAACUACA;
GGCUGUAAGAUAACUACA;
GUUUAACGUUAUCAUUUUCCCA;
GUAAGAGAAAGGGAUGGGCACUUA;
GAGAAAGGGAUGGGCACUUA;
GAAAGGGAUGGGCACUUA;
GUAAAUGAAAACGUUGUU;
GAUAAACAUGACUCAUAAUUUAGU;
GGAACAAAAGCCAGGGACCAUGG;
GAACAAAAGCCAGGGACCAUGG;
GGGAGAAUAGUUUGUUCUGGGUAC;
GGAGAAUAGUUUGUUCUGGGUAC;
GAGAAUAGUUUGUUCUGGGUAC;
GAAUAGUUUGUUCUGGGUAC;
GAAAUAGAGGCUUAUGGAUU;
GUUCUGGGUACAGGGGUAAGAGAA;
GGGUACAGGGGUAAGAGAA;
GGUACAGGGGUAAGAGAA;
GUAAAUUCUCAUCAUUUUUUAUUG;
GGAGAGGAUAGGACAGAGGACAUG;
GAGAGGAUAGGACAGAGGACAUG;

GAGGAUAGGACAGAGGACAUG;
GGAUAGGACAGAGGACAUG;
GAUAGGACAGAGGACAUG;
GAAUAAAUGUAGAAUUUUAAUG;
GUCAAAAGCUACCGGUUACCUG;
GUUUUUAAGGCGGGGAGUCACAU;
GUCUUACAUCCUCCUUACUGCCAC;
GAGUCACAGGGUAGGAUUCAUGUU;
GUCACAGGGUAGGAUUCAUGUU;
GGCACAGAGUUCAAGCUAAUACAU;
GCACAGAGUUCAAGCUAAUACAU;
GAGUUCAAGCUAAUACAU;
GAUGCAGAACUAGUGUAGAC;
GUGUUGAGUAUCUCCUGUUUGGCA;
GUUGAGUAUCUCCUGUUUGGCA;
GAGUAUCUCCUGUUUGGCA;
GAAAAUCAGAUUUCAUGUGUG;
GCCACAAGAAUGAUCAUUCUAAAC;
GGCGGGGAGUCACAUGGGAGUCA;
GCGGGGAGUCACAUGGGAGUCA;
GGGGAGUCACAUGGGAGUCA;
GGGAGUCACAUGGGAGUCA;
GGAGUCACAUGGGAGUCA;
GCUUUUGACAGUUUUUAAGGCG;
GAUCAUUCUUGUGGCAGUAAG;
GAGCAAGAGAUGAACUAG;
GCCUGAACAAGUUUUGAAAC;
GUAGAUUGAGGUAGAAUCAAGAA;
GAUUGAGGUAGAAUCAAGAA;
GGAUGUAAGACUGGAGAUAGAGAC;
GAUGUAAGACUGGAGAUAGAGAC;
GUAAGACUGGAGAUAGAGAC;
GGGAGUCACAUGGGAGUCACAGGG;
GGAGUCACAUGGGAGUCACAGGG;
GAGUCACAUGGGAGUCACAGGG;
GUCACAUGGGAGUCACAGGG;
GUUUACAUAUCUGUCUUCCUUAA; or
GAUUUCAUGUGUGAAGAA.

In an embodiment, when two or more gRNAs are used to position two or more breaks, e.g., two or more single stranded breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Tables 8A-8E**. In an embodiment, when two or more

gRNAs are used to position four breaks, e.g., four single strand breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Tables 8A-8E**.

In an embodiment, the LCA10 target position in the *CEP290* gene is targeted. In an embodiment, the targeting domain comprises a sequence that is the same as, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from, a targeting domain sequence from **Tables 5A-5B**. In some embodiments, the targeting domain is selected from those in **Tables 5A-5B**. For example, in certain embodiments, the targeting domain is:

GAGUUCAAGCUAAUACAUGA;
GUUGUUCUGAGUAGCUU;
GGCAAAAGCAGCAGAAAGCA;
GUUGUUCUGAGUAGCUU; or
GGCAAAAGCAGCAGAAAGCA.

In an embodiment, when two or more gRNAs are used to position two or more breaks, e.g., two or more single stranded breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Tables 5A-5B**. In an embodiment, when two or more gRNAs are used to position four breaks, e.g., four single strand breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Tables 5A-5B**.

In an embodiment, the LCA10 target position in the *CEP290* gene is targeted. In an embodiment, the targeting domain comprises a sequence that is the same as, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from, a targeting domain sequence from **Tables 9A-9B**. In some embodiments, the targeting domain is selected from those in **Tables 9A-9B**. For example, in certain embodiments, the targeting domain is:

GGCAAAAGCAGCAGAAAGCA;
GUGGCUGAAUGACUUCU;
GUUGUUCUGAGUAGCUU;
GACUAGAGGUCACGAAA; or
GAGUUCAAGCUAAUACAUGA.

In an embodiment, when two or more gRNAs are used to position two or more breaks, e.g., two or more single stranded breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Tables 9A-9B**. In an embodiment, when two or more gRNAs are used to position four breaks, e.g., four single strand breaks in the target nucleic acid sequence, each guide RNA is independently selected from one of **Tables 9A-9B**.

In an embodiment, the gRNA, e.g., a gRNA comprising a targeting domain, which is complementary with a target domain from the *CEP290* gene, is a modular gRNA. In other embodiments, the gRNA is a chimeric gRNA.

In an embodiment, when two gRNAs are used to position two breaks, e.g., two single strand breaks, in the target nucleic acid sequence, each guide RNA is independently selected from one or more of **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**.

In an embodiment, the targeting domain which is complementary with a target domain from the *CEP290* gene comprises 16 or more nucleotides in length. In an embodiment, the targeting domain which is complementary with a target domain from the *CEP290* gene is 16 nucleotides or more in length. In an embodiment, the targeting domain is 16 nucleotides in length. In an embodiment, the targeting domain is 17 nucleotides in length. In an embodiment, the targeting domain is 18 nucleotides in length. In an embodiment, the targeting domain is 19 nucleotides in length. In an embodiment, the targeting domain is 20 nucleotides in length. In an embodiment, the targeting domain is 21 nucleotides in length. In an embodiment, the targeting domain is 22 nucleotides in length. In an embodiment, the targeting domain is 23 nucleotides in length. In an embodiment, the targeting domain is 24 nucleotides in length. In an embodiment, the targeting domain is 25 nucleotides in length. In an embodiment, the targeting domain is 26 nucleotides in length.

In an embodiment, the targeting domain comprises 16 nucleotides.

In an embodiment, the targeting domain comprises 17 nucleotides.

In an embodiment, the targeting domain comprises 18 nucleotides.

In an embodiment, the targeting domain comprises 19 nucleotides.

In an embodiment, the targeting domain comprises 20 nucleotides.

In an embodiment, the targeting domain comprises 21 nucleotides.

In an embodiment, the targeting domain comprises 22 nucleotides.

In an embodiment, the targeting domain comprises 23 nucleotides.

In an embodiment, the targeting domain comprises 24 nucleotides.

In an embodiment, the targeting domain comprises 25 nucleotides.

In an embodiment, the targeting domain comprises 26 nucleotides.

A gRNA as described herein may comprise from 5' to 3': a targeting domain (comprising a "core domain", and optionally a "secondary domain"); a first complementarity domain; a linking domain; a second complementarity domain; a proximal domain; and a tail domain. In some embodiments, the proximal domain and tail domain are taken together as a single domain.

In an embodiment, a gRNA comprises a linking domain of no more than 25 nucleotides in length; a proximal and tail domain, that taken together, are at least 20 nucleotides in length; and a targeting domain of equal to or greater than 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 nucleotides in length.

In another embodiment, a gRNA comprises a linking domain of no more than 25 nucleotides in length; a proximal and tail domain, that taken together, are at least 30 nucleotides in length; and a targeting domain of equal to or greater than 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 nucleotides in length.

In another embodiment, a gRNA comprises a linking domain of no more than 25 nucleotides in length; a proximal and tail domain, that taken together, are at least 30 nucleotides in length; and a targeting domain of equal to or greater than 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 nucleotides in length.

In another embodiment, a gRNA comprises a linking domain of no more than 25 nucleotides in length; a proximal and tail domain, that taken together, are at least 40 nucleotides in length; and a targeting domain of equal to or greater than 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 nucleotides in length.

A cleavage event, e.g., a double strand or single strand break, is generated by a Cas9 molecule. The Cas9 molecule may be an enzymatically active Cas9 (eaCas9) molecule, e.g., an eaCas9 molecule that forms a double strand break in a target nucleic acid or an eaCas9 molecule forms a single strand break in a target nucleic acid (e.g., a nickase molecule).

In an embodiment, the eaCas9 molecule catalyzes a double strand break.

In some embodiments, the eaCas9 molecule comprises HNH-like domain cleavage activity but has no, or no significant, N-terminal RuvC-like domain cleavage activity. In this case, the eaCas9 molecule is an HNH-like domain nickase, e.g., the eaCas9 molecule comprises a mutation at D10, e.g., D10A. In other embodiments, the eaCas9 molecule comprises N-terminal RuvC-like domain cleavage activity but has no, or no significant, HNH-like domain

cleavage activity. In an embodiment, the eaCas9 molecule is an N-terminal RuvC-like domain nickase, e.g., the eaCas9 molecule comprises a mutation at H840, e.g., H840A. In an embodiment, the eaCas9 molecule is an N-terminal RuvC-like domain nickase, e.g., the eaCas9 molecule comprises a mutation at H863, e.g., H863A.

In an embodiment, a single strand break is formed in the strand of the target nucleic acid to which the targeting domain of said gRNA is complementary. In another embodiment, a single strand break is formed in the strand of the target nucleic acid other than the strand to which the targeting domain of said gRNA is complementary.

In another aspect, disclosed herein is a nucleic acid, e.g., an isolated or non-naturally occurring nucleic acid, e.g., DNA, that comprises (a) a sequence that encodes a gRNA molecule comprising a targeting domain that is complementary with a target domain in *CEP290* gene as disclosed herein.

In an embodiment, the nucleic acid encodes a gRNA molecule, e.g., the first gRNA molecule, comprising a targeting domain comprising a sequence that is the same as, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from, a targeting domain sequence from any one of **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**. In an embodiment, the nucleic acid encodes a gRNA molecule comprising a targeting domain that is selected from those in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**.

In an embodiment, the nucleic acid encodes a modular gRNA, e.g., one or more nucleic acids encode a modular gRNA. In other embodiments, the nucleic acid encodes a chimeric gRNA. The nucleic acid may encode a gRNA, e.g., the first gRNA molecule, comprising a targeting domain comprising 16 nucleotides or more in length. In one embodiment, the nucleic acid encodes a gRNA, e.g., the first gRNA molecule, comprising a targeting domain that is 16 nucleotides in length. In other embodiments, the nucleic acid encodes a gRNA, e.g., the first gRNA molecule, comprising a targeting domain that is 17 nucleotides in length. In still other embodiments, the nucleic acid encodes a gRNA, e.g., the first gRNA molecule, comprising a targeting domain that is 18 nucleotides in length. In still other embodiments, the nucleic acid encodes a gRNA, e.g., the first gRNA molecule, comprising a targeting domain that is 19 nucleotides in length. In still other embodiments, the nucleic acid encodes a gRNA, e.g., the first

gRNA molecule, comprising a targeting domain that is 20 nucleotides in length. In still other embodiments, the nucleic acid encodes a gRNA, e.g., the first gRNA molecule, comprising a targeting domain that is 21 nucleotides in length. In still other embodiments, the nucleic acid encodes a gRNA, e.g., the first gRNA molecule, comprising a targeting domain that is 22 nucleotides in length. In still other embodiments, the nucleic acid encodes a gRNA, e.g., the first gRNA molecule, comprising a targeting domain that is 23 nucleotides in length. In still other embodiments, the nucleic acid encodes a gRNA, e.g., the first gRNA molecule, comprising a targeting domain that is 24 nucleotides in length. In still other embodiments, the nucleic acid encodes a gRNA, e.g., the first gRNA molecule, comprising a targeting domain that is 25 nucleotides in length. In still other embodiments, the nucleic acid encodes a gRNA, e.g., the first gRNA molecule, comprising a targeting domain that is 26 nucleotides in length.

In an embodiment, a nucleic acid encodes a gRNA comprising from 5' to 3': a targeting domain (comprising a "core domain", and optionally a "secondary domain"); a first complementarity domain; a linking domain; a second complementarity domain; a proximal domain; and a tail domain. In some embodiments, the proximal domain and tail domain are taken together as a single domain.

In an embodiment, a nucleic acid encodes a gRNA e.g., the first gRNA molecule, comprising a linking domain of no more than 25 nucleotides in length; a proximal and tail domain, that taken together, are at least 20 nucleotides in length; and a targeting domain of equal to or greater than 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 nucleotides in length.

In an embodiment, a nucleic acid encodes a gRNA e.g., the first gRNA molecule, comprising a linking domain of no more than 25 nucleotides in length; a proximal and tail domain, that taken together, are at least 30 nucleotides in length; and a targeting domain of equal to or greater than 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 nucleotides in length.

In an embodiment, a nucleic acid encodes a gRNA e.g., the first gRNA molecule, comprising a linking domain of no more than 25 nucleotides in length; a proximal and tail domain, that taken together, are at least 30 nucleotides in length; and a targeting domain of equal to or greater than 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 nucleotides in length.

In an embodiment, a nucleic acid encodes a gRNA comprising e.g., the first gRNA molecule, a linking domain of no more than 25 nucleotides in length; a proximal and tail domain,

that taken together, are at least 40 nucleotides in length; and a targeting domain of equal to or greater than 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 nucleotides in length.

In an embodiment, a nucleic acid comprises (a) a sequence that encodes a gRNA molecule comprising a targeting domain that is complementary with a target domain in the *CEP290* gene as disclosed herein, and further comprising (b) a sequence that encodes a Cas9 molecule.

The Cas9 molecule may be a nickase molecule, a enzymatically activating Cas9 (eaCas9) molecule, e.g., an eaCas9 molecule that forms a double strand break in a target nucleic acid and an eaCas9 molecule forms a single strand break in a target nucleic acid. In an embodiment, a single strand break is formed in the strand of the target nucleic acid to which the targeting domain of said gRNA is complementary. In another embodiment, a single strand break is formed in the strand of the target nucleic acid other than the strand to which the targeting domain of said gRNA is complementary.

In an embodiment, the eaCas9 molecule catalyzes a double strand break.

In some embodiments, the eaCas9 molecule comprises HNH-like domain cleavage activity but has no, or no significant, N-terminal RuvC-like domain cleavage activity. In other embodiments, the said eaCas9 molecule is an HNH-like domain nickase, e.g., the eaCas9 molecule comprises a mutation at D10, e.g., D10A. In other embodiments, the eaCas9 molecule comprises N-terminal RuvC-like domain cleavage activity but has no, or no significant, HNH-like domain cleavage activity. In another embodiment, the eaCas9 molecule is an N-terminal RuvC-like domain nickase, e.g., the eaCas9 molecule comprises a mutation at H840, e.g., H840A. In another embodiment, the eaCas9 molecule is an N-terminal RuvC-like domain nickase, e.g., the eaCas9 molecule comprises a mutation at H863, e.g., H863A.

A nucleic acid disclosed herein may comprise (a) a sequence that encodes a gRNA molecule comprising a targeting domain that is complementary with a target domain in the *CEP290* gene as disclosed herein; and (b) a sequence that encodes a Cas9 molecule.

A nucleic acid disclosed herein may comprise (a) a sequence that encodes a gRNA molecule comprising a targeting domain that is complementary with a target domain in the *CEP290* gene as disclosed herein; (b) a sequence that encodes a Cas9 molecule; and further comprises (c)(i) a sequence that encodes a second gRNA molecule described herein having a targeting domain that is complementary to a second target domain of the *CEP290* gene, and

optionally, (ii) a sequence that encodes a third gRNA molecule described herein having a targeting domain that is complementary to a third target domain of the *CEP290* gene; and optionally, (iii) a sequence that encodes a fourth gRNA molecule described herein having a targeting domain that is complementary to a fourth target domain of the *CEP290* gene.

In an embodiment, a nucleic acid encodes a second gRNA molecule comprising a targeting domain configured to provide a cleavage event, e.g., a double strand break or a single strand break, sufficiently close to a LCA10 target position in the *CEP290* gene to allow alteration, e.g., alteration associated with NHEJ, of the LCA10 target position, either alone or in combination with the break positioned by said first gRNA molecule.

In an embodiment, a nucleic acid encodes a third gRNA molecule comprising a targeting domain configured to provide a cleavage event, e.g., a double strand break or a single strand break, sufficiently close to a LCA10 target position in the *CEP290* gene to allow alteration, e.g., alteration associated with NHEJ, either alone or in combination with the break positioned by the first and/or second gRNA molecule.

In an embodiment, a nucleic acid encodes a fourth gRNA molecule comprising a targeting domain configured to provide a cleavage event, e.g., a double strand break or a single strand break, sufficiently close to a LCA10 target position in the *CEP290* gene to allow alteration, e.g., alteration associated with NHEJ, either alone or in combination with the break positioned by the first gRNA molecule, the second gRNA molecule and/or the third gRNA molecule.

In an embodiment, a nucleic acid encodes a second gRNA molecule comprising a targeting domain configured to provide a cleavage event, e.g., a double strand break or a single strand break, in combination with the break position by said first gRNA molecule, sufficiently close to a LCA10 target position in the *CEP290* gene to allow alteration, e.g., alteration associated with NHEJ, of the a LCA10 target position in the *CEP290* gene, either alone or in combination with the break positioned by said first gRNA molecule.

In an embodiment, a nucleic acid encodes a third gRNA molecule comprising a targeting domain configured to provide a cleavage event, e.g., a double strand break or a single strand break, in combination with the break position by said first and/or second gRNA molecule sufficiently close to a LCA10 target position in the *CEP290* gene to allow alteration, e.g.,

alteration associated with NHEJ, either alone or in combination with the break positioned by the first and/or second gRNA molecule.

In an embodiment, a nucleic acid encodes a fourth gRNA molecule comprising a targeting domain configured to provide a cleavage event, e.g., a double strand break or a single strand break, in combination with the break positioned by the first gRNA molecule, the second gRNA molecule and/or the third gRNA molecule, sufficiently close to a LCA10 target position in the *CEP290* gene to allow alteration, e.g., alteration associated with NHEJ, either alone or in combination with the break positioned by the first gRNA molecule, the second gRNA molecule and/or the third gRNA molecule.

In an embodiment, the nucleic acid encodes a second gRNA molecule. The second gRNA is selected to target the LCA10 target position. Optionally, the nucleic acid may encode a third gRNA, and further optionally, the nucleic acid may encode a fourth gRNA molecule.

In an embodiment, the nucleic acid encodes a second gRNA molecule comprising a targeting domain comprising a sequence that is the same as, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from, a targeting domain sequence from one of **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**. In an embodiment, the nucleic acid encodes a second gRNA molecule comprising a targeting domain selected from those in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**. In an embodiment, when a third or fourth gRNA molecule are present, the third and fourth gRNA molecules may independently comprise a targeting domain comprising a sequence that is the same as, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from, a targeting domain sequence from one of **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**. In a further embodiment, when a third or fourth gRNA molecule are present, the third and fourth gRNA molecules may independently comprise a targeting domain selected from those in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**.

In an embodiment, the nucleic acid encodes a second gRNA which is a modular gRNA, e.g., wherein one or more nucleic acid molecules encode a modular gRNA. In other embodiments, the nucleic acid encoding a second gRNA is a chimeric gRNA. In other

embodiments, when a nucleic acid encodes a third or fourth gRNA, the third and fourth gRNA may be a modular gRNA or a chimeric gRNA. When multiple gRNAs are used, any combination of modular or chimeric gRNAs may be used.

A nucleic acid may encode a second, a third, and/or a fourth gRNA, each independently, comprising a targeting domain comprising 16 nucleotides or more in length. In an embodiment, the nucleic acid encodes a second gRNA comprising a targeting domain that is 16 nucleotides in length. In other embodiments, the nucleic acid encodes a second gRNA comprising a targeting domain that is 17 nucleotides in length. In still other embodiments, the nucleic acid encodes a second gRNA comprising a targeting domain that is 18 nucleotides in length. In still other embodiments, the nucleic acid encodes a second gRNA comprising a targeting domain that is 19 nucleotides in length. In still other embodiments, the nucleic acid encodes a second gRNA comprising a targeting domain that is 20 nucleotides in length. In still other embodiments, the nucleic acid encodes a second gRNA comprising a targeting domain that is 21 nucleotides in length. In still other embodiments, the nucleic acid encodes a second gRNA comprising a targeting domain that is 22 nucleotides in length. In still other embodiments, the nucleic acid encodes a second gRNA comprising a targeting domain that is 23 nucleotides in length. In still other embodiments, the nucleic acid encodes a second gRNA comprising a targeting domain that is 24 nucleotides in length. In still other embodiments, the nucleic acid encodes a second gRNA comprising a targeting domain that is 25 nucleotides in length. In still other embodiments, the nucleic acid encodes a second gRNA comprising a targeting domain that is 26 nucleotides in length.

In an embodiment, the targeting domain comprises 16 nucleotides.

In an embodiment, the targeting domain comprises 17 nucleotides.

In an embodiment, the targeting domain comprises 18 nucleotides.

In an embodiment, the targeting domain comprises 19 nucleotides.

In an embodiment, the targeting domain comprises 20 nucleotides.

In an embodiment, the targeting domain comprises 21 nucleotides.

In an embodiment, the targeting domain comprises 22 nucleotides.

In an embodiment, the targeting domain comprises 23 nucleotides.

In an embodiment, the targeting domain comprises 24 nucleotides.

In an embodiment, the targeting domain comprises 25 nucleotides.

In an embodiment, the targeting domain comprises 26 nucleotides.

In an embodiment, a nucleic acid encodes a second, a third, and/or a fourth gRNA, each independently, comprising from 5' to 3': a targeting domain (comprising a "core domain", and optionally a "secondary domain"); a first complementarity domain; a linking domain; a second complementarity domain; a proximal domain; and a tail domain. In some embodiments, the proximal domain and tail domain are taken together as a single domain.

In an embodiment, a nucleic acid encodes a second, a third, and/or a fourth gRNA, each independently, comprising a linking domain of no more than 25 nucleotides in length; a proximal and tail domain, that taken together, are at least 20 nucleotides in length; and a targeting domain of equal to or greater than 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 nucleotides in length.

In an embodiment, a nucleic acid encodes a second, a third, and/or a fourth gRNA, each independently, comprising a linking domain of no more than 25 nucleotides in length; a proximal and tail domain, that taken together, are at least 30 nucleotides in length; and a targeting domain of equal to or greater than 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 nucleotides in length.

In an embodiment, a nucleic acid encodes a second, a third, and/or a fourth gRNA, each independently, comprising a linking domain of no more than 25 nucleotides in length; a proximal and tail domain, that taken together, are at least 30 nucleotides in length; and a targeting domain of equal to or greater than 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 nucleotides in length.

In an embodiment, a nucleic acid encodes a second, a third, and/or a fourth gRNA, each independently, comprising a linking domain of no more than 25 nucleotides in length; a proximal and tail domain, that taken together, are at least 40 nucleotides in length; and a targeting domain of equal to or greater than 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 nucleotides in length.

In some embodiments, when the *CEP290* gene is altered, e.g., by NHEJ, the nucleic acid encodes (a) a sequence that encodes a gRNA molecule comprising a targeting domain that is complementary with a target domain in the *CEP290* gene as disclosed herein; (b) a sequence that encodes a Cas9 molecule; optionally, (c)(i) a sequence that encodes a second gRNA molecule described herein having a targeting domain that is complementary to a second target domain of the *CEP290* gene, and further optionally, (ii) a sequence that encodes a third gRNA molecule described herein having a targeting domain that is complementary to a third target domain of the *CEP290* gene; and still further optionally, (iii) a sequence that encodes a fourth gRNA molecule

described herein having a targeting domain that is complementary to a fourth target domain of the *CEP290* gene.

As described above, a nucleic acid may comprise (a) a sequence encoding a gRNA molecule comprising a targeting domain that is complementary with a target domain in the *CEP290*, and (b) a sequence encoding a Cas9 molecule. In some embodiments, (a) and (b) are present on the same nucleic acid molecule, e.g., the same vector, e.g., the same viral vector, e.g., the same adeno-associated virus (AAV) vector. In an embodiment, the nucleic acid molecule is an AAV vector, e.g., an AAV vector described herein. Exemplary AAV vectors that may be used in any of the described compositions and methods include an AAV1 vector, a modified AAV1 vector, an AAV2 vector, a modified AAV2 vector, an AAV3 vector, a modified AAV3 vector, an AAV4 vector, a modified AAV4 vector, an AAV5 vector, a modified AAV5 vector, an AAV6 vector, a modified AAV6 vector, an AAV7 vector, a modified AAV7 vector, an AAV8 vector, an AAV9 vector, an AAV.rh10 vector, a modified AAV.rh10 vector, an AAV.rh32/33 vector, a modified AAV.rh32/33 vector, an AAV.rh43 vector, a modified AAV.rh43 vector, an AAV.rh64R1 vector, and a modified AAV.rh64R1 vector.

In other embodiments, (a) is present on a first nucleic acid molecule, e.g. a first vector, e.g., a first viral vector, e.g., a first AAV vector; and (b) is present on a second nucleic acid molecule, e.g., a second vector, e.g., a second vector, e.g., a second AAV vector. The first and second nucleic acid molecules may be AAV vectors, e.g., the AAV vectors described herein.

In other embodiments, the nucleic acid may further comprise (c)(i) a sequence that encodes a second gRNA molecule as described herein. In some embodiments, the nucleic acid comprises (a), (b) and (c)(i). Each of (a) and (c)(i) may be present on the same nucleic acid molecule, e.g., the same vector, e.g., the same viral vector, e.g., the same adeno-associated virus (AAV) vector. In an embodiment, the nucleic acid molecule is an AAV vector, e.g., an AAV vectors described herein.

In other embodiments, (a) and (c)(i) are on different vectors. For example, (a) may be present on a first nucleic acid molecule, e.g. a first vector, e.g., a first viral vector, e.g., a first AAV vector; and (c)(i) may be present on a second nucleic acid molecule, e.g., a second vector, e.g., a second vector, e.g., a second AAV vector. In an embodiment, the first and second nucleic acid molecules are AAV vectors, e.g., the AAV vectors described herein.

In another embodiment, each of (a), (b), and (c)(i) are present on the same nucleic acid molecule, e.g., the same vector, e.g., the same viral vector, e.g., an AAV vector. In an embodiment, the nucleic acid molecule is an AAV vector. In an alternate embodiment, one of (a), (b), and (c)(i) is encoded on a first nucleic acid molecule, e.g., a first vector, e.g., a first viral vector, e.g., a first AAV vector; and a second and third of (a), (b), and (c)(i) is encoded on a second nucleic acid molecule, e.g., a second vector, e.g., a second vector, e.g., a second AAV vector. The first and second nucleic acid molecule may be AAV vectors, e.g., the AAV vectors described herein.

In an embodiment, (a) is present on a first nucleic acid molecule, e.g., a first vector, e.g., a first viral vector, a first AAV vector; and (b) and (c)(i) are present on a second nucleic acid molecule, e.g., a second vector, e.g., a second vector, e.g., a second AAV vector. The first and second nucleic acid molecule may be AAV vectors, e.g., the AAV vectors described herein.

In other embodiments, (b) is present on a first nucleic acid molecule, e.g., a first vector, e.g., a first viral vector, e.g., a first AAV vector; and (a) and (c)(i) are present on a second nucleic acid molecule, e.g., a second vector, e.g., a second vector, e.g., a second AAV vector. The first and second nucleic acid molecule may be AAV vectors, e.g., the AAV vectors described herein.

In other embodiments, (c)(i) is present on a first nucleic acid molecule, e.g., a first vector, e.g., a first viral vector, e.g., a first AAV vector; and (b) and (a) are present on a second nucleic acid molecule, e.g., a second vector, e.g., a second vector, e.g., a second AAV vector. The first and second nucleic acid molecule may be AAV vectors, e.g., the AAV vectors described herein.

In another embodiment, each of (a), (b) and (c)(i) are present on different nucleic acid molecules, e.g., different vectors, e.g., different viral vectors, e.g., different AAV vector. For example, (a) may be on a first nucleic acid molecule, (b) on a second nucleic acid molecule, and (c)(i) on a third nucleic acid molecule. The first, second and third nucleic acid molecule may be AAV vectors, e.g., the AAV vectors described herein.

In another embodiment, when a third and/or fourth gRNA molecule are present, each of (a), (b), (c)(i), (c) (ii) and (c)(iii) may be present on the same nucleic acid molecule, e.g., the same vector, e.g., the same viral vector, e.g., an AAV vector. In an embodiment, the nucleic acid molecule is an AAV vector, e.g., an AAV vector. In an alternate embodiment, each of (a), (b), (c)(i), (c)(ii) and (c)(iii) may be present on the different nucleic acid molecules, e.g., different

vectors, e.g., the different viral vectors, e.g., different AAV vectors. In further embodiments, each of (a), (b), (c)(i), (c) (ii) and (c)(iii) may be present on more than one nucleic acid molecule, but fewer than five nucleic acid molecules, e.g., AAV vectors, e.g., the AAV vectors described herein.

The nucleic acids described herein may comprise a promoter operably linked to the sequence that encodes the gRNA molecule of (a), e.g., a promoter described herein, e.g., a promoter described in **Table 19**. The nucleic acid may further comprise a second promoter operably linked to the sequence that encodes the second, third and/or fourth gRNA molecule of (c), e.g., a promoter described herein. The promoter and second promoter differ from one another. In some embodiments, the promoter and second promoter are the same.

The nucleic acids described herein may further comprise a promoter operably linked to the sequence that encodes the Cas9 molecule of (b), e.g., a promoter described herein, e.g., a promoter described in **Table 19**.

In another aspect, disclosed herein is a composition comprising (a) a gRNA molecule comprising a targeting domain that is complementary with a target domain in the *CEP290* gene, as described herein. The composition of (a) may further comprise (b) a Cas9 molecule, e.g., a Cas9 molecule as described herein. A composition of (a) and (b) may further comprise (c) a second, third and/or fourth gRNA molecule, e.g., a second, third and/or fourth gRNA molecule described herein.

In another aspect, methods and compositions discussed herein, provide for treating or delaying the onset or progression of LCA10 by altering the LCA10 target position in the *CEP290* gene.

In another aspect, disclosed herein is a method of altering a cell, e.g., altering the structure, e.g., altering the sequence, of a target nucleic acid of a cell, comprising contacting said cell with: (a) a gRNA that targets the *CEP290* gene, e.g., a gRNA as described herein; (b) a Cas9 molecule, e.g., a Cas9 molecule as described herein; and optionally, (c) a second, third and/or fourth gRNA that targets *CEP290* gene, e.g., a gRNA as described herein.

In some embodiments, the method comprises contacting said cell with (a) and (b).

In some embodiments, the method comprises contacting said cell with (a), (b), and (c).

The gRNA of (a) may be selected from any of **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables**

9A-9B, or **Table 10**, or a gRNA that differs by no more than 1, 2, 3, 4, or 5 nucleotides from, a targeting domain sequence from any of **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**. The gRNA of (c) may be selected from any of **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**, or a gRNA that differs by no more than 1, 2, 3, 4, or 5 nucleotides from, a targeting domain sequence from any of **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**.

In some embodiments, the method comprises contacting a cell from a subject suffering from or likely to develop LCA10. The cell may be from a subject having a mutation at a LCA10 target position.

In some embodiments, the cell being contacted in the disclosed method is a photoreceptor cell. The contacting may be performed *ex vivo* and the contacted cell may be returned to the subject's body after the contacting step. In other embodiments, the contacting step may be performed *in vivo*.

In some embodiments, the method of altering a cell as described herein comprises acquiring knowledge of the presence of a LCA10 target position in said cell, prior to the contacting step. Acquiring knowledge of the presence of a LCA10 target position in the cell may be by sequencing the *CEP290* gene, or a portion of the *CEP290* gene.

In some embodiments, the contacting step of the method comprises contacting the cell with a nucleic acid, e.g., a vector, e.g., an AAV vector, e.g., an AAV vector described herein, that expresses at least one of (a), (b), and (c). In some embodiments, the contacting step of the method comprises contacting the cell with a nucleic acid, e.g., a vector, e.g., an AAV vector, that expresses each of (a), (b), and (c). In another embodiment, the contacting step of the method comprises delivering to the cell a Cas9 molecule of (b) and a nucleic acid which encodes a gRNA (a) and optionally, a second gRNA (c)(i) (and further optionally, a third gRNA (c)(iv) and/or fourth gRNA (c)(iii)).

In an embodiment, contacting comprises contacting the cell with a nucleic acid, e.g., a vector, e.g., an AAV vector, e.g., an AAV1 vector, a modified AAV1 vector, an AAV2 vector, a modified AAV2 vector, an AAV3 vector, a modified AAV3 vector, an AAV4 vector, a modified

AAV4 vector, an AAV5 vector, a modified AAV5 vector, an AAV6 vector, a modified AAV6 vector, an AAV7 vector, a modified AAV7 vector, an AAV8 vector, an AAV9 vector, an AAV.rh10 vector, a modified AAV.rh10 vector, an AAV.rh32/33 vector, a modified AAV.rh32/33 vector, an AAV.rh43vector, a modified AAV.rh43vector, an AAV.rh64R1vector, and a modified AAV.rh64R1vector, e.g., an AAV vector described herein.

In an embodiment, contacting comprises delivering to said cell said Cas9 molecule of (b), as a protein or an mRNA, and a nucleic acid which encodes and (a) and optionally (c).

In an embodiment, contacting comprises delivering to said cell said Cas9 molecule of (b), as a protein or an mRNA, said gRNA of (a), as an RNA, and optionally said second gRNA of (c), as an RNA.

In an embodiment, contacting comprises delivering to said cell said gRNA of (a) as an RNA, optionally said second gRNA of (c) as an RNA, and a nucleic acid that encodes the Cas9 molecule of (b).

In another aspect, disclosed herein is a method of treating, or preventing a subject suffering from developing, LCA10, e.g., by altering the structure, e.g., sequence, of a target nucleic acid of the subject, comprising contacting the subject (or a cell from the subject) with:

(a) a gRNA that targets the *CEP290* gene, e.g., a gRNA disclosed herein;

(b) a Cas9 molecule, e.g., a Cas9 molecule disclosed herein; and

optionally, (c)(i) a second gRNA that targets the *CEP290* gene, e.g., a second gRNA disclosed herein, and

further optionally, (c)(ii) a third gRNA, and still further optionally, (c)(iii) a fourth gRNA that target the *CEP290*, e.g., a third and fourth gRNA disclosed herein.

In some embodiments, contacting comprises contacting with (a) and (b).

In some embodiments, contacting comprises contacting with (a), (b), and (c)(i).

In some embodiments, contacting comprises contacting with (a), (b), (c)(i) and (c)(ii).

In some embodiments, contacting comprises contacting with (a), (b), (c)(i), (c)(ii) and (c)(iii).

The gRNA of (a) or (c) (e.g., (c)(i), (c)(ii), or (c)(iii)) may be independently selected from any of **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**, or a gRNA that differs by no more than 1, 2, 3, 4, or 5 nucleotides from, a targeting domain sequence from any of **Tables**

1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B, or **Table 10**.

In an embodiment, said subject is suffering from, or likely to develop LCA10. In an embodiment, said subject has a mutation at a LCA10 target position.

In an embodiment, the method comprises acquiring knowledge of the presence of a mutation at a LCA10 target position in said subject.

In an embodiment, the method comprises acquiring knowledge of the presence of a mutation a LCA10 target position in said subject by sequencing the *CEP290* gene or a portion of the *CEP290* gene.

In an embodiment, the method comprises altering the LCA10 target position in the *CEP290* gene.

In an embodiment, a cell of said subject is contacted *ex vivo* with (a), (b) and optionally (c). In an embodiment, said cell is returned to the subject's body.

In an embodiment, the method comprises introducing a cell into said subject's body, wherein said cell subject was contacted *ex vivo* with (a), (b) and optionally (c).

In an embodiment, the method comprises said contacting is performed *in vivo*. In an embodiment, the method comprises sub-retinal delivery. In an embodiment, contacting comprises sub-retinal injection. In an embodiment, contacting comprises intra-vitreal injection.

In an embodiment, contacting comprises contacting the subject with a nucleic acid, e.g., a vector, e.g., an AAV vector described herein, e.g., a nucleic acid that encodes at least one of (a), (b), and optionally (c).

In an embodiment, contacting comprises delivering to said subject said Cas9 molecule of (b), as a protein or mRNA, and a nucleic acid which encodes and (a) and optionally (c).

In an embodiment, contacting comprises delivering to said subject said Cas9 molecule of (b), as a protein or mRNA, said gRNA of (a), as an RNA, and optionally said second gRNA of (c), as an RNA.

In an embodiment, contacting comprises delivering to said subject said gRNA of (a), as an RNA, optionally said second gRNA of (c), as an RNA, and a nucleic acid that encodes the Cas9 molecule of (b).

In another aspect, disclosed herein is a reaction mixture comprising a gRNA, a nucleic acid, or a composition described herein, and a cell, e.g., a cell from a subject having, or likely to develop LCA10, or a subject having a mutation at a LCA10 target position.

In another aspect, disclosed herein is a kit comprising, (a) a gRNA molecule described herein, or a nucleic acid that encodes said gRNA, and one or more of the following:

(b) a Cas9 molecule, e.g., a Cas9 molecule described herein, or a nucleic acid or mRNA that encodes the Cas9;

(c)(i) a second gRNA molecule, e.g., a second gRNA molecule described herein or a nucleic acid that encodes (c)(i);

(c)(ii) a third gRNA molecule, e.g., a second gRNA molecule described herein or a nucleic acid that encodes (c)(ii); or

(c)(iii) a fourth gRNA molecule, e.g., a second gRNA molecule described herein or a nucleic acid that encodes (c)(iii).

In an embodiment, the kit comprises nucleic acid, e.g., an AAV vector, e.g., an AAV vector described herein, that encodes one or more of (a), (b), (c)(i), (c)(ii), and (c)(iii). In an embodiment, the kit further comprises a governing gRNA molecule, or a nucleic acid that encodes a governing gRNA molecule.

In yet another aspect, disclosed herein is a gRNA molecule, e.g., a gRNA molecule described herein, for use in treating LCA10 in a subject, e.g., in accordance with a method of treating LCA10 as described herein.

In an embodiment, the gRNA molecule in used in combination with a Cas9 molecule, e.g., a Cas9 molecule described herein. Additionaly or alternatively, in an embodiment, the gRNA molecule is used in combination with a second, third and/or fouth gRNA molecule, e.g., a second, third and/or fouth gRNA molecule described herein.

In still another aspect, disclosed herein is use of a gRNA molecule, e.g., a gRNA molecule described herein, in the manufacture of a medicament for treating LCA10 in a subject, e.g., in accordance with a method of treating LCA10 as described herein.

In an embodiment, the medicament comprises a Cas9 molecule, e.g., a Cas9 molecule described herein. Additionaly or alternatively, in an embodiment, the medicament comprises a second, third and/or fouth gRNA molecule, e.g., a second, third and/or fouth gRNA molecule described herein.

In one aspect, disclosed herein is a recombinant adenovirus-associated virus (AAV) genome comprising the following components:

left ITR–spacer 1–PIII promoter–gRNA–spacer 2–PII promoter–N-ter NLS–Cas9–C-ter NLS–poly(A) signal–spacer 3–right ITR,

wherein the left ITR component comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, any of the left ITR nucleotide sequences disclosed in **Table 24**, or any of the nucleotide sequences of SEQ ID NOS: 407-415;

wherein the spacer 1 component comprises, or consists of, a nucleotide sequence having 0 to 150 nucleotides in length, e.g., SEQ ID NO: 416;

wherein the PIII promoter component comprises, or consists of, an RNA polymerase III promoter sequence;

wherein the gRNA component comprises a targeting domain and a scaffold domain,

wherein the targeting domain is 16-26 nucleotides in length, and comprises, or consists of, a targeting domain sequence disclosed herein, e.g., in any of **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**; and

wherein the scaffold domain (also referred to as a tracr domain in **Figs. 19A-24F**) comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, or 5 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, a nucleotide sequence of SEQ ID NO: 418;

wherein the spacer 2 component comprises, or consists of, a nucleotide sequence having 0 to 150 nucleotides in length e.g., SEQ ID NO: 419;

wherein the PII promoter component comprises, or consists of, a polymerase II promoter sequence, e.g., a constitutive or tissue specific promoter, e.g., a promoter disclosed in **Table 19**;

wherein the N-ter NLS component comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, or 5 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 420 or a nucleotide sequence that encodes the amino acid sequence of SEQ ID NO: 434;

wherein the Cas9 component comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least

90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 421 or a nucleotide sequence that encodes the amino acid sequence of SEQ ID NO: 435;

wherein the C-ter NLS component comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, or 5 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 422 or a nucleotide sequence that encodes the amino acid sequence of SEQ ID NO: 434;

wherein the poly(A) signal component comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, any of the nucleotide sequences disclosed in **Table 26**, or any of the nucleotide sequences of SEQ ID NOS: 424, 455 or 456;

wherein the spacer 3 component comprises, or consists of, a nucleotide sequence having 0 to 150 nucleotides in length, e.g., SEQ ID NO: 425; and

wherein the right ITR component comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, any of the right ITR nucleotide sequences disclosed in **Table 24**, or any of the nucleotide sequences of SEQ ID NOS: 436-444.

In an embodiment, the left ITR component comprises, or consists of, a nucleotide sequence that is the same as any of the nucleotide sequences of SEQ ID NOS: 407-415.

In an embodiment, the spacer 1 component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 416.

In an embodiment, the PIII promoter component is a U6 promoter component.

In an embodiment, the U6 promoter component comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 417;

In an embodiment, the U6 promoter component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 417.

In an embodiment, the PIII promoter component is an H1 promoter component that comprises an H1 promoter sequence.

In an embodiment, the PIII promoter component is a tRNA promoter component that comprises a tRNA promoter sequence.

In an embodiment, the targeting domain comprises, or consists of, a nucleotide sequence that is the same as a nucleotide sequence selected from **Table 10**.

In an embodiment, the gRNA scaffold domain comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 418.

In an embodiment, the spacer 2 component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 419;

In an embodiment, the PII promoter component is a CMV promoter component, and comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 401. In an embodiment, the PII promoter comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 401.

In an embodiment, the PII promoter component is an EFS promoter component, and comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 402. In an embodiment, the PII promoter comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 402.

In an embodiment, the PII promoter component is a GRK1 promoter (e.g., a human GRK1 promoter) component, and comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 403. In an embodiment, the PII promoter comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 403.

In an embodiment, the PII promoter component is a CRX promoter (e.g., a human CRX promoter) component, and comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO:

404. In an embodiment, the PII promoter comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 404.

In an embodiment, the PII promoter component is an NRL promoter (e.g., a human NRL promoter) component, and comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 405. In an embodiment, the PII promoter comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 405.

In an embodiment, the PII promoter component is an RCVRN promoter (e.g., a human RCVRN promoter) component, and comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 406. In an embodiment, the PII promoter comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 406.

In an embodiment, the N-ter NLS component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 420 or a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 434.

In an embodiment, the Cas9 component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 421 or a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 435.

In an embodiment, the C-ter NLS component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 422 or a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 434.

In an embodiment, the poly(A) signal component comprises, or consists of, a nucleotide sequence that is the same as any of the nucleotide sequences disclosed in **Table 26**, or any of the nucleotide sequences of SEQ ID NOS: 424, 455 or 456. In an embodiment, the poly(A) signal component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 424.

In an embodiment, the spacer 3 component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 425.

In an embodiment, the right ITR component comprises, or consists of, a nucleotide sequence that is the same as any of the nucleotide sequences of SEQ ID NOS: 436-444.

In an embodiment, the recombinant AAV genome further comprises a second gRNA component comprising a targeting domain and a scaffold domain,

wherein the targeting domain consists of a targeting domain sequence disclosed herein, in any of **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**; and

wherein the scaffold domain (also referred to as a tracr domain in **Figs. 19A-24F**) comprises, or consists of, a nucleotide sequence that is the same as, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 418.

In an embodiment, the targeting domain of the second gRNA component comprises, or consists of, a nucleotide sequence that is the same as a nucleotide sequence selected from **Table 10**. In an embodiment, the second gRNA component is between the first gRNA component and the spacer 2 component.

In an embodiment, the second gRNA component has the same nucleotide sequence as the first gRNA component. In another embodiment, the second gRNA component has a nucleotide sequence that is different from the second gRNA component.

In an embodiment, the recombinant AAV genome further comprises a second PIII promoter component that comprises, or consists of, an RNA polymerase III promoter sequence;

In an embodiment, the recombinant AAV genome further comprises a second PIII promoter component (e.g., a second U6 promoter component) between the first gRNA component and the second gRNA component.

In an embodiment, the second PIII promoter component (e.g., the second U6 promoter component) has the same nucleotide sequence as the first PIII promoter component (e.g., the first U6 promoter component). In another embodiment, the second PIII promoter component (e.g., the second U6 promoter component) has a nucleotide sequence that is different from the first PIII promoter component (e.g. the first U6 promoter component).

In an embodiment, the PIII promoter component is an H1 promoter component that comprises an H1 promoter sequence.

In an embodiment, the PIII promoter component is a tRNA promoter component that comprises a tRNA promoter sequence.

In an embodiment, the recombinant AAV genome further comprises a spacer 4 component between the first gRNA component and the second PIII promoter component (e.g., the second U6 promoter component). In an embodiment, the spacer 4 component comprises, or consists of, a nucleotide sequence having 0 to 150 nucleotides in length, e.g., SEQ ID NO: 427. In an embodiment, the spacer 4 component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 427.

In an embodiment, the recombinant AAV genome comprises the following components: left ITR–spacer 1–PIII promoter–gRNA–spacer 4–second PIII promoter–second gRNA–spacer 2–PII promoter–N-ter NLS–Cas9–C-ter NLS–poly(A) signal–spacer 3–right ITR.

In an embodiment, the recombinant AAV genome further comprises an affinity tag component (e.g., 3xFLAG component), wherein the affinity tag component (e.g., 3xFLAG component) comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, or 5 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotides sequence of SEQ ID NO: 423, or a nucleotide sequence encoding any of the amino acid sequences disclosed in **Table 25** or any of the amino acid sequences of SEQ ID NO: 435 or 451-454.

In an embodiment, the affinity tag component (e.g., 3xFLAG component) is between the C-ter NLS component and the poly(A) signal component. In an embodiment, the an affinity tag component (e.g., 3xFLAG component) comprises, or consists of, a nucleotide sequence that is the same as, the nucleotides sequence of SEQ ID NO: 423, or a nucleotide sequence encoding any of the amino acid sequences of SEQ ID NOS: 435 or 451-454.

In an embodiment, the recombinant AAV genome comprises the nucleotide sequences of SEQ ID NOS: 408, 417, 418, 401, 420, 421, 422, 424, and 437.

In an embodiment, the recombinant AAV genome comprises the nucleotide sequences of SEQ ID NOS: 408, 417, 418, 402, 420, 421, 422, 424, and 437.

In an embodiment, the recombinant AAV genome comprises the nucleotide sequences of SEQ ID NOS: 408, 417, 418, 403, 420, 421, 422, 424, and 437.

In an embodiment, the recombinant AAV genome comprises the nucleotide sequences of SEQ ID NOS: 408, 417, 418, 404, 420, 421, 422, 424, and 437.

In an embodiment, the recombinant AAV genome comprises the nucleotide sequences of SEQ ID NOS: 408, 417, 418, 405, 420, 421, 422, 424, and 437.

In an embodiment, the recombinant AAV genome comprises the nucleotide sequences of SEQ ID NOS: 408, 417, 418, 406, 420, 421, 422, 424, and 437.

In an embodiment, the recombinant AAV genome further comprises SEQ ID NOS: 416, 419, and 425, and, optionally, SEQ ID NO 427.

In an embodiment, the recombinant AAV genome further comprises the nucleotide sequence of SEQ ID NO: 423.

In an embodiment, the recombinant AAV genome comprises or consists of one or more (e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 or all) of the component sequences shown in **Figs. 19A-19G, 20A-20F, 21A-21F, 22A-22F, 23A-23F,** or **24A-24F, Tables 19** or **24-26**, or any of the nucleotide sequences of SEQ ID NOS: 428-433 or 436-444.

In another aspect, disclosed herein is a recombinant adenovirus-associated virus (AAV) genome comprising the following components:

left ITR–spacer 1–first PIII promoter–first gRNA–spacer 4–second PIII promoter–second gRNA–spacer 2–PII promoter–N-ter NLS–Cas9–C-ter NLS–poly(A) signal–spacer 3–right ITR.

wherein the left ITR component comprises, or consists of, a nucleotide sequence that is the same as, or differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, any of the left ITR nucleotide sequences disclosed in **Table 24**, or any of the nucleotide sequences of SEQ ID NOS: 407-415;

wherein the spacer 1 component comprises, or consists of, a nucleotide sequence having 0 to 150 nucleotides in length, e.g., SEQ ID NO: 416;

wherein the first PIII promoter component (e.g., a first U6 promoter component) comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 417;

wherein the first gRNA component comprises a targeting domain and a scaffold domain,

wherein the targeting domain is 16-26 nucleotides in length, and comprises, or consists of, a targeting domain sequence disclosed herein, e.g., in any of **Tables 1A-1D,**

Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B, or **Table 10**; and

wherein the scaffold domain (also referred to herein as a tracr domain in **Figs. 19A-24F**) comprises, or consists of, a nucleotide sequence that is the same as, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 418;

wherein the spacer 4 component comprises, or consists of, a nucleotide sequence having 0 to 150 nucleotides in length, e.g., SEQ ID NO: 427.

wherein the second gRNA component comprises a targeting domain and a scaffold domain,

wherein the targeting domain of the second gRNA component is 16-26 nucleotides in length and comprises, or consists of, a targeting domain sequence disclosed herein, e.g., in any of **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**; and

wherein the scaffold domain (also referred to as a tracr domain in **Figs. 19A-24F**) of the second gRNA component comprises, or consists of, a nucleotide sequence that is the same as, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 418.

wherein the spacer 2 component comprises, or consists of, a nucleotide sequence having 0 to 150 nucleotides in length e.g., SEQ ID NO: 419;

wherein the PII promoter component comprises, or consists of, a polymerase II promoter sequence, e.g., a constitutive or tissue specific promoter, e.g., a promoter disclosed in **Table 19**;

wherein the N-ter NLS component comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, or 5 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 420 or a nucleotide sequence that encodes the amino acid sequence of SEQ ID NO: 434;

wherein the Cas9 component comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least

90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 421 or a nucleotide sequence that encodes the amino acid sequence of SEQ ID NO: 435;

wherein the C-ter NLS component comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, or 5 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 422 or a nucleotide sequence that encodes the amino acid sequence of SEQ ID NO: 434;

wherein the poly(A) signal component comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, any of the nucleotide sequences disclosed in **Table 26**, or any of the nucleotide sequence of SEQ ID NO: 424, 455 or 456;

wherein the spacer 3 component comprises, or consists of, a nucleotide sequence having 0 to 150 nucleotides in length, e.g., SEQ ID NO: 425; and

wherein the right ITR component comprises, or consists of, a nucleotide sequence that is the same as, or differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, any of the right ITR nucleotide sequences disclosed in **Table 24**, or SEQ ID NOS: 436-444.

In an embodiment, the left ITR component comprises, or consists of, a nucleotide sequence that is the same as any of the nucleotide sequences of SEQ ID NOS: 407-415.

In an embodiment, the spacer 1 component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 416.

In an embodiment, the first PIII promoter component (e.g., the first U6 promoter component) comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 417.

In an embodiment, the first PIII promoter is an H1 promoter component that comprises an H1 promoter sequence. In another embodiment, the first PIII promoter is a tRNA promoter component that comprises a tRNA promoter sequence.

In an embodiment, the targeting domain of the first gRNA component comprises, or consists of, a nucleotide sequence that is the same as a nucleotide sequence selected from **Table 10**.

In an embodiment, the gRNA scaffold domain of the first gRNA component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 418.

In an embodiment, the spacer 4 component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 427.

In an embodiment, the second PIII promoter component (e.g., the first U6 promoter component) has the same nucleotide sequence as the first PIII promoter component (e.g., the first U6 promoter component). In another embodiment, the second PIII promoter component (e.g., the second U6 promoter component) has a nucleotide sequence that is different from the first PIII promoter component (e.g., the first U6 promoter component).

In an embodiment, the second PIII promoter is an H1 promoter component that comprises an H1 promoter sequence. In another embodiment, the second PIII promoter is a tRNA promoter component that comprises a tRNA promoter sequence.

In an embodiment, the targeting domain of the second gRNA component comprises, or consists of, a nucleotide sequence that is the same as a nucleotide sequence selected from **Table 10**.

In an embodiment, the second gRNA component has the same nucleotide sequence as the first gRNA component. In another embodiment, the second gRNA component has a nucleotide sequence that is different from the second gRNA component.

In an embodiment, the spacer 2 component comprises, or consists of, a nucleotide sequence having 0 to 150 nucleotides in length e.g., SEQ ID NO: 419;

In an embodiment, the PII promoter component is a CMV promoter component, and comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 401. In an embodiment, the PII promoter comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 401.

In an embodiment, the PII promoter component is an EFS promoter component, and comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 402. In an embodiment, the PII

promoter comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 402.

In an embodiment, the PII promoter component is a GRK1 promoter (e.g., a human GRK1 promoter) component, and comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 403. In an embodiment, the PII promoter comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 403.

In an embodiment, the PII promoter component is a CRX promoter (e.g., a human CRX promoter) component, and comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 404. In an embodiment, the PII promoter comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 404.

In an embodiment, the PII promoter component is an NRL promoter (e.g., a human NRL promoter) component, and comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 405. In an embodiment, the PII promoter comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 405.

In an embodiment, the PII promoter component is an RCVRN promoter (e.g., a human RCVRN promoter) component, and comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 406. In an embodiment, the PII promoter comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 406.

In an embodiment, the N-ter NLS component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 420 or a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 434.

In an embodiment, the Cas9 component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 421 or a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 435.

In an embodiment, the C-ter NLS component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 422 or a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 434.

In an embodiment, the poly(A) signal component comprises, or consists of, a nucleotide sequence that is the same as any of the nucleotide sequences disclosed in **Table 26**, or any of the nucleotide sequences of SEQ ID NOS: 424, 455 or 456. In an embodiment, the poly(A) signal component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 424.

In an embodiment, the spacer 3 component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 425.

In an embodiment, the right ITR component comprises, or consists of, a nucleotide sequence that is the same as any of the nucleotide sequences disclosed in **Table 24**, or any of the nucleotide sequences of SEQ ID NOS: 436-444.

In an embodiment, the recombinant AAV genome further comprises an affinity tag component (e.g., a 3xFLAG component). In an embodiment, the affinity tag component (e.g., the 3xFLAG component) comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, or 5 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 423, or a nucleotide sequence encoding any of the amino acid sequences disclosed in **Table 25** or any of the amino acid sequences of SEQ ID NO: 435 or 451-454.

In an embodiment, the affinity tag component (e.g., the 3xFLAG component) is between the C-ter NLS component and the poly(A) signal component. In an embodiment, the affinity tag component (e.g., the 3xFLAG component) comprises, or consists of, a nucleotide sequence that is the same as, the nucleotide sequence of SEQ ID NO: 423 or a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 435.

In an embodiment, the recombinant AAV genome comprises the nucleotide sequences of SEQ ID NOS: 408, 417, 418, 401, 420, 421, 422, 424, and 437.

In an embodiment, the recombinant AAV genome comprises the nucleotide sequences of SEQ ID NOS: 408, 417, 418, 402, 420, 421, 422, 424, and 437.

In an embodiment, the recombinant AAV genome comprises the nucleotide sequences of SEQ ID NOS: 408, 417, 418, 403, 420, 421, 422, 424, and 437.

In an embodiment, the recombinant AAV genome comprises the nucleotide sequences of SEQ ID NOS: 408, 417, 418, 404, 420, 421, 422, 424, and 437.

In an embodiment, the recombinant AAV genome comprises the nucleotide sequences of SEQ ID NOS: 408, 417, 418, 405, 420, 421, 422, 424, and 437.

In an embodiment, the recombinant AAV genome comprises the nucleotide sequences of SEQ ID NOS: 408, 417, 418, 406, 420, 421, 422, 424, and 437.

In an embodiment, the recombinant AAV genome further comprises the nucleotide sequences of SEQ ID NO: 416, 419, 425, and 427.

In an embodiment, the recombinant AAV genome further comprises the nucleotide sequence of SEQ ID NO: 423.

In an embodiment, the recombinant AAV genome comprises any of the nucleotide sequences of SEQ ID NOS: 428-433.

In an embodiment, the recombinant AAV genome comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 100, 200, 300, 400, or 500 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with any of the nucleotide sequences shown in **Figs. 19A-19G, 20A-20F, 21A-21F, 22A-22F, 23A-23F,** or **24A-24F**, or any of the nucleotide sequences of SEQ ID NOS: 428-433 or 436-444.

In an embodiment, the recombinant AAV genome comprises, or consists of, a nucleotide sequence that is the same as any of the nucleotide sequences shown in **Figs. 19A-19G, 20A-20F, 21A-21F, 22A-22F, 23A-23F,** or **24A-24F**, or any of the nucleotide sequences of SEQ ID NOS: 428-433 or 436-444.

In an embodiment, the recombinant AAV genome comprises or consists of one or more (e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 or all) of the component sequences shown in **Figs. 19A-19G, 20A-20F, 21A-21F, 22A-22F, 23A-23F,** or **24A-24F,** or **Tables 19** or **24-26**, or any of the nucleotide sequences of SEQ ID NOS: 428-433 or 436-444.

Unless otherwise indicated, when components of a recombinant AAV genome are described herein, the order can be as provided, but other orders are included as well. In other

words, in an embodiment, the order is as set out in the text, but in other embodiments, the order can be different.

It is understood that the recombinant AAV genomes disclosed herein can be single stranded or double stranded. Disclosed herein are also the reverse, complementary form of any of the recombinant AAV genomes disclosed herein, and the double stranded form thereof.

In another aspect, disclosed herein is a nucleic acid molecule (e.g., an expression vector) that comprises a recombinant AAV genome disclosed herein. In an embodiment, the nucleic acid molecule further comprises a nucleotide sequence that encodes an antibiotic resistant gene (e.g., an Amp resistant gene). In an embodiment, the nucleic acid molecule further comprises replication origin sequence (e.g., a ColE1 origin, an M13 origin, or both).

In another aspect, disclosed herein is a recombinant AAV viral particle comprising a recombinant AAV genome disclosed herein.

In an embodiment, the recombinant AAV viral particle has any of the serotype disclosed herein, e.g., in **Table 24**, or a combination thereof. In another embodiment, the recombinant AAV viral particle has a tissue specificity of retinal pigment epithelium cells, photoreceptors, horizontal cells, bipolar cells, amacrine cells, ganglion cells, or a combination thereof.

In another aspect, disclosed herein is a method of producing a recombinant AAV viral particle disclosed herein comprising providing a recombinant AAV genome disclosed herein and one or more capsid proteins under conditions that allow for assembly of an AAV viral particle.

In another aspect, disclosed herein is a method of altering a cell comprising contacting the cell with a recombinant AAV viral particle disclosed herein.

In another aspect, disclosed herein is a method of treating a subject having or likely to develop LCA10 comprising contacting the subject (or a cell from the subject) with a recombinant viral particle disclosed herein.

In another aspect, disclosed herein is a recombinant AAV viral particle comprising a recombinant AAV genome disclosed herein for use in treating LCA10 in a subject.

In another aspect, disclosed herein is use of a recombinant AAV viral particle comprising a recombinant AAV genome disclosed herein in the manufacture of a medicament for treating LCA10 in a subject.

The gRNA molecules and methods, as disclosed herein, can be used in combination with a governing gRNA molecule, comprising a targeting domain which is complementary to a target

domain on a nucleic acid that encodes a component of the CRISPR/Cas system introduced into a cell or subject. In an embodiment, the governing gRNA molecule targets a nucleic acid that encodes a Cas9 molecule or a nucleic acid that encodes a target gene gRNA molecule. In an embodiment, the governing gRNA comprises a targeting domain that is complementary to a target domain in a sequence that encodes a Cas9 component, e.g., a Cas9 molecule or target gene gRNA molecule. In an embodiment, the target domain is designed with, or has, minimal homology to other nucleic acid sequences in the cell, e.g., to minimize off-target cleavage. For example, the targeting domain on the governing gRNA can be selected to reduce or minimize off-target effects. In an embodiment, a target domain for a governing gRNA can be disposed in the control or coding region of a Cas9 molecule or disposed between a control region and a transcribed region. In an embodiment, a target domain for a governing gRNA can be disposed in the control or coding region of a target gene gRNA molecule or disposed between a control region and a transcribed region for a target gene gRNA. While not wishing to be bound by theory, in an embodiment, it is believed that altering, e.g., inactivating, a nucleic acid that encodes a Cas9 molecule or a nucleic acid that encodes a target gene gRNA molecule can be effected by cleavage of the targeted nucleic acid sequence or by binding of a Cas9 molecule/governing gRNA molecule complex to the targeted nucleic acid sequence.

The compositions, reaction mixtures and kits, as disclosed herein, can also include a governing gRNA molecule, e.g., a governing gRNA molecule disclosed herein,

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

Headings, including numeric and alphabetical headings and subheadings, are for organization and presentation and are not intended to be limiting.

Other features and advantages of the invention will be apparent from the detailed description, drawings, and from the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

Figs. 1A-1G are representations of several exemplary gRNAs.

Fig. 1A depicts a modular gRNA molecule derived in part (or modeled on a sequence in part) from *Streptococcus pyogenes (S. pyogenes)* as a duplexed structure (SEQ ID NOS: 42 and 43, respectively, in order of appearance);

Fig. 1B depicts a unimolecular (or chimeric) gRNA molecule derived in part from *S. pyogenes* as a duplexed structure (SEQ ID NO: 44);

Fig. 1C depicts a unimolecular gRNA molecule derived in part from *S. pyogenes* as a duplexed structure (SEQ ID NO: 45);

Fig. 1D depicts a unimolecular gRNA molecule derived in part from *S. pyogenes* as a duplexed structure (SEQ ID NO: 46);

Fig. 1E depicts a unimolecular gRNA molecule derived in part from *S. pyogenes* as a duplexed structure (SEQ ID NO: 47);

Fig. 1F depicts a modular gRNA molecule derived in part from *Streptococcus thermophilus (S. thermophilus)* as a duplexed structure (SEQ ID NOS: 48 and 49, respectively, in order of appearance);

Fig. 1G depicts an alignment of modular gRNA molecules of *S. pyogenes* and *S. thermophilus* (SEQ ID NOS: 50-53, respectively, in order of appearance).

Figs. 2A-2G depict an alignment of Cas9 sequences from Chylinski *et al.* (RNA Biol. 2013; 10(5): 726–737). The N-terminal RuvC-like domain is boxed and indicated with a "Y". The other two RuvC-like domains are boxed and indicated with a "B". The HNH-like domain is boxed and indicated by a "G". Sm: *S. mutans* (SEQ ID NO: 1); Sp: *S. pyogenes* (SEQ ID NO: 2); St: *S. thermophilus* (SEQ ID NO: 3); Li: *L. innocua* (SEQ ID NO: 4). Motif: this is a motif based on the four sequences: residues conserved in all four sequences are indicated by single letter amino acid abbreviation; "*" indicates any amino acid found in the corresponding position of any of the four sequences; and "-" indicates any amino acid, e.g., any of the 20 naturally occurring amino acids.

Figs. 3A-3B show an alignment of the N-terminal RuvC-like domain from the Cas9 molecules disclosed in Chylinski *et al* (SEQ ID NOS: 54-103, respectively, in order of appearance). The last line of **Fig. 3B** identifies 4 highly conserved residues.

Figs. 4A-4B show an alignment of the N-terminal RuvC-like domain from the Cas9 molecules disclosed in Chylinski *et al.* with sequence outliers removed (SEQ ID NOS: 104-177, respectively, in order of appearance). The last line of **Fig. 4B** identifies 3 highly conserved residues.

Figs. 5A-5C show an alignment of the HNH-like domain from the Cas9 molecules disclosed in Chylinski *et al* (SEQ ID NOS: 178-252, respectively, in order of appearance). The last line of **Fig. 5C** identifies conserved residues.

Figs. 6A-6B show an alignment of the HNH-like domain from the Cas9 molecules disclosed in Chylinski *et al.* with sequence outliers removed (SEQ ID NOS: 253-302, respectively, in order of appearance). The last line of **Fig. 6B** identifies 3 highly conserved residues.

Figs. 7A-7B depict an alignment of Cas9 sequences from *S. pyogenes* and *Neisseria meningitidis (N. meningitidis)*. The N-terminal RuvC-like domain is boxed and indicated with a "Y". The other two RuvC-like domains are boxed and indicated with a "B". The HNH-like domain is boxed and indicated with a "G". Sp: *S. pyogenes*; Nm: *N. meningitidis*. Motif: this is a motif based on the two sequences: residues conserved in both sequences are indicated by a single amino acid designation; "*" indicates any amino acid found in the corresponding position of any of the two sequences; "-" indicates any amino acid, e.g., any of the 20 naturally occurring amino acids, and "-" indicates any amino acid, e.g., any of the 20 naturally occurring amino acids, or absent.

Fig. 8 shows a nucleic acid sequence encoding Cas9 of *N. meningitidis* (SEQ ID NO: 303). Sequence indicated by an "R" is an SV40 NLS; sequence indicated as "G" is an HA tag; and sequence indicated by an "O" is a synthetic NLS sequence; the remaining (unmarked) sequence is the open reading frame (ORF).

Figs. 9A-9B are schematic representations of the domain organization of *S. pyogenes* Cas 9. **Fig. 9A** shows the organization of the Cas9 domains, including amino acid positions, in reference to the two lobes of Cas9 (recognition (REC) and nuclease (NUC) lobes). **Fig. 9B** shows the percent homology of each domain across 83 Cas9 orthologs.)

Fig. 10 shows the nucleotide locations of the *Alu* repeats, cryptic exon and point mutation, c.2991+1655 A to G in the human *CEP290* locus. "X" indicates the cryptic exon. The blue triangle indicates the LCA target position c.2991+1655A to G.

Fig. 11A-11B show the rates of indels induced by various gRNAs at the *CEP290* locus. **Fig. 11A** shows gene editing (% indels) as assessed by sequencing for *S. pyogenes* and *S. aureus* gRNAs when co-expressed with Cas9 in patient-derived IVS26 primary fibroblasts. **Fig. 11B** shows gene editing (% indels) as assessed by sequencing for *S. aureus* gRNAs when co-expressed with Cas9 in HEK293 cells.

Figs. 12A-12B show changes in expression of the wildtype and mutant (including cryptic exon) alleles of *CEP290* in cells transfected with Cas9 and the indicated gRNA pairs. Total RNA was isolated from modified cells and qRT-PCR with Taqman primer-probe sets was used to quantify expression. Expression is normalized to the Beta-Actin housekeeping gene and each sample is normalized to the GFP control sample (cells transfected with only GFP). Error bars represent standard deviation of 4 technical replicates.

Fig. 13 shows changes in gene expression of the wildtype and mutant (including cryptic exon) alleles of *CEP290* in cells transfected with Cas9 and pairs of gRNAs shown to have in initial qRT-PCR screening. Total RNA was isolated from modified cells and qRT-PCR with Taqman primer-probe sets was used to quantify expression. Expression is normalized to the Beta-Actin housekeeping gene and each sample is normalized to the GFP control sample (cells transfected with only GFP). Error bars represent standard error of the mean of two to six biological replicates.

Fig. 14 shows deletion rates in cells transfected with indicated gRNA pairs and Cas9 as measured by droplet digital PCR (ddPCR). % deletion was calculated by dividing the number of positive droplets in deletion assay by the number of positive droplets in a control assay. Three biological replicates are shown for two different gRNA pairs.

Fig. 15 shows deletion rates in 293T cells transfected with exemplary AAV expression plasmids. pSS10 encodes EFS-driven saCas9 without gRNA. pSS15 and pSS17 encode EFS-driven saCas9 and one U6-driven gRNA, CEP290-64 and CEP290-323 respectively. pSS11 encodes EFS-driven saCas9 and two U6-driven gRNAs, CEP290-64 and CEP290-323 in the same vector. Deletion PCR were performed with gDNA exacted from 293T cells post transfection. The size of the PCR amplicons indicates the presence or absence of deletion events, and the deletion ratio was calculated.

Fig. 16 shows the composition of structural proteins in AAV2 viral preps expressing Cas9. Reference AAV2 vectors (lanes 1 & 2) were obtained from Vector Core at University of

North Carolina, Chapel Hill. AAV2-CMV-GFP (lane 3) and AAV2-CMV-saCas9-minpA (lane4) were packaged and purified with "Triple Transfection Protocol" followed by CsCl ultracentrifugation. Titers were obtained by quantitative PCR with primers annealing to the ITR structures on these vectors. Viral preps were denatured and probed with B1 antibody on Western Blots to demonstrate three structural proteins composing AAV2, VP1, VP2, and VP3 respectively.

Fig. 17 depicts the deletion rates in 293T cells transduced with AAV viral vectors at MOI of 1000 viral genome (vg) per cell and 10,000 vg per cell. AAV2 viral vectors were produced with "Triple Transfection Protocol" using pHelper, pRep2Cap2, pSS8 encoding gRNAs CEP290-64 and CEP290-323, and CMV-driven saCas9. Viral preps were titered with primers annealing to ITRs on pSS8. 6 days post transduction, gDNA were extracted from 293T cells. Deletion PCR was carried out on the *CEP290* locus, and deletion rates were calculated based on the predicted amplicons. Western blotting was carried out to show the AAV-mediated saCas9 expression in 293T cells (primary antibody: anti-Flag, M2; loading control: anti-alphaTubulin).

Fig. 18A-18B depicts additional exemplary structures of unimolecular gRNA molecules. **Fig. 18A** shows an exemplary structure of a unimolecular gRNA molecule derived in part from *S. pyogenes* as a duplexed structure. **Fig. 18B** shows an exemplary structure of a unimolecular gRNA molecule derived in part from *S. aureus* as a duplexed structure.

Figs. 19A-19G depicts the nucleotide sequence of an exemplary recombinant AAV genome containg a CMV promoter. Various components of the recombinant AAV genome are also indicated. N = A, T, G or C. The number of N residues can vary, e.g., from 16 to 26 nucleotides. Upper stand: 5'→3' (SEQ ID NO: 428); lower stand: 3'→5' SEQ ID NO: 445).

Figs. 20A-20F depicts the nucleotide sequence of an exemplary recombinant AAV genome containg an EFS promoter. Various components of the recombinant AAV genome are also indicated. N = A, T, G or C. The number of N residues can vary, e.g., from 16 to 26 nucleotides. Upper stand: 5'→3' (SEQ ID NO: 429); lower stand: 3'→5' (SEQ ID NO: 446).

Figs. 21A-21F depicts the nucleotide sequence of an exemplary recombinant AAV genome containg a CRK1 promoter. Various components of the recombinant AAV genome are also indicated. N = A, T, G or C. The number of N residues can vary, e.g., from 16 to 26 nucleotides. Upper stand: 5'→3' (SEQ ID NO: 430); lower stand: 3'→5' (SEQ ID NO: 447).

Figs. 22A-22F depicts the nucleotide sequence of an exemplary recombinant AAV genome containg a CRX promoter. Various components of the recombinant AAV genome are also indicated. N = A, T, G or C. The number of N residues can vary, e.g., from 16 to 26 nucleotides. Upper stand: 5'→3' (SEQ ID NO: 431); lower stand: 3'→5' (SEQ ID NO: 448).

Figs. 23A-23F depicts the nucleotide sequence of an exemplary recombinant AAV genome containg a NRL promoter. Various components of the recombinant AAV genome are also indicated. N = A, T, G or C. The number of N residues can vary, e.g., from 16 to 26 nucleotides. Upper stand: 5'→3' (SEQ ID NO: 432); lower stand: 3'→5' (SEQ ID NO: 449).

Figs. 24A-24F depicts the nucleotide sequence of an exemplary recombinant AAV genome containg a NRL promoter. Various components of the recombinant AAV genome are also indicated. N = A, T, G or C. The number of N residues can vary, e.g., from 16 to 26 nucleotides. Upper stand: 5'→3' (SEQ ID NO: 433); lower stand: 3'→5' (SEQ ID NO: 450).

DETAILED DESCRIPTION

Definitions

"Domain", as used herein, is used to describe segments of a protein or nucleic acid. Unless otherwise indicated, a domain is not required to have any specific functional property.

Calculations of homology or sequence identity between two sequences (the terms are used interchangeably herein) are performed as follows. The sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). The optimal alignment is determined as the best score using the GAP program in the GCG software package with a Blossum 62 scoring matrix with a gap penalty of 12, a gap extend penalty of 4, and a frameshift gap penalty of 5. The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences.

"Governing gRNA molecule", as used herein, refers to a gRNA molecule that comprises a targeting domain that is complementary to a target domain on a nucleic acid that comprises a

sequence that encodes a component of the CRISPR/Cas system that is introduced into a cell or subject. A governing gRNA does not target an endogenous cell or subject sequence. In an embodiment, a governing gRNA molecule comprises a targeting domain that is complementary with a target sequence on: (a) a nucleic acid that encodes a Cas9 molecule; (b) a nucleic acid that encodes a gRNA which comprises a targeting domain that targets the *CEP290* gene (a target gene gRNA); or on more than one nucleic acid that encodes a CRISPR/Cas component, e.g., both (a) and (b). In an embodiment, a nucleic acid molecule that encodes a CRISPR/Cas component, e.g., that encodes a Cas9 molecule or a target gene gRNA, comprises more than one target domain that is complementary with a governing gRNA targeting domain. While not wishing to be bound by theory, it is believed that a governing gRNA molecule complexes with a Cas9 molecule and results in Cas9 mediated inactivation of the targeted nucleic acid, e.g., by cleavage or by binding to the nucleic acid, and results in cessation or reduction of the production of a CRISPR/Cas system component. In an embodiment, the Cas9 molecule forms two complexes: a complex comprising a Cas9 molecule with a target gene gRNA, which complex will alter the *CEP290* gene; and a complex comprising a Cas9 molecule with a governing gRNA molecule, which complex will act to prevent further production of a CRISPR/Cas system component, e.g., a Cas9 molecule or a target gene gRNA molecule. In an embodiment, a governing gRNA molecule/Cas9 molecule complex binds to or promotes cleavage of a control region sequence, e.g., a promoter, operably linked to a sequence that encodes a Cas9 molecule, a sequence that encodes a transcribed region, an exon, or an intron, for the Cas9 molecule. In an embodiment, a governing gRNA molecule/Cas9 molecule complex binds to or promotes cleavage of a control region sequence, e.g., a promoter, operably linked to a gRNA molecule, or a sequence that encodes the gRNA molecule. In an embodiment, the governing gRNA, e.g., a Cas9-targeting governing gRNA molecule, or a target gene gRNA-targeting governing gRNA molecule, limits the effect of the Cas9 molecule/target gene gRNA molecule complex-mediated gene targeting. In an embodiment, a governing gRNA places temporal, level of expression, or other limits, on activity of the Cas9 molecule/target gene gRNA molecule complex. In an embodiment, a governing gRNA reduces off-target or other unwanted activity. In an embodiment, a governing gRNA molecule inhibits, e.g., entirely or substantially entirely inhibits, the production of a component of the Cas9 system and thereby limits, or governs, its activity.

"Modulator", as used herein, refers to an entity, e.g., a drug that can alter the activity (e.g., enzymatic activity, transcriptional activity, or translational activity), amount, distribution, or structure of a subject molecule or genetic sequence. In an embodiment, modulation comprises cleavage, e.g., breaking of a covalent or non-covalent bond, or the forming of a covalent or non-covalent bond, e.g., the attachment of a moiety, to the subject molecule. In an embodiment, a modulator alters the, three dimensional, secondary, tertiary, or quaternary structure, of a subject molecule. A modulator can increase, decrease, initiate, or eliminate a subject activity.

"Large molecule", as used herein, refers to a molecule having a molecular weight of at least 2, 3, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, or 100 kD. Large molecules include proteins, polypeptides, nucleic acids, biologics, and carbohydrates.

"Polypeptide", as used herein, refers to a polymer of amino acids having less than 100 amino acid residues. In an embodiment, it has less than 50, 20, or 10 amino acid residues.

"Non-homologous end joining" or "NHEJ", as used herein, refers to ligation mediated repair and/or non-template mediated repair including, e.g., canonical NHEJ (cNHEJ), alternative NHEJ (altNHEJ), microhomology-mediated end joining (MMEJ), single-strand annealing (SSA), and synthesis-dependent microhomology-mediated end joining (SD-MMEJ).

"Reference molecule", e.g., a reference Cas9 molecule or reference gRNA, as used herein, refers to a molecule to which a subject molecule, e.g., a subject Cas9 molecule of subject gRNA molecule, e.g., a modified or candidate Cas9 molecule is compared. For example, a Cas9 molecule can be characterized as having no more than 10% of the nuclease activity of a reference Cas9 molecule. Examples of reference Cas9 molecules include naturally occurring unmodified Cas9 molecules, e.g., a naturally occurring Cas9 molecule such as a Cas9 molecule of *S. pyogenes*, *S. aureus*, or *S. thermophilus*. In an embodiment, the reference Cas9 molecule is the naturally occurring Cas9 molecule having the closest sequence identity or homology with the Cas9 molecule to which it is being compared. In an embodiment, the reference Cas9 molecule is a sequence, e.g., a naturally occurring or known sequence, which is the parental form on which a change, e.g., a mutation has been made.

"Replacement", or "replaced", as used herein with reference to a modification of a molecule does not require a process limitation but merely indicates that the replacement entity is present.

"Small molecule", as used herein, refers to a compound having a molecular weight less than about 2 kD, e.g., less than about 2 kD, less than about 1.5 kD, less than about 1 kD, or less than about 0.75 kD.

"Subject", as used herein, means either a human or non-human animal. The term includes, but is not limited to, mammals (e.g., humans, other primates, pigs, rodents (e.g., mice and rats or hamsters), rabbits, guinea pigs, cows, horses, cats, dogs, sheep, and goats). In an embodiment, the subject is a human. In other embodiments, the subject is poultry.

"Treat", "treating" and "treatment", as used herein, mean the treatment of a disease in a mammal, e.g., in a human, including (a) inhibiting the disease, i.e., arresting or preventing its development; (b) relieving the disease, i.e., causing regression of the disease state; and (c) curing the disease.

"X" as used herein in the context of an amino acid sequence, refers to any amino acid (e.g., any of the twenty natural amino acids) unless otherwise specified.

Methods of Altering *CEP290*

CEP290 encodes a centrosomal protein that plays a role in centrosome and cilia development. The *CEP290* gene is involved in forming cilia around cells, particularly in the photoreceptors at the back of the retina, which are needed to detect light and color.

Disclosed herein are methods and compositions for altering the LCA10 target position in the *CEP290* gene. *LCA10* target position can be altered (e.g., corrected) by gene editing, e.g., using CRISPR-Cas9 mediated methods. The alteration (e.g., correction) of the mutant *CEP290* gene can be mediated by any mechanism. Exemplary mechanisms that can be associated with the alteration (e.g., correction) of the mutatnt *CEP290* gene include, but are not limited to, non-homologous end joining (e.g., classical or alternative), microhomology-mediated end joining (MMEJ), homology-directed repair (e.g., endogenous donor template mediated), SDSA (synthesis dependent strand annealing), single strand annealing or single strand invasion. Methods described herein introduce one or more breaks near the site of the LCA target position (e.g., c.2991+1655A to G) in at least one allele of the *CEP290* gene. In an embodiment, the one or more breaks are repaired by NHEJ. During repair of the one or more breaks, DNA sequences are inserted and/or deleted resulting in the loss or destruction of the cryptic splice site resulting from the mutation at the LCA10 target position (e.g., c.2991+1655A to G). The method can

include acquiring knowledge of the mutation carried by the subject, e.g., by sequencing the appropriate portion of the *CEP290* gene.

Altering the LCA10 target position refers to (1) break-induced introduction of an indel (also referred to herein as NHEJ-mediated introduction of an indel) in close proximity to or including a LCA10 target position (e.g., c.2991+1655A to G), or (2) break-induced deletion (also referred to herein as NHEJ-mediated deletion) of genomic sequence including the mutation at a LCA10 target position (e.g., c.2991+1655A to G). Both approaches give rise to the loss or destruction of the cryptic splice site.

In an embodiment, the method comprises introducing a break-induced indel in close proximity to or including the LCA10 target position (e.g., c.2991+1655A to G). As described herein, in one embodiment, the method comprises the introduction of a double strand break sufficiently close to (e.g., either 5' or 3' to) the LCA10 target position, e.g., c.2991+1655A to G, such that the break-induced indel could be reasonably expected to span the mutation. A single gRNAs, e.g., unimolecular (or chimeric) or modular gRNA molecules, is configured to position a double strand break sufficiently close to the LCA10 target position in the *CEP290* gene. In an embodiment, the break is positioned to avoid unwanted target chromosome elements, such as repeat elements, e.g., *an Alu repeat*. The double strand break may be positioned within 40 nucleotides (e.g., within 1, 2, 3, 4, 5, 10, 15, 16, 17, 18, 19, 20, 25, 30, 35 or 40 nucleotides) upstream of the LCA10 target position, or within 40 nucleotides (e.g., within 1, 2, 3, 4, 5, 10, 15, 16, 17, 18, 19, 20, 25, 30, 35 or 40 nucleotides) downstream of the LCA10 target position (see **Fig. 9**). While not wishing to be bound by theory, in an embodiment, it is believed that NHEJ-mediated repair of the double strand break allows for the NHEJ-mediated introduction of an indel in close proximity to or including the LCA10 target position.

In another embodiment, the method comprises the introduction of a pair of single strand breaks sufficiently close to (either 5' or 3' to, respectively) the mutation at the LCA10 target position (e.g., c.2991+1655A to G) such that the break-induced indel could be reasonably expected to span the mutation. Two gRNAs, e.g., unimolecular (or chimeric) or modular gRNA molecules, are configured to position the two single strand breaks sufficiently close to the LCA10 target position in the *CEP290* gene. In an embodiment, the breaks are positioned to avoid unwanted target chromosome elements, such as repeat elements, e.g., *an Alu repeat*. In an embodiment, the pair of single strand breaks is positioned within 40 nucleotides (e.g., within 1,

2, 3, 4, 5, 10, 15, 16, 17, 18, 19, 20, 25, 30, 35 or 40 nucleotides) upstream of the LCA10 target position, or within 40 nucleotides (e.g., within 1, 2, 3, 4, 5, 10, 15, 16, 17, 18, 19, 20, 25, 30, 35 or 40 nucleotides) downstream of the LCA10 target position (see **Fig. 9**). While not wishing to be bound by theory, in an embodiment, it is believed that NHEJ mediated repair of the pair of single strand breaks allows for the NHEJ-mediated introduction of an indel in close proximity to or including the LCA10 target position. In an embodiment, the pair of single strand breaks may be accompanied by an additional double strand break, positioned by a third gRNA molecule, as is discussed below. In another embodiment, the pair of single strand breaks may be accompanied by two additional single strand breaks positioned by a third gRNA molecule and a fourth gRNA molecule, as is discussed below.

In an embodiment, the method comprises introducing a break-induced deletion of genomic sequence including the mutation at the LCA10 target position (e.g., c.2991+1655A to G). As described herein, in one embodiment, the method comprises the introduction of two double strand breaks—one 5' and the other 3' to (i.e., flanking) the LCA10 target position. Two gRNAs, e.g., unimolecular (or chimeric) or modular gRNA molecules, are configured to position the two double strand breaks on opposite sides of the LCA10 target position in the *CEP290* gene. In an embodiment, the first double strand break is positioned upstream of the LCA10 target position within intron 26 (e.g., within 1654 nucleotides), and the second double strand break is positioned downstream of the LCA10 target position within intron 26 (e.g., within 4183 nucleotides) (see **Fig. 10**). In an embodiment, the breaks (i.e., the two double strand breaks) are positioned to avoid unwanted target chromosome elements, such as repeat elements, e.g., an *Alu* repeat, or the endogenous *CEP290* splice sites.

The first double strand break may be positioned as follows:

(1) upstream of the 5' end of the *Alu* repeat in intron 26,

(2) between the 3' end of the *Alu* repeat and the LCA10 target position in intron 26, or

(3) within the Alu repeat provided that a sufficient length of the gRNA fall outside of the repeat so as to avoid binding to other Alu repeats in the genome,

and the second double strand break to be paired with the first double strand break may be positioned downstream of the LCA10 target position in intron 26.

For example, the first double strand break may be positioned:

(1) within 1162 nucleotides upstream of the 5' end of the *Alu* repeat,

(2) within 1000 nucleotides upstream of the 5' end of the *Alu* repeat,

(3) within 900 nucleotides upstream of the 5' end of the *Alu* repeat,

(4) within 800 nucleotides upstream of the 5' end of the *Alu* repeat,

(5) within 700 nucleotides upstream of the 5' end of the *Alu* repeat,

(6) within 600 nucleotides upstream of the 5' end of the *Alu* repeat,

(7) within 500 nucleotides upstream of the 5' end of the *Alu* repeat,

(8) within 400 nucleotides upstream of the 5' end of the *Alu* repeat,

(9) within 300 nucleotides upstream of the 5' end of the *Alu* repeat,

(10) within 200 nucleotides upstream of the 5' end of the *Alu* repeat,

(11) within 100 nucleotides upstream of the 5' end of the *Alu* repeat,

(12) within 50 nucleotides upstream of the 5' end of the *Alu* repeat,

(13) within the *Alu* repeat provided that a sufficient length of the gRNA falls outside of the repeat so as to avoid binding to other Alu repeats in the genome,

(14) within 40 nucleotides (e.g., within 1, 2, 3, 4, 5, 10, 15, 16, 17, 18, 19, 20, 25, 30, 35 or 40 nucleotides) upstream of the LCA10 target position, or

(15) within 17 nucleotides (e.g., within 1, 2, 3, 4, 5, 10, 15, 16 or 17 nucleotides) upstream of the LCA10 target position,

and the second double strand breaks to be paired with the first double strand break may be positioned:

(1) within 4183 nucleotides downstream of the LCA10 target position,

(2) within 4000 nucleotides downstream of the LCA10 target position,

(3) within 3000 nucleotides downstream of the LCA10 target position,

(4) within 2000 nucleotides downstream of the LCA10 target position,

(5) within 1000 nucleotides downstream of the LCA10 target position,

(6) within 700 nucleotides downstream of the LCA10 target position,

(7) within 500 nucleotides downstream of the LCA10 target position,

(8) within 300 nucleotides downstream of the LCA10 target position,

(9) within 100 nucleotides downstream of the LCA10 target position,

(10) within 60 nucleotides downstream of the LCA10 target position, or

(11) within 40 (e.g., 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35 or 40 nucleotides) nucleotides downstream of the LCA10 target position.

While not wishing to be bound by theory, in an embodiment, it is believed that the two double strand breaks allow for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene.

The method also comprises the introduction of two sets of breaks, e.g., one double strand break (either 5' or 3' to the mutation at the LCA10 target position, e.g., c.2991+1655A to G) and a pair of single strand breaks (on the other side of the LCA10 target position opposite from the double strand break) such that the two sets of breaks are positioned to flank the LCA10 target position. Three gRNAs, e.g., unimolecular (or chimeric) or modular gRNA molecules, are configured to position the one double strand break and the pair of single strand breaks on opposite sides of the LCA10 target position in the *CEP290* gene. In an embodiment, the first set of breaks (either the double strand break or the pair of single strand breaks) is positioned upstream of the LCA10 target position within intron 26 (e.g., within 1654 nucleotides), and the second set of breaks (either the double strand break or the pair of single strand breaks) are positioned downstream of the LCA10 target position within intron 26 (e.g., within 4183 nucleotides) (see **Fig. 10**). In an embodiment, the two sets of breaks (i.e., the double strand break and the pair of single strand breaks) are positioned to avoid unwanted target chromosome elements, such as repeat elements, e.g., an *Alu* repeat, or the endogenous *CEP290* splice sites.

The first set of breaks (either the double strand break or the pair of single strand breaks) may be positioned:

(1) upstream of the 5' end of the *Alu* repeat in intron 26,

(2) between the 3' end of the *Alu* repeat and the LCA10 target position in intron 26, or

(3) within the *Alu* repeat provided that a sufficient length of the gRNA falls outside of the repeat so as to avoid binding to other *Alu* repeats in the genome,

and the second set of breaks to be paired with the first set of breaks (either the double strand break or the pair of single strand breaks) may be positioned downstream of the LCA10 target position in intron 26.

For example, the first set of breaks (either the double strand break or the pair of single strand breaks) may be positioned:

(1) within 1162 nucleotides upstream of the 5' end of the *Alu* repeat,

(2) within 1000 nucleotides upstream of the 5' end of the *Alu* repeat,

(3) within 900 nucleotides upstream of the 5' end of the *Alu* repeat,

(4) within 800 nucleotides upstream of the 5' end of the *Alu* repeat,

(5) within 700 nucleotides upstream of the 5' end of the *Alu* repeat,

(6) within 600 nucleotides upstream of the 5' end of the *Alu* repeat,

(7) within 500 nucleotides upstream of the 5' end of the *Alu* repeat,

(8) within 400 nucleotides upstream of the 5' end of the *Alu* repeat,

(9) within 300 nucleotides upstream of the 5' end of the *Alu* repeat,

(10) within 200 nucleotides upstream of the 5' end of the *Alu* repeat,

(11) within 100 nucleotides upstream of the 5' end of the *Alu* repeat,

(12) within 50 nucleotides upstream of the 5' end of the *Alu* repeat,

(13) within the *Alu* repeat provided that a sufficient length of the gRNA falls outside of the repeat so as to avoid binding to other *Alu* repeats in the genome,

(14) within 40 nucleotides (e.g., within 1, 2, 3, 4, 5, 10, 15, 16, 17, 18, 19, 20, 25, 30, 35 or 40 nucleotides) upstream of the LCA10 target position, or

(15) within 17 nucleotides (e.g., within 1, 2, 3, 4, 5, 10, 15, 16 or 17 nucleotides) upstream of the LCA10 target position,

and the second set of breaks to be paired with the first set of breaks (either the double strand break or the pair of single strand breaks) may be positioned:

(1) within 4183 nucleotides downstream of the LCA10 target position,

(2) within 4000 nucleotides downstream of the LCA10 target position,

(3) within 3000 nucleotides downstream of the LCA10 target position,

(4) within 2000 nucleotides downstream of the LCA10 target position,

(5) within 1000 nucleotides downstream of the LCA10 target position,

(6) within 700 nucleotides downstream of the LCA10 target position,

(7) within 500 nucleotides downstream of the LCA10 target position,

(8) within 300 nucleotides downstream of the LCA10 target position,

(9) within 100 nucleotides downstream of the LCA10 target position,

(10) within 60 nucleotides downstream of the LCA10 target position, or

(11) within 40 (e.g., 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35 or 40 nucleotides) nucleotides downstream of the LCA10 target position.

While not wishing to be bound by theory, it is believed that the two sets of breaks (either the double strand break or the pair of single strand breaks) allow for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene.

The method also comprises the introduction of two sets of breaks, e.g., two pairs of single strand breaks, wherein the two sets of single-stranded breaks are positioned to flank the LCA10 target position. In an embodiment, the first set of breaks (e.g., the first pair of single strand breaks) is 5' to the mutation at the LCA10 target position (e.g., c.2991+1655A to G) and the second set of breaks (e.g., the second pair of single strand breaks) is 3' to the mutation at the LCA10 target position. Four gRNAs, e.g., unimolecular (or chimeric) or modular gRNA molecules, are configured to position the two pairs of single strand breaks on opposite sides of the LCA10 target position in the *CEP290* gene. In an embodiment, the first set of breaks (e.g., the first pair of single strand breaks) is positioned upstream of the LCA10 target position within intron 26 (e.g., within 1654 nucleotides), and the second set of breaks (e.g., the second pair of single strand breaks) is positioned downstream of the LCA10 target position within intron 26 (e.g., within 4183 nucleotides) (see **Fig. 10**). In an embodiment, the two sets of breaks (i.e., the two pairs of single strand breaks) are positioned to avoid unwanted target chromosome elements, such as repeat elements, e.g., an *Alu* repeat, or the endogenous *CEP290* splice sites.

The first set of breaks (e.g., the first pair of single strand breaks) may be positioned:

(1) upstream of the 5' end of the *Alu* repeat in intron 26,

(2) between the 3' end of the *Alu* repeat and the LCA10 target position in intron 26, or

(3) within the *Alu* repeat provided that a sufficient length of the gRNA falls outside of the repeat so as to avoid binding to other *Alu* repeats in the genome,

and the second set of breaks to be paired with the first set of breaks (e.g., the second pair of single strand breaks) may be positioned downstream of the LCA10 target position in intron 26.

For example, the first set of breaks (e.g., the first pair of single strand breaks) may be positioned:

(1) within 1162 nucleotides upstream of the 5' end of the *Alu* repeat,

(2) within 1000 nucleotides upstream of the 5' end of the *Alu* repeat,

(3) within 900 nucleotides upstream of the 5' end of the *Alu* repeat,

(4) within 800 nucleotides upstream of the 5' end of the *Alu* repeat,

(5) within 700 nucleotides upstream of the 5' end of the *Alu* repeat,

(6) within 600 nucleotides upstream of the 5' end of the *Alu* repeat,

(7) within 500 nucleotides upstream of the 5' end of the *Alu* repeat,

(8) within 400 nucleotides upstream of the 5' end of the *Alu* repeat,

(9) within 300 nucleotides upstream of the 5' end of the *Alu* repeat,

(10) within 200 nucleotides upstream of the 5' end of the *Alu* repeat,

(11) within 100 nucleotides upstream of the 5' end of the *Alu* repeat,

(12) within 50 nucleotides upstream of the 5' end of the *Alu* repeat,

(13) within the *Alu* repeat provided that a sufficient length of the gRNA falls outside of the repeat so as to avoid binding to other *Alu* repeats in the genome,

(14) within 40 nucleotides (e.g., within 1, 2, 3, 4, 5, 10, 15, 16, 17, 18, 19, 20, 25, 30, 35 or 40 nucleotides) upstream of the LCA10 target position, or

(15) within 17 nucleotides (e.g., within 1, 2, 3, 4, 5, 10, 15, 16 or 17 nucleotides) upstream of the LCA10 target position,

and the second set of breaks to be paired with the first set of breaks (e.g., the second pair of single strand breaks) may be positioned:

(1) within 4183 nucleotides downstream of the LCA10 target position,

(2) within 4000 nucleotides downstream of the LCA10 target position,

(3) within 3000 nucleotides downstream of the LCA10 target position,

(4) within 2000 nucleotides downstream of the LCA10 target position,

(5) within 1000 nucleotides downstream of the LCA10 target position,

(6) within 700 nucleotides downstream of the LCA10 target position,

(7) within 500 nucleotides downstream of the LCA10 target position,

(8) within 300 nucleotides downstream of the LCA10 target position,

(9) within 100 nucleotides downstream of the LCA10 target position,

(10) within 60 nucleotides downstream of the LCA10 target position, or

(11) within 40 (e.g., 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35 or 40 nucleotides) nucleotides downstream of the LCA10 target position.

While not wishing to be bound by theory, it is believed that the two sets of breaks (e.g., the two pairs of single strand breaks) allow for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene.

Methods to Treat or Prevent LCA10

Described herein are methods for treating or delaying the onset or progression of Leber's Congenital Amaurosis 10 (LCA10) caused by a c.2991+1655 A to G (adenine to guanine) mutation in the *CEP290* gene. The disclosed methods for treating or delaying the onset or progression of LCA10 alter the *CEP290* gene by genome editing using a gRNA targeting the LCA10 target position and a Cas9 enzyme. Details on gRNAs targeting the LCA10 target position and Cas9 enzymes are provided below.

In an embodiment, treatment is initiated prior to onset of the disease.

In an embodiment, treatment is initiated after onset of the disease.

In an embodiment, treatment is initiated prior to loss of visual acuity and/or sensitivity to glare.

In an embodiment, treatment is initiated at onset of loss of visual acuity.

In an embodiment, treatment is initiated after onset of loss of visual acuity and/or sensitivity to glare.

In an embodiment, treatment is initiated *in utero*.

In an embodiment, treatment is initiated after birth.

In an embodiment, treatment is initiated prior to the age of 1.

In an embodiment, treatment is initiated prior to the age of 2.

In an embodiment, treatment is initiated prior to the age of 5.

In an embodiment, treatment is initiated prior to the age of 10.

In an embodiment, treatment is initiated prior to the age of 15.

In an embodiment, treatment is initiated prior to the age of 20.

A subject's vision can evaluated, e.g., prior to treatment, or after treatment, e.g., to monitor the progress of the treatment. In an embodiment, the subject's vision is evaluated prior to treatment, e.g., to determine the need for treatment. In an embodiment, the subject's vision is evaluated after treatment has been initiated, e.g., to access the effectiveness of the treatment. Vision can be evaluated by one or more of: evaluating changes in function relative to the

contralateral eye, e.g., by utilizing retinal analytical techniques; by evaluating mean, median and distribution of change in best corrected visual acuity (BCVA); evaluation by Optical Coherence Tomography; evaluation of changes in visual field using perimetry; evaluation by full-field electroretinography (ERG); evaluation by slit lamp examination; evaluation of intraocular pressure; evaluation of autofluorescence, evaluation with fundoscopy; evaluation with fundus photography; evaluation with fluorescein angiography (FA); or evaluation of visual field sensitivity (FFST).

In an embodiment, a subject's vision may be assessed by measuring the subject's mobility, e.g., the subject's ability to maneuver in space.

In an embodiment, treatment is initiated in a subject who has tested positive for a mutation in the *CEP290* gene, e.g., prior to disease onset or in the earliest stages of disease.

In an embodiment, a subject has a family member that has been diagnosed with LCA10. For example, the subject has a family member that has been diagnosed with LCA10, and the subject demonstrates a symptom or sign of the disease or has been found to have a mutation in the *CEP290* gene.

In an embodiment, a cell (e.g., a retinal cell, e.g., a photoreceptor cell) from a subject suffering from or likely to develop LCA10 is treated *ex vivo*. In an embodiment, the cell is removed from the subject, altered as described herein, and introduced into, e.g., returned to, the subject.

In an embodiment, a cell (e.g., a retinal cell, e.g., a photoreceptor cell) altered to correct a mutation in the LCA10 target position is introduced into the subject.

In an embodiment, the cell is a retinual cell (e.g., retinal pigment epithelium cell), a photoreceptor cell, a horizontal cell, a bipolar cell, an amacrine cell, or a ganglion cell. In an embodiment, it is contemplated herein that a population of cells (e.g., a population of retinal cells, e.g., a population of photoreceptor cells) from a subject may be contacted *ex vivo* to alter a mutation in *CEP290*, e.g., a 2991+1655 A to G. In an embodiment, such cells are introduced to the subject's body to prevent or treat LCA10.

In an embodiment, the population of cells are a population of retinual cells (e.g., retinal pigment epithelium cells), photoreceptor cells, horizontal cells, bipolar cells, amacrine cells, ganglion cells, or a combination thereof.

In an embodiment, the method described herein comprises delivery of gRNA or other components described herein, e.g., a Cas9 molecule, by one or more AAV vectors, e.g., one or more AAV vectors described herein.

I. gRNA Molecules

A gRNA molecule, as that term is used herein, refers to a nucleic acid that promotes the specific targeting or homing of a gRNA molecule/Cas9 molecule complex to a target nucleic acid. gRNA molecules can be unimolecular (having a single RNA molecule), sometimes referred to herein as "chimeric" gRNAs, or modular (comprising more than one, and typically two, separate RNA molecules). A gRNA molecule comprises a number of domains. The gRNA molecule domains are described in more detail below.

Several exemplary gRNA structures, with domains indicated thereon, are provided in Fig. 1. While not wishing to be bound by theory, with regard to the three dimensional form, or intra- or inter-strand interactions of an active form of a gRNA, regions of high complementarity are sometimes shown as duplexes in Fig. 1 and other depictions provided herein.

In an embodiment, a unimolecular, or chimeric, gRNA comprises, preferably from 5' to 3':

a targeting domain (which is complementary to a target nucleic acid in the *CEP290* gene, e.g., a targeting domain from any of **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**);
a first complementarity domain;
a linking domain;
a second complementarity domain (which is complementary to the first complementarity domain);
a proximal domain; and
optionally, a tail domain.

In an embodiment, a modular gRNA comprises:

a first strand comprising, preferably from 5' to 3';

a targeting domain (which is complementary to a target nucleic acid in the *CEP290* gene, e.g., a targeting domain from **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**); and

a first complementarity domain; and

a second strand, comprising, preferably from 5' to 3':

optionally, a 5' extension domain;

a second complementarity domain;

a proximal domain; and

optionally, a tail domain.

The domains are discussed briefly below:

The Targeting Domain

Figs. 1A-1G provide examples of the placement of targeting domains.

The targeting domain comprises a nucleotide sequence that is complementary, e.g., at least 80, 85, 90, or 95% complementary, e.g., fully complementary, to the target sequence on the target nucleic acid. The targeting domain is part of an RNA molecule and will therefore comprise the base uracil (U), while any DNA encoding the gRNA molecule will comprise the base thymine (T). While not wishing to be bound by theory, in an embodiment, it is believed that the complementarity of the targeting domain with the target sequence contributes to specificity of the interaction of the gRNA molecule/Cas9 molecule complex with a target nucleic acid. It is understood that in a targeting domain and target sequence pair, the uracil bases in the targeting domain will pair with the adenine bases in the target sequence. In an embodiment, the target domain itself comprises in the 5' to 3' direction, an optional secondary domain, and a core domain. In an embodiment, the core domain is fully complementary with the target sequence. In an embodiment, the targeting domain is 5 to 50 nucleotides in length. The strand of the target nucleic acid with which the targeting domain is complementary is referred to herein as the complementary strand. Some or all of the nucleotides of the domain can have a modification, e.g., a modification found in Section VIII herein.

In an embodiment, the targeting domain is 16 nucleotides in length.

In an embodiment, the targeting domain is 17 nucleotides in length.

In an embodiment, the targeting domain is 18 nucleotides in length.
In an embodiment, the targeting domain is 19 nucleotides in length.
In an embodiment, the targeting domain is 20 nucleotides in length.
In an embodiment, the targeting domain is 21 nucleotides in length.
In an embodiment, the targeting domain is 22 nucleotides in length.
In an embodiment, the targeting domain is 23 nucleotides in length.
In an embodiment, the targeting domain is 24 nucleotides in length.
In an embodiment, the targeting domain is 25 nucleotides in length.
In an embodiment, the targeting domain is 26 nucleotides in length.
In an embodiment, the targeting domain comprises 16 nucleotides.
In an embodiment, the targeting domain comprises 17 nucleotides.
In an embodiment, the targeting domain comprises 18 nucleotides.
In an embodiment, the targeting domain comprises 19 nucleotides.
In an embodiment, the targeting domain comprises 20 nucleotides.
In an embodiment, the targeting domain comprises 21 nucleotides.
In an embodiment, the targeting domain comprises 22 nucleotides.
In an embodiment, the targeting domain comprises 23 nucleotides.
In an embodiment, the targeting domain comprises 24 nucleotides.
In an embodiment, the targeting domain comprises 25 nucleotides.
In an embodiment, the targeting domain comprises 26 nucleotides.
Targeting domains are discussed in more detail below.

The First Complementarity Domain

Figs. 1A-1G provide examples of first complementarity domains.

The first complementarity domain is complementary with the second complementarity domain, and in an embodiment, has sufficient complementarity to the second complementarity domain to form a duplexed region under at least some physiological conditions. In an embodiment, the first complementarity domain is 5 to 30 nucleotides in length. In an embodiment, the first complementarity domain is 5 to 25 nucleotides in length. In an embodiment, the first complementary domain is 7 to 25 nucleotides in length. In an embodiment, the first complementary domain is 7 to 22 nucleotides in length. In an embodiment, the first

complementary domain is 7 to 18 nucleotides in length. In an embodiment, the first complementary domain is 7 to 15 nucleotides in length. In an embodiment, the first complementary domain is 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotides in length.

In an embodiment, the first complentarity domain comprises 3 subdomains, which, in the 5' to 3' direction are: a 5' subdomain, a central subdomain, and a 3' subdomain. In an embodiment, the 5' subdomain is 4-9, e.g., 4, 5, 6, 7, 8 or 9 nucleotides in length. In an embodiment, the central subdomain is 1, 2, or 3, e.g., 1, nucleotide in length. In an embodiment, the 3' subdomain is 3 to 25, e.g., 4-22, 4-18, or 4 to 10, or 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25, nucleotides in length.

The first complementarity domain can share homology with, or be derived from, a naturally occurring first complementarity domain. In an embodiment, it has at least 50% homology with a first complementarity domain disclosed herein, e.g., an *S. pyogenes*, *S. aureus,* or *S. thermophilus,* first complementarity domain.

Some or all of the nucleotides of the domain can have a modification, e.g., a modification found in Section VIII herein.

First complementarity domains are discussed in more detail below.

The Linking Domain

Figs. 1A-1G provide examples of linking domains.

A linking domain serves to link the first complementarity domain with the second complementarity domain of a unimolecular gRNA. The linking domain can link the first and second complementarity domains covalently or non-covalently. In an embodiment, the linkage is covalent. In an embodiment, the linking domain covalently couples the first and second complementarity domains, see, e.g., **Figs. 1B-1E**. In an embodiment, the linking domain is, or comprises, a covalent bond interposed between the first complementarity domain and the second complementarity domain. Typically the linking domain comprises one or more, e.g., 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides.

In modular gRNA molecules the two molecules are associated by virtue of the hybridization of the complementarity domains see e.g., **Fig. 1A**.

A wide variety of linking domains are suitable for use in unimolecular gRNA molecules. Linking domains can consist of a covalent bond, or be as short as one or a few nucleotides, e.g., 1, 2, 3, 4, or 5 nucleotides in length. In an embodiment, a linking domain is 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, or 25 or more nucleotides in length. In an embodiment, a linking domain is 2 to 50, 2 to 40, 2 to 30, 2 to 20, 2 to 10, or 2 to 5 nucleotides in length. In an embodiment, a linking domain shares homology with, or is derived from, a naturally occurring sequence, e.g., the sequence of a tracrRNA that is 5' to the second complementarity domain. In an embodiment, the linking domain has at least 50% homology with a linking domain disclosed herein.

Some or all of the nucleotides of the domain can have a modification, e.g., a modification found in Section VIII herein.

Linking domains are discussed in more detail below.

The 5' Extension Domain

In an embodiment, a modular gRNA can comprise additional sequence, 5' to the second complementarity domain, referred to herein as the 5' extension domain, see, e.g., **Fig. 1A**. In an embodiment, the 5' extension domain is, 2-10, 2-9, 2-8, 2-7, 2-6, 2-5, or 2-4 nucleotides in length. In an embodiment, the 5' extension domain is 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more nucleotides in length.

The Second Complementarity Domain

Figs. 1A-1G provide examples of second complementarity domains.

The second complementarity domain is complementary with the first complementarity domain, and in an embodiment, has sufficient complementarity to the second complementarity domain to form a duplexed region under at least some physiological conditions. In an embodiment, e.g., as shown in **Figs. 1A-1B**, the second complementarity domain can include sequence that lacks complementarity with the first complementarity domain, e.g., sequence that loops out from the duplexed region.

In an embodiment, the second complementarity domain is 5 to 27 nucleotides in length. In an embodiment, it is longer than the first complementarity region. In an embodiment the second complementary domain is 7 to 27 nucleotides in length. In an embodiment, the second complementary domain is 7 to 25 nucleotides in length. In an embodiment, the second

complementary domain is 7 to 20 nucleotides in length. In an embodiment, the second complementary domain is 7 to 17 nucleotides in length. In an embodiment, the complementary domain is 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 nucleotides in length.

In an embodiment, the second complentarity domain comprises 3 subdomains, which, in the 5' to 3' direction are: a 5' subdomain, a central subdomain, and a 3' subdomain. In an embodiment, the 5' subdomain is 3 to 25, e.g., 4 to 22, 4 to18, or 4 to 10, or 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotides in length. In an embodiment, the central subdomain is 1, 2, 3, 4 or 5, e.g., 3, nucleotides in length. In an embodiment, the 3' subdomain is 4 to 9, e.g., 4, 5, 6, 7, 8 or 9 nucleotides in length.

In an embodiment, the 5' subdomain and the 3' subdomain of the first complementarity domain, are respectively, complementary, e.g., fully complementary, with the 3' subdomain and the 5' subdomain of the second complementarity domain.

The second complementarity domain can share homology with or be derived from a naturally occurring second complementarity domain. In an embodiment, it has at least 50% homology with a second complementarity domain disclosed herein, e.g., an *S. pyogenes*, *S. aureus*, or *S. thermophilus*, first complementarity domain.

Some or all of the nucleotides of the domain can have a modification, e.g., a modification found in Section VIII herein.

A Proximal domain

Figs. 1A-1G provide examples of proximal domains.

In an embodiment, the proximal domain is 5 to 20 nucleotides in length. In an embodiment, the proximal domain can share homology with or be derived from a naturally occurring proximal domain. In an embodiment, it has at least 50% homology with a proximal domain disclosed herein, e.g., an *S. pyogenes*, *S. aureus*, or *S. thermophilus,* proximal domain.

Some or all of the nucleotides of the domain can have a modification, e.g., a modification found in Section VIII herein.

A Tail Domain

Figs. 1A-1G provide examples of tail domains.

As can be seen by inspection of the tail domains in **Figs. 1A** and **1B-1F**, a broad spectrum of tail domains are suitable for use in gRNA molecules. In an embodiment, the tail domain is 0 (absent), 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides in length. In embodiment, the tail domain nucleotides are from or share homology with sequence from the 5' end of a naturally occurring tail domain, see e.g., **Fig. 1D** or **1E**. In an embodiment, the tail domain includes sequences that are complementary to each other and which, under at least some physiological conditions, form a duplexed region.

In an embodiment, the tail domain is absent or is 1 to 50 nucleotides in length. In an embodiment, the tail domain can share homology with or be derived from a naturally occurring proximal tail domain. In an embodiment, it has at least 50% homology with a tail domain disclosed herein, e.g., an *S. pyogenes*, *S. aureus*, or *S. thermophilus,* tail domain.

In an embodiment, the tail domain includes nucleotides at the 3' end that are related to the method of in vitro or in vivo transcription. When a T7 promoter is used for in vitro transcription of the gRNA, these nucleotides may be any nucleotides present before the 3' end of the DNA template. When a U6 promoter is used for in vivo transcription, these nucleotides may be the sequence UUUUUU. When alternate pol-III promoters are used, these nucleotides may be various numbers or uracil bases or may include alternate bases.

The domains of gRNA molecules are described in more detail below.

The Targeting Domain

The "targeting domain" of the gRNA is complementary to the "target domain" on the target nucleic acid. The strand of the target nucleic acid comprising the core domain target is referred to herein as the "complementary strand" of the target nucleic acid. Guidance on the selection of targeting domains can be found, e.g., in Fu Y *et al.*, Nat Biotechnol 2014 (doi: 10.1038/nbt.2808) and Sternberg SH *et al.*, Nature 2014 (doi: 10.1038/nature13011).

In an embodiment, the targeting domain is 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 nucleotides in length.

In an embodiment, the targeting domain is 16 nucleotides in length.

In an embodiment, the targeting domain is 17 nucleotides in length.

In an embodiment, the targeting domain is 18 nucleotides in length.

In an embodiment, the targeting domain is 19 nucleotides in length.

In an embodiment, the targeting domain is 20 nucleotides in length.

In an embodiment, the targeting domain is 21 nucleotides in length.

In an embodiment, the targeting domain is 22 nucleotides in length.

In an embodiment, the targeting domain is 23 nucleotides in length.

In an embodiment, the targeting domain is 24 nucleotides in length.

In an embodiment, the targeting domain is 25 nucleotides in length.

In an embodiment, the targeting domain is 26 nucleotides in length.

In an embodiment, the targeting domain comprises 16 nucleotides.

In an embodiment, the targeting domain comprises 17 nucleotides.

In an embodiment, the targeting domain comprises 18 nucleotides.

In an embodiment, the targeting domain comprises 19 nucleotides.

In an embodiment, the targeting domain comprises 20 nucleotides.

In an embodiment, the targeting domain comprises 21 nucleotides.

In an embodiment, the targeting domain comprises 22 nucleotides.

In an embodiment, the targeting domain comprises 23 nucleotides.

In an embodiment, the targeting domain comprises 24 nucleotides.

In an embodiment, the targeting domain comprises 25 nucleotides.

In an embodiment, the targeting domain comprises 26 nucleotides.

In an embodiment, the targeting domain is 10 +/-5, 20+/-5, 30+/-5, 40+/-5, 50+/-5, 60+/-5, 70+/-5, 80+/-5, 90+/-5, or 100+/-5 nucleotides, in length.

In an embodiment, the targeting domain is 20+/-5 nucleotides in length.

In an embodiment, the targeting domain is 20+/-10, 30+/-10, 40+/-10, 50+/-10, 60+/-10, 70+/-10, 80+/-10, 90+/-10, or 100+/-10 nucleotides, in length.

In an embodiment, the targeting domain is 30+/-10 nucleotides in length.

In an embodiment, the targeting domain is 10 to 100, 10 to 90, 10 to 80, 10 to 70, 10 to 60, 10 to 50, 10 to 40, 10 to 30, 10 to 20 or 10 to 15 nucleotides in length. In other embodiments, the targeting domain is 20 to 100, 20 to 90, 20 to 80, 20 to 70, 20 to 60, 20 to 50, 20 to 40, 20 to 30, or 20 to 25 nucleotides in length.

Typically the targeting domain has full complementarity with the target sequence. In some embodiments the targeting domain has or includes 1, 2, 3, 4, 5, 6, 7 or 8 nucleotides that are not complementary with the corresponding nucleotide of the targeting domain.

In an embodiment, the target domain includes 1, 2, 3, 4 or 5 nucleotides that are complementary with the corresponding nucleotide of the targeting domain within 5 nucleotides of its 5' end. In an embodiment, the target domain includes 1, 2, 3, 4 or 5 nucleotides that are complementary with the corresponding nucleotide of the targeting domain within 5 nucleotides of its 3' end.

In an embodiment, the target domain includes 1, 2, 3, or 4 nucleotides that are not complementary with the corresponding nucleotide of the targeting domain within 5 nucleotides of its 5' end. In an embodiment, the target domain includes 1, 2, 3, or 4 nucleotides that are not complementary with the corresponding nucleotide of the targeting domain within 5 nucleotides of its 3' end.

In an embodiment, the degree of complementarity, together with other properties of the gRNA, is sufficient to allow targeting of a Cas9 molecule to the target nucleic acid.

In some embodiments, the targeting domain comprises two consecutive nucleotides that are not complementary to the target domain ("non-complementary nucleotides"), e.g., two consecutive noncomplementary nucleotides that are within 5 nucleotides of the 5' end of the targeting domain, within 5 nucleotides of the 3' end of the targeting domain, or more than 5 nucleotides away from one or both ends of the targeting domain.

In an embodiment, no two consecutive nucleotides within 5 nucleotides of the 5' end of the targeting domain, within 5 nucleotides of the 3' end of the targeting domain, or within a region that is more than 5 nucleotides away from one or both ends of the targeting domain, are not complementary to the targeting domain.

In an embodiment, there are no noncomplementary nucleotides within 5 nucleotides of the 5' end of the targeting domain, within 5 nucleotides of the 3' end of the targeting domain, or within a region that is more than 5 nucleotides away from one or both ends of the targeting domain.

In an embodiment, the targeting domain nucleotides do not comprise modifications, e.g., modifications of the type provided in Section VIII. However, in an embodiment, the targeting domain comprises one or more modifications, e.g., modifications that it render it less susceptible to degradation or more bio-compatible, e.g., less immunogenic. By way of example, the backbone of the targeting domain can be modified with a phosphorothioate, or other modification(s) from Section VIII. In an embodiment, a nucleotide of the targeting domain can

comprise a 2' modification, e.g., a 2-acetylation, e.g., a 2' methylation, or other modification(s) from Section VIII.

In some embodiments, the targeting domain includes 1, 2, 3, 4, 5, 6, 7 or 8 or more modifications. In an embodiment, the targeting domain includes 1, 2, 3, or 4 modifications within 5 nucleotides of its 5' end. In an embodiment, the targeting domain comprises as many as 1, 2, 3, or 4 modifications within 5 nucleotides of its 3' end.

In some embodiments, the targeting domain comprises modifications at two consecutive nucleotides, e.g., two consecutive nucleotides that are within 5 nucleotides of the 5' end of the targeting domain, within 5 nucleotides of the 3' end of the targeting domain, or more than 5 nucleotides away from one or both ends of the targeting domain.

In an embodiment, no two consecutive nucleotides are modified within 5 nucleotides of the 5' end of the targeting domain, within 5 nucleotides of the 3' end of the targeting domain, or within a region that is more than 5 nucleotides away from one or both ends of the targeting domain. In an embodiment, no nucleotide is modified within 5 nucleotides of the 5' end of the targeting domain, within 5 nucleotides of the 3' end of the targeting domain, or within a region that is more than 5 nucleotides away from one or both ends of the targeting domain.

Modifications in the targeting domain can be selected to not interfere with targeting efficacy, which can be evaluated by testing a candidate modification in the system described in Section IV gRNA's having a candidate targeting domain having a selected length, sequence, degree of complementarity, or degree of modification, can be evaluated in a system in Section IV. The candidate targeting domain can be placed, either alone, or with one or more other candidate changes in a gRNA molecule/Cas9 molecule system known to be functional with a selected target and evaluated.

In some embodiments, all of the modified nucleotides are complementary to and capable of hybridizing to corresponding nucleotides present in the target domain. In other embodiments, 1, 2, 3, 4, 5, 6, 7 or 8 or more modified nucleotides are not complementary to or capable of hybridizing to corresponding nucleotides present in the target domain.

In an embodiment, the targeting domain comprises, preferably in the 5'→3' direction: a secondary domain and a core domain. These domains are discussed in more detail below.

The Core Domain and Secondary Domain of the Targeting Domain

The "core domain" of the targeting domain is complementary to the "core domain target" on the target nucleic acid. In an embodiment, the core domain comprises about 8 to about 13 nucleotides from the 3' end of the targeting domain (e.g., the most 3' 8 to 13 nucleotides of the targeting domain).

In an embodiment, the secondary domain is absent or optional.

In an embodiment, the core domain and targeting domain, are independently, 6 +/-2, 7+/-2, 8+/-2, 9+/-2, 10+/-2, 11+/-2, 12+/-2, 13+/-2, 14+/-2, 15+/-2, 16+-2, 17+/-2, or 18+/-2, nucleotides in length.

In an embodiment, the core domain and targeting domain, are independently, 10+/-2 nucleotides in length.

In an embodiment, the core domain and targeting domain, are independently, 10+/-4 nucleotides in length.

In an embodiment, the core domain and targeting domain, are independently, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, or 18, nucleotides in length.

In an embodiment, the core domain and targeting domain, are independently 3 to 20, 4 to 20, 5 to 20, 6 to 20, 7 to 20, 8 to 20, 9 to 20 10 to 20 or 15 to 20 nucleotides in length.

In an embodiment, the core domain and targeting domain, are independently 3 to 15, e.g., 6 to 15, 7 to 14, 7 to 13, 6 to 12, 7 to 12, 7 to 11, 7 to 10, 8 to 14, 8 to 13, 8 to 12, 8 to 11, 8 to 10 or 8 to 9 nucleotides in length.

The core domain is complementary with the core domain target. Typically the core domain has exact complementarity with the core domain target. In some embodiments, the core domain can have 1, 2, 3, 4 or 5 nucleotides that are not complementary with the corresponding nucleotide of the core domain. In an embodiment, the degree of complementarity, together with other properties of the gRNA, is sufficient to allow targeting of a Cas9 molecule to the target nucleic acid.

The "secondary domain" of the targeting domain of the gRNA is complementary to the "secondary domain target" of the target nucleic acid.

In an embodiment, the secondary domain is positioned 5' to the core domain.

In an embodiment, the secondary domain is absent or optional.

In an embodiment, if the targeting domain is 26 nucleotides in length and the core domain (counted from the 3' end of the targeting domain) is 8 to 13 nucleotides in length, the secondary domain is 12 to 17 nucleotides in length.

In an embodiment, if the targeting domain is 25 nucleotides in length and the core domain (counted from the 3' end of the targeting domain) is 8 to 13 nucleotides in length, the secondary domain is 12 to 17 nucleotides in length.

In an embodiment, if the targeting domain is 24 nucleotides in length and the core domain (counted from the 3' end of the targeting domain) is 8 to 13 nucleotides in length, the secondary domain is 11 to 16 nucleotides in length.

In an embodiment, if the targeting domain is 23 nucleotides in length and the core domain (counted from the 3' end of the targeting domain) is 8 to 13 nucleotides in length, the secondary domain is 10 to 15 nucleotides in length.

In an embodiment, if the targeting domain is 22 nucleotides in length and the core domain (counted from the 3' end of the targeting domain) is 8 to 13 nucleotides in length, the secondary domain is 9 to 14 nucleotides in length.

In an embodiment, if the targeting domain is 21 nucleotides in length and the core domain (counted from the 3' end of the targeting domain) is 8 to 13 nucleotides in length, the secondary domain is 8 to 13 nucleotides in length.

In an embodiment, if the targeting domain is 20 nucleotides in length and the core domain (counted from the 3' end of the targeting domain) is 8 to 13 nucleotides in length, the secondary domain is 7 to 12 nucleotides in length.

In an embodiment, if the targeting domain is 19 nucleotides in length and the core domain (counted from the 3' end of the targeting domain) is 8 to 13 nucleotides in length, the secondary domain is 6 to 11 nucleotides in length.

In an embodiment, if the targeting domain is 18 nucleotides in length and the core domain (counted from the 3' end of the targeting domain) is 8 to 13 nucleotides in length, the secondary domain is 5 to 10 nucleotides in length.

In an embodiment, if the targeting domain is 17 nucleotides in length and the core domain (counted from the 3' end of the targeting domain) is 8 to 13 nucleotides in length, the secondary domain is 4 to 9 nucleotides in length.

In an embodiment, if the targeting domain is 16 nucleotides in length and the core domain (counted from the 3' end of the targeting domain) is 8 to 13 nucleotides in length, the secondary domain is 3 to 8 nucleotides in length.

In an embodiment, the secondary domain is 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15 nucleotides in length.

The secondary domain is complementary with the secondary domain target. Typically the secondary domain has exact complementarity with the secondary domain target. In some embodiments the secondary domain can have 1, 2, 3, 4 or 5 nucleotides that are not complementary with the corresponding nucleotide of the secondary domain. In an embodiment, the degree of complementarity, together with other properties of the gRNA, is sufficient to allow targeting of a Cas9 molecule to the target nucleic acid.

In an embodiment, the core domain nucleotides do not comprise modifications, e.g., modifications of the type provided in Section VIII. However, in an embodiment, the core domain comprise one or more modifications, e.g., modifications that it render it less susceptible to degradation or more bio-compatible, e.g., less immunogenic. By way of example, the backbone of the core domain can be modified with a phosphorothioate, or other modification(s) from Section VIII. In an embodiment a nucleotide of the core domain can comprise a 2' modification, e.g., a 2-acetylation, e.g., a 2' methylation, or other modification(s) from Section VIII. Typically, a core domain will contain no more than 1, 2, or 3 modifications.

Modifications in the core domain can be selected to not interfere with targeting efficacy, which can be evaluated by testing a candidate modification in the system described in Section IV. gRNA's having a candidate core domain having a selected length, sequence, degree of complementarity, or degree of modification, can be evaluated in the system described at Section IV. The candidate core domain can be placed, either alone, or with one or more other candidate changes in a gRNA molecule/Cas9 molecule system known to be functional with a selected target and evaluated.

In an embodiment, the secondary domain nucleotides do not comprise modifications, e.g., modifications of the type provided in Section VIII. However, in an embodiment, the secondary domain comprises one or more modifications, e.g., modifications that it render it less susceptible to degradation or more bio-compatible, e.g., less immunogenic. By way of example, the backbone of the secondary domain can be modified with a phosphorothioate, or other

modification(s) from Section VIII. In an embodiment a nucleotide of the secondary domain can comprise a 2' modification, e.g., a 2-acetylation, e.g., a 2' methylation, or other modification(s) from Section VIII. Typically, a secondary domain will contain no more than 1, 2, or 3 modifications.

Modifications in the secondary domain can be selected to not interfere with targeting efficacy, which can be evaluated by testing a candidate modification in the system described in Section IV. gRNA's having a candidate secondary domain having a selected length, sequence, degree of complementarity, or degree of modification, can be evaluated in the system described at Section IV. The candidate secondary domain can be placed, either alone, or with one or more other candidate changes in a gRNA molecule/Cas9 molecule system known to be functional with a selected target and evaluated.

In an embodiment, (1) the degree of complementarity between the core domain and its target, and (2) the degree of complementarity between the secondary domain and its target, may differ. In an embodiment, (1) may be greater (2). In an embodiment, (1) may be less than (2). In an embodiment, (1) and (2) are the same, e.g., each may be completely complementary with its target.

In an embodiment, (1) the number of modification (e.g., modifications from Section VIII) of the nucleotides of the core domain and (2) the number of modification (e.g., modifications from Section VIII) of the nucleotides of the secondary domain, may differ. In an embodiment, (1) may be less than (2). In an embodiment, (1) may be greater than (2). In an embodiment, (1) and (2) may be the same, e.g., each may be free of modifications.

The First and Second Complementarity Domains

The first complementarity domain is complementary with the second complementarity domain.

Typically the first domain does not have exact complementarity with the second complementarity domain target. In some embodiments, the first complementarity domain can have 1, 2, 3, 4 or 5 nucleotides that are not complementary with the corresponding nucleotide of the second complementarity domain. In an embodiment, 1, 2, 3, 4, 5 or 6, e.g., 3 nucleotides, will not pair in the duplex, and, e.g., form a non-duplexed or looped-out region. In an embodiment, an unpaired, or loop-out, region, e.g., a loop-out of 3 nucleotides, is present on the

second complementarity domain. In an embodiment, the unpaired region begins 1, 2, 3, 4, 5, or 6, e.g., 4, nucleotides from the 5' end of the second complementarity domain.

In an embodiment, the degree of complementarity, together with other properties of the gRNA, is sufficient to allow targeting of a Cas9 molecule to the target nucleic acid.

In an embodiment, the first and second complementarity domains are:

independently, 6 +/-2, 7+/-2, 8+/-2, 9+/-2, 10+/-2, 11+/-2, 12+/-2, 13+/-2, 14+/-2, 15+/-2, 16+/-2, 17+/-2, 18+/-2, 19+/-2, or 20+/-2, 21+/-2, 22+/-2, 23+/-2, or 24+/-2 nucleotides in length;

independently, 6, 7, 8, 9, 10, 11, 12, 13, 14, 14, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, or 26, nucleotides in length; or

independently, 5 to 24, 5 to 23, 5 to 22, 5 to 21, 5 to 20, 7 to 18, 9 to 16, or 10 to 14 nucleotides in length.

In an embodiment, the second complementarity domain is longer than the first complementarity domain, e.g., 2, 3, 4, 5, or 6, e.g., 6, nucleotides longer.

In an embodiment, the first and second complementary domains, independently, do not comprise modifications, e.g., modifications of the type provided in Section VIII.

In an embodiment, the first and second complementary domains, independently, comprise one or more modifications, e.g., modifications that the render the domain less susceptible to degradation or more bio-compatible, e.g., less immunogenic. By way of example, the backbone of the domain can be modified with a phosphorothioate, or other modification(s) from Section VIII. In an embodiment a nucleotide of the domain can comprise a 2' modification, e.g., a 2-acetylation, e.g., a 2' methylation, or other modification(s) from Section VIII.

In an embodiment, the first and second complementary domains, independently, include 1, 2, 3, 4, 5, 6, 7 or 8 or more modifications. In an embodiment, the first and second complementary domains, independently, include 1, 2, 3, or 4 modifications within 5 nucleotides of its 5' end. In an embodiment, the first and second complementary domains, independently, include as many as 1, 2, 3, or 4 modifications within 5 nucleotides of its 3' end.

In an embodiment, the first and second complementary domains, independently, include modifications at two consecutive nucleotides, e.g., two consecutive nucleotides that are within 5 nucleotides of the 5' end of the domain, within 5 nucleotides of the 3' end of the domain, or

more than 5 nucleotides away from one or both ends of the domain. In an embodiment, the first and second complementary domains, independently, include no two consecutive nucleotides that are modified, within 5 nucleotides of the 5' end of the domain, within 5 nucleotides of the 3' end of the domain, or within a region that is more than 5 nucleotides away from one or both ends of the domain. In an embodiment, the first and second complementary domains, independently, include no nucleotide that is modified within 5 nucleotides of the 5' end of the domain, within 5 nucleotides of the 3' end of the domain, or within a region that is more than 5 nucleotides away from one or both ends of the domain.

Modifications in a complementarity domain can be selected to not interfere with targeting efficacy, which can be evaluated by testing a candidate modification in the system described in Section IV. gRNA's having a candidate complementarity domain having a selected length, sequence, degree of complementarity, or degree of modification, can be evaluated in the system described in Section IV. The candidate complementarity domain can be placed, either alone, or with one or more other candidate changes in a gRNA molecule/Cas9 molecule system known to be functional with a selected target and evaluated.

In an embodiment, the first complementarity domain has at least 60, 70, 80, 85%, 90% or 95% homology with, or differs by no more than 1, 2, 3, 4, 5, or 6 nucleotides from, a reference first complementarity domain, e.g., a naturally occurring, e.g., an *S. pyogenes*, *S. aureus*, or *S. thermophilus,* first complementarity domain, or a first complementarity domain described herein, e.g., from **Figs. 1A-1G**.

In an embodiment, the second complementarity domain has at least 60, 70, 80, 85%, 90%, or 95% homology with, or differs by no more than 1, 2, 3, 4, 5, or 6 nucleotides from, a reference second complementarity domain, e.g., a naturally occurring, e.g., an *S. pyogenes*, *S. aureus*, or *S. thermophilus,* second complementarity domain, or a second complementarity domain described herein, e.g., from **Fig. 1A-1G**.

The duplexed region formed by first and second complementarity domains is typically 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22 base pairs in length (excluding any looped out or unpaired nucleotides).

In some embodiments, the first and second complementarity domains, when duplexed, comprise 11 paired nucleotides, for example, in the gRNA sequence (one paired strand underlined, one bolded):

NNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAGAAA**UAGC**AAG**UUAAAA**
UAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 5).

In some embodiments, the first and second complementarity domains, when duplexed, comprise 15 paired nucleotides, for example in the gRNA sequence (one paired strand underlined, one bolded):

NNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAUGCUGAAA**AGCAUAGC**A
AG**UUAAAAU**AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUG
C (SEQ ID NO: 27).

In some embodiments the first and second complementarity domains, when duplexed, comprise 16 paired nucleotides, for example in the gRNA sequence (one paired strand underlined, one bolded):

NNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAUGCUGGAAA**CAGCAUAG**
CAAG**UUAAAAU**AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGG
UGC (SEQ ID NO: 28).

In some embodiments the first and second complementarity domains, when duplexed, comprise 21 paired nucleotides, for example in the gRNA sequence (one paired strand underlined, one bolded):

NNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAUGCUGUUUUGGAAA**CAAA**
ACAGCAUAGCAAG**UUAAAAU**AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCA
CCGAGUCGGUGC (SEQ ID NO: 29).

In some embodiments, nucleotides are exchanged to remove poly-U tracts, for example in the gRNA sequences (exchanged nucleotides underlined):

NNNNNNNNNNNNNNNNNNNNGU**A**UUAGAGCUAGAAAUAGCAAGUUAA**U**AU
AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 30);

NNNNNNNNNNNNNNNNNNNNGUUU**A**AGAGCUAGAAAUAGCAAGUU**U**AAAU
AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 31); and

NNNNNNNNNNNNNNNNNNNNGUAUUAGAGCUAUGCUGUAUUGGAAACAAUACAGCAUAGCAAGUUAAUAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 32).

The 5' Extension Domain

In an embodiment, a modular gRNA can comprise additional sequence, 5' to the second complementarity domain. In an embodiment, the 5' extension domain is 2 to 10, 2 to 9, 2 to 8, 2 to 7, 2 to 6, 2 to 5, or 2 to 4 nucleotides in length. In an embodiment, the 5' extension domain is 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more nucleotides in length.

In an embodiment, the 5' extension domain nucleotides do not comprise modifications, e.g., modifications of the type provided in Section VIII. However, in an embodiment, the 5' extension domain comprises one or more modifications, e.g., modifications that it render it less susceptible to degradation or more bio-compatible, e.g., less immunogenic. By way of example, the backbone of the 5' extension domain can be modified with a phosphorothioate, or other modification(s) from Section VIII. In an embodiment, a nucleotide of the 5' extension domain can comprise a 2' modification, e.g., a 2-acetylation, e.g., a 2' methylation, or other modification(s) from Section VIII.

In some embodiments, the 5' extension domain can comprise as many as 1, 2, 3, 4, 5, 6, 7 or 8 modifications. In an embodiment, the 5' extension domain comprises as many as 1, 2, 3, or 4 modifications within 5 nucleotides of its 5' end, e.g., in a modular gRNA molecule. In an embodiment, the 5' extension domain comprises as many as 1, 2, 3, or 4 modifications within 5 nucleotides of its 3' end, e.g., in a modular gRNA molecule.

In some embodiments, the 5' extension domain comprises modifications at two consecutive nucleotides, e.g., two consecutive nucleotides that are within 5 nucleotides of the 5' end of the 5' extension domain, within 5 nucleotides of the 3' end of the 5' extension domain, or more than 5 nucleotides away from one or both ends of the 5' extension domain. In an embodiment, no two consecutive nucleotides are modified within 5 nucleotides of the 5' end of the 5' extension domain, within 5 nucleotides of the 3' end of the 5' extension domain, or within a region that is more than 5 nucleotides away from one or both ends of the 5' extension domain. In an embodiment, no nucleotide is modified within 5 nucleotides of the 5' end of the 5'

extension domain, within 5 nucleotides of the 3' end of the 5' extension domain, or within a region that is more than 5 nucleotides away from one or both ends of the 5' extension domain.

Modifications in the 5' extension domain can be selected to not interfere with gRNA molecule efficacy, which can be evaluated by testing a candidate modification in the system described in Section IV. gRNA's having a candidate 5' extension domain having a selected length, sequence, degree of complementarity, or degree of modification, can be evaluated in the system described at Section IV. The candidate 5' extension domain can be placed, either alone, or with one or more other candidate changes in a gRNA molecule/Cas9 molecule system known to be functional with a selected target and evaluated.

In an embodiment, the 5' extension domain has at least 60, 70, 80, 85, 90 or 95% homology with, or differs by no more than 1, 2, 3, 4, 5, or 6 nucleotides from, a reference 5' extension domain, e.g., a naturally occurring, e.g., an *S. pyogenes*, *S. aureus,* or *S. thermophilus,* 5' extension domain, or a 5' extension domain described herein, e.g., from **Figs. 1A-1G**.

The Linking Domain

In a unimolecular gRNA molecule the linking domain is disposed between the first and second complementarity domains. In a modular gRNA molecule, the two molecules are associated with one another by the complementarity domains.

In an embodiment, the linking domain is 10 +/-5, 20+/-5, 30+/-5, 40+/-5, 50+/-5, 60+/-5, 70+/-5, 80+/-5, 90+/-5, or 100+/-5 nucleotides, in length.

In an embodiment, the linking domain is 20+/-10, 30+/-10, 40+/-10, 50+/-10, 60+/-10, 70+/-10, 80+/-10, 90+/-10, or 100+/-10 nucleotides, in length.

In an embodiment, the linking domain is 10 to 100, 10 to 90, 10 to 80, 10 to 70, 10 to 60, 10 to 50, 10 to 40, 10 to 30, 10 to 20 or 10 to 15 nucleotides in length. In other embodiments, the linking domain is 20 to 100, 20 to 90, 20 to 80, 20 to 70, 20 to 60, 20 to 50, 20 to 40, 20 to 30, or 20 to 25 nucleotides in length.

In an embodiment, the linking domain is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 17, 18, 19, or 20 nucleotides in length.

In an embodiment, the linking domain is a covalent bond.

In an embodiment, the linking domain comprises a duplexed region, typically adjacent to or within 1, 2, or 3 nucleotides of the 3' end of the first complementarity domain and/or the 5-

end of the second complementarity domain. In an embodiment, the duplexed region can be 20+/-10 base pairs in length. In an embodiment, the duplexed region can be 10+/-5, 15+/-5, 20+/-5, or 30+/-5 base pairs in length. In an embodiment, the duplexed region can be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 base pairs in length.

Typically the sequences forming the duplexed region have exact complementarity with one another, though in some embodiments as many as 1, 2, 3, 4, 5, 6, 7 or 8 nucleotides are not complementary with the corresponding nucleotides.

In an embodiment, the linking domain nucleotides do not comprise modifications, e.g., modifications of the type provided in Section VIII. However, in an embodiment, the linking domain comprises one or more modifications, e.g., modifications that it render it less susceptible to degradation or more bio-compatible, e.g., less immunogenic. By way of example, the backbone of the linking domain can be modified with a phosphorothioate, or other modification(s) from Section VIII. In an embodiment a nucleotide of the linking domain can comprise a 2' modification, e.g., a 2-acetylation, e.g., a 2' methylation, or other modification(s) from Section VIII. In some embodiments, the linking domain can comprise as many as 1, 2, 3, 4, 5, 6, 7 or 8 modifications.

Modifications in a linking domain can be selected to not interfere with targeting efficacy, which can be evaluated by testing a candidate modification in the system described in Section IV. gRNA's having a candidate linking domain having a selected length, sequence, degree of complementarity, or degree of modification, can be evaluated a system described in Section IV. A candidate linking domain can be placed, either alone, or with one or more other candidate changes in a gRNA molecule/Cas9 molecule system known to be functional with a selected target and evaluated.

In an embodiment, the linking domain has at least 60, 70, 80, 85, 90 or 95% homology with, or differs by no more than 1, 2, 3, 4, 5 ,or 6 nucleotides from, a reference linking domain, e.g., a linking domain described herein, e.g., from **Figs. 1A-1G**.

The Proximal Domain

In an embodiment, the proximal domain is 6 +/-2, 7+/-2, 8+/-2, 9+/-2, 10+/-2, 11+/-2, 12+/-2, 13+/-2, 14+/-2, 14+/-2, 16+/-2, 17+/-2, 18+/-2, 19+/-2, or 20+/-2 nucleotides in length.

In an embodiment, the proximal domain is 6, 7, 8, 9, 10, 11, 12, 13, 14, 14, 16, 17, 18, 19, or 20 nucleotides in length.

In an embodiment, the proximal domain is 5 to 20, 7, to 18, 9 to 16, or 10 to 14 nucleotides in length.

In an embodiment, the proximal domain nucleotides do not comprise modifications, e.g., modifications of the type provided in Section VIII. However, in an embodiment, the proximal domain comprises one or more modifications, e.g., modifications that it render it less susceptible to degradation or more bio-compatible, e.g., less immunogenic. By way of example, the backbone of the proximal domain can be modified with a phosphorothioate, or other modification(s) from Section VIII. In an embodiment a nucleotide of the proximal domain can comprise a 2' modification, e.g., a 2-acetylation, e.g., a 2' methylation, or other modification(s) from Section VIII.

In some embodiments, the proximal domain can comprise as many as 1, 2, 3, 4, 5, 6, 7 or 8 modifications. In an embodiment, the proximal domain comprises as many as 1, 2, 3, or 4 modifications within 5 nucleotides of its 5' end, e.g., in a modular gRNA molecule. In an embodiment, the target domain comprises as many as 1, 2, 3, or 4 modifications within 5 nucleotides of its 3' end, e.g., in a modular gRNA molecule.

In some embodiments, the proximal domain comprises modifications at two consecutive nucleotides, e.g., two consecutive nucleotides that are within 5 nucleotides of the 5' end of the proximal domain, within 5 nucleotides of the 3' end of the proximal domain, or more than 5 nucleotides away from one or both ends of the proximal domain. In an embodiment, no two consecutive nucleotides are modified within 5 nucleotides of the 5' end of the proximal domain, within 5 nucleotides of the 3' end of the proximal domain, or within a region that is more than 5 nucleotides away from one or both ends of the proximal domain. In an embodiment, no nucleotide is modified within 5 nucleotides of the 5' end of the proximal domain, within 5 nucleotides of the 3' end of the proximal domain, or within a region that is more than 5 nucleotides away from one or both ends of the proximal domain.

Modifications in the proximal domain can be selected to not interfere with gRNA molecule efficacy, which can be evaluated by testing a candidate modification in the system described in Section IV. gRNA's having a candidate proximal domain having a selected length, sequence, degree of complementarity, or degree of modification, can be evaluated in the system

described at Section IV. The candidate proximal domain can be placed, either alone, or with one or more other candidate changes in a gRNA molecule/Cas9 molecule system known to be functional with a selected target and evaluated.

In an embodiment, the proximal domain has at least 60, 70, 80, 85 90 or 95% homology with, or differs by no more than 1, 2, 3, 4, 5, or 6 nucleotides from, a reference proximal domain, e.g., a naturally occurring, e.g., an *S. pyogenes*, *S. aureus*, or *S. thermophilus,* proximal domain, or a proximal domain described herein, e.g., from **Figs. 1A-1G**.

The Tail Domain

In an embodiment, the tail domain is 10 +/-5, 20+/-5, 30+/-5, 40+/-5, 50+/-5, 60+/-5, 70+/-5, 80+/-5, 90+/-5, or 100+/-5 nucleotides, in length.

In an embodiment, the tail domain is 20+/-5 nucleotides in length.

In an embodiment, the tail domain is 20+/-10, 30+/-10, 40+/-10, 50+/-10, 60+/-10, 70+/-10, 80+/-10, 90+/-10, or 100+/-10 nucleotides, in length.

In an embodiment, the tail domain is 25+/-10 nucleotides in length.

In an embodiment, the tail domain is 10 to 100, 10 to 90, 10 to 80, 10 to 70, 10 to 60, 10 to 50, 10 to 40, 10 to 30, 10 to 20 or 10 to 15 nucleotides in length.

In other embodiments, the tail domain is 20 to 100, 20 to 90, 20 to 80, 20 to 70, 20 to 60, 20 to 50, 20 to 40, 20 to 30, or 20 to 25 nucleotides in length.

In an embodiment, the tail domain is 1 to 20, 1 to 1, 1 to 10, or 1 to 5 nucleotides in length.

In an embodiment, the tail domain nucleotides do not comprise modifications, e.g., modifications of the type provided in Section VIII. However, in an embodiment, the tail domain comprises one or more modifications, e.g., modifications that it render it less susceptible to degradation or more bio-compatible, e.g., less immunogenic. By way of example, the backbone of the tail domain can be modified with a phosphorothioate, or other modification(s) from Section VIII. In an embodiment a nucleotide of the tail domain can comprise a 2' modification, e.g., a 2-acetylation, e.g., a 2' methylation, or other modification(s) from Section VIII.

In some embodiments, the tail domain can have as many as 1, 2, 3, 4, 5, 6, 7 or 8 modifications. In an embodiment, the target domain comprises as many as 1, 2, 3, or 4

modifications within 5 nucleotides of its 5' end. In an embodiment, the target domain comprises as many as 1, 2, 3, or 4 modifications within 5 nucleotides of its 3' end.

In an embodiment, the tail domain comprises a tail duplex domain, which can form a tail duplexed region. In an embodiment, the tail duplexed region can be 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 base pairs in length. In an embodiment, a further single stranded domain, exists 3' to the tail duplexed domain. In an embodiment, this domain is 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides in length. In an embodiment it is 4 to 6 nucleotides in length.

In an embodiment, the tail domain has at least 60, 70, 80, or 90% homology with, or differs by no more than 1, 2, 3, 4, 5 ,or 6 nucleotides from, a reference tail domain, e.g., a naturally occurring, e.g., an *S. pyogenes*, or *S. thermophilus,* tail domain, or a tail domain described herein, e.g., from **Figs. 1A-1G**.

In an embodiment, the proximal and tail domain, taken together comprise the following sequences:

AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU (SEQ ID NO: 33),

AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGGUGC (SEQ ID NO: 34),

AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGGAUC (SEQ ID NO: 35),

AAGGCUAGUCCGUUAUCAACUUGAAAAAGUG (SEQ ID NO: 36),

AAGGCUAGUCCGUUAUCA (SEQ ID NO: 37), or

AAGGCUAGUCCG (SEQ ID NO: 38).

In an embodiment, the tail domain comprises the 3' sequence UUUUUU, e.g., if a U6 promoter is used for transcription.

In an embodiment, the tail domain comprises the 3' sequence UUUU, e.g., if an H1 promoter is used for transcription.

In an embodiment, tail domain comprises variable numbers of 3' Us depending, e.g., on the termination signal of the pol-III promoter used.

In an embodiment, the tail domain comprises variable 3' sequence derived from the DNA template if a T7 promoter is used.

In an embodiment, the tail domain comprises variable 3' sequence derived from the DNA template, e.g., if in vitro transcription is used to generate the RNA molecule.

In an embodiment, the tail domain comprises variable 3' sequence derived from the DNA template, e., if a pol-II promoter is used to drive transcription.

Modifications in the tail domain can be selected to not interfere with targeting efficacy, which can be evaluated by testing a candidate modification in the system described in Section IV. gRNAs having a candidate tail domain having a selected length, sequence, degree of complementarity, or degree of modification, can be evaluated in the system described in Section IV. The candidate tail domain can be placed, either alone, or with one or more other candidate changes in a gRNA molecule/Cas9 molecule system known to be functional with a selected target and evaluated.

In some embodiments, the tail domain comprises modifications at two consecutive nucleotides, e.g., two consecutive nucleotides that are within 5 nucleotides of the 5' end of the tail domain, within 5 nucleotides of the 3' end of the tail domain, or more than 5 nucleotides away from one or both ends of the tail domain. In an embodiment, no two consecutive nucleotides are modified within 5 nucleotides of the 5' end of the tail domain, within 5 nucleotides of the 3' end of the tail domain, or within a region that is more than 5 nucleotides away from one or both ends of the tail domain. In an embodiment, no nucleotide is modified within 5 nucleotides of the 5' end of the tail domain, within 5 nucleotides of the 3' end of the tail domain, or within a region that is more than 5 nucleotides away from one or both ends of the tail domain.

In an embodiment a gRNA has the following structure:

5' [targeting domain]-[first complementarity domain]-[linking domain]-[second complementarity domain]-[proximal domain]-[tail domain]-3'

wherein, the targeting domain comprises a core domain and optionally a secondary domain, and is 10 to 50 nucleotides in length;

the first complementarity domain is 5 to 25 nucleotides in length and, In an embodiment has at least 50, 60, 70, 80, 85, 90 or 95% homology with a reference first complementarity domain disclosed herein;

the linking domain is 1 to 5 nucleotides in length;

the proximal domain is 5 to 20 nucleotides in length and, in an embodiment has at least 50, 60, 70, 80, 85, 90 or 95% homology with a reference proximal domain disclosed herein; and

the tail domain is absent or a nucleotide sequence is 1 to 50 nucleotides in length and, in an embodiment has at least 50, 60, 70, 80, 85, 90 or 95% homology with a reference tail domain disclosed herein.

Exemplary Chimeric gRNAs

In an embodiment, a unimolecular, or chimeric, gRNA comprises, preferably from 5' to 3':

a targeting domain, e.g., comprising 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, or 26 nucleotides (which is complementary to a target nucleic acid);

a first complementarity domain;

a linking domain;

a second complementarity domain (which is complementary to the first complementarity domain);

a proximal domain; and

a tail domain,

wherein,

(a) the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides;

(b) there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain; or

(c) there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the sequence from (a), (b), or (c), has at least 60, 75, 80, 85, 90, 95, or 99% homology with the corresponding sequence of a naturally occurring gRNA, or with a gRNA described herein.

In an embodiment, the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 nucleotides (e.g., 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 16 nucleotides (e.g., 16 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 16 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 17 nucleotides (e.g., 17 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 17 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 18 nucleotides (e.g., 18 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 18 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 19 nucleotides (e.g., 19 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 19 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 20 nucleotides (e.g., 20 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 20 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 21 nucleotides (e.g., 21 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 21 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 22 nucleotides (e.g., 22 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 22 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 23 nucleotides (e.g., 23 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 23 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 24 nucleotides (e.g., 24 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 24 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 25 nucleotides (e.g., 25 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 25 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 26 nucleotides (e.g., 26 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 26 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 16 nucleotides (e.g., 16 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 16 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 16 nucleotides (e.g., 16 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 16 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 16 nucleotides (e.g., 16 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 16 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 17 nucleotides (e.g., 17 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 17 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 17 nucleotides (e.g., 17 consecutive nucleotides) having complementarity with the target domain, e.g., the

targeting domain is 17 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 17 nucleotides (e.g., 17 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 17 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 18 nucleotides (e.g., 18 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 18 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 18 nucleotides (e.g., 18 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 18 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 18 nucleotides (e.g., 18 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 18 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 19 nucleotides (e.g., 19 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 19 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 19 nucleotides (e.g., 19 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 19 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 19 nucleotides (e.g., 19 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 19 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41,

46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 20 nucleotides (e.g., 20 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 20 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 20 nucleotides (e.g., 20 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 20 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 20 nucleotides (e.g., 20 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 20 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 21 nucleotides (e.g., 21 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 21 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 21 nucleotides (e.g., 21 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 21 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 21 nucleotides (e.g., 21 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 21 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 22 nucleotides (e.g., 22 consecutive nucleotides) having complementarity with the target domain, e.g., the

targeting domain is 22 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 22 nucleotides (e.g., 22 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 22 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 22 nucleotides (e.g., 22 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 22 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 23 nucleotides (e.g., 23 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 23 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 23 nucleotides (e.g., 23 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 23 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 23 nucleotides (e.g., 23 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 23 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 24 nucleotides (e.g., 24 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 24 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 24 nucleotides (e.g., 24 consecutive nucleotides) having complementarity with the target domain, e.g., the

targeting domain is 24 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 24 nucleotides (e.g., 24 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 24 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 25 nucleotides (e.g., 25 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 25 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 25 nucleotides (e.g., 25 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 25 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 25 nucleotides (e.g., 25 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 25 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 26 nucleotides (e.g., 26 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 26 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 26 nucleotides (e.g., 26 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 26 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 26 nucleotides (e.g., 26 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 26 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41,

46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the unimolecular, or chimeric, gRNA molecule (comprising a targeting domain, a first complementary domain, a linking domain, a second complementary domain, a proximal domain and, optionally, a tail domain) comprises the following sequence in which the targeting domain is depicted as 20 Ns but could be any sequence and range in length from 16 to 26 nucleotides and in which the gRNA sequence is followed by 6 Us, which serve as a termination signal for the U6 promoter, but which could be either absent or fewer in number: NNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGG CUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 480). In an embodiment, the unimolecular, or chimeric, gRNA molecule is a *S. pyogenes* gRNA molecule.

In some embodiments, the unimolecular, or chimeric, gRNA molecule (comprising a targeting domain, a first complementary domain, a linking domain, a second complementary domain, a proximal domain and, optionally, a tail domain) comprises the following sequence in which the targeting domain is depicted as 20 Ns but could be any sequence and range in length from 16 to 26 nucleotides and in which the gRNA sequence is followed by 6 Us, which serve as a termination signal for the U6 promoter, but which could be either absent or fewer in number: NNNNNNNNNNNNNNNNNNNGUUUUAGUACUCUGGAAACAGAAUCUACUAAAAC AAGGCAAAAUGCCGUGUUUAUCUCGUCAACUUGUUGGCGAGAUUUUUU (SEQ ID NO: 481). In an embodiment, the unimolecular, or chimeric, gRNA molecule is a *S. aureus* gRNA molecule.

The sequences and structures of exemplary chimeric gRNAs are also shown in **Figs. 18A-18B**.

Exemplary Modular gRNAs

In an embodiment, a modular gRNA comprises:

a first strand comprising, preferably from 5' to 3':

a targeting domain, e.g., comprising 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, or 26 nucleotides;

a first complementarity domain; and

a second strand, comprising, preferably from 5' to 3':

optionally a 5' extension domain;

a second complementarity domain;

a proximal domain; and

a tail domain,

wherein:

(a) the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides;

(b) there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain; or

(c) there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the sequence from (a), (b), or (c), has at least 60, 75, 80, 85, 90, 95, or 99% homology with the corresponding sequence of a naturally occurring gRNA, or with a gRNA described herein.

In an embodiment, the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 nucleotides (e.g., 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or 26 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 16 nucleotides (e.g., 16 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 16 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 17 nucleotides (e.g., 17 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 17 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 18 nucleotides (e.g., 18 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 18 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 19 nucleotides (e.g., 19 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 19 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 20 nucleotides (e.g., 20 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 20 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 21 nucleotides (e.g., 21 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 21 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 22 nucleotides (e.g., 22 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 22 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 23 nucleotides (e.g., 23 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 23 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 24 nucleotides (e.g., 24 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 24 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 25 nucleotides (e.g., 25 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 25 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 26 nucleotides (e.g., 26 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 26 nucleotides in length.

In an embodiment, the targeting domain comprises, has, or consists of, 16 nucleotides (e.g., 16 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 16 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 16 nucleotides (e.g., 16 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 16 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 16 nucleotides (e.g., 16 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 16 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 17 nucleotides (e.g., 17 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 17 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 17 nucleotides (e.g., 17 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 17 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 17 nucleotides (e.g., 17 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 17 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 18 nucleotides (e.g., 18 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 18 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 18 nucleotides (e.g., 18 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 18 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 18 nucleotides (e.g., 18 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 18 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 19 nucleotides (e.g., 19 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 19 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 19 nucleotides (e.g., 19 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 19 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 19 nucleotides (e.g., 19 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 19 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 20 nucleotides (e.g., 20 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 20 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 20 nucleotides (e.g., 20 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 20 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 20 nucleotides (e.g., 20 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 20 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 21 nucleotides (e.g., 21 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 21 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 21 nucleotides (e.g., 21 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 21 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 21 nucleotides (e.g., 21 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 21 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 22 nucleotides (e.g., 22 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 22 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 22 nucleotides (e.g., 22 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 22 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 22 nucleotides (e.g., 22 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 22 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 23 nucleotides (e.g., 23 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 23 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 23 nucleotides (e.g., 23 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 23 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 23 nucleotides (e.g., 23 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 23 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 24 nucleotides (e.g., 24 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 24 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 24 nucleotides (e.g., 24 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 24 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 24 nucleotides (e.g., 24 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 24 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 25 nucleotides (e.g., 25 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 25 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 25 nucleotides (e.g., 25 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 25 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 25 nucleotides (e.g., 25 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 25 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 26 nucleotides (e.g., 26 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 26 nucleotides in length; and the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides.

In an embodiment, the targeting domain comprises, has, or consists of, 26 nucleotides (e.g., 26 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 26 nucleotides in length; and there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain.

In an embodiment, the targeting domain comprises, has, or consists of, 26 nucleotides (e.g., 26 consecutive nucleotides) having complementarity with the target domain, e.g., the targeting domain is 26 nucleotides in length; and there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain.

II. Methods for Designing gRNAs

Methods for designing gRNAs are described herein, including methods for selecting, designing and validating target domains. Exemplary targeting domains are also provided herein. Targeting Domains discussed herein can be incorporated into the gRNAs described herein.

Methods for selection and validation of target sequences as well as off-target analyses are described, e.g., in Mali et al., 2013 Science 339(6121): 823-826; Hsu et al. Nat Biotechnol, 31(9): 827-32; Fu et al., 2014 Nat Biotechnol, doi: 10.1038/nbt.2808. PubMed PMID: 24463574; Heigwer et al., 2014 Nat Methods 11(2):122-3. doi: 10.1038/nmeth.2812. PubMed PMID:

24481216; Bae et al., 2014 Bioinformatics PubMed PMID: 24463181; Xiao A et al., 2014 Bioinformatics PubMed PMID: 24389662.

For example, a software tool can be used to optimize the choice of gRNA within a user's target sequence, e.g., to minimize total off-target activity across the genome. Off target activity may be other than cleavage. For each possible gRNA choice using *S. pyogenes* Cas9, software tools can identify all potential off-target sequences (preceding either NAG or NGG PAMs) across the genome that contain up to a certain number (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10) of mismatched base-pairs. The cleavage efficiency at each off-target sequence can be predicted, e.g., using an experimentally-derived weighting scheme. Each possible gRNA can then ranked according to its total predicted off-target cleavage; the top-ranked gRNAs represent those that are likely to have the greatest on-target and the least off-target cleavage. Other functions, e.g., automated reagent design for gRNA vector construction, primer design for the on-target Surveyor assay, and primer design for high-throughput detection and quantification of off-target cleavage via next-generation sequencing, can also be included in the tool. Candidate gRNA molecules can be evaluated by art-known methods or as described in Section IV herein.

Guide RNAs (gRNAs) for use with *S. pyogenes, S. aureus and N. meningitidis* Cas9s were identified using a DNA sequence searching algorithm. Guide RNA design was carried out using a custom guide RNA design software based on the public tool cas-offinder (Bae et al. Bioinformatics. 2014; 30(10): 1473-1475). Said custom guide RNA design software scores guides after calculating their genomewide off-target propensity. Typically matches ranging from perfect matches to 7 mismatches are considered for guides ranging in length from 17 to 24. Once the off-target sites are computationally determined, an aggregate score is calculated for each guide and summarized in a tabular output using a web-interface. In addition to identifying potential gRNA sites adjacent to PAM sequences, the software also identifies all PAM adjacent sequences that differ by 1, 2, 3 or more nucleotides from the selected gRNA sites. Genomic DNA sequence for each gene was obtained from the UCSC Genome browser and sequences were screened for repeat elements using the publically available RepeatMasker program. RepeatMasker searches input DNA sequences for repeated elements and regions of low complexity. The output is a detailed annotation of the repeats present in a given query sequence.

Following identification, gRNAs were ranked into tiers based on their distance to the target site, their orthogonality and presence of a 5' G (based on identification of close matches in

the human genome containing a relavant PAM, e.g., in the case of *S. pyogenes*, a NGG PAM, in the case of *S. aureus*, NNGRR (e.g, a NNGRRT or NNGRRV) PAM, and in the case of *N. meningitides*, a NNNNGATT or NNNNGCTT PAM. Orthogonality refers to the number of sequences in the human genome that contain a minimum number of mismatches to the target sequence. A "high level of orthogonality" or "good orthogonality" may, for example, refer to 20-mer gRNAs that have no identical sequences in the human genome besides the intended target, nor any sequences that contain one or two mismatches in the target sequence. Targeting domains with good orthogonality are selected to minimize off-target DNA cleavage.

As an emxaple, for *S. pyogenes* and *N. meningitides* targets, 17-mer, or 20-mer gRNAs were designed. As another example, for *S. aureus* targets, 18-mer, 19-mer, 20-mer, 21-mer, 22-mer, 23-mer and 24-mer gRNAs were designed. Tarteting domains, disclosed herein, may comprises the 17-mer described in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**, e.g., the targeting domains of 18 or more nucleotides may comprise the 17-mer gRNAs described in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**. Tarteting domains, disclosed herein, may comprises the 18-mer described in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**, e.g., the targeting domains of 19 or more nucleotides may comprise the 18-mer gRNAs described in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**. Tarteting domains, disclosed herein, may comprises the 19-mer described in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**, e.g., the targeting domains of 20 or more nucleotides may comprise the 19-mer gRNAs described in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**. Tarteting domains, disclosed herein, may comprises the 20-mer gRNAs described in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**, e.g., the targeting domains of 21 or more nucleotides may comprise the 20-mer gRNAs described in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E,**

Tables 9A-9B, or **Table 10**. Tarteting domains, disclosed herein, may comprises the 21-mer described in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**, e.g., the targeting domains of 22 or more nucleotides may comprise the 21-mer gRNAs described in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**. Tarteting domains, disclosed herein, may comprises the 22-mer described in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**, e.g., the targeting domains of 23 or more nucleotides may comprise the 22-mer gRNAs described in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**. Tarteting domains, disclosed herein, may comprises the 23-mer described in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**, e.g., the targeting domains of 24 or more nucleotides may comprise the 23-mer gRNAs described in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**. Tarteting domains, disclosed herein, may comprises the 24-mer described in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**, e.g., the targeting domains of 25 or more nucleotides may comprise the 24-mer gRNAs described in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**.

gRNAs were identified for both single-gRNA nuclease cleavage and for a dual-gRNA paired "nickase" strategy. Criteria for selecting gRNAs and the determination for which gRNAs can be used for the dual-gRNA paired "nickase" strategy is based on two considerations:

1. gRNA pairs should be oriented on the DNA such that PAMs are facing out and cutting with the D10A Cas9 nickase will result in 5' overhangs.
2. An assumption that cleaving with dual nickase pairs will result in deletion of the entire intervening sequence at a reasonable frequency. However, cleaving with dual nickase pairs can also result in indel mutations at the site of only one of the gRNAs.

Candidate pair members can be tested for how efficiently they remove the entire sequence versus causing indel mutations at the site of one gRNA.

The Targeting Domains discussed herein can be incorporated into the gRNAs described herein.

Three strategies were utilized to identify gRNAs for use with *S. pyogenes, S. aureus and N. meningitidis* Cas9 enzymes.

In one strategy, gRNAs were designed for use with *S. pyogenes* and *S. aureus* Cas9 enzymes to induce an indel mediated by NHEJ in close proximity to or including the LCA10 target position (e.g., c.2991+1655A to G). The gRNAs were identified and ranked into 4 tiers for *S. pyogenes* (**Tables 1A-1D**). The targeting domain for tier 1 gRNA molecules to be used with *S. pyogenes* Cas9 molecules were selected based on (1) a short distance to the target position, e.g., within 40bp upstream and 40bp downstream of the mutation, (2) a high level of orthogonality, and (3) the presence of a 5' G. For selection of tier 2 gRNAs, a short distance and high orthogonality were required but the presence of a 5'G was not required. Tier 3 uses the same distance restriction and the requirement for a 5'G, but removes the requirement of good orthogonality. Tier 4 uses the same distance restriction but removes the requirement of good orthogonality and the 5'G. The gRNAs were identified and ranked into 4 tiers for *S. aureus,* when the relavent PAM was NNGRR (**Tables 2A-2C**). The targeting domain for tier 1 gRNA molecules to be used with *S. pyogenes* Cas9 molecules were selected based on (1) a short distance to the target position, e.g., within 40 bp upstream and 40 bp downstream of the mutation, (2) a high level of orthogonality, and (3) the presence of a 5' G. For selection of tier 2 gRNAs, a short distance and high orthogonality were required but the presence of a 5'G was not required. Tier 3 uses the same distance restriction and the requirement for a 5'G, but removes the requirement of good orthogonality. Tier 4 uses the same distance restriction but removes the requirement of good orthogonality and the 5'G. The gRNAs were identified and ranked into 5 tiers for *S. aureus* when the relavent PAM was NNGRRT or NNGRRV (**Tables 6A-6D**). The targeting domain for tier 1 gRNA molecules to be used with *S. aureus* Cas9 molecules were selected based on (1) a short distance to the target position, e.g., within 40 bp upstream and 40 bp downstream of the mutation, (2) a high level of orthogonality, (3) the presence of a 5' G and (4) PAM was NNGRRT. For selection of tier 2 gRNAs, a short distance and high orthogonality were required but the presence of a 5'G was not required, and PAM was NNGRRT. Tier 3 uses

the same distance restriction and the requirement for a 5'G, but removes the requirement of good orthogonality, and PAM was NNGRRT. Tier 4 uses the same distance restriction but removes the requirement of good orthogonality and the 5'G, ang PAM was NNGRRT. Tier 5 required a short distance to the target position, e.g., within 40 bp upstream and 40 bp downstream of the mutation and PAM was NNGRRV. Note that tiers are non-inclusive (each gRNA is listed only once for the strategy). In certain instances, no gRNA was identified based on the criteria of the particular tier.

In a second strategy, gRNAs were designed for use with *S. pyogenes, S. aureus* and *N. meningitidis* Cas9 molecules to delete a genomic sequence including the mutation at the LCA10 target position (e.g., c.2991+1655A to G), e.g., mediated by NHEJ. The gRNAs were identified and ranked into 4 tiers for *S. pyogenes* (**Tables 3A-3D**). The targeting domain to be used with *S. pyogenes* Cas9 molecules for tier 1 gRNA molecules were selected based on (1) flanking the mutation without targeting unwanted chromosome elements, such as an *Alu* repeat, e.g., within 400bp upstream of an *Alu* repeat or 700 bp downstream of mutation, (2) a high level of orthogonality, and (3) the presence of a 5' G. For selection of tier 2 gRNAs, a reasonable distance and high orthogonality were required but the presence of a 5'G was not required. Tier 3 uses the same distance restriction and the requirement for a 5'G, but removes the requirement of good orthogonality. Tier 4 uses the same distance restriction but removes the requirement of good orthogonality and the 5'G. The gRNAs were identified and ranked into 4 tiers for *S. aureus,* when the relavent PAM was NNGRR (**Tables 4A-4D**). The targeting domain to be used with *S. aureus* Cas9 molecules for tier 1 gRNA molecules were selected based on (1) flanking the mutation without targeting unwanted chromosome elements, such as an *Alu* repeat, e.g., within 400 bp upstream of an *Alu* repeat or 700 bp downstream of mutation, (2) a high level of orthogonality, and (3) the presence of a 5' G. For selection of tier 2 gRNAs, a reasonable distance and high orthogonality were required but the presence of a 5'G was not required. Tier 3 uses the same distance restriction and the requirement for a 5'G, but removes the requirement of good orthogonality. Tier 4 uses the same distance restriction but removes the requirement of good orthogonality and the 5'G. The gRNAs were identified and ranked into 2 tiers for *N. meningitides* (**Tables 5A-5B**). The targeting domain to be used with *N. meningitides* Cas9 molecules for tier 1 gRNA molecules were selected based on (1) flanking the mutation without targeting unwanted chromosome elements, such as an *Alu* repeat, e.g., within 400bp upstream of

an *Alu* repeat or 700 bp downstream of mutation, (2) a high level of orthogonality, and (3) the presence of a 5' G. For selection of tier 2 gRNAs, a reasonable distance and high orthogonality were required but the presence of a 5'G was not required. Note that tiers are non-inclusive (each gRNA is listed only once for the strategy). In certain instances, no gRNA was identified based on the criteria of the particular tier.In a third strategy, gRNAs were designed for use with *S. pyogenes, S. aureus* and *N. meningitidis* Cas9 molecules to delete a genomic sequence including the mutation at the LCA10 target position (e.g., c.2991+1655A to G), e.g., mediated by NHEJ. The gRNAs were identified and ranked into 4 tiers for *S. pyogenes* (**Tables 7A-7D**). The targeting domain to be used with *S. pyogenes* Cas9 enzymes for tier 1 gRNA molecules were selected based on (1) flanking the mutation without targeting unwanted chromosome elements, such as an *Alu* repeat, e.g., within 1000 bp upstream of an *Alu* repeat or 1000 bp downstream of mutation, (2) a high level of orthogonality, (3) the presence of a 5' G and (4) and PAM was NNGRRT. For selection of tier 2 gRNAs, a reasonable distance and high orthogonality were required but the presence of a 5'G was not required, and and PAM was NNGRRT. Tier 3 uses the same distance restriction and the requirement for a 5'G, but removes the requirement of good orthogonality, and and PAM was NNGRRT. Tier 4 uses the same distance restriction but removes the requirement of good orthogonality and the 5'G, and and PAM was NNGRRT. The gRNAs were identified and ranked into 4 tiers for *S. aureus,* when the relavent PAM was NNGRRT or NNGRRV (**Tables 8A-8E**). The targeting domain to be used with *S. aureus* Cas9 enzymes for tier 1 gRNA molecules were selected based on (1) flanking the mutation without targeting unwanted chromosome elements, such as an *Alu* repeat, e.g., within 1000 bp upstream of an *Alu* repeat or 1000bp downstream of mutation, (2) a high level of orthogonality, and (3) the presence of a 5' G. For selection of tier 2 gRNAs, a reasonable distance and high orthogonality were required but the presence of a 5'G was not required. Tier 3 uses the same distance restriction and the requirement for a 5'G, but removes the requirement of good orthogonality. Tier 4 uses the same distance restriction but removes the requirement of good orthogonality and the 5'G. Tier 5 used the same distance restriction and PAM was NNGRRV. The gRNAs were identified and ranked into 2 tiers for *N. meningitides* (**Tables 9A-9B**). The targeting domain to be used with *N. meningitides* Cas9 molecules for tier 1 gRNA molecules were selected based on (1) flanking the mutation without targeting unwanted chromosome elements, such as an *Alu* repeat, e.g., within 1000bp upstream of an *Alu* repeat or 1000 bp downstream of mutation, (2) a

high level of orthogonality, and (3) the presence of a 5' G. For selection of tier 2 gRNAs, a reasonable distance and high orthogonality were required but the presence of a 5'G was not required. Note that tiers are non-inclusive (each gRNA is listed only once for the strategy). In certain instances, no gRNA was identified based on the criteria of the particular tier.

In an embodiment, when a single gRNA molecule is used to target a Cas9 nickase to create a single strand break to introduce a break-induced indel in close proximity to or including the LCA10 target position, the gRNA is used to target either upstream of (e.g., within 40 bp upstream of the LCA10 target position), or downstream of (e.g., within 40 bp downstream of the LCA10 target position) in the *CEP290* gene.

In an embodiment, when a single gRNA molecule is used to target a Cas9 nuclease to create a double strand break to introduce a break-induced indel in close proximity to or including the LCA10 target position, the gRNA is used to target either upstream of (e.g., within 40 bp upstream of the LCA10 target position), or downstream of (e.g., within 40 bp downstream of the LCA10 target position) in the *CEP290* gene.

In an embodiment, dual targeting is used to create two double strand breaks to delete a genomic sequence including the mutation at the LCA10 target position, e.g., mediated by NHEJ. In an embodiment, the first and second gRNAs are used target two Cas9 nucleases to flank, e.g., the first of gRNA is used to target upstream of (e.g., within 400 bp upstream of the *Alu* repeat, or within 40 bp upstream of the LCA10 target position), and the second gRNA is used to target downstream of (e.g., within 700 bp downstream of the LCA10 target position) in the *CEP290* gene.

In an embodiment, dual targeting is used to create a double strand break and a pair of single strand breaks to delete a genomic sequence including the mutation at the LCA10 target position, e.g., mediated by NHEJ. In an embodiment, the first, second and third gRNAs are used to target one Cas9 nuclease and two Cas9 nickases to flank, e.g., the first gRNA that will be used with the Cas9 nuclease is used to target upstream of (e.g., within 400 bp upstream of the *Alu* repeat, or within 40 bp upstream of the LCA10 target position) or downstream of (e.g., within 700 bp downstream) of the LCA10 target position, and the second and third gRNAs that will be used with the Cas9 nickase pair are used to target the opposite side of the LCA10 target position (e.g., within 400 bp upstream of the *Alu* repeat, within 40 bp upstream of the LCA10 target position, or within 700 bp downstream of the LCA10 target position) in the *CEP290* gene.

In an embodiment, when four gRNAs (e.g., two pairs) are used to target four Cas9 nickases to create four single strand breaks to delete genomic sequence including the mutation at the LCA10 target position, e.g., mediated by NHEJ, the first pair and second pair of gRNAs are used to target four Cas9 nickases to flank, e.g., the first pair of gRNAs are used to target upstream of (e.g., within 400 bp upstream of the *Alu* repeat, or within 40 bp upstream of the LCA10 target position), and the second pair of gRNAs are used to target downstream of (e.g., within 700 bp downstream of the LCA10 target position) in the *CEP290* gene.

In an embodiment, dual targeting is utilized to delete genomic sequence including the mutation at the LCA10 target position mediated by NHEJ. It is contemplated herein that in an embodiment any upstream gRNA (e.g., within 400 bp upstream of an Alu repeat, or within 40bp upstream of the LCA10 target position) in **Tables 1A-1C** and **Tables 3A-3D** can be paired with any downstream gRNA (e.g., within 700 downstream of LCA10 target position) in **Tables 3A-3D** to be used with a *S. pyogenes* Cas9 molecule to generate dual targeting. Exemplary pairs including selecting a targeting domain that is labeled as upstream from **Tables 1A-1C** or **Tables 3A-3D** and a second targeting domain that is labeled as downstream from **Tables 3A-3D**. In an embodiment, a targeting domain that is labeled as upstream in **Tables 1A-1C** or **Tables 3A-3D** can be combined with any of the targeting domains that is labeled as downstream in **Tables 3A-3D**.

In an embodiment, dual targeting is utilized to delete genomic sequence including the mutation at the LCA10 target position mediated by NHEJ. It is contemplated herein that in an embodiment any upstream gRNA (e.g., within 400 bp upstream of an *Alu* repeat, or within 40bp upstream of the LCA10 target position) in **Tables 2A-2C** and **Tables 4A-4D** can be paired with any downstream gRNA (e.g., within 700 downstream of LCA10 target position) in **Tables 4A-4D** to be used with a *S. aureus* Cas9 molecule to generate dual targeting. Exemplary pairs include selecting a targeting domain that is labeled as upstream from **Tables 2A-2C** or **Tables 4A-4D** and a second targeting domain that is labeled as downstream from **Tables 4A-4D**. In an embodiment, a targeting domain that is labeled as upstream in **Tables 2A-2C** or **Tables 4A-4D** can be combined with any of the targeting domains that is labeled as downstream in **Tables 4A-4D**.

In an embodiment, dual targeting is utilized to delete genomic sequence including the mutation at the LCA10 target position mediated by NHEJ. It is contemplated herein that in an

embodiment any upstream gRNA (e.g., within 400 bp upstream of an *Alu* repeat, or within 40 bp upstream of the LCA10 target position) in **Tables 5A-5B** can be paired with any downstream gRNA (e.g., within 700 downstream of LCA10 target position) in **Tables 5A-5B** to be used with a *N. meningitidis* Cas9 molecule to generate dual targeting. Exemplary pairs include selecting a targeting domain that is labeled as upstream from **Tables 5A-5B** and a second targeting domain that is labeled as downstream from **Tables 5A-5B**. In an embodiment, a targeting domain that is labeled as upstream in **Tables 5A-5B** can be combined with any of the targeting domains that is labeled as downstream in **Tables 5A-5B**.

In an embodiment, dual targeting (e.g., dual double strand cleavage) is used to create two double strand breaks to delete a genomic sequence including the mutation at the LCA10 target position, e.g., mediated by NHEJ. In an embodiment, the first and second gRNAs are used target two Cas9 nucleases to flank, e.g., the first of gRNA is used to target upstream of (e.g., within 1000 bp upstream of the *Alu* repeat, or within 40 bp upstream of the LCA10 target position), and the second gRNA is used to target downstream of (e.g., within 1000 bp downstream of the LCA10 target position) in the *CEP290* gene.

In an embodiment, dual targeting (e.g., dual double strand cleavage) is used to create a double strand break and a pair of single strand breaks to delete a genomic sequence including the mutation at the LCA10 target position, e.g., mediated by NHEJ. In an embodiment, the first, second and third gRNAs are used to target one Cas9 nuclease and two Cas9 nickases to flank, e.g., the first gRNA that will be used with the Cas9 nuclease is used to target upstream of (e.g., within 1000 bp upstream of the *Alu* repeat, or within 40bp upstream of the LCA10 target position) or downstream of (e.g., within 1000 bp downstream) of the LCA10 target position, and the second and third gRNAs that will be used with the Cas9 nickase pair are used to target the opposite side of the LCA10 target position (e.g., within 1000 bp upstream of the *Alu* repeat, or within 40bp upstream of the LCA10 target position or within 1000 bp downstream of the LCA10 target position) in the *CEP290* gene.

In an embodiment, when four gRNAs (e.g., two pairs) are used to target four Cas9 nickases to create four single strand breaks to delete genomic sequence including the mutation at the LCA10 target position, e.g., mediated by NHEJ, the first pair and second pair of gRNAs are used to target four Cas9 nickases to flank, e.g., the first pair of gRNAs are used to target upstream of (e.g., within 1000 bp upstream of the *Alu* repeat, or within 40bp upstream of the

LCA10 target position), and the second pair of gRNAs are used to target downstream of (e.g., within 1000 bp downstream of the LCA10 target position) in the *CEP290* gene.

In an embodiment, dual targeting is utilized to delete genomic sequence including the mutation at the LCA10 target position, e.g., mediated by NHEJ. It is contemplated herein that in an embodiment any upstream gRNA (e.g., within 1000 bp upstream of an Alu repeat, or within 40bp upstream of the LCA10 target position) in **Tables 1A-1C, Tables 3A-3D,** or **Tables 7A-7D** can be paired with any downstream gRNA (e.g., within 1000 downstream of LCA10 target position) in **Tables 1A-1C, Tables 3A-3D,** or **Tables 7A-7D** to be used with a *S. pyogenes* Cas9 molecule to generate dual targeting. Exemplary pairs including selecting a targeting domain that is labeled as upstream from **Tables 1A-1C, Tables 3A-3D,** or **Tables 7A-7D** and a second targeting domain that is labeled as downstream from **Tables 1A-1C, Tables 3A-3D,** or **Tables 7A-7D**. In an embodiment, a targeting domain that is labeled as upstream in **Tables 1A-1C, Tables 3A-3D,** or **Tables 7A-7D** can be combined with any of the targeting domains that is labeled as downstream in **Tables 1A-1C, Tables 3A-3D,** or **Tables 7A-7D**.

In an embodiment, dual targeting is utilized to delete genomic sequence including the mutation at the LCA10 target position mediated by NHEJ. It is contemplated herein that in an embodiment any upstream gRNA (e.g., within 1000 bp upstream of an *Alu* repeat, or within 40bp upstream of the LCA10 target position) in **Tables 2A-2C**, **Tables 4A-4D, Tables 6A-6D,** or **Tables 8A-8E** can be paired with any downstream gRNA (e.g., within 1000 downstream of LCA10 target position) in **Tables 2A-2C**, **Tables 4A-4D, Tables 6A-6D,** or **Tables 8A-8E** to be used with a *S. aureus* Cas9 molecule to generate dual targeting. Exemplary pairs include selecting a targeting domain that is labeled as upstream from **Tables 2A-2C**, **Tables 4A-4D, Tables 6A-6D,** or **Tables 8A-8E** and a second targeting domain that is labeled as downstream from **Tables 2A-2C**, **Tables 4A-4D, Tables 6A-6D,** or **Tables 8A-8E**. In an embodiment, a targeting domain that is labeled as upstream in **Tables 2A-2C**, **Tables 4A-4D, Tables 6A-6D,** or **Tables 8A-8E** can be combined with any of the targeting domains that is labeled as downstream in **Tables 2A-2C**, **Tables 4A-4D, Tables 6A-6D,** or **Tables 8A-8E**.

In an embodiment, dual targeting is utilized to delete genomic sequence including the mutation at the LCA10 target position, e.g., mediated by NHEJ. It is contemplated herein that in an embodiment any upstream gRNA (e.g., within 1000 bp upstream of an *Alu* repeat, or within 40bp upstream of the LCA10 target position) in **Tables 5A-5B** or **Tables 9A-9B** can be paired

with any downstream gRNA (e.g., within 1000 downstream of LCA10 target position) in **Tables 5A-5D** to be used with a *N. meningitidis* Cas9 molecule to generate dual targeting. Exemplary pairs include selecting a targeting domain that is labeled as upstream from **Tables 5A-5B** or **Tables 9A-9B** and a second targeting domain that is labeled as downstream from **Tables 5A-5B** or **Tables 9A-9B**. In an embodiment, a targeting domain that is labeled as upstream in **Tables 5A-5B** or **Tables 9A-9B** and can be combined with any of the targeting domains that is labeled as downstream in **Tables 5A-5B** or **Tables 9A-9B**.

Any of the targeting domains in the tables described herein can be used with a Cas9 nickase molecule to generate a single strand break.

Any of the targeting domains in the tables described herein can be used with a Cas9 nuclease molecule to generate a double strand break.

In an embodiment, dual targeting (e.g., dual nicking) is used to create two nicks on opposite DNA strands by using *S. pyogenes, S. aureus* and *N. meningitidis* Cas9 nickases with two targeting domains that are complementary to opposite DNA strands, e.g., a gRNA comprising any minus strand targeting domain may be paired any gRNA comprising a plus strand targeting domain provided that the two gRNAs are oriented on the DNA such that PAMs face outward and the distance between the 5' ends of the gRNAs is 0-50bp. Exemplary nickase pairs including selecting a targeting domain from Group A and a second targeting domain from Group B in **Table 1D** (for *S. pyogenes*), or selecting a targeting domain from Group A and a second targeting domain from Group B in **Table 6D** (for *S. aureus).* It is contemplated herein that in an embodiment a targeting domain of Group A can be combined with any of the targeting domains of Group B in **Table 1D** (for *S. pyogenes*). For example, CEP290-B5 or CEP290-B10 can be combined with CEP290-B1 or CEP290-B6. It is contemplated herein that in an embodiment a targeting domain of Group A can be combined with any of the targeting domains of Group B in **Table 6D** (for *S. aureus*). For example, CEP290-12 or CEP290-17 can be combined with CEP290-11 or CEP290-16.

In an embodiment, dual targeting (e.g., dual double strand cleavage) is used to create two double strand breaks by using *S. pyogenes, S. aureus* and *N. meningitidis* Cas9 nucleases with two targeting domains. It is contemplated herein that in an embodiment any upstream gRNA of any of **Tables 1A-1C, 2A-2C, 3A-3D, 4A-4D, 5A-5B, 6A-6C, 7A-7D, 8A-8E,** or **9A-9B** can be paired with any downstream gRNA of any of **Tables 1A-1C, 2A-2C, 3A-3D, 4A-4D, 5A-5B,**

6A-6C, 7A-7D, 8A-8E, or **9A-9B**. Exemplary nucleases pairs are shown in **Table 10**, e.g., CEP290-323 can be combined with CEP290-11, CEP290-323 can be combined with CEP290-64, CEP290-490 can be combined with CEP290-496, CEP290-490 can be combined with CEP290-502, CEP290-490 can be combined with CEP290-504, CEP290-492 can be combined with CEP290-502, or CEP290-492 can be combined with CEP290-504.

It is contemplated herein that any upstream gRNA described herein may be paired with any downstream gRNA described herein. When an upstream gRNA designed for use with one species of Cas9 is paired with a downstream gRNA designed for use from a different species of Cas9, both Cas9 species are used to generate a single or double-strand break, as desired.

Exemplary Targeting Domains

Table 1A provides targeting domains for NHEJ-mediated introduction of an indel in close proximity to or including the LCA10 target position in the *CEP290* gene selected according to the first tier parameters. The targeting domains are within 40 bases of the LCA10 target position, have good orthogonality, and start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. pyogenes* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 1A

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-B4	+	GAGAUACUCACAAUUACAAC	20	upstream
CEP290-B28	+	GAUACUCACAAUUACAACUG	20	upstream

Table 1B provides targeting domains for NHEJ-mediated introduction of an indel in close proximity to or including the LCA10 target position in the *CEP290* gene selected according to the second tier parameters. The targeting domains are within 40 bases of the LCA10 target position, have good orthogonality, and do not start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. pyogenes* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 1B

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-B6	-	CUCAUACCUAUCCCUAU	17	downstream
CEP290-B20	+	ACACUGCCAAUAGGGAU	17	downstream
CEP290-B10	+	CAAUUACAACUGGGGCC	17	upstream
CEP290-B21	+	CUAAGACACUGCCAAUA	17	downstream
CEP290-B9	+	AUACUCACAAUUACAAC	17	upstream
CEP290-B1	-	UAUCUCAUACCUAUCCCUAU	20	downstream
CEP290-B29	+	AAGACACUGCCAAUAGGGAU	20	downstream
CEP290-B5	+	UCACAAUUACAACUGGGGCC	20	upstream
CEP290-B30	+	AGAUACUCACAAUUACAACU	20	upstream

Table 1C provides targeting domains for NHEJ-mediated introduction of an indel in close proximity to or including the LCA10 target position in the *CEP290* gene selected according to the fourth tier parameters. The targeting domains are within 40 bases of the LCA10 target position and do not start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. pyogenes* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 1C

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-B22	+	ACUAAGACACUGCCAAU	17	downstream
CEP290-B23	+	UACUCACAAUUACAACU	17	upstream
CEP290-B24	+	ACUCACAAUUACAACUG	17	upstream
CEP290-B25	+	ACAACUGGGGCCAGGUG	17	upstream
CEP290-B26	+	ACUGGGGCCAGGUGCGG	17	upstream
CEP290-B27	-	AUGUGAGCCACCGCACC	17	upstream
CEP290-B31	+	AAACUAAGACACUGCCAAUA	20	downstream
CEP290-B32	+	AAAACUAAGACACUGCCAAU	20	upstream
CEP290-B33	+	AUUACAACUGGGGCCAGGUG	20	upstream
CEP290-B34	+	ACAACUGGGGCCAGGUGCGG	20	upstream

Table 1D provides targeting domains for NHEJ-mediated introduction of an indel in close proximity to or including the LCA10 target position in the *CEP290* gene that can be used for dual targeting. Any of the targeting domains in the table can be used with a *S. pyogenes* Cas9 (nickase) molecule to generate a single stranded break.

Exemplary nickase pairs including selecting a targeting domain from Group A and a second targeting domain from Group B. It is contemplated herein that a targeting domain of Group A can be combined with any of the targeting domains of Group B. For example, the *CEP290-B*5 or *CEP290-B*10 can be combined with *CEP290-B*1 or *CEP290-B*6.

Table 1D

Group A	Group B
CEP290-B5	CEP290-B1
CEP290-B10	CEP290-B6

Table 2A provides targeting domains for NHEJ-mediated introduction of an indel in close proximity to or including the LCA10 target position in the *CEP290* gene selected according to the first tier parameters. The targeting domains are within 40 bases of the LCA10 target position, have good orthogonality, and start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. aureus* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 2A

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-B1000	+	GAGAUACUCACAAUUACAAC	20	upstream
CEP290-B1001	+	GAUACUCACAAUUACAA	17	upstream

Table 2B provides targeting domains for NHEJ-mediated introduction of an indel in close proximity to or including the LCA10 target position in the *CEP290* gene selected according to the second tier parameters. The targeting domains are within 40 bases of the LCA10 target position, have good orthogonality, and do not start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. aureus* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 2B

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-B1002	+	CACUGCCAAUAGGGAUAGGU	20	downstream

CEP290-B1003	+	UGCCAAUAGGGAUAGGU	17	downstream
CEP290-B1004	+	UGAGAUACUCACAAUUACAA	20	upstream
CEP290-B1005	+	AUACUCACAAUUACAAC	17	upstream

Table 2C provides targeting domains for NHEJ-mediated introduction of an indel in close proximity to or including the LCA10 target position in the *CEP290* gene selected according to the fourth tier parameters. The targeting domains are within 40 bases of the LCA10 target position, and do not start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. aureus* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 2C

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-B1006	-	ACCUGGCCCCAGUUGUAAUU	20	upstream
CEP290-B1007	-	UGGCCCCAGUUGUAAUU	17	upstream

Table 3A provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the first tier parameters. The targeting domains are within 400bp upstream of an *Alu* repeat or 700bp downstream of the mutation, have good orthogonality, and start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. pyogenes* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 3A

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-B8	-	GCUACCGGUUACCUGAA	17	downstream
CEP290-B217	+	GCAGAACUAGUGUAGAC	17	downstream
CEP290-B69	-	GUUGAGUAUCUCCUGUU	17	downstream
CEP290-B115	+	GAUGCAGAACUAGUGUAGAC	20	downstream
CEP290-B187	+	GCUUGAACUCUGUGCCAAAC	20	downstream

Table 3B provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the second tier parameters. The targeting domains are within 400bp upstream of an *Alu* repeat or 700bp downstream of the mutation, have good orthogonality, and do not start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. pyogenes* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 3B

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
cep290-B269	-	AGCUACCGGUUACCUGA	17	downstream
cep290-B285	+	UUUAAGGCGGGGAGUCACAU	20	downstream
CEP290-B3	-	AAAGCUACCGGUUACCUGAA	20	downstream
cep290-B207	-	AAAAGCUACCGGUUACCUGA	20	downstream
cep290-B106	-	CUCAUACCUAUCCCUAU	17	downstream
cep290-B55	+	ACACUGCCAAUAGGGAU	17	downstream
cep290-B138	-	UAUCUCAUACCUAUCCCUAU	20	downstream
cep290-B62	-	ACGUGCUCUUUUCUAUAUAU	20	downstream
cep290-B121	+	AUUUGACACCACAUGCACUG	20	downstream
cep290-B120	-	CGUGCUCUUUUCUAUAUAUA	20	downstream
cep290-B36	-	UGGUGUCAAAUAUGGUGCUU	20	downstream
cep290-B236	+	ACUUUUACCCUUCAGGUAAC	20	downstream
cep290-B70	-	AGUGCAUGUGGUGUCAAAUA	20	downstream
cep290-B177	-	UACAUGAGAGUGAUUAGUGG	20	downstream
cep290-B451	-	CGUUGUUCUGAGUAGCUUUC	20	upstream
cep290-B452	+	CCACAAGAUGUCUCUUGCCU	20	upstream
cep290-B453	-	CCUAGGCAAGAGACAUCUUG	20	upstream
cep290-B454	+	UGCCUAGGACUUUCUAAUGC	20	upstream
cep290-B498	-	CGUUGUUCUGAGUAGCUUUC	20	upstream
cep290-B523	-	AUUAGCUCAAAAGCUUUUGC	20	upstream

Table 3C provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the third tier parameters. The targeting domains are within 400bp upstream of an *Alu* repeat or 700bp downstream of the mutation, and start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the

targeting domains in the table can be used with a *S. pyogenes* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 3C

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	
CEP290-B87	-	GCAUGUGGUGUCAAAUA	17	downstream
CEP290-B50	+	GAUGACAUGAGGUAAGU	17	downstream
CEP290-B260	+	GUCACAUGGGAGUCACA	17	downstream
CEP290-B283	-	GAGAGCCACAGUGCAUG	17	downstream
CEP290-B85	-	GCUCUUUUCUAUAUAUA	17	downstream
CEP290-B78	+	GCUUUUGACAGUUUUUA	17	downstream
CEP290-B292	+	GAUAGAGACAGGAAUAA	17	downstream
CEP290-B278	+	GGACUUGACUUUUACCCUUC	20	downstream
CEP290-B227	+	GGGAGUCACAUGGGAGUCAC	20	downstream
CEP290-B261	-	GUGGAGAGCCACAGUGCAUG	20	downstream
CEP290-B182	+	GCCUGAACAAGUUUUGAAAC	20	downstream
CEP290-B67	+	GGAGUCACAUGGGAGUCACA	20	downstream
CEP290-B216	+	GUAAGACUGGAGAUAGAGAC	20	downstream
CEP290-B241	+	GCUUUUGACAGUUUUUAAGG	20	downstream
CEP290-B161	+	GUUUAGAAUGAUCAUUCUUG	20	downstream
CEP290-B259	+	GUAGCUUUUGACAGUUUUUA	20	downstream
CEP290-B79	+	GGAGAUAGAGACAGGAAUAA	20	downstream
CEP290-B436	+	GUUCUGUCCUCAGUAAA	17	upstream
CEP290-B444	+	GGAUAGGACAGAGGACA	17	upstream
CEP290-B445	+	GAUGAAAAAUACUCUUU	17	upstream
CEP290-B459	-	GAACUCUAUACCUUUUACUG	20	upstream
CEP290-B465	+	GUAACAUAAUCACCUCUCUU	20	upstream
CEP290-B473	+	GAAAGAUGAAAAAUACUCUU	20	upstream
CEP290-B528	+	GUAACAUAAUCACCUCUCUU	20	upstream

Table 3D provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the fourth tier parameters. The targeting domains are within 400bp upstream of an *Alu* repeat or 700bp downstream of the mutation, and do not start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. pyogenes* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 3D

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	
CEP290-B233	+	AAGGCGGGGAGUCACAU	17	downstream
CEP290-B175	+	UAAGGCGGGGAGUCACA	17	downstream
CEP290-B280	+	UGAACUCUGUGCCAAAC	17	downstream
CEP290-B92	+	CUAAGACACUGCCAAUA	17	downstream
CEP290-B268	+	UUUACCCUUCAGGUAAC	17	downstream
CEP290-B154	+	UGACACCACAUGCACUG	17	downstream
CEP290-B44	+	ACUAAGACACUGCCAAU	17	downstream
CEP290-B231	+	UUGCUCUAGAUGACAUG	17	downstream
CEP290-B242	+	UGACAGUUUUUAAGGCG	17	downstream
CEP290-B226	-	UGUCAAAUAUGGUGCUU	17	downstream
CEP290-B159	+	AGUCACAUGGGAGUCAC	17	downstream
CEP290-B222	-	AUGAGAGUGAUUAGUGG	17	downstream
CEP290-B274	+	UGACAUGAGGUAAGUAG	17	downstream
CEP290-B68	-	UACAUGAGAGUGAUUAG	17	downstream
CEP290-B212	+	UAAGGAGGAUGUAAGAC	17	downstream
CEP290-B270	+	CUUGACUUUUACCCUUC	17	downstream
CEP290-B96	+	UCACUGAGCAAAACAAC	17	downstream
CEP290-B104	+	AGACUUAUAUUCCAUUA	17	downstream
CEP290-B122	+	CAUGGGAGUCACAGGGU	17	downstream
CEP290-B229	+	UAGAAUGAUCAUUCUUG	17	downstream
CEP290-B99	+	UUGACAGUUUUUAAGGC	17	downstream
CEP290-B7	-	AAACUGUCAAAAGCUAC	17	downstream
CEP290-B41	+	UCAUUCUUGUGGCAGUA	17	downstream
CEP290-B37	+	AUGACAUGAGGUAAGUA	17	downstream
CEP290-B97	-	UGUUUCAAAACUUGUUC	17	downstream
CEP290-B173	-	AUAUCUGUCUUCCUUAA	17	downstream
CEP290-B136	+	UGAACAAGUUUUGAAAC	17	downstream
CEP290-B71	-	UUCUGCAUCUUAUACAU	17	downstream
CEP290-B172	-	AUAAGUCUUUUGAUAUA	17	downstream
CEP290-B238	+	UUUGACAGUUUUUAAGG	17	downstream
CEP290-B148	-	UGCUCUUUUCUAUAUAU	17	downstream
CEP290-B208	+	AGACUGGAGAUAGAGAC	17	downstream
CEP290-B53	+	CAUAAGAAAGAACACUG	17	downstream
CEP290-B166	+	UUCUUGUGGCAGUAAGG	17	downstream
CEP290-B247	-	AAGCAUACUUUUUUUAA	17	downstream
CEP290-B245	+	CAACUGGAAGAGAGAAA	17	downstream
CEP290-B167	+	UAUGCUUAAGAAAAAAA	17	downstream
CEP290-B171	-	UUUUAUUAUCUUUAUUG	17	downstream
CEP290-B140	+	CUAGAUGACAUGAGGUAAGU	20	downstream
CEP290-B147	+	UUUUAAGGCGGGGAGUCACA	20	downstream
CEP290-B253	+	AAGACACUGCCAAUAGGGAU	20	downstream

CEP290-B73	-	UCCUGUUUCAAAACUUGUUC	20	downstream
CEP290-B206	-	UGUGUUGAGUAUCUCCUGUU	20	downstream
CEP290-B57	+	CUCUUGCUCUAGAUGACAUG	20	downstream
CEP290-B82	+	CAGUAAGGAGGAUGUAAGAC	20	downstream
CEP290-B265	+	AGAUGACAUGAGGUAAGUAG	20	downstream
CEP290-B105	+	AAUUCACUGAGCAAAACAAC	20	downstream
CEP290-B239	+	UCACAUGGGAGUCACAGGGU	20	downstream
CEP290-B180	+	UAGAUGACAUGAGGUAAGUA	20	downstream
CEP290-B103	+	UUUUGACAGUUUUUAAGGCG	20	downstream
CEP290-B254	-	UAAUACAUGAGAGUGAUUAG	20	downstream
CEP290-B134	-	UAGUUCUGCAUCUUAUACAU	20	downstream
CEP290-B151	+	AAACUAAGACACUGCCAAUA	20	downstream
CEP290-B196	+	AAAACUAAGACACUGCCAAU	20	downstream
CEP290-B2	-	UAAAAACUGUCAAAAGCUAC	20	downstream
CEP290-B240	+	CUUUUGACAGUUUUUAAGGC	20	downstream
CEP290-B116	+	AAAAGACUUAUAUUCCAUUA	20	downstream
CEP290-B39	+	AUACAUAAGAAAGAACACUG	20	downstream
CEP290-B91	-	AAUAUAAGUCUUUUGAUAUA	20	downstream
CEP290-B126	+	UGAUCAUUCUUGUGGCAGUA	20	downstream
CEP290-B202	-	UACAUAUCUGUCUUCCUUAA	20	downstream
CEP290-B152	-	CUUAAGCAUACUUUUUUUAA	20	downstream
CEP290-B77	+	AAACAACUGGAAGAGAGAAA	20	downstream
CEP290-B145	+	UCAUUCUUGUGGCAGUAAGG	20	downstream
CEP290-B72	+	AAGUAUGCUUAAGAAAAAAA	20	downstream
CEP290-B221	-	AUUUUUUAUUAUCUUUAUUG	20	downstream
CEP290-B424	+	CUAGGACUUUCUAAUGC	17	upstream
CEP290-B425	-	AUCUAAGAUCCUUUCAC	17	upstream
CEP290-B426	+	UUAUCACCACACUAAAU	17	upstream
CEP290-B427	-	AGCUCAAAAGCUUUUGC	17	upstream
CEP290-B428	-	UGUUCUGAGUAGCUUUC	17	upstream
CEP290-B429	+	ACUUUCUAAUGCUGGAG	17	upstream
CEP290-B430	-	CUCUAUACCUUUUACUG	17	upstream
CEP290-B431	+	CAAGAUGUCUCUUGCCU	17	upstream
CEP290-B432	-	AUUAUGCCUAUUUAGUG	17	upstream
CEP290-B433	+	AUGACUCAUAAUUUAGU	17	upstream
CEP290-B434	-	UAGAGGCUUAUGGAUUU	17	upstream
CEP290-B435	+	UAUUCUACUCCUGUGAA	17	upstream
CEP290-B437	+	CUAAUGCUGGAGAGGAU	17	upstream
CEP290-B438	-	AGGCAAGAGACAUCUUG	17	upstream
CEP290-B439	+	AGCCUCUAUUUCUGAUG	17	upstream
CEP290-B440	-	CAGCAUUAGAAAGUCCU	17	upstream
CEP290-B441	-	CUGCUUUUGCCAAAGAG	17	upstream

CEP290-B442	+	ACAUAAUCACCUCUCUU	17	upstream
CEP290-B443	-	UCAGAAAUAGAGGCUUA	17	upstream
CEP290-B446	-	UUCCUCAUCAGAAAUAG	17	upstream
CEP290-B447	+	ACAGAGGACAUGGAGAA	17	upstream
CEP290-B448	+	UGGAGAGGAUAGGACAG	17	upstream
CEP290-B449	+	AGGAAGAUGAACAAAUC	17	upstream
CEP290-B450	+	AGAUGAAAAAUACUCUU	17	upstream
CEP290-B455	+	AGGACUUUCUAAUGCUGGAG	20	upstream
CEP290-B456	-	AUUAGCUCAAAAGCUUUUGC	20	upstream
CEP290-B457	-	CUCCAGCAUUAGAAAGUCCU	20	upstream
CEP290-B458	+	AACAUGACUCAUAAUUUAGU	20	upstream
CEP290-B460	-	AUCUUCCUCAUCAGAAAUAG	20	upstream
CEP290-B461	+	AUAAGCCUCUAUUUCUGAUG	20	upstream
CEP290-B462	+	UCUUAUUCUACUCCUGUGAA	20	upstream
CEP290-B463	-	CUGCUGCUUUUGCCAAAGAG	20	upstream
CEP290-B464	+	UUUCUAAUGCUGGAGAGGAU	20	upstream
CEP290-B466	+	AAAUUAUCACCACACUAAAU	20	upstream
CEP290-B467	+	CUUGUUCUGUCCUCAGUAAA	20	upstream
CEP290-B468	-	AAAAUUAUGCCUAUUUAGUG	20	upstream
CEP290-B469	-	UCAUCAGAAAUAGAGGCUUA	20	upstream
CEP290-B470	-	AAAUAGAGGCUUAUGGAUUU	20	upstream
CEP290-B471	+	UGCUGGAGAGGAUAGGACAG	20	upstream
CEP290-B472	+	AUGAGGAAGAUGAACAAAUC	20	upstream
CEP290-B474	-	CUUAUCUAAGAUCCUUUCAC	20	upstream
CEP290-B475	+	AGAGGAUAGGACAGAGGACA	20	upstream
CEP290-B476	+	AGGACAGAGGACAUGGAGAA	20	upstream
CEP290-B477	+	AAAGAUGAAAAAUACUCUUU	20	upstream
CEP290-B495	-	AGCUCAAAAGCUUUUGC	17	upstream
CEP290-B529	-	UGUUCUGAGUAGCUUUC	17	upstream
CEP290-B513	+	AUGACUCAUAAUUUAGU	17	upstream
CEP290-B490	+	UAUUCUACUCCUGUGAA	17	upstream
CEP290-B485	-	CUGCUUUUGCCAAAGAG	17	upstream
CEP290-B492	+	ACAUAAUCACCUCUCUU	17	upstream
CEP290-B506	+	AACAUGACUCAUAAUUUAGU	20	upstream
CEP290-B500	+	UCUUAUUCUACUCCUGUGAA	20	upstream
CEP290-B521	-	CUGCUGCUUUUGCCAAAGAG	20	upstream

Table 4A provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the first tier parameters. The targeting domains are within 400bp upstream of an *Alu* repeat or 700bp downstream of the mutation, have good orthogonality, and start with G. It is

contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. aureus* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 4A

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-B1008	+	GAAUCCUGAAAGCUACU	17	upstream

Table 4B provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the second tier parameters. The targeting domains are within 400bp upstream of an *Alu* repeat or 700bp downstream of the mutation, have good orthogonality, and do not start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. aureus* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 4B

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-B1009	-	CCUACUUACCUCAUGUCAUC	20	downstream
CEP290-B1010	+	CUAUGAGCCAGCAAAAGCUU	20	upstream
CEP290-B1011	-	ACGUUGUUCUGAGUAGCUUU	20	upstream
CEP290-B1012	-	CAUAGAGACACAUUCAGUAA	20	upstream
CEP290-B1013	-	ACUUACCUCAUGUCAUC	17	downstream
CEP290-B1014	+	UGAGCCAGCAAAAGCUU	17	upstream
CEP290-B1015	-	UUGUUCUGAGUAGCUUU	17	upstream
CEP290-B1016	-	AGAGACACAUUCAGUAA	17	upstream
CEP290-B1017	+	UUUAAGGCGGGGAGUCACAU	20	downstream
CEP290-B1018	-	CAAAAGCUACCGGUUACCUG	20	downstream
CEP290-B1019	+	UUUUAAGGCGGGGAGUCACA	20	downstream
CEP290-B1020	-	UGUCAAAAGCUACCGGUUAC	20	downstream
CEP290-B1021	+	AAGGCGGGGAGUCACAU	17	downstream
CEP290-B1022	-	AAGCUACCGGUUACCUG	17	downstream
CEP290-B1023	+	UAAGGCGGGGAGUCACA	17	downstream

CEP290-B1024	-	CAAAAGCUACCGGUUAC	17	downstream
CEP290-B1025	+	UAGGAAUCCUGAAAGCUACU	20	upstream
CEP290-B1026	+	CAGAACAACGUUUUCAUUUA	20	upstream
CEP290-B1027	-	CAAAAGCUUUUGCUGGCUCA	20	upstream
CEP290-B1028	+	AGCAAAAGCUUUUGAGCUAA	20	upstream
CEP290-B1029	+	AUCUUAUUCUACUCCUGUGA	20	upstream
CEP290-B1030	+	AACAACGUUUUCAUUUA	17	upstream
CEP290-B1031	-	AAGCUUUUGCUGGCUCA	17	upstream
CEP290-B1032	+	AAAAGCUUUUGAGCUAA	17	upstream
CEP290-B1033	+	UUAUUCUACUCCUGUGA	17	upstream

Table 4C provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the third tier parameters. The targeting domains are within 400bp upstream of an *Alu* repeat or 700bp downstream of the mutation, and start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. aureus* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 4C

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-B1034	+	GAAACAGGAAUAGAAAUUCA	20	downstream
CEP290-B1035	+	GAAAGAUGAAAAAUACUCUU	20	upstream
CEP290-B1036	-	GAAAUAGAGGCUUAUGGAUU	20	upstream
CEP290-B1037	-	GAAUAUAAGUCUUUUGAUAU	20	downstream
CEP290-B1038	+	GAGAAAUGGUUCCCUAUAUA	20	downstream
CEP290-B1039	+	GAGAGGAUAGGACAGAGGAC	20	upstream
CEP290-B1040	+	GAUGAGGAAGAUGAACAAAU	20	upstream
CEP290-B1041	+	GAUGCAGAACUAGUGUAGAC	20	downstream
CEP290-B1042	-	GAUUUGUUCAUCUUCCUCAU	20	upstream
CEP290-B1043	+	GCAGUAAGGAGGAUGUAAGA	20	downstream
CEP290-B1044	+	GCCUGAACAAGUUUUGAAAC	20	downstream
CEP290-B1045	+	GCUUGAACUCUGUGCCAAAC	20	downstream
CEP290-B1046	-	GCUUUCUGCUGCUUUUGCCA	20	upstream
CEP290-B1047	-	GCUUUCUGCUGCUUUUGCCA	20	upstream
CEP290-B1048	+	GCUUUUGACAGUUUUUAAGG	20	downstream
CEP290-B1049	+	GGAAAGAUGAAAAAUACUCU	20	upstream
CEP290-B1050	+	GGAGGAUGUAAGACUGGAGA	20	downstream

CEP290-B1051	+	GGGGAGUCACAUGGGAGUCA	20	downstream
CEP290-B1052	-	GGUGAUUAUGUUACUUUUUA	20	upstream
CEP290-B1053	-	GGUGAUUAUGUUACUUUUUA	20	upstream
CEP290-B1054	+	GUAAGACUGGAGAUAGAGAC	20	downstream
CEP290-B1055	+	GUCACAUGGGAGUCACAGGG	20	downstream
CEP290-B1056	-	GUGGUGUCAAAUAUGGUGCU	20	downstream
CEP290-B1057	+	GAAAAAAAAGGUAAUGC	17	downstream
CEP290-B1058	+	GAAAAGAGCACGUACAA	17	downstream
CEP290-B1059	+	GAAUCCUGAAAGCUACU	17	upstream
CEP290-B1060	-	GAAUGAUCAUUCUAAAC	17	downstream
CEP290-B1061	+	GACAGAGGACAUGGAGA	17	upstream
CEP290-B1062	+	GACUUUCUAAUGCUGGA	17	upstream
CEP290-B1063	-	GAGAGUGAUUAGUGGUG	17	downstream
CEP290-B1064	+	GAGCAAAACAACUGGAA	17	downstream
CEP290-B1065	+	GAGGAAGAUGAACAAAU	17	upstream
CEP290-B1066	+	GAGUCACAUGGGAGUCA	17	downstream
CEP290-B1067	+	GAUCUUAUUCUACUCCU	17	upstream
CEP290-B1068	+	GAUCUUAUUCUACUCCU	17	upstream
CEP290-B1069	+	GAUGAAAAAUACUCUUU	17	upstream
CEP290-B1070	+	GAUGACAUGAGGUAAGU	17	downstream
CEP290-B1071	-	GAUUAUGUUACUUUUUA	17	upstream
CEP290-B1072	-	GAUUAUGUUACUUUUUA	17	upstream
CEP290-B1073	+	GCAAAACAACUGGAAGA	17	downstream
CEP290-B1074	+	GCAGAACUAGUGUAGAC	17	downstream
CEP290-B1075	-	GCUCUUUUCUAUAUAUA	17	downstream
CEP290-B1076	+	GGAUAGGACAGAGGACA	17	upstream
CEP290-B1077	+	GGAUGUAAGACUGGAGA	17	downstream
CEP290-B1078	+	GUAAGGAGGAUGUAAGA	17	downstream
CEP290-B1079	-	GUAUCUCCUGUUUGGCA	17	downstream
CEP290-B1080	-	GUCAUCUAGAGCAAGAG	17	downstream
CEP290-B1081	+	GUCCUCAGUAAAAGGUA	17	upstream
CEP290-B1082	+	GUGAAAGGAUCUUAGAU	17	upstream
CEP290-B1083	-	GUGCUCUUUUCUAUAUA	17	downstream
CEP290-B1084	-	GUGUCAAAUAUGGUGCU	17	downstream
CEP290-B1085	+	GUUCCCUAUAUAUAGAA	17	downstream

Table 4D provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the fourth tier parameters. The targeting domains are within 400bp upstream of an *Alu* repeat or 700bp downstream of the mutation, and do not start with G. It is contemplated herein that the

targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. aureus* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 4D

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-B1086	+	AAAACUAAGACACUGCCAAU	20	downstream
CEP290-B1087	+	AAAAGACUUAUAUUCCAUUA	20	downstream
CEP290-B1088	+	AAACAUGACUCAUAAUUUAG	20	upstream
CEP290-B1089	+	AAACAUGACUCAUAAUUUAG	20	upstream
CEP290-B1090	+	AAAGAUGAAAAAUACUCUUU	20	upstream
CEP290-B1091	+	AAAUUCACUGAGCAAAACAA	20	downstream
CEP290-B1092	+	AACAAGUUUUGAAACAGGAA	20	downstream
CEP290-B1093	+	AACAGGAGAUACUCAACACA	20	downstream
CEP290-B1094	+	AACAUGACUCAUAAUUUAGU	20	upstream
CEP290-B1095	+	AACAUGACUCAUAAUUUAGU	20	upstream
CEP290-B1096	-	AAUAUAAGUCUUUUGAUAUA	20	downstream
CEP290-B1097	+	AAUCACUCUCAUGUAUUAGC	20	downstream
CEP290-B1098	+	AAUUCACUGAGCAAAACAAC	20	downstream
CEP290-B1099	+	ACAAAAGAACAUACAUAAGA	20	downstream
CEP290-B1100	+	ACGUACAAAAGAACAUACAU	20	downstream
CEP290-B1101	-	ACGUGCUCUUUUCUAUAUAU	20	downstream
CEP290-B1102	-	ACGUUGUUCUGAGUAGCUUU	20	upstream
CEP290-B1103	+	ACUGAGCAAAACAACUGGAA	20	downstream
CEP290-B1104	+	AGAGGAUAGGACAGAGGACA	20	upstream
CEP290-B1105	+	AGAUGCAGAACUAGUGUAGA	20	downstream
CEP290-B1106	+	AGCAAAAGCUUUUGAGCUAA	20	upstream
CEP290-B1107	-	AGCAUUAGAAAGUCCUAGGC	20	upstream
CEP290-B1108	+	AGCUUGAACUCUGUGCCAAA	20	downstream
CEP290-B1109	+	AGCUUUUGACAGUUUUUAAG	20	downstream
CEP290-B1110	+	AGGACAGAGGACAUGGAGAA	20	upstream
CEP290-B1111	+	AGGAUAGGACAGAGGACAUG	20	upstream
CEP290-B1112	+	AGGUAAUGCCUGAACAAGUU	20	downstream
CEP290-B1113	+	AUAAGAAAGAACACUGUGGU	20	downstream
CEP290-B1114	+	AUAAGCCUCUAUUUCUGAUG	20	upstream
CEP290-B1115	-	AUACAUGAGAGUGAUUAGUG	20	downstream
CEP290-B1116	+	AUAGAAAAGAGCACGUACAA	20	downstream
CEP290-B1117	+	AUCAUUCUUGUGGCAGUAAG	20	downstream
CEP290-B1118	+	AUCUUAUUCUACUCCUGUGA	20	upstream
CEP290-B1119	-	AUCUUGUGGAUAAUGUAUCA	20	upstream

CEP290-B1120	+	AUGAGGAAGAUGAACAAAUC	20	upstream
CEP290-B1121	+	AUGAUCAUUCUUGUGGCAGU	20	downstream
CEP290-B1122	+	AUGCUGGAGAGGAUAGGACA	20	upstream
CEP290-B1123	+	AUGGUUCCCUAUAUAUAGAA	20	downstream
CEP290-B1124	-	AUUUAAUUUGUUUCUGUGUG	20	downstream
CEP290-B1125	+	CAAAACCUAUGUAUAAGAUG	20	downstream
CEP290-B1126	+	CAAAAGACUUAUAUUCCAUU	20	downstream
CEP290-B1127	-	CAAAAGCUUUUGCUGGCUCA	20	upstream
CEP290-B1128	-	CAAGAAUGAUCAUUCUAAAC	20	downstream
CEP290-B1129	-	CACAGAGUUCAAGCUAAUAC	20	downstream
CEP290-B1130	+	CACAGGGUAGGAUUCAUGUU	20	downstream
CEP290-B1131	+	CACUGCCAAUAGGGAUAGGU	20	downstream
CEP290-B1132	+	CAGAACAACGUUUUCAUUUA	20	upstream
CEP290-B1133	-	CAGAGUUCAAGCUAAUACAU	20	downstream
CEP290-B1134	-	CAGUAAAUGAAAACGUUGUU	20	upstream
CEP290-B1135	-	CAGUAAAUGAAAACGUUGUU	20	upstream
CEP290-B1136	+	CAGUAAGGAGGAUGUAAGAC	20	downstream
CEP290-B1137	+	CAUAAGCCUCUAUUUCUGAU	20	upstream
CEP290-B1138	-	CAUAGAGACACAUUCAGUAA	20	upstream
CEP290-B1139	+	CAUCUCUUGCUCUAGAUGAC	20	downstream
CEP290-B1140	-	CAUGAGAGUGAUUAGUGGUG	20	downstream
CEP290-B1141	-	CAUGUCAUCUAGAGCAAGAG	20	downstream
CEP290-B1142	+	CAUUUACUGAAUGUGUCUCU	20	upstream
CEP290-B1143	+	CAUUUACUGAAUGUGUCUCU	20	upstream
CEP290-B1144	+	CCAUUAAAAAAAGUAUGCUU	20	downstream
CEP290-B1145	+	CCUAGGACUUUCUAAUGCUG	20	upstream
CEP290-B1146	+	CCUCUCUUUGGCAAAAGCAG	20	upstream
CEP290-B1147	+	CCUCUCUUUGGCAAAAGCAG	20	upstream
CEP290-B1148	+	CCUGUGAAAGGAUCUUAGAU	20	upstream
CEP290-B1149	-	CGUGCUCUUUUCUAUAUAUA	20	downstream
CEP290-B1150	-	CUAAGAUCCUUUCACAGGAG	20	upstream
CEP290-B1151	+	CUAGAUGACAUGAGGUAAGU	20	downstream
CEP290-B1152	+	CUAUGAGCCAGCAAAAGCUU	20	upstream
CEP290-B1153	+	CUCAUAAUUUAGUAGGAAUC	20	upstream
CEP290-B1154	+	CUCAUAAUUUAGUAGGAAUC	20	upstream
CEP290-B1155	-	CUCAUCAGAAAUAGAGGCUU	20	upstream
CEP290-B1156	+	CUCUAUUUCUGAUGAGGAAG	20	upstream
CEP290-B1157	-	CUUAAGCAUACUUUUUUUAA	20	downstream
CEP290-B1158	-	CUUAUCUAAGAUCCUUUCAC	20	upstream
CEP290-B1159	+	CUUUCUAAUGCUGGAGAGGA	20	upstream
CEP290-B1160	+	CUUUUGACAGUUUUUAAGGC	20	downstream
CEP290-B1161	+	UAAAACUAAGACACUGCCAA	20	downstream

CEP290-B1162	+	UAAGAAAAAAAAGGUAAUGC	20	downstream
CEP290-B1163	+	UAAUGCUGGAGAGGAUAGGA	20	upstream
CEP290-B1164	-	UACAUAUCUGUCUUCCUUAA	20	downstream
CEP290-B1165	-	UACAUCCUCCUUACUGCCAC	20	downstream
CEP290-B1166	-	UACAUGAGAGUGAUUAGUGG	20	downstream
CEP290-B1167	-	UACCUCAUGUCAUCUAGAGC	20	downstream
CEP290-B1168	-	UACGUGCUCUUUUCUAUAUA	20	downstream
CEP290-B1169	-	UAGAGCAAGAGAUGAACUAG	20	downstream
CEP290-B1170	+	UAGAUGACAUGAGGUAAGUA	20	downstream
CEP290-B1171	+	UAGGAAUCCUGAAAGCUACU	20	upstream
CEP290-B1172	+	UAGGACAGAGGACAUGGAGA	20	upstream
CEP290-B1173	+	UAGGACUUUCUAAUGCUGGA	20	upstream
CEP290-B1174	+	UCACUGAGCAAAACAACUGG	20	downstream
CEP290-B1175	-	UCAUGUUUAUCAAUAUUAUU	20	upstream
CEP290-B1176	-	UCAUGUUUAUCAAUAUUAUU	20	upstream
CEP290-B1177	+	UCCACAAGAUGUCUCUUGCC	20	upstream
CEP290-B1178	+	UCCAUAAGCCUCUAUUUCUG	20	upstream
CEP290-B1179	-	UCCUAGGCAAGAGACAUCUU	20	upstream
CEP290-B1180	+	UCUAGAUGACAUGAGGUAAG	20	downstream
CEP290-B1181	-	UCUAUACCUUUUACUGAGGA	20	upstream
CEP290-B1182	+	UCUGUCCUCAGUAAAAGGUA	20	upstream
CEP290-B1183	-	UCUUAAGCAUACUUUUUUUA	20	downstream
CEP290-B1184	-	UCUUAUCUAAGAUCCUUUCA	20	upstream
CEP290-B1185	-	UCUUCCAGUUGUUUUGCUCA	20	downstream
CEP290-B1186	+	UGAGCAAAACAACUGGAAGA	20	downstream
CEP290-B1187	-	UGAGUAUCUCCUGUUUGGCA	20	downstream
CEP290-B1188	+	UGAUCAUUCUUGUGGCAGUA	20	downstream
CEP290-B1189	+	UGCCUAGGACUUUCUAAUGC	20	upstream
CEP290-B1190	+	UGCCUGAACAAGUUUUGAAA	20	downstream
CEP290-B1191	-	UGGUGUCAAAUAUGGUGCUU	20	downstream
CEP290-B1192	+	UGUAAGACUGGAGAUAGAGA	20	downstream
CEP290-B1193	-	UGUCCUAUCCUCUCCAGCAU	20	upstream
CEP290-B1194	-	UUAACGUUAUCAUUUUCCCA	20	upstream
CEP290-B1195	-	UUACAUAUCUGUCUUCCUUA	20	downstream
CEP290-B1196	+	UUAGAUCUUAUUCUACUCCU	20	upstream
CEP290-B1197	+	UUAGAUCUUAUUCUACUCCU	20	upstream
CEP290-B1198	-	UUCAGGAUUCCUACUAAAUU	20	upstream
CEP290-B1199	-	UUCAGGAUUCCUACUAAAUU	20	upstream
CEP290-B1200	-	UUCAUCUUCCUCAUCAGAAA	20	upstream
CEP290-B1201	+	UUGCCUAGGACUUUCUAAUG	20	upstream
CEP290-B1202	-	UUUCUGCUGCUUUUGCCAAA	20	upstream
CEP290-B1203	-	UUUCUGCUGCUUUUGCCAAA	20	upstream

CEP290-B1204	+	UUUUGACAGUUUUUAAGGCG	20	downstream
CEP290-B1205	+	UUUUUAAGGCGGGGAGUCAC	20	downstream
CEP290-B1206	+	AAAAGCUUUUGAGCUAA	17	upstream
CEP290-B1207	+	AAAGAACAUACAUAAGA	17	downstream
CEP290-B1208	+	AAAUGGUUCCCUAUAUA	17	downstream
CEP290-B1209	+	AACAACGUUUUCAUUUA	17	upstream
CEP290-B1210	+	AACCUAUGUAUAAGAUG	17	downstream
CEP290-B1211	+	AACUAAGACACUGCCAA	17	downstream
CEP290-B1212	+	AAGACUGGAGAUAGAGA	17	downstream
CEP290-B1213	+	AAGACUUAUAUUCCAUU	17	downstream
CEP290-B1214	+	AAGAUGAAAAAUACUCU	17	upstream
CEP290-B1215	-	AAGCAUACUUUUUUUAA	17	downstream
CEP290-B1216	+	AAGCCUCUAUUUCUGAU	17	upstream
CEP290-B1217	-	AAGCUUUUGCUGGCUCA	17	upstream
CEP290-B1218	+	AAGUUUUGAAACAGGAA	17	downstream
CEP290-B1219	+	ACAAGAUGUCUCUUGCC	17	upstream
CEP290-B1220	+	ACAGAGGACAUGGAGAA	17	upstream
CEP290-B1221	+	ACAGGAAUAGAAAUUCA	17	downstream
CEP290-B1222	+	ACAUGGGAGUCACAGGG	17	downstream
CEP290-B1223	-	ACGUUAUCAUUUUCCCA	17	upstream
CEP290-B1224	+	ACUAAGACACUGCCAAU	17	downstream
CEP290-B1225	+	AGAAAGAACACUGUGGU	17	downstream
CEP290-B1226	+	AGACUGGAGAUAGAGAC	17	downstream
CEP290-B1227	+	AGACUUAUAUUCCAUUA	17	downstream
CEP290-B1228	-	AGAGACACAUUCAGUAA	17	upstream
CEP290-B1229	-	AGAGUUCAAGCUAAUAC	17	downstream
CEP290-B1230	-	AGAUCCUUUCACAGGAG	17	upstream
CEP290-B1231	+	AGAUGAAAAAUACUCUU	17	upstream
CEP290-B1232	+	AGAUGACAUGAGGUAAG	17	downstream
CEP290-B1233	-	AGCAAGAGAUGAACUAG	17	downstream
CEP290-B1234	+	AGCCUCUAUUUCUGAUG	17	upstream
CEP290-B1235	+	AGGAAGAUGAACAAAUC	17	upstream
CEP290-B1236	+	AGGACUUUCUAAUGCUG	17	upstream
CEP290-B1237	+	AGGAGAUACUCAACACA	17	downstream
CEP290-B1238	+	AGGAUAGGACAGAGGAC	17	upstream
CEP290-B1239	-	AGGAUUCCUACUAAAUU	17	upstream
CEP290-B1240	-	AGGAUUCCUACUAAAUU	17	upstream
CEP290-B1241	+	AGGGUAGGAUUCAUGUU	17	downstream
CEP290-B1242	-	AGUUCAAGCUAAUACAU	17	downstream
CEP290-B1243	+	AUAAGCCUCUAUUUCUG	17	upstream
CEP290-B1244	-	AUAAGUCUUUUGAUAUA	17	downstream
CEP290-B1245	+	AUAAUUUAGUAGGAAUC	17	upstream

CEP290-B1246	+	AUAAUUUAGUAGGAAUC	17	upstream
CEP290-B1247	-	AUACCUUUUACUGAGGA	17	upstream
CEP290-B1248	-	AUAGAGGCUUAUGGAUU	17	upstream
CEP290-B1249	+	AUAGGACAGAGGACAUG	17	upstream
CEP290-B1250	-	AUAUCUGUCUUCCUUAA	17	downstream
CEP290-B1251	-	AUCAGAAAUAGAGGCUU	17	upstream
CEP290-B1252	+	AUCAUUCUUGUGGCAGU	17	downstream
CEP290-B1253	-	AUCCUCCUUACUGCCAC	17	downstream
CEP290-B1254	-	AUCUAAGAUCCUUUCAC	17	upstream
CEP290-B1255	-	AUCUUCCUCAUCAGAAA	17	upstream
CEP290-B1256	+	AUGACAUGAGGUAAGUA	17	downstream
CEP290-B1257	+	AUGACUCAUAAUUUAGU	17	upstream
CEP290-B1258	+	AUGACUCAUAAUUUAGU	17	upstream
CEP290-B1259	-	AUGAGAGUGAUUAGUGG	17	downstream
CEP290-B1260	-	AUUAGAAAGUCCUAGGC	17	upstream
CEP290-B1261	+	AUUCUUGUGGCAGUAAG	17	downstream
CEP290-B1262	+	CACUCUCAUGUAUUAGC	17	downstream
CEP290-B1263	-	CAUAUCUGUCUUCCUUA	17	downstream
CEP290-B1264	+	CAUGACUCAUAAUUUAG	17	upstream
CEP290-B1265	+	CAUGACUCAUAAUUUAG	17	upstream
CEP290-B1266	-	CAUGAGAGUGAUUAGUG	17	downstream
CEP290-B1267	+	CCUAGGACUUUCUAAUG	17	upstream
CEP290-B1268	-	CCUAUCCUCUCCAGCAU	17	upstream
CEP290-B1269	+	CUAGGACUUUCUAAUGC	17	upstream
CEP290-B1270	-	CUCAUGUCAUCUAGAGC	17	downstream
CEP290-B1271	+	CUCUUGCUCUAGAUGAC	17	downstream
CEP290-B1272	+	CUCUUUGGCAAAAGCAG	17	upstream
CEP290-B1273	+	CUCUUUGGCAAAAGCAG	17	upstream
CEP290-B1274	+	CUGAACAAGUUUUGAAA	17	downstream
CEP290-B1275	+	CUGAGCAAAACAACUGG	17	downstream
CEP290-B1276	-	CUGCUGCUUUUGCCAAA	17	upstream
CEP290-B1277	-	CUGCUGCUUUUGCCAAA	17	upstream
CEP290-B1278	+	CUGGAGAGGAUAGGACA	17	upstream
CEP290-B1279	-	UAAAUGAAAACGUUGUU	17	upstream
CEP290-B1280	-	UAAAUGAAAACGUUGUU	17	upstream
CEP290-B1281	-	UAAGCAUACUUUUUUUA	17	downstream
CEP290-B1282	+	UAAGGAGGAUGUAAGAC	17	downstream
CEP290-B1283	+	UAAUGCCUGAACAAGUU	17	downstream
CEP290-B1284	-	UAAUUUGUUUCUGUGUG	17	downstream
CEP290-B1285	+	UACAAAAGAACAUACAU	17	downstream
CEP290-B1286	-	UAGGCAAGAGACAUCUU	17	upstream
CEP290-B1287	-	UAUAAGUCUUUUGAUAU	17	downstream

CEP290-B1288	-	UAUCUAAGAUCCUUUCA	17	upstream
CEP290-B1289	+	UAUUUCUGAUGAGGAAG	17	upstream
CEP290-B1290	+	UCACUGAGCAAAACAAC	17	downstream
CEP290-B1291	+	UCAUUCUUGUGGCAGUA	17	downstream
CEP290-B1292	-	UCCAGUUGUUUUGCUCA	17	downstream
CEP290-B1293	+	UCUAAUGCUGGAGAGGA	17	upstream
CEP290-B1294	+	UGAACAAGUUUUGAAAC	17	downstream
CEP290-B1295	+	UGAACUCUGUGCCAAAC	17	downstream
CEP290-B1296	+	UGACAGUUUUUAAGGCG	17	downstream
CEP290-B1297	+	UGAGCCAGCAAAAGCUU	17	upstream
CEP290-B1298	+	UGCAGAACUAGUGUAGA	17	downstream
CEP290-B1299	+	UGCCAAUAGGGAUAGGU	17	downstream
CEP290-B1300	-	UGCUCUUUUCUAUAUAU	17	downstream
CEP290-B1301	+	UGCUGGAGAGGAUAGGA	17	upstream
CEP290-B1302	-	UGUCAAAUAUGGUGCUU	17	downstream
CEP290-B1303	-	UGUUUAUCAAUAUUAUU	17	upstream
CEP290-B1304	-	UGUUUAUCAAUAUUAUU	17	upstream
CEP290-B1305	+	UUAAAAAAAGUAUGCUU	17	downstream
CEP290-B1306	+	UUAAGGCGGGGAGUCAC	17	downstream
CEP290-B1307	+	UUACUGAAUGUGUCUCU	17	upstream
CEP290-B1308	+	UUACUGAAUGUGUCUCU	17	upstream
CEP290-B1309	+	UUAUUCUACUCCUGUGA	17	upstream
CEP290-B1310	+	UUCACUGAGCAAAACAA	17	downstream
CEP290-B1311	-	UUCUGCUGCUUUUGCCA	17	upstream
CEP290-B1312	-	UUCUGCUGCUUUUGCCA	17	upstream
CEP290-B1313	+	UUGAACUCUGUGCCAAA	17	downstream
CEP290-B1314	+	UUGACAGUUUUUAAGGC	17	downstream
CEP290-B1315	-	UUGUGGAUAAUGUAUCA	17	upstream
CEP290-B1316	-	UUGUUCAUCUUCCUCAU	17	upstream
CEP290-B1317	-	UUGUUCUGAGUAGCUUU	17	upstream
CEP290-B1318	+	UUUGACAGUUUUUAAGG	17	downstream
CEP290-B1319	+	UUUUGACAGUUUUUAAG	17	downstream

Table 5A provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the first tier parameters. The targeting domains are within 400bp upstream of an *Alu* repeat or 700bp downstream of the mutation, have good orthogonality, and start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *N.*

meningitidis Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 5A

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-B65	-	GAGUUCAAGCUAAUACAUGA	20	downstream
CEP290-B296	-	GUUGUUCUGAGUAGCUU	17	upstream
CEP290-B308	+	GGCAAAAGCAGCAGAAAGCA	20	upstream
CEP290-B536	-	GUUGUUCUGAGUAGCUU	17	upstream
CEP290-B482	+	GGCAAAAGCAGCAGAAAGCA	20	upstream

Table 5B provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the second tier parameters. The targeting domains are within 400 bp upstream of an *Alu* repeat or 700 bp downstream of the mutation, have good orthogonality, and do not start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *N. meningitidis* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 5B

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-B235	-	UUCAAGCUAAUACAUGA	17	downstream
CEP290-B109	+	CACAUGGGAGUCACAGG	17	downstream
CEP290-B129	+	AGUCACAUGGGAGUCACAGG	20	downstream
CEP290-B295	-	AAUAGAGGCUUAUGGAU	17	upstream
CEP290-B297	-	CUGAGGACAGAACAAGC	17	upstream
CEP290-B298	-	CAUCAGAAAUAGAGGCU	17	upstream
CEP290-B299	-	CUGCUUUUGCCAAAGAG	17	upstream
CEP290-B300	+	AGCAGAAAGCAAACUGA	17	upstream
CEP290-B301	+	AAAAGCAGCAGAAAGCA	17	upstream
CEP290-B302	-	UUACUGAGGACAGAACAAGC	20	upstream
CEP290-B303	-	AACGUUGUUCUGAGUAGCUU	20	upstream
CEP290-B304	-	CUGCUGCUUUUGCCAAAGAG	20	upstream
CEP290-B305	-	AGAAAUAGAGGCUUAUGGAU	20	upstream
CEP290-B306	-	CCUCAUCAGAAAUAGAGGCU	20	upstream

CEP290-B307	+	AGCAGCAGAAAGCAAACUGA	20	upstream
CEP290-B531	-	CUGCUUUUGCCAAAGAG	17	upstream
CEP290-B522	+	AGCAGAAAGCAAACUGA	17	upstream
CEP290-B537	+	AAAAGCAGCAGAAAGCA	17	upstream
CEP290-B504	-	AACGUUGUUCUGAGUAGCUU	20	upstream
CEP290-B478	-	CUGCUGCUUUUGCCAAAGAG	20	upstream
CEP290-B526	+	AGCAGCAGAAAGCAAACUGA	20	upstream

Table 6A provides targeting domains for introduction of an indel (e.g., mediated by NHEJ) in close proximity to or including the LCA10 target position in the *CEP290* gene selected according to the first tier parameters. The targeting domains are within 40 bases of the LCA10 target position, have good orthogonality, start with G and PAM is NNGRRT. It is contemplated herein that in an embodiment the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. aureus* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 6A

gRNA Name	DNA Strand	Targeting Domain	Target Site Length
cep290-12	-	GCACCUGGCCCCAGUUGUAAUU	22

Table 6B provides targeting domains for introduction of an indel (e.g., mediated by NHEJ) in close proximity to or including the LCA10 target position in the *CEP290* gene selected according to the second tier parameters. The targeting domains are within 40 bases of the LCA10 target position, have good orthogonality, and PAM is NNGRRT. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. aureus* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 6B

gRNA Name	DNA Strand	Targeting Domain	Target Site Length
CEP290-35	+	AAAUAAAACUAAGACACUGCCAAU	24
CEP290-36	+	AAUAAAACUAAGACACUGCCAAU	23
CEP290-37	+	AUAAAACUAAGACACUGCCAAU	22
CEP290-38	+	AAAACUAAGACACUGCCAAU	20
CEP290-39	+	AAACUAAGACACUGCCAAU	19

CEP290-40	+	AACUAAGACACUGCCAAU	18
CEP290-512	-	ACCUGGCCCCAGUUGUAAUU	20
CEP290-17	-	CCGCACCUGGCCCCAGUUGUAAUU	24
CEP290-41	-	CGCACCUGGCCCCAGUUGUAAUU	23
CEP290-42	-	CACCUGGCCCCAGUUGUAAUU	21
CEP290-513	-	CCUGGCCCCAGUUGUAAUU	19
CEP290-514	-	CUGGCCCCAGUUGUAAUU	18
CEP290-43	+	UAAAACUAAGACACUGCCAAU	21

Table 6C provides targeting domains for introduction of an indel (e.g., mediated by NHEJ) in close proximity to or including the LCA10 target position in the *CEP290* gene selected according to the fifth tier parameters. The targeting domains are within 40 bases of the LCA10 target position, and PAM is NNGRRV. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. aureus* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 6C

gRNA Name	DNA Strand	Targeting Domain	Target Site Length
CEP290-44	+	AAAAUAAAACUAAGACACUGCCAA	24
CEP290-45	+	AAAUAAAACUAAGACACUGCCAA	23
CEP290-46	+	AAUAAAACUAAGACACUGCCAA	22
CEP290-47	+	AUAAAACUAAGACACUGCCAA	21
CEP290-48	+	AAAACUAAGACACUGCCAA	19
CEP290-49	+	AAACUAAGACACUGCCAA	18
CEP290-16	+	AAGACACUGCCAAUAGGGAUAGGU	24
CEP290-50	+	AGACACUGCCAAUAGGGAUAGGU	23
CEP290-51	+	ACACUGCCAAUAGGGAUAGGU	21
CEP290-510	+	ACUGCCAAUAGGGAUAGGU	19
CEP290-509	+	CACUGCCAAUAGGGAUAGGU	20
CEP290-511	+	CUGCCAAUAGGGAUAGGU	18
CEP290-11	+	GACACUGCCAAUAGGGAUAGGU	22
CEP290-52	+	UAAAACUAAGACACUGCCAA	20
CEP290-13	+	AUGAGAUACUCACAAUUACAAC	22
CEP290-53	+	AGAUACUCACAAUUACAAC	19
CEP290-18	+	GUAUGAGAUACUCACAAUUACAAC	24
CEP290-54	+	GAGAUACUCACAAUUACAAC	20
CEP290-55	+	GAUACUCACAAUUACAAC	18
CEP290-14	+	UAUGAGAUACUCACAAUUACAAC	23

CEP290-57	+	UGAGAUACUCACAAUUACAAC	21
CEP290-58	+	AUGAGAUAUUCACAAUUACAA	21
CEP290-59	+	AGAUAUUCACAAUUACAA	18
CEP290-19	+	GGUAUGAGAUAUUCACAAUUACAA	24
CEP290-61	+	GUAUGAGAUAUUCACAAUUACAA	23
CEP290-63	+	GAGAUAUUCACAAUUACAA	19
CEP290-65	+	UAUGAGAUAUUCACAAUUACAA	22
CEP290-66	+	UGAGAUAUUCACAAUUACAA	20

Table 6D provides targeting domains for introduction of an indel (e.g., mediated by NHEJ) in close proximity to or including the LCA10 target position in the *CEP290* gene that can be used for dual targeting. Any of the targeting domains in the table can be used with a *S. aureus* Cas9 (nickase) molecule to generate a single stranded break. Exemplary nickase pairs including selecting a targeting domain from Group A and a second targeting domain from Group B. It is contemplated herein that a targeting domain of Group A can be combined with any of the targeting domains of Group B. For example, the *CEP290*-12 or *CEP290*-17 can be combined with *CEP290*-11 or *CEP290*-16.

Table 6D

Group A	Group B
CEP290-12	CEP290-11
CEP290-17	CEP290-16

Table 7A provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the first tier parameters. The targeting domains are within 1000 bp upstream of an *Alu* repeat, within 40bp upstream of mutation, or 1000 bp downstream of the mutation, have good orthogonality, and start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. pyogenes* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 7A

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-67	+	GAAAGAUGAAAAAUACUCUU	20	upstream

CEP290-68	-	GAAAUAGAUGUAGAUUG	17	downstream
CEP290-70	-	GAAAUAUUAAGGGCUCUUCC	20	upstream
CEP290-71	+	GAACAAAAGCCAGGGACCAU	20	upstream
CEP290-72	-	GAACUCUAUACCUUUUACUG	20	upstream
CEP290-73	-	GAAGAAUGGAAUAGAUAAUA	20	downstream
CEP290-74	+	GAAUAGUUUGUUCUGGGUAC	20	upstream
CEP290-75	-	GAAUGGAAUAGAUAAUA	17	downstream
CEP290-76	+	GAAUUUACAGAGUGCAUCCA	20	upstream
CEP290-77	-	GAGAAAAAGGAGCAUGAAAC	20	upstream
CEP290-78	-	GAGAGCCACAGUGCAUG	17	downstream
CEP290-79	-	GAGGUAGAAUCAAGAAG	17	downstream
CEP290-80	+	GAGUGCAUCCAUGGUCC	17	upstream
CEP290-81	+	GAUAACUACAAAGGGUC	17	upstream
CEP290-82	+	GAUAGAGACAGGAAUAA	17	downstream
CEP290-83	+	GAUGAAAAAUACUCUUU	17	upstream
CEP290-84	+	GAUGACAUGAGGUAAGU	17	downstream
CEP290-85	+	GAUGCAGAACUAGUGUAGAC	20	downstream
CEP290-86	+	GCAGAACUAGUGUAGAC	17	downstream
CEP290-87	-	GCAUGUGGUGUCAAAUA	17	downstream
CEP290-88	+	GCCUGAACAAGUUUUGAAAC	20	downstream
CEP290-89	-	GCUACCGGUUACCUGAA	17	downstream
CEP290-90	-	GCUCUUUUCUAUAUAUA	17	downstream
CEP290-91	+	GCUUGAACUCUGUGCCAAAC	20	downstream
CEP290-92	+	GCUUUUGACAGUUUUUAAGG	20	downstream
CEP290-93	-	GCUUUUGUUCCUUGGAA	17	upstream
CEP290-94	+	GGAACAAAAGCCAGGGACCA	20	upstream
CEP290-95	+	GGACUUGACUUUUACCCUUC	20	downstream
CEP290-96	+	GGAGAAUAGUUUGUUCU	17	upstream
CEP290-97	+	GGAGUCACAUGGGAGUCACA	20	downstream
CEP290-98	+	GGAUAGGACAGAGGACA	17	upstream
CEP290-99	+	GGCUGUAAGAUAACUACAAA	20	upstream
CEP290-100	+	GGGAGAAUAGUUUGUUC	17	upstream
CEP290-101	+	GGGAGUCACAUGGGAGUCAC	20	downstream
CEP290-102	-	GGGCUCUUCCUGGACCA	17	upstream
CEP290-103	+	GGGUACAGGGGUAAGAGAAA	20	upstream
CEP290-104	-	GGUCCCUGGCUUUUGUUCCU	20	upstream
CEP290-105	-	GUAAAGGUUCAUGAGACUAG	20	downstream
CEP290-106	+	GUAACAUAAUCACCUCUCUU	20	upstream
CEP290-107	+	GUAAGACUGGAGAUAGAGAC	20	downstream
CEP290-108	+	GUACAGGGGUAAGAGAA	17	upstream
CEP290-109	+	GUAGCUUUUGACAGUUUUUA	20	downstream
CEP290-110	+	GUCACAUGGGAGUCACA	17	downstream

CEP290-111	-	GUGGAGAGCCACAGUGCAUG	20	downstream
CEP290-112	-	GUUACAAUCUGUGAAUA	17	upstream
CEP290-113	+	GUUCUGUCCUCAGUAAA	17	upstream
CEP290-114	-	GUUGAGUAUCUCCUGUU	17	downstream
CEP290-115	+	GUUUAGAAUGAUCAUUCUUG	20	downstream
CEP290-116	+	GUUUGUUCUGGGUACAG	17	upstream
CEP290-117	-	UAAAAACUGUCAAAAGCUAC	20	downstream
CEP290-118	+	UAAAAGGUAUAGAGUUCAAG	20	upstream
CEP290-119	+	UAAAUCAUGCAAGUGACCUA	20	upstream
CEP290-120	+	UAAGAUAACUACAAAGGGUC	20	upstream

Table 7B provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the second tier parameters. The targeting domains are within 1000 bp upstream of an *Alu* repeat, within 40bp upstream of mutation, or 1000 bp downstream of the mutation, have good orthogonality, and do not start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. pyogenes* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 7B

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-121	-	AAAAAGGAGCAUGAAAC	17	upstream
CEP290-122	+	AAAACUAAGACACUGCCAAU	20	downstream
CEP290-123	+	AAAAGACUUAUAUUCCAUUA	20	downstream
CEP290-124	-	AAAAGCUACCGGUUACCUGA	20	downstream
CEP290-125	-	AAAAUUAUGCCUAUUUAGUG	20	upstream
CEP290-126	+	AAACAACUGGAAGAGAGAAA	20	downstream
CEP290-127	+	AAACUAAGACACUGCCAAUA	20	downstream
CEP290-128	-	AAACUGUCAAAAGCUAC	17	downstream
CEP290-129	-	AAAGAAAUAGAUGUAGAUUG	20	downstream
CEP290-130	+	AAAGAUGAAAAAUACUCUUU	20	upstream
CEP290-131	-	AAAGCUACCGGUUACCUGAA	20	downstream
CEP290-133	-	AAAUAGAGGCUUAUGGAUUU	20	upstream
CEP290-134	+	AAAUUAUCACCACACUAAAU	20	upstream
CEP290-135	-	AACAAACUAUUCUCCCA	17	upstream
CEP290-136	-	AACAGUAGCUGAAAUAUUAA	20	upstream
CEP290-137	+	AACAUGACUCAUAAUUUAGU	20	upstream

CEP290-138	-	AACUAUUCUCCCAUGGUCCC	20	upstream
CEP290-140	+	AAGACACUGCCAAUAGGGAU	20	downstream
CEP290-141	-	AAGGAAAUACAAAAACUGGA	20	downstream
CEP290-142	+	AAGGUAUAGAGUUCAAG	17	upstream
CEP290-143	-	AAGGUUCAUGAGACUAG	17	downstream
CEP290-144	+	AAUAGUUUGUUCUGGGUACA	20	upstream
CEP290-145	-	AAUAUAAGUCUUUUGAUAUA	20	downstream
CEP290-146	-	AAUAUAUUAUCUAUUUAUAG	20	upstream
CEP290-147	-	AAUAUUGUAAUCAAAGG	17	upstream
CEP290-148	+	AAUAUUUCAGCUACUGU	17	upstream
CEP290-149	-	AAUUAUUGUUGCUUUUUGAG	20	downstream
CEP290-150	+	AAUUCACUGAGCAAAACAAC	20	downstream
CEP290-151	+	ACAAAAGCCAGGGACCA	17	upstream
CEP290-152	+	ACACUGCCAAUAGGGAU	17	downstream
CEP290-153	+	ACAGAGUGCAUCCAUGGUCC	20	upstream
CEP290-154	+	ACAUAAUCACCUCUCUU	17	upstream
CEP290-155	-	ACCAGACAUCUAAGAGAAAA	20	upstream
CEP290-156	-	ACGUGCUCUUUUCUAUAUAU	20	downstream
CEP290-157	+	ACUUUCUAAUGCUGGAG	17	upstream
CEP290-158	+	ACUUUUACCCUUCAGGUAAC	20	downstream
CEP290-159	-	AGAAUAUUGUAAUCAAAGGA	20	upstream
CEP290-160	-	AGACAUCUAAGAGAAAA	17	upstream
CEP290-161	+	AGACUUAUAUUCCAUUA	17	downstream
CEP290-162	+	AGAGGAUAGGACAGAGGACA	20	upstream
CEP290-163	+	AGAUGACAUGAGGUAAGUAG	20	downstream
CEP290-164	+	AGAUGUCUGGUUAAAAG	17	upstream
CEP290-165	+	AGCCUCUAUUUCUGAUG	17	upstream
CEP290-166	-	AGCUACCGGUUACCUGA	17	downstream
CEP290-167	-	AGCUCAAAAGCUUUUGC	17	upstream
CEP290-168	-	AGGAAAUACAAAAACUGGAU	20	downstream
CEP290-169	+	AGGAAGAUGAACAAAUC	17	upstream
CEP290-170	+	AGGACAGAGGACAUGGAGAA	20	upstream
CEP290-171	+	AGGACUUUCUAAUGCUGGAG	20	upstream
CEP290-172	-	AGGCAAGAGACAUCUUG	17	upstream
CEP290-173	-	AGGUAGAAUAUUGUAAUCAA	20	upstream
CEP290-174	-	AGUAGCUGAAAUAUUAA	17	upstream
CEP290-175	+	AGUCACAUGGGAGUCAC	17	downstream
CEP290-176	-	AGUGCAUGUGGUGUCAAAUA	20	downstream
CEP290-177	+	AGUUUGUUCUGGGUACA	17	upstream
CEP290-178	+	AUAAGCCUCUAUUUCUGAUG	20	upstream
CEP290-179	-	AUAAGUCUUUUGAUAUA	17	downstream
CEP290-180	+	AUACAUAAGAAAGAACACUG	20	downstream

CEP290-181	+	AUAGUUUGUUCUGGGUACAG	20	upstream
CEP290-182	-	AUAUCUGUCUUCCUUAA	17	downstream
CEP290-183	-	AUAUUAAGGGCUCUUCC	17	upstream
CEP290-184	-	AUAUUGUAAUCAAAGGA	17	upstream
CEP290-185	+	AUCAUGCAAGUGACCUA	17	upstream
CEP290-186	-	AUCUAAGAUCCUUUCAC	17	upstream
CEP290-187	-	AUCUUCCUCAUCAGAAAUAG	20	upstream
CEP290-188	+	AUGACAUGAGGUAAGUA	17	downstream
CEP290-189	+	AUGACUCAUAAUUUAGU	17	upstream
CEP290-190	-	AUGAGAGUGAUUAGUGG	17	downstream
CEP290-191	+	AUGAGGAAGAUGAACAAAUC	20	upstream
CEP290-192	+	AUGGGAGAAUAGUUUGUUCU	20	upstream
CEP290-193	-	AUUAGCUCAAAAGCUUUUGC	20	upstream
CEP290-194	-	AUUAUGCCUAUUUAGUG	17	upstream
CEP290-195	+	AUUCCAAGGAACAAAAGCCA	20	upstream
CEP290-196	-	AUUGAGGUAGAAUCAAGAAG	20	downstream
CEP290-197	+	AUUUGACACCACAUGCACUG	20	downstream
CEP290-198	+	CAAAAGCCAGGGACCAU	17	upstream
CEP290-199	-	CAACAGUAGCUGAAAUAUUA	20	upstream
CEP290-200	+	CAAGAUGUCUCUUGCCU	17	upstream
CEP290-201	-	CAGAACAAACUAUUCUCCCA	20	upstream
CEP290-202	-	CAGAUUUCAUGUGUGAAGAA	20	downstream
CEP290-204	-	CAGCAUUAGAAAGUCCU	17	upstream
CEP290-205	+	CAGGGGUAAGAGAAAGGGAU	20	upstream
CEP290-206	+	CAGUAAGGAGGAUGUAAGAC	20	downstream
CEP290-207	-	CAGUAGCUGAAAUAUUA	17	upstream
CEP290-208	+	CAUAAGAAAGAACACUG	17	downstream
CEP290-209	+	CAUGGGAGAAUAGUUUGUUC	20	upstream
CEP290-210	+	CAUGGGAGUCACAGGGU	17	downstream
CEP290-211	+	CAUUCCAAGGAACAAAAGCC	20	upstream
CEP290-212	+	CCACAAGAUGUCUCUUGCCU	20	upstream
CEP290-213	-	CCUAGGCAAGAGACAUCUUG	20	upstream
CEP290-214	-	CGUGCUCUUUUCUAUAUAUA	20	downstream
CEP290-215	-	CGUUGUUCUGAGUAGCUUUC	20	upstream
CEP290-216	+	CUAAGACACUGCCAAUA	17	downstream
CEP290-217	+	CUAAUGCUGGAGAGGAU	17	upstream
CEP290-218	+	CUAGAUGACAUGAGGUAAGU	20	downstream
CEP290-219	+	CUAGGACUUUCUAAUGC	17	upstream
CEP290-220	-	CUCAUACCUAUCCCUAU	17	downstream
CEP290-221	-	CUCCAGCAUUAGAAAGUCCU	20	upstream
CEP290-222	-	CUCUAUACCUUUUACUG	17	upstream
CEP290-223	+	CUCUUGCUCUAGAUGACAUG	20	downstream

CEP290-224	-	CUGCUGCUUUUGCCAAAGAG	20	upstream
CEP290-225	-	CUGCUUUUGCCAAAGAG	17	upstream
CEP290-226	-	CUGGCUUUUGUUCCUUGGAA	20	upstream
CEP290-227	+	CUGUAAGAUAACUACAA	17	upstream
CEP290-228	-	CUUAAGCAUACUUUUUUUAA	20	downstream
CEP290-229	+	CUUAAUAUUUCAGCUACUGU	20	upstream
CEP290-231	+	CUUAGAUGUCUGGUUAAAAG	20	upstream
CEP290-232	-	CUUAUCUAAGAUCCUUUCAC	20	upstream
CEP290-233	+	CUUGACUUUUACCCUUC	17	downstream
CEP290-234	+	CUUGUUCUGUCCUCAGUAAA	20	upstream
CEP290-235	+	CUUUUGACAGUUUUUAAGGC	20	downstream

Table 7C provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the third tier parameters. The targeting domains are within 1000bp upstream of an *Alu* repeat, within 40bp upstream of mutation, or 1000bp downstream of the mutation, and start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. pyogenes* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 7C

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-236	-	GAAAUACAAAAACUGGA	17	downstream
CEP290-237	+	GCUUUUGACAGUUUUUA	17	downstream
CEP290-238	+	GGAGAUAGAGACAGGAAUAA	20	downstream
CEP290-239	-	GGAGUGCAGUGGAGUGAUCU	20	downstream
CEP290-240	+	GGGGUAAGAGAAAGGGA	17	upstream
CEP290-241	+	GGGUAAGAGAAAGGGAU	17	upstream
CEP290-242	-	GUCUCACUGUGUUGCCC	17	downstream
CEP290-243	-	GUGCAGUGGAGUGAUCU	17	downstream
CEP290-244	+	GUGUGUGUGUGUGUGUUAUG	20	upstream
CEP290-245	+	GUGUGUGUGUGUUAUGU	17	upstream

Table 7D provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the fourth tier parameters. The targeting domains are within 1000bp upstream of an *Alu* repeat,

within 40bp upstream of mutation, or 1000bp downstream of the mutation, and do not start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. pyogenes* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 7D

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-246	-	AAAUACAAAAACUGGAU	17	downstream
CEP290-247	-	AAGCAUACUUUUUUUAA	17	downstream
CEP290-248	+	AAGGCGGGGAGUCACAU	17	downstream
CEP290-249	+	AAGUAUGCUUAAGAAAAAAA	20	downstream
CEP290-250	+	ACAGAGGACAUGGAGAA	17	upstream
CEP290-251	+	ACAGGGGUAAGAGAAAGGGA	20	upstream
CEP290-253	+	ACUAAGACACUGCCAAU	17	downstream
CEP290-254	+	ACUCCACUGCACUCCAGCCU	20	downstream
CEP290-255	+	AGACUGGAGAUAGAGAC	17	downstream
CEP290-256	-	AGAGUCUCACUGUGUUGCCC	20	downstream
CEP290-257	+	AGAUGAAAAAUACUCUU	17	upstream
CEP290-258	-	AUAUUAUCUAUUUAUAG	17	upstream
CEP290-259	-	AUUUCAUGUGUGAAGAA	17	downstream
CEP290-260	-	AUUUUUUAUUAUCUUUAUUG	20	downstream
CEP290-261	+	CAACUGGAAGAGAGAAA	17	downstream
CEP290-262	+	CACUCCACUGCACUCCAGCC	20	downstream
CEP290-263	-	CACUGUGUUGCCCAGGC	17	downstream
CEP290-264	+	CCAAGGAACAAAAGCCA	17	upstream
CEP290-265	+	CCACUGCACUCCAGCCU	17	downstream
CEP290-266	-	CCCAGGCUGGAGUGCAG	17	downstream
CEP290-267	-	CCCUGGCUUUUGUUCCU	17	upstream
CEP290-268	+	CGCUUGAACCUGGGAGGCAG	20	downstream
CEP290-269	-	UAAGGAAAUACAAAAAC	17	downstream
CEP290-270	-	UAAUAAGGAAAUACAAAAAC	20	downstream
CEP290-271	-	UACUGCAACCUCUGCCUCCC	20	downstream
CEP290-272	+	UAUGCUUAAGAAAAAAA	17	downstream
CEP290-273	+	UCAUUCUUGUGGCAGUAAGG	20	downstream
CEP290-274	+	UCCACUGCACUCCAGCC	17	downstream
CEP290-275	-	UCUCACUGUGUUGCCCAGGC	20	downstream
CEP290-276	+	UGAACAAGUUUUGAAAC	17	downstream
CEP290-277	-	UGCAACCUCUGCCUCCC	17	downstream
CEP290-278	+	UGUGUGUGUGUGUGUUAUGU	20	upstream

CEP290-279	+	UGUGUGUGUGUGUUAUG	17	upstream
CEP290-280	+	UUCUUGUGGCAGUAAGG	17	downstream
CEP290-281	+	UUGAACCUGGGAGGCAG	17	downstream
CEP290-282	-	UUGCCCAGGCUGGAGUGCAG	20	downstream
CEP290-283	-	UUUUAUUAUCUUUAUUG	17	downstream

Table 8A provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the first tier parameters. The targeting domains are within 1000 bp upstream of an *Alu* repeat, within 40bp upstream of mutation, or 1000 bp downstream of the mutation, have good orthogonality, start with G and PAM is NNGRRT. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. aureus* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 8A

1st Tier				
gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-284	+	GCUAAAUCAUGCAAGUGACCUAAG	24	upstream
CEP290-487	-	GGUCACUUGCAUGAUUUAG	19	upstream
CEP290-486	-	GUCACUUGCAUGAUUUAG	18	upstream
CEP290-285	+	GCCUAGGACUUUCUAAUGCUGGA	23	upstream
CEP290-479	+	GGACUUUCUAAUGCUGGA	18	upstream
CEP290-286	+	GGGACCAUGGGAGAAUAGUUUGUU	24	upstream
CEP290-287	+	GGACCAUGGGAGAAUAGUUUGUU	23	upstream
CEP290-288	+	GACCAUGGGAGAAUAGUUUGUU	22	upstream
CEP290-289	-	GGUCCCUGGCUUUUGUUCCUUGGA	24	upstream
CEP290-290	-	GUCCCUGGCUUUUGUUCCUUGGA	23	upstream
CEP290-374	-	GAAAACGUUGUUCUGAGUAGCUUU	24	upstream
CEP290-478	-	GUUGUUCUGAGUAGCUUU	18	upstream
CEP290-489	-	GGUCCCUGGCUUUUGUUCCU	20	upstream
CEP290-488	-	GUCCCUGGCUUUUGUUCCU	19	upstream
CEP290-291	-	GACAUCUUGUGGAUAAUGUAUCA	23	upstream
CEP290-292	-	GUCCUAGGCAAGAGACAUCUU	21	upstream
CEP290-293	+	GCCAGCAAAAGCUUUUGAGCUAA	23	upstream
CEP290-481	+	GCAAAAGCUUUUGAGCUAA	19	upstream
CEP290-294	+	GAUCUUAUUCUACUCCUGUGA	21	upstream
CEP290-295	-	GCUUUCAGGAUUCCUACUAAAUU	23	upstream

CEP290-323	+	GUUCUGUCCUCAGUAAAAGGUA	22	upstream
CEP290-480	+	GAACAACGUUUUCAUUUA	18	upstream
CEP290-296	-	GUAGAAUAUCAUAAGUUACAAUCU	24	upstream
CEP290-297	-	GAAUAUCAUAAGUUACAAUCU	21	upstream
CEP290-298	+	GUGGCUGUAAGAUAACUACA	20	upstream
CEP290-299	+	GGCUGUAAGAUAACUACA	18	upstream
CEP290-300	-	GUUUAACGUUAUCAUUUUCCCA	22	upstream
CEP290-301	+	GUAAGAGAAAGGGAUGGGCACUUA	24	upstream
CEP290-492	+	GAGAAAGGGAUGGGCACUUA	20	upstream
CEP290-491	+	GAAAGGGAUGGGCACUUA	18	upstream
CEP290-483	-	GUAAAUGAAAACGUUGUU	18	upstream
CEP290-302	+	GAUAAACAUGACUCAUAAUUUAGU	24	upstream
CEP290-303	+	GGAACAAAAGCCAGGGACCAUGG	23	upstream
CEP290-304	+	GAACAAAAGCCAGGGACCAUGG	22	upstream
CEP290-305	+	GGGAGAAUAGUUUGUUCUGGGUAC	24	upstream
CEP290-306	+	GGAGAAUAGUUUGUUCUGGGUAC	23	upstream
CEP290-307	+	GAGAAUAGUUUGUUCUGGGUAC	22	upstream
CEP290-490	+	GAAUAGUUUGUUCUGGGUAC	20	upstream
CEP290-482	-	GAAAUAGAGGCUUAUGGAUU	20	upstream
CEP290-308	+	GUUCUGGGUACAGGGGUAAGAGAA	24	upstream
CEP290-494	+	GGGUACAGGGGUAAGAGAA	19	upstream
CEP290-493	+	GGUACAGGGGUAAGAGAA	18	upstream
CEP290-309	-	GUAAAUUCUCAUCAUUUUUUAUUG	24	upstream
CEP290-310	+	GGAGAGGAUAGGACAGAGGACAUG	24	upstream
CEP290-311	+	GAGAGGAUAGGACAGAGGACAUG	23	upstream
CEP290-313	+	GAGGAUAGGACAGAGGACAUG	21	upstream
CEP290-485	+	GGAUAGGACAGAGGACAUG	19	upstream
CEP290-484	+	GAUAGGACAGAGGACAUG	18	upstream
CEP290-314	-	GAAUAAAUGUAGAAUUUUAAUG	22	upstream
CEP290-64	-	GUCAAAAGCUACCGGUUACCUG	22	downstream
CEP290-315	+	GUUUUUAAGGCGGGGAGUCACAU	23	downstream
CEP290-203	-	GUCUUACAUCCUCCUUACUGCCAC	24	downstream
CEP290-316	+	GAGUCACAGGGUAGGAUUCAUGUU	24	downstream
CEP290-317	+	GUCACAGGGUAGGAUUCAUGUU	22	downstream
CEP290-318	-	GGCACAGAGUUCAAGCUAAUACAU	24	downstream
CEP290-319	-	GCACAGAGUUCAAGCUAAUACAU	23	downstream
CEP290-505	-	GAGUUCAAGCUAAUACAU	18	downstream
CEP290-496	+	GAUGCAGAACUAGUGUAGAC	20	downstream
CEP290-320	-	GUGUUGAGUAUCUCCUGUUUGGCA	24	downstream
CEP290-321	-	GUUGAGUAUCUCCUGUUUGGCA	22	downstream
CEP290-504	-	GAGUAUCUCCUGUUUGGCA	19	downstream
CEP290-322	-	GAAAAUCAGAUUUCAUGUGUG	21	downstream

CEP290-324	-	GCCACAAGAAUGAUCAUUCUAAAC	24	downstream
CEP290-325	+	GGCGGGGAGUCACAUGGGAGUCA	23	downstream
CEP290-326	+	GCGGGGAGUCACAUGGGAGUCA	22	downstream
CEP290-499	+	GGGGAGUCACAUGGGAGUCA	20	downstream
CEP290-498	+	GGGAGUCACAUGGGAGUCA	19	downstream
CEP290-497	+	GGAGUCACAUGGGAGUCA	18	downstream
CEP290-327	+	GCUUUUGACAGUUUUUAAGGCG	22	downstream
CEP290-328	+	GAUCAUUCUUGUGGCAGUAAG	21	downstream
CEP290-329	-	GAGCAAGAGAUGAACUAG	18	downstream
CEP290-500	+	GCCUGAACAAGUUUUGAAAC	20	downstream
CEP290-330	-	GUAGAUUGAGGUAGAAUCAAGAA	23	downstream
CEP290-506	-	GAUUGAGGUAGAAUCAAGAA	20	downstream
CEP290-331	+	GGAUGUAAGACUGGAGAUAGAGAC	24	downstream
CEP290-332	+	GAUGUAAGACUGGAGAUAGAGAC	23	downstream
CEP290-503	+	GUAAGACUGGAGAUAGAGAC	20	downstream
CEP290-333	+	GGGAGUCACAUGGGAGUCACAGGG	24	downstream
CEP290-334	+	GGAGUCACAUGGGAGUCACAGGG	23	downstream
CEP290-335	+	GAGUCACAUGGGAGUCACAGGG	22	downstream
CEP290-502	+	GUCACAUGGGAGUCACAGGG	20	downstream
CEP290-336	-	GUUUACAUAUCUGUCUUCCUUAA	23	downstream
CEP290-507	-	GAUUUCAUGUGUGAAGAA	18	downstream

Table 8B provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the second tier parameters. The targeting domains are within 1000 bp upstream of an *Alu* repeat, within 40bp upstream of mutation, or 1000 bp downstream of the mutation, and have good orthogonality. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. aureus* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 8B

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-337	+	AAAUCAUGCAAGUGACCUAAG	21	upstream
CEP290-338	+	AAUCAUGCAAGUGACCUAAG	20	upstream
CEP290-339	+	AUCAUGCAAGUGACCUAAG	19	upstream
CEP290-340	-	AGGUCACUUGCAUGAUUUAG	20	upstream
CEP290-341	-	AAUAUUAAGGGCUCUUCCUGGACC	24	upstream

CEP290-342	-	AUAUUAAGGGCUCUUCCUGGACC	23	upstream
CEP290-343	-	AUUAAGGGCUCUUCCUGGACC	21	upstream
CEP290-344	-	AAGGGCUCUUCCUGGACC	18	upstream
CEP290-345	+	AGGACUUUCUAAUGCUGGA	19	upstream
CEP290-346	+	ACCAUGGGAGAAUAGUUUGUU	21	upstream
CEP290-347	+	AUGGGAGAAUAGUUUGUU	18	upstream
CEP290-348	+	ACUCCUGUGAAAGGAUCUUAGAU	23	upstream
CEP290-349	-	AAAACGUUGUUCUGAGUAGCUUU	23	upstream
CEP290-350	-	AAACGUUGUUCUGAGUAGCUUU	22	upstream
CEP290-351	-	AACGUUGUUCUGAGUAGCUUU	21	upstream
CEP290-352	-	ACGUUGUUCUGAGUAGCUUU	20	upstream
CEP290-353	-	AUUUAUAGUGGCUGAAUGACUU	22	upstream
CEP290-354	-	AUAGUGGCUGAAUGACUU	18	upstream
CEP290-355	-	AUGGUCCCUGGCUUUUGUUCCU	22	upstream
CEP290-356	-	AGACAUCUUGUGGAUAAUGUAUCA	24	upstream
CEP290-357	-	ACAUCUUGUGGAUAAUGUAUCA	22	upstream
CEP290-358	-	AUCUUGUGGAUAAUGUAUCA	20	upstream
CEP290-359	-	AAAGUCCUAGGCAAGAGACAUCUU	24	upstream
CEP290-360	-	AAGUCCUAGGCAAGAGACAUCUU	23	upstream
CEP290-361	-	AGUCCUAGGCAAGAGACAUCUU	22	upstream
CEP290-362	+	AGCCAGCAAAAGCUUUUGAGCUAA	24	upstream
CEP290-363	+	AGCAAAAGCUUUUGAGCUAA	20	upstream
CEP290-364	+	AGAUCUUAUUCUACUCCUGUGA	22	upstream
CEP290-365	+	AUCUUAUUCUACUCCUGUGA	20	upstream
CEP290-366	-	AUCUAAGAUCCUUUCACAGGAG	22	upstream
CEP290-369	-	AAGAUCCUUUCACAGGAG	18	upstream
CEP290-370	-	AGCUUUCAGGAUUCCUACUAAAUU	24	upstream
CEP290-371	+	ACUCAGAACAACGUUUUCAUUUA	23	upstream
CEP290-372	+	AGAACAACGUUUUCAUUUA	19	upstream
CEP290-373	-	AGAAUAUCAUAAGUUACAAUCU	22	upstream
CEP290-375	-	AAUAUCAUAAGUUACAAUCU	20	upstream
CEP290-376	-	AUAUCAUAAGUUACAAUCU	19	upstream
CEP290-377	+	AAGUGGCUGUAAGAUAACUACA	22	upstream
CEP290-378	+	AGUGGCUGUAAGAUAACUACA	21	upstream
CEP290-379	-	AUGUUUAACGUUAUCAUUUUCCCA	24	upstream
CEP290-380	-	AACGUUAUCAUUUUCCCA	18	upstream
CEP290-381	+	AAGAGAAAGGGAUGGGCACUUA	22	upstream
CEP290-382	+	AGAGAAAGGGAUGGGCACUUA	21	upstream
CEP290-383	+	AGAAAGGGAUGGGCACUUA	19	upstream
CEP290-384	-	AUUCAGUAAAUGAAAACGUUGUU	23	upstream
CEP290-385	-	AGUAAAUGAAAACGUUGUU	19	upstream
CEP290-386	+	AUAAACAUGACUCAUAAUUUAGU	23	upstream

CEP290-387	+	AAACAUGACUCAUAAUUUAGU	21	upstream
CEP290-388	+	AACAUGACUCAUAAUUUAGU	20	upstream
CEP290-389	+	ACAUGACUCAUAAUUUAGU	19	upstream
CEP290-390	-	AUUCUUAUCUAAGAUCCUUUCAC	23	upstream
CEP290-391	+	AGGAACAAAAGCCAGGGACCAUGG	24	upstream
CEP290-392	+	AACAAAAGCCAGGGACCAUGG	21	upstream
CEP290-393	+	ACAAAAGCCAGGGACCAUGG	20	upstream
CEP290-394	+	AAAAGCCAGGGACCAUGG	18	upstream
CEP290-395	+	AGAAUAGUUUGUUCUGGGUAC	21	upstream
CEP290-396	+	AAUAGUUUGUUCUGGGUAC	19	upstream
CEP290-397	+	AUAGUUUGUUCUGGGUAC	18	upstream
CEP290-398	-	AUCAGAAAUAGAGGCUUAUGGAUU	24	upstream
CEP290-399	-	AGAAAUAGAGGCUUAUGGAUU	21	upstream
CEP290-400	-	AAAUAGAGGCUUAUGGAUU	19	upstream
CEP290-401	-	AAUAGAGGCUUAUGGAUU	18	upstream
CEP290-402	-	AAUAUAUUAUCUAUUUAUAGUGG	23	upstream
CEP290-403	-	AUAUAUUAUCUAUUUAUAGUGG	22	upstream
CEP290-404	-	AUAUUAUCUAUUUAUAGUGG	20	upstream
CEP290-405	-	AUUAUCUAUUUAUAGUGG	18	upstream
CEP290-406	-	AAAUUCUCAUCAUUUUUUAUUG	22	upstream
CEP290-407	-	AAUUCUCAUCAUUUUUUAUUG	21	upstream
CEP290-408	-	AUUCUCAUCAUUUUUUAUUG	20	upstream
CEP290-409	+	AGAGGAUAGGACAGAGGACAUG	22	upstream
CEP290-410	+	AGGAUAGGACAGAGGACAUG	20	upstream
CEP290-411	-	AGAAUAAAUGUAGAAUUUUAAUG	23	upstream
CEP290-412	-	AAUAAAUGUAGAAUUUUAAUG	21	upstream
CEP290-413	-	AUAAAUGUAGAAUUUUAAUG	20	upstream
CEP290-414	-	AAAUGUAGAAUUUUAAUG	18	upstream
CEP290-415	-	AUUUUUUAUUGUAGAAUAAAUG	22	upstream
CEP290-416	+	CUAAAUCAUGCAAGUGACCUAAG	23	upstream
CEP290-417	-	CCUUAGGUCACUUGCAUGAUUUAG	24	upstream
CEP290-418	-	CUUAGGUCACUUGCAUGAUUUAG	23	upstream
CEP290-419	+	CCUAGGACUUUCUAAUGCUGGA	22	upstream
CEP290-420	+	CUAGGACUUUCUAAUGCUGGA	21	upstream
CEP290-421	+	CCAUGGGAGAAUAGUUUGUU	20	upstream
CEP290-422	+	CAUGGGAGAAUAGUUUGUU	19	upstream
CEP290-423	+	CUCCUGUGAAAGGAUCUUAGAU	22	upstream
CEP290-424	+	CCUGUGAAAGGAUCUUAGAU	20	upstream
CEP290-426	+	CUGUGAAAGGAUCUUAGAU	19	upstream
CEP290-427	-	CCCUGGCUUUUGUUCCUUGGA	21	upstream
CEP290-428	-	CCUGGCUUUUGUUCCUUGGA	20	upstream
CEP290-429	-	CUGGCUUUUGUUCCUUGGA	19	upstream

CEP290-430	-	CGUUGUUCUGAGUAGCUUU	19	upstream
CEP290-431	-	CUAUUUAUAGUGGCUGAAUGACUU	24	upstream
CEP290-432	-	CCAUGGUCCCUGGCUUUUGUUCCU	24	upstream
CEP290-433	-	CAUGGUCCCUGGCUUUUGUUCCU	23	upstream
CEP290-434	-	CAUCUUGUGGAUAAUGUAUCA	21	upstream
CEP290-435	-	CUUGUGGAUAAUGUAUCA	18	upstream
CEP290-437	-	CCUAGGCAAGAGACAUCUU	19	upstream
CEP290-438	-	CUAGGCAAGAGACAUCUU	18	upstream
CEP290-439	+	CCAGCAAAAGCUUUUGAGCUAA	22	upstream
CEP290-440	+	CAGCAAAAGCUUUUGAGCUAA	21	upstream
CEP290-441	+	CAAAAGCUUUUGAGCUAA	18	upstream
CEP290-442	+	CUUAUUCUACUCCUGUGA	18	upstream
CEP290-443	-	CUAAGAUCCUUUCACAGGAG	20	upstream
CEP290-444	-	CUUCCUCAUCAGAAAUAGAGGCUU	24	upstream
CEP290-445	-	CCUCAUCAGAAAUAGAGGCUU	21	upstream
CEP290-446	-	CUCAUCAGAAAUAGAGGCUU	20	upstream
CEP290-447	-	CAUCAGAAAUAGAGGCUU	18	upstream
CEP290-448	-	CUUUCAGGAUUCCUACUAAAUU	22	upstream
CEP290-449	-	CAGGAUUCCUACUAAAUU	18	upstream
CEP290-450	+	CUGUCCUCAGUAAAAGGUA	19	upstream
CEP290-451	+	CUCAGAACAACGUUUUCAUUUA	22	upstream
CEP290-452	+	CAGAACAACGUUUUCAUUUA	20	upstream
CEP290-453	+	CAAGUGGCUGUAAGAUAACUACA	23	upstream
CEP290-454	-	CAUUCAGUAAAUGAAAACGUUGUU	24	upstream
CEP290-457	-	CAGUAAAUGAAAACGUUGUU	20	upstream
CEP290-458	+	CAUGACUCAUAAUUUAGU	18	upstream
CEP290-459	-	CUUAUCUAAGAUCCUUUCAC	20	upstream
CEP290-460	+	CAAAAGCCAGGGACCAUGG	19	upstream
CEP290-461	-	CAGAAAUAGAGGCUUAUGGAUU	22	upstream
CEP290-462	+	CUGGGUACAGGGGUAAGAGAA	21	upstream
CEP290-463	-	CAAUAUAUUAUCUAUUUAUAGUGG	24	upstream
CEP290-464	-	CAUUUUUUAUUGUAGAAUAAAUG	23	upstream
CEP290-465	+	UAAAUCAUGCAAGUGACCUAAG	22	upstream
CEP290-466	+	UCAUGCAAGUGACCUAAG	18	upstream
CEP290-467	-	UUAGGUCACUUGCAUGAUUUAG	22	upstream
CEP290-468	-	UAGGUCACUUGCAUGAUUUAG	21	upstream
CEP290-469	-	UAUUAAGGGCUCUUCCUGGACC	22	upstream
CEP290-470	-	UUAAGGGCUCUUCCUGGACC	20	upstream
CEP290-471	-	UAAGGGCUCUUCCUGGACC	19	upstream
CEP290-472	+	UGCCUAGGACUUUCUAAUGCUGGA	24	upstream
CEP290-473	+	UAGGACUUUCUAAUGCUGGA	20	upstream
CEP290-474	+	UACUCCUGUGAAAGGAUCUUAGAU	24	upstream

CEP290-475	+	UCCUGUGAAAGGAUCUUAGAU	21	upstream
CEP290-476	+	UGUGAAAGGAUCUUAGAU	18	upstream
CEP290-477	-	UCCCUGGCUUUUGUUCCUUGGA	22	upstream
CEP290-515	-	UGGCUUUUGUUCCUUGGA	18	upstream
CEP290-516	-	UAUUUAUAGUGGCUGAAUGACUU	23	upstream
CEP290-517	-	UUUAUAGUGGCUGAAUGACUU	21	upstream
CEP290-518	-	UUAUAGUGGCUGAAUGACUU	20	upstream
CEP290-519	-	UAUAGUGGCUGAAUGACUU	19	upstream
CEP290-520	-	UGGUCCCUGGCUUUUGUUCCU	21	upstream
CEP290-521	-	UCCCUGGCUUUUGUUCCU	18	upstream
CEP290-522	-	UCUUGUGGAUAAUGUAUCA	19	upstream
CEP290-523	-	UCCUAGGCAAGAGACAUCUU	20	upstream
CEP290-524	+	UUAGAUCUUAUUCUACUCCUGUGA	24	upstream
CEP290-525	+	UAGAUCUUAUUCUACUCCUGUGA	23	upstream
CEP290-526	+	UCUUAUUCUACUCCUGUGA	19	upstream
CEP290-527	-	UUAUCUAAGAUCCUUUCACAGGAG	24	upstream
CEP290-528	-	UAUCUAAGAUCCUUUCACAGGAG	23	upstream
CEP290-529	-	UCUAAGAUCCUUUCACAGGAG	21	upstream
CEP290-530	-	UAAGAUCCUUUCACAGGAG	19	upstream
CEP290-531	-	UUCCUCAUCAGAAAUAGAGGCUU	23	upstream
CEP290-532	-	UCCUCAUCAGAAAUAGAGGCUU	22	upstream
CEP290-533	-	UCAUCAGAAAUAGAGGCUU	19	upstream
CEP290-534	-	UUUCAGGAUUCCUACUAAAUU	21	upstream
CEP290-535	-	UUCAGGAUUCCUACUAAAUU	20	upstream
CEP290-536	-	UCAGGAUUCCUACUAAAUU	19	upstream
CEP290-537	+	UUGUUCUGUCCUCAGUAAAAGGUA	24	upstream
CEP290-538	+	UGUUCUGUCCUCAGUAAAAGGUA	23	upstream
CEP290-539	+	UUCUGUCCUCAGUAAAAGGUA	21	upstream
CEP290-540	+	UCUGUCCUCAGUAAAAGGUA	20	upstream
CEP290-541	+	UGUCCUCAGUAAAAGGUA	18	upstream
CEP290-542	+	UACUCAGAACAACGUUUUCAUUUA	24	upstream
CEP290-543	+	UCAGAACAACGUUUUCAUUUA	21	upstream
CEP290-544	-	UAGAAUAUCAUAAGUUACAAUCU	23	upstream
CEP290-545	-	UAUCAUAAGUUACAAUCU	18	upstream
CEP290-546	+	UCAAGUGGCUGUAAGAUAACUACA	24	upstream
CEP290-547	+	UGGCUGUAAGAUAACUACA	19	upstream
CEP290-548	-	UGUUUAACGUUAUCAUUUUCCCA	23	upstream
CEP290-549	-	UUUAACGUUAUCAUUUUCCCA	21	upstream
CEP290-550	-	UUAACGUUAUCAUUUUCCCA	20	upstream
CEP290-551	-	UAACGUUAUCAUUUUCCCA	19	upstream
CEP290-552	+	UAAGAGAAAGGGAUGGGCACUUA	23	upstream
CEP290-553	-	UUCAGUAAAUGAAAACGUUGUU	22	upstream

CEP290-554	-	UCAGUAAAUGAAAACGUUGUU	21	upstream
CEP290-555	+	UAAACAUGACUCAUAAUUUAGU	22	upstream
CEP290-556	-	UAUUCUUAUCUAAGAUCCUUUCAC	24	upstream
CEP290-557	-	UUCUUAUCUAAGAUCCUUUCAC	22	upstream
CEP290-558	-	UCUUAUCUAAGAUCCUUUCAC	21	upstream
CEP290-559	-	UUAUCUAAGAUCCUUUCAC	19	upstream
CEP290-560	-	UAUCUAAGAUCCUUUCAC	18	upstream
CEP290-561	-	UCAGAAAUAGAGGCUUAUGGAUU	23	upstream
CEP290-562	+	UUCUGGGUACAGGGGUAAGAGAA	23	upstream
CEP290-563	+	UCUGGGUACAGGGGUAAGAGAA	22	upstream
CEP290-564	+	UGGGUACAGGGGUAAGAGAA	20	upstream
CEP290-565	-	UAUAUUAUCUAUUUAUAGUGG	21	upstream
CEP290-566	-	UAUUAUCUAUUUAUAGUGG	19	upstream
CEP290-567	-	UAAAUUCUCAUCAUUUUUUAUUG	23	upstream
CEP290-568	-	UUCUCAUCAUUUUUUAUUG	19	upstream
CEP290-569	-	UCUCAUCAUUUUUUAUUG	18	upstream
CEP290-570	-	UAGAAUAAAUGUAGAAUUUUAAUG	24	upstream
CEP290-571	-	UAAAUGUAGAAUUUUAAUG	19	upstream
CEP290-572	-	UCAUUUUUUAUUGUAGAAUAAAUG	24	upstream
CEP290-573	-	UUUUUUAUUGUAGAAUAAAUG	21	upstream
CEP290-574	-	UUUUUAUUGUAGAAUAAAUG	20	upstream
CEP290-575	-	UUUUAUUGUAGAAUAAAUG	19	upstream
CEP290-576	-	UUUAUUGUAGAAUAAAUG	18	upstream
CEP290-577	-	AAAAGCUACCGGUUACCUG	19	downstream
CEP290-578	-	AAAGCUACCGGUUACCUG	18	downstream
CEP290-579	+	AGUUUUUAAGGCGGGGAGUCACAU	24	downstream
CEP290-580	-	ACAUCCUCCUUACUGCCAC	19	downstream
CEP290-581	+	AGUCACAGGGUAGGAUUCAUGUU	23	downstream
CEP290-582	+	ACAGGGUAGGAUUCAUGUU	19	downstream
CEP290-583	-	ACAGAGUUCAAGCUAAUACAU	21	downstream
CEP290-584	-	AGAGUUCAAGCUAAUACAU	19	downstream
CEP290-585	+	AUAAGAUGCAGAACUAGUGUAGAC	24	downstream
CEP290-586	+	AAGAUGCAGAACUAGUGUAGAC	22	downstream
CEP290-587	+	AGAUGCAGAACUAGUGUAGAC	21	downstream
CEP290-588	+	AUGCAGAACUAGUGUAGAC	19	downstream
CEP290-589	-	AGUAUCUCCUGUUUGGCA	18	downstream
CEP290-590	-	ACGAAAAUCAGAUUUCAUGUGUG	23	downstream
CEP290-591	-	AAAAUCAGAUUUCAUGUGUG	20	downstream
CEP290-592	-	AAAUCAGAUUUCAUGUGUG	19	downstream
CEP290-593	-	AAUCAGAUUUCAUGUGUG	18	downstream
CEP290-594	-	ACAAGAAUGAUCAUUCUAAAC	21	downstream
CEP290-595	-	AAGAAUGAUCAUUCUAAAC	19	downstream

CEP290-596	-	AGAAUGAUCAUUCUAAAC	18	downstream
CEP290-597	+	AGGCGGGGAGUCACAUGGGAGUCA	24	downstream
CEP290-598	+	AGCUUUUGACAGUUUUUAAGGCG	23	downstream
CEP290-599	+	AAUGAUCAUUCUUGUGGCAGUAAG	24	downstream
CEP290-600	+	AUGAUCAUUCUUGUGGCAGUAAG	23	downstream
CEP290-601	+	AUCAUUCUUGUGGCAGUAAG	20	downstream
CEP290-602	-	AUCUAGAGCAAGAGAUGAACUAG	23	downstream
CEP290-603	-	AGAGCAAGAGAUGAACUAG	19	downstream
CEP290-604	+	AAUGCCUGAACAAGUUUUGAAAC	23	downstream
CEP290-605	+	AUGCCUGAACAAGUUUUGAAAC	22	downstream
CEP290-606	-	AGAUUGAGGUAGAAUCAAGAA	21	downstream
CEP290-607	-	AUUGAGGUAGAAUCAAGAA	19	downstream
CEP290-608	+	AUGUAAGACUGGAGAUAGAGAC	22	downstream
CEP290-609	+	AAGACUGGAGAUAGAGAC	18	downstream
CEP290-610	+	AGUCACAUGGGAGUCACAGGG	21	downstream
CEP290-611	-	ACAUAUCUGUCUUCCUUAA	19	downstream
CEP290-612	-	AAAUCAGAUUUCAUGUGUGAAGAA	24	downstream
CEP290-613	-	AAUCAGAUUUCAUGUGUGAAGAA	23	downstream
CEP290-614	-	AUCAGAUUUCAUGUGUGAAGAA	22	downstream
CEP290-615	-	AGAUUUCAUGUGUGAAGAA	19	downstream
CEP290-616	+	AAAUAAAACUAAGACACUGCCAAU	24	downstream
CEP290-617	+	AAUAAAACUAAGACACUGCCAAU	23	downstream
CEP290-618	+	AUAAAACUAAGACACUGCCAAU	22	downstream
CEP290-619	+	AAAACUAAGACACUGCCAAU	20	downstream
CEP290-620	+	AAACUAAGACACUGCCAAU	19	downstream
CEP290-621	+	AACUAAGACACUGCCAAU	18	downstream
CEP290-622	-	AACUAUUUAAUUUGUUUCUGUGUG	24	downstream
CEP290-623	-	ACUAUUUAAUUUGUUUCUGUGUG	23	downstream
CEP290-624	-	AUUUAAUUUGUUUCUGUGUG	20	downstream
CEP290-625	-	CUGUCAAAAGCUACCGGUUACCUG	24	downstream
CEP290-626	-	CAAAAGCUACCGGUUACCUG	20	downstream
CEP290-627	-	CUUACAUCCUCCUUACUGCCAC	22	downstream
CEP290-628	-	CAUCCUCCUUACUGCCAC	18	downstream
CEP290-629	+	CACAGGGUAGGAUUCAUGUU	20	downstream
CEP290-630	+	CAGGGUAGGAUUCAUGUU	18	downstream
CEP290-631	-	CACAGAGUUCAAGCUAAUACAU	22	downstream
CEP290-632	-	CAGAGUUCAAGCUAAUACAU	20	downstream
CEP290-633	-	CACGAAAAUCAGAUUUCAUGUGUG	24	downstream
CEP290-634	-	CGAAAAUCAGAUUUCAUGUGUG	22	downstream
CEP290-635	-	CCACAAGAAUGAUCAUUCUAAAC	23	downstream
CEP290-636	-	CACAAGAAUGAUCAUUCUAAAC	22	downstream
CEP290-637	-	CAAGAAUGAUCAUUCUAAAC	20	downstream

CEP290-638	+	CGGGGAGUCACAUGGGAGUCA	21	downstream
CEP290-639	+	CUUUUGACAGUUUUUAAGGCG	21	downstream
CEP290-640	+	CAUUCUUGUGGCAGUAAG	18	downstream
CEP290-641	-	CAUCUAGAGCAAGAGAUGAACUAG	24	downstream
CEP290-642	-	CUAGAGCAAGAGAUGAACUAG	21	downstream
CEP290-643	+	CCUGAACAAGUUUUGAAAC	19	downstream
CEP290-644	+	CUGAACAAGUUUUGAAAC	18	downstream
CEP290-645	-	CUCUCUUCCAGUUGUUUUGCUCA	23	downstream
CEP290-646	-	CUCUUCCAGUUGUUUUGCUCA	21	downstream
CEP290-647	-	CUUCCAGUUGUUUUGCUCA	19	downstream
CEP290-648	+	CACAUGGGAGUCACAGGG	18	downstream
CEP290-649	-	CAUAUCUGUCUUCCUUAA	18	downstream
CEP290-650	-	CAGAUUUCAUGUGUGAAGAA	20	downstream
CEP290-651	-	CUAUUUAAUUUGUUUCUGUGUG	22	downstream
CEP290-652	-	UGUCAAAAGCUACCGGUUACCUG	23	downstream
CEP290-653	-	UCAAAAGCUACCGGUUACCUG	21	downstream
CEP290-654	+	UUUUUAAGGCGGGGAGUCACAU	22	downstream
CEP290-655	+	UUUUAAGGCGGGGAGUCACAU	21	downstream
CEP290-656	+	UUUAAGGCGGGGAGUCACAU	20	downstream
CEP290-657	+	UUAAGGCGGGGAGUCACAU	19	downstream
CEP290-658	+	UAAGGCGGGGAGUCACAU	18	downstream
CEP290-659	-	UCUUACAUCCUCCUUACUGCCAC	23	downstream
CEP290-660	-	UUACAUCCUCCUUACUGCCAC	21	downstream
CEP290-661	-	UACAUCCUCCUUACUGCCAC	20	downstream
CEP290-662	+	UCACAGGGUAGGAUUCAUGUU	21	downstream
CEP290-663	+	UAAGAUGCAGAACUAGUGUAGAC	23	downstream
CEP290-664	+	UGCAGAACUAGUGUAGAC	18	downstream
CEP290-665	-	UGUUGAGUAUCUCCUGUUUGGCA	23	downstream
CEP290-666	-	UUGAGUAUCUCCUGUUUGGCA	21	downstream
CEP290-667	-	UGAGUAUCUCCUGUUUGGCA	20	downstream
CEP290-668	+	UAGCUUUUGACAGUUUUUAAGGCG	24	downstream
CEP290-669	+	UUUUGACAGUUUUUAAGGCG	20	downstream
CEP290-670	+	UUUGACAGUUUUUAAGGCG	19	downstream
CEP290-671	+	UUGACAGUUUUUAAGGCG	18	downstream
CEP290-672	+	UGAUCAUUCUUGUGGCAGUAAG	22	downstream
CEP290-673	+	UCAUUCUUGUGGCAGUAAG	19	downstream
CEP290-674	-	UCUAGAGCAAGAGAUGAACUAG	22	downstream
CEP290-675	-	UAGAGCAAGAGAUGAACUAG	20	downstream
CEP290-676	+	UAAUGCCUGAACAAGUUUUGAAAC	24	downstream
CEP290-677	+	UGCCUGAACAAGUUUUGAAAC	21	downstream
CEP290-678	-	UGUAGAUUGAGGUAGAAUCAAGAA	24	downstream
CEP290-679	-	UAGAUUGAGGUAGAAUCAAGAA	22	downstream

CEP290-680	-	UUGAGGUAGAAUCAAGAA	18	downstream
CEP290-681	+	UGUAAGACUGGAGAUAGAGAC	21	downstream
CEP290-682	+	UAAGACUGGAGAUAGAGAC	19	downstream
CEP290-683	-	UCUCUCUUCCAGUUGUUUUGCUCA	24	downstream
CEP290-684	-	UCUCUUCCAGUUGUUUUGCUCA	22	downstream
CEP290-685	-	UCUUCCAGUUGUUUUGCUCA	20	downstream
CEP290-686	-	UUCCAGUUGUUUUGCUCA	18	downstream
CEP290-687	+	UCACAUGGGAGUCACAGGG	19	downstream
CEP290-688	-	UGUUUACAUAUCUGUCUUCCUUAA	24	downstream
CEP290-689	-	UUUACAUAUCUGUCUUCCUUAA	22	downstream
CEP290-690	-	UUACAUAUCUGUCUUCCUUAA	21	downstream
CEP290-691	-	UACAUAUCUGUCUUCCUUAA	20	downstream
CEP290-692	-	UCAGAUUUCAUGUGUGAAGAA	21	downstream
CEP290-693	+	UAAAACUAAGACACUGCCAAU	21	downstream
CEP290-694	-	UAUUUAAUUUGUUUCUGUGUG	21	downstream
CEP290-695	-	UUUAAUUUGUUUCUGUGUG	19	downstream
CEP290-696	-	UUAAUUUGUUUCUGUGUG	18	downstream

Table 8C provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the third tier parameters. The targeting domains are within 1000 bp upstream of an *Alu* repeat, within 40bp upstream of mutation, or 1000 bp downstream of the mutation, start with G and PAM is NNGRRT. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. aureus* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 8C

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-697	-	GUAGAAUAAAUUUAUUUAAUG	21	upstream
CEP290-495	-	GAAUAAAUUUAUUUAAUG	18	upstream
CEP290-698	-	GAGAAAAAGGAGCAUGAAACAGG	23	upstream
CEP290-699	-	GAAAAAGGAGCAUGAAACAGG	21	upstream
CEP290-700	-	GUAGAAUAAAAAAUAAAAAAAC	22	upstream
CEP290-701	-	GAAUAAAAAAUAAAAAAAC	19	upstream
CEP290-702	-	GAAUAAAAAAUAAAAAAACUAGAG	24	upstream
CEP290-508	-	GAAAUAGAUGUAGAUUGAGG	20	downstream
CEP290-703	-	GAUAAUAAGGAAAUACAAAAA	21	downstream

CEP290-704	-	GUGUUGCCCAGGCUGGAGUGCAG	23	downstream
CEP290-705	-	GUUGCCCAGGCUGGAGUGCAG	21	downstream
CEP290-706	-	GCCCAGGCUGGAGUGCAG	18	downstream
CEP290-707	-	GUUGUUUUUUUUUUUGAAA	19	downstream
CEP290-708	-	GAGUCUCACUGUGUUGCCCAGGC	23	downstream
CEP290-709	-	GUCUCACUGUGUUGCCCAGGC	21	downstream

Table 8D provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the fourth tier parameters. The targeting domains are within 1000 bp upstream of an *Alu* repeat, within 40bp upstream of mutation, or 1000 bp downstream of the mutation and PAM is NNGRRT. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. aureus* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 8D

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-710	-	AAUGUAGAAUAAAUUUAUUUAAUG	24	upstream
CEP290-711	-	AUGUAGAAUAAAUUUAUUUAAUG	23	upstream
CEP290-712	-	AGAAUAAAUUUAUUUAAUG	19	upstream
CEP290-713	+	AUUUAUUCUACAAUAAAAAAUGAU	24	upstream
CEP290-714	+	AUUCUACAAUAAAAAAUGAU	20	upstream
CEP290-715	-	AGAGAAAAAGGAGCAUGAAACAGG	24	upstream
CEP290-716	-	AGAAAAAGGAGCAUGAAACAGG	22	upstream
CEP290-717	-	AAAAAGGAGCAUGAAACAGG	20	upstream
CEP290-718	-	AAAAGGAGCAUGAAACAGG	19	upstream
CEP290-719	-	AAAGGAGCAUGAAACAGG	18	upstream
CEP290-720	+	ACAAUAAAAAAUGAUGAGAAUUUA	24	upstream
CEP290-721	+	AAUAAAAAAUGAUGAGAAUUUA	22	upstream
CEP290-722	+	AUAAAAAAUGAUGAGAAUUUA	21	upstream
CEP290-723	+	AAAAAAUGAUGAGAAUUUA	19	upstream
CEP290-724	+	AAAAAUGAUGAGAAUUUA	18	upstream
CEP290-725	-	AUGUAGAAUAAAAAAUAAAAAAAC	24	upstream
CEP290-726	-	AGAAUAAAAAAUAAAAAAAC	20	upstream
CEP290-727	-	AAUAAAAAAUAAAAAAAC	18	upstream
CEP290-728	-	AAUAAAAAAUAAAAAAACUAGAG	23	upstream
CEP290-729	-	AUAAAAAAUAAAAAAACUAGAG	22	upstream

CEP290-730	-	AAAAAAUAAAAAAACUAGAG	20	upstream
CEP290-731	-	AAAAAUAAAAAAACUAGAG	19	upstream
CEP290-732	-	AAAAUAAAAAAACUAGAG	18	upstream
CEP290-733	+	CAAUAAAAAAUGAUGAGAAUUUA	23	upstream
CEP290-734	-	UGUAGAAUAAAUUUAUUUAAUG	22	upstream
CEP290-735	-	UAGAAUAAAUUUAUUUAAUG	20	upstream
CEP290-736	+	UUUAUUCUACAAUAAAAAAUGAU	23	upstream
CEP290-737	+	UUAUUCUACAAUAAAAAAUGAU	22	upstream
CEP290-738	+	UAUUCUACAAUAAAAAAUGAU	21	upstream
CEP290-739	+	UUCUACAAUAAAAAAUGAU	19	upstream
CEP290-740	+	UCUACAAUAAAAAAUGAU	18	upstream
CEP290-741	+	UAAAAAAUGAUGAGAAUUUA	20	upstream
CEP290-742	-	UGUAGAAUAAAAAAUAAAAAAAC	23	upstream
CEP290-743	-	UAGAAUAAAAAAUAAAAAAAC	21	upstream
CEP290-744	-	UAAAAAAUAAAAAAACUAGAG	21	upstream
CEP290-745	-	AAAAGAAAUAGAUGUAGAUUGAGG	24	downstream
CEP290-746	-	AAAGAAAUAGAUGUAGAUUGAGG	23	downstream
CEP290-747	-	AAGAAAUAGAUGUAGAUUGAGG	22	downstream
CEP290-748	-	AGAAAUAGAUGUAGAUUGAGG	21	downstream
CEP290-749	-	AAAUAGAUGUAGAUUGAGG	19	downstream
CEP290-750	-	AAUAGAUGUAGAUUGAGG	18	downstream
CEP290-751	-	AUAAUAAGGAAAUACAAAAACUGG	24	downstream
CEP290-752	-	AAUAAGGAAAUACAAAAACUGG	22	downstream
CEP290-753	-	AUAAGGAAAUACAAAAACUGG	21	downstream
CEP290-754	-	AAGGAAAUACAAAAACUGG	19	downstream
CEP290-755	-	AGGAAAUACAAAAACUGG	18	downstream
CEP290-756	-	AUAGAUAAUAAGGAAAUACAAAAA	24	downstream
CEP290-757	-	AGAUAAUAAGGAAAUACAAAAA	22	downstream
CEP290-758	-	AUAAUAAGGAAAUACAAAAA	20	downstream
CEP290-759	-	AAUAAGGAAAUACAAAAA	18	downstream
CEP290-760	+	AAAAAAAAAAACAACAAAAA	20	downstream
CEP290-761	+	AAAAAAAAAACAACAAAAA	19	downstream
CEP290-762	+	AAAAAAAAACAACAAAAA	18	downstream
CEP290-763	-	AGAGUCUCACUGUGUUGCCCAGGC	24	downstream
CEP290-764	-	AGUCUCACUGUGUUGCCCAGGC	22	downstream
CEP290-765	+	CAAAAAAAAAAACAACAAAAA	21	downstream
CEP290-766	-	CUCACUGUGUUGCCCAGGC	19	downstream
CEP290-767	-	UAAUAAGGAAAUACAAAAACUGG	23	downstream
CEP290-768	-	UAAGGAAAUACAAAAACUGG	20	downstream
CEP290-769	-	UAGAUAAUAAGGAAAUACAAAAA	23	downstream
CEP290-770	-	UAAUAAGGAAAUACAAAAA	19	downstream
CEP290-771	-	UGUGUUGCCCAGGCUGGAGUGCAG	24	downstream

CEP290-772	-	UGUUGCCCAGGCUGGAGUGCAG	22	downstream
CEP290-773	-	UUGCCCAGGCUGGAGUGCAG	20	downstream
CEP290-774	-	UGCCCAGGCUGGAGUGCAG	19	downstream
CEP290-775	+	UUUCAAAAAAAAAAACAACAAAAA	24	downstream
CEP290-776	+	UUCAAAAAAAAAAACAACAAAAA	23	downstream
CEP290-777	+	UCAAAAAAAAAAACAACAAAAA	22	downstream
CEP290-778	-	UUUUUGUUGUUUUUUUUUUUGAAA	24	downstream
CEP290-779	-	UUUUGUUGUUUUUUUUUUUGAAA	23	downstream
CEP290-780	-	UUUGUUGUUUUUUUUUUUGAAA	22	downstream
CEP290-781	-	UUGUUGUUUUUUUUUUUGAAA	21	downstream
CEP290-782	-	UGUUGUUUUUUUUUUUGAAA	20	downstream
CEP290-783	-	UUGUUUUUUUUUUUGAAA	18	downstream
CEP290-784	-	UCUCACUGUGUUGCCCAGGC	20	downstream
CEP290-785	-	UCACUGUGUUGCCCAGGC	18	downstream

Table 8E provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the fifth tier parameters. The targeting domains are within 1000 bp upstream of an *Alu* repeat, within 40bp upstream of mutation, or 1000 bp downstream of the mutation and PAM is NNGRRV. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *S. aureus* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 8E

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	
CEP290-786	+	ACUGUUGGCUACAUCCAUUCC	21	upstream
CEP290-787	+	AAUUUACAGAGUGCAUCCAUGGUC	24	upstream
CEP290-788	+	AUUUACAGAGUGCAUCCAUGGUC	23	upstream
CEP290-789	+	ACAGAGUGCAUCCAUGGUC	19	upstream
CEP290-790	-	AGCAUUAGAAAGUCCUAGGC	20	upstream
CEP290-791	-	AUGGUCCCUGGCUUUUGUUCC	21	upstream
CEP290-792	-	AUAGAGACACAUUCAGUAA	19	upstream
CEP290-793	-	AGCUCAAAAGCUUUUGCUGGCUCA	24	upstream
CEP290-794	-	AAAAGCUUUUGCUGGCUCA	19	upstream
CEP290-795	-	AAAGCUUUUGCUGGCUCA	18	upstream
CEP290-796	+	AAUCCAUAAGCCUCUAUUUCUGAU	24	upstream
CEP290-797	+	AUCCAUAAGCCUCUAUUUCUGAU	23	upstream
CEP290-798	+	AUAAGCCUCUAUUUCUGAU	19	upstream

CEP290-799	+	AGCUAAAUCAUGCAAGUGACCUA	23	upstream
CEP290-800	+	AAAUCAUGCAAGUGACCUA	19	upstream
CEP290-801	+	AAUCAUGCAAGUGACCUA	18	upstream
CEP290-802	-	AAACCUCUUUUAACCAGACAUCU	23	upstream
CEP290-803	-	AACCUCUUUUAACCAGACAUCU	22	upstream
CEP290-804	-	ACCUCUUUUAACCAGACAUCU	21	upstream
CEP290-805	+	AGUUUGUUCUGGGUACAGGGGUAA	24	upstream
CEP290-806	+	AUGACUCAUAAUUUAGUAGGAAUC	24	upstream
CEP290-807	+	ACUCAUAAUUUAGUAGGAAUC	21	upstream
CEP290-808	-	AAUGGAUGUAGCCAACAGUAG	21	upstream
CEP290-809	-	AUGGAUGUAGCCAACAGUAG	20	upstream
CEP290-810	+	AUCACCUCUCUUUGGCAAAAGCAG	24	upstream
CEP290-811	+	ACCUCUCUUUGGCAAAAGCAG	21	upstream
CEP290-812	-	AGGUAGAAUAUUGUAAUCAAAGG	23	upstream
CEP290-813	-	AGAAUAUUGUAAUCAAAGG	19	upstream
CEP290-814	+	AAGGAACAAAAGCCAGGGACC	21	upstream
CEP290-815	+	AGGAACAAAAGCCAGGGACC	20	upstream
CEP290-816	+	ACAUCCAUUCCAAGGAACAAAAGC	24	upstream
CEP290-817	+	AUCCAUUCCAAGGAACAAAAGC	22	upstream
CEP290-818	+	AUUCCAAGGAACAAAAGC	18	upstream
CEP290-819	+	AGAAUUAGAUCUUAUUCUACUCCU	24	upstream
CEP290-820	+	AAUUAGAUCUUAUUCUACUCCU	22	upstream
CEP290-821	+	AUUAGAUCUUAUUCUACUCCU	21	upstream
CEP290-822	+	AGAUCUUAUUCUACUCCU	18	upstream
CEP290-823	-	AUUUGUUCAUCUUCCUCAU	19	upstream
CEP290-824	-	AGAGGUGAUUAUGUUACUUUUUA	23	upstream
CEP290-825	-	AGGUGAUUAUGUUACUUUUUA	21	upstream
CEP290-826	-	AACCUCUUUUAACCAGACAUCUAA	24	upstream
CEP290-827	-	ACCUCUUUUAACCAGACAUCUAA	23	upstream
CEP290-828	+	AUAAACAUGACUCAUAAUUUAG	22	upstream
CEP290-829	+	AAACAUGACUCAUAAUUUAG	20	upstream
CEP290-830	+	AACAUGACUCAUAAUUUAG	19	upstream
CEP290-831	+	ACAUGACUCAUAAUUUAG	18	upstream
CEP290-832	-	ACAGGUAGAAUAUUGUAAUCAAAG	24	upstream
CEP290-833	-	AGGUAGAAUAUUGUAAUCAAAG	22	upstream
CEP290-834	-	AGAAUAUUGUAAUCAAAG	18	upstream
CEP290-835	+	AUAGUUUGUUCUGGGUACAGGGGU	24	upstream
CEP290-836	+	AGUUUGUUCUGGGUACAGGGGU	22	upstream
CEP290-837	-	AGACAUCUAAGAGAAAAAGGAGC	23	upstream
CEP290-838	-	ACAUCUAAGAGAAAAAGGAGC	21	upstream
CEP290-839	-	AUCUAAGAGAAAAAGGAGC	19	upstream
CEP290-840	+	AGAGGAUAGGACAGAGGACA	20	upstream

CEP290-841	+	AGGAUAGGACAGAGGACA	18	upstream
CEP290-842	+	AGGAAAGAUGAAAAAUACUCUU	22	upstream
CEP290-843	+	AAAGAUGAAAAAUACUCUU	19	upstream
CEP290-844	+	AAGAUGAAAAAUACUCUU	18	upstream
CEP290-845	+	AGGAAAGAUGAAAAAUACUCUUU	23	upstream
CEP290-846	+	AAAGAUGAAAAAUACUCUUU	20	upstream
CEP290-847	+	AAGAUGAAAAAUACUCUUU	19	upstream
CEP290-848	+	AGAUGAAAAAUACUCUUU	18	upstream
CEP290-849	+	AGGAAAGAUGAAAAAUACUCU	21	upstream
CEP290-850	+	AAAGAUGAAAAAUACUCU	18	upstream
CEP290-851	+	AUAGGACAGAGGACAUGGAGAA	22	upstream
CEP290-852	+	AGGACAGAGGACAUGGAGAA	20	upstream
CEP290-853	+	AGGAUAGGACAGAGGACAUGGAGA	24	upstream
CEP290-854	+	AUAGGACAGAGGACAUGGAGA	21	upstream
CEP290-855	+	AGGACAGAGGACAUGGAGA	19	upstream
CEP290-856	+	AAGGAACAAAAGCCAGGGACCAU	23	upstream
CEP290-857	+	AGGAACAAAAGCCAGGGACCAU	22	upstream
CEP290-858	+	AACAAAAGCCAGGGACCAU	19	upstream
CEP290-859	+	ACAAAAGCCAGGGACCAU	18	upstream
CEP290-860	+	ACAUUUAUUCUACAAUAAAAAAUG	24	upstream
CEP290-861	+	AUUUAUUCUACAAUAAAAAAUG	22	upstream
CEP290-862	+	AUUCUACAAUAAAAAAUG	18	upstream
CEP290-863	+	AUUGUGUGUGUGUGUGUGUUAU	24	upstream
CEP290-864	+	CUACUGUUGGCUACAUCCAUUCC	23	upstream
CEP290-865	+	CUGUUGGCUACAUCCAUUCC	20	upstream
CEP290-866	+	CAGAGUGCAUCCAUGGUC	18	upstream
CEP290-867	-	CUCCAGCAUUAGAAAGUCCUAGGC	24	upstream
CEP290-868	-	CCAGCAUUAGAAAGUCCUAGGC	22	upstream
CEP290-869	-	CAGCAUUAGAAAGUCCUAGGC	21	upstream
CEP290-870	-	CAUUAGAAAGUCCUAGGC	18	upstream
CEP290-871	-	CCCAUGGUCCCUGGCUUUUGUUCC	24	upstream
CEP290-872	-	CCAUGGUCCCUGGCUUUUGUUCC	23	upstream
CEP290-873	-	CAUGGUCCCUGGCUUUUGUUCC	22	upstream
CEP290-874	-	CUCAUAGAGACACAUUCAGUAA	22	upstream
CEP290-875	-	CAUAGAGACACAUUCAGUAA	20	upstream
CEP290-876	-	CUCAAAAGCUUUUGCUGGCUCA	22	upstream
CEP290-877	-	CAAAAGCUUUUGCUGGCUCA	20	upstream
CEP290-878	+	CCAUAAGCCUCUAUUUCUGAU	21	upstream
CEP290-879	+	CAUAAGCCUCUAUUUCUGAU	20	upstream
CEP290-880	+	CAGCUAAAUCAUGCAAGUGACCUA	24	upstream
CEP290-881	+	CUAAAUCAUGCAAGUGACCUA	21	upstream
CEP290-882	-	CAAACCUCUUUUAACCAGACAUCU	24	upstream

CEP290-883	-	CCUCUUUUAACCAGACAUCU	20	upstream
CEP290-884	-	CUCUUUUAACCAGACAUCU	19	upstream
CEP290-885	+	CUCAUAAUUUAGUAGGAAUC	20	upstream
CEP290-886	+	CAUAAUUUAGUAGGAAUC	18	upstream
CEP290-887	+	CACCUCUCUUUGGCAAAAGCAG	22	upstream
CEP290-888	+	CCUCUCUUUGGCAAAAGCAG	20	upstream
CEP290-889	+	CUCUCUUUGGCAAAAGCAG	19	upstream
CEP290-890	-	CAGGUAGAAUAUUGUAAUCAAAGG	24	upstream
CEP290-891	+	CCAAGGAACAAAAGCCAGGGACC	23	upstream
CEP290-892	+	CAAGGAACAAAAGCCAGGGACC	22	upstream
CEP290-893	+	CAUCCAUUCCAAGGAACAAAAGC	23	upstream
CEP290-894	+	CCAUUCCAAGGAACAAAAGC	20	upstream
CEP290-895	+	CAUUCCAAGGAACAAAAGC	19	upstream
CEP290-896	+	CUCUUGCCUAGGACUUUCUAAUGC	24	upstream
CEP290-897	+	CUUGCCUAGGACUUUCUAAUGC	22	upstream
CEP290-898	+	CCUAGGACUUUCUAAUGC	18	upstream
CEP290-899	-	CCUGAUUUGUUCAUCUUCCUCAU	23	upstream
CEP290-900	-	CUGAUUUGUUCAUCUUCCUCAU	22	upstream
CEP290-901	-	CCUCUUUUAACCAGACAUCUAA	22	upstream
CEP290-902	-	CUCUUUUAACCAGACAUCUAA	21	upstream
CEP290-903	-	CUUUUAACCAGACAUCUAA	19	upstream
CEP290-904	-	CCUCUGUCCUAUCCUCUCCAGCAU	24	upstream
CEP290-905	-	CUCUGUCCUAUCCUCUCCAGCAU	23	upstream
CEP290-906	-	CUGUCCUAUCCUCUCCAGCAU	21	upstream
CEP290-907	-	CAGGUAGAAUAUUGUAAUCAAAG	23	upstream
CEP290-908	+	CUGGGUACAGGGGUAAGAGA	20	upstream
CEP290-909	-	CUUUCUGCUGCUUUUGCCA	19	upstream
CEP290-910	-	CAGACAUCUAAGAGAAAAAGGAGC	24	upstream
CEP290-911	-	CAUCUAAGAGAAAAAGGAGC	20	upstream
CEP290-912	+	CUGGAGAGGAUAGGACAGAGGACA	24	upstream
CEP290-913	+	CAAGGAACAAAAGCCAGGGACCAU	24	upstream
CEP290-914	+	CAUUUAUUCUACAAUAAAAAAUG	23	upstream
CEP290-915	+	GCUACUGUUGGCUACAUCCAUUCC	24	upstream
CEP290-916	+	GUUGGCUACAUCCAUUCC	18	upstream
CEP290-917	-	GCAUUAGAAAGUCCUAGGC	19	upstream
CEP290-918	-	GGUCCCUGGCUUUUGUUCC	19	upstream
CEP290-919	-	GUCCCUGGCUUUUGUUCC	18	upstream
CEP290-920	-	GGCUCAUAGAGACACAUUCAGUAA	24	upstream
CEP290-921	-	GCUCAUAGAGACACAUUCAGUAA	23	upstream
CEP290-922	-	GCUCAAAAGCUUUUGCUGGCUCA	23	upstream
CEP290-923	+	GCUAAAUCAUGCAAGUGACCUA	22	upstream
CEP290-924	+	GUUUGUUCUGGGUACAGGGGUAA	23	upstream

CEP290-925	+	GUUCUGGGUACAGGGGUAA	19	upstream
CEP290-926	+	GACUCAUAAUUUAGUAGGAAUC	22	upstream
CEP290-927	-	GGAAUGGAUGUAGCCAACAGUAG	23	upstream
CEP290-928	-	GAAUGGAUGUAGCCAACAGUAG	22	upstream
CEP290-929	-	GGAUGUAGCCAACAGUAG	18	upstream
CEP290-930	-	GGUAGAAUAUUGUAAUCAAAGG	22	upstream
CEP290-931	-	GUAGAAUAUUGUAAUCAAAGG	21	upstream
CEP290-932	-	GAAUAUUGUAAUCAAAGG	18	upstream
CEP290-933	+	GGAACAAAAGCCAGGGACC	19	upstream
CEP290-934	+	GAACAAAAGCCAGGGACC	18	upstream
CEP290-935	+	GAAUUAGAUCUUAUUCUACUCCU	23	upstream
CEP290-936	+	GCCUAGGACUUUCUAAUGC	19	upstream
CEP290-937	-	GAUUUGUUCAUCUUCCUCAU	20	upstream
CEP290-938	-	GAGAGGUGAUUAUGUUACUUUUUA	24	upstream
CEP290-939	-	GAGGUGAUUAUGUUACUUUUUA	22	upstream
CEP290-940	-	GGUGAUUAUGUUACUUUUUA	20	upstream
CEP290-941	-	GUGAUUAUGUUACUUUUUA	19	upstream
CEP290-942	-	GUCCUAUCCUCUCCAGCAU	19	upstream
CEP290-943	+	GAUAAACAUGACUCAUAAUUUAG	23	upstream
CEP290-944	-	GGUAGAAUAUUGUAAUCAAAG	21	upstream
CEP290-945	-	GUAGAAUAUUGUAAUCAAAG	20	upstream
CEP290-946	+	GUUCUGGGUACAGGGGUAAGAGA	23	upstream
CEP290-947	+	GGGUACAGGGGUAAGAGA	18	upstream
CEP290-948	+	GUUUGUUCUGGGUACAGGGGU	21	upstream
CEP290-949	-	GUUUGCUUUCUGCUGCUUUUGCCA	24	upstream
CEP290-950	-	GCUUUCUGCUGCUUUUGCCA	20	upstream
CEP290-951	-	GACAUCUAAGAGAAAAAGGAGC	22	upstream
CEP290-952	+	GGAGAGGAUAGGACAGAGGACA	22	upstream
CEP290-953	+	GAGAGGAUAGGACAGAGGACA	21	upstream
CEP290-954	+	GAGGAUAGGACAGAGGACA	19	upstream
CEP290-955	+	GGAAAGAUGAAAAAUACUCUU	21	upstream
CEP290-956	+	GAAAGAUGAAAAAUACUCUU	20	upstream
CEP290-957	+	GGAAAGAUGAAAAAUACUCUUU	22	upstream
CEP290-958	+	GAAAGAUGAAAAAUACUCUUU	21	upstream
CEP290-959	+	GGAAAGAUGAAAAAUACUCU	20	upstream
CEP290-960	+	GAAAGAUGAAAAAUACUCU	19	upstream
CEP290-961	+	GGAUAGGACAGAGGACAUGGAGAA	24	upstream
CEP290-962	+	GAUAGGACAGAGGACAUGGAGAA	23	upstream
CEP290-963	+	GGACAGAGGACAUGGAGAA	19	upstream
CEP290-964	+	GACAGAGGACAUGGAGAA	18	upstream
CEP290-965	+	GGAUAGGACAGAGGACAUGGAGA	23	upstream
CEP290-966	+	GAUAGGACAGAGGACAUGGAGA	22	upstream

CEP290-967	+	GGACAGAGGACAUGGAGA	18	upstream
CEP290-968	+	GGAACAAAAGCCAGGGACCAU	21	upstream
CEP290-969	+	GAACAAAAGCCAGGGACCAU	20	upstream
CEP290-970	+	GUGUGUGUGUGUGUGUUAU	21	upstream
CEP290-971	+	GUGUGUGUGUGUGUUAU	19	upstream
CEP290-972	+	GUGUGUGUGUGUGUGUUAUG	22	upstream
CEP290-973	+	GUGUGUGUGUGUGUUAUG	20	upstream
CEP290-974	+	GUGUGUGUGUGUUAUG	18	upstream
CEP290-975	+	UACUGUUGGCUACAUCCAUUCC	22	upstream
CEP290-976	+	UGUUGGCUACAUCCAUUCC	19	upstream
CEP290-977	+	UUUACAGAGUGCAUCCAUGGUC	22	upstream
CEP290-978	+	UUACAGAGUGCAUCCAUGGUC	21	upstream
CEP290-979	+	UACAGAGUGCAUCCAUGGUC	20	upstream
CEP290-980	-	UCCAGCAUUAGAAAGUCCUAGGC	23	upstream
CEP290-981	-	UGGUCCCUGGCUUUUGUUCC	20	upstream
CEP290-982	-	UCAUAGAGACACAUUCAGUAA	21	upstream
CEP290-983	-	UAGAGACACAUUCAGUAA	18	upstream
CEP290-984	-	UCAAAAGCUUUUGCUGGCUCA	21	upstream
CEP290-985	+	UCCAUAAGCCUCUAUUUCUGAU	22	upstream
CEP290-986	+	UAAGCCUCUAUUUCUGAU	18	upstream
CEP290-987	+	UAAAUCAUGCAAGUGACCUA	20	upstream
CEP290-988	-	UCUUUUAACCAGACAUCU	18	upstream
CEP290-989	+	UUUGUUCUGGGUACAGGGGUAA	22	upstream
CEP290-990	+	UUGUUCUGGGUACAGGGGUAA	21	upstream
CEP290-991	+	UGUUCUGGGUACAGGGGUAA	20	upstream
CEP290-992	+	UUCUGGGUACAGGGGUAA	18	upstream
CEP290-993	+	UGACUCAUAAUUUAGUAGGAAUC	23	upstream
CEP290-994	+	UCAUAAUUUAGUAGGAAUC	19	upstream
CEP290-995	-	UGGAAUGGAUGUAGCCAACAGUAG	24	upstream
CEP290-996	-	UGGAUGUAGCCAACAGUAG	19	upstream
CEP290-997	+	UCACCUCUCUUUGGCAAAAGCAG	23	upstream
CEP290-998	+	UCUCUUUGGCAAAAGCAG	18	upstream
CEP290-999	-	UAGAAUAUUGUAAUCAAAGG	20	upstream
CEP290-1000	+	UCCAAGGAACAAAAGCCAGGGACC	24	upstream
CEP290-1001	+	UCCAUUCCAAGGAACAAAAGC	21	upstream
CEP290-1002	+	UUAGAUCUUAUUCUACUCCU	20	upstream
CEP290-1003	+	UAGAUCUUAUUCUACUCCU	19	upstream
CEP290-1004	+	UCUUGCCUAGGACUUUCUAAUGC	23	upstream
CEP290-1005	+	UUGCCUAGGACUUUCUAAUGC	21	upstream
CEP290-1006	+	UGCCUAGGACUUUCUAAUGC	20	upstream
CEP290-1007	-	UCCUGAUUUGUUCAUCUUCCUCAU	24	upstream
CEP290-1008	-	UGAUUUGUUCAUCUUCCUCAU	21	upstream

CEP290-1009	-	UUUGUUCAUCUUCCUCAU	18	upstream
CEP290-1010	-	UGAUUAUGUUACUUUUUA	18	upstream
CEP290-1011	-	UCUUUUAACCAGACAUCUAA	20	upstream
CEP290-1012	-	UUUUAACCAGACAUCUAA	18	upstream
CEP290-1013	-	UCUGUCCUAUCCUCUCCAGCAU	22	upstream
CEP290-1014	-	UGUCCUAUCCUCUCCAGCAU	20	upstream
CEP290-1015	-	UCCUAUCCUCUCCAGCAU	18	upstream
CEP290-1016	+	UGAUAAACAUGACUCAUAAUUUAG	24	upstream
CEP290-1017	+	UAAACAUGACUCAUAAUUUAG	21	upstream
CEP290-1018	-	UAGAAUAUUGUAAUCAAAG	19	upstream
CEP290-1019	+	UGUUCUGGGUACAGGGGUAAGAGA	24	upstream
CEP290-1020	+	UUCUGGGUACAGGGGUAAGAGA	22	upstream
CEP290-1021	+	UCUGGGUACAGGGGUAAGAGA	21	upstream
CEP290-1022	+	UGGGUACAGGGGUAAGAGA	19	upstream
CEP290-1023	+	UAGUUUGUUCUGGGUACAGGGGU	23	upstream
CEP290-1024	+	UUUGUUCUGGGUACAGGGGU	20	upstream
CEP290-1025	+	UUGUUCUGGGUACAGGGGU	19	upstream
CEP290-1026	+	UGUUCUGGGUACAGGGGU	18	upstream
CEP290-1027	-	UUUGCUUUCUGCUGCUUUUGCCA	23	upstream
CEP290-1028	-	UUGCUUUCUGCUGCUUUUGCCA	22	upstream
CEP290-1029	-	UGCUUUCUGCUGCUUUUGCCA	21	upstream
CEP290-1030	-	UUUCUGCUGCUUUUGCCA	18	upstream
CEP290-1031	-	UCUAAGAGAAAAAGGAGC	18	upstream
CEP290-1032	+	UGGAGAGGAUAGGACAGAGGACA	23	upstream
CEP290-1033	+	UUAGGAAAGAUGAAAAAUACUCUU	24	upstream
CEP290-1034	+	UAGGAAAGAUGAAAAAUACUCUU	23	upstream
CEP290-1035	+	UAGGAAAGAUGAAAAAUACUCUUU	24	upstream
CEP290-1036	+	UUUAGGAAAGAUGAAAAAUACUCU	24	upstream
CEP290-1037	+	UUAGGAAAGAUGAAAAAUACUCU	23	upstream
CEP290-1038	+	UAGGAAAGAUGAAAAAUACUCU	22	upstream
CEP290-1039	+	UAGGACAGAGGACAUGGAGAA	21	upstream
CEP290-1040	+	UAGGACAGAGGACAUGGAGA	20	upstream
CEP290-1041	+	UUUAUUCUACAAUAAAAAAUG	21	upstream
CEP290-1042	+	UUAUUCUACAAUAAAAAAUG	20	upstream
CEP290-1043	+	UAUUCUACAAUAAAAAAUG	19	upstream
CEP290-1044	+	UUGUGUGUGUGUGUGUGUGUUAU	23	upstream
CEP290-1045	+	UGUGUGUGUGUGUGUGUGUUAU	22	upstream
CEP290-1046	+	UGUGUGUGUGUGUGUGUUAU	20	upstream
CEP290-1047	+	UGUGUGUGUGUGUGUUAU	18	upstream
CEP290-1048	+	UUGUGUGUGUGUGUGUGUGUUAUG	24	upstream
CEP290-1049	+	UGUGUGUGUGUGUGUGUGUUAUG	23	upstream
CEP290-1050	+	UGUGUGUGUGUGUGUGUUAUG	21	upstream

CEP290-1051	+	UGUGUGUGUGUGUUAUG	19	upstream
CEP290-1052	+	ACUGUUGGCUACAUCCAUUCCA	22	upstream
CEP290-1053	+	AUUAUCCACAAGAUGUCUCUUGCC	24	upstream
CEP290-1054	+	AUCCACAAGAUGUCUCUUGCC	21	upstream
CEP290-1055	+	AUGAGCCAGCAAAAGCUU	18	upstream
CEP290-1056	+	ACAGAGUGCAUCCAUGGUCCAGG	23	upstream
CEP290-1057	+	AGAGUGCAUCCAUGGUCCAGG	21	upstream
CEP290-1058	+	AGUGCAUCCAUGGUCCAGG	19	upstream
CEP290-1059	-	AGCUGAAAUAUUAAGGGCUCUUC	23	upstream
CEP290-1060	-	AAAUAUUAAGGGCUCUUC	18	upstream
CEP290-1061	-	AACUCUAUACCUUUUACUGAGGA	23	upstream
CEP290-1062	-	ACUCUAUACCUUUUACUGAGGA	22	upstream
CEP290-1063	-	ACUUGAACUCUAUACCUUUUACU	23	upstream
CEP290-1064	-	AACUCUAUACCUUUUACU	18	upstream
CEP290-1065	+	AGUAGGAAUCCUGAAAGCUACU	22	upstream
CEP290-1066	+	AGGAAUCCUGAAAGCUACU	19	upstream
CEP290-1067	-	AGCCAACAGUAGCUGAAAUAUU	22	upstream
CEP290-1068	-	AACAGUAGCUGAAAUAUU	18	upstream
CEP290-1069	+	AUCCAUUCCAAGGAACAAAAGCC	23	upstream
CEP290-1070	+	AUUCCAAGGAACAAAAGCC	19	upstream
CEP290-1071	-	AUCCCUUUCUCUUACCCCUGUACC	24	upstream
CEP290-1072	+	AGGACUUUCUAAUGCUGGAGAGGA	24	upstream
CEP290-1073	+	ACUUUCUAAUGCUGGAGAGGA	21	upstream
CEP290-1074	+	AAUGCUGGAGAGGAUAGGACA	21	upstream
CEP290-1075	+	AUGCUGGAGAGGAUAGGACA	20	upstream
CEP290-1076	-	AUCAUAAGUUACAAUCUGUGAAU	23	upstream
CEP290-1077	-	AUAAGUUACAAUCUGUGAAU	20	upstream
CEP290-1078	-	AAGUUACAAUCUGUGAAU	18	upstream
CEP290-1079	-	AACCAGACAUCUAAGAGAAAA	21	upstream
CEP290-1080	-	ACCAGACAUCUAAGAGAAAA	20	upstream
CEP290-1081	+	AAGCCUCUAUUUCUGAUGAGGAAG	24	upstream
CEP290-1082	+	AGCCUCUAUUUCUGAUGAGGAAG	23	upstream
CEP290-1083	+	AUGAGGAAGAUGAACAAAUC	20	upstream
CEP290-1084	+	AUUUACUGAAUGUGUCUCU	19	upstream
CEP290-1085	+	ACAGGGGUAAGAGAAAGGG	19	upstream
CEP290-1086	+	CUACUGUUGGCUACAUCCAUUCCA	24	upstream
CEP290-1087	+	CUGUUGGCUACAUCCAUUCCA	21	upstream
CEP290-1088	+	CCACAAGAUGUCUCUUGCC	19	upstream
CEP290-1089	+	CACAAGAUGUCUCUUGCC	18	upstream
CEP290-1090	-	CCUUUGUAGUUAUCUUACAGCCAC	24	upstream
CEP290-1091	-	CUUUGUAGUUAUCUUACAGCCAC	23	upstream
CEP290-1092	+	CUCUAUGAGCCAGCAAAAGCUU	22	upstream

CEP290-1093	+	CUAUGAGCCAGCAAAAGCUU	20	upstream
CEP290-1094	+	CAGAGUGCAUCCAUGGUCCAGG	22	upstream
CEP290-1095	-	CUGAAAUAUUAAGGGCUCUUC	21	upstream
CEP290-1096	-	CUCUAUACCUUUUACUGAGGA	21	upstream
CEP290-1097	-	CUAUACCUUUUACUGAGGA	19	upstream
CEP290-1098	-	CACUUGAACUCUAUACCUUUUACU	24	upstream
CEP290-1099	-	CUUGAACUCUAUACCUUUUACU	22	upstream
CEP290-1100	-	CCAACAGUAGCUGAAAUAUU	20	upstream
CEP290-1101	-	CAACAGUAGCUGAAAUAUU	19	upstream
CEP290-1102	+	CAUCCAUUCCAAGGAACAAAAGCC	24	upstream
CEP290-1103	+	CCAUUCCAAGGAACAAAAGCC	21	upstream
CEP290-1104	+	CAUUCCAAGGAACAAAAGCC	20	upstream
CEP290-1105	-	CCCUUUCUCUUACCCCUGUACC	22	upstream
CEP290-1106	-	CCUUUCUCUUACCCCUGUACC	21	upstream
CEP290-1107	-	CUUUCUCUUACCCCUGUACC	20	upstream
CEP290-1108	+	CUUUCUAAUGCUGGAGAGGA	20	upstream
CEP290-1109	+	CUAAUGCUGGAGAGGAUAGGACA	23	upstream
CEP290-1110	-	CAUAAGUUACAAUCUGUGAAU	21	upstream
CEP290-1111	-	CCAGACAUCUAAGAGAAAA	19	upstream
CEP290-1112	-	CAGACAUCUAAGAGAAAA	18	upstream
CEP290-1113	+	CCUCUAUUUCUGAUGAGGAAG	21	upstream
CEP290-1114	+	CUCUAUUUCUGAUGAGGAAG	20	upstream
CEP290-1115	+	CUAUUUCUGAUGAGGAAG	18	upstream
CEP290-1116	+	CUGAUGAGGAAGAUGAACAAAUC	23	upstream
CEP290-1117	+	CAUUUACUGAAUGUGUCUCU	20	upstream
CEP290-1118	+	CAGGGGUAAGAGAAAGGG	18	upstream
CEP290-1119	+	GUUGGCUACAUCCAUUCCA	19	upstream
CEP290-1120	-	GUAGUUAUCUUACAGCCAC	19	upstream
CEP290-1121	+	GUCUCUAUGAGCCAGCAAAAGCUU	24	upstream
CEP290-1122	+	GAGUGCAUCCAUGGUCCAGG	20	upstream
CEP290-1123	+	GUGCAUCCAUGGUCCAGG	18	upstream
CEP290-1124	-	GCUGAAAUAUUAAGGGCUCUUC	22	upstream
CEP290-1125	-	GAAAUAUUAAGGGCUCUUC	19	upstream
CEP290-1126	-	GAACUCUAUACCUUUUACUGAGGA	24	upstream
CEP290-1127	-	GAACUCUAUACCUUUUACU	19	upstream
CEP290-1128	+	GUAGGAAUCCUGAAAGCUACU	21	upstream
CEP290-1129	+	GGAAUCCUGAAAGCUACU	18	upstream
CEP290-1130	-	GUAGCCAACAGUAGCUGAAAUAUU	24	upstream
CEP290-1131	-	GCCAACAGUAGCUGAAAUAUU	21	upstream
CEP290-1132	+	GGACUUUCUAAUGCUGGAGAGGA	23	upstream
CEP290-1133	+	GACUUUCUAAUGCUGGAGAGGA	22	upstream
CEP290-1134	+	GCUGGAGAGGAUAGGACA	18	upstream

CEP290-1135	+	GCCUCUAUUUCUGAUGAGGAAG	22	upstream
CEP290-1136	+	GAUGAGGAAGAUGAACAAAUC	21	upstream
CEP290-1137	+	GAGGAAGAUGAACAAAUC	18	upstream
CEP290-1138	+	GGGUACAGGGGUAAGAGAAAGGG	23	upstream
CEP290-1139	+	GGUACAGGGGUAAGAGAAAGGG	22	upstream
CEP290-1140	+	GUACAGGGGUAAGAGAAAGGG	21	upstream
CEP290-1141	+	GUGUGUGUGUGUGUGUGUUAUGU	23	upstream
CEP290-1142	+	GUGUGUGUGUGUGUGUUAUGU	21	upstream
CEP290-1143	+	GUGUGUGUGUGUGUUAUGU	19	upstream
CEP290-1144	+	UACUGUUGGCUACAUCCAUUCCA	23	upstream
CEP290-1145	+	UGUUGGCUACAUCCAUUCCA	20	upstream
CEP290-1146	+	UUGGCUACAUCCAUUCCA	18	upstream
CEP290-1147	+	UUAUCCACAAGAUGUCUCUUGCC	23	upstream
CEP290-1148	+	UAUCCACAAGAUGUCUCUUGCC	22	upstream
CEP290-1149	+	UCCACAAGAUGUCUCUUGCC	20	upstream
CEP290-1150	-	UUUGUAGUUAUCUUACAGCCAC	22	upstream
CEP290-1151	-	UUGUAGUUAUCUUACAGCCAC	21	upstream
CEP290-1152	-	UGUAGUUAUCUUACAGCCAC	20	upstream
CEP290-1153	-	UAGUUAUCUUACAGCCAC	18	upstream
CEP290-1154	+	UCUCUAUGAGCCAGCAAAAGCUU	23	upstream
CEP290-1155	+	UCUAUGAGCCAGCAAAAGCUU	21	upstream
CEP290-1156	+	UAUGAGCCAGCAAAAGCUU	19	upstream
CEP290-1157	+	UACAGAGUGCAUCCAUGGUCCAGG	24	upstream
CEP290-1158	-	UAGCUGAAAUAUUAAGGGCUCUUC	24	upstream
CEP290-1159	-	UGAAAUAUUAAGGGCUCUUC	20	upstream
CEP290-1160	-	UCUAUACCUUUUACUGAGGA	20	upstream
CEP290-1161	-	UAUACCUUUUACUGAGGA	18	upstream
CEP290-1162	-	UUGAACUCUAUACCUUUUACU	21	upstream
CEP290-1163	-	UGAACUCUAUACCUUUUACU	20	upstream
CEP290-1164	+	UUAGUAGGAAUCCUGAAAGCUACU	24	upstream
CEP290-1165	+	UAGUAGGAAUCCUGAAAGCUACU	23	upstream
CEP290-1166	+	UAGGAAUCCUGAAAGCUACU	20	upstream
CEP290-1167	-	UAGCCAACAGUAGCUGAAAUAUU	23	upstream
CEP290-1168	+	UCCAUUCCAAGGAACAAAAGCC	22	upstream
CEP290-1169	+	UUCCAAGGAACAAAAGCC	18	upstream
CEP290-1170	-	UCCCUUUCUCUUACCCCUGUACC	23	upstream
CEP290-1171	-	UUUCUCUUACCCCUGUACC	19	upstream
CEP290-1172	-	UUCUCUUACCCCUGUACC	18	upstream
CEP290-1173	+	UUUCUAAUGCUGGAGAGGA	19	upstream
CEP290-1174	+	UUCUAAUGCUGGAGAGGA	18	upstream
CEP290-1175	+	UCUAAUGCUGGAGAGGAUAGGACA	24	upstream
CEP290-1176	+	UAAUGCUGGAGAGGAUAGGACA	22	upstream

CEP290-1177	+	UGCUGGAGAGGAUAGGACA	19	upstream
CEP290-1178	-	UAUCAUAAGUUACAAUCUGUGAAU	24	upstream
CEP290-1179	-	UCAUAAGUUACAAUCUGUGAAU	22	upstream
CEP290-1180	-	UAAGUUACAAUCUGUGAAU	19	upstream
CEP290-1181	-	UUUAACCAGACAUCUAAGAGAAAA	24	upstream
CEP290-1182	-	UUAACCAGACAUCUAAGAGAAAA	23	upstream
CEP290-1183	-	UAACCAGACAUCUAAGAGAAAA	22	upstream
CEP290-1184	+	UCUAUUUCUGAUGAGGAAG	19	upstream
CEP290-1185	+	UCUGAUGAGGAAGAUGAACAAAUC	24	upstream
CEP290-1186	+	UGAUGAGGAAGAUGAACAAAUC	22	upstream
CEP290-1187	+	UGAGGAAGAUGAACAAAUC	19	upstream
CEP290-1188	+	UUUUCAUUUACUGAAUGUGUCUCU	24	upstream
CEP290-1189	+	UUUCAUUUACUGAAUGUGUCUCU	23	upstream
CEP290-1190	+	UUCAUUUACUGAAUGUGUCUCU	22	upstream
CEP290-1191	+	UCAUUUACUGAAUGUGUCUCU	21	upstream
CEP290-1192	+	UUUACUGAAUGUGUCUCU	18	upstream
CEP290-1193	+	UGGGUACAGGGGUAAGAGAAAGGG	24	upstream
CEP290-1194	+	UACAGGGGUAAGAGAAAGGG	20	upstream
CEP290-1195	+	UGUGUGUGUGUGUGUGUUAUGU	24	upstream
CEP290-1196	+	UGUGUGUGUGUGUGUUAUGU	22	upstream
CEP290-1197	+	UGUGUGUGUGUGUUAUGU	20	upstream
CEP290-1198	+	UGUGUGUGUGUUAUGU	18	upstream
CEP290-1199	+	AUUUACAGAGUGCAUCCAUGGUCC	24	upstream
CEP290-1200	+	ACAGAGUGCAUCCAUGGUCC	20	upstream
CEP290-1201	+	AGAGUGCAUCCAUGGUCC	18	upstream
CEP290-1202	-	ACUUGAACUCUAUACCUUUUA	21	upstream
CEP290-1203	+	AGCUAAAUCAUGCAAGUGACCU	22	upstream
CEP290-1204	+	AAAUCAUGCAAGUGACCU	18	upstream
CEP290-1205	+	AUCCAUAAGCCUCUAUUUCUGAUG	24	upstream
CEP290-1206	+	AUAAGCCUCUAUUUCUGAUG	20	upstream
CEP290-1207	+	AAGCCUCUAUUUCUGAUG	18	upstream
CEP290-1208	+	AGAAUAGUUUGUUCUGGGUA	20	upstream
CEP290-1209	+	AAUAGUUUGUUCUGGGUA	18	upstream
CEP290-1210	+	AGGAGAAUGAUCUAGAUAAUCAUU	24	upstream
CEP290-1211	+	AGAAUGAUCUAGAUAAUCAUU	21	upstream
CEP290-1212	+	AAUGAUCUAGAUAAUCAUU	19	upstream
CEP290-1213	+	AUGAUCUAGAUAAUCAUU	18	upstream
CEP290-1214	+	AAUGCUGGAGAGGAUAGGA	19	upstream
CEP290-1215	+	AUGCUGGAGAGGAUAGGA	18	upstream
CEP290-1216	+	AAAAUCCAUAAGCCUCUAUUUCUG	24	upstream
CEP290-1217	+	AAAUCCAUAAGCCUCUAUUUCUG	23	upstream
CEP290-1218	+	AAUCCAUAAGCCUCUAUUUCUG	22	upstream

CEP290-1219	+	AUCCAUAAGCCUCUAUUUCUG	21	upstream
CEP290-1220	-	AAACAGGUAGAAUAUUGUAAUCA	23	upstream
CEP290-1221	-	AACAGGUAGAAUAUUGUAAUCA	22	upstream
CEP290-1222	-	ACAGGUAGAAUAUUGUAAUCA	21	upstream
CEP290-1223	-	AGGUAGAAUAUUGUAAUCA	19	upstream
CEP290-1224	+	AAGGAACAAAAGCCAGGGACCA	22	upstream
CEP290-1225	+	AGGAACAAAAGCCAGGGACCA	21	upstream
CEP290-1226	+	AACAAAAGCCAGGGACCA	18	upstream
CEP290-1227	-	AGGUAGAAUAUUGUAAUCAAAGGA	24	upstream
CEP290-1228	-	AGAAUAUUGUAAUCAAAGGA	20	upstream
CEP290-1229	-	AAUAUUGUAAUCAAAGGA	18	upstream
CEP290-1230	-	AGUCAUGUUUAUCAAUAUUAUU	22	upstream
CEP290-1231	-	AUGUUUAUCAAUAUUAUU	18	upstream
CEP290-1232	-	AACCAGACAUCUAAGAGAAA	20	upstream
CEP290-1233	-	ACCAGACAUCUAAGAGAAA	19	upstream
CEP290-1234	-	AUUCUUAUCUAAGAUCCUUUCA	22	upstream
CEP290-1235	-	AAACAGGUAGAAUAUUGUAAUCAA	24	upstream
CEP290-1236	-	AACAGGUAGAAUAUUGUAAUCAA	23	upstream
CEP290-1237	-	ACAGGUAGAAUAUUGUAAUCAA	22	upstream
CEP290-1238	-	AGGUAGAAUAUUGUAAUCAA	20	upstream
CEP290-1239	+	AUGAGGAAGAUGAACAAAU	19	upstream
CEP290-1240	+	AGAGGAUAGGACAGAGGAC	19	upstream
CEP290-1241	+	CAGAGUGCAUCCAUGGUCC	19	upstream
CEP290-1242	+	CUUGCCUAGGACUUUCUAAUGCUG	24	upstream
CEP290-1243	+	CCUAGGACUUUCUAAUGCUG	20	upstream
CEP290-1244	+	CUAGGACUUUCUAAUGCUG	19	upstream
CEP290-1245	-	CCACUUGAACUCUAUACCUUUUA	23	upstream
CEP290-1246	-	CACUUGAACUCUAUACCUUUUA	22	upstream
CEP290-1247	-	CUUGAACUCUAUACCUUUUA	20	upstream
CEP290-1248	+	CAGCUAAAUCAUGCAAGUGACCU	23	upstream
CEP290-1249	+	CUAAAUCAUGCAAGUGACCU	20	upstream
CEP290-1250	+	CUCUUGCCUAGGACUUUCUAAUG	23	upstream
CEP290-1251	+	CUUGCCUAGGACUUUCUAAUG	21	upstream
CEP290-1252	+	CCAUAAGCCUCUAUUUCUGAUG	22	upstream
CEP290-1253	+	CAUAAGCCUCUAUUUCUGAUG	21	upstream
CEP290-1254	+	CUAAUGCUGGAGAGGAUAGGA	21	upstream
CEP290-1255	+	CCAUAAGCCUCUAUUUCUG	19	upstream
CEP290-1256	+	CAUAAGCCUCUAUUUCUG	18	upstream
CEP290-1257	-	CAGGUAGAAUAUUGUAAUCA	20	upstream
CEP290-1258	-	CUUUCUGCUGCUUUUGCCAAA	21	upstream
CEP290-1259	+	CCAAGGAACAAAAGCCAGGGACCA	24	upstream
CEP290-1260	+	CAAGGAACAAAAGCCAGGGACCA	23	upstream

CEP290-1261	+	CUCUUAGAUGUCUGGUUAA	19	upstream
CEP290-1262	-	CAUGUUUAUCAAUAUUAUU	19	upstream
CEP290-1263	-	CCAGACAUCUAAGAGAAA	18	upstream
CEP290-1264	-	CUUAUCUAAGAUCCUUUCA	19	upstream
CEP290-1265	-	CAGGUAGAAUAUUGUAAUCAA	21	upstream
CEP290-1266	+	CUGAUGAGGAAGAUGAACAAAU	22	upstream
CEP290-1267	+	CUGGAGAGGAUAGGACAGAGGAC	23	upstream
CEP290-1268	-	CAUCUUCCUCAUCAGAAA	18	upstream
CEP290-1269	+	GCCUAGGACUUUCUAAUGCUG	21	upstream
CEP290-1270	-	GCCACUUGAACUCUAUACCUUUUA	24	upstream
CEP290-1271	+	GCUAAAUCAUGCAAGUGACCU	21	upstream
CEP290-1272	+	GCCUAGGACUUUCUAAUG	18	upstream
CEP290-1273	+	GGGAGAAUAGUUUGUUCUGGGUA	23	upstream
CEP290-1274	+	GGAGAAUAGUUUGUUCUGGGUA	22	upstream
CEP290-1275	+	GAGAAUAGUUUGUUCUGGGUA	21	upstream
CEP290-1276	+	GAAUAGUUUGUUCUGGGUA	19	upstream
CEP290-1277	+	GGAGAAUGAUCUAGAUAAUCAUU	23	upstream
CEP290-1278	+	GAGAAUGAUCUAGAUAAUCAUU	22	upstream
CEP290-1279	+	GAAUGAUCUAGAUAAUCAUU	20	upstream
CEP290-1280	-	GAAACAGGUAGAAUAUUGUAAUCA	24	upstream
CEP290-1281	-	GGUAGAAUAUUGUAAUCA	18	upstream
CEP290-1282	-	GCUUUCUGCUGCUUUUGCCAAA	22	upstream
CEP290-1283	+	GGAACAAAAGCCAGGGACCA	20	upstream
CEP290-1284	+	GAACAAAAGCCAGGGACCA	19	upstream
CEP290-1285	-	GGUAGAAUAUUGUAAUCAAAGGA	23	upstream
CEP290-1286	-	GUAGAAUAUUGUAAUCAAAGGA	22	upstream
CEP290-1287	-	GAAUAUUGUAAUCAAAGGA	19	upstream
CEP290-1288	-	GAGUCAUGUUUAUCAAUAUUAUU	23	upstream
CEP290-1289	-	GUCAUGUUUAUCAAUAUUAUU	21	upstream
CEP290-1290	-	GGUAGAAUAUUGUAAUCAA	19	upstream
CEP290-1291	-	GUAGAAUAUUGUAAUCAA	18	upstream
CEP290-1292	+	GAUGAGGAAGAUGAACAAAU	20	upstream
CEP290-1293	+	GCUGGAGAGGAUAGGACAGAGGAC	24	upstream
CEP290-1294	+	GGAGAGGAUAGGACAGAGGAC	21	upstream
CEP290-1295	+	GAGAGGAUAGGACAGAGGAC	20	upstream
CEP290-1296	+	GAGGAUAGGACAGAGGAC	18	upstream
CEP290-1297	-	GUUCAUCUUCCUCAUCAGAAA	21	upstream
CEP290-1298	+	UUUACAGAGUGCAUCCAUGGUCC	23	upstream
CEP290-1299	+	UUACAGAGUGCAUCCAUGGUCC	22	upstream
CEP290-1300	+	UACAGAGUGCAUCCAUGGUCC	21	upstream
CEP290-1301	+	UUGCCUAGGACUUUCUAAUGCUG	23	upstream
CEP290-1302	+	UGCCUAGGACUUUCUAAUGCUG	22	upstream

CEP290-1303	+	UAGGACUUUCUAAUGCUG	18	upstream
CEP290-1304	-	UUGAACUCUAUACCUUUUA	19	upstream
CEP290-1305	-	UGAACUCUAUACCUUUUA	18	upstream
CEP290-1306	+	UCAGCUAAAUCAUGCAAGUGACCU	24	upstream
CEP290-1307	+	UAAAUCAUGCAAGUGACCU	19	upstream
CEP290-1308	+	UCUCUUGCCUAGGACUUUCUAAUG	24	upstream
CEP290-1309	+	UCUUGCCUAGGACUUUCUAAUG	22	upstream
CEP290-1310	+	UUGCCUAGGACUUUCUAAUG	20	upstream
CEP290-1311	+	UGCCUAGGACUUUCUAAUG	19	upstream
CEP290-1312	+	UCCAUAAGCCUCUAUUUCUGAUG	23	upstream
CEP290-1313	+	UAAGCCUCUAUUUCUGAUG	19	upstream
CEP290-1314	+	UGGGAGAAUAGUUUGUUCUGGGUA	24	upstream
CEP290-1315	+	UUUCUAAUGCUGGAGAGGAUAGGA	24	upstream
CEP290-1316	+	UUCUAAUGCUGGAGAGGAUAGGA	23	upstream
CEP290-1317	+	UCUAAUGCUGGAGAGGAUAGGA	22	upstream
CEP290-1318	+	UAAUGCUGGAGAGGAUAGGA	20	upstream
CEP290-1319	+	UCCAUAAGCCUCUAUUUCUG	20	upstream
CEP290-1320	-	UUGCUUUCUGCUGCUUUUGCCAAA	24	upstream
CEP290-1321	-	UGCUUUCUGCUGCUUUUGCCAAA	23	upstream
CEP290-1322	-	UUUCUGCUGCUUUUGCCAAA	20	upstream
CEP290-1323	-	UUCUGCUGCUUUUGCCAAA	19	upstream
CEP290-1324	-	UCUGCUGCUUUUGCCAAA	18	upstream
CEP290-1325	-	UAGAAUAUUGUAAUCAAAGGA	21	upstream
CEP290-1326	+	UUUUUCUCUUAGAUGUCUGGUUAA	24	upstream
CEP290-1327	+	UUUUCUCUUAGAUGUCUGGUUAA	23	upstream
CEP290-1328	+	UUUCUCUUAGAUGUCUGGUUAA	22	upstream
CEP290-1329	+	UUCUCUUAGAUGUCUGGUUAA	21	upstream
CEP290-1330	+	UCUCUUAGAUGUCUGGUUAA	20	upstream
CEP290-1331	+	UCUUAGAUGUCUGGUUAA	18	upstream
CEP290-1332	-	UGAGUCAUGUUUAUCAAUAUUAUU	24	upstream
CEP290-1333	-	UCAUGUUUAUCAAUAUUAUU	20	upstream
CEP290-1334	-	UUUUAACCAGACAUCUAAGAGAAA	24	upstream
CEP290-1335	-	UUUAACCAGACAUCUAAGAGAAA	23	upstream
CEP290-1336	-	UUAACCAGACAUCUAAGAGAAA	22	upstream
CEP290-1337	-	UAACCAGACAUCUAAGAGAAA	21	upstream
CEP290-1338	-	UUAUUCUUAUCUAAGAUCCUUUCA	24	upstream
CEP290-1339	-	UAUUCUUAUCUAAGAUCCUUUCA	23	upstream
CEP290-1340	-	UUCUUAUCUAAGAUCCUUUCA	21	upstream
CEP290-1341	-	UCUUAUCUAAGAUCCUUUCA	20	upstream
CEP290-1342	-	UUAUCUAAGAUCCUUUCA	18	upstream
CEP290-1343	+	UUCUGAUGAGGAAGAUGAACAAAU	24	upstream
CEP290-1344	+	UCUGAUGAGGAAGAUGAACAAAU	23	upstream

CEP290-1345	+	UGAUGAGGAAGAUGAACAAAU	21	upstream
CEP290-1346	+	UGAGGAAGAUGAACAAAU	18	upstream
CEP290-1347	+	UGGAGAGGAUAGGACAGAGGAC	22	upstream
CEP290-1348	-	UUUGUUCAUCUUCCUCAUCAGAAA	24	upstream
CEP290-1349	-	UUGUUCAUCUUCCUCAUCAGAAA	23	upstream
CEP290-1350	-	UGUUCAUCUUCCUCAUCAGAAA	22	upstream
CEP290-1351	-	UUCAUCUUCCUCAUCAGAAA	20	upstream
CEP290-1352	-	UCAUCUUCCUCAUCAGAAA	19	upstream
CEP290-1353	-	ACUUACCUCAUGUCAUCUAGAGC	23	downstream
CEP290-1354	-	ACCUCAUGUCAUCUAGAGC	19	downstream
CEP290-1355	+	ACAGUUUUUAAGGCGGGGAGUCAC	24	downstream
CEP290-1356	+	AGUUUUUAAGGCGGGGAGUCAC	22	downstream
CEP290-1357	-	ACAGAGUUCAAGCUAAUAC	19	downstream
CEP290-1358	+	AUUAGCUUGAACUCUGUGCCAAAC	24	downstream
CEP290-1359	+	AGCUUGAACUCUGUGCCAAAC	21	downstream
CEP290-1360	-	AUGUGGUGUCAAAUAUGGUGCU	22	downstream
CEP290-1361	-	AUGUGGUGUCAAAUAUGGUGCUU	23	downstream
CEP290-1362	+	AGAUGACAUGAGGUAAGU	18	downstream
CEP290-1363	-	AAUACAUGAGAGUGAUUAGUGG	22	downstream
CEP290-1364	-	AUACAUGAGAGUGAUUAGUGG	21	downstream
CEP290-1365	-	ACAUGAGAGUGAUUAGUGG	19	downstream
CEP290-16	+	AAGACACUGCCAAUAGGGAUAGGU	24	downstream
CEP290-1366	+	AGACACUGCCAAUAGGGAUAGGU	23	downstream
CEP290-1367	+	ACACUGCCAAUAGGGAUAGGU	21	downstream
CEP290-510	+	ACUGCCAAUAGGGAUAGGU	19	downstream
CEP290-1368	-	AAAGGUUCAUGAGACUAGAGGUC	23	downstream
CEP290-1369	-	AAGGUUCAUGAGACUAGAGGUC	22	downstream
CEP290-1370	-	AGGUUCAUGAGACUAGAGGUC	21	downstream
CEP290-1371	+	AAACAGGAGAUACUCAACACA	21	downstream
CEP290-1372	+	AACAGGAGAUACUCAACACA	20	downstream
CEP290-1373	+	ACAGGAGAUACUCAACACA	19	downstream
CEP290-1374	+	AGCACGUACAAAAGAACAUACAU	23	downstream
CEP290-1375	+	ACGUACAAAAGAACAUACAU	20	downstream
CEP290-1376	+	AGUAAGGAGGAUGUAAGAC	19	downstream
CEP290-1377	+	AGCUUUUGACAGUUUUUAAGG	21	downstream
CEP290-1378	-	ACGUGCUCUUUUCUAUAUAU	20	downstream
CEP290-1379	+	AAAUUCACUGAGCAAAACAACUGG	24	downstream
CEP290-1380	+	AAUUCACUGAGCAAAACAACUGG	23	downstream
CEP290-1381	+	AUUCACUGAGCAAAACAACUGG	22	downstream
CEP290-1382	+	ACUGAGCAAAACAACUGG	18	downstream
CEP290-1383	+	AACAAGUUUUGAAACAGGAA	20	downstream
CEP290-1384	+	ACAAGUUUUGAAACAGGAA	19	downstream

CEP290-1385	+	AAUGCCUGAACAAGUUUUGAAA	22	downstream
CEP290-1386	+	AUGCCUGAACAAGUUUUGAAA	21	downstream
CEP290-1387	+	AUUCACUGAGCAAAACAACUGGAA	24	downstream
CEP290-1388	+	ACUGAGCAAAACAACUGGAA	20	downstream
CEP290-1389	+	AAAAAGGUAAUGCCUGAACAAGUU	24	downstream
CEP290-1390	+	AAAAGGUAAUGCCUGAACAAGUU	23	downstream
CEP290-1391	+	AAAGGUAAUGCCUGAACAAGUU	22	downstream
CEP290-1392	+	AAGGUAAUGCCUGAACAAGUU	21	downstream
CEP290-1393	+	AGGUAAUGCCUGAACAAGUU	20	downstream
CEP290-1394	-	ACGUGCUCUUUUCUAUAUA	19	downstream
CEP290-1395	+	AUUAUCUAUUCCAUUCUUCACAC	23	downstream
CEP290-1396	+	AUCUAUUCCAUUCUUCACAC	20	downstream
CEP290-1397	+	AAGAGAGAAAUGGUUCCCUAUAUA	24	downstream
CEP290-1398	+	AGAGAGAAAUGGUUCCCUAUAUA	23	downstream
CEP290-1399	+	AGAGAAAUGGUUCCCUAUAUA	21	downstream
CEP290-1400	+	AGAAAUGGUUCCCUAUAUA	19	downstream
CEP290-1401	-	AGGAAAUUAUUGUUGCUUU	19	downstream
CEP290-1402	+	ACUGAGCAAAACAACUGGAAGA	22	downstream
CEP290-1403	+	AGCAAAACAACUGGAAGA	18	downstream
CEP290-1404	+	AUACAUAAGAAAGAACACUGUGGU	24	downstream
CEP290-1405	+	ACAUAAGAAAGAACACUGUGGU	22	downstream
CEP290-1406	+	AUAAGAAAGAACACUGUGGU	20	downstream
CEP290-1407	+	AAGAAAGAACACUGUGGU	18	downstream
CEP290-1408	-	AAGAAUGGAAUAGAUAAU	18	downstream
CEP290-1409	+	AAGGAGGAUGUAAGACUGGAGA	22	downstream
CEP290-1410	+	AGGAGGAUGUAAGACUGGAGA	21	downstream
CEP290-1411	+	AGGAUGUAAGACUGGAGA	18	downstream
CEP290-1412	-	AAAAACUUGAAAUUUGAUAGUAG	23	downstream
CEP290-1413	-	AAAACUUGAAAUUUGAUAGUAG	22	downstream
CEP290-1414	-	AAACUUGAAAUUUGAUAGUAG	21	downstream
CEP290-1415	-	AACUUGAAAUUUGAUAGUAG	20	downstream
CEP290-1416	-	ACUUGAAAUUUGAUAGUAG	19	downstream
CEP290-1417	-	ACAUAUCUGUCUUCCUUA	18	downstream
CEP290-1418	+	AUUAAAAAAAGUAUGCUU	18	downstream
CEP290-1419	+	AUAUCAAAAGACUUAUAUUCCAUU	24	downstream
CEP290-1420	+	AUCAAAAGACUUAUAUUCCAUU	22	downstream
CEP290-1421	+	AAAAGACUUAUAUUCCAUU	19	downstream
CEP290-1422	+	AAAGACUUAUAUUCCAUU	18	downstream
CEP290-1423	-	AAAAUCAGAUUUCAUGUGUGAAGA	24	downstream
CEP290-1424	-	AAAUCAGAUUUCAUGUGUGAAGA	23	downstream
CEP290-1425	-	AAUCAGAUUUCAUGUGUGAAGA	22	downstream
CEP290-1426	-	AUCAGAUUUCAUGUGUGAAGA	21	downstream

CEP290-1427	-	AGAUUUCAUGUGUGAAGA	18	downstream
CEP290-1428	-	AAUGGAAUAUAAGUCUUUUGAUAU	24	downstream
CEP290-1429	-	AUGGAAUAUAAGUCUUUUGAUAU	23	downstream
CEP290-1430	-	AAUAUAAGUCUUUUGAUAU	19	downstream
CEP290-1431	-	AUAUAAGUCUUUUGAUAU	18	downstream
CEP290-1432	-	AAGAAUGGAAUAGAUAAUA	19	downstream
CEP290-1433	-	AGAAUGGAAUAGAUAAUA	18	downstream
CEP290-1434	-	AAAACUGGAUGGGUAAUAAAGCAA	24	downstream
CEP290-1435	-	AAACUGGAUGGGUAAUAAAGCAA	23	downstream
CEP290-1436	-	AACUGGAUGGGUAAUAAAGCAA	22	downstream
CEP290-1437	-	ACUGGAUGGGUAAUAAAGCAA	21	downstream
CEP290-1438	+	AUAGAAAUUCACUGAGCAAAACAA	24	downstream
CEP290-1439	+	AGAAAUUCACUGAGCAAAACAA	22	downstream
CEP290-1440	+	AAAUUCACUGAGCAAAACAA	20	downstream
CEP290-1441	+	AAUUCACUGAGCAAAACAA	19	downstream
CEP290-1442	+	AUUCACUGAGCAAAACAA	18	downstream
CEP290-1443	+	AGGAUGUAAGACUGGAGAUAGAGA	24	downstream
CEP290-1444	+	AUGUAAGACUGGAGAUAGAGA	21	downstream
CEP290-1445	-	AAAUUUGAUAGUAGAAGAAAA	21	downstream
CEP290-1446	-	AAUUUGAUAGUAGAAGAAAA	20	downstream
CEP290-1447	-	AUUUGAUAGUAGAAGAAAA	19	downstream
CEP290-1448	+	AAAAUAAAACUAAGACACUGCCAA	24	downstream
CEP290-1449	+	AAAUAAAACUAAGACACUGCCAA	23	downstream
CEP290-1450	+	AAUAAAACUAAGACACUGCCAA	22	downstream
CEP290-1451	+	AUAAAACUAAGACACUGCCAA	21	downstream
CEP290-1452	+	AAAACUAAGACACUGCCAA	19	downstream
CEP290-1453	+	AAACUAAGACACUGCCAA	18	downstream
CEP290-1454	-	AAUAAAGCAAAAGAAAAAC	19	downstream
CEP290-1455	-	AUAAAGCAAAAGAAAAAC	18	downstream
CEP290-1456	-	AUUCUUUUUUUGUUGUUUUUUUUU	24	downstream
CEP290-1457	+	ACUCCAGCCUGGGCAACACA	20	downstream
CEP290-1458	-	CUUACCUCAUGUCAUCUAGAGC	22	downstream
CEP290-1459	-	CCUCAUGUCAUCUAGAGC	18	downstream
CEP290-1460	+	CAGUUUUUAAGGCGGGGAGUCAC	23	downstream
CEP290-1461	-	CACAGAGUUCAAGCUAAUAC	20	downstream
CEP290-1462	-	CAGAGUUCAAGCUAAUAC	18	downstream
CEP290-1463	+	CUUGAACUCUGUGCCAAAC	19	downstream
CEP290-1464	-	CAUGUGGUGUCAAAUAUGGUGCU	23	downstream
CEP290-1465	-	CAUGUGGUGUCAAAUAUGGUGCUU	24	downstream
CEP290-1466	+	CUCUAGAUGACAUGAGGUAAGU	22	downstream
CEP290-1467	+	CUAGAUGACAUGAGGUAAGU	20	downstream
CEP290-1468	-	CUAAUACAUGAGAGUGAUUAGUGG	24	downstream

CEP290-1469	-	CAUGAGAGUGAUUAGUGG	18	downstream
CEP290-509	+	CACUGCCAAUAGGGAUAGGU	20	downstream
CEP290-511	+	CUGCCAAUAGGGAUAGGU	18	downstream
CEP290-1470	+	CCAAACAGGAGAUACUCAACACA	23	downstream
CEP290-1471	+	CAAACAGGAGAUACUCAACACA	22	downstream
CEP290-1472	+	CAGGAGAUACUCAACACA	18	downstream
CEP290-1473	+	CACGUACAAAAGAACAUACAU	21	downstream
CEP290-1474	+	CGUACAAAAGAACAUACAU	19	downstream
CEP290-1475	+	CAGUAAGGAGGAUGUAAGAC	20	downstream
CEP290-1476	+	CUUUUGACAGUUUUUAAGG	19	downstream
CEP290-1477	-	CGUGCUCUUUUCUAUAUAU	19	downstream
CEP290-1478	+	CACUGAGCAAAACAACUGG	19	downstream
CEP290-1479	+	CCUGAACAAGUUUUGAAACAGGAA	24	downstream
CEP290-1480	+	CUGAACAAGUUUUGAAACAGGAA	23	downstream
CEP290-1481	+	CAAGUUUUGAAACAGGAA	18	downstream
CEP290-1482	+	CCUGAACAAGUUUUGAAA	18	downstream
CEP290-1483	+	CACUGAGCAAAACAACUGGAA	21	downstream
CEP290-1484	+	CUGAGCAAAACAACUGGAA	19	downstream
CEP290-1485	-	CGUGCUCUUUUCUAUAUA	18	downstream
CEP290-1486	+	CUAUUCCAUUCUUCACAC	18	downstream
CEP290-1487	-	CUUAGGAAAUUAUUGUUGCUUU	22	downstream
CEP290-1488	-	CUUUUUGAGAGGUAAAGGUUC	21	downstream
CEP290-1489	+	CACUGAGCAAAACAACUGGAAGA	23	downstream
CEP290-1490	+	CUGAGCAAAACAACUGGAAGA	21	downstream
CEP290-1491	+	CAUAAGAAAGAACACUGUGGU	21	downstream
CEP290-1492	-	CUUGAAAUUUGAUAGUAG	18	downstream
CEP290-1493	+	CCAUUAAAAAAAGUAUGCUU	20	downstream
CEP290-1494	+	CAUUAAAAAAAGUAUGCUU	19	downstream
CEP290-1495	+	CAAAAGACUUAUAUUCCAUU	20	downstream
CEP290-1496	-	CAGAUUUCAUGUGUGAAGA	19	downstream
CEP290-1497	-	CUGGAUGGGUAAUAAAGCAA	20	downstream
CEP290-1498	-	CUUAAGCAUACUUUUUUUA	19	downstream
CEP290-1499	-	CUUUUUUUGUUGUUUUUUUUU	21	downstream
CEP290-1500	+	CUGCACUCCAGCCUGGGCAACACA	24	downstream
CEP290-1501	+	CACUCCAGCCUGGGCAACACA	21	downstream
CEP290-1502	+	CUCCAGCCUGGGCAACACA	19	downstream
CEP290-1503	+	GUUUUUAAGGCGGGGAGUCAC	21	downstream
CEP290-230	-	GGCACAGAGUUCAAGCUAAUAC	22	downstream
CEP290-1504	-	GCACAGAGUUCAAGCUAAUAC	21	downstream
CEP290-1505	+	GCUUGAACUCUGUGCCAAAC	20	downstream
CEP290-139	-	GCAUGUGGUGUCAAAUAUGGUGCU	24	downstream
CEP290-1506	-	GUGGUGUCAAAUAUGGUGCU	20	downstream

CEP290-1507	-	GGUGUCAAAUAUGGUGCU	18	downstream
CEP290-1508	-	GUGGUGUCAAAUAUGGUGCUU	21	downstream
CEP290-1509	-	GGUGUCAAAUAUGGUGCUU	19	downstream
CEP290-1510	-	GUGUCAAAUAUGGUGCUU	18	downstream
CEP290-1511	+	GCUCUAGAUGACAUGAGGUAAGU	23	downstream
CEP290-11	+	GACACUGCCAAUAGGGAUAGGU	22	downstream
CEP290-1512	-	GGUUCAUGAGACUAGAGGUC	20	downstream
CEP290-1513	-	GUUCAUGAGACUAGAGGUC	19	downstream
CEP290-1514	+	GCCAAACAGGAGAUACUCAACACA	24	downstream
CEP290-1515	+	GAGCACGUACAAAAGAACAUACAU	24	downstream
CEP290-1516	+	GCACGUACAAAAGAACAUACAU	22	downstream
CEP290-1517	+	GUACAAAAGAACAUACAU	18	downstream
CEP290-1518	+	GUGGCAGUAAGGAGGAUGUAAGAC	24	downstream
CEP290-1519	+	GGCAGUAAGGAGGAUGUAAGAC	22	downstream
CEP290-1520	+	GCAGUAAGGAGGAUGUAAGAC	21	downstream
CEP290-1521	+	GUAAGGAGGAUGUAAGAC	18	downstream
CEP290-1522	+	GGUAGCUUUUGACAGUUUUUAAGG	24	downstream
CEP290-1523	+	GUAGCUUUUGACAGUUUUUAAGG	23	downstream
CEP290-1524	+	GCUUUUGACAGUUUUUAAGG	20	downstream
CEP290-1525	-	GUACGUGCUCUUUUCUAUAUAU	22	downstream
CEP290-1526	-	GUGCUCUUUUCUAUAUAU	18	downstream
CEP290-1527	+	GAACAAGUUUUGAAACAGGAA	21	downstream
CEP290-1528	+	GUAAUGCCUGAACAAGUUUUGAAA	24	downstream
CEP290-1529	+	GCCUGAACAAGUUUUGAAA	19	downstream
CEP290-1530	+	GGUAAUGCCUGAACAAGUU	19	downstream
CEP290-1531	+	GUAAUGCCUGAACAAGUU	18	downstream
CEP290-1532	-	GUACGUGCUCUUUUCUAUAUA	21	downstream
CEP290-1533	+	GAGAGAAAUGGUUCCCUAUAUA	22	downstream
CEP290-1534	+	GAGAAAUGGUUCCCUAUAUA	20	downstream
CEP290-1535	+	GAAAUGGUUCCCUAUAUA	18	downstream
CEP290-1536	-	GCUUAGGAAAUUAUUGUUGCUUU	23	downstream
CEP290-1537	-	GGAAAUUAUUGUUGCUUU	18	downstream
CEP290-1538	-	GCUUUUUGAGAGGUAAAGGUUC	22	downstream
CEP290-1539	+	GAGCAAAACAACUGGAAGA	19	downstream
CEP290-1540	-	GUGUGAAGAAUGGAAUAGAUAAU	23	downstream
CEP290-1541	-	GUGAAGAAUGGAAUAGAUAAU	21	downstream
CEP290-1542	-	GAAGAAUGGAAUAGAUAAU	19	downstream
CEP290-1543	+	GUAAGGAGGAUGUAAGACUGGAGA	24	downstream
CEP290-1544	+	GGAGGAUGUAAGACUGGAGA	20	downstream
CEP290-1545	+	GAGGAUGUAAGACUGGAGA	19	downstream
CEP290-1546	-	GAAAAACUUGAAAUUUGAUAGUAG	24	downstream
CEP290-1547	-	GUGUUUACAUAUCUGUCUUCCUUA	24	downstream

CEP290-1548	-	GUUUACAUAUCUGUCUUCCUUA	22	downstream
CEP290-1549	+	GUUCCAUUAAAAAAAGUAUGCUU	23	downstream
CEP290-1550	-	GGAAUAUAAGUCUUUUGAUAU	21	downstream
CEP290-1551	-	GAAUAUAAGUCUUUUGAUAU	20	downstream
CEP290-1552	-	GUGUGAAGAAUGGAAUAGAUAAUA	24	downstream
CEP290-1553	-	GUGAAGAAUGGAAUAGAUAAUA	22	downstream
CEP290-1554	-	GAAGAAUGGAAUAGAUAAUA	20	downstream
CEP290-1555	-	GGAUGGGUAAUAAAGCAA	18	downstream
CEP290-1556	+	GAAAUUCACUGAGCAAAACAA	21	downstream
CEP290-1557	+	GGAUGUAAGACUGGAGAUAGAGA	23	downstream
CEP290-1558	+	GAUGUAAGACUGGAGAUAGAGA	22	downstream
CEP290-1559	+	GUAAGACUGGAGAUAGAGA	19	downstream
CEP290-1560	-	GAAAUUUGAUAGUAGAAGAAAA	22	downstream
CEP290-1561	-	GGGUAAUAAAGCAAAAGAAAAAC	23	downstream
CEP290-1562	-	GGUAAUAAAGCAAAAGAAAAAC	22	downstream
CEP290-1563	-	GUAAUAAAGCAAAAGAAAAAC	21	downstream
CEP290-1564	+	GCACUCCAGCCUGGGCAACACA	22	downstream
CEP290-1565	-	UACUUACCUCAUGUCAUCUAGAGC	24	downstream
CEP290-1566	-	UUACCUCAUGUCAUCUAGAGC	21	downstream
CEP290-1567	-	UACCUCAUGUCAUCUAGAGC	20	downstream
CEP290-1568	+	UUUUUAAGGCGGGGAGUCAC	20	downstream
CEP290-1569	+	UUUUAAGGCGGGGAGUCAC	19	downstream
CEP290-1570	+	UUUAAGGCGGGGAGUCAC	18	downstream
CEP290-1571	-	UUGGCACAGAGUUCAAGCUAAUAC	24	downstream
CEP290-1572	-	UGGCACAGAGUUCAAGCUAAUAC	23	downstream
CEP290-1573	+	UUAGCUUGAACUCUGUGCCAAAC	23	downstream
CEP290-1574	+	UAGCUUGAACUCUGUGCCAAAC	22	downstream
CEP290-1575	+	UUGAACUCUGUGCCAAAC	18	downstream
CEP290-1576	-	UGUGGUGUCAAAUAUGGUGCU	21	downstream
CEP290-1577	-	UGGUGUCAAAUAUGGUGCU	19	downstream
CEP290-1578	-	UGUGGUGUCAAAUAUGGUGCUU	22	downstream
CEP290-1579	-	UGGUGUCAAAUAUGGUGCUU	20	downstream
CEP290-1580	+	UGCUCUAGAUGACAUGAGGUAAGU	24	downstream
CEP290-1581	+	UCUAGAUGACAUGAGGUAAGU	21	downstream
CEP290-1582	+	UAGAUGACAUGAGGUAAGU	19	downstream
CEP290-1583	-	UAAUACAUGAGAGUGAUUAGUGG	23	downstream
CEP290-1584	-	UACAUGAGAGUGAUUAGUGG	20	downstream
CEP290-1585	-	UAAAGGUUCAUGAGACUAGAGGUC	24	downstream
CEP290-1586	-	UUCAUGAGACUAGAGGUC	18	downstream
CEP290-1587	+	UGGCAGUAAGGAGGAUGUAAGAC	23	downstream
CEP290-1588	+	UAGCUUUUGACAGUUUUUAAGG	22	downstream
CEP290-1589	+	UUUUGACAGUUUUUAAGG	18	downstream

CEP290-1590	-	UUGUACGUGCUCUUUUCUAUAUAU	24	downstream
CEP290-1591	-	UGUACGUGCUCUUUUCUAUAUAU	23	downstream
CEP290-1592	-	UACGUGCUCUUUUCUAUAUAU	21	downstream
CEP290-1593	+	UUCACUGAGCAAAACAACUGG	21	downstream
CEP290-1594	+	UCACUGAGCAAAACAACUGG	20	downstream
CEP290-1595	+	UGAACAAGUUUUGAAACAGGAA	22	downstream
CEP290-1596	+	UAAUGCCUGAACAAGUUUUGAAA	23	downstream
CEP290-1597	+	UGCCUGAACAAGUUUUGAAA	20	downstream
CEP290-1598	+	UUCACUGAGCAAAACAACUGGAA	23	downstream
CEP290-1599	+	UCACUGAGCAAAACAACUGGAA	22	downstream
CEP290-1600	+	UGAGCAAAACAACUGGAA	18	downstream
CEP290-1601	-	UUUGUACGUGCUCUUUUCUAUAUA	24	downstream
CEP290-1602	-	UUGUACGUGCUCUUUUCUAUAUA	23	downstream
CEP290-1603	-	UGUACGUGCUCUUUUCUAUAUA	22	downstream
CEP290-1604	-	UACGUGCUCUUUUCUAUAUA	20	downstream
CEP290-1605	+	UAUUAUCUAUUCCAUUCUUCACAC	24	downstream
CEP290-1606	+	UUAUCUAUUCCAUUCUUCACAC	22	downstream
CEP290-1607	+	UAUCUAUUCCAUUCUUCACAC	21	downstream
CEP290-1608	+	UCUAUUCCAUUCUUCACAC	19	downstream
CEP290-1609	-	UGCUUAGGAAAUUAUUGUUGCUUU	24	downstream
CEP290-1610	-	UUAGGAAAUUAUUGUUGCUUU	21	downstream
CEP290-1611	-	UAGGAAAUUAUUGUUGCUUU	20	downstream
CEP290-1612	-	UUGCUUUUUGAGAGGUAAAGGUUC	24	downstream
CEP290-1613	-	UGCUUUUUGAGAGGUAAAGGUUC	23	downstream
CEP290-1614	-	UUUUUGAGAGGUAAAGGUUC	20	downstream
CEP290-1615	-	UUUUGAGAGGUAAAGGUUC	19	downstream
CEP290-1616	-	UUUGAGAGGUAAAGGUUC	18	downstream
CEP290-1617	+	UCACUGAGCAAAACAACUGGAAGA	24	downstream
CEP290-1618	+	UGAGCAAAACAACUGGAAGA	20	downstream
CEP290-1619	+	UACAUAAGAAAGAACACUGUGGU	23	downstream
CEP290-1620	+	UAAGAAAGAACACUGUGGU	19	downstream
CEP290-1621	-	UGUGUGAAGAAUGGAAUAGAUAAU	24	downstream
CEP290-1622	-	UGUGAAGAAUGGAAUAGAUAAU	22	downstream
CEP290-1623	-	UGAAGAAUGGAAUAGAUAAU	20	downstream
CEP290-1624	+	UAAGGAGGAUGUAAGACUGGAGA	23	downstream
CEP290-1625	-	UGUUUACAUAUCUGUCUUCCUUA	23	downstream
CEP290-1626	-	UUUACAUAUCUGUCUUCCUUA	21	downstream
CEP290-1627	-	UUACAUAUCUGUCUUCCUUA	20	downstream
CEP290-1628	-	UACAUAUCUGUCUUCCUUA	19	downstream
CEP290-1629	+	UGUUCCAUUAAAAAAAGUAUGCUU	24	downstream
CEP290-1630	+	UUCCAUUAAAAAAAGUAUGCUU	22	downstream
CEP290-1631	+	UCCAUUAAAAAAAGUAUGCUU	21	downstream

CEP290-1632	+	UAUCAAAAGACUUAUAUUCCAUU	23	downstream
CEP290-1633	+	UCAAAAGACUUAUAUUCCAUU	21	downstream
CEP290-1634	-	UCAGAUUUCAUGUGUGAAGA	20	downstream
CEP290-1635	-	UGGAAUAUAAGUCUUUUGAUAU	22	downstream
CEP290-1636	-	UGUGAAGAAUGGAAUAGAUAAUA	23	downstream
CEP290-1637	-	UGAAGAAUGGAAUAGAUAAUA	21	downstream
CEP290-1638	-	UGGAUGGGUAAUAAAGCAA	19	downstream
CEP290-1639	+	UAGAAAUUCACUGAGCAAAACAA	23	downstream
CEP290-1640	+	UGUAAGACUGGAGAUAGAGA	20	downstream
CEP290-1641	+	UAAGACUGGAGAUAGAGA	18	downstream
CEP290-1642	-	UUGAAAUUUGAUAGUAGAAGAAAA	24	downstream
CEP290-1643	-	UGAAAUUUGAUAGUAGAAGAAAA	23	downstream
CEP290-1644	-	UUUGAUAGUAGAAGAAAA	18	downstream
CEP290-1645	+	UAAAACUAAGACACUGCCAA	20	downstream
CEP290-1646	-	UUUUUCUUAAGCAUACUUUUUUUA	24	downstream
CEP290-1647	-	UUUUCUUAAGCAUACUUUUUUUA	23	downstream
CEP290-1648	-	UUUCUUAAGCAUACUUUUUUUA	22	downstream
CEP290-1649	-	UUCUUAAGCAUACUUUUUUUA	21	downstream
CEP290-1650	-	UCUUAAGCAUACUUUUUUUA	20	downstream
CEP290-1651	-	UUAAGCAUACUUUUUUUA	18	downstream
CEP290-1652	-	UGGGUAAUAAAGCAAAAGAAAAAC	24	downstream
CEP290-1653	-	UAAUAAAGCAAAAGAAAAAC	20	downstream
CEP290-1654	-	UUCUUUUUUUGUUGUUUUUUUUUU	23	downstream
CEP290-1655	-	UCUUUUUUUGUUGUUUUUUUUU	22	downstream
CEP290-1656	-	UUUUUUUGUUGUUUUUUUUU	20	downstream
CEP290-1657	-	UUUUUUGUUGUUUUUUUUU	19	downstream
CEP290-1658	-	UUUUUGUUGUUUUUUUUU	18	downstream
CEP290-1659	+	UGCACUCCAGCCUGGGCAACACA	23	downstream
CEP290-1660	+	UCCAGCCUGGGCAACACA	18	downstream
CEP290-1661	+	AUUUUCGUGACCUCUAGUCUC	21	downstream
CEP290-1662	+	ACUAAUCACUCUCAUGUAUUAGC	23	downstream
CEP290-1663	+	AAUCACUCUCAUGUAUUAGC	20	downstream
CEP290-1664	+	AUCACUCUCAUGUAUUAGC	19	downstream
CEP290-1665	+	AGAUGACAUGAGGUAAGUA	19	downstream
CEP290-1666	-	ACCUCAUGUCAUCUAGAGCAAGAG	24	downstream
CEP290-1667	-	AUGUCAUCUAGAGCAAGAG	19	downstream
CEP290-1668	-	AAUACAUGAGAGUGAUUAGUGGUG	24	downstream
CEP290-1669	-	AUACAUGAGAGUGAUUAGUGGUG	23	downstream
CEP290-1670	-	ACAUGAGAGUGAUUAGUGGUG	21	downstream
CEP290-1671	-	AUGAGAGUGAUUAGUGGUG	19	downstream
CEP290-1672	-	ACGUGCUCUUUUCUAUAUAUA	21	downstream
CEP290-1673	+	ACAAAACCUAUGUAUAAGAUG	21	downstream

CEP290-1674	+	AAAACCUAUGUAUAAGAUG	19	downstream
CEP290-1675	+	AAACCUAUGUAUAAGAUG	18	downstream
CEP290-1676	+	AUAUAUAGAAAAGAGCACGUACAA	24	downstream
CEP290-1677	+	AUAUAGAAAAGAGCACGUACAA	22	downstream
CEP290-1678	+	AUAGAAAAGAGCACGUACAA	20	downstream
CEP290-1679	+	AGAAAAGAGCACGUACAA	18	downstream
CEP290-1680	+	AGAAAUGGUUCCCUAUAUAUAGAA	24	downstream
CEP290-1681	+	AAAUGGUUCCCUAUAUAUAGAA	22	downstream
CEP290-1682	+	AAUGGUUCCCUAUAUAUAGAA	21	downstream
CEP290-1683	+	AUGGUUCCCUAUAUAUAGAA	20	downstream
CEP290-1684	-	AUGGAAUAUAAGUCUUUUGAUAUA	24	downstream
CEP290-1685	-	AAUAUAAGUCUUUUGAUAUA	20	downstream
CEP290-1686	-	AUAUAAGUCUUUUGAUAUA	19	downstream
CEP290-1687	+	ACGUACAAAAGAACAUACAUAAGA	24	downstream
CEP290-1688	+	ACAAAAGAACAUACAUAAGA	20	downstream
CEP290-1689	+	AAAAGAACAUACAUAAGA	18	downstream
CEP290-1690	+	AAGAAAAAAAAGGUAAUGC	19	downstream
CEP290-1691	+	AGAAAAAAAAGGUAAUGC	18	downstream
CEP290-1692	+	AAACAGGAAUAGAAAUUCA	19	downstream
CEP290-1693	+	AACAGGAAUAGAAAUUCA	18	downstream
CEP290-1694	+	AAGAUCACUCCACUGCACUCCAGC	24	downstream
CEP290-1695	+	AGAUCACUCCACUGCACUCCAGC	23	downstream
CEP290-1696	+	AUCACUCCACUGCACUCCAGC	21	downstream
CEP290-1697	+	ACUCCACUGCACUCCAGC	18	downstream
CEP290-1698	-	CCCCUACUUACCUCAUGUCAUC	22	downstream
CEP290-1699	-	CCCUACUUACCUCAUGUCAUC	21	downstream
CEP290-1700	-	CCUACUUACCUCAUGUCAUC	20	downstream
CEP290-1701	-	CUACUUACCUCAUGUCAUC	19	downstream
CEP290-1702	+	CUGAUUUUCGUGACCUCUAGUCUC	24	downstream
CEP290-1703	+	CACUAAUCACUCUCAUGUAUUAGC	24	downstream
CEP290-1704	+	CUAAUCACUCUCAUGUAUUAGC	22	downstream
CEP290-1705	+	CUCUAGAUGACAUGAGGUAAGUA	23	downstream
CEP290-1706	+	CUAGAUGACAUGAGGUAAGUA	21	downstream
CEP290-1707	-	CCUCAUGUCAUCUAGAGCAAGAG	23	downstream
CEP290-1708	-	CUCAUGUCAUCUAGAGCAAGAG	22	downstream
CEP290-1709	-	CAUGUCAUCUAGAGCAAGAG	20	downstream
CEP290-1710	-	CAUGAGAGUGAUUAGUGGUG	20	downstream
CEP290-1711	-	CGUGCUCUUUUCUAUAUAUA	20	downstream
CEP290-1712	+	CAAAACCUAUGUAUAAGAUG	20	downstream
CEP290-1713	+	CGUACAAAAGAACAUACAUAAGA	23	downstream
CEP290-1714	+	CAAAAGAACAUACAUAAGA	19	downstream
CEP290-1715	+	CUUAAGAAAAAAAAGGUAAUGC	22	downstream

CEP290-1716	-	CUUAAGCAUACUUUUUUUAA	20	downstream
CEP290-1717	+	CACUCCACUGCACUCCAGC	19	downstream
CEP290-132	-	GUCCCCUACUUACCUCAUGUCAUC	24	downstream
CEP290-1718	+	GAUUUUCGUGACCUCUAGUCUC	22	downstream
CEP290-1719	+	GCUCUAGAUGACAUGAGGUAAGUA	24	downstream
CEP290-1720	+	GAUGACAUGAGGUAAGUA	18	downstream
CEP290-1721	-	GUACGUGCUCUUUUCUAUAUAUA	23	downstream
CEP290-1722	-	GUGCUCUUUUCUAUAUAUA	19	downstream
CEP290-1723	+	GUACAAAACCUAUGUAUAAGAUG	23	downstream
CEP290-1724	+	GAAAUGGUUCCCUAUAUAUAGAA	23	downstream
CEP290-1725	+	GGUUCCCUAUAUAUAGAA	18	downstream
CEP290-1726	-	GGAAUAUAAGUCUUUUGAUAUA	22	downstream
CEP290-1727	-	GAAUAUAAGUCUUUUGAUAUA	21	downstream
CEP290-1728	+	GUACAAAAGAACAUACAUAAGA	22	downstream
CEP290-1729	+	GCUUAAGAAAAAAAAGGUAAUGC	23	downstream
CEP290-1730	+	GAAACAGGAAUAGAAAUUCA	20	downstream
CEP290-1731	+	GAUCACUCCACUGCACUCCAGC	22	downstream
CEP290-1732	-	UCCCCUACUUACCUCAUGUCAUC	23	downstream
CEP290-1733	-	UACUUACCUCAUGUCAUC	18	downstream
CEP290-1734	+	UGAUUUUCGUGACCUCUAGUCUC	23	downstream
CEP290-1735	+	UUUUCGUGACCUCUAGUCUC	20	downstream
CEP290-1736	+	UUUCGUGACCUCUAGUCUC	19	downstream
CEP290-1737	+	UUCGUGACCUCUAGUCUC	18	downstream
CEP290-1738	+	UAAUCACUCUCAUGUAUUAGC	21	downstream
CEP290-1739	+	UCACUCUCAUGUAUUAGC	18	downstream
CEP290-1740	+	UCUAGAUGACAUGAGGUAAGUA	22	downstream
CEP290-1741	+	UAGAUGACAUGAGGUAAGUA	20	downstream
CEP290-1742	-	UCAUGUCAUCUAGAGCAAGAG	21	downstream
CEP290-1743	-	UGUCAUCUAGAGCAAGAG	18	downstream
CEP290-1744	-	UACAUGAGAGUGAUUAGUGGUG	22	downstream
CEP290-1745	-	UGAGAGUGAUUAGUGGUG	18	downstream
CEP290-1746	-	UGUACGUGCUCUUUUCUAUAUAUA	24	downstream
CEP290-1747	-	UACGUGCUCUUUUCUAUAUAUA	22	downstream
CEP290-1748	-	UGCUCUUUUCUAUAUAUA	18	downstream
CEP290-1749	+	UGUACAAAACCUAUGUAUAAGAUG	24	downstream
CEP290-1750	+	UACAAAACCUAUGUAUAAGAUG	22	downstream
CEP290-1751	+	UAUAUAGAAAAGAGCACGUACAA	23	downstream
CEP290-1752	+	UAUAGAAAAGAGCACGUACAA	21	downstream
CEP290-1753	+	UAGAAAAGAGCACGUACAA	19	downstream
CEP290-1754	+	UGGUUCCCUAUAUAUAGAA	19	downstream
CEP290-1755	-	UGGAAUAUAAGUCUUUUGAUAUA	23	downstream
CEP290-1756	-	UAUAAGUCUUUUGAUAUA	18	downstream

CEP290-1757	+	UACAAAAGAACAUACAUAAGA	21	downstream
CEP290-1758	+	UGCUUAAGAAAAAAAAGGUAAUGC	24	downstream
CEP290-1759	+	UUAAGAAAAAAAAGGUAAUGC	21	downstream
CEP290-1760	+	UAAGAAAAAAAAGGUAAUGC	20	downstream
CEP290-1761	+	UUUUGAAACAGGAAUAGAAAUUCA	24	downstream
CEP290-1762	+	UUUGAAACAGGAAUAGAAAUUCA	23	downstream
CEP290-1763	+	UUGAAACAGGAAUAGAAAUUCA	22	downstream
CEP290-1764	+	UGAAACAGGAAUAGAAAUUCA	21	downstream
CEP290-1765	-	UUUUCUUAAGCAUACUUUUUUUAA	24	downstream
CEP290-1766	-	UUUCUUAAGCAUACUUUUUUUAA	23	downstream
CEP290-1767	-	UUCUUAAGCAUACUUUUUUUAA	22	downstream
CEP290-1768	-	UCUUAAGCAUACUUUUUUUAA	21	downstream
CEP290-1769	-	UUAAGCAUACUUUUUUUAA	19	downstream
CEP290-1770	-	UAAGCAUACUUUUUUUAA	18	downstream
CEP290-1771	+	UCACUCCACUGCACUCCAGC	20	downstream
CEP290-1772	+	AGUUUUUAAGGCGGGGAGUCACA	23	downstream
CEP290-1773	-	AAACUGUCAAAAGCUACCGGUUAC	24	downstream
CEP290-1774	-	AACUGUCAAAAGCUACCGGUUAC	23	downstream
CEP290-252	-	ACUGUCAAAAGCUACCGGUUAC	22	downstream
CEP290-1775	+	AGUUCAUCUCUUGCUCUAGAUGAC	24	downstream
CEP290-1776	+	AUCUCUUGCUCUAGAUGAC	19	downstream
CEP290-1777	-	ACGAAAAUCAGAUUUCAUGU	20	downstream
CEP290-1778	-	AAUACAUGAGAGUGAUUAGUG	21	downstream
CEP290-1779	-	AUACAUGAGAGUGAUUAGUG	20	downstream
CEP290-1780	-	ACAUGAGAGUGAUUAGUG	18	downstream
CEP290-1781	+	AUUAGCUUGAACUCUGUGCCAAA	23	downstream
CEP290-1782	+	AGCUUGAACUCUGUGCCAAA	20	downstream
CEP290-1783	-	AUGUAGAUUGAGGUAGAAUCAAG	23	downstream
CEP290-1784	-	AGAUUGAGGUAGAAUCAAG	19	downstream
CEP290-1785	+	AUAAGAUGCAGAACUAGUGUAGA	23	downstream
CEP290-1786	+	AAGAUGCAGAACUAGUGUAGA	21	downstream
CEP290-1787	+	AGAUGCAGAACUAGUGUAGA	20	downstream
CEP290-1788	+	AUGCAGAACUAGUGUAGA	18	downstream
CEP290-1789	-	AUAGAUGUAGAUUGAGGUAGAAUC	24	downstream
CEP290-1790	-	AGAUGUAGAUUGAGGUAGAAUC	22	downstream
CEP290-1791	-	AUGUAGAUUGAGGUAGAAUC	20	downstream
CEP290-1792	+	AGAAUGAUCAUUCUUGUGGCAGUA	24	downstream
CEP290-1793	+	AAUGAUCAUUCUUGUGGCAGUA	22	downstream
CEP290-1794	+	AUGAUCAUUCUUGUGGCAGUA	21	downstream
CEP290-1795	+	AUCAUUCUUGUGGCAGUA	18	downstream
CEP290-1796	+	AGAAUGAUCAUUCUUGUGGCAGU	23	downstream
CEP290-1797	+	AAUGAUCAUUCUUGUGGCAGU	21	downstream

CEP290-1798	+	AUGAUCAUUCUUGUGGCAGU	20	downstream
CEP290-1799	-	AGAGGUAAAGGUUCAUGAGAC	21	downstream
CEP290-1800	-	AGGUAAAGGUUCAUGAGAC	19	downstream
CEP290-1801	+	AGCUUUUGACAGUUUUUAAG	20	downstream
CEP290-1802	+	AGCUUUUGACAGUUUUUAAGGC	22	downstream
CEP290-1803	+	AGAAAUUCACUGAGCAAAACAAC	23	downstream
CEP290-1804	+	AAAUUCACUGAGCAAAACAAC	21	downstream
CEP290-1805	+	AAUUCACUGAGCAAAACAAC	20	downstream
CEP290-1806	+	AUUCACUGAGCAAAACAAC	19	downstream
CEP290-1807	+	AGUAAGGAGGAUGUAAGA	18	downstream
CEP290-1808	+	AUCAAAAGACUUAUAUUCCAUUA	23	downstream
CEP290-1809	+	AAAAGACUUAUAUUCCAUUA	20	downstream
CEP290-1810	+	AAAGACUUAUAUUCCAUUA	19	downstream
CEP290-1811	+	AAGACUUAUAUUCCAUUA	18	downstream
CEP290-1812	-	AGGAAAUUAUUGUUGCUUUUU	21	downstream
CEP290-1813	-	AAAUUAUUGUUGCUUUUU	18	downstream
CEP290-1814	-	AAAGAAAAACUUGAAAUUUGAUAG	24	downstream
CEP290-1815	-	AAGAAAAACUUGAAAUUUGAUAG	23	downstream
CEP290-1816	-	AGAAAAACUUGAAAUUUGAUAG	22	downstream
CEP290-1817	-	AAAAACUUGAAAUUUGAUAG	20	downstream
CEP290-1818	-	AAAACUUGAAAUUUGAUAG	19	downstream
CEP290-1819	-	AAACUUGAAAUUUGAUAG	18	downstream
CEP290-1820	-	AAGAAAAAAGAAAUAGAUGUAGA	23	downstream
CEP290-1821	-	AGAAAAAAGAAAUAGAUGUAGA	22	downstream
CEP290-1822	-	AAAAAAGAAAUAGAUGUAGA	20	downstream
CEP290-1823	-	AAAAAGAAAUAGAUGUAGA	19	downstream
CEP290-1824	-	AAAAGAAAUAGAUGUAGA	18	downstream
CEP290-1825	-	AGAGUCUCACUGUGUUGCCCAGG	23	downstream
CEP290-1826	-	AGUCUCACUGUGUUGCCCAGG	21	downstream
CEP290-1827	+	CAGUUUUUAAGGCGGGGAGUCACA	24	downstream
CEP290-1828	-	CUGUCAAAAGCUACCGGUUAC	21	downstream
CEP290-1829	+	CAUCUCUUGCUCUAGAUGAC	20	downstream
CEP290-1830	-	CACGAAAAUCAGAUUUCAUGU	21	downstream
CEP290-1831	-	CGAAAAUCAGAUUUCAUGU	19	downstream
CEP290-1832	-	CUAAUACAUGAGAGUGAUUAGUG	23	downstream
CEP290-1833	+	CUUGAACUCUGUGCCAAA	18	downstream
CEP290-1834	+	CUCUAGAUGACAUGAGGUAAG	21	downstream
CEP290-1835	+	CUAGAUGACAUGAGGUAAG	19	downstream
CEP290-1836	+	CGGUAGCUUUUGACAGUUUUUAAG	24	downstream
CEP290-1837	+	CUUUUGACAGUUUUUAAG	18	downstream
CEP290-1838	+	CUUUUGACAGUUUUUAAGGC	20	downstream
CEP290-1839	+	CAGUAAGGAGGAUGUAAGA	19	downstream

CEP290-1840	+	CAAAAGACUUAUAUUCCAUUA	21	downstream
CEP290-1841	-	CUUAGGAAAUUAUUGUUGCUUUUU	24	downstream
CEP290-1842	-	CUGUGUUGCCCAGGCUGGAGUGCA	24	downstream
CEP290-1843	-	CAGAGUCUCACUGUGUUGCCCAGG	24	downstream
CEP290-1844	-	CUCACUGUGUUGCCCAGG	18	downstream
CEP290-1845	+	GUUUUUAAGGCGGGGAGUCACA	22	downstream
CEP290-1846	-	GUCAAAAGCUACCGGUUAC	19	downstream
CEP290-1847	+	GUUCAUCUCUUGCUCUAGAUGAC	23	downstream
CEP290-1848	-	GGUCACGAAAAUCAGAUUUCAUGU	24	downstream
CEP290-1849	-	GUCACGAAAAUCAGAUUUCAUGU	23	downstream
CEP290-1850	-	GAAAAUCAGAUUUCAUGU	18	downstream
CEP290-1851	-	GCUAAUACAUGAGAGUGAUUAGUG	24	downstream
CEP290-1852	+	GCUUGAACUCUGUGCCAAA	19	downstream
CEP290-1853	+	GCUCUAGAUGACAUGAGGUAAG	22	downstream
CEP290-1854	-	GAUGUAGAUUGAGGUAGAAUCAAG	24	downstream
CEP290-1855	-	GUAGAUUGAGGUAGAAUCAAG	21	downstream
CEP290-1856	-	GAUUGAGGUAGAAUCAAG	18	downstream
CEP290-1857	+	GAUGCAGAACUAGUGUAGA	19	downstream
CEP290-1858	-	GAUGUAGAUUGAGGUAGAAUC	21	downstream
CEP290-1859	-	GUAGAUUGAGGUAGAAUC	18	downstream
CEP290-1860	+	GAAUGAUCAUUCUUGUGGCAGUA	23	downstream
CEP290-1861	+	GAUCAUUCUUGUGGCAGUA	19	downstream
CEP290-1862	+	GAAUGAUCAUUCUUGUGGCAGU	22	downstream
CEP290-1863	+	GAUCAUUCUUGUGGCAGU	18	downstream
CEP290-1864	-	GAGAGGUAAAGGUUCAUGAGAC	22	downstream
CEP290-1865	-	GAGGUAAAGGUUCAUGAGAC	20	downstream
CEP290-1866	-	GGUAAAGGUUCAUGAGAC	18	downstream
CEP290-1867	+	GGUAGCUUUUGACAGUUUUUAAG	23	downstream
CEP290-1868	+	GUAGCUUUUGACAGUUUUUAAG	22	downstream
CEP290-1869	+	GCUUUUGACAGUUUUUAAG	19	downstream
CEP290-1870	+	GUAGCUUUUGACAGUUUUUAAGGC	24	downstream
CEP290-1871	+	GCUUUUGACAGUUUUUAAGGC	21	downstream
CEP290-1872	+	GAAAUUCACUGAGCAAAACAAC	22	downstream
CEP290-1873	+	GUGGCAGUAAGGAGGAUGUAAGA	23	downstream
CEP290-1874	+	GGCAGUAAGGAGGAUGUAAGA	21	downstream
CEP290-1875	+	GCAGUAAGGAGGAUGUAAGA	20	downstream
CEP290-1876	-	GGAAAUUAUUGUUGCUUUUU	20	downstream
CEP290-1877	-	GAAAUUAUUGUUGCUUUUU	19	downstream
CEP290-1878	-	GAAAAACUUGAAAUUUGAUAG	21	downstream
CEP290-1879	-	GAAGAAAAAAGAAAUAGAUGUAGA	24	downstream
CEP290-1880	-	GAAAAAAGAAAUAGAUGUAGA	21	downstream
CEP290-1881	-	GUGUUGCCCAGGCUGGAGUGCA	22	downstream

CEP290-1882	-	GUUGCCCAGGCUGGAGUGCA	20	downstream
CEP290-1883	-	GAGUCUCACUGUGUUGCCCAGG	22	downstream
CEP290-1884	-	GUCUCACUGUGUUGCCCAGG	20	downstream
CEP290-1885	+	UUUUUAAGGCGGGGAGUCACA	21	downstream
CEP290-1886	+	UUUUAAGGCGGGGAGUCACA	20	downstream
CEP290-1887	+	UUUAAGGCGGGGAGUCACA	19	downstream
CEP290-1888	+	UUAAGGCGGGGAGUCACA	18	downstream
CEP290-1889	-	UGUCAAAAGCUACCGGUUAC	20	downstream
CEP290-1890	-	UCAAAAGCUACCGGUUAC	18	downstream
CEP290-1891	+	UUCAUCUCUUGCUCUAGAUGAC	22	downstream
CEP290-1892	+	UCAUCUCUUGCUCUAGAUGAC	21	downstream
CEP290-1893	+	UCUCUUGCUCUAGAUGAC	18	downstream
CEP290-1894	-	UCACGAAAAUCAGAUUUCAUGU	22	downstream
CEP290-1895	-	UAAUACAUGAGAGUGAUUAGUG	22	downstream
CEP290-1896	-	UACAUGAGAGUGAUUAGUG	19	downstream
CEP290-1897	+	UAUUAGCUUGAACUCUGUGCCAAA	24	downstream
CEP290-1898	+	UUAGCUUGAACUCUGUGCCAAA	22	downstream
CEP290-1899	+	UAGCUUGAACUCUGUGCCAAA	21	downstream
CEP290-1900	+	UUGCUCUAGAUGACAUGAGGUAAG	24	downstream
CEP290-1901	+	UGCUCUAGAUGACAUGAGGUAAG	23	downstream
CEP290-1902	+	UCUAGAUGACAUGAGGUAAG	20	downstream
CEP290-1903	+	UAGAUGACAUGAGGUAAG	18	downstream
CEP290-1904	-	UGUAGAUUGAGGUAGAAUCAAG	22	downstream
CEP290-1905	-	UAGAUUGAGGUAGAAUCAAG	20	downstream
CEP290-1906	+	UAUAAGAUGCAGAACUAGUGUAGA	24	downstream
CEP290-1907	+	UAAGAUGCAGAACUAGUGUAGA	22	downstream
CEP290-1908	-	UAGAUGUAGAUUGAGGUAGAAUC	23	downstream
CEP290-1909	-	UGUAGAUUGAGGUAGAAUC	19	downstream
CEP290-1910	+	UGAUCAUUCUUGUGGCAGUA	20	downstream
CEP290-1911	+	UAGAAUGAUCAUUCUUGUGGCAGU	24	downstream
CEP290-1912	+	UGAUCAUUCUUGUGGCAGU	19	downstream
CEP290-1913	-	UUGAGAGGUAAAGGUUCAUGAGAC	24	downstream
CEP290-1914	-	UGAGAGGUAAAGGUUCAUGAGAC	23	downstream
CEP290-1915	+	UAGCUUUUGACAGUUUUUAAG	21	downstream
CEP290-1916	+	UAGCUUUUGACAGUUUUUAAGGC	23	downstream
CEP290-1917	+	UUUUGACAGUUUUUAAGGC	19	downstream
CEP290-1918	+	UUUGACAGUUUUUAAGGC	18	downstream
CEP290-1919	+	UAGAAAUUCACUGAGCAAAACAAC	24	downstream
CEP290-1920	+	UUCACUGAGCAAAACAAC	18	downstream
CEP290-1921	+	UGUGGCAGUAAGGAGGAUGUAAGA	24	downstream
CEP290-1922	+	UGGCAGUAAGGAGGAUGUAAGA	22	downstream
CEP290-1923	+	UAUCAAAAGACUUAUAUUCCAUUA	24	downstream

CEP290-1924	+	UCAAAAGACUUAUAUUCCAUUA	22	downstream
CEP290-1925	-	UUAGGAAAUUAUUGUUGCUUUUU	23	downstream
CEP290-1926	-	UAGGAAAUUAUUGUUGCUUUUU	22	downstream
CEP290-1927	-	UGUGUUGCCCAGGCUGGAGUGCA	23	downstream
CEP290-1928	-	UGUUGCCCAGGCUGGAGUGCA	21	downstream
CEP290-1929	-	UUGCCCAGGCUGGAGUGCA	19	downstream
CEP290-1930	-	UGCCCAGGCUGGAGUGCA	18	downstream
CEP290-1931	-	UCUCACUGUGUUGCCCAGG	19	downstream
CEP290-13	+	AUGAGAUACUCACAAUUACAAC	22	upstream
CEP290-18	+	GUAUGAGAUACUCACAAUUACAAC	24	upstream
CEP290-14	+	UAUGAGAUACUCACAAUUACAAC	23	upstream
CEP290-19	+	GGUAUGAGAUAUUCACAAUUACAA	24	upstream

Table 9A provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the first tier parameters. The targeting domains are within 1000 bp upstream of an *Alu* repeat, within 40bp upstream of mutation, or 1000 bp downstream of the mutation, have good orthogonality, and start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting domains in the table can be used with a *N. meningitidis* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 9A

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-1932	+	GGCAAAAGCAGCAGAAAGCA	20	upstream
CEP290-1933	-	GUGGCUGAAUGACUUCU	17	upstream
CEP290-1934	-	GUUGUUCUGAGUAGCUU	17	upstream
CEP290-1935	-	GACUAGAGGUCACGAAA	17	downstream
CEP290-1936	-	GAGUUCAAGCUAAUACAUGA	20	downstream

Table 9B provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene selected according to the second tier parameters. The targeting domains are within 1000 bp upstream of an *Alu* repeat, within 40bp upstream of mutation, or 1000 bp downstream of the mutation, have good orthogonality, and do not start with G. It is contemplated herein that the targeting domain hybridizes to the target domain through complementary base pairing. Any of the targeting

domains in the table can be used with a *N. meningitidis* Cas9 molecule that generates a double stranded break (Cas9 nuclease) or a single-stranded break (Cas9 nickase).

Table 9B

gRNA Name	DNA Strand	Targeting Domain	Target Site Length	Position relative to mutation
CEP290-1937	+	AAAAGCAGCAGAAAGCA	17	upstream
CEP290-1938	-	AACGUUGUUCUGAGUAGCUU	20	upstream
CEP290-1939	-	AAUAGAGGCUUAUGGAU	17	upstream
CEP290-1940	+	ACUUAAUGAGUGCUUCCCUC	20	upstream
CEP290-1941	-	AGAAAUAGAGGCUUAUGGAU	20	upstream
CEP290-1942	+	AGCAGAAAGCAAACUGA	17	upstream
CEP290-1943	+	AGCAGCAGAAAGCAAACUGA	20	upstream
CEP290-1944	+	AGGGUCUGGUCCAUAUU	17	upstream
CEP290-1945	-	AUAGUGGCUGAAUGACUUCU	20	upstream
CEP290-1946	+	AUGUCUGGUUAAAAGAG	17	upstream
CEP290-1947	+	CAAAGGGUCUGGUCCAUAUU	20	upstream
CEP290-1948	-	CAUCAGAAAUAGAGGCU	17	upstream
CEP290-1949	-	CCUCAUCAGAAAUAGAGGCU	20	upstream
CEP290-1950	-	CUGAGGACAGAACAAGC	17	upstream
CEP290-1951	-	CUGCUGCUUUUGCCAAAGAG	20	upstream
CEP290-1952	-	CUGCUUUUGCCAAAGAG	17	upstream
CEP290-1953	+	UAAUGAGUGCUUCCCUC	17	upstream
CEP290-1954	+	UAGAUGUCUGGUUAAAAGAG	20	upstream
CEP290-1955	-	UCAUUCUCCUUAGGUCACUU	20	upstream
CEP290-1956	-	UUACUGAGGACAGAACAAGC	20	upstream
CEP290-1957	-	UUCUCCUUAGGUCACUU	17	upstream
CEP290-1958	-	AAGAAAAAAGAAAUAGA	17	downstream
CEP290-1959	-	AGAUUGAGGUAGAAUCAAGA	20	downstream
CEP290-1960	+	AGUCACAUGGGAGUCACAGG	20	downstream
CEP290-1961	+	CAAAAAAAGAAUCCUCU	17	downstream
CEP290-1962	+	CAACAAAAAAAGAAUCCUCU	20	downstream
CEP290-1963	+	CACAUGGGAGUCACAGG	17	downstream
CEP290-1964	+	CAUUCUUCACACAUGAA	17	downstream
CEP290-1965	-	UAGAAGAAAAAAGAAAUAGA	20	downstream
CEP290-1966	-	UGAGACUAGAGGUCACGAAA	20	downstream
CEP290-1967	-	UUCAAGCUAAUACAUGA	17	downstream
CEP290-1968	+	UUCCAUUCUUCACACAUGAA	20	downstream
CEP290-1969	-	UUGAGGUAGAAUCAAGA	17	downstream

Table 10 provides targeting domains for break-induced deletion of genomic sequence including the mutation at the LCA10 target position in the *CEP290* gene by dual targeting (e.g., dual double strand cleavage). Exemplary gRNA pairs to be used with *S. aureus* Cas9 are shown in **Table 10**, e.g., CEP290-323 can be combined with CEP290-11, CEP290-323 can be combined with CEP290-64, CEP290-490 can be combined with CEP290-496, CEP290-490 can be combined with CEP290-502, CEP290-490 can be combined with CEP290-504, CEP290-492 can be combined with CEP290-502, or CEP290-492 can be combined with CEP290-504.

Table 10

Upstream gRNA		Downstream gRNA	
CEP290-323	GTTCTGTCCTCAGTAAAAGGTA	CEP290-11	GACACTGCCAATAGGGATAGG
CEP290-323	GTTCTGTCCTCAGTAAAAGGTA	CEP290-64	GTCAAAAGCTACCGGTTACCTG
CEP290-490	GAATAGTTTGTTCTGGGTAC	CEP290-496	GATGCAGAACTAGTGTAGAC
CEP290-490	GAATAGTTTGTTCTGGGTAC	CEP290-502	GTCACATGGGAGTCACAGGG
CEP290-490	GAATAGTTTGTTCTGGGTAC	CEP290-504	GAGTATCTCCTGTTTGGCA
CEP290-492	GAGAAAGGGATGGGCACTTA	CEP290-502	GTCACATGGGAGTCACAGGG
CEP290-492	GAGAAAGGGATGGGCACTTA	CEP290-504	GAGTATCTCCTGTTTGGCA

III. Cas9 Molecules

Cas9 molecules of a variety of species can be used in the methods and compositions described herein. While the *S. pyogenes*, *S. aureus*, and *S. thermophilus* Cas9 molecules are the subject of much of the disclosure herein, Cas9 molecules of, derived from, or based on the Cas9 proteins of other species listed herein can be used as well. In other words, while the much of the description herein uses *S. pyogenes* and *S. thermophilus* Cas9 molecules Cas9 molecules from the other species can replace them. Such species include: *Acidovorax avenae, Actinobacillus pleuropneumoniae, Actinobacillus succinogenes, Actinobacillus suis, Actinomyces sp., Cycliphilusdenitrificans, Aminomonas paucivorans, Bacillus cereus, Bacillus smithii, Bacillus thuringiensis, Bacteroides sp., Blastopirellula marina, Bradyrhizobium sp., Brevibacillus laterosporus, Campylobacter coli, Campylobacter jejuni, Campylobacter lari, Candidatus puniceispirillum, Clostridium cellulolyticum, Clostridium perfringens, Corynebacterium accolens, Corynebacterium diphtheria, Corynebacterium matruchotii, Dinoroseobacter shibae, Eubacterium dolichum, Gammaproteobacterium, Gluconacetobacter diazotrophicus, Haemophilus parainfluenzae, Haemophilus sputorum, Helicobacter canadensis, Helicobacter cinaedi, Helicobacter mustelae, Ilyobacter polytropus, Kingella kingae, Lactobacillus crispatus,*

Listeria ivanovii, Listeria monocytogenes, Listeriaceae bacterium, Methylocystis sp., Methylosinus trichosporium, Mobiluncus mulieris, Neisseria bacilliformis, Neisseria cinerea, Neisseria flavescens, Neisseria lactamica, Neisseria meningitidis, Neisseria sp., Neisseria wadsworthii, Nitrosomonas sp., Parvibaculum lavamentivorans, Pasteurella multocida, Phascolarctobacterium succinatutens, Ralstonia syzygii, Rhodopseudomonas palustris, Rhodovulum sp., Simonsiella muelleri, Sphingomonas sp., Sporolactobacillus vineae, Staphylococcus aureus, Staphylococcus lugdunensis, Streptococcus sp., Subdoligranulum sp., Tistrella mobilis, Treponema sp., or *Verminephrobacter eiseniae.*

A Cas9 molecule, or Cas9 polypeptide, as that term is used herein, refers to a molecule or polypeptide that can interact with a guide RNA (gRNA) molecule and, in concert with the gRNA molecule, homes or localizes to a site which comprises a target domain and PAM sequence. Cas9 molecule and Cas9 polypeptide, as those terms are used herein, refer to naturally occurring Cas9 molecules and to engineered, altered, or modified Cas9 molecules or Cas9 polypeptides that differ, e.g., by at least one amino acid residue, from a reference sequence, e.g., the most similar naturally occurring Cas9 molecule or a sequence of **Table 11**.

Cas9 Domains

Crystal structures have been determined for two different naturally occurring bacterial Cas9 molecules (Jinek et al., Science, 343(6176):1247997, 2014) and for *S. pyogenes* Cas9 with a guide RNA (e.g., a synthetic fusion of crRNA and tracrRNA) (Nishimasu et al., Cell, 156:935-949, 2014; and Anders et al., Nature, 2014, doi: 10.1038/nature13579).

A naturally occurring Cas9 molecule comprises two lobes: a recognition (REC) lobe and a nuclease (NUC) lobe; each of which further comprises domains described herein. **Figs. 8A-8B** provide a schematic of the organization of important Cas9 domains in the primary structure. The domain nomenclature and the numbering of the amino acid residues encompassed by each domain used throughout this disclosure is as described in Nishimasu et al. The numbering of the amino acid residues is with reference to Cas9 from *S. pyogenes*.

The REC lobe comprises the arginine-rich bridge helix (BH), the REC1 domain, and the REC2 domain. The REC lobe does not share structural similarity with other known proteins, indicating that it is a Cas9-specific functional domain. The BH domain is a long α helix and arginine rich region and comprises amino acids 60-93 of the sequence of *S. pyogenes* Cas9. The

REC1 domain is important for recognition of the repeat:anti-repeat duplex, e.g., of a gRNA or a tracrRNA, and is therefore critical for Cas9 activity by recognizing the target sequence. The REC1 domain comprises two REC1 motifs at amino acids 94 to 179 and 308 to 717 of the sequence of *S. pyogenes* Cas9. These two REC1 domains, though separated by the REC2 domain in the linear primary structure, assemble in the tertiary structure to form the REC1 domain. The REC2 domain, or parts thereof, may also play a role in the recognition of the repeat:anti-repeat duplex. The REC2 domain comprises amino acids 180-307 of the sequence of *S. pyogenes* Cas9.

The NUC lobe comprises the RuvC domain (also referred to herein as RuvC-like domain), the HNH domain (also referred to herein as HNH-like domain), and the PAM-interacting (PI) domain. The RuvC domain shares structural similarity to retroviral integrase superfamily members and cleaves a single strand, e.g., the non-complementary strand of the target nucleic acid molecule. The RuvC domain is assembled from the three split RuvC motifs (RuvC I, RuvCII, and RuvCIII, which are often commonly referred to in the art as RuvCI domain, or N-terminal RuvC domain, RuvCII domain, and RuvCIII domain) at amino acids 1-59, 718-769, and 909-1098, respectively, of the sequence of *S. pyogenes* Cas9. Similar to the REC1 domain, the three RuvC motifs are linearly separated by other domains in the primary structure, however in the tertiary structure, the three RuvC motifs assemble and form the RuvC domain. The HNH domain shares structural similarity with HNH endonucleases, and cleaves a single strand, e.g., the complementary strand of the target nucleic acid molecule. The HNH domain lies between the RuvC II-III motifs and comprises amino acids 775-908 of the sequence of *S. pyogenes* Cas9. The PI domain interacts with the PAM of the target nucleic acid molecule, and comprises amino acids 1099-1368 of the sequence of *S. pyogenes* Cas9.

A RuvC-like domain and an HNH-like domain

In an embodiment, a Cas9 molecule or Cas9 polypeptide comprises an HNH-like domain and a RuvC-like domain. In an embodiment, cleavage activity is dependent on a RuvC-like domain and an HNH-like domain. A Cas9 molecule or Cas9 polypeptide, e.g., an eaCas9 molecule or eaCas9 polypeptide, can comprise one or more of the following domains: a RuvC-like domain and an HNH-like domain. In an embodiment, a Cas9 molecule or Cas9 polypeptide is an eaCas9 molecule or eaCas9 polypeptide and the eaCas9 molecule or eaCas9 polypeptide

comprises a RuvC-like domain, e.g., a RuvC-like domain described below, and/or an HNH-like domain, e.g., an HNH-like domain described below.

RuvC-like domains

In an embodiment, a RuvC-like domain cleaves, a single strand, e.g., the non-complementary strand of the target nucleic acid molecule. The Cas9 molecule or Cas9 polypeptide can include more than one RuvC-like domain (e.g., one, two, three or more RuvC-like domains). In an embodiment, a RuvC-like domain is at least 5, 6, 7, 8 amino acids in length but not more than 20, 19, 18, 17, 16 or 15 amino acids in length. In an embodiment, the Cas9 molecule or Cas9 polypeptide comprises an N-terminal RuvC-like domain of about 10 to 20 amino acids, e.g., about 15 amino acids in length.

N-terminal RuvC-like domains

Some naturally occurring Cas9 molecules comprise more than one RuvC-like domain with cleavage being dependent on the N-terminal RuvC-like domain. Accordingly, Cas9 molecules or Cas9 polypeptide can comprise an N-terminal RuvC-like domain. Exemplary N-terminal RuvC-like domains are described below.

In an embodiment, an eaCas9 molecule or eaCas9 polypeptide comprises an N-terminal RuvC-like domain comprising an amino acid sequence of formula I:

D-X1-G-X2-X3-X4-X5-G-X6-X7-X8-X9 (SEQ ID NO: 8),

wherein,

X1 is selected from I, V, M, L and T (e.g., selected from I, V, and L);

X2 is selected from T, I, V, S, N, Y, E and L (e.g., selected from T, V, and I);

X3 is selected from N, S, G, A, D, T, R, M and F (e.g., A or N);

X4 is selected from S, Y, N and F (e.g., S);

X5 is selected from V, I, L, C, T and F (e.g., selected from V, I and L);

X6 is selected from W, F, V, Y, S and L (e.g., W);

X7 is selected from A, S, C, V and G (e.g., selected from A and S);

X8 is selected from V, I, L, A, M and H (e.g., selected from V, I, M and L); and

X9 is selected from any amino acid or is absent, designated by Δ (e.g., selected from T, V, I, L, Δ, F, S, A, Y, M and R, or, e.g., selected from T, V, I, L and Δ).

In an embodiment, the N-terminal RuvC-like domain differs from a sequence of SEQ ID NO:8, by as many as 1 but no more than 2, 3, 4, or 5 residues.

In embodiment, the N-terminal RuvC-like domain is cleavage competent.

In embodiment, the N-terminal RuvC-like domain is cleavage incompetent.

In an embodiment, a eaCas9 molecule or eaCas9 polypeptide comprises an N-terminal RuvC-like domain comprising an amino acid sequence of formula II:

D-X1-G-X2-X3-S-X5-G-X6-X7-X8-X9, (SEQ ID NO: 9),

wherein

X1 is selected from I, V, M, L and T (e.g., selected from I, V, and L);

X2 is selected from T, I, V, S, N, Y, E and L (e.g., selected from T, V, and I);

X3 is selected from N, S, G, A, D, T, R, M and F (e.g., A or N);

X5 is selected from V, I, L, C, T and F (e.g., selected from V, I and L);

X6 is selected from W, F, V, Y, S and L (e.g., W);

X7 is selected from A, S, C, V and G (e.g., selected from A and S);

X8 is selected from V, I, L, A, M and H (e.g., selected from V, I, M and L); and

X9 is selected from any amino acid or is absent (e.g., selected from T, V, I, L, Δ, F, S, A, Y, M and R or selected from e.g., T, V, I, L and Δ).

In an embodiment, the N-terminal RuvC-like domain differs from a sequence of SEQ ID NO:9 by as many as 1 but no more than 2, 3, 4, or 5 residues.

In an embodiment, the N-terminal RuvC-like domain comprises an amino acid sequence of formula III:

D-I-G-X2-X3-S-V-G-W-A-X8-X9 (SEQ ID NO: 10),

wherein

X2 is selected from T, I, V, S, N, Y, E and L (e.g., selected from T, V, and I);

X3 is selected from N, S, G, A, D, T, R, M and F (e.g., A or N);

X8 is selected from V, I, L, A, M and H (e.g., selected from V, I, M and L); and

X9 is selected from any amino acid or is absent (e.g., selected from T, V, I, L, Δ, F, S, A, Y, M and R or selected from e.g., T, V, I, L and Δ).

In an embodiment, the N-terminal RuvC-like domain differs from a sequence of SEQ ID NO:10 by as many as 1 but no more than, 2, 3, 4, or 5 residues.

In an embodiment, the N-terminal RuvC-like domain comprises an amino acid sequence of formula III:

D-I-G-T-N-S-V-G-W-A-V-X (SEQ ID NO: 11),

wherein

X is a non-polar alkyl amino acid or a hydroxyl amino acid, e.g., X is selected from V, I, L and T (e.g., the eaCas9 molecule can comprise an N-terminal RuvC-like domain shown in **Figs. 2A-2G** (is depicted as Y)).

In an embodiment, the N-terminal RuvC-like domain differs from a sequence of SEQ ID NO:11 by as many as 1 but no more than, 2, 3, 4, or 5 residues.

In an embodiment, the N-terminal RuvC-like domain differs from a sequence of an N-terminal RuvC like domain disclosed herein, e.g., in **Figs. 3A-3B** or **Figs. 7A-7B**, as many as 1 but no more than 2, 3, 4, or 5 residues. In an embodiment, 1, 2, or all 3 of the highly conserved residues identified in **Figs. 3A-3B** or **Figs. 7A-7B** are present.

In an embodiment, the N-terminal RuvC-like domain differs from a sequence of an N-terminal RuvC-like domain disclosed herein, e.g., in **Figs. 4A-4B** or **Figs. 7A-7B**, as many as 1 but no more than 2, 3, 4, or 5 residues. In an embodiment, 1, 2, 3 or all 4 of the highly conserved residues identified in **Figs. 4A-4B** or **Figs. 7A-7B** are present.

Additional RuvC-like domains

In addition to the N-terminal RuvC-like domain, the Cas9 molecule or Cas9 polypeptide, e.g., an eaCas9 molecule or eaCas9 polypeptide, can comprise one or more additional RuvC-like domains. In an embodiment, the Cas9 molecule or Cas9 polypeptide can comprise two additional RuvC-like domains. Preferably, the additional RuvC-like domain is at least 5 amino acids in length and, e.g., less than 15 amino acids in length, e.g., 5 to 10 amino acids in length, e.g., 8 amino acids in length.

An additional RuvC-like domain can comprise an amino acid sequence:

I-X1-X2-E-X3-A-R-E (SEQ ID NO:12), wherein

X1 is V or H,

X2 is I, L or V (e.g., I or V); and

X3 is M or T.

In an embodiment, the additional RuvC-like domain comprises the amino acid sequence:

I-V-X2-E-M-A-R-E (SEQ ID NO:13), wherein

X2 is I, L or V (e.g., I or V) (e.g., the eaCas9 molecule or eaCas9 polypeptide can comprise an additional RuvC-like domain shown in **Fig. 2A-2G** or **Figs. 7A-7B** (depicted as B)).

An additional RuvC-like domain can comprise an amino acid sequence:

H-H-A-X1-D-A-X2-X3 (SEQ ID NO:14), wherein

X1 is H or L;

X2 is R or V; and

X3 is E or V.

In an embodiment, the additional RuvC-like domain comprises the amino acid sequence: H-H-A-H-D-A-Y-L (SEQ ID NO:15).

In an embodiment, the additional RuvC-like domain differs from a sequence of SEQ ID NO:13, 15, 12 or 14 by as many as 1 but no more than 2, 3, 4, or 5 residues.

In some embodiments, the sequence flanking the N-terminal RuvC-like domain is a sequences of formula V:

K-X1'-Y-X2'-X3'-X4'-Z-T-D-X9'-Y, (SEQ ID NO:16).

wherein

X1' is selected from K and P,

X2' is selected from V, L, I, and F (e.g., V, I and L);

X3' is selected from G, A and S (e.g., G),

X4' is selected from L, I, V and F (e.g., L);

X9' is selected from D, E, N and Q; and

Z is an N-terminal RuvC-like domain, e.g., as described above.

HNH-like domains

In an embodiment, an HNH-like domain cleaves a single stranded complementary domain, e.g., a complementary strand of a double stranded nucleic acid molecule. In an embodiment, an HNH-like domain is at least 15, 20, 25 amino acids in length but not more than 40, 35 or 30 amino acids in length, e.g., 20 to 35 amino acids in length, e.g., 25 to 30 amino acids in length. Exemplary HNH-like domains are described below.

In an embodiment, an eaCas9 molecule or eaCas9 polypeptide comprises an HNH-like domain having an amino acid sequence of formula VI:

X1-X2-X3-H-X4-X5-P-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-N-X16-X17-X18-X19-X20-X21-X22-X23-N (SEQ ID NO:17), wherein

X1 is selected from D, E, Q and N (e.g., D and E);

X2 is selected from L, I, R, Q, V, M and K;

X3 is selected from D and E;

X4 is selected from I, V, T, A and L (e.g., A, I and V);

X5 is selected from V, Y, I, L, F and W (e.g., V, I and L);

X6 is selected from Q, H, R, K, Y, I, L, F and W;

X7 is selected from S, A, D, T and K (e.g., S and A);

X8 is selected from F, L, V, K, Y, M, I, R, A, E, D and Q (e.g., F);

X9 is selected from L, R, T, I, V, S, C, Y, K, F and G;

X10 is selected from K, Q, Y, T, F, L, W, M, A, E, G, and S;

X11 is selected from D, S, N, R, L and T (e.g., D);

X12 is selected from D, N and S;

X13 is selected from S, A, T, G and R (e.g., S);

X14 is selected from I, L, F, S, R, Y, Q, W, D, K and H (e.g., I, L and F);

X15 is selected from D, S, I, N, E, A, H, F, L, Q, M, G, Y and V;

X16 is selected from K, L, R, M, T and F (e.g., L, R and K);

X17 is selected from V, L, I, A and T;

X18 is selected from L, I, V and A (e.g., L and I);

X19 is selected from T, V, C, E, S and A (e.g., T and V);

X20 is selected from R, F, T, W, E, L, N, C, K, V, S, Q, I, Y, H and A;

X21 is selected from S, P, R, K, N, A, H, Q, G and L;

X22 is selected from D, G, T, N, S, K, A, I, E, L, Q, R and Y; and

X23 is selected from K, V, A, E, Y, I, C, L, S, T, G, K, M, D and F.

In an embodiment, a HNH-like domain differs from a sequence of SEQ ID NO:16 by at least one but no more than, 2, 3, 4, or 5 residues.

In an embodiment, the HNH-like domain is cleavage competent.

In an embodiment, the HNH-like domain is cleavage incompetent.

In an embodiment, an eaCas9 molecule or eaCas9 polypeptide comprises an HNH-like domain comprising an amino acid sequence of formula VII:

X1-X2-X3-H-X4-X5-P-X6-S-X8-X9-X10-D-D-S-X14-X15-N-K-V-L-X19-X20-X21-X22-X23-N (SEQ ID NO:18),

wherein

X1 is selected from D and E;

X2 is selected from L, I, R, Q, V, M and K;

X3 is selected from D and E;

X4 is selected from I, V, T, A and L (e.g., A, I and V);

X5 is selected from V, Y, I, L, F and W (e.g., V, I and L);

X6 is selected from Q, H, R, K, Y, I, L, F and W;

X8 is selected from F, L, V, K, Y, M, I, R, A, E, D and Q (e.g., F);

X9 is selected from L, R, T, I, V, S, C, Y, K, F and G;

X10 is selected from K, Q, Y, T, F, L, W, M, A, E, G, and S;

X14 is selected from I, L, F, S, R, Y, Q, W, D, K and H (e.g., I, L and F);

X15 is selected from D, S, I, N, E, A, H, F, L, Q, M, G, Y and V;

X19 is selected from T, V, C, E, S and A (e.g., T and V);

X20 is selected from R, F, T, W, E, L, N, C, K, V, S, Q, I, Y, H and A;

X21 is selected from S, P, R, K, N, A, H, Q, G and L;

X22 is selected from D, G, T, N, S, K, A, I, E, L, Q, R and Y; and

X23 is selected from K, V, A, E, Y, I, C, L, S, T, G, K, M, D and F.

In an embodiment, the HNH-like domain differs from a sequence of SEQ ID NO:15 by 1, 2, 3, 4, or 5 residues.

In an embodiment, an eaCas9 molecule or eaCas9 polypeptide comprises an HNH-like domain comprising an amino acid sequence of formula VII:

X1-V-X3-H-I-V-P-X6-S-X8-X9-X10-D-D-S-X14-X15-N-K-V-L-T-X20-X21-X22-X23-N (SEQ ID NO:19),

wherein

X1 is selected from D and E;

X3 is selected from D and E;

X6 is selected from Q, H, R, K, Y, I, L and W;

X8 is selected from F, L, V, K, Y, M, I, R, A, E, D and Q (e.g., F);

X9 is selected from L, R, T, I, V, S, C, Y, K, F and G;

X10 is selected from K, Q, Y, T, F, L, W, M, A, E, G, and S;

X14 is selected from I, L, F, S, R, Y, Q, W, D, K and H (e.g., I, L and F);

X15 is selected from D, S, I, N, E, A, H, F, L, Q, M, G, Y and V;

X20 is selected from R, F, T, W, E, L, N, C, K, V, S, Q, I, Y, H and A;

X21 is selected from S, P, R, K, N, A, H, Q, G and L;

X22 is selected from D, G, T, N, S, K, A, I, E, L, Q, R and Y; and

X23 is selected from K, V, A, E, Y, I, C, L, S, T, G, K, M, D and F.

In an embodiment, the HNH-like domain differs from a sequence of SEQ ID NO:GG by 1, 2, 3, 4, or 5 residues.

In an embodiment, an eaCas9 molecule or eaCas9 polypeptide comprises an HNH-like domain having an amino acid sequence of formula VIII:

D-X2-D-H-I-X5-P-Q-X7-F-X9-X10-D-X12-S-I-D-N-X16-V-L-X19-X20-S-X22-X23-N (SEQ ID NO:20),

wherein

X2 is selected from I and V;

X5 is selected from I and V;

X7 is selected from A and S;

X9 is selected from I and L;

X10 is selected from K and T;

X12 is selected from D and N;

X16 is selected from R, K and L; X19 is selected from T and V;

X20 is selected from S and R;

X22 is selected from K, D and A; and

X23 is selected from E, K, G and N (e.g., the eaCas9 molecule or eaCas9 polypeptide can comprise an HNH-like domain as described herein).

In an embodiment, the HNH-like domain differs from a sequence of SEQ ID NO:19 by as many as 1 but no more than 2, 3, 4, or 5 residues.

In an embodiment, an eaCas9 molecule or eaCas9 polypeptide comprises the amino acid sequence of formula IX:

L-Y-Y-L-Q-N-G-X1'-D-M-Y-X2'-X3'-X4'-X5'-L-D-I—X6'-X7'-L-S-X8'-Y-Z-N-R-X9'-K-X10'-D-X11'-V-P (SEQ ID NO:21),

wherein

X1' is selected from K and R;

X2' is selected from V and T;

X3' is selected from G and D;

X4' is selected from E, Q and D;

X5' is selected from E and D;

X6' is selected from D, N and H;

X7' is selected from Y, R and N;

X8' is selected from Q, D and N; X9' is selected from G and E;

X10' is selected from S and G;

X11' is selected from D and N; and

Z is an HNH-like domain, e.g., as described above.

In an embodiment, the eaCas9 molecule or eaCas9 polypeptide comprises an amino acid sequence that differs from a sequence of SEQ ID NO:21 by as many as 1 but no more than 2, 3, 4, or 5 residues.

In an embodiment, the HNH-like domain differs from a sequence of an HNH-like domain disclosed herein, e.g., in **Figs. 5A-5C** or **Figs. 7A-7B**, as many as 1 but no more than 2, 3, 4, or 5 residues. In an embodiment, 1 or both of the highly conserved residues identified in **Figs. 5A-5C** or **Figs. 7A-7B** are present.

In an embodiment, the HNH -like domain differs from a sequence of an HNH-like domain disclosed herein, e.g., in **Figs. 6A-6B** or **Figs. 7A-7B** , as many as 1 but no more than 2, 3, 4, or 5 residues. In an embodiment, 1, 2, all 3 of the highly conserved residues identified in **Figs. 6A-6B** or **Figs. 7A-7B** are present.

Cas9 Activities

Nuclease and Helicase Activities

In an embodiment, the Cas9 molecule or Cas9 polypeptide is capable of cleaving a target nucleic acid molecule. Typically wild type Cas9 molecules cleave both strands of a target nucleic acid molecule. Cas9 molecules and Cas9 polypeptides can be engineered to alter nuclease cleavage (or other properties), e.g., to provide a Cas9 molecule or Cas9 peolypeptide which is a nickase, or which lacks the ability to cleave target nucleic acid. A Cas9 molecule or Cas9 polypeptide that is capable of cleaving a target nucleic acid molecule is referred to herein as an eaCas9 molecule or eaCas9 polypeptide.

In an embodiment, an eaCas9 molecule or eaCas9 polypeptide comprises one or more of the following activities:

a nickase activity, i.e., the ability to cleave a single strand, e.g., the non-complementary strand or the complementary strand, of a nucleic acid molecule;

a double stranded nuclease activity, i.e., the ability to cleave both strands of a double stranded nucleic acid and create a double stranded break, which in an embodiment is the presence of two nickase activities;

an endonuclease activity;

an exonuclease activity; and

a helicase activity, i.e., the ability to unwind the helical structure of a double stranded nucleic acid.

In an embodiment, an enzymatically active or eaCas9 molecule or eaCas9 polypeptide cleaves both strands and results in a double stranded break. In an embodiment, an eaCas9 molecule cleaves only one strand, e.g., the strand to which the gRNA hybridizes to, or the strand complementary to the strand the gRNA hybridizes with. In an embodiment, an eaCas9 molecule or eaCas9 polypeptide comprises cleavage activity associated with an HNH-like domain. In an embodiment, an eaCas9 molecule or eaCas9 polypeptide comprises cleavage activity associated with an N-terminal RuvC-like domain. In an embodiment, an eaCas9 molecule or eaCas9 polypeptide comprises cleavage activity associated with an HNH-like domain and cleavage activity associated with an N-terminal RuvC-like domain. In an embodiment, an eaCas9 molecule or eaCas9 polypeptide comprises an active, or cleavage competent, HNH-like domain and an inactive, or cleavage incompetent, N-terminal RuvC-like domain. In an embodiment, an eaCas9 molecule or eaCas9 polypeptide comprises an inactive, or cleavage incompetent, HNH-like domain and an active, or cleavage competent, N-terminal RuvC-like domain.

Some Cas9 molecules or Cas9 polypeptides have the ability to interact with a gRNA molecule, and in conjunction with the gRNA molecule localize to a core target domain, but are incapable of cleaving the target nucleic acid, or incapable of cleaving at efficient rates. Cas9 molecules having no, or no substantial, cleavage activity are referred to herein as an eiCas9 molecule or eiCas9 polypeptide. For example, an eiCas9 molecule or eiCas9 polypeptide can lack cleavage activity or have substantially less, e.g., less than 20, 10, 5, 1 or 0.1 % of the cleavage activity of a reference Cas9 molecule or eiCas9 polypeptide, as measured by an assay described herein.

Targeting and PAMs

A Cas9 molecule or Cas9 polypeptide, is a polypeptide that can interact with a guide RNA (gRNA) molecule and, in concert with the gRNA molecule, localizes to a site which comprises a target domain and PAM sequence.

In an embodiment, the ability of an eaCas9 molecule or eaCas9 polypeptide to interact with and cleave a target nucleic acid is PAM sequence dependent. A PAM sequence is a sequence in the target nucleic acid. In an embodiment, cleavage of the target nucleic acid occurs upstream from the PAM sequence. EaCas9 molecules from different bacterial species can recognize different sequence motifs (e.g., PAM sequences). In an embodiment, an eaCas9 molecule of *S. pyogenes* recognizes the sequence motif NGG, NAG, NGA and directs cleavage of a target nucleic acid sequence 1 to 10, e.g., 3 to 5, base pairs upstream from that sequence. See, e.g., Mali *et al.*, Science 2013; 339(6121): 823-826. In an embodiment, an eaCas9 molecule of *S. thermophilus* recognizes the sequence motif NGGNG and NNAGAAW (W = A or T) and directs cleavage of a core target nucleic acid sequence 1 to 10, e.g., 3 to 5, base pairs upstream from these sequences. See, e.g., Horvath *et al.*, Science 2010; 327(5962):167-170, and Deveau *et al.*, J Bacteriol 2008; 190(4): 1390-1400. In an embodiment, an eaCas9 molecule of *S. mutans* recognizes the sequence motif NGG and/or NAAR (R = A or G) and directs cleavage of a core target nucleic acid sequence 1 to 10, e.g., 3 to 5 base pairs, upstream from this sequence. See, e.g., Deveau *et al.*, J Bacteriol 2008; 190(4): 1390-1400. In an embodiment, an eaCas9 molecule of *S. aureus* recognizes the sequence motif NNGRR (R = A or G) and directs cleavage of a target nucleic acid sequence 1 to 10, e.g., 3 to 5, base pairs upstream from that sequence. In an embodiment, an eaCas9 molecule of *S. aureus* recognizes the sequence motif NNGRRN (R = A or G) and directs cleavage of a target nucleic acid sequence 1 to 10, e.g., 3 to 5, base pairs upstream from that sequence. In an embodiment, an eaCas9 molecule of *S. aureus* recognizes the sequence motif NNGRRT (R = A or G) and directs cleavage of a target nucleic acid sequence 1 to 10, e.g., 3 to 5, base pairs upstream from that sequence. In an embodiment, an eaCas9 molecule of *S. aureus* recognizes the sequence motif NNGRRV (R = A or G, V = A, G or C) and directs cleavage of a target nucleic acid sequence 1 to 10, e.g., 3 to 5, base pairs upstream from that sequence. In an embodiment, an eaCas9 molecule of *Neisseria meningitidis* recognizes the sequence motif NNNNGATT or NNNGCTT and directs cleavage of a target nucleic acid sequence 1 to 10, e.g., 3 to 5, base pairs upstream from that sequence. See, e.g., Hou *et al.*,

PNAS Early Edition 2013, 1-6. The ability of a Cas9 molecule to recognize a PAM sequence can be determined, e.g., using a transformation assay described in Jinek *et al.*, Science 2012 337:816. In the aforementioned embodiments, N can be any nucleotide residue, e.g., any of A, G, C or T.

As is discussed herein, Cas9 molecules can be engineered to alter the PAM specificity of the Cas9 molecule.

Exemplary naturally occurring Cas9 molecules are described in Chylinski *et al.*, RNA BIOLOGY 2013 10:5, 727-737. Such Cas9 molecules include Cas9 molecules of a cluster 1 bacterial family, cluster 2 bacterial family, cluster 3 bacterial family, cluster 4 bacterial family, cluster 5 bacterial family, cluster 6 bacterial family, a cluster 7 bacterial family, a cluster 8 bacterial family, a cluster 9 bacterial family, a cluster 10 bacterial family, a cluster 11 bacterial family, a cluster 12 bacterial family, a cluster 13 bacterial family, a cluster 14 bacterial family, a cluster 15 bacterial family, a cluster 16 bacterial family, a cluster 17 bacterial family, a cluster 18 bacterial family, a cluster 19 bacterial family, a cluster 20 bacterial family, a cluster 21 bacterial family, a cluster 22 bacterial family, a cluster 23 bacterial family, a cluster 24 bacterial family, a cluster 25 bacterial family, a cluster 26 bacterial family, a cluster 27 bacterial family, a cluster 28 bacterial family, a cluster 29 bacterial family, a cluster 30 bacterial family, a cluster 31 bacterial family, a cluster 32 bacterial family, a cluster 33 bacterial family, a cluster 34 bacterial family, a cluster 35 bacterial family, a cluster 36 bacterial family, a cluster 37 bacterial family, a cluster 38 bacterial family, a cluster 39 bacterial family, a cluster 40 bacterial family, a cluster 41 bacterial family, a cluster 42 bacterial family, a cluster 43 bacterial family, a cluster 44 bacterial family, a cluster 45 bacterial family, a cluster 46 bacterial family, a cluster 47 bacterial family, a cluster 48 bacterial family, a cluster 49 bacterial family, a cluster 50 bacterial family, a cluster 51 bacterial family, a cluster 52 bacterial family, a cluster 53 bacterial family, a cluster 54 bacterial family, a cluster 55 bacterial family, a cluster 56 bacterial family, a cluster 57 bacterial family, a cluster 58 bacterial family, a cluster 59 bacterial family, a cluster 60 bacterial family, a cluster 61 bacterial family, a cluster 62 bacterial family, a cluster 63 bacterial family, a cluster 64 bacterial family, a cluster 65 bacterial family, a cluster 66 bacterial family, a cluster 67 bacterial family, a cluster 68 bacterial family, a cluster 69 bacterial family, a cluster 70 bacterial family, a cluster 71 bacterial family, a cluster 72 bacterial family, a cluster 73 bacterial family, a cluster 74 bacterial family, a cluster 75 bacterial family, a cluster 76 bacterial family, a cluster 77 bacterial family, or a cluster 78 bacterial family.

Exemplary naturally occurring Cas9 molecules include a Cas9 molecule of a cluster 1 bacterial family. Examples include a Cas9 molecule of: *S. pyogenes* (e.g., strain SF370, MGAS10270, MGAS10750, MGAS2096, MGAS315, MGAS5005, MGAS6180, MGAS9429, NZ131 and SSI-1), *S. thermophilus* (e.g., strain LMD-9), *S. pseudoporcinus* (e.g., strain SPIN 20026), *S. mutans* (e.g., strain UA159, NN2025), *S. macacae* (e.g., strain NCTC11558), *S. gallolyticus* (e.g., strain UCN34, ATCC BAA-2069), *S. equines* (e.g., strain ATCC 9812, MGCS 124), *S. dysdalactiae* (e.g., strain GGS 124), *S. bovis* (e.g., strain ATCC 700338), *S. anginosus* (e.g., strain F0211), *S. agalactiae* (e.g., strain NEM316, A909), *Listeria monocytogenes* (e.g., strain F6854), *Listeria innocua* (*L. innocua*, e.g., strain Clip11262), *Enterococcus italicus* (e.g., strain DSM 15952), or *Enterococcus faecium* (e.g., strain 1,231,408). Another exemplary Cas9 molecule is a Cas9 molecule of *Neisseria meningitidis* (Hou *et al.*, PNAS Early Edition 2013, 1-6).

In an embodiment, a Cas9 molecule or Cas9 polypeptide, e.g., an eaCas9 molecule or eaCas9 polypeptide, comprises an amino acid sequence:

having 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% homology with;

differs at no more than, 2, 5, 10, 15, 20, 30, or 40% of the amino acid residues when compared with;

differs by at least 1, 2, 5, 10 or 20 amino acids but by no more than 100, 80, 70, 60, 50, 40 or 30 amino acids from; or

is identical to any Cas9 molecule sequence described herein, or a naturally occurring Cas9 molecule sequence, e.g., a Cas9 molecule from a species listed herein or described in Chylinski *et al.*, RNA BIOLOGY 2013 10:5, 727-737; Hou *et al.*, PNAS Early Edition 2013, 1-6; SEQ ID NOS:1-4. In an embodiment, the Cas9 molecule or Cas9 polypeptide comprises one or more of the following activities: a nickase activity; a double stranded cleavage activity (e.g., an endonuclease and/or exonuclease activity); a helicase activity; or the ability, together with a gRNA molecule, to home to a target nucleic acid.

In an embodiment, a Cas9 molecule or Cas9 polypeptide comprises the amino acid sequence of the consensus sequence of **Figs. 2A-2G**, wherein "*" indicates any amino acid found in the corresponding position in the amino acid sequence of a Cas9 molecule of *S. pyogenes, S. thermophilus, S. mutans* and *L. innocua*, and "-" indicates any amino acid. In an embodiment, a

Cas9 molecule or Cas9 polypeptide differs from the sequence of the consensus sequence disclosed in **Figs. 2A-2G** by at least 1, but no more than 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acid residues. In an embodiment, a Cas9 molecule or Cas9 polypeptide comprises the amino acid sequence of SEQ ID NO:7 of **Figs. 7A-7B**, wherein "*" indicates any amino acid found in the corresponding position in the amino acid sequence of a Cas9 molecule of *S. pyogenes, or N. meningitidis*, "-" indicates any amino acid, and "-" indicates any amino acid or absent. In an embodiment, a Cas9 molecule or Cas9 polypeptide differs from the sequence of SEQ ID NO:6 or 7 disclosed in **Figs. 7A-7B** by at least 1, but no more than 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acid residues.

A comparison of the sequence of a number of Cas9 molecules indicate that certain regions are conserved. These are identified below as:

region 1 (residues1 to 180, or in the case of region 1'residues 120 to 180)

region 2 (residues360 to 480);

region 3 (residues 660 to 720);

region 4 (residues 817 to 900); and

region 5 (residues 900 to 960);

In an embodiment, a Cas9 molecule or Cas9 polypeptide comprises regions 1-5, together with sufficient additional Cas9 molecule sequence to provide a biologically active molecule, e.g., a Cas9 molecule having at least one activity described herein. In an embodiment, each of regions 1-6, independently, have, 50%, 60%, 70%, or 80% homology with the corresponding residues of a Cas9 molecule or Cas9 polypeptide described herein, e.g., a sequence from **Figs. 2A-2G** or from **Figs. 7A-7B**.

In an embodiment, a Cas9 molecule or Cas9 polypeptide, e.g., an eaCas9 molecule or eaCas9 polypeptide, comprises an amino acid sequence referred to as region 1:

having 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% homology with amino acids 1-180 (the numbering is according to the motif sequence in **Figs. 2A-2G**; 52% of residues in the four Cas9 sequences in **Figs. 2A-2G** are conserved) of the amino acid sequence of Cas9 of *S. pyogenes*;

differs by at least 1, 2, 5, 10 or 20 amino acids but by no more than 90, 80, 70, 60, 50, 40 or 30 amino acids from amino acids 1-180 of the amino acid sequence of Cas9 of *S. pyogenes, S. thermophilus, S. mutans* or *L. innocua*; or

is identical to 1-180 of the amino acid sequence of Cas9 of *S. pyogenes, S. thermophilus, S. mutans* or *L. innocua*.

In an embodiment, a Cas9 molecule or Cas9 polypeptide, e.g., an eaCas9 molecule or eaCas9 polypeptide, comprises an amino acid sequence referred to as region 1':

having 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% homology with amino acids 120-180 (55% of residues in the four Cas9 sequences in **Figs. 2A-2G** are conserved) of the amino acid sequence of Cas9 of *S. pyogenes, S. thermophilus, S. mutans* or *L. innocua*;

differs by at least 1, 2, or 5 amino acids but by no more than 35, 30, 25, 20 or 10 amino acids from amino acids 120-180 of the amino acid sequence of Cas9 of *S. pyogenes, S. thermophilus, S. mutans* or *L. innocua* ; or

is identical to 120-180 of the amino acid sequence of Cas9 of *S. pyogenes, S. thermophilus, S. mutans* or *L. innocua.*

In an embodiment, a Cas9 molecule or Cas9 polypeptide, e.g., an eaCas9 molecule or eaCas9 polypeptide, comprises an amino acid sequence referred to as region 2:

having 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% homology with amino acids 360-480 (52% of residues in the four Cas9 sequences in **Figs. 2A-2G** are conserved) of the amino acid sequence of Cas9 of *S. pyogenes, S. thermophilus, S. mutans* or *L. innocua*;

differs by at least 1, 2, or 5 amino acids but by no more than 35, 30, 25, 20 or 10 amino acids from amino acids 360-480 of the amino acid sequence of Cas9 of *S. pyogenes, S. thermophilus, S. mutans* or *L. innocua*; or

is identical to 360-480 of the amino acid sequence of Cas9 of *S. pyogenes, S. thermophilus, S. mutans* or *L. innocua.*

In an embodiment, a Cas9 molecule or Cas9 polypeptide, e.g., an eaCas9 molecule or eaCas9 polypeptide, comprises an amino acid sequence referred to as region 3:

having 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% homology with amino acids 660-720 (56% of residues in the four Cas9 sequences in **Figs. 2A-2G** are conserved) of the amino acid sequence of Cas9 of *S. pyogenes, S. thermophilus, S. mutans* or *L. innocua*;

differs by at least 1, 2, or 5 amino acids but by no more than 35, 30, 25, 20 or 10 amino

acids from amino acids 660-720 of the amino acid sequence of Cas9 of *S. pyogenes, S. thermophilus, S. mutans* or *L. innocua*; or

is identical to 660-720 of the amino acid sequence of Cas9 of *S. pyogenes, S. thermophilus, S. mutans* or *L. innocua.*

In an embodiment, a Cas9 molecule or Cas9 polypeptide, e.g., an eaCas9 molecule or eaCas9 polypeptide, comprises an amino acid sequence referred to as region 4:

having 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% homology with amino acids 817-900 (55% of residues in the four Cas9 sequences in **Figs. 2A-2G** are conserved) of the amino acid sequence of Cas9 of *S. pyogenes, S. thermophilus, S. mutans* or *L. innocua*;

differs by at least 1, 2, or 5 amino acids but by no more than 35, 30, 25, 20 or 10 amino acids from amino acids 817-900 of the amino acid sequence of Cas9 of *S. pyogenes, S. thermophilus, S. mutans* or *L. innocua*; or

is identical to 817-900 of the amino acid sequence of Cas9 of *S. pyogenes, S. thermophilus, S. mutans* or *L. innocua.*

In an embodiment, a Cas9 molecule or Cas9 polypeptide, e.g., an eaCas9 molecule or eaCas9 polypeptide, comprises an amino acid sequence referred to as region 5:

having 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% homology with amino acids 900-960 (60% of residues in the four Cas9 sequences in Figs. 2A-2G are conserved) of the amino acid sequence of Cas9 of *S. pyogenes, S. thermophilus, S. mutans* or *L. innocua*;

differs by at least 1, 2, or 5 amino acids but by no more than 35, 30, 25, 20 or 10 amino acids from amino acids 900-960 of the amino acid sequence of Cas9 of *S. pyogenes, S. thermophilus, S. mutans* or *L. innocua*; or

is identical to 900-960 of the amino acid sequence of Cas9 of *S. pyogenes, S. thermophilus, S. mutans* or *L. innocua.*

Engineered or Altered Cas9 Molecules and Cas9 Polypeptides

Cas9 molecules and Cas9 polypeptides described herein, e.g., naturally occurring Cas9 molecules, can possess any of a number of properties, including: nickase activity, nuclease activity (e.g., endonuclease and/or exonuclease activity); helicase activity; the ability to associate

functionally with a gRNA molecule; and the ability to target (or localize to) a site on a nucleic acid (e.g., PAM recognition and specificity). In an embodiment, a Cas9 molecule or Cas9 polypeptide can include all or a subset of these properties. In typical embodiments, a Cas9 molecule or Cas9 polypeptide has the ability to interact with a gRNA molecule and, in concert with the gRNA molecule, localize to a site in a nucleic acid. Other activities, e.g., PAM specificity, cleavage activity, or helicase activity can vary more widely in Cas9 molecules and Cas9 polypeptides.

Cas9 molecules include engineered Cas9 molecules and engineered Cas9 polypeptides (engineered, as used in this context, means merely that the Cas9 molecule or Cas9 polypeptide differs from a reference sequences, and implies no process or origin limitation). An engineered Cas9 molecule or Cas9 polypeptide can comprise altered enzymatic properties, e.g., altered nuclease activity, (as compared with a naturally occurring or other reference Cas9 molecule) or altered helicase activity. As discussed herein, an engineered Cas9 molecule or Cas9 polypeptide can have nickase activity (as opposed to double strand nuclease activity). In an embodiment an engineered Cas9 molecule or Cas9 polypeptide can have an alteration that alters its size, e.g., a deletion of amino acid sequence that reduces its size, e.g., without significant effect on one or more, or any Cas9 activity. In an embodiment, an engineered Cas9 molecule or Cas9 polypeptide can comprise an alteration that affects PAM recognition. E.g., an engineered Cas9 molecule can be altered to recognize a PAM sequence other than that recognized by the endogenous wild-type PI domain. In an embodiment, a Cas9 molecule or Cas9 polypeptide can differ in sequence from a naturally occurring Cas9 molecule but not have significant alteration in one or more Cas9 activities.

Cas9 molecules or Cas9 polypeptides with desired properties can be made in a number of ways, e.g., by alteration of a parental, e.g., naturally occurring, Cas9 molecules or Cas9 polypeptides, to provide an altered Cas9 molecule or Cas9 polypeptide having a desired property. For example, one or more mutations or differences relative to a parental Cas9 molecule, e.g., a naturally occurring or engineered Cas9 molecule, can be introduced. Such mutations and differences comprise: substitutions (e.g., conservative substitutions or substitutions of non-essential amino acids); insertions; or deletions. In an embodiment, a Cas9 molecule or Cas9 polypeptide can comprises one or more mutations or differences, e.g., at least

1, 2, 3, 4, 5, 10, 15, 20, 30, 40 or 50 mutations, but less than 200, 100, or 80 mutations relative to a reference, e.g., a parental, Cas9 molecule.

In an embodiment, a mutation or mutations do not have a substantial effect on a Cas9 activity, e.g. a Cas9 activity described herein. In an embodiment, a mutation or mutations have a substantial effect on a Cas9 activity, e.g. a Cas9 activity described herein.

Non-Cleaving and Modified-Cleavage Cas9 Molecules and Cas9 Polypeptides

In an embodiment, a Cas9 molecule or Cas9 polypeptide comprises a cleavage property that differs from naturally occurring Cas9 molecules, e.g., that differs from the naturally occurring Cas9 molecule having the closest homology. For example, a Cas9 molecule or Cas9 polypeptide can differ from naturally occurring Cas9 molecules, e.g., a Cas9 molecule of *S. pyogenes*, as follows: its ability to modulate, e.g., decreased or increased, cleavage of a double stranded nucleic acid (endonuclease and/or exonuclease activity), e.g., as compared to a naturally occurring Cas9 molecule (e.g., a Cas9 molecule of *S. pyogenes)*; its ability to modulate, e.g., decreased or increased, cleavage of a single strand of a nucleic acid, e.g., a non-complementary strand of a nucleic acid molecule or a complementary strand of a nucleic acid molecule (nickase activity) , e.g., as compared to a naturally occurring Cas9 molecule (e.g., a Cas9 molecule of *S. pyogenes)*; or the ability to cleave a nucleic acid molecule, e.g., a double stranded or single stranded nucleic acid molecule, can be eliminated.

Modified Cleavage eaCas9 Molecules and eaCas9 Polypeptides

In an embodiment, an eaCas9 molecule or eaCas9 polypeptide comprises one or more of the following activities: cleavage activity associated with an N-terminal RuvC-like domain; cleavage activity associated with an HNH-like domain; cleavage activity associated with an HNH-like domain and cleavage activity associated with an N-terminal RuvC-like domain.

In an embodiment, an eaCas9 molecule or eaCas9 polypeptide comprises an active, or cleavage competent, HNH-like domain (e.g., an HNH-like domain described herein, e.g., SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO: 20, or SEQ ID NO:21) and an inactive, or cleavage incompetent, N-terminal RuvC-like domain. An exemplary inactive, or cleavage incompetent N-terminal RuvC-like domain can have a mutation of an aspartic acid in an N-terminal RuvC-like domain, e.g., an aspartic acid at position 9 of the consensus sequence

disclosed in **Figs. 2A-2G** or an aspartic acid at position 10 of SEQ ID NO:7, e.g., can be substituted with an alanine. In an embodiment, the eaCas9 molecule or eaCas9 polypeptide differs from wild type in the N-terminal RuvC-like domain and does not cleave the target nucleic acid, or cleaves with significantly less efficiency, e.g., less than 20, 10, 5, 1 or .1 % of the cleavage activity of a reference Cas9 molecule, e.g., as measured by an assay described herein. The reference Cas9 molecule can by a naturally occurring unmodified Cas9 molecule, e.g., a naturally occurring Cas9 molecule such as a Cas9 molecule of *S. pyogenes*, or *S. thermophilus*. In an embodiment, the reference Cas9 molecule is the naturally occurring Cas9 molecule having the closest sequence identity or homology.

In an embodiment, an eaCas9 molecule or eaCas9 polypeptide comprises an inactive, or cleavage incompetent, HNH domain and an active, or cleavage competent, N-terminal RuvC-like domain (e.g., an N-terminal RuvC-like domain described herein, e.g., SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO: 12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, or SEQ ID NO:16). Exemplary inactive, or cleavage incompetent HNH-like domains can have a mutation at one or more of: a histidine in an HNH-like domain, e.g., a histidine shown at position 856 of **Figs. 2A-2G**, e.g., can be substituted with an alanine; and one or more asparagines in an HNH-like domain, e.g., an asparagine shown at position 870 of **Figs. 2A-2G** and/or at position 879 of **Figs. 2A-2G**, e.g., can be substituted with an alanine. In an embodiment, the eaCas9 differs from wild type in the HNH-like domain and does not cleave the target nucleic acid, or cleaves with significantly less efficiency, e.g., less than 20, 10, 5, 1 or 0.1% of the cleavage activity of a reference Cas9 molecule, e.g., as measured by an assay described herein. The reference Cas9 molecule can by a naturally occurring unmodified Cas9 molecule, e.g., a naturally occurring Cas9 molecule such as a Cas9 molecule of *S. pyogenes*, or *S. thermophilus*. In an embodiment, the reference Cas9 molecule is the naturally occurring Cas9 molecule having the closest sequence identity or homology.

Alterations in the Ability to Cleave One or Both Strands of a Target Nucleic Acid

In an embodiment, exemplary Cas9 activities comprise one or more of PAM specificity, cleavage activity, and helicase activity. A mutation(s) can be present, e.g., in one or more RuvC-like domain, e.g., an N-terminal RuvC-like domain; an HNH-like domain; a region outside the RuvC-like domains and the HNH-like domain. In some embodiments, a mutation(s) is present in

a RuvC-like domain, e.g., an N-terminal RuvC-like domain. In some embodiments, a mutation(s) is present in an HNH-like domain. In some embodiments, mutations are present in both a RuvC-like domain, e.g., an N-terminal RuvC-like domain, and an HNH-like domain.

Exemplary mutations that may be made in the RuvC domain or HNH domain with reference to the *S. pyogenes* sequence include: D10A, E762A, H840A, N854A, N863A and/or D986A.

In an embodiment, a Cas9 molecule or Cas9 polypeptide is an eiCas9 molecule or eiCas9 polypeptide comprising one or more differences in a RuvC domain and/or in an HNH domain as compared to a reference Cas9 molecule, and the eiCas9 molecule or eiCas9 polypeptide does not cleave a nucleic acid, or cleaves with significantly less efficiency than does wildype, e.g., when compared with wild type in a cleavage assay, e.g., as described herein, cuts with less than 50, 25, 10, or 1% of a reference Cas9 molecule, as measured by an assay described herein.

Whether or not a particular sequence, e.g., a substitution, may affect one or more activity, such as targeting activity, cleavage activity, etc., can be evaluated or predicted, e.g., by evaluating whether the mutation is conservative or by the method described in Section IV. In an embodiment, a "non-essential" amino acid residue, as used in the context of a Cas9 molecule, is a residue that can be altered from the wild-type sequence of a Cas9 molecule, e.g., a naturally occurring Cas9 molecule, e.g., an eaCas9 molecule, without abolishing or more preferably, without substantially altering a Cas9 activity (e.g., cleavage activity), whereas changing an "essential" amino acid residue results in a substantial loss of activity (e.g., cleavage activity).

In an embodiment, a Cas9 molecule or Cas9 polypeptide comprises a cleavage property that differs from naturally occurring Cas9 molecules, e.g., that differs from the naturally occurring Cas9 molecule having the closest homology. For example, a Cas9 molecule or Cas9 polypeptide can differ from naturally occurring Cas9 molecules, e.g., a Cas9 molecule of *S aureus*, *S. pyogenes*, or *C. jejuni* as follows: its ability to modulate, e.g., decreased or increased, cleavage of a double stranded break (endonuclease and/or exonuclease activity), e.g., as compared to a naturally occurring Cas9 molecule (e.g., a Cas9 molecule of *S aureus*, *S. pyogenes,* or *C. jejuni)*; its ability to modulate, e.g., decreased or increased, cleavage of a single strand of a nucleic acid, e.g., a non-complimentary strand of a nucleic acid molecule or a complementary strand of a nucleic acid molecule (nickase activity), e.g., as compared to a naturally occurring Cas9 molecule (e.g., a Cas9 molecule of *S aureus*, *S. pyogenes,* or *C. jejuni*);

or the ability to cleave a nucleic acid molecule, e.g., a double stranded or single stranded nucleic acid molecule, can be eliminated.

In an embodiment, the altered Cas9 molecule or Cas9 polypeptide is an eaCas9 molecule or eaCas9 polypeptide comprising one or more of the following activities: cleavage activity associated with a RuvC domain; cleavage activity associated with an HNH domain; cleavage activity associated with an HNH domain and cleavage activity associated with a RuvC domain.

In an embodiment, the altered Cas9 molecule or Cas9 polypeptide is an eiCas9 molecule or eiCas9 polypeptide which does not cleave a nucleic acid molecule (either double stranded or single stranded nucleic acid molecules) or cleaves a nucleic acid molecule with significantly less efficiency, e.g., less than 20, 10, 5, 1 or 0.1% of the cleavage activity of a reference Cas9 molecule, e.g., as measured by an assay described herein. The reference Cas9 molecule can be a naturally occurring unmodified Cas9 molecule, e.g., a naturally occurring Cas9 molecule such as a Cas9 molecule of *S. pyogenes*, *S. thermophilus, S. aureus, C. jejuni* or *N. meningitidis*. In an embodiment, the reference Cas9 molecule is the naturally occurring Cas9 molecule having the closest sequence identity or homology. In an embodiment, the eiCas9 molecule or eiCas9 polypeptide lacks substantial cleavage activity associated with a RuvC domain and cleavage activity associated with an HNH domain.

In an embodiment, the altered Cas9 molecule or Cas9 polypeptide is an eaCas9 molecule or eaCas9 polypeptide comprising the fixed amino acid residues of *S. pyogenes* shown in the consensus sequence disclosed in **Figs. 2A-2G**, and has one or more amino acids that differ from the amino acid sequence of *S. pyogenes* (e.g., has a substitution) at one or more residue (e.g., 2, 3, 5, 10, 15, 20, 30, 50, 70, 80, 90, 100, 200 amino acid residues) represented by an "-" in the consensus sequence disclosed in **Figs. 2A-2G** or SEQ ID NO:7.

In an embodiment, the altered Cas9 molecule or Cas9 polypeptide comprises a sequence in which:

the sequence corresponding to the fixed sequence of the consensus sequence disclosed in **Figs. 2A-2G** differs at no more than 1, 2, 3, 4, 5, 10, 15, or 20% of the fixed residues in the consensus sequence disclosed in **Figs. 2A-2G**;

the sequence corresponding to the residues identified by "*" in the consensus sequence disclosed in **Figs. 2A-2G** differ at no more than 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, or 40% of the

"*" residues from the corresponding sequence of naturally occurring Cas9 molecule, e.g., an *S. pyogenes* Cas9 molecule; and,

the sequence corresponding to the residues identified by "-" in the consensus sequence disclosed in **Figs. 2A-2G** differ at no more than 5, 10, 15, 20, 25, 30, 35, 40, 45, 55, or 60% of the "-" residues from the corresponding sequence of naturally occurring Cas9 molecule, e.g., an *S. pyogenes* Cas9 molecule.

In an embodiment, the altered Cas9 molecule or Cas9 polypeptide is an eaCas9 molecule or eaCas9 polypeptide comprising the fixed amino acid residues of *S. thermophilus* shown in the consensus sequence disclosed in **Figs. 2A-2G**, and has one or more amino acids that differ from the amino acid sequence of *S. thermophilus* (e.g., has a substitution) at one or more residue (e.g., 2, 3, 5, 10, 15, 20, 30, 50, 70, 80, 90, 100, 200 amino acid residues) represented by an "-" in the consensus sequence disclosed in **Figs. 2A-2G**.

In an embodiment the altered Cas9 molecule or Cas9 polypeptide comprises a sequence in which:

the sequence corresponding to the fixed sequence of the consensus sequence disclosed in **Figs. 2A-2G** differs at no more than 1, 2, 3, 4, 5, 10, 15, or 20% of the fixed residues in the consensus sequence disclosed in **Figs. 2A-2G**;

the sequence corresponding to the residues identified by "*"in the consensus sequence disclosed in **Figs. 2A-2G** differ at no more than 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, or 40% of the "*" residues from the corresponding sequence of naturally occurring Cas9 molecule, e.g., an *S. thermophilus* Cas9 molecule; and,

the sequence corresponding to the residues identified by "-" in the consensus sequence disclosed in **Figs. 2A-2G** differ at no more than 5, 10, 15, 20, 25, 30, 35, 40, 45, 55, or 60% of the "-" residues from the corresponding sequence of naturally occurring Cas9 molecule, e.g., an *S. thermophilus* Cas9 molecule.

In an embodiment, the altered Cas9 molecule or Cas9 polypeptide is an eaCas9 molecule or eaCas9 polypeptide comprising the fixed amino acid residues of *S. mutans* shown in the consensus sequence disclosed in **Figs. 2A-2G**, and has one or more amino acids that differ from the amino acid sequence of *S. mutans* (e.g., has a substitution) at one or more residue (e.g., 2, 3, 5, 10, 15, 20, 30, 50, 70, 80, 90, 100, 200 amino acid residues) represented by an "-" in the consensus sequence disclosed in **Figs. 2A-2G**.

In an embodiment, the altered Cas9 molecule or Cas9 polypeptide comprises a sequence in which:

the sequence corresponding to the fixed sequence of the consensus sequence disclosed in **Figs. 2A-2G** differs at no more than 1, 2, 3, 4, 5, 10, 15, or 20% of the fixed residues in the consensus sequence disclosed in **Figs. 2A-2G**;

the sequence corresponding to the residues identified by "*" in the consensus sequence disclosed in **Figs. 2A-2G** differ at no more than 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, or 40% of the "*" residues from the corresponding sequence of naturally occurring Cas9 molecule, e.g., an *S. mutans* Cas9 molecule; and,

the sequence corresponding to the residues identified by "-" in the consensus sequence disclosed in **Figs. 2A-2G** differ at no more than 5, 10, 15, 20, 25, 30, 35, 40, 45, 55, or 60% of the "-" residues from the corresponding sequence of naturally occurring Cas9 molecule, e.g., an *S. mutans* Cas9 molecule.

In an embodiment, the altered Cas9 molecule or Cas9 polypeptide is an eaCas9 molecule or eaCas9 polypeptide comprising the fixed amino acid residues of *L. innocula* shown in the consensus sequence disclosed in **Figs. 2A-2G**, and has one or more amino acids that differ from the amino acid sequence of *L. innocula* (e.g., has a substitution) at one or more residue (e.g., 2, 3, 5, 10, 15, 20, 30, 50, 70, 80, 90, 100, 200 amino acid residues) represented by an "-"in the consensus sequence disclosed in **Figs. 2A-2G**.

In an embodiment, the altered Cas9 molecule or Cas9 polypeptide comprises a sequence in which:

the sequence corresponding to the fixed sequence of the consensus sequence disclosed in **Figs. 2A-2G** differs at no more than 1, 2, 3, 4, 5, 10, 15, or 20% of the fixed residues in the consensus sequence disclosed in **Figs. 2A-2G**;

the sequence corresponding to the residues identified by "*" in the consensus sequence disclosed in **Figs. 2A-2G** differ at no more than 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, or 40% of the "*" residues from the corresponding sequence of naturally occurring Cas9 molecule, e.g., an *L. innocula* Cas9 molecule; and,

the sequence corresponding to the residues identified by "-" in the consensus sequence disclosed in **Figs. 2A-2G** differ at no more than 5, 10, 15, 20, 25, 30, 35, 40, 45, 55, or 60% of

the "-" residues from the corresponding sequence of naturally occurring Cas9 molecule, e.g., an *L. innocula* Cas9 molecule.

In an embodiment, the altered Cas9 molecule or Cas9 polypeptide, e.g., an eaCas9 molecule, can be a fusion, e.g., of two of more different Cas9 molecules or Cas9 polypeptides, e.g., of two or more naturally occurring Cas9 molecules of different species. For example, a fragment of a naturally occurring Cas9 molecule of one species can be fused to a fragment of a Cas9 molecule of a second species. As an example, a fragment of Cas9 molecule of *S. pyogenes* comprising an N-terminal RuvC-like domain can be fused to a fragment of Cas9 molecule of a species other than *S. pyogenes* (e.g., *S. thermophilus*) comprising an HNH-like domain.

Cas9 Molecules and Cas9 Polypeptides with Altered PAM Recognition or No PAM Recognition

Naturally occurring Cas9 molecules can recognize specific PAM sequences, for example, the PAM recognition sequences described above for *S. pyogenes*, *S. thermophilus, S. mutans, S. aureus* and *N. meningitidis*.

In an embodiment, a Cas9 molecule or Cas9 polypeptide has the same PAM specificities as a naturally occurring Cas9 molecule. In other embodiments, a Cas9 molecule or Cas9 polypeptide has a PAM specificity not associated with a naturally occurring Cas9 molecule, or a PAM specificity not associated with the naturally occurring Cas9 molecule to which it has the closest sequence homology. For example, a naturally occurring Cas9 molecule can be altered, e.g., to alter PAM recognition, e.g., to alter the PAM sequence that the Cas9 molecule recognizes to decrease off target sites and/or improve specificity; or eliminate a PAM recognition requirement. In an embodiment, a Cas9 molecule or Cas9 polypeptide can be altered, e.g., to increase length of PAM recognition sequence and/or improve Cas9 specificity to high level of identity, e.g., to decrease off target sites and increase specificity. In an embodiment, the length of the PAM recognition sequence is at least 4, 5, 6, 7, 8, 9, 10 or 15 amino acids in length. Cas9 molecules or Cas9 polypeptides that recognize different PAM sequences and/or have reduced off-target activity can be generated using directed evolution. Exemplary methods and systems that can be used for directed evolution of Cas9 molecules are described, e.g., in Esvelt *et al.* NATURE 2011, 472(7344): 499-503. Candidate Cas9 molecules can be evaluated, e.g., by methods described in Section IV.

Alterations of the PI domain, which mediates PAM recognition, are discussed below.

Synthetic Cas9 Molecules and Cas9 Polypeptides with Altered PI Domains

Current genome-editing methods are limited in the diversity of target sequences that can be targeted by the PAM sequence that is recognized by the Cas9 molecule utilized. A synthetic Cas9 molecule (or Syn-Cas9 molecule), or synthetic Cas9 polypeptide (or Syn-Cas9 polypeptide), as that term is used herein, refers to a Cas9 molecule or Cas9 polypeptide that comprises a Cas9 core domain from one bacterial species and a functional altered PI domain, i.e., a PI domain other than that naturally associated with the Cas9 core domain, e.g., from a different bacterial species.

In an embodiment, the altered PI domain recognizes a PAM sequence that is different from the PAM sequence recognized by the naturally-occurring Cas9 from which the Cas9 core domain is derived. In an embodiment, the altered PI domain recognizes the same PAM sequence recognized by the naturally-occurring Cas9 from which the Cas9 core domain is derived, but with different affinity or specificity. A Syn-Cas9 molecule or Syn-Cas9 polypetide can be, respectively, a Syn-eaCas9 molecule or Syn-eaCas9 polypeptide or a Syn-eiCas9 molecule Syn-eiCas9 polypeptide.

An exemplary Syn-Cas9 molecule or Syn-Cas9 polypetide comprises:

a) a Cas9 core domain, e.g., a Cas9 core domain from **Table 11** or **12**, e.g., a *S. aureus*, *S. pyogenes,* or *C. jejuni* Cas9 core domain; and

b) an altered PI domain from a species X Cas9 sequence selected from **Tables 14** and **15**.

In an embodiment, the RKR motif (the PAM binding motif) of said altered PI domain comprises: differences at 1, 2, or 3 amino acid residues; a difference in amino acid sequence at the first, second, or third position; differences in amino acid sequence at the first and second positions, the first and third positions, or the second and third positions; as compared with the sequence of the RKR motif of the native or endogenous PI domain associated with the Cas9 core domain.

In an embodiment, the Cas9 core domain comprises the Cas9 core domain from a species X Cas9 from **Table 11** and said altered PI domain comprises a PI domain from a species Y Cas9 from **Table 11**.

In an embodiment, the RKR motif of the species X Cas9 is other than the RKR motif of the species Y Cas9.

In an embodiment, the RKR motif of the altered PI domain is selected from XXY, XNG, and XNQ.

In an embodiment, the altered PI domain has at least 60, 70, 80, 90, 95, or 100% homology with the amino acid sequence of a naturally occurring PI domain of said species Y from **Table 11**.

In an embodiment, the altered PI domain differs by no more than 50, 40, 30, 25, 20, 15, 10, 5, 4, 3, 2, or 1 amino acid residue from the amino acid sequence of a naturally occurring PI domain of said second species from **Table 11**.

In an embodiment, the Cas9 core domain comprises a *S. aureus* core domain and altered PI domain comprises: an *A. denitrificans* PI domain; a *C. jejuni* PI domain; a *H. mustelae* PI domain; or an altered PI domain of species X PI domain, wherein species X is selected from **Table 15**.

In an embodiment, the Cas9 core domain comprises a *S. pyogenes* core domain and the altered PI domain comprises: an *A. denitrificans* PI domain; a *C. jejuni* PI domain; a *H. mustelae* PI domain; or an altered PI domain of species X PI domain, wherein species X is selected from **Table 15**.

In an embodiment, the Cas9 core domain comprises a *C. jejuni* core domain and the altered PI domain comprises: an *A. denitrificans* PI domain; a *H. mustelae* PI domain; or an altered PI domain of species X PI domain, wherein species X is selected from **Table 15**.

In an embodiment, the Cas9 molecule or Cas9 polypeptide further comprises a linker disposed between said Cas9 core domain and said altered PI domain.

In an embodiment, the linker comprises: a linker described elsewhere herein disposed between the Cas9 core domain and the heterologous PI domain. Suitable linkers are further described in Section V.

Exemplary altered PI domains for use in Syn-Cas9 molecules are described in **Tables 14** and **15**. The sequences for the 83 Cas9 orthologs referenced in **Tables 14** and **15** are provided in **Table 11**. **Table 13** provides the Cas9 orthologs with known PAM sequences and the corresponding RKR motif.

In an embodiment, a Syn-Cas9 molecule or Syn-Cas9 polypeptide may also be size-optimized, e.g., the Syn-Cas9 molecule or Syn-Cas9 polypeptide comprises one or more deletions, and optionally one or more linkers disposed between the amino acid residues flanking

the deletions. In an embodiment, a Syn-Cas9 molecule or Syn-Cas9 polypeptide comprises a REC deletion.

Size-Optimized Cas9 Molecules and Cas9 Polypeptides

Engineered Cas9 molecules and engineered Cas9 polypeptides described herein include a Cas9 molecule or Cas9 polypeptide comprising a deletion that reduces the size of the molecule while still retaining desired Cas9 properties, e.g., essentially native conformation, Cas9 nuclease activity, and/or target nucleic acid molecule recognition. Provided herein are Cas9 molecules or Cas9 polypeptides comprising one or more deletions and optionally one or more linkers, wherein a linker is disposed between the amino acid residues that flank the deletion. Methods for identifying suitable deletions in a reference Cas9 molecule, methods for generating Cas9 molecules with a deletion and a linker, and methods for using such Cas9 molecules will be apparent to one of ordinary skill in the art upon review of this document.

A Cas9 molecule, e.g., a *S. aureus*, *S. pyogenes*, or *C. jejuni,* Cas9 molecule, having a deletion is smaller, e.g., has reduced number of amino acids, than the corresponding naturally-occurring Cas9 molecule. The smaller size of the Cas9 molecules allows increased flexibility for delivery methods, and thereby increases utility for genome-editing. A Cas9 molecule or Cas9 polypeptide can comprise one or more deletions that do not substantially affect or decrease the activity of the resultant Cas9 molecules or Cas9 polypeptides described herein. Activities that are retained in the Cas9 molecules or Cas9 polypeptides comprising a deletion as described herein include one or more of the following:

a nickase activity, i.e., the ability to cleave a single strand, e.g., the non-complementary strand or the complementary strand, of a nucleic acid molecule; a double stranded nuclease activity, i.e., the ability to cleave both strands of a double stranded nucleic acid and create a double stranded break, which in an embodiment is the presence of two nickase activities;

an endonuclease activity;

an exonuclease activity;

a helicase activity, i.e., the ability to unwind the helical structure of a double stranded nucleic acid;

and recognition activity of a nucleic acid molecule, e.g., a target nucleic acid or a gRNA.

Activity of the Cas9 molecules or Cas9 polypeptides described herein can be assessed using the activity assays described herein or in the art.

Identifying regions suitable for deletion

Suitable regions of Cas9 molecules for deletion can be identified by a variety of methods. Naturally-occurring orthologous Cas9 molecules from various bacterial species, e.g., any one of those listed in **Table 11**, can be modeled onto the crystal structure of *S. pyogenes* Cas9 (Nishimasu et al., Cell, 156:935-949, 2014) to examine the level of conservation across the selected Cas9 orthologs with respect to the three-dimensional conformation of the protein. Less conserved or unconserved regions that are spatially located distant from regions involved in Cas9 activity, e.g., interface with the target nucleic acid molecule and/or gRNA, represent regions or domains are candidates for deletion without substantially affecting or decreasing Cas9 activity.

REC-Optimized Cas9 Molecules and Cas9 Polypeptides

A REC-optimized Cas9 molecule, or a REC-optimized Cas9 polypeptide, as that term is used herein, refers to a Cas9 molecule or Cas9 polypeptide that comprises a deletion in one or both of the REC2 domain and the $RE1_{CT}$ domain (collectively a REC deletion), wherein the deletion comprises at least 10% of the amino acid residues in the cognate domain. A REC-optimized Cas9 molecule or Cas9 polypeptide can be an eaCas9 molecule or eaCas9 polypetide, or an eiCas9 molecule or eiCas9 polypeptide. An exemplary REC-optimized Cas9 molecule or REC-optimized Cas9 polypeptide comprises:

a) a deletion selected from:

i) a REC2 deletion;

ii) a $REC1_{CT}$ deletion; or

iii) a $REC1_{SUB}$ deletion.

Optionally, a linker is disposed between the amino acid residues that flank the deletion. In an embodiment, a Cas9 molecule or Cas9 polypeptide includes only one deletion, or only two deletions. A Cas9 molecule or Cas9 polypeptide can comprise a REC2 deletion and a $REC1_{CT}$ deletion. A Cas9 molecule or Cas9 polypeptide can comprise a REC2 deletion and a $REC1_{SUB}$ deletion.

Generally, the deletion will contain at least 10% of the amino acids in the cognate domain, e.g., a REC2 deletion will include at least 10% of the amino acids in the REC2 domain.

A deletion can comprise: at least 10, 20, 30, 40, 50, 60, 70, 80, or 90% of the amino acid residues of its cognate domain; all of the amino acid residues of its cognate domain; an amino acid residue outside its cognate domain; a plurality of amino acid residues outside its cognate domain; the amino acid residue immediately N terminal to its cognate domain; the amino acid residue immediately C terminal to its cognate domain; the amino acid residue immediately N terminal to its cognate and the amino acid residue immediately C terminal to its cognate domain; a plurality of, e.g., up to 5, 10, 15, or 20, amino acid residues N terminal to its cognate domain; a plurality of, e.g., up to 5, 10, 15, or 20, amino acid residues C terminal to its cognate domain; a plurality of, e.g., up to 5, 10, 15, or 20, amino acid residues N terminal to to its cognate domain and a plurality of e.g., up to 5, 10, 15, or 20, amino acid residues C terminal to its cognate domain.

In an embodiment, a deletion does not extend beyond: its cognate domain; the N terminal amino acid residue of its cognate domain; the C terminal amino acid residue of its cognate domain.

A REC-optimized Cas9 molecule or REC-optimized Cas9 polypeptide can include a linker disposed between the amino acid residues that flank the deletion. Suitable linkers for use between the amino acid resides that flank a REC deletion in a REC-optimized Cas9 molecule or REC-optimized Cas9 polypeptide is disclosed in Section V.

In an embodiment, a REC-optimized Cas9 molecule or REC-optimized Cas9 polypeptide comprises an amino acid sequence that, other than any REC deletion and associated linker, has at least 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 99, or 100% homology with the amino acid sequence of a naturally occurring Cas 9, e.g., a Cas9 molecule described in **Table 11**, e.g., a *S. aureus* Cas9 molecule, a *S. pyogenes* Cas9 molecule, or a *C. jejuni* Cas9 molecule.

In an embodiment, a a REC-optimized Cas9 molecule or REC-optimized Cas9 polypeptide comprises an amino acid sequence that, other than any REC deletion and associated linker, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, or 25, amino acid residues from the amino acid sequence of a naturally occurring Cas 9, e.g., a Cas9 molecule described in **Table 11**, e.g., a *S. aureus* Cas9 molecule, a *S. pyogenes* Cas9 molecule, or a *C. jejuni* Cas9 molecule.

In an embodiment, a REC-optimized Cas9 molecule or REC-optimized Cas9 polypeptide comprises an amino acid sequence that, other than any REC deletion and associate linker, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, or 25% of the, amino acid residues from the amino acid sequence of a naturally occurring Cas 9, e.g., a Cas9 molecule described in **Table 11**, e.g., a *S. aureus* Cas9 molecule, a *S. pyogenes* Cas9 molecule, or a *C. jejuni* Cas9 molecule.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters. Methods of alignment of sequences for comparison are well known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith and Waterman, (1970) Adv. Appl. Math. 2:482c, by the homology alignment algorithm of Needleman and Wunsch, (1970) J. Mol. Biol. 48:443, by the search for similarity method of Pearson and Lipman, (1988) Proc. Nat'l. Acad. Sci. USA 85:2444, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (see, e.g., Brent et al., (2003) Current Protocols in Molecular Biology).

Two examples of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al., (1977) Nuc. Acids Res. 25:3389-3402; and Altschul et al., (1990) J. Mol. Biol. 215:403-410, respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information.

The percent identity between two amino acid sequences can also be determined using the algorithm of E. Meyers and W. Miller, (1988) Comput. Appl. Biosci. 4:11-17) which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4. In addition, the percent identity between two amino acid sequences can be determined using the Needleman and Wunsch (1970) J. Mol. Biol. 48:444-453) algorithm which has been incorporated into the GAP program in the GCG software

package (available at www.gcg.com), using either a Blossom 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6.

Sequence information for exemplary REC deletions are provided for 83 naturally-occurring Cas9 orthologs in **Table 11**.

5

The amino acid sequences of exemplary Cas9 molecules from different bacterial species are shown below.

Table 11. Amino Acid Sequence of Cas9 Orthologs

		REC2			$REC1_{CT}$			Rec_{sub}		
Species / Composite ID	**Amino acid sequence**	start (AA pos)	stop (AA pos)	# AA deleted (n)	start (AA pos)	stop (AA pos)	# AA deleted (n)	start (AA pos)	stop (AA pos)	# AA deleted (n)
Staphylococcus Aureus tr|J7RUA5|J7RUA5_STAAU	SEQ ID NO: 304	126	166	41	296	352	57	296	352	57
Streptococcus Pyogenes sp|Q99ZW2|CAS9_STRP1	SEQ ID NO: 305	176	314	139	511	592	82	511	592	82
Campylobacter jejuni NCTC 11168 gi|218563121|ref|YP_002344900.1	SEQ ID NO: 306	137	181	45	316	360	45	316	360	45
Bacteroides fragilis NCTC 9343 gi|60683389|ref|YP_213533.1|	SEQ ID NO: 307	148	339	192	524	617	84	524	617	84
Bifidobacterium bifidum S17 gi|310286728|ref|YP_003937986.	SEQ ID NO: 308	173	335	163	516	607	87	516	607	87
Veillonella atypica ACS-134-V-Col7a gi|303229466|ref|ZP_07316256.1	SEQ ID NO: 309	185	339	155	574	663	79	574	663	79
Lactobacillus rhamnosus GG gi|258509199|ref|YP_003171950.1	SEQ ID NO: 310	169	320	152	559	645	78	559	645	78
Filifactor alocis ATCC 35896 gi|374307738|ref|YP_005054169.1	SEQ ID NO: 311	166	314	149	508	592	76	508	592	76
Oenococcus kitaharae DSM 17330 gi|366983953|gb|EHN59352.1|	SEQ ID NO: 312	169	317	149	555	639	80	555	639	80
Fructobacillus fructosus KCTC 3544 gi|339625081|ref|ZP_08660870.1	SEQ ID NO: 313	168	314	147	488	571	76	488	571	76
Catenibacterium mitsuokai DSM 15897 gi|224543312|ref|ZP_03683851.1	SEQ ID NO: 314	173	318	146	511	594	78	511	594	78

Finegoldia magna ATCC 29328 gi|169823755|ref|YP_001691366.1	SEQ ID NO: 315	168	313	146	452	534	77	452	534	77
CoriobacteriumglomeransPW2 gi|328956315|ref|YP_004373648.1	SEQ ID NO: 316	175	318	144	511	592	82	511	592	82
Eubacterium yurii ATCC 43715 gi|306821691|ref|ZP_07455288.1	SEQ ID NO: 317	169	310	142	552	633	76	552	633	76
Peptoniphilus duerdenii ATCC BAA-1640 gi|304438954|ref|ZP_07398877.1	SEQ ID NO: 318	171	311	141	535	615	76	535	615	76
Acidaminococcus sp. D21 gi|227824983|ref|ZP_03989815.1	SEQ ID NO: 319	167	306	140	511	591	75	511	591	75
Lactobacillus farciminis KCTC 3681 gi|336394882|ref|ZP_08576281.1	SEQ ID NO: 320	171	310	140	542	621	85	542	621	85
Streptococcus sanguinis SK49 gi|422884106|ref|ZP_16930555.1	SEQ ID NO: 321	185	324	140	411	490	85	411	490	85
Coprococcus catus GD-7 gi|291520705|emb|CBK78998.1|	SEQ ID NO: 322	172	310	139	556	634	76	556	634	76
Streptococcus mutans UA159 gi|24379809|ref|NP_721764.1|	SEQ ID NO: 323	176	314	139	392	470	84	392	470	84
Streptococcus pyogenes M1 GAS gi|13622193|gb|AAK33936.1|	SEQ ID NO: 324	176	314	139	523	600	82	523	600	82
Streptococcus thermophilus LMD-9 gi|116628213|ref|YP_820832.1|	SEQ ID NO: 325	176	314	139	481	558	81	481	558	81
Fusobacteriumnucleatum ATCC49256 gi|34762592|ref|ZP_00143587.1|	SEQ ID NO: 326	171	308	138	537	614	76	537	614	76
Planococcus antarcticus DSM 14505 gi|389815359|ref|ZP_10206685.1	SEQ ID NO: 327	162	299	138	538	614	94	538	614	94
Treponema denticola ATCC 35405 gi|42525843|ref|NP_970941.1|	SEQ ID NO: 328	169	305	137	524	600	81	524	600	81
Solobacterium moorei F0204 gi|320528778|ref|ZP_08029929.1	SEQ ID NO: 329	179	314	136	544	619	77	544	619	77
Staphylococcus pseudintermedius ED99 gi|323463801|gb|ADX75954.1|	SEQ ID NO: 330	164	299	136	531	606	92	531	606	92
Flavobacterium branchiophilum FL-15 gi|347536497|ref|YP_004843922.1	SEQ ID NO: 331	162	286	125	538	613	63	538	613	63
Ignavibacterium album JCM 16511 gi|385811609|ref|YP_005848005	SEQ ID NO: 332	223	329	107	357	432	90	357	432	90

.1										
Bergeyella zoohelcum ATCC 43767 gi|423317190|ref|ZP_17295095.1	SEQ ID NO: 333	165	261	97	529	604	56	529	604	56
Nitrobacter hamburgensis X14 gi|92109262|ref|YP_571550.1|	SEQ ID NO: 334	169	253	85	536	611	48	536	611	48
Odoribacter laneus YIT 12061 gi|374384763|ref|ZP_09642280.1	SEQ ID NO: 335	164	242	79	535	610	63	535	610	63
Legionella pneumophila str. Paris gi|54296138|ref|YP_122507.1|	SEQ ID NO: 336	164	239	76	402	476	67	402	476	67
Bacteroides sp. 20 3 gi|301311869|ref|ZP_07217791.1	SEQ ID NO: 337	198	269	72	530	604	83	530	604	83
Akkermansia muciniphila ATCC BAA-835 gi|187736489|ref|YP_001878601 .	SEQ ID NO: 338	136	202	67	348	418	62	348	418	62
Prevotella sp. C561 gi|345885718|ref|ZP_08837074.1	SEQ ID NO: 339	184	250	67	357	425	78	357	425	78
Wolinella succinogenes DSM 1740 gi|34557932|ref|NP_907747.1|	SEQ ID NO: 340	157	218	36	401	468	60	401	468	60
Alicyclobacillus hesperidum URH17-3-68 gi|403744858|ref|ZP_10953934.1	SEQ ID NO: 341	142	196	55	416	482	61	416	482	61
Caenispirillum salinarum AK4 gi|427429481|ref|ZP_18919511.1	SEQ ID NO: 342	161	214	54	330	393	68	330	393	68
Eubacterium rectale ATCC 33656 gi|238924075|ref|YP_002937591 .1	SEQ ID NO: 343	133	185	53	322	384	60	322	384	60
Mycoplasma synoviae 53 gi|71894592|ref|YP_278700.1|	SEQ ID NO: 344	187	239	53	319	381	80	319	381	80
Porphyromonas sp. oral taxon 279 str. F0450 gi|402847315|ref|ZP_10895610.1	SEQ ID NO: 345	150	202	53	309	371	60	309	371	60
Streptococcus thermophilus LMD-9 gi|116627542|ref|YP_820161.1|	SEQ ID NO: 346	127	178	139	424	486	81	424	486	81
Roseburia inulinivorans DSM 16841 gi|225377804|ref|ZP_03755025.1	SEQ ID NO: 347	154	204	51	318	380	69	318	380	69
Methylosinus trichosporium OB3b gi|296446027|ref|ZP_06887976.1	SEQ ID NO: 348	144	193	50	426	488	64	426	488	64
Ruminococcus albus 8 gi|325677756|ref|ZP_08157403.1	SEQ ID NO: 349	139	187	49	351	412	55	351	412	55
Bifidobacterium longum DJO10A gi|189440764|ref|YP_001955845 .	SEQ ID NO: 350	183	230	48	370	431	44	370	431	44

Enterococcus faecalis TX0012 gi|315149830|gb|EFT93846.1|	SEQ ID NO: 351	123	170	48	327	387	60	327	387	60
Mycoplasma mobile 163K gi|47458868|ref|YP_015730.1|	SEQ ID NO: 352	179	226	48	314	374	79	314	374	79
Actinomyces coleocanis DSM 15436 gi|227494853|ref|ZP_03925169.1	SEQ ID NO: 353	147	193	47	358	418	40	358	418	40
Dinoroseobacter shibae DFL 12 gi|159042956|ref|YP_001531750.1	SEQ ID NO: 354	138	184	47	338	398	48	338	398	48
Actinomyces sp. oral taxon 180 str. F0310 gi|315605738|ref|ZP_07880770.1	SEQ ID NO: 355	183	228	46	349	409	40	349	409	40
Alcanivorax sp. W11-5 gi|407803669|ref|ZP_11150502.1	SEQ ID NO: 356	139	183	45	344	404	61	344	404	61
Aminomonas paucivorans DSM 12260 gi|312879015|ref|ZP_07738815.1	SEQ ID NO: 357	134	178	45	341	401	63	341	401	63
Mycoplasma canis PG 14 gi|384393286|gb|EIE39736.1|	SEQ ID NO: 358	139	183	45	319	379	76	319	379	76
Lactobacillus coryniformis KCTC 3535 gi|336393381|ref|ZP_08574780.1	SEQ ID NO: 359	141	184	44	328	387	61	328	387	61
Elusimicrobium minutum Pei191 gi|187250660|ref|YP_001875142.1	SEQ ID NO: 360	177	219	43	322	381	47	322	381	47
Neisseria meningitidis Z2491 gi|218767588|ref|YP_002342100.1	SEQ ID NO: 361	147	189	43	360	419	61	360	419	61
Pasteurella multocida str. Pm70 gi|15602992|ref|NP_246064.1|	SEQ ID NO: 362	139	181	43	319	378	61	319	378	61
Rhodovulum sp. PH10 gi|402849997|ref|ZP_10898214.1	SEQ ID NO: 363	141	183	43	319	378	48	319	378	48
Eubacterium dolichum DSM 3991 gi|160915782|ref|ZP_02077990.1	SEQ ID NO: 364	131	172	42	303	361	59	303	361	59
Nitratifractor salsuginis DSM 16511 gi|319957206|ref|YP_004168469.1	SEQ ID NO: 365	143	184	42	347	404	61	347	404	61
Rhodospirillum rubrum ATCC 11170 gi|83591793|ref|YP_425545.1|	SEQ ID NO: 366	139	180	42	314	371	55	314	371	55
Clostridium cellulolyticum H10 gi|220930482|ref|YP_002507391.1	SEQ ID NO: 367	137	176	40	320	376	61	320	376	61
Helicobacter mustelae 12198 gi|291276265|ref|YP_003516037.1	SEQ ID NO: 368	148	187	40	298	354	48	298	354	48
Ilyobacter polytropus DSM 2926 gi|310780384|ref|YP_003968716	SEQ ID NO: 369	134	173	40	462	517	63	462	517	63

.1										
Sphaerochaeta globus str. Buddy gi|325972003|ref|YP_004248194.1	SEQ ID NO: 370	163	202	40	335	389	45	335	389	45
Staphylococcus lugdunensis M23590 gi|315659848|ref|ZP_07912707.1	SEQ ID NO: 371	128	167	40	337	391	57	337	391	57
Treponema sp. JC4 gi|384109266|ref|ZP_10010146.1	SEQ ID NO: 372	144	183	40	328	382	63	328	382	63
uncultured delta proteobacterium HF0070 07E19 gi|297182908|gb|ADI19058.1|	SEQ ID NO: 373	154	193	40	313	365	55	313	365	55
Alicycliphilus denitrificans K601 gi|330822845|ref|YP_004386148.1	SEQ ID NO: 374	140	178	39	317	366	48	317	366	48
Azospirillum sp. B510 gi|288957741|ref|YP_003448082.1	SEQ ID NO: 375	205	243	39	342	389	46	342	389	46
Bradyrhizobium sp. BTAi1 gi|148255343|ref|YP_001239928.1	SEQ ID NO: 376	143	181	39	323	370	48	323	370	48
Parvibaculum lavamentivorans DS-1 gi|154250555|ref|YP_001411379.1	SEQ ID NO: 377	138	176	39	327	374	58	327	374	58
Prevotella timonensis CRIS 5C-B1 gi|282880052|ref|ZP_06288774.1	SEQ ID NO: 378	170	208	39	328	375	61	328	375	61
Bacillus smithii 7 3 47FAA gi|365156657|ref|ZP_09352959.1	SEQ ID NO: 379	134	171	38	401	448	63	401	448	63
Cand. Puniceispirillum marinum IMCC1322 gi|294086111|ref|YP_003552871.1	SEQ ID NO: 380	135	172	38	344	391	53	344	391	53
Barnesiella intestinihominis YIT 11860 gi|404487228|ref|ZP_11022414.1	SEQ ID NO: 381	140	176	37	371	417	60	371	417	60
Ralstonia syzygii R24 gi|344171927|emb|CCA84553.1|	SEQ ID NO: 382	140	176	37	395	440	50	395	440	50
Wolinella succinogenes DSM 1740 gi|34557790|ref|NP_907605.1|	SEQ ID NO: 383	145	180	36	348	392	60	348	392	60
Mycoplasma gallisepticum str. F gi|284931710|gb|ADC31648.1|	SEQ ID NO: 384	144	177	34	373	416	71	373	416	71
Acidothermus cellulolyticus 11B gi|117929158|ref|YP_873709.1|	SEQ ID NO: 385	150	182	33	341	380	58	341	380	58
Mycoplasma ovipneumoniae SC01 gi|363542550|ref|ZP_09312133.1	SEQ ID NO: 386	156	184	29	381	420	62	381	420	62

Table 12. Amino Acid Sequence of Cas9 Core Domains

Strain Name	Cas9 Start (AA pos)	Cas9 Stop (AA pos)
	Start and Stop numbers refer to the sequence in **Table 11**	
Staphylococcus Aureus	1	772
Streptococcus Pyogenes	1	1099
Campulobacter Jejuni	1	741

Table 13. Identified PAM sequences and corresponding RKR motifs.

Strain Name	PAM sequence (NA)	RKR motif (AA)
Streptococcus pyogenes	NGG	RKR
Streptococcus mutans	NGG	RKR
Streptococcus thermophilus A	NGGNG (SEQ ID NO: 90)	RYR
Treponema denticola	NAAAAN (SEQ ID NO: 96)	VAK
Streptococcus thermophilus B	NNAAAAW (SEQ ID NO: 97)	IYK
Campylobacter jejuni	NNNNACA (SEQ ID NO: 98)	NLK
Pasteurella multocida	GNNNCNNA (SEQ ID NO: 99)	KDG
Neisseria meningitidis	NNNNGATT (SEQ ID NO: 94) or	IGK
Staphylococcus aureus	NNGRRV (R = A or G; V = A. G or C) (SEQ ID NO: 93) NNGRRT (R = A or G) (SEQ ID NO: 95)	NDK

PI domains are provided in **Tables 14** and **15**.

Table 14. Altered PI Domains

Strain Name	PI Start (AA pos)	PI Stop (AA pos)	Length of PI (AA)	RKR motif (AA)
	Start and Stop numbers refer to the sequences in Table 100			
Alicycliphilus denitrificans K601	837	1029	193	--Y
Campylobacter jejuni NCTC 11168	741	984	244	-NG
Helicobacter mustelae 12198	771	1024	254	-NQ

Table 15. Other Altered PI Domains

Strain Name	PI Start (AA pos)	PI Stop (AA pos)	Length of PI (AA)	RKR motif (AA)
	Start and Stop numbers refer to the sequences in Table 11			
Akkermansia muciniphila ATCC BAA-835	871	1101	231	ALK
Ralstonia syzygii R24	821	1062	242	APY
Cand. Puniceispirillum marinum IMCC1322	815	1035	221	AYK
Fructobacillus fructosus KCTC 3544	1074	1323	250	DGN
Eubacterium yurii ATCC 43715	1107	1391	285	DGY
Eubacterium dolichum DSM 3991	779	1096	318	DKK
Dinoroseobacter shibae DFL 12	851	1079	229	DPI
Clostridium cellulolyticum H10	767	1021	255	EGK
Pasteurella multocida str. Pm70	815	1056	242	ENN
Mycoplasma canis PG 14	907	1233	327	EPK
Porphyromonas sp. oral taxon 279 str. F0450	935	1197	263	EPT
Filifactor alocis ATCC 35896	1094	1365	272	EVD
Aminomonas paucivorans DSM 12260	801	1052	252	EVY
Wolinella succinogenes DSM 1740	1034	1409	376	EYK
Oenococcus kitaharae DSM 17330	1119	1389	271	GAL
CoriobacteriumglomeransPW2	1126	1384	259	GDR
Peptoniphilus duerdenii ATCC BAA-1640	1091	1364	274	GDS
Bifidobacterium bifidum S17	1138	1420	283	GGL
Alicyclobacillus hesperidum URH17-3-68	876	1146	271	GGR
Roseburia inulinivorans DSM 16841	895	1152	258	GGT
Actinomyces coleocanis DSM 15436	843	1105	263	GKK
Odoribacter laneus YIT 12061	1103	1498	396	GKV
Coprococcus catus GD-7	1063	1338	276	GNQ
Enterococcus faecalis TX0012	829	1150	322	GRK
Bacillus smithii 7 3 47FAA	809	1088	280	GSK
Legionella pneumophila str. Paris	1021	1372	352	GTM
Bacteroides fragilis NCTC 9343	1140	1436	297	IPV
Mycoplasma ovipneumoniae SC01	923	1265	343	IRI
Actinomyces sp. oral taxon 180 str. F0310	895	1181	287	KEK
Treponema sp. JC4	832	1062	231	KIS
Fusobacteriumnucleatum ATCC49256	1073	1374	302	KKV
Lactobacillus farciminis KCTC 3681	1101	1356	256	KKV
Nitratifractor salsuginis DSM 16511	840	1132	293	KMR
Lactobacillus coryniformis KCTC 3535	850	1119	270	KNK
Mycoplasma mobile 163K	916	1236	321	KNY

Flavobacterium branchiophilum FL-15	1182	1473	292	KQK
Prevotella timonensis CRIS 5C-B1	957	1218	262	KQQ
Methylosinus trichosporium OB3b	830	1082	253	KRP
Prevotella sp. C561	1099	1424	326	KRY
Mycoplasma gallisepticum str. F	911	1269	359	KTA
Lactobacillus rhamnosus GG	1077	1363	287	KYG
Wolinella succinogenes DSM 1740	811	1059	249	LPN
Streptococcus thermophilus LMD-9	1099	1388	290	MLA
Treponema denticola ATCC 35405	1092	1395	304	NDS
Bergeyella zoohelcum ATCC 43767	1098	1415	318	NEK
Veillonella atypica ACS-134-V-Col7a	1107	1398	292	NGF
Neisseria meningitidis Z2491	835	1082	248	NHN
Ignavibacterium album JCM 16511	1296	1688	393	NKK
Ruminococcus albus 8	853	1156	304	NNF
Streptococcus thermophilus LMD-9	811	1121	311	NNK
Barnesiella intestinihominis YIT 11860	871	1153	283	NPV
Azospirillum sp. B510	911	1168	258	PFH
Rhodospirillum rubrum ATCC 11170	863	1173	311	PRG
Planococcus antarcticus DSM 14505	1087	1333	247	PYY
Staphylococcus pseudintermedius ED99	1073	1334	262	QIV
Alcanivorax sp. W11-5	843	1113	271	RIE
Bradyrhizobium sp. BTAi1	811	1064	254	RIY
Streptococcus pyogenes M1 GAS	1099	1368	270	RKR
Streptococcus mutans UA159	1078	1345	268	RKR
Streptococcus Pyogenes	1099	1368	270	RKR
Bacteroides sp. 20 3	1147	1517	371	RNI
S. aureus	772	1053	282	RNK
Solobacterium moorei F0204	1062	1327	266	RSG
Finegoldia magna ATCC 29328	1081	1348	268	RTE
uncultured delta proteobacterium HF0070 07E19	770	1011	242	SGG
Acidaminococcus sp. D21	1064	1358	295	SIG
Eubacterium rectale ATCC 33656	824	1114	291	SKK
Caenispirillum salinarum AK4	1048	1442	395	SLV
Acidothermus cellulolyticus 11B	830	1138	309	SPS
Catenibacterium mitsuokai DSM 15897	1068	1329	262	SPT
Parvibaculum lavamentivorans DS-1	827	1037	211	TGN
Staphylococcus lugdunensis M23590	772	1054	283	TKK
Streptococcus sanguinis SK49	1123	1421	299	TRM
Elusimicrobium minutum Pei191	910	1195	286	TTG

Nitrobacter hamburgensis X14	914	1166	253	VAY
Mycoplasma synoviae 53	991	1314	324	VGF
Sphaerochaeta globus str. Buddy	877	1179	303	VKG
Ilyobacter polytropus DSM 2926	837	1092	256	VNG
Rhodovulum sp. PH10	821	1059	239	VPY
Bifidobacterium longum DJO10A	904	1187	284	VRK

Amino acid sequences described in Table 11 (in order of appearance):

SEQ ID NO: 304

MKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRI
QRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDT
GNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQ
LDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLY
NALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGK
PEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQIS
NLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSP
VVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTT
GKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVK
QEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKD
FINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAED
ALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKD
YKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHH
DPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDD
YPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQA
EFIASFYNNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKT
QSIKKYSTDILGNLYEVKSKKHPQIIKKG

SEQ ID NO: 305

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRL
KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY
HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTY
NQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF
DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS
MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRI
PYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHS
LLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
SVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA
HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQ
TTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINR
LSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK

SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKG
KSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS
AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRV
ILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD
ATLIHQSITGLYETRIDLSQLGGD

SEQ ID NO: 306
MARILAFDIGISSIGWAFSENDELKDCGVRIFTKVENPKTGESLALPRRLARSARKRLARRKAR
LNHLKHLIANEFKLNYEDYQSFDESLAKAYKGSLISPYELRFRALNELLSKQDFARVILHIAKR
RGYDDIKNSDDKEKGAILKAIKQNEEKLANYQSVGEYLYKEYFQKFKENSKEFTNVRNKKESYE
RCIAQSFLKDELKLIFKKQREFGFSFSKKFEEEVLSVAFYKRALKDFSHLVGNCSFFTDEKRAP
KNSPLAFMFVALTRIINLLNNLKNTEGILYTKDDLNALLNEVLKNGTLTYKQTKKLLGLSDDYE
FKGEKGTYFIEFKKYKEFIKALGEHNLSQDDLNEIAKDITLIKDEIKLKKALAKYDLNQNQIDS
LSKLEFKDHLNISFKALKLVTPLMLEGKKYDEACNELNLKVAINEDKKDFLPAFNETYYKDEVT
NPVVLRAIKEYRKVLNALLKKYGKVHKINIELAREVGKNHSQRAKIEKEQNENYKAKKDAELEC
EKLGLKINSKNILKLRLFKEQKEFCAYSGEKIKISDLQDEKMLEIDHIYPYSRSFDDSYMNKVL
VFTKQNQEKLNQTPFEAFGNDSAKWQKIEVLAKNLPTKKQKRILDKNYKDKEQKNFKDRNLNDT
RYIARLVLNYTKDYLDFLPLSDDENTKLNDTQKGSKVHVEAKSGMLTSALRHTWGFSAKDRNNH
LHHAIDAVIIAYANNSIVKAFSDFKKEQESNSAELYAKKISELDYKNKRKFFEPFSGFRQKVLD
KIDEIFVSKPERKKPSGALHEETFRKEEEFYQSYGGKEGVLKALELGKIRKVNGKIVKNGDMFR
VDIFKHKKTNKFYAVPIYTMDFALKVLPNKAVARSKKGEIKDWILMDENYEFCFSLYKDSLILI
QTKDMQEPEFVYYNAFTSSTVSLIVSKHDNKFETLSKNQKILFKNANEKEVIAKSIGIQNLKVF
EKYIVSALGEVTKAEFRQREDFKK

SEQ ID NO: 307
MKRILGLDLGTNSIGWALVNEAENKDERSSIVKLGVRVNPLTVDELTNFEKGKSITTNADRTLK
RGMRRNLQRYKLRRETLTEVLKEHKLITEDTILSENGNRTTFETYRLRAKAVTEEISLEEFARV
LLMINKKRGYKSSRKAKGVEEGTLIDGMDIARELYNNNLTPGELCLQLLDAGKKFLPDFYRSDL
QNELDRIWEKQKEYYPEILTDVLKEELRGKKRDAVWAICAKYFVWKENYTEWNKEKGKTEQQER
EHKLEGIYSKRKRDEAKRENLQWRVNGLKEKLSLEQLVIVFQEMNTQINNSSGYLGAISDRSKE
LYFNKQTVGQYQMEMLDKNPNASLRNMVFYRQDYLDEFNMLWEKQAVYHKELTEELKKEIRDII
IFYQRRLKSQKGLIGFCEFESRQIEVDIDGKKKIKTVGNRVISRSSPLFQEFKIWQILNNIEVT
VVGKKRKRRKLKENYSALFEELNDAEQLELNGSRRLCQEEKELLAQELFIRDKMTKSEVLKLLF
DNPQELDLNFKTIDGNKTGYALFQAYSKMIEMSGHEPVDFKKPVEKVVEYIKAVFDLLNWNTDI
LGFNSNEELDNQPYYKLWHLLYSFEGDNTPTGNGRLIQKMTELYGFEKEYATILANVSFQDDYG
SLSAKAIHKILPHLKEGNRYDVACVYAGYRHSESSLTREEIANKVLKDRLMLLPKNSLHNPVVE
KILNQMVNVINVIIDIYGKPDEIRVELARELKKNAKEREELTKSIAQTTKAHEEYKTLLQTEFG
LTNVSRTDILRYKLYKELESCGYKTLYSNTYISREKLFSKEFDIEHIIPQARLFDDSFSNKTLE
ARSVNIEKGNKTAYDFVKEKFGESGADNSLEHYLNNIEDLFKSGKISKTKYNKLKMAEQDIPDG
FIERDLRNTQYIAKKALSMLNEISHRVVATSGSVTDKLREDWQLIDVMKELNWEKYKALGLVEY
FEDRDGRQIGRIKDWTKRNDHRHHAMDALTVAFTKDVFIQYFNNKNASLDPNANEHAIKNKYFQ
NGRAIAPMPLREFRAEAKKHLENTLISIKAKNKVITGNINKTRKKGGVNKNMQQTPRGQLHLET
IYGSGKQYLTKEEKVNASFDMRKIGTVSKSAYRDALLKRLYENDNDPKKAFAGKNSLDKQPIWL
DKEQMRKVPEKVKIVTLEAIYTIRKEISPDLKVDKVIDVGVRKILIDRLNEYGNDAKKAFSNLD
KNPIWLNKEKGISIKRVTISGISNAQSLHVKKDKDGKPILDENGRNIPVDFVNTGNNHHVAVYY
RPVIDKRGQLVVDEAGNPKYELEEVVVSFFEAVTRANLGLPIIDKDYKTTEGWQFLFSMKQNEY

FVFPNEKTGFNPKEIDLLDVENYGLISPNLFRVQKFSLKNYVFRHHLETTIKDTSSILRGITWI
DFRSSKGLDTIVKVRVNHIGQIVSVGEY

SEQ ID NO: 308

MSRKNYVDDYAISLDIGNASVGWSAFTPNYRLVRAKGHELIGVRLFDPADTAESRRMARTTRRR
YSRRRWRLRLLDALFDQALSEIDPSFLARRKYSWVHPDDENNADCWYGSVLFDSNEQDKRFYEK
YPTIYHLRKALMEDDSQHDIREIYLAIHHMVKYRGNFLVEGTLESSNAFKEDELLKLLGRITRY
EMSEGEQNSDIEQDDENKLVAPANGQLADALCATRGSRSMRVDNALEALSAVNDLSREQRAIVK
AIFAGLEGNKLDLAKIFVSKEFSSENKKILGIYFNKSDYEEKCVQIVDSGLLDDEEREFLDRMQ
GQYNAIALKQLLGRSTSVSDSKCASYDAHRANWNLIKLQLRTKENEKDINENYGILVGWKIDSG
QRKSVRGESAYENMRKKANVFFKKMIETSDLSETDKNRLIHDIEEDKLFPIQRDSDNGVIPHQL
HQNELKQIIKKQGKYYPFLLDAFEKDGKQINKIEGLLTFRVPYFVGPLVVPEDLQKSDNSENHW
MVRKKKGEITPWNFDEMVDKDASGRKFIERLVGTDSYLLGEPTLPKNSLLYQEYEVLNELNNVR
LSVRTGNHWNDKRRMRLGREEKTLLCQRLFMKGQTVTKRTAENLLRKEYGRTYELSGLSDESKF
TSSLSTYGKMCRIFGEKYVNEHRDLMEKIVELQTVFEDKETLLHQLRQLEGISEADCALLVNTH
YTGWGRLSRKLLTTKAGECKISDDFAPRKHSIIEIMRAEDRNLMEIITDKQLGFSDWIEQENLG
AENGSSLMEVVDDLRVSPKVKRGIIQSIRLIDDISKAVGKRPSRIFLELADDIQPSGRTISRKS
RLQDLYRNANLGKEFKGIADELNACSDKDLQDDRLFLYYTQLGKDMYTGEELDLDRLSSAYDID
HIIPQAVTQNDSIDNRVLVARAENARKTDSFTYMPQIADRMRNFWQILLDNGLISRVKFERLTR
QNEFSEREKERFVQRSLVETRQIMKNVATLMRQRYGNSAAVIGLNAELTKEMHRYLGFSHKNRD
INDYHHAQDALCVGIAGQFAANRGFFADGEVSDGAQNSYNQYLRDYLRGYREKLSAEDRKQGRA
FGFIVGSMRSQDEQKRVNPRTGEVVWSEEDKDYLRKVMNYRKMLVTQKVGDDFGALYDETRYAA
TDPKGIKGIPFDGAKQDTSLYGGFSSAKPAYAVLIESKGKTRLVNVTMQEYSLLGDRPSDDELR
KVLAKKKSEYAKANILLRHVPKMQLIRYGGGLMVIKSAGELNNAQQLWLPYEEYCYFDDLSQGK
GSLEKDDLKKLLDSILGSVQCLYPWHRFTEEELADLHVAFDKLPEDEKKNVITGIVSALHADAK
TANLSIVGMTGSWRRMNNKSGYTFSDEDEFIFQSPSGLFEKRVTVGELKRKAKKEVNSKYRTNE
KRLPTLSGASQP

SEQ ID NO: 309

METQTSNQLITSHLKDYPKQDYFVGLDIGTNSVGWAVTNTSYELLKFHSHKMWGSRLFEEGESA
VTRRGFRSMRRRLERRKLRLKLLEELFADAMAQVDSTFFIRLHESKYHYEDKTTGHSSKHILFI
DEDYTDQDYFTEYPTIYHLRKDLMENGTDDIRKLFLAVHHILKYRGNFLYEGATFNSNAFTFED
VLKQALVNITFNCFDTNSAISSISNILMESGKTKSDKAKAIERLVDTYTVFDEVNTPDKPQKEQ
VKEDKKTLKAFANLVLGLSANLIDLFGSVEDIDDDLKKLQIVGDTYDEKRDELAKVWGDEIHII
DDCKSVYDAIILMSIKEPGLTISQSKVKAFDKHKEDLVILKSLLKLDRNVYNEMFKSDKKGLHN
YVHYIKQGRTEETSCSREDFYKYTKKIVEGLADSKDKEYILNEIELQTLLPLQRIKDNGVIPYQ
LHLEELKVILDKCGPKFPFLHTVSDGFSVTEKLIKMLEFRIPYYVGPLNTHHNIDNGGFSWAVR
KQAGRVTPWNFEEKIDREKSAAAFIKNLTNKCTYLFGEDVLPKSSLLYSEFMLLNELNNVRIDG
KALAQGVKQHLIDSIFKQDHKKMTKNRIELFLKDNNYITKKHKPEITGLDGEIKNDLTSYRDMV
RILGNNFDVSMAEDIITDITIFGESKKMLRQTLRNKFGSQLNDETIKKLSKLRYRDWGRLSKKL
LKGIDGCDKAGNGAPKTIIELMRNDSYNLMEILGDKFSFMECIEEENAKLAQGQVVNPHDIIDE
LALSPAVKRAVWQALRIVDEVAHIKKALPSRIFVEVARTNKSEKKKKDSRQKRLSDLYSAIKKD
DVLQSGLQDKEFGALKSGLANYDDAALRSKKLYLYYTQMGRCAYTGNIIDLNQLNTDNYDIDHI
YPRSLTKDDSFDNLVLCERTANAKKSDIYPIDNRIQTKQKPFWAFLKHQGLISERKYERLTRIA
PLTADDLSGFIARQLVETNQSVKATTTLLRRLYPDIDVVFVKAENVSDFRHNNNFIKVRSLNHH
HHAKDAYLNIVVGNVYHEKFTRNFRLFFKKNGANRTYNLAKMFNYDVICTNAQDGKAWDVKTSM
NTVKKMMASNDVRVTRRLLEQSGALADATIYKASVAAKAKDGAYIGMKTKYSVFADVTKYGGMT

KIKNAYSIIVQYTGKKGEEIKEIVPLPIYLINRNATDIELIDYVKSVIPKAKDISIKYRKLCIN
QLVKVNGFYYYLGGKTNDKIYIDNAIELVVPHDIATYIKLLDKYDLLRKENKTLKASSITTSIY
NINTSTVVSLNKVGIDVFDYFMSKLRTPLYMKMKGNKVDELSSTGRSKFIKMTLEEQSIYLLEV
LNLLTNSKTTFDVKPLGITGSRSTIGVKIHNLDEFKIINESITGLYSNEVTIV

SEQ ID NO: 310
MTKLNQPYGIGLDIGSNSIGFAVVDANSHLLRLKGETAIGARLFREGQSAADRRGSRTTRRRLS
RTRWRLSFLRDFFAPHITKIDPDFFLRQKYSEISPKDKDRFKYEKRLFNDRTDAEFYEDYPSMY
HLRLHLMTHTHKADPREIFLAIHHILKSRGHFLTPGAAKDFNTDKVDLEDIFPALTEAYAQVYP
DLELTFDLAKADDFKAKLLDEQATPSDTQKALVNLLLSSDGEKEIVKKRKQVLTEFAKAITGLK
TKFNLALGTEVDEADASNWQFSMGQLDDKWSNIETSMTDQGTEIFEQIQELYRARLLNGIVPAG
MSLSQAKVADYGQHKEDLELFKTYLKKLNDHELAKTIRGLYDRYINGDDAKPFLREDFVKALTK
EVTAHPNEVSEQLLNRMGQANFMLKQRTKANGAIPIQLQQRELDQIIANQSKYYDWLAAPNPVE
AHRWKMPYQLDELLNFHIPYYVGPLITPKQQAESGENVFAWMVRKDPSGNITPYNFDEKVDREA
SANTFIQRMKTTDTYLIGEDVLPKQSLLYQKYEVLNELNNVRINNECLGTDQKQRLIREVFERH
SSVTIKQVADNLVAHGDFARRPEIRGLADEKRFLSSLSTYHQLKEILHEAIDDPTKLLDIENII
TWSTVFEDHTIFETKLAEIEWLDPKKINELSGIRYRGWGQFSRKLLDGLKLGNGHTVIQELMLS
NHNLMQILADETLKETMTELNQDKLKTDDIEDVINDAYTSPSNKKALRQVLRVVEDIKHAANGQ
DPSWLFIETADGTGTAGKRTQSRQKQIQTVYANAAQELIDSAVRGELEDKIADKASFTDRLVLY
FMQGGRDIYTGAPLNIDQLSHYDIDHILPQSLIKDDSLDNRVLVNATINREKNNVFASTLFAGK
MKATWRKWHEAGLISGRKLRNLMLRPDEIDKFAKGFVARQLVETRQIIKLTEQIAAAQYPNTKI
IAVKAGLSHQLREELDFPKNRDVNHYHHAFDAFLAARIGTYLLKRYPKLAPFFTYGEFAKVDVK
KFREFNFIGALTHAKKNIIAKDTGEIVWDKERDIRELDRIYNFKRMLITHEVYFETADLFKQTI
YAAKDSKERGGSKQLIPKKQGYPTQVYGGYTQESGSYNALVRVAEADTTAYQVIKISAQNASKI
ASANLKSREKGKQLLNEIVVKQLAKRRKNWKPSANSFKIVIPRFGMGTLFQNAKYGLFMVNSDT
YYRNYQELWLSRENQKLLKKLFSIKYEKTQMNHDALQVYKAIIDQVEKFFKLYDINQFRAKLSD
AIERFEKLPINTDGNKIGKTETLRQILIGLQANGTRSNVKNLGIKTDLGLLQVGSGIKLDKDTQ
IVYQSPSGLFKRRIPLADL

SEQ ID NO: 311
MTKEYYLGLDVGTNSVGWAVTDSQYNLCKFKKKDMWGIRLFESANTAKDRRLQRGNRRRLERKK
QRIDLLQEIFSPEICKIDPTFFIRLNESRLHLEDKSNDFKYPLFIEKDYSDIEYYKEFPTIFHL
RKHLIESEEKQDIRLIYLALHNIIKTRGHFLIDGDLQSAKQLRPILDTFLLSLQEEQNLSVSLS
ENQKDEYEEILKNRSIAKSEKVKKLKNLFEISDELEKEEKKAQSAVIENFCKFIVGNKGDVCKF
LRVSKEELEIDSFSFSEGKYEDDIVKNLEEKVPEKVYLFEQMKAMYDWNILVDILETEEYISFA
KVKQYEKHKTNLRLLRDIILKYCTKDEYNRMFNDEKEAGSYTAYVGKLKKNNKKYWIEKKRNPE
EFYKSLGKLLDKIEPLKEDLEVLTMMIEECKNHTLLPIQKNKDNGVIPHQVHEVELKKILENAK
KYYSFLTETDKDGYSVVQKIESIFRFRIPYYVGPLSTRHQEKGSNVWMVRKPGREDRIYPWNME
EIIDFEKSNENFITRMTNKCTYLIGEDVLPKHSLLYSKYMVLNELNNVKVRGKKLPTSLKQKVF
EDLFENKSKVTGKNLLEYLQIQDKDIQIDDLSGFDKDFKTSLKSYLDFKKQIFGEEIEKESIQN
MIEDIIKWITIYGNDKEMLKRVIRANYSNQLTEEQMKKITGFQYSGWGNFSKMFLKGISGSDVS
TGETFDIITAMWETDNNLMQILSKKFTFMDNVEDFNSGKVGKIDKITYDSTVKEMFLSPENKRA
VWQTIQVAEEIKKVMGCEPKKIFIEMARGGEKVKKRTKSRKAQLLELYAACEEDCRELIKEIED
RDERDFNSMKLFLYYTQFGKCMYSGDDIDINELIRGNSKWDRDHIYPQSKIKDDSIDNLVLVNK
TYNAKKSNELLSEDIQKKMHSFWLSLLNKKLITKSKYDRLTRKGDFTDEELSGFIARQLVETRQ
STKAIADIFKQIYSSEVVYVKSSLVSDFRKKPLNYLKSRRVNDYHHAKDAYLNIVVGNVYNKKF
TSNPIQWMKKNRDTNYSLNKVFEHDVVINGEVIWEKCTYHEDTNTYDGGTLDRIRKIVERDNIL

YTEYAYCEKGELFNATIQNKNGNSTVSLKKGLDVKKYGGYFSANTSYFSLIEFEDKKGDRARHI
IGVPIYIANMLEHSPSAFLEYCEQKGYQNVRILVEKIKKNSLLIINGYPLRIRGENEVDTSFKR
AIQLKLDQKNYELVRNIEKFLEKYVEKKGNYPIDENRDHITHEKMNQLYEVLLSKMKKFNKKGM
ADPSDRIEKSKPKFIKLEDLIDKINVINKMLNLLRCDNDTKADLSLIELPKNAGSFVVKKNTIG
KSKIILVNQSVTGLYENRREL

SEQ ID NO: 312
MARDYSVGLDIGTSSVGWAAIDNKYHLIRAKSKNLIGVRLFDSAVTAEKRRGYRTTRRRLSRRH
WRLRLLNDIFAGPLTDFGDENFLARLKYSWVHPQDQSNQAHFAAGLLFDSKEQDKDFYRKYPTI
YHLRLALMNDDQKHDLREVYLAIHHLVKYRGHFLIEGDVKADSAFDVHTFADAIQRYAESNNSD
ENLLGKIDEKKLSAALTDKHGSKSQRAETAETAFDILDLQSKKQIQAILKSVVGNQANLMAIFG
LDSSAISKDEQKNYKFSFDDADIDEKIADSEALLSDTEFEFLCDLKAAFDGLTLKMLLGDDKTV
SAAMVRRFNEHQKDWEYIKSHIRNAKNAGNGLYEKSKKFDGINAAYLALQSDNEDDRKKAKKIF
QDEISSADIPDDVKADFLKKIDDDQFLPIQRTKNNGTIPHQLHRNELEQIIEKQGIYYPFLKDT
YQENSHELNKITALINFRVPYYVGPLVEEEQKIADDGKNIPDPTNHWMVRKSNDTITPWNLSQV
VDLDKSGRRFIERLTGTDTYLIGEPTLPKNSLLYQKFDVLQELNNIRVSGRRLDIRAKQDAFEH
LFKVQKTVSATNLKDFLVQAGYISEDTQIEGLADVNGKNFNNALTTYNYLVSVLGREFVENPSN
EELLEEITELQTVFEDKKVLRRQLDQLDGLSDHNREKLSRKHYTGWGRISKKLLTTKIVQNADK
IDNQTFDVPRMNQSIIDTLYNTKMNLMEIINNAEDDFGVRAWIDKQNTTDGDEQDVYSLIDELA
GPKEIKRGIVQSFRILDDITKAVGYAPKRVYLEFARKTQESHLTNSRKNQLSTLLKNAGLSELV
TQVSQYDAAALQNDRLYLYFLQQGKDMYSGEKLNLDNLSNYDIDHIIPQAYTKDNSLDNRVLVS
NITNRRKSDSSNYLPALIDKMRPFWSVLSKQGLLSKHKFANLTRTRDFDDMEKERFIARSLVET
RQIIKNVASLIDSHFGGETKAVAIRSSLTADMRRYVDIPKNRDINDYHHAFDALLFSTVGQYTE
NSGLMKKGQLSDSAGNQYNRYIKEWIHAARLNAQSQRVNPFGFVVGSMRNAAPGKLNPETGEIT
PEENADWSIADLDYLHKVMNFRKITVTRRLKDQKGQLYDESRYPSVLHDAKSKASINFDKHKPV
DLYGGFSSAKPAYAALIKFKNKFRLVNVLRQWTYSDKNSEDYILEQIRGKYPKAEMVLSHIPYG
QLVKKDGALVTISSATELHNFEQLWLPLADYKLINTLLKTKEDNLVDILHNRLDLPEMTIESAF
YKAFDSILSFAFNRYALHQNALVKLQAHRDDFNALNYEDKQQTLERILDALHASPASSDLKKIN
LSSGFGRLFSPSHFTLADTDEFIFQSVTGLFSTQKTVAQLYQETK

SEQ ID NO: 313
MVYDVGLDIGTGSVGWVALDENGKLARAKGKNLVGVRLFDTAQTAADRRGFRTTRRRLSRRKWR
LRLLDELFSAEINEIDSSFFQRLKYSYVHPKDEENKAHYYGGYLFPTEEETKKFHRSYPTIYHL
RQELMAQPNKRFDIREIYLAIHHLVKYRGHFLSSQEKITIGSTYNPEDLANAIEVYADEKGLSW
ELNNPEQLTEIISGEAGYGLNKSMKADEALKLFEFDNNQDKVAIKTLLAGLTGNQIDFAKLFGK
DISDKDEAKLWKLKLDDEALEEKSQTILSQLTDEEIELFHAVVQAYDGFVLIGLLNGADSVSAA
MVQLYDQHREDRKLLKSLAQKAGLKHKRFSEIYEQLALATDEATIKNGISTARELVEESNLSKE
VKEDTLRRLDENEFLPKQRTKANSVIPHQLHLAELQKILQNQGQYYPFLLDTFEKEDGQDNKIE
ELLRFRIPYYVGPLVTKKDVEHAGGDADNHWVERNEGFEKSRVTPWNFDKVFNRDKAARDFIER
LTGNDTYLIGEKTLPQNSLRYQLFTVLNELNNVRVNGKKFDSKTKADLINDLFKARKTVSLSAL
KDYLKAQGKGDVTITGLADESKFNSSLSSYNDLKKTFDAEYLENEDNQETLEKIIEIQTVFEDS
KIASRELSKLPLDDDQVKKLSQTHYTGWGRLSEKLLDSKIIDERGQKVSILDKLKSTSQNFMSI
INNDKYGVQAWITEQNTGSSKLTFDEKVNELTTSPANKRGIKQSFAVLNDIKKAMKEEPRRVYL
EFAREDQTSVRSVPRYNQLKEKYQSKSLSEEAKVLKKTLDGNKNKMSDDRYFLYFQQQGKDMYT
GRPINFERLSQDYDIDHIIPQAFTKDDSLDNRVLVSRPENARKSDSFAYTDEVQKQDGSLWTSL
LKSGFINRKKYERLTKAGKYLDGQKTGFIARQLVETRQIIKNVASLIEGEYENSKAVAIRSEIT
ADMRLLVGIKKHREINSFHHAFDALLITAAGQYMQNRYPDRDSTNVYNEFDRYTNDYLKNLRQL

SSRDEVRRLKSFGFVVGTMRKGNEDWSEENTSYLRKVMMFKNILTTKKTEKDRGPLNKETIFSP
KSGKKLIPLNSKRSDTALYGGYSNVYSAYMTLVRANGKNLLIKIPISIANQIEVGNLKINDYIV
NNPAIKKFEKILISKLPLGQLVNEDGNLIYLASNEYRHNAKQLWLSTTDADKIASISENSSDEE
LLEAYDILTSENVKNRFPFFKKDIDKLSQVRDEFLDSDKRIAVIQTILRGLQIDAAYQAPVKII
SKKVSDWHKLQQSGGIKLSDNSEMIYQSATGIFETRVKISDLL

SEQ ID NO: 314
IVDYCIGLDLGTGSVGWAVVDMNHRLMKRNGKHLWGSRLFSNAETAANRRASRSIRRRYNKRRE
RIRLLRAILQDMVLEKDPTFFIRLEHTSFLDEEDKAKYLGTDYKDNYNLFIDEDFNDYTYYHKY
PTIYHLRKALCESTEKADPRLIYLALHHIVKYRGNFLYEGQKFNMDASNIEDKLSDIFTQFTSF
NNIPYEDDEKKNLEILEILKKPLSKKAKVDEVMTLIAPEKDYKSAFKELVTGIAGNKMNVTKMI
LCEPIKQGDSEIKLKFSDSNYDDQFSEVEKDLGEYVEFVDALHNVYSWVELQTIMGATHTDNAS
ISEAMVSRYNKHHDDLKLLKDCIKNNVPNKYFDMFRNDSEKSKGYYNYINRPSKAPVDEFYKYV
KKCIEKVDTPEAKQILNDIELENFLLKQNSRTNGSVPYQMQLDEMIKIIDNQAEYYPILKEKRE
QLLSILTFRIPYYFGPLNETSEHAWIKRLEGKENQRILPWNYQDIVDVDATAEGFIKRMRSYCT
YFPDEEVLPKNSLIVSKYEVYNELNKIRVDDKLLEVDVKNDIYNELFMKNKTVTEKKLKNWLVN
NQCCSKDAEIKGFQKENQFSTSLTPWIDFTNIFGKIDQSNFDLIENIIYDLTVFEDKKIMKRRL
KKKYALPDDKVKQILKLKYKDWSRLSKKLLDGIVADNRFGSSVTVLDVLEMSRLNLMEIINDKD
LGYAQMIEEATSCPEDGKFTYEEVERLAGSPALKRGIWQSLQIVEEITKVMKCRPKYIYIEFER
SEEAKERTESKIKKLENVYKDLDEQTKKEYKSVLEELKGFDNTKKISSDSLFLYFTQLGKCMYS
GKKLDIDSLDKYQIDHIVPQSLVKDDSFDNRVLVVPSENQRKLDDLVVPFDIRDKMYRFWKLLF
DHELISPKKFYSLIKTEYTERDEERFINRQLVETRQITKNVTQIIEDHYSTTKVAAIRANLSHE
FRVKNHIYKNRDINDYHHAHDAYIVALIGGFMRDRYPNMHDSKAVYSEYMKMFRKNKNDQKRWK
DGFVINSMNYPYEVDGKLIWNPDLINEIKKCFYYKDCYCTTKLDQKSGQLFNLTVLSNDAHADK
GVTKAVVPVNKNRSDVHKYGGFSGLQYTIVAIEGQKKKGKKTELVKKISGVPLHLKAASINEKI
NYIEEKEGLSDVRIIKDNIPVNQMIEMDGGEYLLTSPTEYVNARQLVLNEKQCALIADIYNAIY
KQDYDNLDDILMIQLYIELTNKMKVLYPAYRGIAEKFESMNENYVVISKEEKANIIKQMLIVMH
RGPQNGNIVYDDFKISDRIGRLKTKNHNLNNIVFISQSPTGIYTKKYKL

SEQ ID NO: 315
MKSEKKYYIGLDVGTNSVGWAVTDEFYNILRAKGKDLWGVRLFEKADTAANTRIFRSGRRRNDR
KGMRLQILREIFEDEIKKVDKDFYDRLDESKFWAEDKKVSGKYSLFNDKNFSDKQYFEKFPTIF
HLRKYLMEEHGKVDIRYYFLAINQMMKRRGHFLIDGQISHVTDDKPLKEQLILLINDLLKIELE
EELMDSIFEILADVNEKRTDKKNNLKELIKGQDFNKQEGNILNSIFESIVTGKAKIKNIISDED
ILEKIKEDNKEDFVLTGDSYEENLQYFEEVLQENITLFNTLKSTYDFLILQSILKGKSTLSDAQ
VERYDEHKKDLEILKKVIKKYDEDGKLFKQVFKEDNGNGYVSYIGYYLNKNKKITAKKKISNIE
FTKYVKGILEKQCDCEDEDVKYLLGKIEQENFLLKQISSINSVIPHQIHLFELDKILENLAKNY
PSFNNKKEEFTKIEKIRKTFTFRIPYYVGPLNDYHKNNGGNAWIFRNKGEKIRPWNFEKIVDLH
KSEEEFIKRMLNQCTYLPEETVLPKSSILYSEYMVLNELNNLRINGKPLDTDVKLKLIEELFKK
KTKVTLKSIRDYMVRNNFADKEDFDNSEKNLEIASNMKSYIDFNNILEDKFDVEMVEDLIEKIT
IHTGNKKLLKKYIEETYPDLSSSQIQKIINLKYKDWGRLSRKLLDGIKGTKKETEKTDTVINFL
RNSSDNLMQIIGSQNYSFNEYIDKLRKKYIPQEISYEVVENLYVSPSVKKMIWQVIRVTEEITK
VMGYDPDKIFIEMAKSEEEKKTTISRKNKLLDLYKAIKKDERDSQYEKLLTGLNKLDDSDLRSR
KLYLYYTQMGRDMYTGEKIDLDKLFDSTHYDKDHIIPQSMKKDDSIINNLVLVNKNANQTTKGN
IYPVPSSIRNNPKIYNYWKYLMEKEFISKEKYNRLIRNTPLTNEELGGFINRQLVETRQSTKAI
KELFEKFYQKSKIIPVKASLASDLRKDMNTLKSREVNDLHHAHDAFLNIVAGDVWNREFTSNPI
NYVKENREGDKVKYSLSKDFTRPRKSKGKVIWTPEKGRKLIVDTLNKPSVLISNESHVKKGELF

NATIAGKKDYKKGKIYLPLKKDDRLQDVSKYGGYKAINGAFFFLVEHTKSKKRIRSIELFPLHL
LSKFYEDKNTVLDYAINVLQLQDPKIIIDKINYRTEIIIDNFSYLISTKSNDGSITVKPNEQMY
WRVDEISNLKKIENKYKKDAILTEEDRKIMESYIDKIYQQFKAGKYKNRRTTDTIIEKYEIIDL
DTLDNKQLYQLLVAFISLSYKTSNNAVDFTVIGLGTECGKPRITNLPDNTYLVYKSITGIYEKR
IRIK

SEQ ID NO: 316
MKLRGIEDDYSIGLDMGTSSVGWAVTDERGTLAHFKRKPTWGSRLFREAQTAAVARMPRGQRRR
YVRRRWRLDLLQKLFEQQMEQADPDFFIRLRQSRLLRDDRAEEHADYRWPLFNDCKFTERDYYQ
RFPTIYHVRSWLMETDEQADIRLIYLALHNIVKHRGNFLREGQSLSAKSARPDEALNHLRETLR
VWSSERGFECSIADNGSILAMLTHPDLSPSDRRKKIAPLFDVKSDDAAADKKLGIALAGAVIGL
KTEFKNIFGDFPCEDSSIYLSNDEAVDAVRSACPDDCAELFDRLCEVYSAYVLQGLLSYAPGQT
ISANMVEKYRRYGEDLALLKKLVKIYAPDQYRMFFSGATYPGTGIYDAAQARGYTKYNLGPKKS
EYKPSESMQYDDFRKAVEKLFAKTDARADERYRMMMDRFDKQQFLRRLKTSDNGSIYHQLHLEE
LKAIVENQGRFYPFLKRDADKLVSLVSFRIPYYVGPLSTRNARTDQHGENRFAWSERKPGMQDE
PIFPWNWESIIDRSKSAEKFILRMTGMCTYLQQEPVLPKSSLLYEEFCVLNELNGAHWSIDGDD
EHRFDAADREGIIEELFRRKRTVSYGDVAGWMERERNQIGAHVCGGQGEKGFESKLGSYIFFCK
DVFKVERLEQSDYPMIERIILWNTLFEDRKILSQRLKEEYGSRLSAEQIKTICKKRFTGWGRLS
EKFLTGITVQVDEDSVSIMDVLREGCPVSGKRGRAMVMMEILRDEELGFQKKVDDFNRAFFAEN
AQALGVNELPGSPAVRRSLNQSIRIVDEIASIAGKAPANIFIEVTRDEDPKKKGRRTKRRYNDL
KDALEAFKKEDPELWRELCETAPNDMDERLSLYFMQRGKCLYSGRAIDIHQLSNAGIYEVDHII
PRTYVKDDSLENKALVYREENQRKTDMLLIDPEIRRRMSGYWRMLHEAKLIGDKKFRNLLRSRI
DDKALKGFIARQLVETGQMVKLVRSLLEARYPETNIISVKASISHDLRTAAELVKCREANDFHH
AHDAFLACRVGLFIQKRHPCVYENPIGLSQVVRNYVRQQADIFKRCRTIPGSSGFIVNSFMTSG
FDKETGEIFKDDWDAEAEVEGIRRSLNFRQCFISRMPFEDHGVFWDATIYSPRAKKTAALPLKQ
GLNPSRYGSFSREQFAYFFIYKARNPRKEQTLFEFAQVPVRLSAQIRQDENALERYARELAKDQ
GLEFIRIERSKILKNQLIEIDGDRLCITGKEEVRNACELAFAQDEMRVIRMLVSEKPVSRECVI
SLFNRILLHGDQASRRLSKQLKLALLSEAFSEASDNVQRNVVLGLIAIFNGSTNMVNLSDIGGS
KFAGNVRIKYKKELASPKVNVHLIDQSVTGMFERRTKIGL

SEQ ID NO: 317
MENKQYYIGLDVGTNSVGWAVTDTSYNLLRAKGKDMWGARLFEKANTAAERRTKRTSRRRSERE
KARKAMLKELFADEINRVDPSFFIRLEESKFFLDDRSENNRQRYTLFNDATFTDKDYYEKYKTI
FHLRSALINSDEKFDVRLVFLAILNLFSHRGHFLNASLKGDGDIQGMDVFYNDLVESCEYFEIE
LPRITNIDNFEKILSQKGKSRTKILEELSEELSISKKDKSKYNLIKLISGLEASVVELYNIEDI
QDENKKIKIGFRESDYEESSLKVKEIIGDEYFDLVERAKSVHDMGLLSNIIGNSKYLCEARVEA
YENHHKDLLKIKELLKKYDKKAYNDMFRKMTDKNYSAYVGSVNSNIAKERRSVDKRKIEDLYKY
IEDTALKNIPDDNKDKIEILEKIKLGEFLKKQLTASNGVIPNQLQSRELRAILKKAENYLPFLK
EKGEKNLTVSEMIIQLFEFQIPYYVGPLDKNPKKDNKANSWAKIKQGGRILPWNFEDKVDVKGS
RKEFIEKMVRKCTYISDEHTLPKQSLLYEKFMVLNEINNIKIDGEKISVEAKQKIYNDLFVKGK
KVSQKDIKKELISLNIMDKDSVLSGTDTVCNAYLSSIGKFTGVFKEEINKQSIVDMIEDIIFLK
TVYGDEKRFVKEEIVEKYGDEIDKDKIKRILGFKFSNWGNLSKSFLELEGADVGTGEVRSIIQS
LWETNFNLMELLSSRFTYMDELEKRVKKLEKPLSEWTIEDLDDMYLSSPVKRMIWQSMKIVDEI
QTVIGYAPKRIFVEMTRSEGEKVRTKSRKDRLKELYNGIKEDSKQWVKELDSKDESYFRSKKMY
LYYLQKGRCMYSGEVIELDKLMDDNLYDIDHIYPRSFVKDDSLDNLVLVKKEINNRKQNDPITP
QIQASCQGFWKILHDQGFMSNEKYSRLTRKTQEFSDEEKLSFINRQIVETGQATKCMAQILQKS
MGEDVDVVFSKARLVSEFRHKFELFKSRLINDFHHANDAYLNIVVGNSYFVKFTRNPANFIKDA

RKNPDNPVYKYHMDRFFERDVKSKSEVAWIGQSEGNSGTIVIVKKTMAKNSPLITKKVEEGHGS
ITKETIVGVKEIKFGRNKVEKADKTPKKPNLQAYRPIKTSDERLCNILRYGGRTSISISGYCLV
EYVKKRKTIRSLEAIPVYLGRKDSLSEEKLLNYFRYNLNDGGKDSVSDIRLCLPFISTNSLVKI
DGYLYYLGGKNDDRIQLYNAYQLKMKKEEVEYIRKIEKAVSMSKFDEIDREKNPVLTEEKNIEL
YNKIQDKFENTVFSKRMSLVKYNKKDLSFGDFLKNKKSKFEEIDLEKQCKVLYNIIFNLSNLKE
VDLSDIGGSKSTGKCRCKKNITNYKEFKLIQQSITGLYSCEKDLMTI

SEQ ID NO: 318
MKNLKEYYIGLDIGTASVGWAVTDESYNIPKFNGKKMWGVRLFDDAKTAEERRTQRGSRRRLNR
RKERINLLQDLFATEISKVDPNFFLRLDNSDLYREDKDEKLKSKYTLFNDKDFKDRDYHKKYPT
IHHLIMDLIEDEGKKDIRLLYLACHYLLKNRGHFIFEGQKFDTKNSFDKSINDLKIHLRDEYNI
DLEFNNEDLIEIITDTTLNKTNKKKELKNIVGDTKFLKAISAIMIGSSQKLVDLFEDGEFEETT
VKSVDFSTTAFDDKYSEYEEALGDTISLLNILKSIYDSSILENLLKDADKSKDGNKYISKAFVK
KFNKHGKDLKTLKRIIKKYLPSEYANIFRNKSINDNYVAYTKSNITSNKRTKASKFTKQEDFYK
FIKKHLDTIKETKLNSSENEDLKLIDEMLTDIEFKTFIPKLKSSDNGVIPYQLKLMELKKILDN
QSKYYDFLNESDEYGTVKDKVESIMEFRIPYYVGPLNPDSKYAWIKRENTKITPWNFKDIVDLD
SSREEFIDRLIGRCTYLKEEKVLPKASLIYNEFMVLNELNNLKLNEFLITEEMKKAIFEELFKT
KKKVTLKAVSNLLKKEFNLTGDILLSGTDGDFKQGLNSYIDFKNIIGDKVDRDDYRIKIEEIIK
LIVLYEDDKTYLKKKIKSAYKNDFTDDEIKKIAALNYKDWGRLSKRFLTGIEGVDKTTGEKGSI
IYFMREYNLNLMELMSGHYTFTEEVEKLNPVENRELCYEMVDELYLSPSVKRMLWQSLRVVDEI
KRIIGKDPKKIFIEMARAKEAKNSRKESRKNKLLEFYKFGKKAFINEIGEERYNYLLNEINSEE
ESKFRWDNLYLYYTQLGRCMYSLEPIDLADLKSNNIYDQDHIYPKSKIYDDSLENRVLVKKNLN
HEKGNQYPIPEKVLNKNAYGFWKILFDKGLIGQKKYTRLTRRTPFEERELAEFIERQIVETRQA
TKETANLLKNICQDSEIVYSKAENASRFRQEFDIIKCRTVNDLHHMHDAYLNIVVGNVYNTKFT
KNPLNFIKDKDNVRSYNLENMFKYDVVRGSYTAWIADDSEGNVKAATIKKVKRELEGKNYRFTR
MSYIGTGGLYDQNLMRKGKGQIPQKENTNKSNIEKYGGYNKASSAYFALIESDGKAGRERTLET
IPIMVYNQEKYGNTEAVDKYLKDNLELQDPKILKDKIKINSLIKLDGFLYNIKGKTGDSLSIAG
SVQLIVNKEEQKLIKKMDKFLVKKKDNKDIKVTSFDNIKEEELIKLYKTLSDKLNNGIYSNKRN
NQAKNISEALDKFKEISIEEKIDVLNQIILLFQSYNNGCNLKSIGLSAKTGVVFIPKKLNYKEC
KLINQSITGLFENEVDLLNL

SEQ ID NO: 319
MGKMYYLGLDIGTNSVGYAVTDPSYHLLKFKGEPMWGAHVFAAGNQSAERRSFRTSRRRLDRRQ
QRVKLVQEIFAPVISPIDPRFFIRLHESALWRDDVAETDKHIFFNDPTYTDKEYYSDYPTIHHL
IVDLMESSEKHDPRLVYLAVAWLVAHRGHFLNEVDKDNIGDVLSFDAFYPEFLAFLSDNGVSPW
VCESKALQATLLSRNSVNDKYKALKSLIFGSQKPEDNFDANISEDGLIQLLAGKKVKVNKLFPQ
ESNDASFTLNDKEDAIEEILGTLTPDECEWIAHIRRLFDWAIMKHALKDGRTISESKVKLYEQH
HHDLTQLKYFVKTYLAKEYDDIFRNVDSETTKNYVAYSYHVKEVKGTLPKNKATQEEFCKYVLG
KVKNIECSEADKVDFDEMIQRLTDNSFMPKQVSGENRVIPYQLYYYELKTILNKAASYLPFLTQ
CGKDAISNQDKLLSIMTFRIPYFVGPLRKDNSEHAWLERKAGKIYPWNFNDKVDLDKSEEAFIR
RMTNTCTYYPGEDVLPLDSLIYEKFMILNEINNIRIDGYPISVDVKQQVFGLFEKKRRVTVKDI
QNLLLSLGALDKHGKLTGIDTTIHSNYNTYHHFKSLMERGVLTRDDVERIVERMTYSDDTKRVR
LWLNNNYGTLTADDVKHISRLRKHDFGRLSKMFLTGLKGVHKETGERASILDFMWNTNDNLMQL
LSECYTFSDEITKLQEAYYAKAQLSLNDFLDSMYISNAVKRPIYRTLAVVNDIRKACGTAPKRI
FIEMARDGESKKKRSVTRREQIKNLYRSIRKDFQQEVDFLEKILENKSDGQLQSDALYLYFAQL
GRDMYTGDPIKLEHIKDQSFYNIDHIYPQSMVKDDSLDNKVLVQSEINGEKSSRYPLDAAIRNK
MKPLWDAYYNHGLISLKKYQRLTRSTPFTDDEKWDFINRQLVETRQSTKALAILLKRKFPDTEI

VYSKAGLSSDFRHEFGLVKSRNINDLHHAKDAFLAIVTGNVYHERFNRRWFMVNQPYSVKTKTL
FTHSIKNGNFVAWNGEEDLGRIVKMLKQNKNTIHFTRFSFDRKEGLFDIQPLKASTGLVPRKAG
LDVVKYGGYDKSTAAYYLLVRFTLEDKKTQHKLMMIPVEGLYKARIDHDKEFLTDYAQTTISEI
LQKDKQKVINIMFPMGTRHIKLNSMISIDGFYLSIGGKSSKGKSVLCHAMVPLIVPHKIECYIK
AMESFARKFKENNKLRIVEKFDKITVEDNLNLYELFLQKLQHNPYNKFFSTQFDVLTNGRSTFT
KLSPEEQVQTLLNILSIFKTCRSSGCDLKSINGSAQAARIMISADLTGLSKKYSDIRLVEQSAS
GLFVSKSQNLLEYL

SEQ ID NO: 320
MTKKEQPYNIGLDIGTSSVGWAVTNDNYDLLNIKKKNLWGVRLFEEAQTAKETRLNRSTRRRYR
RRKNRINWLNEIFSEELAKTDPSFLIRLQNSWVSKKDPDRKRDKYNLFIDGPYTDKEYYREFPT
IFHLRKELILNKDKADIRLIYLALHNILKYRGNFTYEHQKFNISNLNNNLSKELIELNQQLIKY
DISFPDDCDWNHISDILIGRGNATQKSSNILKDFTLDKETKKLLKEVINLILGNVAHLNTIFKT
SLTKDEEKLNFSGKDIESKLDDLDSILDDDQFTVLDAANRIYSTITLNEILNGESYFSMAKVNQ
YENHAIDLCKLRDMWHTTKNEEAVEQSRQAYDDYINKPKYGTKELYTSLKKFLKVALPTNLAKE
AEEKISKGTYLVKPRNSENGVVPYQLNKIEMEKIIDNQSQYYPFLKENKEKLLSILSFRIPYYV
GPLQSAEKNPFAWMERKSNGHARPWNFDEIVDREKSSNKFIRRMTVTDSYLVGEPVLPKNSLIY
QRYEVLNELNNIRITENLKTNPIGSRLTVETKQRIYNELFKKYKKVTVKKLTKWLIAQGYYKNP
ILIGLSQKDEFNSTLTTYLDMKKIFGSSFMEDNKNYDQIEELIEWLTIFEDKQILNEKLHSSKY
SYTPDQIKKISNMRYKGWGRLSKKILMDITTETNTPQLLQLSNYSILDLMWATNNNFISIMSND
KYDFKNYIENHNLNKNEDQNISDLVNDIHVSPALKRGITQSIKIVQEIVKFMGHAPKHIFIEVT
RETKKSEITTSREKRIKRLQSKLLNKANDFKPQLREYLVPNKKIQEELKKHKNDLSSERIMLYF
LQNGKSLYSEESLNINKLSDYQVDHILPRTYIPDDSLENKALVLAKENQRKADDLLLNSNVIDR
NLERWTYMLNNNMIGLKKFKNLTRRVITDKDKLGFIHRQLVQTSQMVKGVANILDNMYKNQGTT
CIQARANLSTAFRKALSGQDDTYHFKHPELVKNRNVNDFHHAQDAYLASFLGTYRLRRFPTNEM
LLMNGEYNKFYGQVKELYSKKKKLPDSRKNGFIISPLVNGTTQYDRNTGEIIWNVGFRDKILKI
FNYHQCNVTRKTEIKTGQFYDQTIYSPKNPKYKKLIAQKKDMDPNIYGGFSGDNKSSITIVKID
NNKIKPVAIPIRLINDLKDKKTLQNWLEENVKHKKSIQIIKNNVPIGQIIYSKKVGLLSLNSDR
EVANRQQLILPPEHSALLRLLQIPDEDLDQILAFYDKNILVEILQELITKMKKFYPFYKGEREF
LIANIENFNQATTSEKVNSLEELITLLHANSTSAHLIFNNIEKKAFGRKTHGLTLNNTDFIYQS
VTGLYETRIHIE

SEQ ID NO: 321
MTKFNKNYSIGLDIGVSSVGYAVVTEDYRVPAFKFKVLGNTEKEKIKKNLIGSTTFVSAQPAKG
TRVFRVNRRRIDRRNHRITYLRDIFQKEIEKVDKNFYRRLDESFRVLGDKSEDLQIKQPFFGDK
ELETAYHKKYPTIYHLRKHLADADKNSPVADIREVYMAISHILKYRGHFLTLDKINPNNININMQN
SWIDFIESCQEVFDLEISDESKNIADIFKSSENRQEKVKKILPYFQQELLKKDKSIFKQLLQLL
FGLKTKFKDCFELEEEPDLNFSKENYDENLENFLGSLEEDFSDVFAKLKVLRDTILLSGMLTYT
GATHARFSATMVERYEEHRKDLQRFKFFIKQNLSEQDYLDIFGRKTQNGFDVDKETKGYVGYIT
NKMVLTNPQKQKTIQQNFYDYISGKITGIEGAEYFLNKISDGTFLRKLRTSDNGAIPNQIHAYE
LEKIIERQGKDYPFLLENKDKLLSILTFKIPYYVGPLAKGSNSRFAWIKRATSSDILDDNDEDT
RNGKIRPWNYQKLINMDETRDAFITNLIGNDIILLNEKVLPKRSLIYEEVMLQNELTRVKYKDK
YGKAHFFDSELRQNIINGLFKNNSKRVNAKSLIKYLSDNHKDLNAIEIVSGVEKGKSFNSTLKT
YNDLKTIFSEELLDSEIYQKELEEIIKVITVFDDKKSIKNYLTKFFGHLEILDEEKINQLSKLR
YSGWGRYSAKLLLDIRDEDTGFNLLQFLRNDEENRNLTKLISDNTLSFEPKIKDIQSKSTIEDD
IFDEIKKLAGSPAIKRGILNSIKIVDELVQIIGYPPHNIVIEMARENMTTEEGQKKAKTRKTKL
ESALKNIENSLLENGKVPHSDEQLQSEKLYLYYLQNGKDMYTLDKTGSPAPLYLDQLDQYEVDH

IIPYSFLPIDSIDNKVLTHRENNQQKLNNIPDKETVANMKPFWEKLYNAKLISQTKYQRLTTSE
RTPDGVLTESMKAGFIERQLVETRQIIKHVARILDNRFSDTKIITLKSQLITNFRNTFHIAKIR
ELNDYHHAHDAYLAVVVGQTLLKVYPKLAPELIYGHHAHFNRHEENKATLRKHLYSNIMRFFNN
PDSKVSKDIWDCNRDLPIIKDVIYNSQINFVKRTMIKKGAFYNQNPVGKFNKQLAANNRYPLKT
KALCLDTSIYGGYGPMNSALSIIIIAERFNEKKGKIETVKEFHDIFIIDYEKFNNNPFQFLNDT
SENGFLKKNNINRVLGFYRIPKYSLMQKIDGTRMLFESKSNLHKATQFKLTKTQNELFFHMKRL
LTKSNLMDLKSKSAIKESQNFILKHKEEFDNISNQLSAFSQKMLGNTTSLKNLIKGYNERKIKE
IDIRDETIKYFYDNFIKMFSFVKSGAPKDINDFFDNKCTVARMRPKPDKKLLNATLIHQSITGL
YETRIDLSKLGED

SEQ ID NO: 322
MKQEYFLGLDMGTGSLGWAVTDSTYQVMRKHGKALWGTRLFESASTAEERRMFRTARRRLDRRN
WRIQVLQEIFSEEISKVDPGFFLRMKESKYYPEDKRDAEGNCPELPYALFVDDNYTDKNYHKDY
PTIYHLRKMLMETTEIPDIRLVYLVLHHMMKHRGHFLLSGDISQIKEFKSTFEQLIQNIQDEEL
EWHISLDDAAIQFVEHVLKDRNLTRSTKKSRLIKQLNAKSACEKAILNLLSGGTVKLSDIFNNK
ELDESERPKVSFADSGYDDYIGIVEAELAEQYYIIASAKAVYDWSVLVEILGNSVSISEAKIKV
YQKHQADLKTLKKIVRQYMTKEDYKRVFVDTEEKLNNYSAYIGMTKKNGKKVDLKSKQCTQADF
YDFLKKNVIKVIDHKEITQEIESEIEKENFLPKQVTKDNGVIPYQVHDYELKKILDNLGTRMPF
IKENAEKIQQLFEFRIPYYVGPLNRVDDGKDGKFTWSVRKSDARIYPWNFTEVIDVEASAEKFI
RRMTNKCTYLVGEDVLPKDSLVYSKFMVLNELNNLRLNGEKISVELKQRIYEELFCKYRKVTRK
KLERYLVIEGIAKKGVEITGIDGDFKASLTAYHDFKERLTDVQLSQRAKEAIVLNVVLFGDDKK
LLKQRLSKMYPNLTTGQLKGICSLSYQGWGRLSKTFLEEITVPAPGTGEVWNIMTALWQTNDNL
MQLLSRNYGFTNEVEEFNTLKKETDLSYKTVDELYVSPAVKRQIWQTLKVVKEIQKVMGNAPKR
VFVEMAREKQEGKRSDSRKKQLVELYRACKNEERDWITELNAQSDQQLRSDKLFLYYIQKGRCM
YSGETIQLDELWDNTKYDIDHIYPQSKTMDDSLNNRVLVKKNYNAIKSDTYPLSLDIQKKMMSF
WKMLQQQGFITKEKYVRLVRSDELSADELAGFIERQIVETRQSTKAVATILKEALPDTEIVYVK
AGNVSNFRQTYELLKVREMNDLHHAKDAYLNIVVGNAYFVKFTKNAAWFIRNNPGRSYNLKRMF
EFDIERSGEIAWKAGNKGSIVTVKKVMQKNNILVTRKAYEVKGGLFDQQIMKKGKGQVPIKGND
ERLADIEKYGGYNKAAGTYFMLVKSLDKKGKEIRTIEFVPLYLKNQIEINHESAIQYLAQERGL
NSPEILLSKIKIDTLFKVDGFKMWLSGRTGNQLIFKGANQLILSHQEAAILKGVVKYVNRKNEN
KDAKLSERDGMTEEKLLQLYDTFLDKLSNTVYSIRLSAQIKTLTEKRAKFIGLSNEDQCIVLNE
ILHMFQCQSGSANLKLIGGPGSAGILVMNNNITACKQISVINQSPTGIYEKEIDLIKL

SEQ ID NO: 323
MKKPYSIGLDIGTNSVGWAVVTDDYKVPAKKMKVLGNTDKSHIEKNLLGALLFDSGNTAEDRRL
KRTARRRYTRRRNRILYLQEIFSEEMGKVDDSFFHRLEDSFLVTEDKRGERHPIFGNLEEEVKY
HENFPTIYHLRQYLADNPEKVDLRLVYLALAHIIKFRGHFLIEGKFDTRNNDVQRLFQEFLAVY
DNTFENSSLQEQNVQVEEILTDKISKSAKKDRVLKLFPNEKSNGRFAEFLKLIVGNQADFKKHF
ELEEKAPLQFSKDTYEEELEVLLAQIGDNYAELFLSAKKLYDSILLSGILTVTDVGTKAPLSAS
MIQRYNEHQMDLAQLKQFIRQKLSDKYNEVFSDVSKDGYAGYIDGKTNQEAFYKYLKGLLNKIE
GSGYFLDKIEREDFLRKQRTFDNGSIPHQIHLQEMRAIIRRQAEFYPFLADNQDRIEKLLTFRI
PYYVGPLARGKSDFAWLSRKSADKITPWNFDEIVDKESSAEAFINRMTNYDLYLPNQKVLPKHS
LLYEKFTVYNELTKVKYKTEQGKTAFFDANMKQEIFDGVFKVYRKVTKDKLMDFLEKEFDEFRI
VDLTGLDKENKVFNASYGTYHDLCKILDKDFLDNSKNEKILEDIVLTLTLFEDREMIRKRLENY
SDLLTKEQVKKLERRHYTGWGRLSAELIHGIRNKESRKTILDYLIDDGNSNRNFMQLINDDALS
FKEEIAKAQVIGETDNLNQVVSDIAGSPAIKKGILQSLKIVDELVKIMGHQPENIVVEMARENQ
FTNQGRRNSQQRLKGLTDSIKEFGSQILKEHPVENSQLQNDRLFLYYLQNGRDMYTGEELDIDY

LSQYDIDHIIPQAFIKDNSIDNRVLTSSKENRGKSDDVPSKDVVRKMKSYWSKLLSAKLITQRK
FDNLTKAERGGLTDDDKAGFIKRQLVETRQITKHVARILDERFNTETDENNKKIRQVKIVTLKS
NLVSNFRKEFELYKVREINDYHHAHDAYLNAVIGKALLGVYPQLEPEFVYGDYPHFHGHKENKA
TAKKFFYSNIMNFFKKDDVRTDKNGEIIWKKDEHISNIKKVLSYPQVNIVKKVEEQTGGFSKES
ILPKGNSDKLIPRKTKKFYWDTKKYGGFDSPIVAYSILVIADIEKGKSKKLKTVKALVGVTIME
KMTFERDPVAFLERKGYRNVQEENIIKLPKYSLFKLENGRKRLLASARELQKGNEIVLPNHLGT
LLYHAKNIHKVDEPKHLDYVDKHKDEFKELLDVVSNFSKKYTLAEGNLEKIKELYAQNNGEDLK
ELASSFINLLTFTAIGAPATFKFFDKNIDRKRYTSTTEILNATLIHQSITGLYETRIDLNKLGG
D

SEQ ID NO: 324
MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRL
KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY
HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTY
NQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF
DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS
MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRI
PYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHS
LLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
SVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA
HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQ
TTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINR
LSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKG
KSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS
AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRV
ILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD
ATLIHQSITGLYETRIDLSQLGGD

SEQ ID NO: 325
MTKPYSIGLDIGTNSVGWAVTTDNYKVPSKKMKVLGNTSKKYIKKNLLGVLLFDSGITAEGRRL
KRTARRRYTRRRNRILYLQEIFSTEMATLDDAFFQRLDDSFLVPDDKRDSKYPIFGNLVEEKAY
HDEFPTIYHLRKYLADSTKKADLRLVYLALAHMIKYRGHFLIEGEFNSKNNDIQKNFQDFLDTY
NAIFESDLSLENSKQLEEIVKDKISKLEKKDRILKLFPGEKNSGIFSEFLKLIVGNQADFRKCF
NLDEKASLHFSKESYDEDLETLLGYIGDDYSDVFLKAKKLYDAILLSGFLTVTDNETEAPLSSA
MIKRYNEHKEDLALLKEYIRNISLKTYNEVFKDDTKNGYAGYIDGKTNQEDFYVYLKKLLAEFE
GADYFLEKIDREDFLRKQRTFDNGSIPYQIHLQEMRAILDKQAKFYPFLAKNKERIEKILTFRI
PYYVGPLARGNSDFAWSIRKRNEKITPWNFEDVIDKESSAEAFINRMTSFDLYLPEEKVLPKHS
LLYETFNVYNELTKVRFIAESMRDYQFLDSKQKKDIVRLYFKDKRKVTDKDIIEYLHAIYGYDG
IELKGIEKQFNSSLSTYHDLLNIINDKEFLDDSSNEAIIEEIIHTLTIFEDREMIKQRLSKFEN
IFDKSVLKKLSRRHYTGWGKLSAKLINGIRDEKSGNTILDYLIDDGISNRNFMQLIHDDALSFK
KKIQKAQIIGDEDKGNIKEVVKSLPGSPAIKKGILQSIKIVDELVKVMGGRKPESIVVEMAREN

QYTNQGKSNSQQRLKRLEKSLKELGSKILKENIPAKLSKIDNNALQNDRLYLYYLQNGKDMYTG
DDLDIDRLSNYDIDHIIPQAFLKDNSIDNKVLVSSASNRGKSDDVPSLEVVKKRKTFWYQLLKS
KLISQRKFDNLTKAERGGLSPEDKAGFIQRQLVETRQITKHVARLLDEKFNNKKDENNRAVRTV
KIITLKSTLVSQFRKDFELYKVREINDFHHAHDAYLNAVVASALLKKYPKLEPEFVYGDYPKYN
SFRERKSATEKVYFYSNIMNIFKKSISLADGRVIERPLIEVNEETGESVWNKESDLATVRRVLS
YPQVNVVKKVEEQNHGLDRGKPKGLFNANLSSKPKPNSNENLVGAKEYLDPKKYGGYAGISNSF
TVLVKGTIEKGAKKKITNVLEFQGISILDRINYRKDKLNFLLEKGYKDIELIIELPKYSLFELS
DGSRRMLASILSTNNKRGEIHKGNQIFLSQKFVKLLYHAKRISNTINENHRKYVENHKKEFEEL
FYYILEFNENYVGAKKNGKLLNSAFQSWQNHSIDELCSSFIGPTGSERKGLFELTSRGSAADFE
FLGVKIPRYRDYTPSSLLKDATLIHQSVTGLYETRIDLAKLGEG

SEQ ID NO: 326
MKKQKFSDYYLGFDIGTNSVGWCVTDLDYNVLRFNKKDMWGSRLFDEAKTAAERRVQRNSRRRL
KRRKWRLNLLEEIFSDEIMKIDSNFFRRLKESSLWLEDKNSKEKFTLFNDDNYKDYDFYKQYPT
IFHLRDELIKNPEKKDIRLIYLALHSIFKSRGHFLFEGQNLKEIKNFETLYNNLISFLEDNGIN
KSIDKDNIEKLEKIICDSGKGLKDKEKEFKGIFNSDKQLVAIFKLSVGSSVSLNDLFDTDEYKK
EEVEKEKISFREQIYEDDKPIYYSILGEKIELLDIAKSFYDFMVLNNILSDSNYISEAKVKLYE
EHKKDLKNLKYIIRKYNKENYDKLFKDKNENNYPAYIGLNKEKDKKEVVEKSRLKIDDLIKVIK
GYLPKPERIEEKDKTIFNEILNKIELKTILPKQRISDNGTLPYQIHEVELEKILENQSKYYDFL
NYEENGVSTKDKLLKTFKFRIPYYVGPLNSYHKDKGGNSWIVRKEEGKILPWNFEQKVDIEKSA
EEFIKRMTNKCTYLNGEDVIPKDSFLYSEYIILNELNKVQVNDEFLNEENKRKIIDELFKENKK
VSEKKFKEYLLVNQIANRTVELKGIKDSFNSNYVSYIKFKDIFGEKLNLDIYKEISEKSILWKC
LYGDDKKIFEKKIKNEYGDILNKDEIKKINSFKFNTWGRLSEKLLTGIEFINLETGECYSSVME
ALRRTNYNLMELLSSKFTLQESIDNENKEMNEVSYRDLIEESYVSPSLKRAILQTLKIYEEIKK
ITGRVPKKVFIEMARGGDESMKNKKIPARQEQLKKLYDSCGNDIANFSIDIKEMKNSLSSYDNN
SLRQKKLYLYYLQFGKCMYTGREIDLDRLLQNNDTYDIDHIYPRSKVIKDDSFDNLVLVLKNEN
AEKSNEYPVKKEIQEKMKSFWRFLKEKNFISDEKYKRLTGKDDFELRGFMARQLVNVRQTTKEV
GKILQQIEPEIKIVYSKAEIASSFREMFDFIKVRELNDTHHAKDAYLNIVAGNVYNTKFTEKPY
RYLQEIKENYDVKKIYNYDIKNAWDKENSLEIVKKNMEKNTVNITRFIKEEKGELFNLNPIKKG
ETSNEIISIKPKLYDGKDNKLNEKYGYYTSLKAAYFIYVEHEKKNKKVKTFERITRIDSTLIKN
EKNLIKYLVSQKKLLNPKIIKKIYKEQTLIIDSYPYTFTGVDSNKKVELKNKKQLYLEKKYEQI
LKNALKFVEDNQGETEENYKFIYLKKRNNNEKNETIDAVKERYNIEFNEMYDKFLEKLSSKDYK
NYINNKLYTNFLNSKEKFKKLKLWEKSLILREFLKIFNKNTYGKYEIKDSQTKEKLFSFPEDTG
RIRLGQSSLGNNKELLEESVTGLFVKKIKL

SEQ ID NO: 327
MKNYTIGLDIGVASVGWVCIDENYKILNYNNRHAFGVHEFESAESAAGRRLKRGMRRRYNRRKK
RLQLLQSLFDSYITDSGFFSKTDSQHFWKNNNEFENRSLTEVLSSLRISSRKYPTIYHLRSDLI
ESNKKMDLRLVYLALHNLVKYRGHFLQEGNWSEAASAEGMDDQLLELVTRYAELENLSPLDLSE
SQWKAAETLLLNRNLTKTDQSKELTAMFGKEYEPFCKLVAGLGVSLHQLFPSSEQALAYKETKT
KVQLSNENVEEVMELLLEEESALLEAVQPFYQQVVLYELLKGETYVAKAKVSAFKQYQKDMASL
KNLLDKTFGEKVYRSYFISDKNSQREYQKSHKVEVLCKLDQFNKEAKFAETFYKDLKKLLEDKS
KTSIGTTEKDEMLRIIKAIDSNQFLQKQKGIQNAAIPHQNSLYEAEKILRNQQAHYPFITTEWI
EKVKQILAFRIPYYIGPLVKDTTQSPFSWVERKGDAPITPWNFDEQIDKAASAEAFISRMRKTC
TYLKGQEVLPKSSLTYERFEVLNELNGIQLRTTGAESDFRHRLSYEMKCWIIDNVFKQYKTVST
KRLLQELKKSPYADELYDEHTGEIKEVFGTQKENAFATSLSGYISMKSILGAVVDDNPAMTEEL
IYWIAVFEDREILHLKIQEKYPSITDVQRQKLALVKLPGWGRFSRLLIDGLPLDEQGQSVLDHM

EQYSSVFMEVLKNKGFGLEKKIQKMNQHQVDGTKKIRYEDIEELAGSPALKRGIWRSVKIVEEL
VSIFGEPANIVLEVAREDGEKKRTKSRKDQWEELTKTTLKNDPDLKSFIGEIKSQGDQRFNEQR
FWLYVTQQGKCLYTGKALDIQNLSMYEVDHILPQNFVKDDSLDNLALVMPEANQRKNQVGQNKM
PLEIIEANQQYAMRTLWERLHELKLISSGKLGRLKKPSFDEVDKDKFIARQLVETRQIIKHVRD
LLDERFSKSDIHLVKAGIVSKFRRFSEIPKIRDYNNKHHAMDALFAAALIQSILGKYGKNFLAF
DLSKKDRQKQWRSVKGSNKEFFLFKNFGNLRLQSPVTGEEVSGVEYMKHVYFELPWQTTKMTQT
GDGMFYKESIFSPKVKQAKYVSPKTEKFVHDEVKNHSICLVEFTFMKKEKEVQETKFIDLKVIE
HHQFLKEPESQLAKFLAEKETNSPIIHARIIRTIPKYQKIWIEHFPYYFISTRELHNARQFEIS
YELMEKVKQLSERSSVEELKIVFGLLIDQMNDNYPIYTKSSIQDRVQKFVDTQLYDFKSFEIGF
EELKKAVAANAQRSDTFGSRISKKPKPEEVAIGYESITGLKYRKPRSVVGTKR

SEQ ID NO: 328
MKKEIKDYFLGLDVGTGSVGWAVTDTDYKLLKANRKDLWGMRCFETAETAEVRRLHRGARRRIE
RRKKRIKLLQELFSQEIAKTDEGFFQRMKESPFYAEDKTILQENTLFNDKDFADKTYHKAYPTI
NHLIKAWIENKVKPDPRLLYLACHNIIKKRGHFLFEGDFDSENQFDTSIQALFEYLREDMEVDI
DADSQKVKEILKDSSLKNSEKQSRLNKILGLKPSDKQKKAITNLISGNKINFADLYDNPDLKDA
EKNSISFSKDDFDALSDDLASILGDSFELLLKAKAVYNCSVLSKVIGDEQYLSFAKVKIYEKHK
TDLTKLKNVIKKHFPKDYKKVFGYNKNEKNNNNYSGYVGVCKTKSKKLIINNSVNQEDFYKFLK
TILSAKSEIKEVNDILTEIETGTFLPKQISKSNAEIPYQLRKMELEKILSNAEKHFSFLKQKDE
KGLSHSEKIIMLLTFKIPYYIGPINDNHKKFFPDRCWVVKKEKSPSGKTTPWNFFDHIDKEKTA
EAFITSRTNFCTYLVGESVLPKSSLLYSEYTVLNEINNLQIIIDGKNICDIKLKQKIYEDLFKK
YKKITQKQISTFIKHEGICNKTDEVIILGIDKECTSSLKSYIELKNIFGKQVDEISTKNMLEEI
IRWATIYDEGEGKTILKTKIKAEYGKYCSDEQIKKILNLKFSGWGRLSRKFLETVTSEMPGFSE
PVNIITAMRETQNNLMELLSSEFTFTENIKKINSGFEDAEKQFSYDGLVKPLFLSPSVKKMLWQ
TLKLVKEISHITQAPPKKIFIEMAKGAELEPARTKTRLKILQDLYNNCKNDADAFSSEIKDLSG
KIENEDNLRLRSDKLYLYYTQLGKCMYCGKPIEIGHVFDTSNYDIDHIYPQSKIKDDSISNRVL
VCSSCNKNKEDKYPLKSEIQSKQRGFWNFLQRNNFISLEKLNRLTRATPISDDETAKFIARQLV
ETRQATKVAAKVLEKMFPETKIVYSKAETVSMFRNKFDIVKCREINDFHHAHDAYLNIVVGNVY
NTKFTNNPWNFIKEKRDNPKIADTYNYYKVFDYDVKRNNITAWEKGKTIITVKDMLKRNTPIYT
RQAACKKGELFNQTIMKKGLGQHPLKKEGPFSNISKYGGYNKVSAAYYTLIEYEEKGNKIRSLE
TIPLYLVKDIQKDQDVLKSYLTDLLGKKEFKILVPKIKINSLLKINGFPCHITGKTNDSFLLRP
AVQFCCSNNEVLYFKKIIRFSEIRSQREKIGKTISPYEDLSFRSYIKENLWKKTKNDEIGEKEF
YDLLQKKNLEIYDMLLTKHKDTIYKKRPNSATIDILVKGKEKFKSLIIENQFEVILEILKLFSA
TRNVSDLQHIGGSKYSGVAKIGNKISSLDNCILIYQSITGIFEKRIDLLKV

SEQ ID NO: 329
MEGQMKNNGNNLQQGNYYLGLDVGTSSVGWAVTDTDYNVLKFRGKSMWGARLFDEASTAEERRT
HRGNRRRLARRKYRLLLLEQLFEKEIRKIDDNFFVRLHESNLWADDKSKPSKFLLFNDTNFTDK
DYLKKYPTIYHLRSDLIHNSTEHDIRLVFLALHHLIKYRGHFIYDNSANGDVKTLDEAVSDFEE
YLNENDIEFNIENKKEFINVLSDKHLTKKEKKISLKKLYGDITDSENINISVLIEMLSGSSISL
SNLFKDIEFDGKQNLSLDSDIEETLNDVVDILGDNIDLLIHAKEVYDIAVLTSSLGKHKYLCDA
KVELFEKNKKDLMILKKYIKKNHPEDYKKIFSSPTEKKNYAAYSQTNSKNVCSQEEFCLFIKPY
IRDMVKSENEDEVRIAKEVEDKSFLTKLKGTNNSVVPYQIHERELNQILKNIVAYLPFMNDEQE
DISVVDKIKLIFKFKIPYYVGPLNTKSTRSWVYRSDEKIYPWNFSNVIDLDKTAHEFMNRLIGR
CTYTNDPVLPMDSLLYSKYNVLNEINPIKVNGKAIPVEVKQAIYTDLFENSKKKVTRKSIYIYL
LKNGYIEKEDIVSGIDIEIKSKLKSHHDFTQIVQENKCTPEEIERIIKGILVYSDDKSMLRRWL
KNNIKGLSENDVKYLAKLNYKEWGRLSKTLLTDIYTINPEDGEACSILDIMWNTNATLMEILSN

EKYQFKQNIENYKAENYDEKQNLHEELDDMYISPAARRSIWQALRIVDEIVDIKKSAPKKIFIE
MAREKKSAMKKKRTESRKDTLLELYKSCKSQADGFYDEELFEKLSNESNSRLRRDQLYLYYTQM
GRSMYTGKRIDFDKLINDKNTYDIDHIYPRSKIKDDSITNRVLVEKDINGEKTDIYPISEDIRQ
KMQPFWKILKEKGLINEEKYKRLTRNYELTDEELSSFVARQLVETQQSTKALATLLKKEYPSAK
IVYSKAGNVSEFRNRKDKELPKFREINDLHHAKDAYLNIVVGNVYDTKFTEKFFNNIRNENYSL
KRVFDFSVPGAWDAKGSTFNTIKKYMAKNNPIIAFAPYEVKGELFDQQIVPKGKGQFPIKQGKD
IEKYGGYNKLSSAFLFAVEYKGKKARERSLETVYIKDVELYLQDPIKYCESVLGLKEPQIIKPK
ILMGSLFSINNKKLVVTGRSGKQYVCHHIYQLSINDEDSQYLKNIAKYLQEEPDGNIERQNILN
ITSVNNIKLFDVLCTKFNSNTYEIILNSLKNDVNEGREKFSELDILEQCNILLQLLKAFKCNRE
SSNLEKLNNKKQAGVIVIPHLFTKCSVFKVIHQSITGLFEKEMDLLK

SEQ ID NO: 330
MGRKPYILSLDIGTGSVGYACMDKGFNVLKYHDKDALGVYLFDGALTAQERRQFRTSRRRKNRR
IKRLGLLQELLAPLVQNPNFYQFQRQFAWKNDNMDFKNKSLSEVLSFLGYESKKYPTIYHLQEA
LLLKDEKFDPELIYMALYHLVKYRGHFLFDHLKIENLTNNDNMHDFVELIETYENLNNIKLNLD
YEKTKVIYEILKDNEMTKNDRAKRVKNMEKKLEQFSIMLLGLKFNEGKLFNHADNAEELKGANQ
SHTFADNYEENLTPFLTVEQSEFIERANKIYLSLTLQDILKGKKSMAMSKVAAYDKFRNELKQV
KDIVYKADSTRTQFKKIFVSSKKSLKQYDATPNDQTFSSLCLFDQYLIRPKKQYSLLIKELKKI
IPQDSELYFEAENDTLLKVLNTTDNASIPMQINLYEAETILRNQQKYHAEITDEMIEKVLSLIQ
FRIPYYVGPLVNDHTASKFGWMERKSNESIKPWNFDEVVDRSKSATQFIRRMTNKCSYLINEDV
LPKNSLLYQEMEVLNELNATQIRLQTDPKNRKYRMMPQIKLFAVEHIFKKYKTVSHSKFLEIML
NSNHRENFMNHGEKLSIFGTQDDKKFASKLSSYQDMTKIFGDIEGKRAQIEEIIQWITIFEDKK
ILVQKLKECYPELTSKQINQLKKLNYSGWGRLSEKLLTHAYQGHSIIELLRHSDENFMEILTND
VYGFQNFIKEENQVQSNKIQHQDIANLTTSPALKKGIWSTIKLVRELTSIFGEPEKIIMEFATE
DQQKGKKQKSRKQLWDDNIKKNKLKSVDEYKYIIDVANKLNNEQLQQEKLWLYLSQNGKCMYSG
QSIDLDALLSPNATKHYEVDHIFPRSFIKDDSIDNKVLVIKKMNQTKGDQVPLQFIQQPYERIA
YWKSLNKAGLISDSKLHKLMKPEFTAMDKEGFIQRQLVETRQISVHVRDFLKEEYPNTKVIPMK
AKMVSEFRKKFDIPKIRQMNDAHHAIDAYLNGVVYHGAQLAYPNVDLFDFNFKWEKVREKWKAL
GEFNTKQKSRELFFFKKLEKMEVSQGERLISKIKLDMNHFKINYSRKLANIPQQFYNQTAVSPK
TAELKYESNKSNEVVYKGLTPYQTYVVAIKSVNKKGKEKMEYQMIDHYVFDFYKFQNGNEKELA
LYLAQRENKDEVLDAQIVYSLNKGDLLYINNHPCYFVSRKEVINAKQFELTVEQQLSLYNVMNN
KETNVEKLLIEYDFIAEKVINEYHHYLNSKLKEKRVRTFFSESNQTHEDFIKALDELFKVVTAS
ATRSDKIGSRKNSMTHRAFLGKGKDVKIAYTSISGLKTTKPKSLFKLAESRNEL

SEQ ID NO: 331
MAKILGLDLGTNSIGWAVVERENIDFSLIDKGVRIFSEGVKSEKGIESSRAAERTGYRSARKIK
YRRKLRKYETLKVLSLNRMCPLSIEEVEEWKKSGFKDYPLNPEFLKWLSTDEESNVNPYFFRDR
ASKHKVSLFELGRAFYHIAQRRGFLSNRLDQSAEGILEEHCPKIEAIVEDLISIDEISTNITDY
FFETGILDSNEKNGYAKDLDEGDKKLVSLYKSLLAILKKNESDFENCKSEIIERLNKKDVLGKV
KGKIKDISQAMLDGNYKTLGQYFYSLYSKEKIRNQYTSREEHYLSEFITICKVQGIDQINEEEK
INEKKFDGLAKDLYKAIFFQRPLKSQKGLIGKCSFEKSKSRCAISHPDFEEYRMWTYLNTIKIG
TQSDKKLRFLTQDEKLKLVPKFYRKNDFNFDVLAKELIEKGSSFGFYKSSKKNDFFYWFNYKPT
DTVAACQVAASLKNAIGEDWKTKSFKYQTINSNKEQVSRTVDYKDLWHLLTVATSDVYLYEFAI
DKLGLDEKNAKAFSKTKLKKDFASLSLSAINKILPYLKEGLLYSHAVFVANIENIVDENIWKDE
KQRDYIKTQISEIIENYTLEKSRFEIINGLLKEYKSENEDGKRVYYSKEAEQSFENDLKKKLVL
FYKSNEIENKEQQETIFNELLPIFIQQLKDYEFIKIQRLDQKVLIFLKGKNETGQIFCTEEKGT
AEEKEKKIKNRLKKLYHPSDIEKFKKKIIKDEFGNEKIVLGSPLTPSIKNPMAMRALHQLRKVL

NALILEGQIDEKTIIHIEMARELNDANKRKGIQDYQNDNKKFREDAIKEIKKLYFEDCKKEVEP
TEDDILRYQLWMEQNRSEIYEEGKNISICDIIGSNPAYDIEHTIPRSRSQDNSQMNKTLCSQRF
NREVKKQSMPIELNNHLEILPRIAHWKEEADNLTREIEIISRSIKAAATKEIKDKKIRRRHYLT
LKRDYLQGKYDRFIWEEPKVGFKNSQIPDTGIITKYAQAYLKSYFKKVESVKGGMVAEFRKIWG
IQESFIDENGMKHYKVKDRSKHTHHTIDAITIACMTKEKYDVLAHAWTLEDQQNKKEARSIIEA
SKPWKTFKEDLLKIEEEILVSHYTPDNVKKQAKKIVRVRGKKQFVAEVERDVNGKAVPKKAASG
KTIYKLDGEGKKLPRLQQGDTIRGSLHQDSIYGAIKNPLNTDEIKYVIRKDLESIKGSDVESIV
DEVVKEKIKEAIANKVLLLSSNAQQKNKLVGTVWMNEEKRIAINKVRIYANSVKNPLHIKEHSL
LSKSKHVHKQKVYGQNDENYAMAIYELDGKRDFELINIFNLAKLIKQGQGFYPLHKKKEIKGKI
VFVPIEKRNKRDVVLKRGQQVVFYDKEVENPKDISEIVDFKGRIYIIEGLSIQRIVRPSGKVDE
YGVIMLRYFKEARKADDIKQDNFKPDGVFKLGENKPTRKMNHQFTAFVEGIDFKVLPSGKFEKI

SEQ ID NO: 332
MEFKKVLGLDIGTNSIGCALLSLPKSIQDYGKGGRLEWLTSRVIPLDADYMKAFIDGKNGLPQV
ITPAGKRRQKRGSRRLKHRYKLRRSRLIRVFKTLNWLPEDFPLDNPKRIKETISTEGKFSFRIS
DYVPISDESYREFYREFGYPENEIEQVIEEINFRRKTKGKNKNPMIKLLPEDWVVYYLRKKALI
KPTTKEELIRIIYLFNQRRGFKSSRKDLTETAILDYDEFAKRLAEKEKYSAENYETKFVSITKV
KEVVELKTDGRKGKKRFKVILEDSRIEPYEIERKEKPDWEGKEYTFLVTQKLEKGKFKQNKPDL
PKEEDWALCTTALDNRMGSKHPGEFFFDELLKAFKEKRGYKIRQYPVNRWRYKKELEFIWTKQC
QLNPELNNLNINKEILRKLATVLYPSQSKFFGPKIKEFENSDVLHIISEDIIYYQRDLKSQKSL
ISECRYEKRKGIDGEIYGLKCIPKSSPLYQEFRIWQDIHNIKVIRKESEVNGKKKINIDETQLY
INENIKEKLFELFNSKDSLSEKDILELISLNIINSGIKISKKEEETTHRINLFANRKELKGNET
KSRYRKVFKKLGFDGEYILNHPSKLNRLWHSDYSNDYADKEKTEKSILSSLGWKNRNGKWEKSK
NYDVFNLPLEVAKAIANLPPLKKEYGSYSALAIRKMLVVMRDGKYWQHPDQIAKDQENTSLMLF
DKNLIQLTNNQRKVLNKYLLTLAEVQKRSTLIKQKLNEIEHNPYKLELVSDQDLEKQVLKSFLE
KKNESDYLKGLKTYQAGYLIYGKHSEKDVPIVNSPDELGEYIRKKLPNNSLRNPIVEQVIRETI
FIVRDVWKSFGIIDEIHIELGRELKNNSEERKKTSESQEKNFQEKERARKLLKELLNSSNFEHY
DENGNKIFSSFTVNPNPDSPLDIEKFRIWKNQSGLTDEELNKKLKDEKIPTEIEVKKYILWLTQ
KCRSPYTGKIIPLSKLFDSNVYEIEHIIPRSKMKNDSTNNLVICELGVNKAKGDRLAANFISES
NGKCKFGEVEYTLLKYGDYLQYCKDTFKYQKAKYKNLLATEPPEDFIERQINDTRYIGRKLAEL
LTPVVKDSKNIIFTIGSITSELKITWGLNGVWKDILRPRFKRLESIINKKLIFQDEDDPNKYHF
DLSINPQLDKEGLKRLDHRHHALDATIIAATTREHVRYLNSLNAADNDEEKREYFLSLCNHKIR
DFKLPWENFTSEVKSKLLSCVVSYKESKPILSDPFNKYLKWEYKNGKWQKVFAIQIKNDRWKAV
RRSMFKEPIGTVWIKKIKEVSLKEAIKIQAIWEEVKNDPVRKKKEKYIYDDYAQKVIAKIVQEL
GLSSSMRKQDDEKLNKFINEAKVSAGVNKNLNTTNKTIYNLEGRFYEKIKVAEYVLYKAKRMPL
NKKEYIEKLSLQKMFNDLPNFILEKSILDNYPEILKELESDNKYIIEPHKKNNPVNRLLLEHIL
EYHNNPKEAFSTEGLEKLNKKAINKIGKPIKYITRLDGDINEEEIFRGAVFETDKGSNVYFVMY
ENNQTKDREFLKPNPSISVLKAIEHKNKIDFFAPNRLGFSRIILSPGDLVYVPTNDQYVLIKDN
SSNETIINWDDNEFISNRIYQVKKFTGNSCYFLKNDIASLILSYSASNGVGEFGSQNISEYSVD
DPPIRIKDVCIKIRVDRLGNVRPL

SEQ ID NO: 333
MKHILGLDLGTNSIGWALIERNIEEKYGKIIGMGSRIVPMGAELSKFEQGQAQTKNADRRTNRG
ARRLNKRYKQRRNKLIYILQKLDMLPSQIKLKEDFSDPNKIDKITILPISKKQEQLTAFDLVSL
RVKALTEKVGLEDLGKIIYKYNQLRGYAGGSLEPEKEDIFDEEQSKDKKNKSFIAFSKIVFLGE
PQEEIFKNKKLNRRAIIVETEEGNFEGSTFLENIKVGDSLELLINISASKSGDTITIKLPNKTN
WRKKMENIENQLKEKSKEMGREFYISEFLLELLKENRWAKIRNNTILRARYESEFEAIWNEQVK

HYPFLENLDKKTLIEIVSFIFPGEKESQKKYRELGLEKGLKYIIKNQVVFYQRELKDQSHLISD
CRYEPNEKAIAKSHPVFQEYKVWEQINKLIVNTKIEAGTNRKGEKKYKYIDRPIPTALKEWIFE
ELQNKKEITFSAIFKKLKAEFDLREGIDFLNGMSPKDKLKGNETKLQLQKSLGELWDVLGLDSI
NRQIELWNILYNEKGNEYDLTSDRTSKVLEFINKYGNNIVDDNAEETAIRISKIKFARAYSSLS
LKAVERILPLVRAGKYFNNDFSQQLQSKILKLLNENVEDPFAKAAQTYLDNNQSVLSEGGVGNS
IATILVYDKHTAKEYSHDELYKSYKEINLLKQGDLRNPLVEQIINEALVLIRDIWKNYGIKPNE
IRVELARDLKNSAKERATIHKRNKDNQTINNKIKETLVKNKKELSLANIEKVKLWEAQRHLSPY
TGQPIPLSDLFDKEKYDVDHIIPISRYFDDSFTNKVISEKSVNQEKANRTAMEYFEVGSLKYSI
FTKEQFIAHVNEYFSGVKRKNLLATSIPEDPVQRQIKDTQYIAIRVKEELNKIVGNENVKTTTG
SITDYLRNHWGLTDKFKLLLKERYEALLESEKFLEAEYDNYKKDFDSRKKEYEEKEVLFEEQEL
TREEFIKEYKENYIRYKKNKLIIKGWSKRIDHRHHAIDALIVACTEPAHIKRLNDLNKVLQDWL
VEHKSEFMPNFEGSNSELLEEILSLPENERTEIFTQIEKFRAIEMPWKGFPEQVEQKLKEIIIS
HKPKDKLLLQYNKAGDRQIKLRGQLHEGTLYGISQGKEAYRIPLTKFGGSKFATEKNIQKIVSP
FLSGFIANHLKEYNNKKEEAFSAEGIMDLNNKLAQYRNEKGELKPHTPISTVKIYYKDPSKNKK
KKDEEDLSLQKLDREKAFNEKLYVKTGDNYLFAVLEGEIKTKKTSQIKRLYDIISFFDATNFLK
EEFRNAPDKKTFDKDLLFRQYFEERNKAKLLFTLKQGDFVYLPNENEEVILDKESPLYNQYWGD
LKERGKNIYVVQKFSKKQIYFIKHTIADIIKKDVEFGSQNCYETVEGRSIKENCFKLEIDRLGN
IVKVIKR

SEQ ID NO: 334
MHVEIDFPHFSRGDSHLAMNKNEILRGSSVLYRLGLDLGSNSLGWFVTHLEKRGDRHEPVALGP
GGVRIFPDGRDPQSGTSNAVDRRMARGARKRRDRFVERRKELIAALIKYNLLPDDARERRALEV
LDPYALRKTALTDTLPAHHVGRALFHLNQRRGFQSNRKTDSKQSEDGAIKQAASRLATDKGNET
LGVFFADMHLRKSYEDRQTAIRAELVRLGKDHLTGNARKKIWAKVRKRLFGDEVLPRADAPHGV
RARATITGTKASYDYYPTRDMLRDEFNAIWAGQSAHHATITDEARTEIEHIIFYQRPLKPAIVG
KCTLDPATRPFKEDPEGYRAPWSHPLAQRFRILSEARNLEIRDTGKGSRRLTKEQSDLVVAALL
ANREVKFDKLRTLLKLPAEARFNLESDRRAALDGDQTAARLSDKKGFNKAWRGFPPERQIAIVA
RLEETEDENELIAWLEKECALDGAAAARVANTTLPDGHCRLGLRAIKKIVPIMQDGLDEDGVAG
AGYHIAAKRAGYDHAKLPTGEQLGRLPYYGQWLQDAVVGSGDARDQKEKQYGQFPNPTVHIGLG
QLRRVVNDLIDKYGPPTEISIEFTRALKLSEQQKAERQREQRRNQDKNKARAEELAKFGRPANP
RNLLKMRLWEELAHDPLDRKCVYTGEQISIERLLSDEVDIDHILPVAMTLDDSPANKIICMRYA
NRHKRKQTPSEAFGSSPTLQGHRYNWDDIAARATGLPRNKRWRFDANAREEFDKRGGFLARQLN
ETGWLARLAKQYLGAVTDPNQIWVVPGRLTSMLRGKWGLNGLLPSDNYAGVQDKAEEFLASTDD
MEFSGVKNRADHRHHAIDGLVTALTDRSLLWKMANAYDEEHEKFVIEPPWPTMRDDLKAALEKM
VVSHKPDHGIEGKLHEDSAYGFVKPLDATGLKEEEAGNLVYRKAIESLNENEVDRIRDIQLRTI
VRDHVNVEKTKGVALADALRQLQAPSDDYPQFKHGLRHVRILKKEKGDYLVPIANRASGVAYKA
YSAGENFCVEVFETAGGKWDGEAVRRFDANKKNAGPKIAHAPQWRDANEGAKLVMRIHKGDLIR
LDHEGRARIMVVHRLDAAAGRFKLADHNETGNLDKRHATNNDIDPFRWLMASYNTLKKLAAVPV
RVDELGRVWRVMPN

SEQ ID NO: 335
METTLGIDLGTNSIGLALVDQEEHQILYSGVRIFPEGINKDTIGLGEKEESRNATRRAKRQMRR
QYFRKKLRKAKLLELLIAYDMCPLKPEDVRRWKNWDKQQKSTVRQFPDTPAFREWLKQNPYELR
KQAVTEDVTRPELGRILYQMIQRRGFLSSRKGKEEGKIFTGKDRMVGIDETRKNLQKQTLGAYL
YDIAPKNGEKYRFRTERVRARYTLRDMYIREFEIIWQRQAGHLGLAHEQATRKKNIFLEGSATN
VRNSKLITHLQAKYGRGHVLIEDTRITVTFQLPLKEVLGGKIEIEEEQLKFKSNESVLFWQRPL
RSQKSLLSKCVFEGRNFYDPVHQKWIIAGPTPAPLSHPEFEEFRAYQFINNIIYGKNEHLTAIQ

REAVFELMCTESKDFNFEKIPKHLKLFEKFNFDDTTKVPACTTISQLRKLFPHPVWEEKREEIW
HCFYFYDDNTLLFEKLQKDYALQTNDLEKIKKIRLSESYGNVSLKAIRRINPYLKKGYAYSTAV
LLGGIRNSFGKRFEYFKEYEPEIEKAVCRILKEKNAEGEVIRKIKDYLVHNRFGFAKNDRAFQK
LYHHSQAITTQAQKERLPETGNLRNPIVQQGLNELRRTVNKLLATCREKYGPSFKFDHIHVEMG
RELRSSKTEREKQSRQIRENEKKNEAAKVKLAEYGLKAYRDNIQKYLLYKEIEEKGGTVCCPYT
GKTLNISHTLGSDNSVQIEHIIPYSISLDDSLANKTLCDATFNREKGELTPYDFYQKDPSPEKW
GASSWEEIEDRAFRLLPYAKAQRFIRRKPQESNEFISRQLNDTRYISKKAVEYLSAICSDVKAF
PGQLTAELRHLWGLNNILQSAPDITFPLPVSATENHREYYVITNEQNEVIRLFPKQGETPRTEK
GELLLTGEVERKVFRCKGMQEFQTDVSDGKYWRRIKLSSSVTWSPLFAPKPISADGQIVLKGRI
EKGVFVCNQLKQKLKTGLPDGSYWISLPVISQTFKEGESVNNSKLTSQQVQLFGRVREGIFRCH
NYQCPASGADGNFWCTLDTDTAQPAFTPIKNAPPGVGGGQIILTGDVDDKGIFHADDDLHYELP
ASLPKGKYYGIFTVESCDPTLIPIELSAPKTSKGENLIEGNIWVDEHTGEVRFDPKKNREDQRH
HAIDAIVIALSSQSLFQRLSTYNARRENKKRGLDSTEHFPSPWPGFAQDVRQSVVPLLVSYKQN
PKTLCKISKTLYKDGKKIHSCGNAVRGQLHKETVYGQRTAPGATEKSYHIRKDIRELKTSKHIG
KVVDITIRQMLLKHLQENYHIDITQEFNIPSNAFFKEGVYRIFLPNKHGEPVPIKKIRMKEELG
NAERLKDNINQYVNPRNNHHVMIYQDADGNLKEEIVSFWSVIERQNQGQPIYQLPREGRNIVSI
LQINDTFLIGLKEEEPEVYRNDLSTLSKHLYRVQKLSGMYYTFRHHLASTLNNEREEFRIQSLE
AWKRANPVKVQIDEIGRITFLNGPLC

SEQ ID NO: 336
MESSQILSPIGIDLGGKFTGVCLSHLEAFAELPNHANTKYSVILIDHNNFQLSQAQRRATRHRV
RNKKRNQFVKRVALQLFQHILSRDLNAKEETALCHYLNNRGYTYVDTDLDEYIKDETTINLLKE
LLPSESEHNFIDWFLQKMQSSEFRKILVSKVEEKKDDKELKNAVKNIKNFITGFEKNSVEGHRH
RKVYFENIKSDITKDNQLDSIKKKIPSVCLSNLLGHLSNLQWKNLHRYLAKNPKQFDEQTFGNE
FLRMLKNFRHLKGSQESLAVRNLIQQLEQSQDYISILEKTPPEITIPPYEARTNTGMEKDQSLL
LNPEKLNNLYPNWRNLIPGIIDAHPFLEKDLEHTKLRDRKRIISPSKQDEKRDSYILQRYLDLN
KKIDKFKIKKQLSFLGQGKQLPANLIETQKEMETHFNSSLVSVLIQIASAYNKEREDAAQGIWF
DNAFSLCELSNINPPRKQKILPLLVGAILSEDFINNKDKWAKFKIFWNTHKIGRTSLKSKCKEI
EEARKNSGNAFKIDYEEALNHPEHSNNKALIKIIQTIPDIIQAIQSHLGHNDSQALIYHNPFSL
SQLYTILETKRDGFHKNCVAVTCENYWRSQKTEIDPEISYASRLPADSVRPFDGVLARMMQRLA
YEIAMAKWEQIKHIPDNSSLLIPIYLEQNRFEFEESFKKIKGSSSDKTLEQAIEKQNIQWEEKF
QRIINASMNICPYKGASIGGQGEIDHIYPRSLSKKHFGVIFNSEVNLIYCSSQGNREKKEEHYL
LEHLSPLYLKHQFGTDNVSDIKNFISQNVANIKKYISFHLLTPEQQKAARHALFLDYDDEAFKT
ITKFLMSQQKARVNGTQKFLGKQIMEFLSTLADSKQLQLEFSIKQITAEEVHDHRELLSKQEPK
LVKSRQQSFPSHAIDATLTMSIGLKEFPQFSQELDNSWFINHLMPDEVHLNPVRSKEKYNKPNI
SSTPLFKDSLYAERFIPVWVKGETFAIGFSEKDLFEIKPSNKEKLFTLLKTYSTKNPGESLQEL
QAKSKAKWLYFPINKTLALEFLHHYFHKEIVTPDDTTVCHFINSLRYYTKKESITVKILKEPMP
VLSVKFESSKKNVLGSFKHTIALPATKDWERLFNHPNFLALKANPAPNPKEFNEFIRKYFLSDN
NPNSDIPNNGHNIKPQKHKAVRKVFSLPVIPGNAGTMMRIRRKDNKGQPLYQLQTIDDTPSMGI
QINEDRLVKQEVLMDAYKTRNLSTIDGINNSEGQAYATFDNWLTLPVSTFKPEIIKLEMKPHSK
TRRYIRITQSLADFIKTIDEALMIKPSDSIDDPLNMPNEIVCKNKLFGNELKPRDGKMKIVSTG
KIVTYEFESDSTPQWIQTLYVTQLKKQP

SEQ ID NO: 337
MKKIVGLDLGTNSIGWALINAYINKEHLYGIEACGSRIIPMDAAILGNFDKGNSISQTADRTSY
RGIRRLRERHLLRRERLHRILDLLGFLPKHYSDSLNRYGKFLNDIECKLPWVKDETGSYKFIFQ
ESFKEMLANFTEHHPILIANNKKVPYDWTIYYLRKKALTQKISKEELAWILLNFNQKRGYYQLR

GEEEETPNKLVEYYSLKVEKVEDSGERKGKDTWYNVHLENGMIYRRTSNIPLDWEGKTKEFIVT
TDLEADGSPKKDKEGNIKRSFRAPKDDDWTLIKKKTEADIDKIKMTVGAYIYDTLLQKPDQKIR
GKLVRTIERKYYKNELYQILKTQSEFHEELRDKQLYIACLNELYPNNEPRRNSISTRDFCHLFI
EDIIFYQRPLKSKKSLIDNCPYEENRYIDKESGEIKHASIKCIAKSHPLYQEFRLWQFIVNLRI
YRKETDVDVTQELLPTEADYVTLFEWLNEKKEIDQKAFFKYPPFGFKKTTSNYRWNYVEDKPYP
CNETHAQIIARLGKAHIPKAFLSKEKEETLWHILYSIEDKQEIEKALHSFANKNNLSEEFIEQF
KNFPPFKKEYGSYSAKAIKKLLPLMRMGKYWSIENIDNGTRIRINKIIDGEYDENIRERVRQKA
INLTDITHFRALPLWLACYLVYDRHSEVKDIVKWKTPKDIDLYLKSFKQHSLRNPIVEQVITET
LRTVRDIWQQVGHIDEIHIELGREMKNPADKRARMSQQMIKNENTNLRIKALLTEFLNPEFGIE
NVRPYSPSQQDLLRIYEEGVLNSILELPEDIGIILGKFNQTDTLKRPTRSEILRYKLWLEQKYR
SPYTGEMIPLSKLFTPAYEIEHIIPQSRYFDDSLSNKVICESEINKLKDRSLGYEFIKNHHGEK
VELAFDKPVEVLSVEAYEKLVHESYSHNRSKMKKLLMEDIPDQFIERQLNDSRYISKVVKSLLS
NIVREENEQEAISKNVIPCTGGITDRLKKDWGINDVWNKIVLPRFIRLNELTESTRFTSINTNN
TMIPSMPLELQKGFNKKRIDHRHHAMDAIIIACANRNIVNYLNNVSASKNTKITRRDLQTLLCH
KDKTDNNGNYKWVIDKPWETFTQDTLTALQKITVSFKQNLRVINKTTNHYQHYENGKKIVSNQS
KGDSWAIRKSMHKETVHGEVNLRMIKTVSFNEALKKPQAIVEMDLKKKILAMLELGYDTKRIKN
YFEENKDTWQDINPSKIKVYYFTKETKDRYFAVRKPIDTSFDKKKIKESITDTGIQQIMLRHLE
TKDNDPTLAFSPDGIDEMNRNILILNKGKKHQPIYKVRVYEKAEKFTVGQKGNKRTKFVEAAKG
TNLFFAIYETEEIDKDTKKVIRKRSYSTIPLNVVIERQKQGLSSAPEDENGNLPKYILSPNDLV
YVPTQEEINKGEVVMPIDRDRIYKMVDSSGITANFIPASTANLIFALPKATAEIYCNGENCIQN
EYGIGSPQSKNQKAITGEMVKEICFPIKVDRLGNIIQVGSCILTN

SEQ ID NO: 338
MSRSLTFSFDIGYASIGWAVIASASHDDADPSVCGCGTVLFPKDDCQAFKRREYRRLRRNIRSR
RVRIERIGRLLVQAQIITPEMKETSGHPAPFYLASEALKGHRTLAPIELWHVLRWYAHNRGYDN
NASWSNSLSEDGGNGEDTERVKHAQDLMDKHGTATMAETICRELKLEEGKADAPMEVSTPAYKN
LNTAFPRLIVEKEVRRILELSAPLIPGLTAEIIELIAQHHPLTTEQRGVLLQHGIKLARRYRGS
LLFGQLIPRFDNRIISRCPVTWAQVYEAELKKGNSEQSARERAEKLSKVPTANCPEFYEYRMAR
ILCNIRADGEPLSAEIRRELMNQARQEGKLTKASLEKAISSRLGKETETNVSNYFTLHPDSEEA
LYLNPAVEVLQRSGIGQILSPSVYRIAANRLRRGKSVTPNYLLNLLKSRGESGEALEKKIEKES
KKKEADYADTPLKPKYATGRAPYARTVLKKVVEEILDGEDPTRPARGEAHPDGELKAHDGCLYC
LLDTDSSVNQHQKERRLDTMTNNHLVRHRMLILDRLLKDLIQDFADGQKDRISRVCVEVGKELT
TFSAMDSKKIQRELTLRQKSHTDAVNRLKRKLPGKALSANLIRKCRIAMDMNWTCPFTGATYGD
HELENLELEHIVPHSFRQSNALSSLVLTWPGVNRMKGQRTGYDFVEQEQENPVPDKPNLHICSL
NNYRELVEKLDDKKGHEDDRRRKKKRKALLMVRGLSHKHQSQNHEAMKEIGMTEGMMTQSSHLM
KLACKSIKTSLPDAHIDMIPGAVTAEVRKAWDVFGVFKELCPEAADPDSGKILKENLRSLTHLH
HALDACVLGLIPYIIPAHHNGLLRRVLAMRRIPEKLIPQVRPVANQRHYVLNDDGRMMLRDLSA
SLKENIREQLMEQRVIQHVPADMGGALLKETMQRVLSVDGSGEDAMVSLSKKKDGKKEKNQVKA
SKLVGVFPEGPSKLKALKAAIEIDGNYGVALDPKPVVIRHIKVFKRIMALKEQNGGKPVRILKK
GMLIHLTSSKDPKHAGVWRIESIQDSKGGVKLDLQRAHCAVPKNKTHECNWREVDLISLLKKYQ
MKRYPTSYTGTPR

SEQ ID NO: 339
MTQKVLGLDLGTNSIGSAVRNLDLSDDLQWQLEFFSSDIFRSSVNKESNGREYSLAAQRSAHRR
SRGLNEVRRRRLWATLNLLIKHGFCPMSSESLMRWCTYDKRKGLFREYPIDDKDFNAWILLDFN
GDGRPDYSSPYQLRRELVTRQFDFEQPIERYKLGRALYHIAQHRGFKSSKGETLSQQETNSKPS
STDEIPDVAGAMKASEEKLSKGLSTYMKEHNLLTVGAAFAQLEDEGVRVRNNNDYRAIRSQFQH

EIETIFKFQQGLSVESELYERLISEKKNVGTIFYKRPLRSQRGNVGKCTLERSKPRCAIGHPLF
EKFRAWTLINNIKVRMSVDTLDEQLPMKLRLDLYNECFLAFVRTEFKFEDIRKYLEKRLGIHFS
YNDKTINYKDSTSVAGCPITARFRKMLGEEWESFRVEGQKERQAHSKNNISFHRVSYSIEDIWH
FCYDAEEPEAVLAFAQETLRLERKKAEELVRIWSAMPQGYAMLSQKAIRNINKILMLGLKYSDA
VILAKVPELVDVSDEELLSIAKDYYLVEAQVNYDKRINSIVNGLIAKYKSVSEEYRFADHNYEY
LLDESDEKDIIRQIENSLGARRWSLMDANEQTDILQKVRDRYQDFFRSHERKFVESPKLGESFE
NYLTKKFPMVEREQWKKLYHPSQITIYRPVSVGKDRSVLRLGNPDIGAIKNPTVLRVLNTLRRR
VNQLLDDGVISPDETRVVVETARELNDANRKWALDTYNRIRHDENEKIKKILEEFYPKRDGIST
DDIDKARYVIDQREVDYFTGSKTYNKDIKKYKFWLEQGGQCMYTGRTINLSNLFDPNAFDIEHT
IPESLSFDSSDMNLTLCDAHYNRFIKKNHIPTDMPNYDKAITIDGKEYPAITSQLQRWVERVER
LNRNVEYWKGQARRAQNKDRKDQCMREMHLWKMELEYWKKKLERFTVTEVTDGFKNSQLVDTRV
ITRHAVLYLKSIFPHVDVQRGDVTAKFRKILGIQSVDEKKDRSLHSHHAIDATTLTIIPVSAKR
DRMLELFAKIEEINKMLSFSGSEDRTGLIQELEGLKNKLQMEVKVCRIGHNVSEIGTFINDNII
VNHHIKNQALTPVRRRLRKKGYIVGGVDNPRWQTGDALRGEIHKASYYGAITQFAKDDEGKVLM
KEGRPQVNPTIKFVIRRELKYKKSAADSGFASWDDLGKAIVDKELFALMKGQFPAETSFKDACE
QGIYMIKKGKNGMPDIKLHHIRHVRCEAPQSGLKIKEQTYKSEKEYKRYFYAAVGDLYAMCCYT
NGKIREFRIYSLYDVSCHRKSDIEDIPEFITDKKGNRLMLDYKLRTGDMILLYKDNPAELYDLD
NVNLSRRLYKINRFESQSNLVLMTHHLSTSKERGRSLGKTVDYQNLPESIRSSVKSLNFLIMGE
NRDFVIKNGKIIFNHR

SEQ ID NO: 340

MLVSPISVDLGGKNTGFFSFTDSLDNSQSGTVIYDESFVLSQVGRRSKRHSKRNNLRNKLVKRL
FLLILQEHHGLSIDVLPDEIRGLFNKRGYTYAGFELDEKKKDALESDTLKEFLSEKLQSIDRDS
DVEDFLNQIASNAESFKDYKKGFEAVFASATHSPNKKLELKDELKSEYGENAKELLAGLRVTKE
ILDEFDKQENQGNLPRAKYFEELGEYIATNEKVKSFFDSNSLKLTDMTKLIGNISNYQLKELRR
YFNDKEMEKGDIWIPNKLHKITERFVRSWHPKNDADRQRRAELMKDLKSKEIMELLTTTEPVMT
IPPYDDMNNRGAVKCQTLRLNEEYLDKHLPNWRDIAKRLNHGKFNDDLADSTVKGYSEDSTLLH
RLLDTSKEIDIYELRGKKPNELLVKTLGQSDANRLYGFAQNYYELIRQKVRAGIWVPVKNKDDS
LNLEDNSNMLKRCNHNPPHKKNQIHNLVAGILGVKLDEAKFAEFEKELWSAKVGNKKLSAYCKN
IEELRKTHGNTFKIDIEELRKKDPAELSKEEKAKLRLTDDVILNEWSQKIANFFDIDDKHRQRF
NNLFSMAQLHTVIDTPRSGFSSTCKRCTAENRFRSETAFYNDETGEFHKKATATCQRLPADTQR
PFSGKIERYIDKLGYELAKIKAKELEGMEAKEIKVPIILEQNAFEYEESLRKSKTGSNDRVINS
KKDRDGKKLAKAKENAEDRLKDKDKRIKAFSSGICPYCGDTIGDDGEIDHILPRSHTLKIYGTV
FNPEGNLIYVHQKCNQAKADSIYKLSDIKAGVSAQWIEEQVANIKGYKTFSVLSAEQQKAFRYA
LFLQNDNEAYKKVVDWLRTDQSARVNGTQKYLAKKIQEKLTKMLPNKHLSFEFILADATEVSEL
RRQYARQNPLLAKAEKQAPSSHAIDAVMAFVARYQKVFKDGTPPNADEVAKLAMLDSWNPASNE
PLTKGLSTNQKIEKMIKSGDYGQKNMREVFGKSIFGENAIGERYKPIVVQEGGYYIGYPATVKK
GYELKNCKVVTSKNDIAKLEKIIKNQDLISLKENQYIKIFSINKQTISELSNRYFNMNYKNLVE
RDKEIVGLLEFIVENCRYYTKKVDVKFAPKYIHETKYPFYDDWRRFDEAWRYLQENQNKTSSKD
RFVIDKSSLNEYYQPDKNEYKLDVDTQPIWDDFCRWYFLDRYKTANDKKSIRIKARKTFSLLAE
SGVQGKVFRAKRKIPTGYAYQALPMDNNVIAGDYANILLEANSKTLSLVPKSGISIEKQLDKKL
DVIKKTDVRGLAIDNNSFFNADFDTHGIRLIVENTSVKVGNFPISAIDKSAKRMIFRALFEKEK
GKRKKKTTISFKESGPVQDYLKVFLKKIVKIQLRTDGSISNIVVRKNAADFTLSFRSEHIQKLL
K

SEQ ID NO: 341

MAYRLGLDIGITSVGWAVVALEKDESGLKPVRIQDLGVRIFDKAEDSKTGASLALPRREARSAR
RRTRRRRHRLWRVKRLLEQHGILSMEQIEALYAQRTSSPDVYALRVAGLDRCLIAEEIARVLIH
IAHRRGFQSNRKSEIKDSDAGKLLKAVQENENLMQSKGYRTVAEMLVSEATKTDAEGKLVHGKK
HGYVSNVRNKAGEYRHTVSRQAIVDEVRKIFAAQRALGNDVMSEELEDSYLKILCSQRNFDDGP
GGDSPYGHGSVSPDGVRQSIYERMVGSCTFETGEKRAPRSSYSFERFQLLTKVVNLRIYRQQED
GGRYPCELTQTERARVIDCAYEQTKITYGKLRKLLDMKDTESFAGLTYGLNRSRNKTEDTVFVE
MKFYHEVRKALQRAGVFIQDLSIETLDQIGWILSVWKSDDNRRKKLSTLGLSDNVIEELLPLNG
SKFGHLSLKAIRKILPFLEDGYSYDVACELAGYQFQGKTEYVKQRLLPPLGEGEVTNPVVRRAL
SQAIKVVNAVIRKHGSPESIHIELARELSKNLDERRKIEKAQKENQKNNEQIKDEIREILGSAH
VTGRDIVKYKLFKQQQEFCMYSGEKLDVTRLFEPGYAEVDHIIPYGISFDDSYDNKVLVKTEQN
RQKGNRTPLEYLRDKPEQKAKFIALVESIPLSQKKKNHLLMDKRAIDLEQEGFRERNLSDTRYI
TRALMNHIQAWLLFDETASTRSKRVVCVNGAVTAYMRARWGLTKDRDAGDKHHAADAVVVACIG
DSLIQRVTKYDKFKRNALADRNRYVQQVSKSEGITQYVDKETGEVFTWESFDERKFLPNEPLEP
WPFFRDELLARLSDDPSKNIRAIGLLTYSETEQIDPIFVSRMPTRKVTGAAHKETIRSPRIVKV
DDNKGTEIQVVVSKVALTELKLTKDGEIKDYFRPEDDPRLYNTLRERLVQFGGDAKAAFKEPVY
KISKDGSVRTPVRKVKIQEKLTLGVPVHGGRGIAENGGMVRIDVFAKGGKYYFVPIYVADVLKR
ELPNRLATAHKPYSEWRVVDDSYQFKFSLYPNDAVMIKPSREVDITYKDRKEPVGCRIMYFVSA
NIASASISLRTHDNSGELEGLGIQGLEVFEKYVVGPLGDTHPVYKERRMPFRVERKMN

SEQ ID NO: 342
MPVLSPLSPNAAQGRRRWSLALDIGEGSIGWAVAEVDAEGRVLQLTGTGVTLFPSAWSNENGTY
VAHGAADRAVRGQQQRHDSRRRRLAGLARLCAPVLERSPEDLKDLTRTPPKADPRAIFFLRADA
ARRPLDGPELFRVLHHMAAHRGIRLAELQEVDPPPESDADDAAPAATEDEDGTRRAAADERAFR
RLMAEHMHRHGTQPTCGEIMAGRLRETPAGAQPVTRARDGLRVGGGVAVPTRALIEQEFDAIRA
IQAPRHPDLPWDSLRRLVLDQAPIAVPPATPCLFLEELRRRGETFQGRTITREAIDRGLTVDPL
IQALRIRETVGNLRLHERITEPDGRQRYVPRAMPELGLSHGELTAPERDTLVRALMHDPDGLAA
KDGRIPYTRLRKLIGYDNSPVCFAQERDTSGGGITVNPTDPLMARWIDGWVDLPLKARSLYVRD
VVARGADSAALARLLAEGAHGVPPVAAAAVPAATAAILESDIMQPGRYSVCPWAAEAILDAWAN
APTEGFYDVTRGLFGFAPGEIVLEDLRRARGALLAHLPRTMAAARTPNRAAQQRGPLPAYESVI
PSQLITSLRRAHKGRAADWSAADPEERNPFLRTWTGNAATDHILNQVRKTANEVITKYGNRRGW
DPLPSRITVELAREAKHGVIRRNEIAKENRENEGRRKKESAALDTFCQDNTVSWQAGGLPKERA
ALRLRLAQRQEFFCPYCAERPKLRATDLFSPAETEIDHVIERRMGGDGPDNLVLAHKDCNNAKG
KKTPHEHAGDLLDSPALAALWQGWRKENADRLKGKGHKARTPREDKDFMDRVGWRFEEDARAKA
EENQERRGRRMLHDTARATRLARLYLAAAVMPEDPAEIGAPPVETPPSPEDPTGYTAIYRTISR
VQPVNGSVTHMLRQRLLQRDKNRDYQTHHAEDACLLLLAGPAVVQAFNTEAAQHGADAPDDRPV
DLMPTSDAYHQQRRARALGRVPLATVDAALADIVMPESDRQDPETGRVHWRLTRAGRGLKRRID
DLTRNCVILSRPRRPSETGTPGALHNATHYGRREITVDGRTDTVVTQRMNARDLVALLDNAKIV
PAARLDAAAPGDTILKEICTEIADRHDRVVDPEGTHARRWISARLAALVPAHAEAVARDIAELA
DLDALADADRTPEQEARRSALRQSPYLGRAISAKKADGRARAREQEILTRALLDPHWGPRGLRH
LIMREARAPSLVRIRANKTDAFGRPVPDAAVWVKTDGNAVSQLWRLTSVVTDDGRRIPLPKPIE
KRIEISNLEYARLNGLDEGAGVTGNNAPPRPLRQDIDRLTPLWRDHGTAPGGYLGTAVGELEDK
ARSALRGKAMRQTLTDAGITAEAGWRLDSEGAVCDLEVAKGDTVKKDGKTYKVGVITQGIFGMP
VDAAGSAPRTPEDCEKFEEQYGIKPWKAKGIPLA

SEQ ID NO: 343
MNYTEKEKLFMKYILALDIGIASVGWAILDKESETVIEAGSNIFPEASAADNQLRRDMRGAKRN
NRRLKTRINDFIKLWENNNLSIPQFKSTEIVGLKVRAITEEITLDELYLILYSYLKHRGISYLE

DALDDTVSGSSAYANGLKLNAKELETHYPCEIQQERLNTIGKYRGQSQIINENGEVLDLSNVFT
IGAYRKEIQRVFEIQKKYHPELTDEFCDGYMLIFNRKRKYYEGPGNEKSRTDYGRFTTKLDANG
NYITEDNIFEKLIGKCSVYPDELRAAAASYTAQEYNVLNDLNNLTINGRKLEENEKHEIVERIK
SSNTINMRKIISDCMGENIDDFAGARIDKSGKEIFHKFEVYNKMRKALLEIGIDISNYSREELD
EIGYIMTINTDKEAMMEAFQKSWIDLSDDVKQCLINMRKTNGALFNKWQSFSLKIMNELIPEMY
AQPKEQMTLLTEMGVTKGTQEEFAGLKYIPVDVVSEDIFNPVVRRSVRISFKILNAVLKKYKAL
DTIVIEMPRDRNSEEQKKRINDSQKLNEKEMEYIEKKLAVTYGIKLSPSDFSSQKQLSLKLKLW
NEQDGICLYSGKTIDPNDIINNPQLFEIDHIIPRSISFDDARSNKVLVYRSENQKKGNQTPYYY
LTHSHSEWSFEQYKATVMNLSKKKEYAISRKKIQNLLYSEDITKMDVLKGFINRNINDTSYASR
LVLNTIQNFFMANEADTKVKVIKGSYTHQMRCNLKLDKNRDESYSHHAVDAMLIGYSELGYEAY
HKLQGEFIDFETGEILRKDMWDENMSDEVYADYLYGKKWANIRNEVVKAEKNVKYWHYVMRKSN
RGLCNQTIRGTREYDGKQYKINKLDIRTKEGIKVFAKLAFSKKDSDRERLLVYLNDRRTFDDLC
KIYEDYSDAANPFVQYEKETGDIIRKYSKKHNGPRIDKLKYKDGEVGACIDISHKYGFEKGSKK
VILESLVPYRMDVYYKEENHSYYLVGVKQSDIKFEKGRNVIDEEAYARILVNEKMIQPGQSRAD
LENLGFKFKLSFYKNDIIEYEKDGKIYTERLVSRTMPKQRNYIETKPIDKAKFEKQNLVGLGKT
KFIKKYRYDILGNKYSCSEEKFTSFC

SEQ ID NO: 344
MLRLYCANNLVLNNVQNLWKYLLLLIFDKKIIFLFKIKVILIRRYMENNNKEKIVIGFDLGVAS
VGWSIVNAETKEVIDLGVRLFSEPEKADYRRAKRTTRRLLRRKKFKREKFHKLILKNAEIFGLQ
SRNEILNVYKDQSSKYRNILKLKINALKEEIKPSELVWILRDYLQNRGYFYKNEKLTDEFVSNS
FPSKKLHEHYEKYGFFRGSVKLDNKLDNKKDKAKEKDEEEESDAKKESEELIFSNKQWINEIVK
VFENQSYLTESFKEEYLKLFNYVRPFNKGPGSKNSRTAYGVFSTDIDPETNKFKDYSNIWDKTI
GKCSLFEEEIRAPKNLPSALIFNLQNEICTIKNEFTEFKNWWLNAEQKSEILKFVFTELFNWKD
KKYSDKKFNKNLQDKIKKYLLNFALENFNLNEEILKNRDLENDTVLGLKGVKYYEKSNATADAA
LEFSSLKPLYVFIKFLKEKKLDLNYLLGLENTEILYFLDSIYLAISYSSDLKERNEWFKKLLKE
LYPKIKNNNLEIIENVEDIFEITDQEKFESFSKTHSLSREAFNHIIPLLLSNNEGKNYESLKHS
NEELKKRTEKAELKAQQNQKYLKDNFLKEALVPLSVKTSVLQAIKIFNQIIKNFGKKYEISQVV
IEMARELTKPNLEKLLNNATNSNIKILKEKLDQTEKFDDFTKKKFIDKIENSVVFRNKLFLWFE
QDRKDPYTQLDIKINEIEDETEIDHVIPYSKSADDSWFNKLLVKKSTNQLKKNKTVWEYYQNES
DPEAKWNKFVAWAKRIYLVQKSDKESKDNSEKNSIFKNKKPNLKFKNITKKLFDPYKDLGFLAR
NLNDTRYATKVFRDQLNNYSKHHSKDDENKLFKVVCMNGSITSFLRKSMWRKNEEQVYRFNFWK
KDRDQFFHHAVDASIIAIFSLLTKTLYNKLRVYESYDVQRREDGVYLINKETGEVKKADKDYWK
DQHNFLKIRENAIEIKNVLNNVDFQNQVRYSRKANTKLNTQLFNETLYGVKEFENNFYKLEKVN
LFSRKDLRKFILEDLNEESEKNKKNENGSRKRILTEKYIVDEILQILENEEFKDSKSDINALNK
YMDSLPSKFSEFFSQDFINKCKKENSLILTFDAIKHNDPKKVIKIKNLKFFREDATLKNKQAVH
KDSKNQIKSFYESYKCVGFIWLKNKNDLEESIFVPINSRVIHFGDKDKDIFDFDSYNKEKLLNE
INLKRPENKKFNSINEIEFVKFVKPGALLLNFENQQIYYISTLESSSLRAKIKLLNKMDKGKAV
SMKKITNPDEYKIIEHVNPLGINLNWTKKLENNN

SEQ ID NO: 345
MLMSKHVLGLDLGVGSIGWCLIALDAQGDPAEILGMGSRVVPLNNATKAIEAFNAGAAFTASQE
RTARRTMRRGFARYQLRRYRLRRELEKVGMLPDAALIQLPLLELWELRERAATAGRRLTLPELG
RVLCHINQKRGYRHVKSDAAAIVGDEGEKKKDSNSAYLAGIRANDEKLQAEHKTVGQYFAEQLR
QNQSESPTGGISYRIKDQIFSRQCYIDEYDQIMAVQRVHYPDILTDEFIRMLRDEVIFMQRPLK
SCKHLVSLCEFEKQERVMRVQQDDGKGGWQLVERRVKFGPKVAPKSSPLFQLCCIYEAVNNIRL
TRPNGSPCDITPEERAKIVAHLQSSASLSFAALKKLLKEKALIADQLTSKSGLKGNSTRVALAS

ALQPYPQYHHLLDMELETRMMTVQLTDEETGEVTEREVAVVTDSYVRKPLYRLWHILYSIEERE
AMRRALITQLGMKEEDLDGGLLDQLYRLDFVKPGYGNKSAKFICKLLPQLQQGLGYSEACAAVG
YRHSNSPTSEEITERTLLEKIPLLQRNELRQPLVEKILNQMINLVNALKAEYGIDEVRVELARE
LKMSREERERMARNNKDREERNKGVAAKIRECGLYPTKPRIQKYMLWKEAGRQCLYCGRSIEEE
QCLREGGMEVEHIIPKSVLYDDSYGNKTCACRRCNKEKGNRTALEYIRAKGREAEYMKRINDLL
KEKKISYSKHQRLRWLKEDIPSDFLERQLRLTQYISRQAMAILQQGIRRVSASEGGVTARLRSL
WGYGKILHTLNLDRYDSMGETERVSREGEATEELHITNWSKRMDHRHHAIDALVVACTRQSYIQ
RLNRLSSEFGREDKKKEDQEAQEQQATETGRLSNLERWLTQRPHFSVRTVSDKVAEILISYRPG
QRVVTRGRNIYRKKMADGREVSCVQRGVLVPRGELMEASFYGKILSQGRVRIVKRYPLHDLKGE
VVDPHLRELITTYNQELKSREKGAPIPPLCLDKDKKQEVRSVRCYAKTLSLDKAIPMCFDEKGE
PTAFVKSASNHHLALYRTPKGKLVESIVTFWDAVDRARYGIPLVITHPREVMEQVLQRGDIPEQ
VLSLLPPSDWVFVDSLQQDEMVVIGLSDEELQRALEAQNYRKISEHLYRVQKMSSSYYVFRYHL
ETSVADDKNTSGRIPKFHRVQSLKAYEERNIRKVRVDLLGRISLL

SEQ ID NO: 346
MSDLVLGLDIGIGSVGVGILNKVTGEIIHKNSRIFPAAQAENNLVRRTNRQGRRLARRKKHRRV
RLNRLFEESGLITDFTKISINLNPYQLRVKGLTDELSNEELFIALKNMVKHRGISYLDDASDDG
NSSVGDYAQIVKENSKQLETKTPGQIQLERYQTYGQLRGDFTVEKDGKKHRLINVFPTSAYRSE
ALRILQTQQEFNPQITDEFINRYLEILTGKRKYYHGPGNEKSRTDYGRYRTSGETLDNIFGILI
GKCTFYPDEFRAAKASYTAQEFNLLNDLNNLTVPTETKKLSKEQKNQIINYVKNEKAMGPAKLF
KYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQMDRETLDKLAYVLTLNTE
REGIQEALEHEFADGSFSQKQVDELVQFRKANSSIFGKGWHNFSVKLMMELIPELYETSEEQMT
ILTRLGKQKTTSSSNKTKYIDEKLLTEEIYNPVVAKSVRQAIKIVNAAIKEYGDFDNIVIEMAR
ETNEDDEKKAIQKIQKANKDEKDAAMLKAANQYNGKAELPHSVFHGHKQLATKIRLWHQQGERC
LYTGKTISIHDLINNSNQFEVDHILPLSITFDDSLANKVLVYATANQEKGQRTPYQALDSMDDA
WSFRELKAFVRESKTLSNKKKEYLLTEEDISKFDVRKKFIERNLVDTRYASRVVLNALQEHFRA
HKIDTKVSVVRGQFTSQLRRHWGIEKTRDTYHHHAVDALIIAASSQLNLWKKQKNTLVSYSEDQ
LLDIETGELISDDEYKESVFKAPYQHFVDTLKSKEFEDSILFSYQVDSKFNRKISDATIYATRQ
AKVGKDKADETYVLGKIKDIYTQDGYDAFMKIYKKDKSKFLMYRHDPQTFEKVIEPILENYPNK
QINEKGKEVPCNPFLKYKEEHGYIRKYSKKGNGPEIKSLKYYDSKLGNHIDITPKDSNNKVVLQ
SVSPWRADVYFNKTTGKYEILGLKYADLQFEKGTGTYKISQEKYNDIKKKEGVDSDSEFKFTLY
KNDLLLVKDTETKEQQLFRFLSRTMPKQKHYVELKPYDKQKFEGGEALIKVLGNVANSGQCKKG
LGKSNISIYKVRTDVLGNQHIIKNEGDKPKLDF

SEQ ID NO: 347
MNAEHGKEGLLIMEENFQYRIGLDIGITSVGWAVLQNNSQDEPVRITDLGVRIFDVAENPKNGD
ALAAPRRDARTTRRRLRRRRHRLERIKFLLQENGLIEMDSFMERYYKGNLPDVYQLRYEGLDRK
LKDEELAQVLIHIAKHRGFRSTRKAETKEKEGGAVLKATTENQKIMQEKGYRTVGEMLYLDEAF
HTECLWNEKGYVLTPRNRPDDYKHTILRSMLVEEVHAIFAAQRAHGNQKATEGLEEAYVEIMTS
QRSFDMGPGLQPDGKPSPYAMEGFGDRVGKCTFEKDEYRAPKATYTAELFVALQKINHTKLIDE
FGTGRFFSEEERKTIIGLLLSSKELKYGTIRKKLNIDPSLKFNSLNYSAKKEGETEEERVLDTE
KAKFASMFWTYEYSKCLKDRTEEMPVGEKADLFDRIGEILTAYKNDDSRSSRLKELGLSGEEID
GLLDLSPAKYQRVSLKAMRKMQPYLEDGLIYDKACEAAGYDFRALNDGNKKHLLKGEEINAIVN
DITNPVVKRSVSQTIKVINAIIQKYGSPQAVNIELAREMSKNFQDRTNLEKEMKKRQQENERAK
QQIIELGKQNPTGQDILKYRLWNDQGGYCLYSGKKIPLEELFDGGYDIDHILPYSITFDDSYRN
KVLVTAQENRQKGNRTPYEYFGADEKRWEDYEASVRLLVRDYKKQQKLLKKNFTEEERKEFKER
NLNDTKYITRVVYNMIRQNLELEPFNHPEKKKQVWAVNGAVTSYLRKRWGLMQKDRSTDRHHAM

DAVVIACCTDGMIHKISRYMQGRELAYSRNFKFPDEETGEILNRDNFTREQWDEKFGVKVPLPW
NSFRDELDIRLLNEDPKNFLLTHADVQRELDYPGWMYGEEESPIEEGRYINYIRPLFVSRMPNH
KVTGSAHDATIRSARDYETRGVVITKVPLTDLKLNKDNEIEGYYDKDSDRLLYQALVRQLLLHG
NDGKKAFAEDFHKPKADGTEGPVVRKVKIEKKQTSGVMVRGGTGIAANGEMVRIDVFRENGKYY
FVPVYTADVVRKVLPNRAATHTKPYSEWRVMDDANFVFSLYSRDLIHVKSKKDIKTNLVNGGLL
LQKEIFAYYTGADIATASIAGFANDSNFKFRGLGIQSLEIFEKCQVDILGNISVVRHENRQEFH

SEQ ID NO: 348
MRVLGLDAGIASLGWALIEIEESNRGELSQGTIIGAGTWMFDAPEEKTQAGAKLKSEQRRTFRG
QRRVVRRRRQRMNEVRRILHSHGLLPSSDRDALKQPGLDPWRIRAEALDRLLGPVELAVALGHI
ARHRGFKSNSKGAKTNDPADDTSKMKRAVNETREKLARFGSAAKMLVEDESFVLRQTPTKNGAS
EIVRRFRNREGDYSRSLLRDDLAAEMRALFTAQARFQSAIATADLQTAFTKAAFFQRPLQDSEK
LVGPCPFEVDEKRAPKRGYSFELFRFLSRLNHVTLRDGKQERTLTRDELALAAADFGAAAKVSF
TALRKKLKLPETTVFVGVKADEESKLDVVARSGKAAEGTARLRSVIVDALGELAWGALLCSPEK
LDKIAEVISFRSDIGRISEGLAQAGCNAPLVDALTAAASDGRFDPFTGAGHISSKAARNILSGL
RQGMTYDKACCAADYDHTASRERGAFDVGGHGREALKRILQEERISRELVGSPTARKALIESIK
QVKAIVERYGVPDRIHVELARDVGKSIEEREEITRGIEKRNRQKDKLRGLFEKEVGRPPQDGAR
GKEELLRFELWSEQMGRCLYTDDYISPSQLVATDDAVQVDHILPWSRFADDSYANKTLCMAKAN
QDKKGRTPYEWFKAEKTDTEWDAFIVRVEALADMKGFKKRNYKLRNAEEAAAKFRNRNLNDTRW
ACRLLAEALKQLYPKGEKDKDGKERRRVFSRPGALTDRLRRAWGLQWMKKSTKGDRIPDDRHHA
LDAIVIAATTESLLQRATREVQEIEDKGLHYDLVKNVTPPWPGFREQAVEAVEKVFVARAERRR
ARGKAHDATIRHIAVREGEQRVYERRKVAELKLADLDRVKDAERNARLIEKLRNWIEAGSPKDD
PPLSPKGDPIFKVRLVTKSKVNIALDTGNPKRPGTVDRGEMARVDVFRKASKKGKYEYYLVPIY
PHDIATMKTPPIRAVQAYKPEDEWPEMDSSYEFCWSLVPMTYLQVISSKGEIFEGYYRGMNRSV
GAIQLSAHSNSSDVVQGIGARTLTEFKKFNVDRFGRKHEVERELRTWRGETWRGKAYI

SEQ ID NO: 349
MGNYYLGLDVGIGSIGWAVINIEKKRIEDFNVRIFKSGEIQEKNRNSRASQQCRRSRGLRRLYR
RKSHRKLRLKNYLSIIGLTTSEKIDYYYETADNNVIQLRNKGLSEKLTPEEIAACLIHICNNRG
YKDFYEVNVEDIEDPDERNEYKEEHDSIVLISNLMNEGGYCTPAEMICNCREFDEPNSVYRKFH
NSAASKNHYLITRHMLVKEVDLILENQSKYYGILDDKTIAKIKDIIFAQRDFEIGPGKNERFRR
FTGYLDSIGKCQFFKDQERGSRFTVIADIYAFVNVLSQYTYTNNRGESVFDTSFANDLINSALK
NGSMDKRELKAIAKSYHIDISDKNSDTSLTKCFKYIKVVKPLFEKYGYDWDKLIENYTDTDNNV
LNRIGIVLSQAQTPKRRREKLKALNIGLDDGLINELTKLKLSGTANVSYKYMQGSIEAFCEGDL
YGKYQAKFNKEIPDIDENAKPQKLPPFKNEDDCEFFKNPVVFRSINETRKLINAIIDKYGYPAA
VNIETADELNKTFEDRAIDTKRNNDNQKENDRIVKEIIECIKCDEVHARHLIEKYKLWEAQEGK
CLYSGETITKEDMLRDKDKLFEVDHIVPYSLILDNTINNKALVYAEENQKKGQRTPLMYMNEAQ
AADYRVRVNTMFKSKKCSKKKYQYLMLPDLNDQELLGGWRSRNLNDTRYICKYLVNYLRKNLRF
DRSYESSDEDDLKIRDHYRVFPVKSRFTSMFRRWWLNEKTWGRYDKAELKKLTYLDHAADAIII
ANCRPEYVVLAGEKLKLNKMYHQAGKRITPEYEQSKKACIDNLYKLFRMDRRTAEKLLSGHGRL
TPIIPNLSEEVDKRLWDKNIYEQFWKDDKDKKSCEELYRENVASLYKGDPKFASSLSMPVISLK
PDHKYRGTITGEEAIRVKEIDGKLIKLKRKSISEITAESINSIYTDDKILIDSLKTIFEQADYK
DVGDYLKKTNQHFFTTSSGKRVNKVTVIEKVPSRWLRKEIDDNNFSLLNDSSYYCIELYKDSKG
DNNLQGIAMSDIVHDRKTKKLYLKPDFNYPDDYYTHVMYIFPGDYLRIKSTSKKSGEQLKFEGY
FISVKNVNENSFRFISDNKPCAKDKRVSITKKDIVIKLAVDLMGKVQGENNGKGISCGEPLSLL
KEKN

SEQ ID NO: 350

MLSRQLLGASHLARPVSYSYNVQDNDVHCSYGERCFMRGKRYRIGIDVGLNSVGLAAVEVSDEN
SPVRLLNAQSVIHDGGVDPQKNKEAITRKNMSGVARRTRRMRRRKRERLHKLDMLLGKFGYPVI
EPESLDKPFEEWHVRAELATRYIEDDELRRESISIALRHMARHRGWRNPYRQVDSLISDNPYSK
QYGELKEKAKAYNDDATAAEEESTPAQLVVAMLDAGYAEAPRLRWRTGSKKPDAEGYLPVRLMQ
EDNANELKQIFRVQRVPADEWKPLFRSVFYAVSPKGSAEQRVGQDPLAPEQARALKASLAFQEY
RIANVITNLRIKDASAELRKLTVDEKQSIYDQLVSPSSEDITWSDLCDFLGFKRSQLKGVGSLT
EDGEERISSRPPRLTSVQRIYESDNKIRKPLVAWWKSASDNEHEAMIRLLSNTVDIDKVREDVA
YASAIEFIDGLDDDALTKLDSVDLPSGRAAYSVETLQKLTRQMLTTDDDLHEARKTLFNVTDSW
RPPADPIGEPLGNPSVDRVLKNVNRYLMNCQQRWGNPVSVNIEHVRSSFSSVAFARKDKREYEK
NNEKRSIFRSSLSEQLRADEQMEKVRESDLRRLEAIQRQNGQCLYCGRTITFRTCEMDHIVPRK
GVGSTNTRTNFAAVCAECNRMKSNTPFAIWARSEDAQTRGVSLAEAKKRVTMFTFNPKSYAPRE
VKAFKQAVIARLQQTEDDAAIDNRSIESVAWMADELHRRIDWYFNAKQYVNSASIDDAEAETMK
TTVSVFQGRVTASARRAAGIEGKIHFIGQQSKTRLDRRHHAVDASVIAMMNTAAAQTLMERESL
RESQRLIGLMPGERSWKEYPYEGTSRYESFHLWLDNMDVLLELLNDALDNDRIAVMQSQRYVLG
NSIAHDATIHPLEKVPLGSAMSADLIRRASTPALWCALTRLPDYDEKEGLPEDSHREIRVHDTR
YSADDEMGFFASQAAQIAVQEGSADIGSAIHHARVYRCWKTNAKGVRKYFYGMIRVFQTDLLRA
CHDDLFTVPLPPQSISMRYGEPRVVQALQSGNAQYLGSLVVGDEIEMDFSSLDVDGQIGEYLQF
FSQFSGGNLAWKHWVVDGFFNQTQLRIRPRYLAAEGLAKAFSDDVVPDGVQKIVTKQGWLPPVN
TASKTAVRIVRRNAFGEPRLSSAHHMPCSWQWRHE

SEQ ID NO: 351

MYSIGLDLGISSVGWSVIDERTGNVIDLGVRLFSAKNSEKNLERRTNRGGRRLIRRKTNRLKDA
KKILAAVGFYEDKSLKNSCPYQLRVKGLTEPLSRGEIYKVTLHILKKRGISYLDEVDTEAAKES
QDYKEQVRKNAQLLTKYTPGQIQLQRLKENNRVKTGINAQGNYQLNVFKVSAYANELATILKTQ
QAFYPNELTDDWIALFVQPGIAEEAGLIYRKRPYYHGPGNEANNSPYGRWSDFQKTGEPATNIF
DKLIGKDFQGELRASGLSLSAQQYNLLNDLTNLKIDGEVPLSSEQKEYILTELMTKEFTRFGVN
DVVKLLGVKKERLSGWRLDKKGKPEIHTLKGYRNWRKIFAEAGIDLATLPTETIDCLAKVLTLN
TEREGIENTLAFELPELSESVKLLVLDRYKELSQSISTQSWHRFSLKTLHLLIPELMNATSEQN
TLLEQFQLKSDVRKRYSEYKKLPTKDVLAEIYNPTVNKTVSQAFKVIDALLVKYGKEQIRYITI
EMPRDDNEEDEKKRIKELHAKNSQRKNDSQSYFMQKSGWSQEKFQTTIQKNRRFLAKLLYYYEQ
DGICAYTGLPISPELLVSDSTEIDHIIPISISLDDSINNKVLVLSKANQVKGQQTPYDAWMDGS
FKKINGKFSNWDDYQKWVESRHFSHKKENNLLETRNIFDSEQVEKFLARNLNDTRYASRLVLNT
LQSFFTNQETKVRVVNGSFTHTLRKKWGADLDKTRETHHHHAVDATLCAVTSFVKVSRYHYAVK
EETGEKVMREIDFETGEIVNEMSYWEFKKSKKYERKTYQVKWPNFREQLKPVNLHPRIKFSHQV
DRKANRKLSDATIYSVREKTEVKTLKSGKQKITTDEYTIGKIKDIYTLDGWEAFKKKQDKLLMK
DLDEKTYERLLSIAETTPDFQEVEEKNGKVKRVKRSPFAVYCEENDIPAIQKYAKKNNGPLIRS
LKYYDGKLNKHINITKDSQGRPVEKTKNGRKVTLQSLKPYRYDIYQDLETKAYYTVQLYYSDLR
FVEGKYGITEKEYMKKVAEQTKGQVVRFCFSLQKNDGLEIEWKDSQRYDVRFYNFQSANSINFK
GLEQEMMPAENQFKQKPYNNGAINLNIAKYGKEGKKLRKFNTDILGKKHYLFYEKEPKNIIK

SEQ ID NO: 352

MYFYKNKENKLNKKVVLGLDLGIASVGWCLTDISQKEDNKFPIILHGVRLFETVDDSDDKLLNE
TRRKKRGQRRRNRRLFTRKRDFIKYLIDNNIIELEFDKNPKILVRNFIEKYINPFSKNLELKYK
SVTNLPIGFHNLRKAAINEKYKLDKSELIVLLYFYLSLRGAFFDNPEDTKSKEMNKNEIEIFDK
NESIKNAEFPIDKIIEFYKISGKIRSTINLKFGHQDYLKEIKQVFEKQNIDFMNYEKFAMEEKS
FFSRIRNYSEGPGNEKSFSKYGLYANENGNPELIINEKGQKIYTKIFKTLWESKIGKCSYDKKL

YRAPKNSFSAKVFDITNKLTDWKHKNEYISERLKRKILLSRFLNKDSKSAVEKILKEENIKFEN
LSEIAYNKDDNKINLPIINAYHSLTTIFKKHLINFENYLISNENDLSKLMSFYKQQSEKLFVPN
EKGSYEINQNNNVLHIFDAISNILNKFSTIQDRIRILEGYFEFSNLKKDVKSSEIYSEIAKLRE
FSGTSSLSFGAYYKFIPNLISEGSKNYSTISYEEKALQNQKNNFSHSNLFEKTWVEDLIASPTV
KRSLRQTMNLLKEIFKYSEKNNLEIEKIVVEVTRSSNNKHERKKIEGINKYRKEKYEELKKVYD
LPNENTTLLKKLWLLRQQQGYDAYSLRKIEANDVINKPWNYDIDHIVPRSISFDDSFSNLVIVN
KLDNAKKSNDLSAKQFIEKIYGIEKLKEAKENWGNWYLRNANGKAFNDKGKFIKLYTIDNLDEF
DNSDFINRNLSDTSYITNALVNHLTFSNSKYKYSVVSVNGKQTSNLRNQIAFVGIKNNKETERE
WKRPEGFKSINSNDFLIREEGKNDVKDDVLIKDRSFNGHHAEDAYFITIISQYFRSFKRIERLN
VNYRKETRELDDLEKNNIKFKEKASFDNFLLINALDELNEKLNQMRFSRMVITKKNTQLFNETL
YSGKYDKGKNTIKKVEKLNLLDNRTDKIKKIEEFFDEDKLKENELTKLHIFNHDKNLYETLKII
WNEVKIEIKNKNLNEKNYFKYFVNKKLQEGKISFNEWVPILDNDFKIIRKIRYIKFSSEEKETD
EIIFSQSNFLKIDQRQNFSFHNTLYWVQIWVYKNQKDQYCFISIDARNSKFEKDEIKINYEKLK
TQKEKLQIINEEPILKINKGDLFENEEKELFYIVGRDEKPQKLEIKYILGKKIKDQKQIQKPVK
KYFPNWKKVNLTYMGEIFKK

SEQ ID NO: 353
MDNKNYRIGIDVGLNSIGFCAVEVDQHDTPLGFLNLSVYRHDAGIDPNGKKTNTTRLAMSGVAR
RTRRLFRKRKRRLAALDRFIEAQGWTLPDHADYKDPYTPWLVRAELAQTPIRDENDLHEKLAIA
VRHIARHRGWRSPWVPVRSLHVEQPPSDQYLALKERVEAKTLLQMPEGATPAEMVVALDLSVDV
NLRPKNREKTDTRPENKKPGFLGGKLMQSDNANELRKIAKIQGLDDALLRELIELVFAADSPKG
ASGELVGYDVLPGQHGKRRAEKAHPAFQRYRIASIVSNLRIRHLGSGADERLDVETQKRVFEYL
LNAKPTADITWSDVAEEIGVERNLLMGTATQTADGERASAKPPVDVTNVAFATCKIKPLKEWWL
NADYEARCVMVSALSHAEKLTEGTAAEVEVAEFLQNLSDEDNEKLDSFSLPIGRAAYSVDSLER
LTKRMIENGEDLFEARVNEFGVSEDWRPPAEPIGARVGNPAVDRVLKAVNRYLMAAEAEWGAPL
SVNIEHVREGFISKRQAVEIDRENQKRYQRNQAVRSQIADHINATSGVRGSDVTRYLAIQRQNG
ECLYCGTAITFVNSEMDHIVPRAGLGSTNTRDNLVATCERCNKSKSNKPFAVWAAECGIPGVSV
AEALKRVDFWIADGFASSKEHRELQKGVKDRLKRKVSDPEIDNRSMESVAWMARELAHRVQYYF
DEKHTGTKVRVFRGSLTSAARKASGFESRVNFIGGNGKTRLDRRHHAMDAATVAMLRNSVAKTL
VLRGNIRASERAIGAAETWKSFRGENVADRQIFESWSENMRVLVEKFNLALYNDEVSIFSSLRL
QLGNGKAHDDTITKLQMHKVGDAWSLTEIDRASTPALWCALTRQPDFTWKDGLPANEDRTIIVN
GTHYGPLDKVGIFGKAAASLLVRGGSVDIGSAIHHARIYRIAGKKPTYGMVRVFAPDLLRYRNE
DLFNVELPPQSVSMRYAEPKVREAIREGKAEYLGWLVVGDELLLDLSSETSGQIAELQQDFPGT
THWTVAGFFSPSRLRLRPVYLAQEGLGEDVSEGSKSIIAGQGWRPAVNKVFGSAMPEVIRRDGL
GRKRRFSYSGLPVSWQG

SEQ ID NO: 354
MRLGLDIGTSSIGWWLYETDGAGSDARITGVVDGGVRIFSDGRDPKSGASLAVDRRAARAMRRR
RDRYLRRRATLMKVLAETGLMPADPAEAKALEALDPFALRAAGLDEPLPLPHLGRALFHLNQRR
GFKSNRKTDRGDNESGKIKDATARLDMEMMANGARTYGEFLHKRRQKATDPRHVPSVRTRLSIA
NRGGPDGKEEAGYDFYPDRRHLEEEFHKLWAAQGAHHPELTETLRDLLFEKIFFQRPLKEPEVG
LCLFSGHHGVPPKDPRLPKAHPLTQRRVLYETVNQLRVTADGREARPLTREERDQVIHALDNKK
PTKSLSSMVLKLPALAKVLKLRDGERFTLETGVRDAIACDPLRASPAHPDRFGPRWSILDADAQ
WEVISRIRRVQSDAEHAALVDWLTEAHGLDRAHAEATAHAPLPDGYGRLGLTATTRILYQLTAD
VVTYADAVKACGWHHSDGRTGECFDRLPYYGEVLERHVIPGSYHPDDDDITRFGRITNPTVHIG
LNQLRRLVNRIIETHGKPHQIVVELARDLKKSEEQKRADIKRIRDTTEAAKKRSEKLEELEIED
NGRNRMLLRLWEDLNPDDAMRRFCPYTGTRISAAMIFDGSCDVDHILPYSRTLDDSFPNRTLCL

REANRQKRNQTPWQAWGDTPHWHAIAANLKNLPENKRWRFAPDAMTRFEGENGFLDRALKDTQY
LARISRSYLDTLFTKGGHVWVVPGRFTEMLRRHWGLNSLLSDAGRGAVKAKNRTDHRHHAIDAA
VIAATDPGLLNRISRAAGQGEAAGQSAELIARDTPPPWEGFRDDLRVRLDRIIVSHRADHGRID
HAARKQGRDSTAGQLHQETAYSIVDDIHVASRTDLLSLKPAQLLDEPGRSGQVRDPQLRKALRV
ATGGKTGKDFENALRYFASKPGPYQAIRRVRIIKPLQAQARVPVPAQDPIKAYQGGSNHLFEIW
RLPDGEIEAQVITSFEAHTLEGEKRPHPAAKRLLRVHKGDMVALERDGRRVVGHVQKMDIANGL
FIVPHNEANADTRNNDKSDPFKWIQIGARPAIASGIRRVSVDEIGRLRDGGTRPI

SEQ ID NO: 355
MLHCIAVIRVPPSEEPGFFETHADSCALCHHGCMTYAANDKAIRYRVGIDVGLRSIGFCAVEVD
DEDHPIRILNSVVHVHDAGTGGPGETESLRKRSGVAARARRRGRAEKQRLKKLDVLLEELGWGV
SSNELLDSHAPWHIRKRLVSEYIEDETERRQCLSVAMAHIARHRGWRNSFSKVDTLLLEQAPSD
RMQGLKERVEDRTGLQFSEEVTQGELVATLLEHDGDVTIRGFVRKGGKATKVHGVLEGKYMQSD
LVAELRQICRTQRVSETTFEKLVLSIFHSKEPAPSAARQRERVGLDELQLALDPAAKQPRAERA
HPAFQKFKVVATLANMRIREQSAGERSLTSEELNRVARYLLNHTESESPTWDDVARKLEVPRHR
LRGSSRASLETGGGLTYPPVDDTTVRVMSAEVDWLADWWDCANDESRGHMIDAISNGCGSEPDD
VEDEEVNELISSATAEDMLKLELLAKKLPSGRVAYSLKTLREVTAAILETGDDLSQAITRLYGV
DPGWVPTPAPIEAPVGNPSVDRVLKQVARWLKFASKRWGVPQTVNIEHTREGLKSASLLEEERE
RWERFEARREIRQKEMYKRLGISGPFRRSDQVRYEILDLQDCACLYCGNEINFQTFEVDHIIPR
VDASSDSRRTNLAAVCHSCNSAKGGLAFGQWVKRGDCPSGVSLENAIKRVRSWSKDRLGLTEKA
MGKRKSEVISRLKTEMPYEEFDGRSMESVAWMAIELKKRIEGYFNSDRPEGCAAVQVNAYSGRL
TACARRAAHVDKRVRLIRLKGDDGHHKNRFDRRNHAMDALVIALMTPAIARTIAVREDRREAQQ
LTRAFESWKNFLGSEERMQDRWESWIGDVEYACDRLNELIDADKIPVTENLRLRNSGKLHADQP
ESLKKARRGSKRPRPQRYVLGDALPADVINRVTDPGLWTALVRAPGFDSQLGLPADLNRGLKLR
GKRISADFPIDYFPTDSPALAVQGGYVGLEFHHARLYRIIGPKEKVKYALLRVCAIDLCGIDCD
DLFEVELKPSSISMRTADAKLKEAMGNGSAKQIGWLVLGDEIQIDPTKFPKQSIGKFLKECGPV
SSWRVSALDTPSKITLKPRLLSNEPLLKTSRVGGHESDLVVAECVEKIMKKTGWVVEINALCQS
GLIRVIRRNALGEVRTSPKSGLPISLNLR

SEQ ID NO: 356
MRYRVGLDLGTASVGAAVFSMDEQGNPMELIWHYERLFSEPLVPDMGQLKPKKAARRLARQQRR
QIDRRASRLRRIAIVSRRLGIAPGRNDSGVHGNDVPTLRAMAVNERIELGQLRAVLLRMGKKRG
YGGTFKAVRKVGEAGEVASGASRLEEEMVALASVQNKDSVTVGEYLAARVEHGLPSKLKVAANN
EYYAPEYALFRQYLGLPAIKGRPDCLPNMYALRHQIEHEFERIWATQSQFHDVMKDHGVKEEIR
NAIFFQRPLKSPADKVGRCSLQTNLPRAPRAQIAAQNFRIEKQMADLRWGMGRRAEMLNDHQKA
VIRELLNQQKELSFRKIYKELERAGCPGPEGKGLNMDRAALGGRDDLSGNTTLAAWRKLGLEDR
WQELDEVTQIQVINFLADLGSPEQLDTDDWSCRFMGKNGRPRNFSDEFVAFMNELRMTDGFDRL
SKMGFEGGRSSYSIKALKALTEWMIAPHWRETPETHRVDEEAAIRECYPESLATPAQGGRQSKL
EPPPLTGNEVVDVALRQVRHTINMMIDDLGSVPAQIVVEMAREMKGGVTRRNDIEKQNKRFASE
RKKAAQSIEENGKTPTPARILRYQLWIEQGHQCPYCESNISLEQALSGAYTNFEHILPRTLTQI
GRKRSELVLAHRECNDEKGNRTPYQAFGHDDRRWRIVEQRANALPKKSSRKTRLLLLKDFEGEA
LTDESIDEFADRQLHESSWLAKVTTQWLSSLGSDVYVSRGSLTAELRRRWGLDTVIPQVRFESG
MPVVDEEGAEITPEEFEKFRLQWEGHRVTREMRTDRRPDKRIDHRHHLVDAIVTALTSRSLYQQ
YAKAWKVADEKQRHGRVDVKVELPMPILTIRDIALEAVRSVRISHKPDRYPDGRFFEATAYGIA
QRLDERSGEKVDWLVSRKSLTDLAPEKKSIDVDKVRANISRIVGEAIRLHISNIFEKRVSKGMT
PQQALREPIEFQGNILRKVRCFYSKADDCVRIEHSSRRGHHYKMLLNDGFAYMEVPCKEGILYG

VPNLVRPSEAVGIKRAPESGDFIRFYKGDTVKNIKTGRVYTIKQILGDGGGKLILTPVTETKPA
DLLSAKWGRLKVGGRNIHLLRLCAE

SEQ ID NO: 357
MIGEHVRGGCLFDDHWTPNWGAFRLPNTVRTFTKAENPKDGSSLAEPRRQARGLRRRLRRKTQR
LEDLRRLLAKEGVLSLSDLETLFRETPAKDPYQLRAEGLDRPLSFPEWVRVLYHITKHRGFQSN
RRNPVEDGQERSRQEEEGKLLSGVGENERLLREGGYRTAGEMLARDPKFQDHRRNRAGDYSHTL
SRSLLLEEARRLFQSQRTLGNPHASSNLEEAFLHLVAFQNPFASGEDIRNKAGHCSLEPDQIRA
PRRSASAETFMLLQKTGNLRLIHRRTGEERPLTDKEREQIHLLAWKQEKVTHKTLRRHLEIPEE
WLFTGLPYHRSGDKAEEKLFVHLAGIHEIRKALDKGPDPAVWDTLRSRRDLLDSIADTLTFYKN
EDEILPRLESLGLSPENARALAPLSFSGTAHLSLSALGKLLPHLEEGKSYTQARADAGYAAPPP
DRHPKLPPLEEADWRNPVVFRALTQTRKVVNALVRRYGPPWCIHLETARELSQPAKVRRRIETE
QQANEKKKQQAEREFLDIVGTAPGPGDLLKMRLWREQGGFCPYCEEYLNPTRLAEPGYAEMDHI
LPYSRSLDNGWHNRVLVHGKDNRDKGNRTPFEAFGGDTARWDRLVAWVQASHLSAPKKRNLLRE
DFGEEAERELKDRNLTDTRFITKTAATLLRDRLTFHPEAPKDPVMTLNGRLTAFLRKQWGLHKN
RKNGDLHHALDAAVLAVASRSFVYRLSSHNAAWGELPRGREAENGFSLPYPAFRSEVLARLCPT
REEILLRLDQGGVGYDEAFRNGLRPVFVSRAPSRRLRGKAHMETLRSPKWKDHPEGPRTASRIP
LKDLNLEKLERMVGKDRDRKLYEALRERLAAFGGNGKKAFVAPFRKPCRSGEGPLVRSLRIFDS
GYSGVELRDGGEVYAVADHESMVRVDVYAKKNRFYLVPVYVADVARGIVKNRAIVAHKSEEEWD
LVDGSFDFRFSLFPGDLVEIEKKDGAYLGYYKSCHRGDGRLLLDRHDRMPRESDCGTFYVSTRK
DVLSMSKYQVDPLGEIRLVGSEKPPFVL

SEQ ID NO: 358
MEKKRKVTLGFDLGIASVGWAIVDSETNQVYKLGSRLFDAPDTNLERRTQRGTRRLLRRRKYRN
QKFYNLVKRTEVFGLSSREAIENRFRELSIKYPNIIELKTKALSQEVCPDEIAWILHDYLKNRG
YFYDEKETKEDFDQQTVESMPSYKLNEFYKKYGYFKGALSQPTESEMKDNKDLKEAFFFDFSNK
EWLKEINYFFNVQKNILSETFIEEFKKIFSFTRDISKGPGSDNMPSPYGIFGEFGDNGQGGRYE
HIWDKNIGKCSIFTNEQRAPKYLPSALIFNFLNELANIRLYSTDKKNIQPLWKLSSVDKLNILL
NLFNLPISEKKKKLTSTNINDIVKKESIKSIMISVEDIDMIKDEWAGKEPNVYGVGLSGLNIEE
SAKENKFKFQDLKILNVLINLLDNVGIKFEFKDRNDIIKNLELLDNLYLFLIYQKESNNKDSSI
DLFIAKNESLNIENLKLKLKEFLLGAGNEFENHNSKTHSLSKKAIDEILPKLLDNNEGWNLEAI
KNYDEEIKSQIEDNSSLMAKQDKKYLNDNFLKDAILPPNVKVTFQQAILIFNKIIQKFSKDFEI
DKVVIELAREMTQDQENDALKGIAKAQKSKKSLVEERLEANNIDKSVFNDKYEKLIYKIFLWIS
QDFKDPYTGAQISVNEIVNNKVEIDHIIPYSLCFDDSSANKVLVHKQSNQEKSNSLPYEYIKQG
HSGWNWDEFTKYVKRVFVNNVDSILSKKERLKKSENLLTASYDGYDKLGFLARNLNDTRYATIL
FRDQLNNYAEHHLIDNKKMFKVIAMNGAVTSFIRKNMSYDNKLRLKDRSDFSHHAYDAAIIALF
SNKTKTLYNLIDPSLNGIISKRSEGYWVIEDRYTGEIKELKKEDWTSIKNNVQARKIAKEIEEY
LIDLDDEVFFSRKTKRKTNRQLYNETIYGIATKTDEDGITNYYKKEKFSILDDKDIYLRLLRER
EKFVINQSNPEVIDQIIEIIESYGKENNIPSRDEAINIKYTKNKINYNLYLKQYMRSLTKSLDQ
FSEEFINQMIANKTFVLYNPTKNTTRKIKFLRLVNDVKINDIRKNQVINKFNGKNNEPKAFYEN
INSLGAIVFKNSANNFKTLSINTQIAIFGDKNWDIEDFKTYNMEKIEKYKEIYGIDKTYNFHSF
IFPGTILLDKQNKEFYYISSIQTVRDIIEIKFLNKIEFKDENKNQDTSKTPKRLMFGIKSIMNN
YEQVDISPFGINKKIFE

SEQ ID NO: 359
MGYRIGLDVGITSTGYAVLKTDKNGLPYKILTLDSVIYPRAENPQTGASLAEPRRIKRGLRRRT
RRTKFRKQRTQQLFIHSGLLSKPEIEQILATPQAKYSVYELRVAGLDRRLTNSELFRVLYFFIG

HRGFKSNRKAELNPENEADKKQMGQLLNSIEEIRKAIAEKGYRTVGELYLKDPKYNDHKRNKGY
IDGYLSTPNRQMLVDEIKQILDKQRELGNEKLTDEFYATYLLGDENRAGIFQAQRDFDEGPGAG
PYAGDQIKKMVGKDIFEPTEDRAAKATYTFQYFNLLQKMTSLNYQNTTGDTWHTLNGLDRQAII
DAVFAKAEKPTKTYKPTDFGELRKLLKLPDDARFNLVNYGSLQTQKEIETVEKKTRFVDFKAYH
DLVKVLPEEMWQSRQLLDHIGTALTLYSSDKRRRRYFAEELNLPAELIEKLLPLNFSKFGHLSI
KSMQNIIPYLEMGQVYSEATTNTGYDFRKKQISKDTIREEITNPVVRRAVTKTIKIVEQIIRRY
GKPDGINIELARELGRNFKERGDIQKRQDKNRQTNDKIAAELTELGIPVNGQNIIRYKLHKEQN
GVDPYTGDQIPFERAFSEGYEVDHIIPYSISWDDSYTNKVLTSAKCNREKGNRIPMVYLANNEQ
RLNALTNIADNIIRNSRKRQKLLKQKLSDEELKDWKQRNINDTRFITRVLYNYFRQAIEFNPEL
EKKQRVLPLNGEVTSKIRSRWGFLKVREDGDLHHAIDATVIAAITPKFIQQVTKYSQHQEVKNN
QALWHDAEIKDAEYAAEAQRMDADLFNKIFNGFPLPWPEFLDELLARISDNPVEMMKSRSWNTY
TPIEIAKLKPVFVVRLANHKISGPAHLDTIRSAKLFDEKGIVLSRVSITKLKINKKGQVATGDG
IYDPENSNNGDKVVYSAIRQALEAHNGSGELAFPDGYLEYVDHGTKKLVRKVRVAKKVSLPVRL
KNKAAADNGSMVRIDVFNTGKKFVFVPIYIKDTVEQVLPNKAIARGKSLWYQITESDQFCFSLY
PGDMVHIESKTGIKPKYSNKENNTSVVPIKNFYGYFDGADIATASILVRAHDSSYTARSIGIAG
LLKFEKYQVDYFGRYHKVHEKKRQLFVKRDE

SEQ ID NO: 360
MQKNINTKQNHIYIKQAQKIKEKLGDKPYRIGLDLGVGSIGFAIVSMEENDGNVLLPKEIIMVG
SRIFKASAGAADRKLSRGQRNNHRHTRERMRYLWKVLAEQKLALPVPADLDRKENSSEGETSAK
RFLGDVLQKDIYELRVKSLDERLSLQELGYVLYHIAGHRGSSAIRTFENDSEEAQKENTENKKI
AGNIKRLMAKKNYRTYGEYLYKEFFENKEKHKREKISNAANNHKFSPTRDLVIKEAEAILKKQA
GKDGFHKELTEEYIEKLTKAIGYESEKLIPESGFCPYLKDEKRLPASHKLNEERRLWETLNNAR
YSDPIVDIVTGEITGYYEKQFTKEQKQKLFDYLLTGSELTPAQTKKLLGLKNTNFEDIILQGRD
KKAQKIKGYKLIKLESMPFWARLSEAQQDSFLYDWNSCPDEKLLTEKLSNEYHLTEEEIDNAFN
EIVLSSSYAPLGKSAMLIILEKIKNDLSYTEAVEEALKEGKLTKEKQAIKDRLPYYGAVLQEST
QKIIAKGFSPQFKDKGYKTPHTNKYELEYGRIANPVVHQTLNELRKLVNEIIDILGKKPCEIGL
ETARELKKSAEDRSKLSREQNDNESNRNRIYEIYIRPQQQVIITRRENPRNYILKFELLEEQKS
QCPFCGGQISPNDIINNQADIEHLFPIAESEDNGRNNLVISHSACNADKAKRSPWAAFASAAKD
SKYDYNRILSNVKENIPHKAWRFNQGAFEKFIENKPMAARFKTDNSYISKVAHKYLACLFEKPN
IICVKGSLTAQLRMAWGLQGLMIPFAKQLITEKESESFNKDVNSNKKIRLDNRHHALDAIVIAY
ASRGYGNLLNKMAGKDYKINYSERNWLSKILLPPNNIVWENIDADLESFESSVKTALKNAFISV
KHDHSDNGELVKGTMYKIFYSERGYTLTTYKKLSALKLTDPQKKKTPKDFLETALLKFKGRESE
MKNEKIKSAIENNKRLFDVIQDNLEKAKKLLEEENEKSKAEGKKEKNINDASIYQKAISLSGDK
YVQLSKKEPGKFFAISKPTPTTTGYGYDTGDSLCVDLYYDNKGKLCGEIIRKIDAQQKNPLKYK
EQGFTLFERIYGGDILEVDFDIHSDKNSFRNNTGSAPENRVFIKVGTFTEITNNNIQIWFGNII
KSTGGQDDSFTINSMQQYNPRKLILSSCGFIKYRSPILKNKEG

SEQ ID NO: 361
MAAFKPNPINYILGLDIGIASVGWAMVEIDEDENPICLIDLGVRVFERAEVPKTGDSLAMARRL
ARSVRRLTRRRAHRLLRARRLLKREGVLQAADFDENGLIKSLPNTPWQLRAAALDRKLTPLEWS
AVLLHLIKHRGYLSQRKNEGETADKELGALLKGVADNAHALQTGDFRTPAELALNKFEKESGHI
RNQRGDYSHTFSRKDLQAELILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLG
HCTFEPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSERPLTDTERATLMDEPYRKSKLTYAQA
RKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSPELQDEIGT
AFSLFKTDEDITGRLKDRIQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEI
YGDHYGKKNTEEKIYLPPIPADEIRNPVVLRALSQARKVINGVVRRYGSPARIHIETAREVGKS

FKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEINLG
RLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVE
TSRFPRSKKQRILLQKFDEDGFKERNLNDTRYVNRFLCQFVADRMRLTGKGKKRVFASNGQITN
LLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGEVLHQ
KTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSR
APNRKMSGQGHMETVKSAKRLDEGVSVLRVPLTQLKLKDLEKMVNREREPKLYEALKARLEAHK
DDPAKAFAEPFYKYDKAGNRTQQVKAVRVEQVQKTGVWVRNHNGIADNATMVRVDVFEKGDKYY
LVPIYSWQVAKGILPDRAVVQGKDEEDWQLIDDSFNFKFSLHPNDLVEVITKKARMFGYFASCH
RGTGNINIRIHDLDHKIGKNGILEGIGVKTALSFQKYQIDELGKEIRPCRLKKRPPVR

SEQ ID NO: 362
MQTTNLSYILGLDLGIASVGWAVVEINENEDPIGLIDVGVRIFERAEVPKTGESLALSRRLARS
TRRLIRRRAHRLLLAKRFLKREGILSTIDLEKGLPNQAWELRVAGLERRLSAIEWGAVLLHLIK
HRGYLSKRKNESQTNNKELGALLSGVAQNHQLLQSDDYRTPAELALKKFAKEEGHIRNQRGAYT
HTFNRLDLLAELNLLFAQQHQFGNPHCKEHIQQYMTELLMWQKPALSGEAILKMLGKCTHEKNE
FKAAKHTYSAERFVWLTKLNNLRILEDGAERALNEEERQLLINHPYEKSKLTYAQVRKLLGLSE
QAIFKHLRYSKENAESATFMELKAWHAIRKALENQGLKDTWQDLAKKPDLLDEIGTAFSLYKTD
EDIQQYLTNKVPNSVINALLVSLNFDKFIELSLKSLRKILPLMEQGKRYDQACREIYGHHYGEA
NQKTSQLLPAIPAQEIRNPVVLRTLSQARKVINAIIRQYGSPARVHIETGRELGKSFKERREIQ
KQQEDNRTKRESAVQKFKELFSDFSSEPKSKDILKFRLYEQQHGKCLYSGKEINIHRLNEKGYV
EIDHALPFSRTWDDSFNNKVLVLASENQNKGNQTPYEWLQGKINSERWKNFVALVLGSQCSAAK
KQRLLTQVIDDNKFIDRNLNDTRYIARFLSNYIQENLLLVGKNKKNVFTPNGQITALLRSRWGL
IKARENNNRHHALDAIVVACATPSMQQKITRFIRFKEVHPYKIENRYEMVDQESGEIISPHFPE
PWAYFRQEVNIRVFDNHPDTVLKEMLPDRPQANHQFVQPLFVSRAPTRKMSGQGHMETIKSAKR
LAEGISVLRIPLTQLKPNLLENMVNKEREPALYAGLKARLAEFNQDPAKAFATPFYKQGGQQVK
AIRVEQVQKSGVLVRENNGVADNASIVRTDVFIKNNKFFLVPIYTWQVAKGILPNKAIVAHKNE
DEWEEMDEGAKFKFSLFPNDLVELKTKKEYFFGYYIGLDRATGNISLKEHDGEISKGKDGVYRV
GVKLALSFEKYQVDELGKNRQICRPQQRQPVR

SEQ ID NO: 363
MGIRFAFDLGTNSIGWAVWRTGPGVFGEDTAASLDGSGVLIFKDGRNPKDGQSLATMRRVPRQS
RKRRDRFVLRRRDLLAALRKAGLFPVDVEEGRRLAATDPYHLRAKALDESLTPHEMGRVIFHLN
QRRGFRSNRKADRQDREKGKIAEGSKRLAETLAATNCRTLGEFLWSRHRGTPRTRSPTRIRMEG
EGAKALYAFYPTREMVRAEFERLWTAQSRFAPDLLTPERHEEIAGILFRQRDLAPPKIGCCTFE
PSERRLPRALPSVEARGIYERLAHLRITTGPVSDRGLTRPERDVLASALLAGKSLTFKAVRKTL
KILPHALVNFEEAGEKGLDGALTAKLLSKPDHYGAAWHGLSFAEKDTFVGKLLDEADEERLIRR
LVTENRLSEDAARRCASIPLADGYGRLGRTANTEILAALVEETDETGTVVTYAEAVRRAGERTG
RNWHHSDERDGVILDRLPYYGEILQRHVVPGSGEPEEKNEAARWGRLANPTVHIGLNQLRKVVN
RLIAAHGRPDQIVVELARELKLNREQKERLDRENRKNREENERRTAILAEHGQRDTAENKIRLR
LFEEQARANAGIALCPYTGRAIGIAELFTSEVEIDHILPVSLTLDDSLANRVLCRREANREKRR
QTPFQAFGATPAWNDIVARAAKLPPNKRWRFDPAALERFEREGGFLGRQLNETKYLSRLAKIYL
GKICDPDRVYVTPGTLTGLLRARWGLNSILSDSNFKNRSDHRHHAVDAVVIGVLTRGMIQRIAH
DAARAEDQDLDRVFRDVPVPFEDFRDHVRERVSTITVAVKPEHGKGGALHEDTSYGLVPDTDPN
AALGNLVVRKPIRSLTAGEVDRVRDRALRARLGALAAPFRDESGRVRDAKGLAQALEAFGAENG
IRRVRILKPDASVVTIADRRTGVPYRAVAPGENHHVDIVQMRDGSWRGFAASVFEVNRPGWRPE
WEVKKLGGKLVMRLHKGDMVELSDKDGQRRVKVVQQIEISANRVRLSPHNDGGKLQDRHADADD
PFRWDLATIPLLKDRGCVAVRVDPIGVVTLRRSNV

SEQ ID NO: 364
MMEVFMGRLVLGLDIGITSVGFGIIDLDESEIVDYGVRLFKEGTAAENETRRTKRGGRRLKRRR
VTRREDMLHLLKQAGIISTSFHPLNNPYDVRVKGLNERLNGEELATALLHLCKHRGSSVETIED
DEAKAKEAGETKKVLSMNDQLLKSGKYVCEIQKERLRTNGHIRGHENNFKTRAYVDEAFQILSH
QDLSNELKSAIITIISRKRMYYDGPGGPLSPTPYGRYTYFGQKEPIDLIEKMRGKCSLFPNEPR
APKLAYSAELFNLLNDLNNLSIEGEKLTSEQKAMILKIVHEKGKITPKQLAKEVGVSLEQIRGF
RIDTKGSPLLSELTGYKMIREVLEKSNDEHLEDHVFYDEIAEILTKTKDIEGRKKQISELSSDL
NEESVHQLAGLTKFTAYHSLSFKALRLINEEMLKTELNQMQSITLFGLKQNNELSVKGMKNIQA
DDTAILSPVAKRAQRETFKVVNRLREIYGEFDSIVVEMAREKNSEEQRKAIRERQKFFEMRNKQ
VADIIGDDRKINAKLREKLVLYQEQDGKTAYSLEPIDLKLLIDDPNAYEVDHIIPISISLDDSI
TNKVLVTHRENQEKGNLTPISAFVKGRFTKGSLAQYKAYCLKLKEKNIKTNKGYRKKVEQYLLN
ENDIYKYDIQKEFINRNLVDTSYASRVVLNTLTTYFKQNEIPTKVFTVKGSLTNAFRRKINLKK
DRDEDYGHHAIDALIIASMPKMRLLSTIFSRYKIEDIYDESTGEVFSSGDDSMYYDDRYFAFIA
SLKAIKVRKFSHKIDTKPNRSVADETIYSTRVIDGKEKVVKKYKDIYDPKFTALAEDILNNAYQ
EKYLMALHDPQTFDQIVKVVNYYFEEMSKSEKYFTKDKKGRIKISGMNPLSLYRDEHGMLKKYS
KKGDGPAITQMKYFDGVLGNHIDISAHYQVRDKKVVLQQISPYRTDFYYSKENGYKFVTIRYKD
VRWSEKKKKYVIDQQDYAMKKAEKKIDDTYEFQFSMHRDELIGITKAEGEALIYPDETWHNFNF
FFHAGETPEILKFTATNNDKSNKIEVKPIHCYCKMRLMPTISKKIVRIDKYATDVVGNLYKVKK
NTLKFEFD

SEQ ID NO: 365
MKKILGVDLGITSFGYAILQETGKDLYRCLDNSVVMRNNPYDEKSGESSQSIRSTQKSMRRLIE
KRKKRIRCVAQTMERYGILDYSETMKINDPKNNPIKNRWQLRAVDAWKRPLSPQELFAIFAHMA
KHRGYKSIATEDLIYELELELGLNDPEKESEKKADERRQVYNALRHLEELRKKYGGETIAQTIH
RAVEAGDLRSYRNHDDYEKMIRREDIEEEIEKVLLRQAELGALGLPEEQVSELIDELKACITDQ
EMPTIDESLFGKCTFYKDELAAPAYSYLYDLYRLYKKLADLNIDGYEVTQEDREKVIEWVEKKI
AQGKNLKKITHKDLRKILGLAPEQKIFGVEDERIVKGKKEPRTFVPFFFLADIAKFKELFASIQ
KHPDALQIFRELAEILQRSKTPQEALDRLRALMAGKGIDTDDRELLELFKNKRSGTRELSHRYI
LEALPLFLEGYDEKEVQRILGFDDREDYSRYPKSLRHLHLREGNLFEKEENPINNHAVKSLASW
ALGLIADLSWRYGPFDEIILETTRDALPEKIRKEIDKAMREREKALDKIIGKYKKEFPSIDKRL
ARKIQLWERQKGLDLYSGKVINLSQLLDGSADIEHIVPQSLGGLSTDYNTIVTLKSVNAAKGNR
LPGDWLAGNPDYRERIGMLSEKGLIDWKKRKNLLAQSLDEIYTENTHSKGIRATSYLEALVAQV
LKRYYPFPDPELRKNGIGVRMIPGKVTSKTRSLLGIKSKSRETNFHHAEDALILSTLTRGWQNR
LHRMLRDNYGKSEAELKELWKKYMPHIEGLTLADYIDEAFRRFMSKGEESLFYRDMFDTIRSIS
YWVDKKPLSASSHKETVYSSRHEVPTLRKNILEAFDSLNVIKDRHKLTTEEFMKRYDKEIRQKL
WLHRIGNTNDESYRAVEERATQIAQILTRYQLMDAQNDKEIDEKFQQALKELITSPIEVTGKLL
RKMRFVYDKLNAMQIDRGLVETDKNMLGIHISKGPNEKLIFRRMDVNNAHELQKERSGILCYLN
EMLFIFNKKGLIHYGCLRSYLEKGQGSKYIALFNPRFPANPKAQPSKFTSDSKIKQVGIGSATG
IIKAHLDLDGHVRSYEVFGTLPEGSIEWFKEESGYGRVEDDPHH

SEQ ID NO: 366
MRPIEPWILGLDIGTDSLGWAVFSCEEKGPPTAKELLGGGVRLFDSGRDAKDHTSRQAERGAFR
RARRQTRTWPWRRDRLIALFQAAGLTPPAAETRQIALALRREAVSRPLAPDALWAALLHLAHHR
GFRSNRIDKRERAAAKALAKAKPAKATAKATAPAKEADDEAGFWEGAEAALRQRMAASGAPTVG
ALLADDLDRGQPVRMRYNQSDRDGVVAPTRALIAEELAEIVARQSSAYPGLDWPAVTRLVLDQR
PLRSKGAGPCAFLPGEDRALRALPTVQDFIIRQTLANLRLPSTSADEPRPLTDEEHAKALALLS

TARFVEWPALRRALGLKRGVKFTAETERNGAKQAARGTAGNLTEAILAPLIPGWSGWDLDRKDR
VFSDLWAARQDRSALLALIGDPRGPTRVTEDETAEAVADAIQIVLPTGRASLSAKAARAIAQAM
APGIGYDEAVTLALGLHHSHRPRQERLARLPYYAAALPDVGLDGDPVGPPPAEDDGAAAEAYYG
RIGNISVHIALNETRKIVNALLHRHGPILRLVMVETTRELKAGADERKRMIAEQAERERENAEI
DVELRKSDRWMANARERRQRVRLARRQNNLCPYTSTPIGHADLLGDAYDIDHVIPLARGGRDSL
DNMVLCQSDANKTKGDKTPWEAFHDKPGWIAQRDDFLARLDPQTAKALAWRFADDAGERVARKS
AEDEDQGFLPRQLTDTGYIARVALRYLSLVTNEPNAVVATNGRLTGLLRLAWDITPGPAPRDLL
PTPRDALRDDTAARRFLDGLTPPPLAKAVEGAVQARLAALGRSRVADAGLADALGLTLASLGGG
GKNRADHRHHFIDAAMIAVTTRGLINQINQASGAGRILDLRKWPRTNFEPPYPTFRAEVMKQWD
HIHPSIRPAHRDGGSLHAATVFGVRNRPDARVLVQRKPVEKLFLDANAKPLPADKIAEIIDGFA
SPRMAKRFKALLARYQAAHPEVPPALAALAVARDPAFGPRGMTANTVIAGRSDGDGEDAGLITP
FRANPKAAVRTMGNAVYEVWEIQVKGRPRWTHRVLTRFDRTQPAPPPPPENARLVMRLRRGDLV
YWPLESGDRLFLVKKMAVDGRLALWPARLATGKATALYAQLSCPNINLNGDQGYCVQSAEGIRK
EKIRTTSCTALGRLRLSKKAT

SEQ ID NO: 367
MKYTLGLDVGIASVGWAVIDKDNNKIIDLGVRCFDKAEESKTGESLATARRIARGMRRRISRRS
QRLRLVKKLFVQYEIIKDSSEFNRIFDTSRDGWKDPWELRYNALSRILKPYELVQVLTHITKRR
GFKSNRKEDLSTTKEGVVITSIKNNSEMLRTKNYRTIGEMIFMETPENSNKRNKVDEYIHTIAR
EDLLNEIKYIFSIQRKLGSPFVTEKLEHDFLNIWEFQRPFASGDSILSKVGKCTLLKEELRAPT
SCYTSEYFGLLQSINNLVLVEDNNTLTLNNDQRAKIIEYAHFKNEIKYSEIRKLLDIEPEILFK
AHNLTHKNPSGNNESKKFYEMKSYHKLKSTLPTDIWGKLHSNKESLDNLFYCLTVYKNDNEIKD
YLQANNLDYLIEYIAKLPTFNKFKHLSLVAMKRIIPFMEKGYKYSDACNMAELDFTGSSKLEKC
NKLTVEPIIENVTNPVVIRALTQARKVINAIIQKYGLPYMVNIELAREAGMTRQDRDNLKKEHE
NNRKAREKISDLIRQNGRVASGLDILKWRLWEDQGGRCAYSGKPIPVCDLLNDSLTQIDHIYPY
SRSMDDSYMNKVLVLTDENQNKRSYTPYEVWGSTEKWEDFEARIYSMHLPQSKEKRLLNRNFIT
KDLDSFISRNLNDTRYISRFLKNYIESYLQFSNDSPKSCVVCVNGQCTAQLRSRWGLNKNREES
DLHHALDAAVIACADRKIIKEITNYYNERENHNYKVKYPLPWHSFRQDLMETLAGVFISRAPRR
KITGPAHDETIRSPKHFNKGLTSVKIPLTTVTLEKLETMVKNTKGGISDKAVYNVLKNRLIEHN
NKPLKAFAEKIYKPLKNGTNGAIIRSIRVETPSYTGVFRNEGKGISDNSLMVRVDVFKKKDKYY
LVPIYVAHMIKKELPSKAIVPLKPESQWELIDSTHEFLFSLYQNDYLVIKTKKGITEGYYRSCH
RGTGSLSLMPHFANNKNVKIDIGVRTAISIEKYNVDILGNKSIVKGEPRRGMEKYNSFKSN

SEQ ID NO: 368
MIRTLGIDIGIASIGWAVIEGEYTDKGLENKEIVASGVRVFTKAENPKNKESLALPRTLARSAR
RRNARKKGRIQQVKHYLSKALGLDLECFVQGEKLATLFQTSKDFLSPWELRERALYRVLDKEEL
ARVILHIAKRRGYDDITYGVEDNDSGKIKKAIAENSKRIKEEQCKTIGEMMYKLYFQKSLNVRN
KKESYNRCVGRSELREELKTIFQIQQELKSPWVNEELIYKLLGNPDAQSKQEREGLIFYQRPLK
GFGDKIGKCSHIKKGENSPYRACKHAPSAEEFVALTKSINFLKNLTNRHGLCFSQEDMCVYLGK
ILQEAQKNEKGLTYSKLKLLLDLPSDFEFLGLDYSGKNPEKAVFLSLPSTFKLNKITQDRKTQD
KIANILGANKDWEAILKELESLQLSKEQIQTIKDAKLNFSKHINLSLEALYHLLPLMREGKRYD
EGVEILQERGIFSKPQPKNRQLLPPLSELAKEESYFDIPNPVLRRALSEFRKVVNALLEKYGGF
HYFHIELTRDVCKAKSARMQLEKINKKNKSENDAASQLLEVLGLPNTYNNRLKCKLWKQQEEYC
LYSGEKITIDHLKDQRALQIDHAFPLSRSLDDSQSNKVLCLTSSNQEKSNKTPYEWLGSDEKKW
DMYVGRVYSSNFSPSKKRKLTQKNFKERNEEDFLARNLVDTGYIGRVTKEYIKHSLSFLPLPDG
KKEHIRIISGSMTSTMRSFWGVQEKNRDHHLHHAQDAIIIACIEPSMIQKYTTYLKDKETHRLK
SHQKAQILREGDHKLSLRWPMSNFKDKIQESIQNIIPSHHVSHKVTGELHQETVRTKEFYYQAF

GGEEGVKKALKFGKIREINQGIVDNGAMVRVDIFKSKDKGKFYAVPIYTYDFAIGKLPNKAIVQ
GKKNGIIKDWLEMDENYEFCFSLFKNDCIKIQTKEMQEAVLAIYKSTNSAKATIELEHLSKYAL
KNEDEEKMFTDTDKEKNKTMTRESCGIQGLKVFQKVKLSVLGEVLEHKPRNRQNIALKTTPKHV

SEQ ID NO: 369
MKYSIGLDIGIASVGWSVINKDKERIEDMGVRIFQKAENPKDGSSLASSRREKRGSRRRNRRKK
HRLDRIKNILCESGLVKKNEIEKIYKNAYLKSPWELRAKSLEAKISNKEIAQILLHIAKRRGFK
SFRKTDRNADDTGKLLSGIQENKKIMEEKGYLTIGDMVAKDPKFNTHVRNKAGSYLFSFSRKLL
EDEVRKIQAKQKELGNTHFTDDVLEKYIEVFNSQRNFDEGPSKPSPYYSEIGQIAKMIGNCTFE
SSEKRTAKNTWSGERFVFLQKLNNFRIVGLSGKRPLTEEERDIVEKEVYLKKEVRYEKLRKILY
LKEEERFGDLNYSKDEKQDKKTEKTKFISLIGNYTIKKLNLSEKLKSEIEEDKSKLDKIIEILT
FNKSDKTIESNLKKLELSREDIEILLSEEFSGTLNLSLKAIKKILPYLEKGLSYNEACEKADYD
YKNNGIKFKRGELLPVVDKDLIANPVVLRAISQTRKVVNAIIRKYGTPHTIHVEVARDLAKSYD
DRQTIIKENKKRELENEKTKKFISEEFGIKNVKGKLLLKYRLYQEQEGRCAYSRKELSLSEVIL
DESMTDIDHIIPYSRSMDDSYSNKVLVLSGENRKKSNLLPKEYFDRQGRDWDTFVLNVKAMKIH
PRKKSNLLKEKFTREDNKDWKSRALNDTRYISRFVANYLENALEYRDDSPKKRVFMIPGQLTAQ
LRARWRLNKVRENGDLHHALDAAVVAVTDQKAINNISNISRYKELKNCKDVIPSIEYHADEETG
EVYFEEVKDTRFPMPWSGFDLELQKRLESENPREEFYNLLSDKRYLGWFNYEEGFIEKLRPVFV
SRMPNRGVKGQAHQETIRSSKKISNQIAVSKKPLNSIKLKDLEKMQGRDTDRKLYEALKNRLEE
YDDKPEKAFAEPFYKPTNSGKRGPLVRGIKVEEKQNVGVYVNGGQASNGSMVRIDVFRKNGKFY
TVPIYVHQTLLKELPNRAINGKPYKDWDLIDGSFEFLYSFYPNDLIEIEFGKSKSIKNDNKLTK
TEIPEVNLSEVLGYYRGMDTSTGAATIDTQDGKIQMRIGIKTVKNIKKYQVDVLGNVYKVKREK
RQTF

SEQ ID NO: 370
MSKKVSRRYEEQAQEICQRLGSRPYSIGLDLGVGSIGVAVAAYDPIKKQPSDLVFVSSRIFIPS
TGAAERRQKRGQRNSLRHRANRLKFLWKLLAERNLMLSYSEQDVPDPARLRFEDAVVRANPYEL
RLKGLNEQLTLSELGYALYHIANHRGSSSVRTFLDEEKSSDDKKLEEQQAMTEQLAKEKGISTF
IEVLTAFNTNGLIGYRNSESVKSKGVPVPTRDIISNEIDVLLQTQKQFYQEILSDEYCDRIVSA
ILFENEKIVPEAGCCPYFPDEKKLPRCHFLNEERRLWEAINNARIKMPMQEGAAKRYQSASFSD
EQRHILFHIARSGTDITPKLVQKEFPALKTSIIVLQGKEKAIQKIAGFRFRRLEEKSFWKRLSE
EQKDDFFSAWTNTPDDKRLSKYLMKHLLLTENEVVDALKTVSLIGDYGPIGKTATQLLMKHLED
GLTYTEALERGMETGEFQELSVWEQQSLLPYYGQILTGSTQALMGKYWHSAFKEKRDSEGFFKP
NTNSDEEKYGRIANPVVHQTLNELRKLMNELITILGAKPQEITVELARELKVGAEKREDIIKQQ
TKQEKEAVLAYSKYCEPNNLDKRYIERFRLLEDQAFVCPYCLEHISVADIAAGRADVDHIFPRD
DTADNSYGNKVVAHRQCNDIKGKRTPYAAFSNTSAWGPIMHYLDETPGMWRKRRKFETNEEEYA
KYLQSKGFVSRFESDNSYIAKAAKEYLRCLFNPNNVTAVGSLKGMETSILRKAWNLQGIDDLLG
SRHWSKDADTSPTMRKNRDDNRHHGLDAIVALYCSRSLVQMINTMSEQGKRAVEIEAMIPIPGY
ASEPNLSFEAQRELFRKKILEFMDLHAFVSMKTDNDANGALLKDTVYSILGADTQGEDLVFVVK
KKIKDIGVKIGDYEEVASAIRGRITDKQPKWYPMEMKDKIEQLQSKNEAALQKYKESLVQAAAV
LEESNRKLIESGKKPIQLSEKTISKKALELVGGYYYLISNNKRTKTFVVKEPSNEVKGFAFDTG
SNLCLDFYHDAQGKLCGEIIRKIQAMNPSYKPAYMKQGYSLYVRLYQGDVCELRASDLTEAESN
LAKTTHVRLPNAKPGRTFVIIITFTEMGSGYQIYFSNLAKSKKGQDTSFTLTTIKNYDVRKVQL
SSAGLVRYVSPLLVDKIEKDEVALCGE

SEQ ID NO: 371

MNQKFILGLDIGITSVGYGLIDYETKNIIDAGVRLFPEANVENNEGRRSKRGSRRLKRRRIHRL
ERVKKLLEDYNLLDQSQIPQSTNPYAIRVKGLSEALSKDELVIALLHIAKRRGIHKIDVIDSND
DVGNELSTKEQLNKNSKLLKDKFVCQIQLERMNEGQVRGEKNRFKTADIIKEIIQLLNVQKNFH
QLDENFINKYIELVEMRREYFEGPGKGSPYGWEGDPKAWYETLMGHCTYFPDELRSVKYAYSAD
LFNALNDLNNLVIQRDGLSKLEYHEKYHIIENVFKQKKKPTLKQIANEINVNPEDIKGYRITKS
GKPQFTEFKLYHDLKSVLFDQSILENEDVLDQIAEILTIYQDKDSIKSKLTELDILLNEEDKEN
IAQLTGYTGTHRLSLKCIRLVLEEQWYSSRNQMEIFTHLNIKPKKINLTAANKIPKAMIDEFIL
SPVVKRTFGQAINLINKIIEKYGVPEDIIIELARENNSKDKQKFINEMQKKNENTRKRINEIIG
KYGNQNAKRLVEKIRLHDEQEGKCLYSLESIPLEDLLNNPNHYEVDHIIPRSVSFDNSYHNKVL
VKQSENSKKSNLTPYQYFNSGKSKLSYNQFKQHILNLSKSQDRISKKKKEYLLEERDINKFEVQ
KEFINRNLVDTRYATRELTNYLKAYFSANNMNVKVKTINGSFTDYLRKVWKFKKERNHGYKHHA
EDALIIANADFLFKENKKLKAVNSVLEKPEIESKQLDIQVDSEDNYSEMFIIPKQVQDIKDFRN
FKYSHRVDKKPNRQLINDTLYSTRKKDNSTYIVQTIKDIYAKDNTTLKKQFDKSPEKFLMYQHD
PRTFEKLEVIMKQYANEKNPLAKYHEETGEYLTKYSKKNNGPIVKSLKYIGNKLGSHLDVTHQF
KSSTKKLVKLSIKPYRFDVYLTDKGYKFITISYLDVLKKDNYYYIPEQKYDKLKLGKAIDKNAK
FIASFYKNDLIKLDGEIYKIIGVNSDTRNMIELDLPDIRYKEYCELNNIKGEPRIKKTIGKKVN
SIEKLTTDVLGNVFTNTQYTKPQLLFKRGN

SEQ ID NO: 372
MIMKLEKWRLGLDLGTNSIGWSVFSLDKDNSVQDLIDMGVRIFSDGRDPKTKEPLAVARRTARS
QRKLIYRRKLRRKQVFKFLQEQGLFPKTKEECMTLKSLNPYELRIKALDEKLEPYELGRALFNL
AVRRGFKSNRKDGSREEVSEKKSPDEIKTQADMQTHLEKAIKENGCRTITEFLYKNQGENGGIR
FAPGRMTYYPTRKMYEEEFNLIRSKQEKYYPQVDWDDIYKAIFYQRPLKPQQRGYCIYENDKER
TFKAMPCSQKLRILQDIGNLAYYEGGSKKRVELNDNQDKVLYELLNSKDKVTFDQMRKALCLAD
SNSFNLEENRDFLIGNPTAVKMRSKNRFGKLWDEIPLEEQDLIIETIITADEDDAVYEVIKKYD
LTQEQRDFIVKNTILQSGTSMLCKEVSEKLVKRLEEIADLKYHEAVESLGYKFADQTVEKYDLL
PYYGKVLPGSTMEIDLSAPETNPEKHYGKISNPTVHVALNQTRVVVNALIKEYGKPSQIAIELS
RDLKNNVEKKAEIARKQNQRAKENIAINDTISALYHTAFPGKSFYPNRNDRMKYRLWSELGLGN
KCIYCGKGISGAELFTKEIEIEHILPFSRTLLDAESNLTVAHSSCNAFKAERSPFEAFGTNPSG
YSWQEIIQRANQLKNTSKKNKFSPNAMDSFEKDSSFIARQLSDNQYIAKAALRYLKCLVENPSD
VWTTNGSMTKLLRDKWEMDSILCRKFTEKEVALLGLKPEQIGNYKKNRFDHRHHAIDAVVIGLT
DRSMVQKLATKNSHKGNRIEIPEFPILRSDLIEKVKNIVVSFKPDHGAEGKLSKETLLGKIKLH
GKETFVCRENIVSLSEKNLDDIVDEIKSKVKDYVAKHKGQKIEAVLSDFSKENGIKKVRCVNRV
QTPIEITSGKISRYLSPEDYFAAVIWEIPGEKKTFKAQYIRRNEVEKNSKGLNVVKPAVLENGK
PHPAAKQVCLLHKDDYLEFSDKGKMYFCRIAGYAATNNKLDIRPVYAVSYCADWINSTNETMLT
GYWKPTPTQNWVSVNVLFDKQKARLVTVSPIGRVFRK

SEQ ID NO: 373
MSSKAIDSLEQLDLFKPQEYTLGLDLGIKSIGWAILSGERIANAGVYLFETAEELNSTGNKLIS
KAAERGRKRRIRRMLDRKARRGRHIRYLLEREGLPTDELEEVVVHQSNRTLWDVRAEAVERKLT
KQELAAVLFHLVRHRGYFPNTKKLPPDDESDSADEEQGKINRATSRLREELKASDCKTIGQFLA
QNRDRQRNREGDYSNLMARKLVFEEALQILAFQRKQGHELSKDFEKTYLDVLMGQRSGRSPKLG
NCSLIPSELRAPSSAPSTEWFKFLQNLGNLQISNAYREEWSIDAPRRAQIIDACSQRSTSSYWQ
IRRDFQIPDEYRFNLVNYERRDPDVDLQEYLQQQERKTLANFRNWKQLEKIIGTGHPIQTLDEA
ARLITLIKDDEKLSDQLADLLPEASDKAITQLCELDFTTAAKISLEAMYRILPHMNQGMGFFDA
CQQESLPEIGVPPAGDRVPPFDEMYNPVVNRVLSQSRKLINAVIDEYGMPAKIRVELARDLGKG
RELRERIKLDQLDKSKQNDQRAEDFRAEFQQAPRGDQSLRYRLWKEQNCTCPYSGRMIPVNSVL

SEDTQIDHILPISQSFDNSLSNKVLCFTEENAQKSNRTPFEYLDAADFQRLEAISGNWPEAKRN
KLLHKSFGKVAEEWKSRALNDTRYLTSALADHLRHHLPDSKIQTVNGRITGYLRKQWGLEKDRD
KHTHHAVDAIVVACTTPAIVQQVTLYHQDIRRYKKLGEKRPTPWPETFRQDVLDVEEEIFITRQ
PKKVSGGIQTKDTLRKHRSKPDRQRVALTKVKLADLERLVEKDASNRNLYEHLKQCLEESGDQP
TKAFKAPFYMPSGPEAKQRPILSKVTLLREKPEPPKQLTELSGGRRYDSMAQGRLDIYRYKPGG
KRKDEYRVVLQRMIDLMRGEENVHVFQKGVPYDQGPEIEQNYTFLFSLYFDDLVEFQRSADSEV
IRGYYRTFNIANGQLKISTYLEGRQDFDFFGANRLAHFAKVQVNLLGKVIK

SEQ ID NO: 374
MRSLRYRLALDLGSTSLGWALFRLDACNRPTAVIKAGVRIFSDGRNPKDGSSLAVTRRAARAMR
RRRDRLLKRKTRMQAKLVEHGFFPADAGKRKALEQLNPYALRAKGLQEALLPGEFARALFHINQ
RRGFKSNRKTDKKDNDSGVLKKAIGQLRQQMAEQGSRTVGEYLWTRLQQGQGVRARYREKPYTT
EEGKKRIDKSYDLYIDRAMIEQEFDALWAAQAAFNPTLFHEAARADLKDTLLHQRPLRPVKPGR
CTLLPEEERAPLALPSTQRFRIHQEVNHLRLLDENLREVALTLAQRDAVVTALETKAKLSFEQI
RKLLKLSGSVQFNLEDAKRTELKGNATSAALARKELFGAAWSGFDEALQDEIVWQLVTEEGEGA
LIAWLQTHTGVDEARAQAIVDVSLPEGYGNLSRKALARIVPALRAAVITYDKAVQAAGFDHHSQ
LGFEYDASEVEDLVHPETGEIRSVFKQLPYYGKALQRHVAFGSGKPEDPDEKRYGKIANPTVHI
GLNQVRMVVNALIRRYGRPTEVVIELARDLKQSREQKVEAQRRQADNQRRNARIRRSIAEVLGI
GEERVRGSDIQKWICWEELSFDAADRRCPYSGVQISAAMLLSDEVEVEHILPFSKTLDDSLNNR
TVAMRQANRIKRNRTPWDARAEFEAQGWSYEDILQRAERMPLRKRYRFAPDGYERWLGDDKDFL
ARALNDTRYLSRVAAEYLRLVCPGTRVIPGQLTALLRGKFGLNDVLGLDGEKNRNDHRHHAVDA
CVIGVTDQGLMQRFATASAQARGDGLTRLVDGMPMPWPTYRDHVERAVRHIWVSHRPDHGFEGA
MMEETSYGIRKDGSIKQRRKADGSAGREISNLIRIHEATQPLRHGVSADGQPLAYKGYVGGSNY
CIEITVNDKGKWEGEVISTFRAYGVVRAGGMGRLRNPHEGQNGRKLIMRLVIGDSVRLEVDGAE
RTMRIVKISGSNGQIFMAPIHEANVDARNTDKQDAFTYTSKYAGSLQKAKTRRVTISPIGEVRD
PGFKG

SEQ ID NO: 375
MARPAFRAPRREHVNGWTPDPHRISKPFFILVSWHLLSRVVIDSSSGCFPGTSRDHTDKFAEWE
CAVQPYRLSFDLGTNSIGWGLLNLDRQGKPREIRALGSRIFSDGRDPQDKASLAVARRLARQMR
RRRDRYLTRRTRLMGALVRFGLMPADPAARKRLEVAVDPYLARERATRERLEPFEIGRALFHLN
QRRGYKPVRTATKPDEEAGKVKEAVERLEAAIAAAGAPTLGAWFAWRKTRGETLRARLAGKGKE
AAYPFYPARRMLEAEFDTLWAEQARHHPDLLTAEAREILRHRIFHQRPLKPPPVGRCTLYPDDG
RAPRALPSAQRLRLFQELASLRVIHLDLSERPLTPAERDRIVAFVQGRPPKAGRKPGKVQKSVP
FEKLRGLLELPPGTGFSLESDKRPELLGDETGARIAPAFGPGWTALPLEEQDALVELLLTEAEP
ERAIAALTARWALDEATAAKLAGATLPDFHGRYGRRAVAELLPVLERETRGDPDGRVRPIRLDE
AVKLLRGGKDHSDFSREGALLDALPYYGAVLERHVAFGTGNPADPEEKRVGRVANPTVHIALNQ
LRHLVNAILARHGRPEEIVIELARDLKRSAEDRRREDKRQADNQKRNEERKRLILSLGERPTPR
NLLKLRLWEEQGPVENRRCPYSGETISMRMLLSEQVDIDHILPFSVSLDDSAANKVVCLREANR
IKRNRSPWEAFGHDSERWAGILARAEALPKNKRWRFAPDALEKLEGEGGLRARHLNDTRHLSRL
AVEYLRCVCPKVRVSPGRLTALLRRRWGIDAILAEADGPPPEVPAETLDPSPAEKNRADHRHHA
LDAVVIGCIDRSMVQRVQLAAASAEREAAAREDNIRRVLEGFKEEPWDGFRAELERRARTIVVS
HRPEHGIGGALHKETAYGPVDPPEEGFNLVVRKPIDGLSKDEINSVRDPRLRRALIDRLAIRRR
DANDPATALAKAAEDLAAQPASRGIRRVRVLKKESNPIRVEHGGNPSGPRSGGPFHKLLLAGEV
HHVDVALRADGRRWVGHWVTLFEAHGGRGADGAAAPPRLGDGERFLMRLHKGDCLKLEHKGRVR
VMQVVKLEPSSNSVVVVEPHQVKTDRSKHVKISCDQLRARGARRVTVDPLGRVRVHAPGARVGI
GGDAGRTAMEPAEDIS

SEQ ID NO: 376

MKRTSLRAYRLGVDLGANSLGWFVVWLDDHGQPEGLGPGGVRIFPDGRNPQSKQSNAAGRRLAR
SARRRRDRYLQRRGKLMGLLVKHGLMPADEPARKRLECLDPYGLRAKALDEVLPLHHVGRALFH
LNQRRGLFANRAIEQGDKDASAIKAAAGRLQTSMQACGARTLGEFLNRRHQLRATVRARSPVGG
DVQARYEFYPTRAMVDAEFEAIWAAQAPHHPTMTAEAHDTIREAIFSQRAMKRPSIGKCSLDPA
TSQDDVDGFRCAWSHPLAQRFRIWQDVRNLAVVETGPTSSRLGKEDQDKVARALLQTDQLSFDE
IRGLLGLPSDARFNLESDRRDHLKGDATGAILSARRHFGPAWHDRSLDRQIDIVALLESALDEA
AIIASLGTTHSLDEAAAQRALSALLPDGYCRLGLRAIKRVLPLMEAGRTYAEAASAAGYDHALL
PGGKLSPTGYLPYYGQWLQNDVVGSDDERDTNERRWGRLPNPTVHIGIGQLRRVVNELIRWHGP
PAEITVELTRDLKLSPRRLAELEREQAENQRKNDKRTSLLRKLGLPASTHNLLKLRLWDEQGDV
ASECPYTGEAIGLERLVSDDVDIDHLIPFSISWDDSAANKVVCMRYANREKGNRTPFEAFGHRQ
GRPYDWADIAERAARLPRGKRWRFGPGARAQFEELGDFQARLLNETSWLARVAKQYLAAVTHPH
RIHVLPGRLTALLRATWELNDLLPGSDDRAAKSRKDHRHHAIDALVAALTDQALLRRMANAHDD
TRRKIEVLLPWPTFRIDLETRLKAMLVSHKPDHGLQARLHEDTAYGTVEHPETEDGANLVYRKT
FVDISEKEIDRIRDRRLRDLVRAHVAGERQQGKTLKAAVLSFAQRRDIAGHPNGIRHVRLTKSI
KPDYLVPIRDKAGRIYKSYNAGENAFVDILQAESGRWIARATTVFQANQANESHDAPAAQPIMR
VFKGDMLRIDHAGAEKFVKIVRLSPSNNLLYLVEHHQAGVFQTRHDDPEDSFRWLFASFDKLRE
WNAELVRIDTLGQPWRRKRGLETGSEDATRIGWTRPKKWP

SEQ ID NO: 377

MERIFGFDIGTTSIGFSVIDYSSTQSAGNIQRLGVRIFPEARDPDGTPLNQQRRQKRMMRRQLR
RRRIRRKALNETLHEAGFLPAYGSADWPVVMADEPYELRRRGLEEGLSAYEFGRAIYHLAQHRH
FKGRELEESDTPDPDVDDEKEAANERAATLKALKNEQTTLGAWLARRPPSDRKRGIHAHRNVVA
EEFERLWEVQSKFHPALKSEEMRARISDTIFAQRPVFWRKNTLGECRFMPGEPLCPKGSWLSQQ
RRMLEKLNNLAIAGGNARPLDAEERDAILSKLQQQASMSWPGVRSALKALYKQRGEPGAEKSLK
FNLELGGESKLLGNALEAKLADMFGPDWPAHPRKQEIRHAVHERLWAADYGETPDKKRVIILSE
KDRKAHREAAANSFVADFGITGEQAAQLQALKLPTGWEPYSIPALNLFLAELEKGERFGALVNG
PDWEGWRRTNFPHRNQPTGEILDKLPSPASKEERERISQLRNPTVVRTQNELRKVVNNLIGLYG
KPDRIRIEVGRDVGKSKREREEIQSGIRRNEKQRKKATEDLIKNGIANPSRDDVEKWILWKEGQ
ERCPYTGDQIGFNALFREGRYEVEHIWPRSRSFDNSPRNKTLCRKDVNIEKGNRMPFEAFGHDE
DRWSAIQIRLQGMVSAKGGTGMSPGKVKRFLAKTMPEDFAARQLNDTRYAAKQILAQLKRLWPD
MGPEAPVKVEAVTGQVTAQLRKLWTLNNILADDGEKTRADHRHHAIDALTVACTHPGMTNKLSR
YWQLRDDPRAEKPALTPPWDTIRADAEKAVSEIVVSHRVRKKVSGPLHKETTYGDTGTDIKTKS
GTYRQFVTRKKIESLSKGELDEIRDPRIKEIVAAHVAGRGGDPKKAFPPYPCVSPGGPEIRKVR
LTSKQQLNLMAQTGNGYADLGSNHHIAIYRLPDGKADFEIVSLFDASRRLAQRNPIVQRTRADG
ASFVMSLAAGEAIMIPEGSKKGIWIVQGVWASGQVVLERDTDADHSTTTRPMPNPILKDDAKKV
SIDPIGRVRPSND

SEQ ID NO: 378

MNKRILGLDTGTNSLGWAVVDWDEHAQSYELIKYGDVIFQEGVKIEKGIESSKAAERSGYKAIR
KQYFRRRLRKIQVLKVLVKYHLCPYLSDDDLRQWHLQKQYPKSDELMLWQRTSDEEGKNPYYDR
HRCLHEKLDLTVEADRYTLGRALYHLTQRRGFLSNRLDTSADNKEDGVVKSGISQLSTEMEEAG
CEYLGDYFYKLYDAQGNKVRIRQRYTDRNKHYQHEFDAICEKQELSSELIEDLQRAIFFQLPLK
SQRHGVGRCTFERGKPRCADSHPDYEEFRMLCFVNNIQVKGPHDLELRPLTYEEREKIEPLFFR
KSKPNFDFEDIAKALAGKKNYAWIHDKEERAYKFNYRMTQGVPGCPTIAQLKSIFGDDWKTGIA
ETYTLIQKKNGSKSLQEMVDDVWNVLYSFSSVEKLKEFAHHKLQLDEESAEKFAKIKLSHSFAA

LSLKAIRKFLPFLRKGMYYTHASFFANIPTIVGKEIWNKEQNRKYIMENVGELVFNYQPKHREV
QGTIEMLIKDFLANNFELPAGATDKLYHPSMIETYPNAQRNEFGILQLGSPRTNAIRNPMAMRS
LHILRRVVNQLLKESIIDENTEVHVEYARELNDANKRRAIADRQKEQDKQHKKYGDEIRKLYKE
ETGKDIEPTQTDVLKFQLWEEQNHHCLYTGEQIGITDFIGSNPKFDIEHTIPQSVGGDSTQMNL
TLCDNRFNREVKKAKLPTELANHEEILTRIEPWKNKYEQLVKERDKQRTFAGMDKAVKDIRIQK
RHKLQMEIDYWRGKYERFTMTEVPEGFSRRQGTGIGLISRYAGLYLKSLFHQADSRNKSNVYVV
KGVATAEFRKMWGLQSEYEKKCRDNHSHHCMDAITIACIGKREYDLMAEYYRMEETFKQGRGSK
PKFSKPWATFTEDVLNIYKNLLVVHDTPNNMPKHTKKYVQTSIGKVLAQGDTARGSLHLDTYYG
AIERDGEIRYVVRRPLSSFTKPEELENIVDETVKRTIKEAIADKNFKQAIAEPIYMNEEKGILI
KKVRCFAKSVKQPINIRQHRDLSKKEYKQQYHVMNENNYLLAIYEGLVKNKVVREFEIVSYIEA
AKYYKRSQDRNIFSSIVPTHSTKYGLPLKTKLLMGQLVLMFEENPDEIQVDNTKDLVKRLYKVV
GIEKDGRIKFKYHQEARKEGLPIFSTPYKNNDDYAPIFRQSINNINILVDGIDFTIDILGKVTL
KE

SEQ ID NO: 379
MNYKMGLDIGIASVGWAVINLDLKRIEDLGVRIFDKAEHPQNGESLALPRRIARSARRRLRRRK
HRLERIRRLLVSENVLTKEEMNLLFKQKKQIDVWQLRVDALERKLNNDELARVLLHLAKRRGFK
SNRKSERNSKESSEFLKNIEENQSILAQYRSVGEMIVKDSKFAYHKRNKLDSYSNMIARDDLER
EIKLIFEKQREFNNPVCTERLEEKYLNIWSSQRPFASKEDIEKKVGFCTFEPKEKRAPKATYTF
QSFIVWEHINKLRLVSPDETRALTEIERNLLYKQAFSKNKMTYYDIRKLLNLSDDIHFKGLLYD
PKSSLKQIENIRFLELDSYHKIRKCIENVYGKDGIRMFNETDIDTFGYALTIFKDDEDIVAYLQ
NEYITKNGKRVSNLANKVYDKSLIDELLNLSFSKFAHLSMKAIRNILPYMEQGEIYSKACELAG
YNFTGPKKKEKALLLPVIPNIANPVVMRALTQSRKVVNAIIKKYGSPVSIHIELARDLSHSFDE
RKKIQKDQTENRKKNETAIKQLIEYELTKNPTGLDIVKFKLWSEQQGRCMYSLKPIELERLLEP
GYVEVDHILPYSRSLDDSYANKVLVLTKENREKGNHTPVEYLGLGSERWKKFEKFVLANKQFSK
KKKQNLLRLRYEETEEKEFKERNLNDTRYISKFFANFIKEHLKFADGDGGQKVYTINGKITAHL
RSRWDFNKNREESDLHHAVDAVIVACATQGMIKKITEFYKAREQNKESAKKKEPIFPQPWPHFA
DELKARLSKFPQESIEAFALGNYDRKKLESLRPVFVSRMPKRSVTGAAHQETLRRCVGIDEQSG
KIQTAVKTKLSDIKLDKDGHFPMYQKESDPRTYEAIRQRLLEHNNDPKKAFQEPLYKPKKNGEP
GPVIRTVKIIDTKNKVVHLDGSKTVAYNSNIVRTDVFEKDGKYYCVPVYTMDIMKGTLPNKAIE
ANKPYSEWKEMTEEYTFQFSLFPNDLVRIVLPREKTIKTSTNEEIIIKDIFAYYKTIDSATGGL
ELISHDRNFSLRGVGSKTLKRFEKYQVDVLGNIHKVKGEKRVGLAAPTNQKKGKTVDSLQSVSD

SEQ ID NO: 380
MRRLGLDLGTNSIGWCLLDLGDDGEPVSIFRTGARIFSDGRDPKSLGSLKATRREARLTRRRRD
RFIQRQKNLINALVKYGLMPADEIQRQALAYKDPYPIRKKALDEAIDPYEMGRAIFHINQRRGF
KSNRKSADNEAGVVKQSIADLEMKLGEAGARTIGEFLADRQATNDTVRARRLSGTNALYEFYPD
RYMLEQEFDTLWAKQAAFNPSLYIEAARERLKEIVFFQRKLKPQEVGRCIFLSDEDRISKALPS
FQRFRIYQELSNLAWIDHDGVAHRITASLALRDHLFDELEHKKKLTFKAMRAILRKQGVVDYPV
GFNLESDNRDHLIGNLTSCIMRDAKKMIGSAWDRLDEEEQDSFILMLQDDQKGDDEVRSILTQQ
YGLSDDVAEDCLDVRLPDGHGSLSKKAIDRILPVLRDQGLIYYDAVKEAGLGEANLYDPYAALS
DKLDYYGKALAGHVMGASGKFEDSDEKRYGTISNPTVHIALNQVRAVVNELIRLHGKPDEVVIE
IGRDLPMGADGKRELERFQKEGRAKNERARDELKKLGHIDSRESRQKFQLWEQLAKEPVDRCCP
FTGKMMSISDLFSDKVEIEHLLPFSLTLDDSMANKTVCFRQANRDKGNRAPFDAFGNSPAGYDW
QEILGRSQNLPYAKRWRFLPDAMKRFEADGGFLERQLNDTRYISRYTTEYISTIIPKNKIWVVT
GRLTSLLRGFWGLNSILRGHNTDDGTPAKKSRDDHRHHAIDAIVVGMTSRGLLQKVSKAARRSE
DLDLTRLFEGRIDPWDGFRDEVKKHIDAIIVSHRPRKKSQGALHNDTAYGIVEHAENGASTVVH

RVPITSLGKQSDIEKVRDPLIKSALLNETAGLSGKSFENAVQKWCADNSIKSLRIVETVSIIPI
TDKEGVAYKGYKGDGNAYMDIYQDPTSSKWKGEIVSRFDANQKGFIPSWQSQFPTARLIMRLRI
NDLLKLQDGEIEEIYRVQRLSGSKILMAPHTEANVDARDRDKNDTFKLTSKSPGKLQSASARKV
HISPTGLIREG

SEQ ID NO: 381

MKNILGLDLGLSSIGWSVIRENSEEQELVAMGSRVVSLTAAELSSFTQGNGVSINSQRTQKRTQ
RKGYDRYQLRRTLLRNKLDTLGMLPDDSLSYLPKLQLWGLRAKAVTQRIELNELGRVLLHLNQK
RGYKSIKSDFSGDKKITDYVKTVKTRYDELKEMRLTIGELFFRRLTENAFFRCKEQVYPRQAYV
EEFDCIMNCQRKFYPDILTDETIRCIRDEIIYYQRPLKSCKYLVSRCEFEKRFYLNAAGKKTEA
GPKVSPRTSPLFQVCRLWESINNIVVKDRRNEIVFISAEQRAALFDFLNTHEKLKGSDLLKLLG
LSKTYGYRLGEQFKTGIQGNKTRVEIERALGNYPDKKRLLQFNLQEESSSMVNTETGEIIPMIS
LSFEQEPLYRLWHVLYSIDDREQLQSVLRQKFGIDDDEVLERLSAIDLVKAGFGNKSSKAIRRI
LPFLQLGMNYAEACEAAGYNHSNNYTKAENEARALLDRLPAIKKNELRQPVVEKILNQMVNVVN
ALMEKYGRFDEIRVELARELKQSKEERSNTYKSINKNQRENEQIAKRIVEYGVPTRSRIQKYKM
WEESKHCCIYCGQPVDVGDFLRGFDVEVEHIIPKSLYFDDSFANKVCSCRSCNKEKNNRTAYDY
MKSKGEKALSDYVERVNTMYTNNQISKTKWQNLLTPVDKISIDFIDRQLRESQYIARKAKEILT
SICYNVTATSGSVTSFLRHVWGWDTVLHDLNFDRYKKVGLTEVIEVNHRGSVIRREQIKDWSKR
FDHRHHAIDALTIACTKQAYIQRLNNLRAEEGPDFNKMSLERYIQSQPHFSVAQVREAVDRILV
SFRAGKRAVTPGKRYIRKNRKRISVQSVLIPRGALSEESVYGVIHVWEKDEQGHVIQKQRAVMK
YPITSINREMLDKEKVVDKRIHRILSGRLAQYNDNPKEAFAKPVYIDKECRIPIRTVRCFAKPA
INTLVPLKKDDKGNPVAWVNPGNNHHVAIYRDEDGKYKERTVTFWEAVDRCRVGIPAIVTQPDT
IWDNILQRNDISENVLESLPDVKWQFVLSLQQNEMFILGMNEEDYRYAMDQQDYALLNKYLYRV
QKLSKSDYSFRYHTETSVEDKYDGKPNLKLSMQMGKLKRVSIKSLLGLNPHKVHISVLGEIKEI
S

SEQ ID NO: 382

MAEKQHRWGLDIGTNSIGWAVIALIEGRPAGLVATGSRIFSDGRNPKDGSSLAVERRGPRQMRR
RRDRYLRRRDRFMQALINVGLMPGDAAARKALVTENPYVLRQRGLDQALTLPEFGRALFHLNQR
RGFQSNRKTDRATAKESGKVKNAIAAFRAGMGNARTVGEALARRLEDGRPVRARMVGQGKDEHY
ELYIAREWIAQEFDALWASQQRFHAEVLADAARDRLRAILLFQRKLLPVPVGKCFLEPNQPRVA
AALPSAQRFRLMQELNHLRVMTLADKRERPLSFQERNDLLAQLVARPKCGFDMLRKIVFGANKE
AYRFTIESERRKELKGCDTAAKLAKVNALGTRWQALSLDEQDRLVCLLLDGENDAVLADALREH
YGLTDAQIDTLLGLSFEDGHMRLGRSALLRVLDALESGRDEQGLPLSYDKAVVAAGYPAHTADL
ENGERDALPYYGELLWRYTQDAPTAKNDAERKFGKIANPTVHIGLNQLRKLVNALIQRYGKPAQ
IVVELARNLKAGLEEKERIKKQQTANLERNERIRQKLQDAGVPDNRENRLRMRLFEELGQGNGL
GTPCIYSGRQISLQRLFSNDVQVDHILPFSKTLDDSFANKVLAQHDANRYKGNRGPFEAFGANR
DGYAWDDIRARAAVLPRNKRNRFAETAMQDWLHNETDFLARQLTDTAYLSRVARQYLTAICSKD
DVYVSPGRLTAMLRAKWGLNRVLDGVMEEQGRPAVKNRDDHRHHAIDAVVIGATDRAMLQQVAT
LAARAREQDAERLIGDMPTPWPNFLEDVRAAVARCVVSHKPDHGPEGGLHNDTAYGIVAGPFED
GRYRVRHRVSLFDLKPGDLSNVRCDAPLQAELEPIFEQDDARAREVALTALAERYRQRKVWLEE
LMSVLPIRPRGEDGKTLPDSAPYKAYKGDSNYCYELFINERGRWDGELISTFRANQAAYRRFRN
DPARFRRYTAGGRPLLMRLCINDYIAVGTAAERTIFRVVKMSENKITLAEHFEGGTLKQRDADK
DDPFKYLTKSPGALRDLGARRIFVDLIGRVLDPGIKGD

SEQ ID NO: 383

MIERILGVDLGISSLGWAIVEYDKDDEAANRIIDCGVRLFTAAETPKKKESPNKARREARGIRR
VLNRRRVRMNMIKKLFLRAGLIQDVDLDGEGGMFYSKANRADVWELRHDGLYRLLKGDELARVL
IHIAKHRGYKFIGDDEADEESGKVKKAGVVLRQNFEAAGCRTVGEWLWRERGANGKKRNKHGDY
EISIHRDLLVEEVEAIFVAQQEMRSTIATDALKAAYREIAFFVRPMQRIEKMVGHCTYFPEERR
APKSAPTAEKFIAISKFFSTVIIDNEGWEQKIIERKTLEELLDFAVSREKVEFRHLRKFLDLSD
NEIFKGLHYKGKPKTAKKREATLFDPNEPTELEFDKVEAEKKAWISLRGAAKLREALGNEFYGR
FVALGKHADEATKILTYYKDEGQKRRELTKLPLEAEMVERLVKIGFSDFLKLSLKAIRDILPAM
ESGARYDEAVLMLGVPHKEKSAILPPLNKTDIDILNPTVIRAFAQFRKVANALVRKYGAFDRVH
FELAREINTKGEIEDIKESQRKNEKERKEAADWIAETSFQVPLTRKNILKKRLYIQQDGRCAYT
GDVIELERLFDEGYCEIDHILPRSRSADDSFANKVLCLARANQQKTDRTPYEWFGHDAARWNAF
ETRTSAPSNRVRTGKGKIDRLLKKNFDENSEMAFKDRNLNDTRYMARAIKTYCEQYWVFKNSHT
KAPVQVRSGKLTSVLRYQWGLESKDRESHTHHAVDAIIIAFSTQGMVQKLSEYYRFKETHREKE
RPKLAVPLANFRDAVEEATRIENTETVKEGVEVKRLLISRPPRARVTGQAHEQTAKPYPRIKQV
KNKKKWRLAPIDEEKFESFKADRVASANQKNFYETSTIPRVDVYHKKGKFHLVPIYLHEMVLNE
LPNLSLGTNPEAMDENFFKFSIFKDDLISIQTQGTPKKPAKIIMGYFKNMHGANMVLSSINNSP
CEGFTCTPVSMDKKHKDKCKLCPEENRIAGRCLQGFLDYWSQEGLRPPRKEFECDQGVKFALDV
KKYQIDPLGYYYEVKQEKRLGTIPQMRSAKKLVKK

SEQ ID NO: 384

MNNSIKSKPEVTIGLDLGVGSVGWAIVDNETNIIHHLGSRLFSQAKTAEDRRSFRGVRRLIRRR
KYKLKRFVNLIWKYNSYFGFKNKEDILNNYQEQQKLHNTVLNLKSEALNAKIDPKALSWILHDY
LKNRGHFYEDNRDFNVYPTKELAKYFDKYGYYKGIIDSKEDNDNKLEEELTKYKFSNKHWLEEV
KKVLSNQTGLPEKFKEEYESLFSYVRNYSEGPGSINSVSPYGIYHLDEKEGKVVQKYNNIWDKT
IGKCNIFPDEYRAPKNSPIAMIFNEINELSTIRSYSIYLTGWFINQEFKKAYLNKLLDLLIKTN
GEKPIDARQFKKLREETIAESIGKETLKDVENEEKLEKEDHKWKLKGLKLNTNGKIQYNDLSSL
AKFVHKLKQHLKLDFLLEDQYATLDKINFLQSLFVYLGKHLRYSNRVDSANLKEFSDSNKLFER
ILQKQKDGLFKLFEQTDKDDEKILAQTHSLSTKAMLLAITRMTNLDNDEDNQKNNDKGWNFEAI
KNFDQKFIDITKKNNNLSLKQNKRYLDDRFINDAILSPGVKRILREATKVFNAILKQFSEEYDV
TKVVIELARELSEEKELENTKNYKKLIKKNGDKISEGLKALGISEDEIKDILKSPTKSYKFLLW
LQQDHIDPYSLKEIAFDDIFTKTEKFEIDHIIPYSISFDDSSSNKLLVLAESNQAKSNQTPYEF
ISSGNAGIKWEDYEAYCRKFKDGDSSLLDSTQRSKKFAKMMKTDTSSKYDIGFLARNLNDTRYA
TIVFRDALEDYANNHLVEDKPMFKVVCINGSVTSFLRKNFDDSSYAKKDRDKNIHHAVDASIIS
IFSNETKTLFNQLTQFADYKLFKNTDGSWKKIDPKTGVVTEVTDENWKQIRVRNQVSEIAKVIE
KYIQDSNIERKARYSRKIENKTNISLFNDTVYSAKKVGYEDQIKRKNLKTLDIHESAKENKNSK
VKRQFVYRKLVNVSLLNNDKLADLFAEKEDILMYRANPWVINLAEQIFNEYTENKKIKSQNVFE
KYMLDLTKEFPEKFSEFLVKSMLRNKTAIIYDDKKNIVHRIKRLKMLSSELKENKLSNVIIRSK
NQSGTKLSYQDTINSLALMIMRSIDPTAKKQYIRVPLNTLNLHLGDHDFDLHNMDAYLKKPKFV
KYLKANEIGDEYKPWRVLTSGTLLIHKKDKKLMYISSFQNLNDVIEIKNLIETEYKENDDSDSK
KKKKANRFLMTLSTILNDYILLDAKDNFDILGLSKNRIDEILNSKLGLDKIVK

SEQ ID NO: 385

MGGSEVGTVPVTWRLGVDVGERSIGLAAVSYEEDKPKEILAAVSWIHDGGVGDERSGASRLALR
GMARRARRLRRFRRARLRDLDMLLSELGWTPLPDKNVSPVDAWLARKRLAEEYVVDETERRRLL
GYAVSHMARHRGWRNPWTTIKDLKNLPQPSDSWERTRESLEARYSVSLEPGTVGQWAGYLLQRA
PGIRLNPTQQSAGRRAELSNATAFETRLRQEDVLWELRCIADVQGLPEDVVSNVIDAVFCQKRP
SVPAERIGRDPLDPSQLRASRACLEFQEYRIVAAVANLRIRDGSGSRPLSLEERNAVIEALLAQ
TERSLTWSDIALEILKLPNESDLTSVPEEDGPSSLAYSQFAPFDETSARIAEFIAKNRRKIPTF

AQWWQEQDRTSRSDLVAALADNSIAGEEEQELLVHLPDAELEALEGLALPSGRVAYSRLTLSGL
TRVMRDDGVDVHNARKTCFGVDDNWRPPLPALHEATGHPVVDRNLAILRKFLSSATMRWGPPQS
IVVELARGASESRERQAEEEAARRAHRKANDRIRAELRASGLSDPSPADLVRARLLELYDCHCM
YCGAPISWENSELDHIVPRTDGGSNRHENLAITCGACNKEKGRRPFASWAETSNRVQLRDVIDR
VQKLKYSGNMYWTRDEFSRYKKSVVARLKRRTSDPEVIQSIESTGYAAVALRDRLLSYGEKNGV
AQVAVFRGGVTAEARRWLDISIERLFSRVAIFAQSTSTKRLDRRHHAVDAVVLTTLTPGVAKTL
ADARSRRVSAEFWRRPSDVNRHSTEEPQSPAYRQWKESCSGLGDLLISTAARDSIAVAAPLRLR
PTGALHEETLRAFSEHTVGAAWKGAELRRIVEPEVYAAFLALTDPGGRFLKVSPSEDVLPADEN
RHIVLSDRVLGPRDRVKLFPDDRGSIRVRGGAAYIASFHHARVFRWGSSHSPSFALLRVSLADL
AVAGLLRDGVDVFTAELPPWTPAWRYASIALVKAVESGDAKQVGWLVPGDELDFGPEGVTTAAG
DLSMFLKYFPERHWVVTGFEDDKRINLKPAFLSAEQAEVLRTERSDRPDTLTEAGEILAQFFPR
CWRATVAKVLCHPGLTVIRRTALGQPRWRRGHLPYSWRPWSADPWSGGTP

SEQ ID NO: 386
MHNKKNITIGFDLGIASIGWAIIDSTTSKILDWGTRTFEERKTANERRAFRSTRRNIRRKAYRN
QRFINLILKYKDLFELKNISDIQRANKKDTENYEKIISFFTEIYKKCAAKHSNILEVKVKALDS
KIEKLDLIWILHDYLENRGFFYDLEEENVADKYEGIEHPSILLYDFFKKNGFFKSNSSIPKDLG
GYSFSNLQWVNEIKKLFEVQEINPEFSEKFLNLFTSVRDYAKGPGSEHSASEYGIFQKDEKGKV
FKKYDNIWDKTIGKCSFFVEENRSPVNYPSYEIFNLLNQLINLSTDLKTTNKKIWQLSSNDRNE
LLDELLKVKEKAKIISISLKKNEIKKIILKDFGFEKSDIDDQDTIEGRKIIKEEPTTKLEVTKH
LLATIYSHSSDSNWININNILEFLPYLDAICIILDREKSRGQDEVLKKLTEKNIFEVLKIDREK
QLDFVKSIFSNTKFNFKKIGNFSLKAIREFLPKMFEQNKNSEYLKWKDEEIRRKWEEQKSKLGK
TDKKTKYLNPRIFQDEIISPGTKNTFEQAVLVLNQIIKKYSKENIIDAIIIESPREKNDKKTIE
EIKKRNKKGKGKTLEKLFQILNLENKGYKLSDLETKPAKLLDRLRFYHQQDGIDLYTLDKINID
QLINGSQKYEIEHIIPYSMSYDNSQANKILTEKAENLKKGKLIASEYIKRNGDEFYNKYYEKAK
ELFINKYKKNKKLDSYVDLDEDSAKNRFRFLTLQDYDEFQVEFLARNLNDTRYSTKLFYHALVE
HFENNEFFTYIDENSSKHKVKISTIKGHVTKYFRAKPVQKNNGPNENLNNNKPEKIEKNRENNE
HHAVDAAIVAIIGNKNPQIANLLTLADNKTDKKFLLHDENYKENIETGELVKIPKFEVDKLAKV
EDLKKIIQEKYEEAKKHTAIKFSRKTRTILNGGLSDETLYGFKYDEKEDKYFKIIKKKLVTSKN
EELKKYFENPFGKKADGKSEYTVLMAQSHLSEFNKLKEIFEKYNGFSNKTGNAFVEYMNDLALK
EPTLKAEIESAKSVEKLLYYNFKPSDQFTYHDNINNKSFKRFYKNIRIIEYKSIPIKFKILSKH
DGGKSFKDTLFSLYSLVYKVYENGKESYKSIPVTSQMRNFGIDEFDFLDENLYNKEKLDIYKSD
FAKPIPVNCKPVFVLKKGSILKKKSLDIDDFKETKETEEGNYYFISTISKRFNRDTAYGLKPLK
LSVVKPVAEPSTNPIFKEYIPIHLDELGNEYPVKIKEHTDDEKLMCTIK

Nucleic Acids Encoding Cas9 Molecules

Nucleic acids encoding the Cas9 molecules or Cas9 polypeptides, e.g., an eaCas9 molecule or eaCas9 polypeptide, are provided herein.

Exemplary nucleic acids encoding Cas9 molecules or Cas9 polypeptides are described in Cong *et al.*, SCIENCE 2013, 399(6121):819-823; Wang *et al.*, CELL 2013, 153(4):910-918; Mali *et al.*, SCIENCE 2013, 399(6121):823-826; Jinek *et al.*, SCIENCE 2012, 337(6096):816-821. Another

exemplary nucleic acid encoding a Cas9 molecule or Cas9 polypeptide is shown in black in **Fig. 8**.

In an embodiment, a nucleic acid encoding a Cas9 molecule or Cas9 polypeptide can be a synthetic nucleic acid sequence. For example, the synthetic nucleic acid molecule can be chemically modified, e.g., as described in Section VIII. In an embodiment, the Cas9 mRNA has one or more (e.g., all of the following properties: it is capped, polyadenylated, substituted with 5-methylcytidine and/or pseudouridine.

In addition, or alternatively, the synthetic nucleic acid sequence can be codon optimized, e.g., at least one non-common codon or less-common codon has been replaced by a common codon. For example, the synthetic nucleic acid can direct the synthesis of an optimized messenger mRNA, e.g., optimized for expression in a mammalian expression system, e.g., described herein.

In addition, or alternatively, a nucleic acid encoding a Cas9 molecule or Cas9 polypeptide may comprise a nuclear localization sequence (NLS). Nuclear localization sequences are known in the art.

Provided below is an exemplary codon optimized nucleic acid sequence encoding a Cas9 molecule of *S. pyogenes*.

```
ATGGATAAAA AGTACAGCAT CGGGCTGGAC ATCGGTACAA ACTCAGTGGG
GTGGGCCGTG ATTACGGACG AGTACAAGGT ACCCTCCAAA AAATTTAAAG
TGCTGGGTAA CACGGACAGA CACTCTATAA AGAAAAATCT TATTGGAGCC
TTGCTGTTCG ACTCAGGCGA GACAGCCGAA GCCACAAGGT TGAAGCGGAC
CGCCAGGAGG CGGTATACCA GGAGAAAGAA CCGCATATGC TACCTGCAAG
AAATCTTCAG TAACGAGATG GCAAAGGTTG ACGATAGCTT TTTCCATCGC
CTGGAAGAAT CCTTTCTTGT TGAGGAAGAC AAGAAGCACG AACGGCACCC
CATCTTTGGC AATATTGTCG ACGAAGTGGC ATATCACGAA AAGTACCCGA
CTATCTACCA CCTCAGGAAG AAGCTGGTGG ACTCTACCGA TAAGGCGGAC
CTCAGACTTA TTTATTTGGC ACTCGCCCAC ATGATTAAAT TTAGAGGACA
TTTCTTGATC GAGGGCGACC TGAACCCGGA CAACAGTGAC GTCGATAAGC
TGTTCATCCA ACTTGTGCAG ACCTACAATC AACTGTTCGA AGAAAACCCT
ATAAATGCTT CAGGAGTCGA CGCTAAAGCA ATCCTGTCCG CGCGCCTCTC
AAAATCTAGA AGACTTGAGA ATCTGATTGC TCAGTTGCCC GGGGAAAAGA
AAAATGGATT GTTTGGCAAC CTGATCGCCC TCAGTCTCGG ACTGACCCCA
AATTTCAAAA GTAACTTCGA CCTGGCCGAA GACGCTAAGC TCCAGCTGTC
CAAGGACACA TACGATGACG ACCTCGACAA TCTGCTGGCC CAGATTGGGG
ATCAGTACGC CGATCTCTTT TTGGCAGCAA AGAACCTGTC CGACGCCATC
CTGTTGAGCG ATATCTTGAG AGTGAACACC GAAATTACTA AAGCACCCCT
TAGCGCATCT ATGATCAAGC GGTACGACGA GCATCATCAG GATCTGACCC
```

TGCTGAAGGC TCTTGTGAGG CAACAGCTCC CCGAAAAATA CAAGGAAATC
TTCTTTGACC AGAGCAAAAA CGGCTACGCT GGCTATATAG ATGGTGGGGC
CAGTCAGGAG GAATTCTATA AATTCATCAA GCCCATTCTC GAGAAAATGG
ACGGCACAGA GGAGTTGCTG GTCAAACTTA ACAGGGAGGA CCTGCTGCGG
AAGCAGCGGA CCTTTGACAA CGGGTCTATC CCCCACCAGA TTCATCTGGG
CGAACTGCAC GCAATCCTGA GGAGGCAGGA GGATTTTTAT CCTTTTCTTA
AAGATAACCG CGAGAAAATA GAAAAGATTC TTACATTCAG GATCCCGTAC
TACGTGGGAC CTCTCGCCCG GGGCAATTCA CGGTTTGCCT GGATGACAAG
GAAGTCAGAG GAGACTATTA CACCTTGGAA CTTCGAAGAA GTGGTGGACA
AGGGTGCATC TGCCCAGTCT TTCATCGAGC GGATGACAAA TTTTGACAAG
AACCTCCCTA ATGAGAAGGT GCTGCCCAAA CATTCTCTGC TCTACGAGTA
CTTTACCGTC TACAATGAAC TGACTAAAGT CAAGTACGTC ACCGAGGGAA
TGAGGAAGCC GGCATTCCTT AGTGGAGAAC AGAAGAAGGC GATTGTAGAC
CTGTTGTTCA AGACCAACAG GAAGGTGACT GTGAAGCAAC TTAAAGAAGA
CTACTTTAAG AAGATCGAAT GTTTTGACAG TGTGGAAATT TCAGGGGTTG
AAGACCGCTT CAATGCGTCA TTGGGGACTT ACCATGATCT TCTCAAGATC
ATAAAGGACA AAGACTTCCT GGACAACGAA GAAAATGAGG ATATTCTCGA
AGACATCGTC CTCACCCTGA CCCTGTTCGA AGACAGGGAA ATGATAGAAG
AGCGCTTGAA AACCTATGCC CACCTCTTCG ACGATAAAGT TATGAAGCAG
CTGAAGCGCA GGAGATACAC AGGATGGGGA AGATTGTCAA GGAAGCTGAT
CAATGGAATT AGGGATAAAC AGAGTGGCAA GACCATACTG GATTTCCTCA
AATCTGATGG CTTCGCCAAT AGGAACTTCA TGCAACTGAT TCACGATGAC
TCTCTTACCT TCAAGGAGGA CATTCAAAAG GCTCAGGTGA GCGGGCAGGG
AGACTCCCTT CATGAACACA TCGCGAATTT GGCAGGTTCC CCCGCTATTA
AAAAGGGCAT CCTTCAAACT GTCAAGGTGG TGGATGAATT GGTCAAGGTA
ATGGGCAGAC ATAAGCCAGA AAATATTGTG ATCGAGATGG CCCGCGAAAA
CCAGACCACA CAGAAGGGCC AGAAAAATAG TAGAGAGCGG ATGAAGAGGA
TCGAGGAGGG CATCAAAGAG CTGGGATCTC AGATTCTCAA AGAACACCCC
GTAGAAAACA CACAGCTGCA GAACGAAAAA TTGTACTTGT ACTATCTGCA
GAACGGCAGA GACATGTACG TCGACCAAGA ACTTGATATT AATAGACTGT
CCGACTATGA CGTAGACCAT ATCGTGCCCC AGTCCTTCCT GAAGGACGAC
TCCATTGATA ACAAAGTCTT GACAAGAAGC GACAAGAACA GGGGTAAAAG
TGATAATGTG CCTAGCGAGG AGGTGGTGAA AAAAATGAAG AACTACTGGC
GACAGCTGCT TAATGCAAAG CTCATTACAC AACGGAAGTT CGATAATCTG
ACGAAAGCAG AGAGAGGTGG CTTGTCTGAG TTGGACAAGG CAGGGTTTAT
TAAGCGGCAG CTGGTGGAAA CTAGGCAGAT CACAAAGCAC GTGGCGCAGA
TTTTGGACAG CCGGATGAAC ACAAAATACG ACGAAAATGA TAAACTGATA
CGAGAGGTCA AAGTTATCAC GCTGAAAAGC AAGCTGGTGT CCGATTTTCG
GAAAGACTTC CAGTTCTACA AAGTTCGCGA GATTAATAAC TACCATCATG
CTCACGATGC GTACCTGAAC GCTGTTGTCG GGACCGCCTT GATAAAGAAG
TACCCAAAGC TGGAATCCGA GTTCGTATAC GGGGATTACA AAGTGTACGA
TGTGAGGAAA ATGATAGCCA AGTCCGAGCA GGAGATTGGA AAGGCCACAG
CTAAGTACTT CTTTTATTCT AACATCATGA ATTTTTTTAA GACGGAAATT
ACCCTGGCCA ACGGAGAGAT CAGAAAGCGG CCCCTTATAG AGACAAATGG
TGAAACAGGT GAAATCGTCT GGGATAAGGG CAGGGATTTC GCTACTGTGA
GGAAGGTGCT GAGTATGCCA CAGGTAAATA TCGTGAAAAA AACCGAAGTA
CAGACCGGAG GATTTTCCAA GGAAAGCATT TTGCCTAAAA GAAACTCAGA

CAAGCTCATC GCCCGCAAGA AAGATTGGGA CCCTAAGAAA TACGGGGGAT
TTGACTCACC CACCGTAGCC TATTCTGTGC TGGTGGTAGC TAAGGTGGAA
AAAGGAAAGT CTAAGAAGCT GAAGTCCGTG AAGGAACTCT TGGGAATCAC
TATCATGGAA AGATCATCCT TTGAAAAGAA CCCTATCGAT TTCCTGGAGG
CTAAGGGTTA CAAGGAGGTC AAGAAAGACC TCATCATTAA ACTGCCAAAA
TACTCTCTCT TCGAGCTGGA AAATGGCAGG AAGAGAATGT TGGCCAGCGC
CGGAGAGCTG CAAAAGGGAA ACGAGCTTGC TCTGCCCTCC AAATATGTTA
ATTTTCTCTA TCTCGCTTCC CACTATGAAA AGCTGAAAGG GTCTCCCGAA
GATAACGAGC AGAAGCAGCT GTTCGTCGAA CAGCACAAGC ACTATCTGGA
TGAAATAATC GAACAAATAA GCGAGTTCAG CAAAAGGGTT ATCCTGGCGG
ATGCTAATTT GGACAAAGTA CTGTCTGCTT ATAACAAGCA CCGGGATAAG
CCTATTAGGG AACAAGCCGA GAATATAATT CACCTCTTTA CACTCACGAA
TCTCGGAGCC CCCGCCGCCT TCAAATACTT TGATACGACT ATCGACCGGA
AACGGTATAC CAGTACCAAA GAGGTCCTCG ATGCCACCCT CATCCACCAG
TCAATTACTG GCCTGTACGA AACACGGATC GACCTCTCTC AACTGGGCGG
CGACTAG
(SEQ ID NO: 22)

Provided below is the corresponding amino acid sequence of a *S. pyogenes* Cas9 molecule.

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRL
KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY
HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTY
NQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF
DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS
MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRI
PYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHS
LLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
SVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA
HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF
KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQ
TTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINR
LSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKG
KSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS
AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRV
ILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD
ATLIHQSITGLYETRIDLSQLGGD*
(SEQ ID NO: 23)

Provided below is an exemplary codon optimized nucleic acid sequence encoding a Cas9 molecule of *N. meningitidis.*

```
ATGGCCGCCTTCAAGCCCAACCCCATCAACTACATCCTGGGCCTGGACATCGGCATCGCCAGCG
TGGGCTGGGCCATGGTGGAGATCGACGAGGACGAGAACCCCATCTGCCTGATCGACCTGGGTGT
GCGCGTGTTCGAGCGCGCTGAGGTGCCCAAGACTGGTGACAGTCTGGCTATGGCTCGCCGGCTT
GCTCGCTCTGTTCGGCGCCTTACTCGCCGGCGCGCTCACCGCCTTCTGCGCGCTCGCCGCCTGC
TGAAGCGCGAGGGTGTGCTGCAGGCTGCCGACTTCGACGAGAACGGCCTGATCAAGAGCCTGCC
CAACACTCCTTGGCAGCTGCGCGCTGCCGCTCTGGACCGCAAGCTGACTCCTCTGGAGTGGAGC
GCCGTGCTGCTGCACCTGATCAAGCACCGCGGCTACCTGAGCCAGCGCAAGAACGAGGGCGAGA
CCGCCGACAAGGAGCTGGGTGCTCTGCTGAAGGGCGTGGCCGACAACGCCCACGCCCTGCAGAC
TGGTGACTTCCGCACTCCTGCTGAGCTGGCCCTGAACAAGTTCGAGAAGGAGAGCGGCCACATC
CGCAACCAGCGCGGCGACTACAGCCACACCTTCAGCCGCAAGGACCTGCAGGCCGAGCTGATCC
TGCTGTTCGAGAAGCAGAAGGAGTTCGGCAACCCCCACGTGAGCGGCGGCCTGAAGGAGGGCAT
CGAGACCCTGCTGATGACCCAGCGCCCCGCCCTGAGCGGCGACGCCGTGCAGAAGATGCTGGGC
CACTGCACCTTCGAGCCAGCCGAGCCCAAGGCCGCCAAGAACACCTACACCGCCGAGCGCTTCA
TCTGGCTGACCAAGCTGAACAACCTGCGCATCCTGGAGCAGGGCAGCGAGCGCCCCCTGACCGA
CACCGAGCGCGCCACCCTGATGGACGAGCCCTACCGCAAGAGCAAGCTGACCTACGCCCAGGCC
CGCAAGCTGCTGGGTCTGGAGGACACCGCCTTCTTCAAGGGCCTGCGCTACGGCAAGGACAACG
CCGAGGCCAGCACCCTGATGGAGATGAAGGCCTACCACGCCATCAGCCGCGCCCTGGAGAAGGA
GGGCCTGAAGGACAAGAAGAGTCCTCTGAACCTGAGCCCCGAGCTGCAGGACGAGATCGGCACC
GCCTTCAGCCTGTTCAAGACCGACGAGGACATCACCGGCCGCCTGAAGGACCGCATCCAGCCCG
AGATCCTGGAGGCCCTGCTGAAGCACATCAGCTTCGACAAGTTCGTGCAGATCAGCCTGAAGGC
CCTGCGCCGCATCGTGCCCCTGATGGAGCAGGGCAAGCGCTACGACGAGGCCTGCGCCGAGATC
TACGGCGACCACTACGGCAAGAAGAACACCGAGGAGAAGATCTACCTGCCTCCTATCCCCGCCG
ACGAGATCCGCAACCCCGTGGTGCTGCGCGCCCTGAGCCAGGCCCGCAAGGTGATCAACGGCGT
GGTGCGCCGCTACGGCAGCCCCGCCCGCATCCACATCGAGACCGCCCGCGAGGTGGGCAAGAGC
TTCAAGGACCGCAAGGAGATCGAGAAGCGCCAGGAGGAGAACCGCAAGGACCGCGAGAAGGCCG
CCGCCAAGTTCCGCGAGTACTTCCCCAACTTCGTGGGCGAGCCCAAGAGCAAGGACATCCTGAA
GCTGCGCCTGTACGAGCAGCAGCACGGCAAGTGCCTGTACAGCGGCAAGGAGATCAACCTGGGC
CGCCTGAACGAGAAGGGCTACGTGGAGATCGACCACGCCCTGCCCTTCAGCCGCACCTGGGACG
ACAGCTTCAACAACAAGGTGCTGGTGCTGGGCAGCGAGAACCAGAACAAGGGCAACCAGACCCC
CTACGAGTACTTCAACGGCAAGGACAACAGCCGCGAGTGGCAGGAGTTCAAGGCCCGCGTGGAG
ACCAGCCGCTTCCCCCGCAGCAAGAAGCAGCGCATCCTGCTGCAGAAGTTCGACGAGGACGGCT
TCAAGGAGCGCAACCTGAACGACACCCGCTACGTGAACCGCTTCCTGTGCCAGTTCGTGGCCGA
CCGCATGCGCCTGACCGGCAAGGGCAAGAAGCGCGTGTTCGCCAGCAACGGCCAGATCACCAAC
CTGCTGCGCGGCTTCTGGGGCCTGCGCAAGGTGCGCGCCGAGAACGACCGCCACCACGCCCTGG
ACGCCGTGGTGGTGGCCTGCAGCACCGTGGCCATGCAGCAGAAGATCACCCGCTTCGTGCGCTA
CAAGGAGATGAACGCCTTCGACGGTAAAACCATCGACAAGGAGACCGGCGAGGTGCTGCACCAG
AAGACCCACTTCCCCCAGCCCTGGGAGTTCTTCGCCCAGGAGGTGATGATCCGCGTGTTCGGCA
AGCCCGACGGCAAGCCCGAGTTCGAGGAGGCCGACACCCCCGAGAAGCTGCGCACCCTGCTGGC
CGAGAAGCTGAGCAGCCGCCCTGAGGCCGTGCACGAGTACGTGACTCCTCTGTTCGTGAGCCGC
GCCCCCAACCGCAAGATGAGCGGTCAGGGTCACATGGAGACCGTGAAGAGCGCCAAGCGCCTGG
ACGAGGGCGTGAGCGTGCTGCGCGTGCCCCTGACCCAGCTGAAGCTGAAGGACCTGGAGAAGAT
GGTGAACCGCGAGCGCGAGCCCAAGCTGTACGAGGCCCTGAAGGCCCGCCTGGAGGCCCACAAG
GACGACCCCGCCAAGGCCTTCGCCGAGCCCTTCTACAAGTACGACAAGGCCGGCAACCGCACCC
AGCAGGTGAAGGCCGTGCGCGTGGAGCAGGTGCAGAAGACCGGCGTGTGGGTGCGCAACCACAA
```

CGGCATCGCCGACAACGCCACCATGGTGCGCGTGGACGTGTTCGAGAAGGGCGACAAGTACTAC
CTGGTGCCCATCTACAGCTGGCAGGTGGCCAAGGGCATCCTGCCCGACCGCGCCGTGGTGCAGG
GCAAGGACGAGGAGGACTGGCAGCTGATCGACGACAGCTTCAACTTCAAGTTCAGCCTGCACCC
CAACGACCTGGTGGAGGTGATCACCAAGAAGGCCCGCATGTTCGGCTACTTCGCCAGCTGCCAC
CGCGGCACCGGCAACATCAACATCCGCATCCACGACCTGGACCACAAGATCGGCAAGAACGGCA
TCCTGGAGGGCATCGGCGTGAAGACCGCCCTGAGCTTCCAGAAGTACCAGATCGACGAGCTGGG
CAAGGAGATCCGCCCCTGCCGCCTGAAGAAGCGCCCTCCTGTGCGCTAA
(SEQ ID NO: 24)

Provided below is the corresponding amino acid sequence of a *N. meningitidis* Cas9 molecule.

MAAFKPNPINYILGLDIGIASVGWAMVEIDEDENPICLIDLGVRVFERAEVPKTGDSLAMARRL
ARSVRRLTRRRAHRLLRARRLLKREGVLQAADFDENGLIKSLPNTPWQLRAAALDRKLTPLEWS
AVLLHLIKHRGYLSQRKNEGETADKELGALLKGVADNAHALQTGDFRTPAELALNKFEKESGHI
RNQRGDYSHTFSRKDLQAELILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLG
HCTFEPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSERPLTDTERATLMDEPYRKSKLTYAQA
RKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSPELQDEIGT
AFSLFKTDEDITGRLKDRIQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEI
YGDHYGKKNTEEKIYLPPIPADEIRNPVVLRALSQARKVINGVVRRYGSPARIHIETAREVGKS
FKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEINLG
RLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVE
TSRFPRSKKQRILLQKFDEDGFKERNLNDTRYVNRFLCQFVADRMRLTGKGKKRVFASNGQITN
LLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGEVLHQ
KTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSR
APNRKMSGQGHMETVKSAKRLDEGVSVLRVPLTQLKLKDLEKMVNREREPKLYEALKARLEAHK
DDPAKAFAEPFYKYDKAGNRTQQVKAVRVEQVQKTGVWVRNHNGIADNATMVRVDVFEKGDKYY
LVPIYSWQVAKGILPDRAVVQGKDEEDWQLIDDSFNFKFSLHPNDLVEVITKKARMFGYFASCH
RGTGNINIRIHDLDHKIGKNGILEGIGVKTALSFQKYQIDELGKEIRPCRLKKRPPVR*
(SEQ ID NO: 25)

Provided below is an amino acid sequence of a *S. aureus* Cas9 molecule.

MKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRI
QRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDT
GNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQ
LDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLY
NALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGK
PEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQIS
NLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSP
VVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTT
GKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVK
QEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKD
FINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAED
ALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKD
YKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHH
DPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDD
YPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQA

EFIASFYNNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKT
QSIKKYSTDILGNLYEVKSKKHPQIIKKG*
(SEQ ID NO: 26)

Provided below is an exemplary codon optimized nucleic acid sequence encoding a Cas9 molecule of *S. aureus* Cas9.

ATGAAAAGGAACTACATTCTGGGGCTGGACATCGGGATTACAAGCGTGGGGTATGGGATTATTG
ACTATGAAACAAGGGACGTGATCGACGCAGGCGTCAGACTGTTCAAGGAGGCCAACGTGGAAAA
CAATGAGGGACGGAGAAGCAAGAGGGGAGCCAGGCGCCTGAAACGACGGAGAAGGCACAGAATC
CAGAGGGTGAAGAAACTGCTGTTCGATTACAACCTGCTGACCGACCATTCTGAGCTGAGTGGAA
TTAATCCTTATGAAGCCAGGGTGAAAGGCCTGAGTCAGAAGCTGTCAGAGGAAGAGTTTTCCGC
AGCTCTGCTGCACCTGGCTAAGCGCCGAGGAGTGCATAACGTCAATGAGGTGGAAGAGGACACC
GGCAACGAGCTGTCTACAAAGGAACAGATCTCACGCAATAGCAAAGCTCTGGAAGAGAAGTATG
TCGCAGAGCTGCAGCTGGAACGGCTGAAGAAAGATGGCGAGGTGAGAGGGTCAATTAATAGGTT
CAAGACAAGCGACTACGTCAAAGAAGCCAAGCAGCTGCTGAAAGTGCAGAAGGCTTACCACCAG
CTGGATCAGAGCTTCATCGATACTTATATCGACCTGCTGGAGACTCGGAGAACCTACTATGAGG
GACCAGGAGAAGGGAGCCCCTTCGGATGGAAAGACATCAAGGAATGGTACGAGATGCTGATGGG
ACATTGCACCTATTTTCCAGAAGAGCTGAGAAGCGTCAAGTACGCTTATAACGCAGATCTGTAC
AACGCCCTGAATGACCTGAACAACCTGGTCATCACCAGGGATGAAAACGAGAAACTGGAATACT
ATGAGAAGTTCCAGATCATCGAAAACGTGTTTAAGCAGAAGAAAAAGCCTACACTGAAACAGAT
TGCTAAGGAGATCCTGGTCAACGAAGAGGACATCAAGGGCTACCGGGTGACAAGCACTGGAAAA
CCAGAGTTCACCAATCTGAAAGTGTATCACGATATTAAGGACATCACAGCACGGAAAGAAATCA
TTGAGAACGCCGAACTGCTGGATCAGATTGCTAAGATCCTGACTATCTACCAGAGCTCCGAGGA
CATCCAGGAAGAGCTGACTAACCTGAACAGCGAGCTGACCCAGGAAGAGATCGAACAGATTAGT
AATCTGAAGGGGTACACCGGAACACACAACCTGTCCCTGAAAGCTATCAATCTGATTCTGGATG
AGCTGTGGCATACAAACGACAATCAGATTGCAATCTTTAACCGGCTGAAGCTGGTCCCAAAAAA
GGTGGACCTGAGTCAGCAGAAAGAGATCCCAACCACACTGGTGGACGATTTCATTCTGTCACCC
GTGGTCAAGCGGAGCTTCATCCAGAGCATCAAAGTGATCAACGCCATCATCAAGAAGTACGGCC
TGCCCAATGATATCATTATCGAGCTGGCTAGGGAGAAGAACAGCAAGGACGCACAGAAGATGAT
CAATGAGATGCAGAAACGAAACCGGCAGACCAATGAACGCATTGAAGAGATTATCCGAACTACC
GGGAAAGAGAACGCAAAGTACCTGATTGAAAAAATCAAGCTGCACGATATGCAGGAGGGAAAGT
GTCTGTATTCTCTGGAGGCCATCCCCCTGGAGGACCTGCTGAACAATCCATTCAACTACGAGGT
CGATCATATTATCCCCAGAAGCGTGTCCTTCGACAATTCCTTTAACAACAAGGTGCTGGTCAAG
CAGGAAGAGAACTCTAAAAAGGGCAATAGGACTCCTTTCCAGTACCTGTCTAGTTCAGATTCCA
AGATCTCTTACGAAACCTTTAAAAAGCACATTCTGAATCTGGCCAAAGGAAAGGGCCGCATCAG
CAAGACCAAAAAGGAGTACCTGCTGGAAGAGCGGGACATCAACAGATTCTCCGTCCAGAAGGAT
TTTATTAACCGGAATCTGGTGGACACAAGATACGCTACTCGCGGCCTGATGAATCTGCTGCGAT
CCTATTTCCGGGTGAACAATCTGGATGTGAAAGTCAAGTCCATCAACGGCGGGTTCACATCTTT
TCTGAGGCGCAAATGGAAGTTTAAAAAGGAGCGCAACAAAGGGTACAAGCACCATGCCGAAGAT
GCTCTGATTATCGCAAATGCCGACTTCATCTTTAAGGAGTGGAAAAAGCTGGACAAAGCCAAGA
AAGTGATGGAGAACCAGATGTTCGAAGAGAAGCAGGCCGAATCTATGCCCGAAATCGAGACAGA
ACAGGAGTACAAGGAGATTTTCATCACTCCTCACCAGATCAAGCATATCAAGGATTTCAAGGAC
TACAAGTACTCTCACCGGGTGGATAAAAAGCCCAACAGAGAGCTGATCAATGACACCCTGTATA
GTACAAGAAAAGACGATAAGGGGAATACCCTGATTGTGAACAATCTGAACGGACTGTACGACAA
AGATAATGACAAGCTGAAAAAGCTGATCAACAAAAGTCCCGAGAAGCTGCTGATGTACCACCAT

GATCCTCAGACATATCAGAAACTGAAGCTGATTATGGAGCAGTACGGCGACGAGAAGAACCCAC
TGTATAAGTACTATGAAGAGACTGGGAACTACCTGACCAAGTATAGCAAAAAGGATAATGGCCC
CGTGATCAAGAAGATCAAGTACTATGGGAACAAGCTGAATGCCCATCTGGACATCACAGACGAT
TACCCTAACAGTCGCAACAAGGTGGTCAAGCTGTCACTGAAGCCATACAGATTCGATGTCTATC
TGGACAACGGCGTGTATAAATTTGTGACTGTCAAGAATCTGGATGTCATCAAAAAGGAGAACTA
CTATGAAGTGAATAGCAAGTGCTACGAAGAGGCTAAAAAGCTGAAAAAGATTAGCAACCAGGCA
GAGTTCATCGCCTCCTTTTACAACAACGACCTGATTAAGATCAATGGCGAACTGTATAGGGTCA
TCGGGGTGAACAATGATCTGCTGAACCGCATTGAAGTGAATATGATTGACATCACTTACCGAGA
GTATCTGGAAAACATGAATGATAAGCGCCCCCCTCGAATTATCAAAACAATTGCCTCTAAGACT
CAGAGTATCAAAAAGTACTCAACCGACATTCTGGGAAACCTGTATGAGGTGAAGAGCAAAAAGC
ACCCTCAGATTATCAAAAAGGGC
(SEQ ID NO: 39)

If any of the above Cas9 sequences are fused with a peptide or polypeptide at the C-terminus, it is understood that the stop codon will be removed.

Other Cas Molecules and Cas Polypeptides

Various types of Cas molecules or Cas polypeptides can be used to practice the inventions disclosed herein. In some embodiments, Cas molecules of Type II Cas systems are used. In other embodiments, Cas molecules of other Cas systems are used. For example, Type I or Type III Cas molecules may be used. Exemplary Cas molecules (and Cas systems) are described, e.g., in Haft *et al.*, PLoS COMPUTATIONAL BIOLOGY 2005, 1(6): e60 and Makarova *et al.*, NATURE REVIEW MICROBIOLOGY 2011, 9:467-477, the contents of both references are incorporated herein by reference in their entirety. Exemplary Cas molecules (and Cas systems) are also shown in **Table 16.**

Table 16: Cas Systems

Gene name‡	System type or subtype	Name from Haft *et al.*§	Structure of encoded protein (PDB accessions)¶	Families (and superfamily) of encoded protein#**	Representatives
cas1	• Type I • Type II • Type III	*cas1*	3GOD, 3LFX and 2YZS	COG1518	SERP2463, SPy1047 and *ygbT*
cas2	• Type I • Type II • Type III	*cas2*	2IVY, 2I8E and 3EXC	COG1343 and COG3512	SERP2462, SPy1048, SPy1723 (N-terminal domain) and *ygbF*
cas3′	• Type I‡‡	*cas3*	NA	COG1203	APE1232 and *ygcB*
cas3″	• Subtype I-A • Subtype I-B	NA	NA	COG2254	APE1231 and BH0336

Table 16: Cas Systems

Gene name[‡]	System type or subtype	Name from Haft *et al.*[§]	Structure of encoded protein (PDB accessions)[¶]	Families (and superfamily) of encoded protein[#**]	Representatives
cas4	• Subtype I-A • Subtype I-B • Subtype I-C • Subtype I-D • Subtype II-B	*cas4* and *csa1*	NA	COG1468	APE1239 and BH0340
cas5	• Subtype I-A • Subtype I-B • Subtype I-C • Subtype I-E	*cas5a*, *cas5d*, *cas5e*, *cas5h*, *cas5p*, *cas5t* and *cmx5*	3KG4	COG1688 (RAMP)	APE1234, BH0337, *devS* and *ygcI*
cas6	• Subtype I-A • Subtype I-B • Subtype I-D • Subtype III-A• Subtype III-B	*cas6* and *cmx6*	3I4H	COG1583 and COG5551 (RAMP)	PF1131 and slr7014
cas6e	• Subtype I-E	*cse3*	1WJ9	(RAMP)	*ygcH*
cas6f	• Subtype I-F	*csy4*	2XLJ	(RAMP)	y1727
cas7	• Subtype I-A • Subtype I-B • Subtype I-C • Subtype I-E	*csa2*, *csd2*, *cse4*, *csh2*, *csp1* and *cst2*	NA	COG1857 and COG3649 (RAMP)	*devR* and *ygcJ*
cas8a1	• Subtype I-A[‡‡]	*cmx1*, *cst1*, *csx8*, *csx13* and CXXC-CXXC	NA	BH0338-like	LA3191[§§] and PG2018[§§]
cas8a2	• Subtype I-A[‡‡]	*csa4* and *csx9*	NA	PH0918	AF0070, AF1873, MJ0385, PF0637, PH0918 and SSO1401
cas8b	• Subtype I-B[‡‡]	*csh1* and TM1802	NA	BH0338-like	MTH1090 and TM1802
cas8c	• Subtype I-C[‡‡]	*csd1* and *csp2*	NA	BH0338-like	BH0338
cas9	• Type II[‡‡]	*csn1* and *csx12*	NA	COG3513	FTN_0757 and SPy1046
cas10	• Type III[‡‡]	*cmr2*, *csm1* and *csx11*	NA	COG1353	MTH326, Rv2823c[§§] and TM1794[§§]
cas10d	• Subtype I-D[‡‡]	*csc3*	NA	COG1353	slr7011
csy1	• Subtype I-F[‡‡]	*csy1*	NA	y1724-like	y1724
csy2	• Subtype I-F	*csy2*	NA	(RAMP)	y1725

Table 16: Cas Systems

Gene name[‡]	System type or subtype	Name from Haft *et al.*[§]	Structure of encoded protein (PDB accessions)[¶]	Families (and superfamily) of encoded protein[#**]	Representatives
csy3	• Subtype I-F	*csy3*	NA	(RAMP)	y1726
cse1	• Subtype I-E[‡‡]	*cse1*	NA	YgcL-like	*ygcL*
cse2	• Subtype I-E	*cse2*	2ZCA	YgcK-like	*ygcK*
csc1	• Subtype I-D	*csc1*	NA	alr1563-like (RAMP)	alr1563
csc2	• Subtype I-D	*csc1* and *csc2*	NA	COG1337 (RAMP)	slr7012
csa5	• Subtype I-A	*csa5*	NA	AF1870	AF1870, MJ0380, PF0643 and SSO1398
csn2	• Subtype II-A	*csn2*	NA	SPy1049-like	SPy1049
csm2	• Subtype III-A[‡‡]	*csm2*	NA	COG1421	MTH1081 and SERP2460
csm3	• Subtype III-A	*csc2* and *csm3*	NA	COG1337 (RAMP)	MTH1080 and SERP2459
csm4	• Subtype III-A	*csm4*	NA	COG1567 (RAMP)	MTH1079 and SERP2458
csm5	• Subtype III-A	*csm5*	NA	COG1332 (RAMP)	MTH1078 and SERP2457
csm6	• Subtype III-A	APE2256 and *csm6*	2WTE	COG1517	APE2256 and SSO1445
cmr1	• Subtype III-B	*cmr1*	NA	COG1367 (RAMP)	PF1130
cmr3	• Subtype III-B	*cmr3*	NA	COG1769 (RAMP)	PF1128
cmr4	• Subtype III-B	*cmr4*	NA	COG1336 (RAMP)	PF1126
cmr5	• Subtype III-B[‡‡]	*cmr5*	2ZOP and 2OEB	COG3337	MTH324 and PF1125
cmr6	• Subtype III-B	*cmr6*	NA	COG1604 (RAMP)	PF1124
csb1	• Subtype I-U	GSU0053	NA	(RAMP)	Balac_1306 and GSU0053
csb2	• Subtype I-U[§§]	NA	NA	(RAMP)	Balac_1305 and GSU0054
csb3	• Subtype I-U	NA	NA	(RAMP)	Balac_1303[§§]
csx17	• Subtype I-U	NA	NA	NA	Btus_2683
csx14	• Subtype I-U	NA	NA	NA	GSU0052

Table 16: Cas Systems					
Gene name[‡]	**System type or subtype**	**Name from Haft *et al.*[§]**	**Structure of encoded protein (PDB accessions)[¶]**	**Families (and superfamily) of encoded protein[#**]**	**Representatives**
csx10	• Subtype I-U	*csx10*	NA	(RAMP)	Caur_2274
csx16	• Subtype III-U	VVA1548	NA	NA	VVA1548
csaX	• Subtype III-U	*csaX*	NA	NA	SSO1438
csx3	• Subtype III-U	*csx3*	NA	NA	AF1864
csx1	• Subtype III-U	*csa3*, *csx1*, *csx2*, DXTHG, NE0113 and TIGR02710	1XMX and 2I71	COG1517 and COG4006	MJ1666, NE0113, PF1127 and TM1812
csx15	• Unknown	NA	NA	TTE2665	TTE2665
csf1	• Type U	*csf1*	NA	NA	AFE_1038
csf2	• Type U	*csf2*	NA	(RAMP)	AFE_1039
csf3	• Type U	*csf3*	NA	(RAMP)	AFE_1040
csf4	• Type U	*csf4*	NA	NA	AFE_1037

IV. Functional Analysis of Candidate Molecules

Candidate Cas9 molecules, candidate gRNA molecules, candidate Cas9 molecule/gRNA molecule complexes, can be evaluated by art-known methods or as described herein. For example, exemplary methods for evaluating the endonuclease activity of Cas9 molecule are described, e.g., in Jinek *et al.*, SCIENCE 2012, 337(6096):816-821.

Binding and Cleavage Assay: Testing the endonuclease activity of Cas9 molecule

The ability of a Cas9 molecule/gRNA molecule complex to bind to and cleave a target nucleic acid can be evaluated in a plasmid cleavage assay. In this assay, synthetic or *in vitro*-transcribed gRNA molecule is pre-annealed prior to the reaction by heating to 95°C and slowly cooling down to room temperature. Native or restriction digest-linearized plasmid DNA (300 ng (~8 nM)) is incubated for 60 min at 37°C with purified Cas9 protein molecule (50-500 nM) and gRNA (50-500 nM, 1:1) in a Cas9 plasmid cleavage buffer (20 mM HEPES pH 7.5, 150 mM KCl, 0.5 mM DTT, 0.1 mM EDTA) with or without 10 mM $MgCl_2$. The reactions are stopped with 5X DNA loading buffer (30% glycerol, 1.2% SDS, 250 mM EDTA), resolved by a 0.8 or 1% agarose gel electrophoresis and visualized by ethidium bromide staining. The resulting

cleavage products indicate whether the Cas9 molecule cleaves both DNA strands, or only one of the two strands. For example, linear DNA products indicate the cleavage of both DNA strands. Nicked open circular products indicate that only one of the two strands is cleaved.

Alternatively, the ability of a Cas9 molecule/gRNA molecule complex to bind to and cleave a target nucleic acid can be evaluated in an oligonucleotide DNA cleavage assay. In this assay, DNA oligonucleotides (10 pmol) are radiolabeled by incubating with 5 units T4 polynucleotide kinase and ~3–6 pmol (~20–40 mCi) [γ-32P]-ATP in 1X T4 polynucleotide kinase reaction buffer at 37°C for 30 min, in a 50 μL reaction. After heat inactivation (65°C for 20 min), reactions are purified through a column to remove unincorporated label. Duplex substrates (100 nM) are generated by annealing labeled oligonucleotides with equimolar amounts of unlabeled complementary oligonucleotide at 95°C for 3 min, followed by slow cooling to room temperature. For cleavage assays, gRNA molecules are annealed by heating to 95°C for 30 s, followed by slow cooling to room temperature. Cas9 (500 nM final concentration) is pre-incubated with the annealed gRNA molecules (500 nM) in cleavage assay buffer (20 mM HEPES pH 7.5, 100 mM KCl, 5 mM MgCl2, 1 mM DTT, 5% glycerol) in a total volume of 9 μl. Reactions are initiated by the addition of 1 μl target DNA (10 nM) and incubated for 1 h at 37°C. Reactions are quenched by the addition of 20 μl of loading dye (5 mM EDTA, 0.025% SDS, 5% glycerol in formamide) and heated to 95°C for 5 min. Cleavage products are resolved on 12% denaturing polyacrylamide gels containing 7 M urea and visualized by phosphorimaging. The resulting cleavage products indicate that whether the complementary strand, the non-complementary strand, or both, are cleaved.

One or both of these assays can be used to evaluate the suitability of a candidate gRNA molecule or candidate Cas9 molecule.

Binding Assay: Testing the binding of Cas9 molecule to target DNA

Exemplary methods for evaluating the binding of Cas9 molecule to target DNA are described, e.g., in Jinek *et al.*, SCIENCE 2012; 337(6096):816-821.

For example, in an electrophoretic mobility shift assay, target DNA duplexes are formed by mixing of each strand (10 nmol) in deionized water, heating to 95°C for 3 min and slow cooling to room temperature. All DNAs are purified on 8% native gels containing 1X TBE. DNA bands are visualized by UV shadowing, excised, and eluted by soaking gel pieces in

DEPC-treated H_2O. Eluted DNA is ethanol precipitated and dissolved in DEPC-treated H_2O. DNA samples are 5' end labeled with [γ-32P]-ATP using T4 polynucleotide kinase for 30 min at 37°C. Polynucleotide kinase is heat denatured at 65°C for 20 min, and unincorporated radiolabel is removed using a column. Binding assays are performed in buffer containing 20 mM HEPES pH 7.5, 100 mM KCl, 5 mM $MgCl_2$, 1 mM DTT and 10% glycerol in a total volume of 10 μl. Cas9 protein molecule is programmed with equimolar amounts of pre-annealed gRNA molecule and titrated from 100 pM to 1 μM. Radiolabeled DNA is added to a final concentration of 20 pM. Samples are incubated for 1 h at 37°C and resolved at 4°C on an 8% native polyacrylamide gel containing 1X TBE and 5 mM $MgCl_2$. Gels are dried and DNA visualized by phosphorimaging.

Differential Scanning Flourimetry (DSF)

The thermostability of Cas9-gRNA ribonucleoprotein (RNP) complexes can be measured via DSF. This technique measures the thermostability of a protein, which can increase under favorable conditions such as the addition of a binding RNA molecule, e.g., a gRNA.

The assay is performed using two different protocols, one to test the best stoichiometric ratio of gRNA:Cas9 protein and another to determine the best solution conditions for RNP formation.

To determine the best solution to form RNP complexes, a 2uM solution of Cas9 in water+10x SYPRO Orange® (Life Techonologies cat#S-6650) and dispensed into a 384 well plate. An equimolar amount of gRNA diluted in solutions with varied pH and salt is then added. After incubating at room temperature for 10'and brief centrifugation to remove any bubbles,a Bio-Rad CFX384™ Real-Time System C1000 Touch™ Thermal Cycler with the Bio-Rad CFX Manager software is used to run a gradient from 20°C to 90°C with a 1° increase in temperature every 10seconds.

The second assay consists of mixing various concentrations of gRNA with 2uM Cas9 in optimal buffer from assay 1 above and incubating at RT for 10' in a 384 well plate. An equal volume of optimal buffer + 10x SYPRO Orange® (Life Techonologies cat#S-6650) is added and the plate sealed with Microseal® B adhesive (MSB-1001). Following brief centrifugation to remove any bubbles, a Bio-Rad CFX384™ Real-Time System C1000 Touch™ Thermal Cycler

with the Bio-Rad CFX Manager software is used to run a gradient from 20°C to 90°C with a 1° increase in temperature every 10seconds.

V. Genome Editing Approaches

While not wishing to be bound by theory, altering the LCA10 target position may be achieved using one of the approaches discussed herein.

V.1 NHEJ Approaches for Gene Targeting

As described herein, nuclease-induced non-homologous end-joining (NHEJ) can be used to introduce indels at a target position. Nuclease-induced NHEJ can also be used to remove (e.g., delete) genomic sequence including the mutation at a target position in a gene of interest.

While not wishing to be bound by theory, it is believed that, in an embodiment, the genomic alterations associated with the methods described herein rely on nuclease-induced NHEJ and the error-prone nature of the NHEJ repair pathway. NHEJ repairs a double-strand break in the DNA by joining together the two ends; however, generally, the original sequence is restored only if two compatible ends, exactly as they were formed by the double-strand break, are perfectly ligated. The DNA ends of the double-strand break are frequently the subject of enzymatic processing, resulting in the addition or removal of nucleotides, at one or both strands, prior to rejoining of the ends. This results in the presence of insertion and/or deletion (indel) mutations in the DNA sequence at the site of the NHEJ repair.

The indel mutations generated by NHEJ are unpredictable in nature; however, at a given break site certain indel sequences are favored and are over represented in the population, likely due to small regions of microhomology. The lengths of deletions can vary widely; most commonly in the 1-50 bp range, but they can easily reach greater than 100-200 bp. Insertions tend to be shorter and often include short duplications of the sequence immediately surrounding the break site. However, it is possible to obtain large insertions, and in these cases, the inserted sequence has often been traced to other regions of the genome or to plasmid DNA present in the cells.

Because NHEJ is a mutagenic process, it can also be used to delete small sequence motifs as long as the generation of a specific final sequence is not required. If a double-strand break is targeted near to a short target sequence, the deletion mutations caused by the NHEJ repair often span, and therefore remove, the unwanted nucleotides. For the deletion of larger DNA segments,

introducing two double-strand breaks, one on each side of the sequence, can result in NHEJ between the ends with removal of the entire intervening sequence. Both of these approaches can be used to delete specific DNA sequences; however, the error-prone nature of NHEJ may still produce indel mutations at the site of deletion.

Both double strand cleaving eaCas9 molecules and single strand, or nickase, eaCas9 molecules can be used in the methods and compositions described herein to generate break-induced indels.

Double strand break

In an embodiment, double strand cleavage is effected by a Cas9 molecule having cleavage activity associated with an HNH-like domain and cleavage activity associated with a RuvC-like domain, e.g., an N-terminal RuvC-like domain, e.g., a wild type Cas9. Such embodiments require only a single gRNA.

Single strand break

In other embodiments, two single strand breaks are effected by a Cas9 molecule having nickase activity, e.g., cleavage activity associated with an HNH-like domain or cleavage activity associated with an N-terminal RuvC-like domain. Such embodiments require two gRNAs, one for placement of each single strand break. In an embodiment, the Cas9 molecule having nickase activity cleaves the strand to which the gRNA hybridizes, but not the strand that is complementary to the strand to which the gRNA hybridizes. In an embodiment, the Cas9 molecule having nickase activity does not cleave the strand to which the gRNA hybridizes, but rather cleaves the strand that is complementary to the strand to which the gRNA hybridizes.

In an embodiment, the nickase has HNH activity, e.g., a Cas9 molecule having the RuvC activity inactivated, e.g., a Cas9 molecule having a mutation at D10, e.g., the D10A mutation. D10A inactivates RuvC therefore the Cas9 nickase has (only) HNH activity and will cut on the strand to which the gRNA hybridizes (the complementary strand, which does not have the NGG PAM on it). In other embodiments, a Cas9 molecule having an H840, e.g., an H840A, mutation can be used as a nickase. H840A inactivates HNH therefore the Cas9 nickase has (only) RuvC activity and cuts on the non-complementary strand (the strand that has the NGG PAM and whose sequence is identical to the gRNA). In other embodiments, a Cas9 molecule having an H863, e.g., an H863A, mutation can be used as a nickase. H863A inactivates HNH therefore the Cas9

nickase has (only) RuvC activity and cuts on the non-complementary strand (the strand that has the NGG PAM and whose sequence is identical to the gRNA).

In an embodiment, in which a nickase and two gRNAs are used to position two single strand breaks, one nick is on the + strand and one nick is on the – strand of the target nucleic acid. The PAMs can be outwardly facing. The gRNAs can be selected such that the gRNAs are separated by, from 0-50, 0-100, or 0-200 nucleotides. In an embodiment, there is no overlap between the target sequences that are complementary to the targeting domains of the two gRNAs. In an embodiment, the gRNAs do not overlap and are separated by as much as 50, 100, or 200 nucleotides. In an embodiment, the use of two gRNAs can increase specificity, e.g., by decreasing off-target binding (Ran *et al.*, Cell 2013; 154(6):1380-1389).

Placement of double strand or single strand breaks relative to the target position

In an embodiment, in which a gRNA and Cas9 nuclease generate a double strand break for the purpose of inducing break-induced indels, a gRNA, e.g., a unimolecular (or chimeric) or modular gRNA molecule, is configured to position one double-strand break in close proximity to a nucleotide of the target position. In an embodiment, the cleavage site is between 0-40 bp away from the target position (e.g., less than 40, 35, 30, 25, 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 bp from the target position).

In an embodiment, in which two gRNAs complexing with a Cas9 nickase induce two single strand breaks for the purpose of introducing break-induced indels, two gRNAs, e.g., independently, unimolecular (or chimeric) or modular gRNA, are configured to position two single-strand breaks to provide for NHEJ-mediated alteration of a nucleotide of the target position. In an embodiment, the gRNAs are configured to position cuts at the same position, or within a few nucleotides of one another, on different strands, essentially mimicking a double strand break. In an embodiment, the two nicks are between 0-40 bp away from the target position (e.g., less than 40, 35, 30, 25, 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 bp from the target position) respectively, and the two single strand breaks are within 25-55 bp of each other (e.g., between 25 to 50, 25 to 45, 25 to 40, 25 to 35, 25 to 30, 50 to 55, 45 to 55, 40 to 55, 35 to 55, 30 to 55, 30 to 50, 35 to 50, 40 to 50 , 45 to 50, 35 to 45, or 40 to 45 bp) and no more than 100 bp away from each other (e.g., no more than 90, 80, 70, 60, 50, 40, 30, 20 or 10 bp). In an embodiment, the gRNAs are configured to place a single strand break on either side of the target

position. In an embodiment, the gRNAs are configured to place a single strand break on the same side (either 5' or 3') of the target position.

Regardless of whether a break is a double strand or a single strand break, the gRNA should be configured to avoid unwanted target chromosome elements, such as repeated elements, e.g., an *Alu* repeat, in the target domain. In addition, a break, whether a double strand or a single strand break, should be sufficiently distant from any sequence that should not be altered. For example, cleavage sites positioned within introns should be sufficiently distant from any intron/exon border, or naturally occurring splice signal, to avoid alteration of the exonic sequence or unwanted splicing events.

V.2 Single-Strand Annealing

Single strand annealing (SSA) is another DNA repair process that repairs a double-strand break between two repeat sequences present in a target nucleic acid. Repeat sequences utilized by the SSA pathway are generally greater than 30 nucleotides in length. Resection at the break ends occurs to reveal repeat sequences on both strands of the target nucleic acid. After resection, single strand overhangs containing the repeat sequences are coated with RPA protein to prevent the repeats sequences from inappropriate annealing, e.g., to themselves. RAD52 binds to and each of the repeat sequences on the overhangs and aligns the sequences to enable the annealing of the complementary repeat sequences. After annealing, the single-strand flaps of the overhangs are cleaved. New DNA synthesis fills in any gaps, and ligation restores the DNA duplex. As a result of the processing, the DNA sequence between the two repeats is deleted. The length of the deletion can depend on many factors including the location of the two repeats utilized, and the pathway or processivity of the resection.

In contrast to HDR pathways, SSA does not require a template nucleic acid to alter or correct a target nucleic acid sequence. Instead, the complementary repeat sequence is utilized.

V. 3 Other DNA Repair Pathways

SSBR (single strand break repair)

Single-stranded breaks (SSB) in the genome are repaired by the SSBR pathway, which is a distinct mechanism from the DSB repair mechanisms discussed above. The SSBR pathway has four major stages: SSB detection, DNA end processing, DNA gap filling, and DNA ligation.

A more detailed explanation is given in Caldecott, Nature Reviews Genetics 9, 619-631 (August 2008), and a summary is given here.

In the first stage, when a SSB forms, PARP1 and/or PARP2 recognize the break and recruit repair machinery. The binding and activity of PARP1 at DNA breaks is transient and it seems to accelerate SSBr by promoting the focal accumulation or stability of SSBr protein complexes at the lesion. Arguably the most important of these SSBr proteins is XRCC1, which functions as a molecular scaffold that interacts with, stabilizes, and stimulates multiple enzymatic components of the SSBr process including the protein responsible for cleaning the DNA 3' and 5' ends. For instance, XRCC1 interacts with several proteins (DNA polymerase beta, PNK, and three nucleases, APE1, APTX, and APLF) that promote end processing. APE1 has endonuclease activity. APLF exhibits endonuclease and 3' to 5' exonuclease activities. APTX has endonuclease and 3' to 5' exonuclease activity.

This end processing is an important stage of SSBR since the 3'- and/or 5'-termini of most, if not all, SSBs are 'damaged'. End processing generally involves restoring a damaged 3'-end to a hydroxylated state and and/or a damaged 5' end to a phosphate moiety, so that the ends become ligation-competent. Enzymes that can process damaged 3' termini include PNKP, APE1, and TDP1. Enzymes that can process damaged 5' termini include PNKP, DNA polymerase beta, and APTX. LIG3 (DNA ligase III) can also participate in end processing. Once the ends are cleaned, gap filling can occur.

At the DNA gap filling stage, the proteins typically present are PARP1, DNA polymerase beta, XRCC1, FEN1 (flap endonculease 1), DNA polymerase delta/epsilon, PCNA, and LIG1. There are two ways of gap filling, the short patch repair and the long patch repair. Short patch repair involves the insertion of a single nucleotide that is missing. At some SSBs, "gap filling" might continue displacing two or more nucleotides (displacement of up to 12 bases have been reported). FEN1 is an endonuclease that removes the displaced 5'-residues. Multiple DNA polymerases, including Pol β, are involved in the repair of SSBs, with the choice of DNA polymerase influenced by the source and type of SSB.

In the fourth stage, a DNA ligase such as LIG1 (Ligase I) or LIG3 (Ligase III) catalyzes joining of the ends. Short patch repair uses Ligase III and long patch repair uses Ligase I.

Sometimes, SSBR is replication-coupled. This pathway can involve one or more of CtIP, MRN, ERCC1, and FEN1. Additional factors that may promote SSBR include: aPARP,

PARP1, PARP2, PARG, XRCC1, DNA polymerase b, DNA polymerase d, DNA polymerase e, PCNA, LIG1, PNK, PNKP, APE1, APTX, APLF, TDP1, LIG3, FEN1, CtIP, MRN, and ERCC1.

MMR (mismatch repair)

Cells contain three excision repair pathways: MMR, BER, and NER. The excision repair pathways hace a common feature in that they typically recognize a lesion on one strand of the DNA, then exo/endonucleaseases remove the lesion and leave a 1-30 nucleotide gap that is sub-sequentially filled in by DNA polymerase and finally sealed with ligase. A more complete picture is given in Li, Cell Research (2008) 18:85–98, and a summary is provided here.

Mismatch repair (MMR) operates on mispaired DNA bases.

The MSH2/6 or MSH2/3 complexes both have ATPases activity that plays an important role in mismatch recognition and the initiation of repair. MSH2/6 preferentially recognizes base-base mismatches and identifies mispairs of 1 or 2 nucleotides, while MSH2/3 preferentially recognizes larger ID mispairs.

hMLH1 heterodimerizes with hPMS2 to form hMutL α which possesses an ATPase activity and is important for multiple steps of MMR. It possesses a PCNA/replication factor C (RFC)-dependent endonuclease activity which plays an important role in 3′ nick-directed MMR involving EXO1. (EXO1 is a participant in both HR and MMR.) It regulates termination of mismatch-provoked excision. Ligase I is the relevant ligase for this pathway. Additional factors that may promote MMR include: EXO1, MSH2, MSH3, MSH6, MLH1, PMS2, MLH3, DNA Pol d, RPA, HMGB1, RFC, and DNA ligase I.

Base excision repair (BER)

The base excision repair (BER) pathway is active throughout the cell cycle; it is responsible primarily for removing small, non-helix-distorting base lesions from the genome. In contrast, the related Nucleotide Excision Repair pathway (discussed in the next section) repairs bulky helix-distorting lesions. A more detailed explanation is given in Caldecott, Nature Reviews Genetics 9, 619-631 (August 2008), and a summary is given here.

Upon DNA base damage, base excision repair (BER) is initiated and the process can be simplified into five major steps: (a) removal of the damaged DNA base; (b) incision of the subsequent a basic site; (c) clean-up of the DNA ends; (d) insertion of the correct nucleotide into

the repair gap; and (e) ligation of the remaining nick in the DNA backbone. These last steps are similar to the SSBR.

In the first step, a damage-specific DNA glycosylase excises the damaged base through cleavage of the N-glycosidic bond linking the base to the sugar phosphate backbone. Then AP endonuclease-1 (APE1) or bifunctional DNA glycosylases with an associated lyase activity incised the phosphodiester backbone to create a DNA single strand break (SSB). The third step of BER involves cleaning-up of the DNA ends. The fourth step in BER is conducted by Pol β that adds a new complementary nucleotide into the repair gap and in the final step XRCC1/Ligase III seals the remaining nick in the DNA backbone. This completes the short-patch BER pathway in which the majority (~80%) of damaged DNA bases are repaired. However, if the 5′ -ends in step 3 are resistant to end processing activity, following one nucleotide insertion by Pol β there is then a polymerase switch to the replicative DNA polymerases, Pol δ/ε, which then add ~2–8 more nucleotides into the DNA repair gap. This creates a 5′ -flap structure, which is recognized and excised by flap endonuclease-1 (FEN-1) in association with the processivity factor proliferating cell nuclear antigen (PCNA). DNA ligase I then seals the remaining nick in the DNA backbone and completes long-patch BER. Additional factors that may promote the BER pathway include: DNA glycosylase, APE1, Polb, Pold, Pole, XRCC1, Ligase III, FEN-1, PCNA, RECQL4, WRN, MYH, PNKP, and APTX.

Nucleotide excision repair (NER)

Nucleotide excision repair (NER) is an important excision mechanism that removes bulky helix-distorting lesions from DNA. Additional details about NER are given in Marteijn et al., Nature Reviews Molecular Cell Biology 15, 465–481 (2014), and a summary is given here. NER a broad pathway encompassing two smaller pathways: global genomic NER (GG-NER) and transcription coupled repair NER (TC-NER). GG-NER and TC-NER use different factors for recognizing DNA damage. However, they utilize the same machinery for lesion incision, repair, and ligation.

Once damage is recognized, the cell removes a short single-stranded DNA segment that contains the lesion. Endonucleases XPF/ERCC1 and XPG (encoded by ERCC5) remove the lesion by cutting the damaged strand on either side of the lesion, resulting in a single-strand gap of 22–30 nucleotides. Next, the cell performs DNA gap filling synthesis and ligation. Involved

in this process are: PCNA, RFC, DNA Pol δ, DNA Pol ε or DNA Pol κ, and DNA ligase I or XRCC1/Ligase III. Replicating cells tend to use DNA pol ε and DNA ligase I, while non-replicating cells tend to use DNA Pol δ, DNA Pol κ, and the XRCC1/ Ligase III complex to perform the ligation step.

NER can involve the following factors: XPA-G, POLH, XPF, ERCC1, XPA-G, and LIG1. Transcription-coupled NER (TC-NER) can involve the following factors: CSA, CSB, XPB, XPD, XPG, ERCC1, and TTDA. Additional factors that may promote the NER repair pathway include XPA-G, POLH, XPF, ERCC1, XPA-G, LIG1, CSA, CSB, XPA, XPB, XPC, XPD, XPF, XPG, TTDA, UVSSA, USP7, CETN2, RAD23B, UV-DDB, CAK subcomplex, RPA, and PCNA.

Interstrand Crosslink (ICL)

A dedicated pathway called the ICL repair pathway repairs interstrand crosslinks. Interstrand crosslinks, or covalent crosslinks between bases in different DNA strand, can occur during replication or transcription. ICL repair involves the coordination of multiple repair processes, in particular, nucleolytic activity, translesion synthesis (TLS), and HDR. Nucleases are recruited to excise the ICL on either side of the crosslinked bases, while TLS and HDR are coordinated to repair the cut strands. ICL repair can involve the following factors: endonucleases, e.g., XPF and RAD51C, endonucleases such as RAD51, translesion polymerases, e.g., DNA polymerase zeta and Rev1), and the Fanconi anemia (FA) proteins, e.g., FancJ.

Other pathways

Several other DNA repair pathways exist in mammals.

Translesion synthesis (TLS) is a pathway for repairing a single stranded break left after a defective replication event and involves translesion polymerases, e.g., DNA polζ and Rev1..

Error-free postreplication repair (PRR) is another pathway for repairing a single stranded break left after a defective replication event.

V.4 Examples of gRNAs in Genome Editing Methods

gRNA molecules as described herein can be used with Cas9 molecules that cleave both or a single strand to alter the sequence of a target nucleic acid, e.g., of a target position or target genetic signature. gRNA molecules useful in these method are described below.

In an embodiment, the gRNA, e.g., a chimeric gRNA, molecule is configured such that it comprises one or more of the following properties;

a) it can position, e.g., when targeting a Cas9 molecule that makes double strand breaks, a double strand break (i) within 50, 100, 150 or 200 nucleotides of a target position, or (ii) sufficiently close that the target position is within the region of end resection;

b) it has a targeting domain of at least 17 nucleotides, e.g., a targeting domain of (i) 17, (ii) 18, or (iii) 20 nucleotides; and

c)

(i) the proximal and tail domain, when taken together, comprise at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides, e.g., at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides from a naturally occurring *S. pyogenes*, *S. thermophilus*, *S. aureus*, or *N. meningitidis* tail and proximal domain, or a sequence that differs by no more than 1, 2, 3, 4, 5; 6, 7, 8, 9 or 10 nucleotides therefrom;

(ii) there are at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides 3' to the last nucleotide of the second complementarity domain, e.g., at least 15, 18, 20, 25, 30, 31, 35, 40, 45, 49, 50, or 53 nucleotides from the corresponding sequence of a naturally occurring *S. pyogenes*, *S. thermophilus*, *S. aureus*, or *N. meningitidis gRNA*, or a sequence that differs by no more than 1, 2, 3, 4, 5; 6, 7, 8, 9 or 10 nucleotides therefrom;

(iii) there are at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides 3' to the last nucleotide of the second complementarity domain that is complementary to its corresponding nucleotide of the first complementarity domain, e.g., at least 16, 19, 21, 26, 31, 32, 36, 41, 46, 50, 51, or 54 nucleotides from the corresponding sequence of a naturally occurring *S. pyogenes*, *S. thermophilus*, *S. aureus*, or *N. meningitidis gRNA*, or a sequence that differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 nucleotides therefrom;

iv) the tail domain is at least 10, 15, 20, 25, 30, 35 or 40 nucleotides in length, e.g., it comprises at least 10, 15, 20, 25, 30, 35 or 40 nucleotides from a naturally occurring *S. pyogenes*,

S. thermophilus, *S. aureus*, or *N. meningitidis* tail domain; or, or a sequence that differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 nucleotides therefrom; or

(v) the tail domain comprises 15, 20, 25, 30, 35, 40 nucleotides or all of the corresponding portions of a naturally occurring tail domain, e.g., a naturally occurring *S. pyogenes*, *S. thermophilus*, *S. aureus*, or *N. meningitidis* tail domain.

In an embodiment, the gRNA molecule is configured such that it comprises properties: a(i); and b(i).

In an embodiment, the gRNA molecule is configured such that it comprises properties: a(i); and b(ii).

In an embodiment, the gRNA molecule is configured such that it comprises properties: a(i); and b(iii).

In an embodiment, the gRNA molecule is configured such that it comprises properties: a(ii); and b(i).

In an embodiment, the gRNA molecule is configured such that it comprises properties: a(ii); and b(ii).

In an embodiment, the gRNA molecule is configured such that it comprises properties: a(ii); and b(iii).

In an embodiment, the gRNA molecule is configured such that it comprises properties: b(i); and c(i).

In an embodiment, the gRNA molecule is configured such that it comprises properties: b(i); and c(ii).

In an embodiment, the gRNA molecule is configured such that it comprises properties: b(ii); and c(i).

In an embodiment, the gRNA molecule is configured such that it comprises properties: b(ii); and c(ii).

In an embodiment, the gRNA molecule is configured such that it comprises properties: b(iii); and c(i).

In an embodiment, the gRNA molecule is configured such that it comprises properties: b(iii); and c(ii).

In an embodiment, the gRNA is used with a Cas9 nickase molecule having HNH activity, e.g., a Cas9 molecule having the RuvC activity inactivated, e.g., a Cas9 molecule having a mutation at D10, e.g., the D10A mutation.

In an embodiment, the gRNA is used with a Cas9 nickase molecule having RuvC activity, e.g., a Cas9 molecule having the HNH activity inactivated, e.g., a Cas9 molecule having a mutation at H840, e.g., a H840A.

In an embodiment, the gRNA is used with a Cas9 nickase molecule having RuvC activity, e.g., a Cas9 molecule having the HNH activity inactivated, e.g., a Cas9 molecule having a mutation at H863, e.g., a H863A.

In an embodiment, a pair of gRNA molecules, e.g., a pair of chimeric gRNA molecules, comprising a first and a second gRNA molecule, is configured such that they comprises one or more of the following properties:

a) the first and second gRNA molecules position, e.g., when targeting a Cas9 molecule that makes single strand or double strand breaks:

(i) as positioned by a first and second gRNA molecule described herein; or

(ii) sufficiently close that the target position is altered when the break is repaired;

b) one or both, independently, has a targeting domain of at least 17 nucleotides, e.g., a targeting domain of (i) 17, (ii) 18, or (iii) 20 nucleotides; and

c) one or both, independently, has a the tail domain is (i) at least 10, 15, 20, 25, 30, 35 or 40 nucleotides in length or (ii) the tail domain comprises, 15, 20, 25, 30, 35, 40, or all of the corresponding portions of a naturally occurring tail domain, e.g., a naturally occurring *S. pyogenes, S. aureus, or S. thermophilus* tail domain.

In an embodiment, one or both of the gRNA molecules is configured such that it comprises properties: a(i); and b(i).

In an embodiment, one or both of the gRNA molecules is configured such that it comprises properties: a(i); and b(ii).

In an embodiment, one or both of the gRNA molecules is configured such that it comprises properties: a(i); and b(iii).

In an embodiment, one or both of the gRNA molecules is configured such that it comprises properties: a(ii); and b(i).

In an embodiment, one or both of the gRNA molecules is configured such that it comprises properties: a(ii); and b(ii).

In an embodiment, one or both of the gRNA molecules is configured such that it comprises properties: a(ii); and b(iii).

In an embodiment, one or both of the gRNA molecules is configured such that it comprises properties: b(i); and c(i).

In an embodiment, one or both of the gRNA molecules is configured such that it comprises properties: b(i); and c(ii).

In an embodiment, one or both of the gRNA molecules is configured such that it comprises properties: b(ii); and c(i).

In an embodiment, one or both of the gRNA molecules is configured such that it comprises properties: b(ii); and c(ii).

In an embodiment, one or both of the gRNA molecules is configured such that it comprises properties: b(iii); and c(i).

In an embodiment, one or both of the gRNA molecules is configured such that it comprises properties: b(iii); and c(ii).

In an embodiment the gRNA is used with a Cas9 nickase molecule having HNH activity, e.g., a Cas9 molecule having the RuvC activity inactivated, e.g., a Cas9 molecule having a mutation at D10, e.g., the D10A mutation.

In an embodiment, the gRNA is used with a Cas9 nickase molecule having RuvC activity, e.g., a Cas9 molecule having the HNH activity inactivated, e.g., a Cas9 molecule having a mutation at H840, e.g., a H840A.

In an embodiment the gRNA is used with a Cas9 nickase molecule having RuvC activity, e.g., a Cas9 molecule having the HNH activity inactivated, e.g., a Cas9 molecule having a mutation at H863, e.g., a H863A.

VI. Targets: Cells

Cas9 molecules and gRNA molecules, e.g., a Cas9 molecule/gRNA molecule complex, can be used to manipulate a cell, e.g., to edit a target nucleic acid, in a wide variety of cells.

In some embodiments, a cell is manipulated by altering one or more target genes, e.g., as described herein. In some embodiments, the expression of one or more target genes (e.g., one or more target genes described herein) is modulated, e.g., *in vivo*.

In an embodiment, the target cell is a retinal cell, e.g., a cell of the retinal pigment epithelium cell or a photoreceptor cell. In another embodiment, the target cell is a horizontal cell, a bipolar cell, an amacrine cell, or a ganglion cell. In an embodiment, the target cell is a cone photoreceptor cell or cone cell, a rod photoreceptor cell or rod cell, or a macular cone photoreceptor cell. In an exemplary embodiment, cone photoreceptors in the macula are targeted, i.e., cone photoreceptors in the macula are the target cells.

In an embodiment, the target cell is removed from the subject, the gene altered *ex vivo*, and the cell returned to the subject. In an embodiment, a photoreceptor cell is removed from the subject, the gene altered *ex vivo*, and the photoreceptor cell returned to the subject. In an embodiment, a cone photoreceptor cell is removed from the subject, the gene altered *ex vivo*, and the cone photoreceptor cell returned to the subject.

In an embodiment, the cells are induced pluripotent stem cells (iPS) cells or cells derived from iPS cells, e.g., iPS cells from the subject, modified to alter the gene and differentiated into retinal progenitor cells or retinal cells, e.g., retinal photoreceptors, and injected into the eye of the subject, e.g., subretinally, e.g., in the submacular region of the retina.

In an embodiment, the cells are targeted *in vivo*, e.g., by delivery of the components, e.g., a Cas9 molecule and a gRNA molecule, to the target cells. In an embodiment, the target cells are retinal pigment epithelium, photoreceptor cells, or a combination thereof. In an embodiment, AAV is used to deliver the components, e.g., a Cas9 molecule and a gRNA molecule, e.g., by transducing the target cells.

VII. Delivery, Formulations and Routes of Administration

The components, e.g., a Cas9 molecule and gRNA molecule can be delivered, formulated, or administered in a variety of forms, see, e.g., **Table 17**. In an embodiment, one Cas9 molecule and two or more (e.g., 2, 3, 4, or more) different gRNA molecules are delivered, e.g., by an AAV vector. In an embodiment, the sequence encoding the Cas9 molecule and the sequence(s) encoding the two or more (e.g., 2, 3, 4, or more) different gRNA molecules are present on the same nucleic acid molecule, e.g., an AAV vector. When a Cas9 or gRNA

component is delivered encoded in DNA the DNA will typically include a control region, e.g., comprising a promoter, to effect expression. Useful promoters for Cas9 molecule sequences include CMV, EFS, EF-1a, MSCV, PGK, CAG, hGRK1, hCRX, hNRL, and hRCVRN control promoters. In an embodiment, the promoter is a constitutive promoter. In another embodiment, the promoter is a tissue specific promoter. Exemplary promoter sequences are disclosed in **Table 19**. Useful promoters for gRNAs include H1, 7SK, and U6 promoters. Promoters with similar or dissimilar strengths can be selected to tune the expression of components. Sequences encoding a Cas9 molecule can comprise a nuclear localization signal (NLS), e.g., an SV40 NLS. In an embodiment, the sequence encoding a Cas9 molecule comprises at least two nuclear localization signals. In an embodiment a promoter for a Cas9 molecule or a gRNA molecule can be, independently, inducible, tissue specific, or cell specific. To detect the expression of a Cas9, an affinity tag can be used. Useful affinity tag sequences include, but are not limted to, 3xFlag tag, single Flag tag, HA tage, Myc tag or HIS tage. Exemplary affinity tage sequences are disclosed in **Table 25**. To regulate Cas9 expression, e.g., in mammalian cells, polyadenylation signals (poly(A) signals) can be used. Exemplary polyadenylation signals are disclosed in **Table 26**.

Table 17 provides examples of how the components can be formulated, delivered, or administered.

Table 17

Elements		
Cas9 Molecule(s)	gRNA molecule(s)	Comments
DNA	DNA	In this embodiment a Cas9 molecule, typically an eaCas9 molecule, and a gRNA are transcribed from DNA. In this embodiment they are encoded on separate molecules.
DNA		In this embodiment a Cas9 molecule, typically an eaCas9 molecule, and a gRNA are transcribed from DNA, here from a single molecule.
DNA	RNA	In this embodiment a Cas9 molecule, typically

		an eaCas9 molecule, is transcribed from DNA.
mRNA	RNA	In this embodiment a Cas9 molecule, typically an eaCas9 molecule, is transcribed from DNA.
Protein	DNA	In this embodiment a Cas9 molecule, typically an eaCas9 molecule, is provided as a protein. A gRNA is transcribed from DNA.
Protein	RNA	In this embodiment an eaCas9 molecule is provided as a protein.

Table 18 summarizes various delivery methods for the components of a Cas system, e.g., the Cas9 molecule component and the gRNA molecule component, as described herein.

Table 18

Delivery Vector/Mode		Delivery into Non-Dividing Cells	Duration of Expression	Genome Integration	Type of Molecule Delivered
Physical (e.g., electroporation, particle gun, Calcium Phosphate transfection)		YES	Transient	NO	Nucleic Acids and Proteins
Viral	Retrovirus	NO	Stable	YES	RNA
	Lentivirus	YES	Stable	YES/NO with modifications	RNA
	Adenovirus	YES	Transient	NO	DNA
	Adeno-Associated Virus (AAV)	YES	Stable	NO	DNA
	Vaccinia Virus	YES	Very	NO	DNA

			Transient		
	Herpes Simplex Virus	YES	Stable	NO	DNA
Non-Viral	Cationic Liposomes	YES	Transient	Depends on what is delivered	Nucleic Acids and Proteins
	Polymeric Nanoparticles	YES	Transient	Depends on what is delivered	Nucleic Acids and Proteins
Biological Non-Viral Delivery Vehicles	Attenuated Bacteria	YES	Transient	NO	Nucleic Acids
	Engineered Bacteriophages	YES	Transient	NO	Nucleic Acids
	Mammalian Virus-like Particles	YES	Transient	NO	Nucleic Acids
	Biological liposomes: Erythrocyte Ghosts and Exosomes	YES	Transient	NO	Nucleic Acids

Table 19 describes exemplary promoter sequences that can be used in AAV vectors, e.g., for Cas9 expression.

Table 19

Promoter	Length (bp)	DNA Sequence
CMV	617	CATTGATTATTGACTAGTTATTAATAGTAATCAATTACGG GGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTAC ATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAA CGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCC ATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGG GTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATC AAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAA

		TGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACAT GACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTA TTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGT ACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGAT TTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTT TTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAA CAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGT ACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAA CCGTCAGATCCGCTAGAGATCCGC (SEQ ID NO: 401)
EFS	252	TCGAGTGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACA TCGCCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCGGC AATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAAC TGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCG AGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCG TGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACA GGTGTCGTGACCGCGG (SEQ ID NO: 402)
Human GRK1 (rhodopsin kinase)	292	GGGCCCCAGAAGCCTGGTGGTTGTTTGTCCTTCTCAGGGG AAAAGTGAGGCGGCCCCTTGGAGGAAGGGGCCGGGCAG AATGATCTAATCGGATTCCAAGCAGCTCAGGGGATTGTCT TTTTCTAGCACCTTCTTGCCACTCCTAAGCGTCCTCCGTGA CCCCGGCTGGGATTTCGCCTGGTGCTGTGTCAGCCCCGGT CTCCCAGGGGCTTCCCAGTGGTCCCCAGGAACCCTCGAC AGGGCCCGGTCTCTCTCGTCCAGCAAGGGCAGGGACGGG CCACAGGCCAAGGGC (SEQ ID NO: 403)
Human CRX (cone rod homeobox transcriptio n factor)	113	GCCTGTAGCCTTAATCTCTCCTAGCAGGGGGTTTGGGGGA GGGAGGAGGAGAAAGAAAGGGCCCCTTATGGCTGAGAC ACAATGACCCAGCCACAAGGAGGGATTACCGGGCG (SEQ ID NO: 404)
Human NRL (neural retina leucine zipper transcriptio n factor enhance upstream of the human TK terminal promoter)	281	AGGTAGGAAGTGGCCTTTAACTCCATAGACCCTATTTAAA CAGCTTCGGACAGGTTTAAACATCTCCTTGGATAATTCCT AGTATCCCTGTTCCCACTCCTACTCAGGGATGATAGCTCT AAGAGGTGTTAGGGGATTAGGCTGAAAATGTAGGTCACC CCTCAGCCATCTGGGAACTAGAATGAGTGAGAGAGGAGA GAGGGGCAGAGACACACACATTCGCATATTAAGGTGACG CGTGTGGCCTCGAACACCGAGCGACCCTGCAGCGACCCG CTTAA (SEQ ID NO: 405)

Human RCVRN (recoverin)	235	ATTTTAATCTCACTAGGGTTCTGGGAGCACCCCCCCCCAC CGCTCCCGCCCTCCACAAAGCTCCTGGGCCCCTCCTCCCT TCAAGGATTGCGAAGAGCTGGTCGCAAATCCTCCTAAGC CACCAGCATCTCGGTCTTCAGCTCACACCAGCCTTGAGCC CAGCCTGCGGCCAGGGGACCACGCACGTCCCACCCACCC AGCGACTCCCCAGCCGCTGCCCACTCTTCCTCACTCA (SEQ ID NO: 406)

Table 25 describes exemplary affinity tag sequences that can be used in AAV vectors, e.g., for Cas9 expression.

Affinity tag	Amino Acid Sequence
3XFlag tag	DYKDHDGDYKDHDIDYKDDDDK (SEQ ID NO: 435)
Flag tag (single)	DYKDDDDK (SEQ ID NO: 451)
HA tag	YPYDVPDYA (SEQ ID NO: 452)
Myc tag	EQKLISEEDL (SEQ ID NO: 453)
HIS tag	HHHHHH (SEQ ID NO: 454)

Table 26 describes exemplary polyA sequences that can be used in AAV vectors, e.g., for Cas9 expression.

PolyA	DNA sequence
mini polyA	TAGCAATAAAGGATCGTTTATTTTCATTGGAAGCGTGTGT TGGTTTTTTGATCAGGCGCG (SEQ ID NO: 424)
bGH polyA	GCTGCAGGATGACCGGTCATCATCACCATCACCATTGAG TTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTG CCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGA CCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAA TGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCT ATTCTGGGGGGTGGGGTGGGGCAGGACA (SEQ ID NO: 455)
SV40 polyA	ATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTG TAACCATTATAAGCTGCAATAAACAAGTTAACAACAACA ATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTG GGAGGTTTTTTAAA (SEQ ID NO: 456)

Table 24 describes exemplary Inverted Terminal Repeat (ITR) sequences that can be used in AAV vectors.

Table 24. Sequences of ITRs from exemplary AAV Serotypes

AAV Serotype	Left ITR Sequence	Right ITR Sequence
AAV1	TTGCCCACTCCCTCTCTGCGCGCTCGCT CGCTCGGTGGGGCCTGCGGACCAAAGG TCCGCAGACGGCAGAGCTCTGCTCTGC CGGCCCCACCGAGCGAGCGAGCGCGC AGAGAGGGAGTGGGCAACTCCATCACT AGGGGTAA (SEQ ID NO: 407)	TTACCCCTAGTGATGGAGTTGCCCACT CCCTCTCTGCGCGCTCGCTCGCTCGGTG GGGCCGGCAGAGCAGAGCTCTGCCGTC TGCGGACCTTTGGTCCGCAGGCCCCAC CGAGCGAGCGAGCGCGCAGAGAGGGA GTGGGCAA (SEQ ID NO: 436)
AAV2	TTGGCCACTCCCTCTCTGCGCGCTCGCT CGCTCACTGAGGCCGGGCGACCAAAG GTCGCCCGACGCCCGGGCTTTGCCCGG GCGGCCTCAGTGAGCGAGCGAGCGCG CAGAGAGGGAGTGGCCAACTCCATCAC TAGGGGTTCCT (SEQ ID NO: 408)	AGGAACCCCTAGTGATGGAGTTGGCCA CTCCCTCTCTGCGCGCTCGCTCGCTCAC TGAGGCCGCCCGGGCAAAGCCCGGGC GTCGGGCGACCTTTGGTCGCCCGGCCT CAGTGAGCGAGCGAGCGCGCAGAGAG GGAGTGGCCAA (SEQ ID NO: 437)
AAV3B	TGGCCACTCCCTCTATGCGCACTCGCTC GCTCGGTGGGGCCTGGCGACCAAAGGT CGCCAGACGGACGTGCTTTGCACGTCC GGCCCCACCGAGCGAGCGAGTGCGCAT AGAGGGAGTGGCCAACTCCATCACTAG AGGTAT (SEQ ID NO: 409)	ATACCTCTAGTGATGGAGTTGGCCACT CCCTCTATGCGCACTCGCTCGCTCGGT GGGGCCGGACGTGCAAAGCACGTCCGT CTGGCGACCTTTGGTCGCCAGGCCCCA CCGAGCGAGCGAGTGCGCATAGAGGG AGTGGCCA (SEQ ID NO: 438)
AAV4	TTGGCCACTCCCTCTATGCGCGCTCGCT CACTCACTCGGCCCTGGAGACCAAAGG TCTCCAGACTGCCGGCCTCTGGCCGGC AGGGCCGAGTGAGTGAGCGAGCGCGC ATAGAGGGAGTGGCCAACTCCATCATC TAGGTTTGCCC (SEQ ID NO: 410)	GGGCAAACCTAGATGATGGAGTTGGCC ACTCCCTCTATGCGCGCTCGCTCACTCA CTCGGCCCTGCCGGCCAGAGGCCGGCA GTCTGGAGACCTTTGGTCTCCAGGGCC GAGTGAGTGAGCGAGCGCGCATAGAG GGAGTGGCCAA (SEQ ID NO: 439)
AAV5	CTCTCCCCCCTGTCGCGTTCGCTCGCTC GCTGGCTCGTTTGGGGGGGTGGCAGCT CAAAGAGCTGCCAGACGACGGCCCTCT GGCCGTCGCCCCCCCAAACGAGCCAGC GAGCGAGCGAACGCGACAGGGGGGAG AGTGCCACACTCTCAAGCAA (SEQ ID	TTGCTTGAGAGTGTGGCACTCTCCCCC CTGTCGCGTTCGCTCGCTCGCTGGCTCG TTTGGGGGGGCGACGGCCAGAGGGCC GTCGTCTGGCAGCTCTTTGAGCTGCCA CCCCCCCAAACGAGCCAGCGAGCGAG CGAACGCGACAGGGGGGAGAG (SEQ ID

	NO: 411)	NO: 440)
AAV6	ATACCCCTAGTGATGGAGTTGCCCACT CCCTCTATGCGCGCTCGCTCGCTCGGT GGGGCCGGCAGAGCAGAGCTCTGCCGT CTGCGGACCTTTGGTCCGCAGGCCCCA CCGAGCGAGCGAGCGCGCATAGAGGG AGTGGGCAA (SEQ ID NO: 412)	TTGCCCACTCCCTCTATGCGCGCTCGCT CGCTCGGTGGGGCCTGCGGACCAAAGG TCCGCAGACGGCAGAGCTCTGCTCTGC CGGCCCCACCGAGCGAGCGAGCGCGC ATAGAGGGAGTGGGCAACTCCATCACT AGGGGTAT (SEQ ID NO: 441)
AAV7	TTGGCCACTCCCTCTATGCGCGCTCGCT CGCTCGGTGGGGCCTGCGGACCAAAGG TCCGCAGACGGCAGAGCTCTGCTCTGC CGGCCCCACCGAGCGAGCGAGCGCGC ATAGAGGGAGTGGCCAACTCCATCACT AGGGGTACCG (SEQ ID NO: 413)	CGGTACCCCTAGTGATGGAGTTGGCCA CTCCCTCTATGCGCGCTCGCTCGCTCGG TGGGGCCGGCAGAGCAGAGCTCTGCCG TCTGCGGACCTTTGGTCCGCAGGCCCC ACCGAGCGAGCGAGCGCGCATAGAGG GAGTGGCCAA (SEQ ID NO: 442)
AAV8	CAGAGAGGGAGTGGCCAACTCCATCAC TAGGGGTAGCGCGAAGCGCCTCCCACG CTGCCGCGTCAGCGCTGACGTAAATTA CGTCATAGGGGAGTGGTCCTGTATTAG CTGTCACGTGAGTGCTTTTGCGGCATTT TGCGACACC (SEQ ID NO: 414)	GGTGTCGCAAAATGCCGCAAAAGCACT CACGTGACAGCTAATACAGGACCACTC CCCTATGACGTAATTTACGTCAGCGCT GACGCGGCAGCGTGGGAGGCGCTTCGC GCTACCCCTAGTGATGGAGTTGGCCAC TCCCTCTCTG (SEQ ID NO: 443)
AAV9	CAGAGAGGGAGTGGCCAACTCCATCAC TAGGGGTAATCGCGAAGCGCCTCCCAC GCTGCCGCGTCAGCGCTGACGTAGATT ACGTCATAGGGGAGTGGTCCTGTATTA GCTGTCACGTGAGTGCTTTTGCGACAT TTTGCGACAC (SEQ ID NO: 415)	GTGTCGCAAAATGTCGCAAAAGCACTC ACGTGACAGCTAATACAGGACCACTCC CCTATGACGTAATCTACGTCAGCGCTG ACGCGGCAGCGTGGGAGGCGCTTCGCG ATTACCCCTAGTGATGGAGTTGGCCAC TCCCTCTCTG (SEQ ID NO: 444)

Additional exemplary sequences for the recombinant AAV genome components described herein are provided below.

Exemplary Left and right ITR sequences are provided in **Table 24** (SEQ ID NOS: 407-415 and 436-444).

Exemplary spacer 1 sequence: CAGATCTGAATTCGGTACC (SEQ ID NO: 416).

Exemplary U6 promoter sequence:

AAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGA TACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTA GTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAA

TTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCT
TGGCTTTATATATCTTGTGGAAAGGACGAAACACC (SEQ ID NO: 417)

Exemplary gRNA targeting domain sequences are described herein, e.g., in **Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B,** or **Table 10**.

Exemplary gRNA scaffold domain sequence:
GTTTTAGTACTCTGGAAACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCT
CGTCAACTTGTTGGCGAGATTTTTT (SEQ ID NO: 418).

Exemplary spacer 2 domain sequence:
GGTACCGCTAGCGCTTAAGTCGCGATGTACGGGCCAGATATACGCGTTGA (SEQ ID NO: 419).

Exemplary Polymerase II promoter sequences are provided in **Table 19**.

Exemplary N-ter NLS nucleotide sequence:
CCGAAGAAAAAGCGCAAGGTCGAAGCGTCC (SEQ ID NO: 420)

Exemplary N-ter NLS amino acid sequence: PKKKRKV (SEQ ID NO: 434)

Exemplary Cas9 nucleotide sequence:
ATGAAAAGGAACTACATTCTGGGGCTGGACATCGGGATTACAAGCGTGGGGTATGG
GATTATTGACTATGAAACAAGGGACGTGATCGACGCAGGCGTCAGACTGTTCAAGG
AGGCCAACGTGGAAAACAATGAGGGACGGAGAAGCAAGAGGGGAGCCAGGCGCCT
GAAACGACGGAGAAGGCACAGAATCCAGAGGGTGAAGAAACTGCTGTTCGATTACA
ACCTGCTGACCGACCATTCTGAGCTGAGTGGAATTAATCCTTATGAAGCCAGGGTGA
AAGGCCTGAGTCAGAAGCTGTCAGAGGAAGAGTTTTCCGCAGCTCTGCTGCACCTG
GCTAAGCGCCGAGGAGTGCATAACGTCAATGAGGTGGAAGAGGACACCGGCAACG
AGCTGTCTACAAAGGAACAGATCTCACGCAATAGCAAAGCTCTGGAAGAGAAGTAT
GTCGCAGAGCTGCAGCTGGAACGGCTGAAGAAAGATGGCGAGGTGAGAGGGTCAA
TTAATAGGTTCAAGACAAGCGACTACGTCAAAGAAGCCAAGCAGCTGCTGAAAGTG
CAGAAGGCTTACCACCAGCTGGATCAGAGCTTCATCGATACTTATATCGACCTGCTG
GAGACTCGGAGAACCTACTATGAGGGACCAGGAGAAGGGAGCCCCTTCGGATGGAA
AGACATCAAGGAATGGTACGAGATGCTGATGGGACATTGCACCTATTTTCCAGAAG
AGCTGAGAAGCGTCAAGTACGCTTATAACGCAGATCTGTACAACGCCCTGAATGAC
CTGAACAACCTGGTCATCACCAGGGATGAAAACGAGAAACTGGAATACTATGAGAA

GTTCCAGATCATCGAAAACGTGTTTAAGCAGAAGAAAAAGCCTACACTGAAACAGA
TTGCTAAGGAGATCCTGGTCAACGAAGAGGACATCAAGGGCTACCGGGTGACAAGC
ACTGGAAAACCAGAGTTCACCAATCTGAAAGTGTATCACGATATTAAGGACATCAC
AGCACGGAAAGAAATCATTGAGAACGCCGAACTGCTGGATCAGATTGCTAAGATCC
TGACTATCTACCAGAGCTCCGAGGACATCCAGGAAGAGCTGACTAACCTGAACAGC
GAGCTGACCCAGGAAGAGATCGAACAGATTAGTAATCTGAAGGGGTACACCGGAAC
ACACAACCTGTCCCTGAAAGCTATCAATCTGATTCTGGATGAGCTGTGGCATACAAA
CGACAATCAGATTGCAATCTTTAACCGGCTGAAGCTGGTCCCAAAAAAGGTGGACC
TGAGTCAGCAGAAAGAGATCCCAACCACACTGGTGGACGATTTCATTCTGTCACCCG
TGGTCAAGCGGAGCTTCATCCAGAGCATCAAAGTGATCAACGCCATCATCAAGAAG
TACGGCCTGCCCAATGATATCATTATCGAGCTGGCTAGGGAGAAGAACAGCAAGGA
CGCACAGAAGATGATCAATGAGATGCAGAAACGAAACCGGCAGACCAATGAACGC
ATTGAAGAGATTATCCGAACTACCGGGAAAGAGAACGCAAAGTACCTGATTGAAAA
AATCAAGCTGCACGATATGCAGGAGGGAAAGTGTCTGTATTCTCTGGAGGCCATCCC
CCTGGAGGACCTGCTGAACAATCCATTCAACTACGAGGTCGATCATATTATCCCCAG
AAGCGTGTCCTTCGACAATTCCTTTAACAACAAGGTGCTGGTCAAGCAGGAAGAGA
ACTCTAAAAAGGGCAATAGGACTCCTTTCCAGTACCTGTCTAGTTCAGATTCCAAGA
TCTCTTACGAAACCTTTAAAAAGCACATTCTGAATCTGGCCAAAGGAAAGGGCCGC
ATCAGCAAGACCAAAAAGGAGTACCTGCTGGAAGAGCGGGACATCAACAGATTCTC
CGTCCAGAAGGATTTTATTAACCGGAATCTGGTGGACACAAGATACGCTACTCGCGG
CCTGATGAATCTGCTGCGATCCTATTTCCGGGTGAACAATCTGGATGTGAAAGTCAA
GTCCATCAACGGCGGGTTCACATCTTTTCTGAGGCGCAAATGGAAGTTTAAAAAGGA
GCGCAACAAAGGGTACAAGCACCATGCCGAAGATGCTCTGATTATCGCAAATGCCG
ACTTCATCTTTAAGGAGTGGAAAAAGCTGGACAAAGCCAAGAAAGTGATGGAGAAC
CAGATGTTCGAAGAGAAGCAGGCCGAATCTATGCCCGAAATCGAGACAGAACAGGA
GTACAAGGAGATTTTCATCACTCCTCACCAGATCAAGCATATCAAGGATTTCAAGGA
CTACAAGTACTCTCACCGGGTGGATAAAAAGCCCAACAGAGAGCTGATCAATGACA
CCCTGTATAGTACAAGAAAAGACGATAAGGGGAATACCCTGATTGTGAACAATCTG
AACGGACTGTACGACAAAGATAATGACAAGCTGAAAAAGCTGATCAACAAAAGTCC
CGAGAAGCTGCTGATGTACCACCATGATCCTCAGACATATCAGAAACTGAAGCTGA
TTATGGAGCAGTACGGCGACGAGAAGAACCCACTGTATAAGTACTATGAAGAGACT

GGGAACTACCTGACCAAGTATAGCAAAAAGGATAATGGCCCCGTGATCAAGAAGAT
CAAGTACTATGGGAACAAGCTGAATGCCCATCTGGACATCACAGACGATTACCCTA
ACAGTCGCAACAAGGTGGTCAAGCTGTCACTGAAGCCATACAGATTCGATGTCTATC
TGGACAACGGCGTGTATAAATTTGTGACTGTCAAGAATCTGGATGTCATCAAAAAG
GAGAACTACTATGAAGTGAATAGCAAGTGCTACGAAGAGGCTAAAAAGCTGAAAA
AGATTAGCAACCAGGCAGAGTTCATCGCCTCCTTTTACAACAACGACCTGATTAAGA
TCAATGGCGAACTGTATAGGGTCATCGGGGTGAACAATGATCTGCTGAACCGCATTG
AAGTGAATATGATTGACATCACTTACCGAGAGTATCTGGAAAACATGAATGATAAG
CGCCCCCCTCGAATTATCAAAACAATTGCCTCTAAGACTCAGAGTATCAAAAAGTAC
TCAACCGACATTCTGGGAAACCTGTATGAGGTGAAGAGCAAAAAGCACCCTCAGAT
TATCAAAAAGGGC (SEQ ID NO: 421)

Exemplary Cas9 amino acid sequence:

MKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRR
RHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVH
NVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVK
EAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGH
CTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPT
LKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIY
QSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNR
LKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKN
SKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPL
EDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETF
KKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYF
RVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKEWKKLD
KAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNR
ELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQK
LKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDY
PNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLK
KISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPR
IIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG (SEQ ID NO: 435)

Exemplary C-ter NLS sequence: CCCAAGAAGAAGAGGAAAGTC (SEQ ID NO: 422).

Exemplary C-ter NLS amino acid sequence: PKKKRKV (SEQ ID NO: 434)

Exemplary poly(A) signal sequence:

TAGCAATAAAGGATCGTTTATTTTCATTGGAAGCGTGTGTTGGTTTTTTGATCAGGCG CG (SEQ ID NO: 424).

Exemplary Spacer 3 sequence:

TCCAAGCTTCGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCGTTAACTC TAGATTTAAATGCATGCTGGGGAGAGATCT (SEQ ID NO: 425)

Exemplary 3xFLAG nucleotide sequence:

GACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGA CGATGACAAG (SEQ ID NO: 423).

Exemplary 3xFLAG amino acid sequence:

DYKDHDGDYKDHDIDYKDDDDK (SEQ ID NO: 435)

Exemplary Spacer 4 sequence: CGACTTAGTTCGATCGAAGG (SEQ ID NO: 427).

Exemplary recombinant AAV genome sequences are provided in **Figs. 19A-24F** (SEQ IDNOS: 428-433 and 445-450). Exemplary sequences of the recombinant AAV genome components (e.g., one or more of the components described above) are also shown in **Figs. 19A-24F** (SEQ IDNOS: 428-433 and 445-450).

DNA-based Delivery of a Cas9 molecule and or a gRNA molecule

Nucleic acids encoding Cas9 molecules (e.g., eaCas9 molecules) and/or gRNA molecules, can be administered to subjects or delivered into cells by art-known methods or as described herein. For example, Cas9-encoding and/or gRNA-encoding DNA can be delivered, e.g., by vectors (e.g., viral or non-viral vectors), non-vector based methods (e.g., using naked DNA or DNA complexes), or a combination thereof.

DNA encoding Cas9 molecules (e.g., eaCas9 molecules) and/or gRNA molecules can be conjugated to molecules (e.g., N-acetylgalactosamine) promoting uptake by the target cells (e.g., the target cells described herein). Donor template molecules can be conjugated to molecules (e.g., N-acetylgalactosamine) promoting uptake by the target cells (e.g., the target cells described herein).

In some embodiments, the Cas9- and/or gRNA-encoding DNA is delivered by a vector (e.g., viral vector/virus or plasmid).

A vector can comprise a sequence that encodes a Cas9 molecule and/or a gRNA molecule. A vector can also comprise a sequence encoding a signal peptide (*e.g.,* for nuclear localization, nucleolar localization, mitochondrial localization), fused, e.g., to a Cas9 molecule sequence. For example, a vector can comprise a nuclear localization sequence (e.g., from SV40) fused to the sequence encoding the Cas9 molecule.

One or more regulatory/control elements, e.g., a promoter, an enhancer, an intron, a polyadenylation signal, a Kozak consensus sequence, internal ribosome entry sites (IRES), a 2A sequence, and splice acceptor or donor can be included in the vectors. In some embodiments, the promoter is recognized by RNA polymerase II (e.g., a CMV promoter). In other embodiments, the promoter is recognized by RNA polymerase III (e.g., a U6 promoter). In some embodiments, the promoter is a regulated promoter (e.g., inducible promoter). In other embodiments, the promoter is a constitutive promoter. In some embodiments, the promoter is a tissue specific promoter. In some embodiments, the promoter is a viral promoter. In other embodiments, the promoter is a non-viral promoter.

In some embodiments, the vector or delivery vehicle is a viral vector (e.g., for generation of recombinant viruses). In some embodiments, the virus is a DNA virus (e.g., dsDNA or ssDNA virus). In other embodiments, the virus is an RNA virus (e.g., an ssRNA virus). Exemplary viral vectors/viruses include, e.g., retroviruses, lentiviruses, adenovirus, adeno-associated virus (AAV), vaccinia viruses, poxviruses, and herpes simplex viruses.

In some embodiments, the virus infects dividing cells. In other embodiments, the virus infects non-dividing cells. In some embodiments, the virus infects both dividing and non-dividing cells. In some embodiments, the virus can integrate into the host genome. In some embodiments, the virus is engineered to have reduced immunity, e.g., in human. In some embodiments, the virus is replication-competent. In other embodiments, the virus is replication-defective, e.g., having one or more coding regions for the genes necessary for additional rounds of virion replication and/or packaging replaced with other genes or deleted. In some embodiments, the virus causes transient expression of the Cas9 molecule and/or the gRNA molecule. In other embodiments, the virus causes long-lasting, e.g., at least 1 week, 2 weeks, 1 month, 2 months, 3 months, 6 months, 9 months, 1 year, 2 years, or permanent expression, of the

Cas9 molecule and/or the gRNA molecule. The packaging capacity of the viruses may vary, e.g., from at least about 4 kb to at least about 30 kb, e.g., at least about 5 kb, 10 kb, 15 kb, 20 kb, 25 kb, 30 kb, 35 kb, 40 kb, 45 kb, or 50 kb.

In an embodiment, the viral vector recognizes a specific cell type or tissue. For example, the viral vector can be pseudotyped with a different/alternative viral envelope glycoprotein; engineered with a cell type-specific receptor (e.g., genetic modification(s) of one or more viral envelope glycoproteins to incorporate a targeting ligand such as a peptide ligand, a single chain antibody, or a growth factor); and/or engineered to have a molecular bridge with dual specificities with one end recognizing a viral glycoprotein and the other end recognizing a moiety of the target cell surface (e.g., a ligand-receptor, monoclonal antibody, avidin-biotin and chemical conjugation).

In some embodiments, the Cas9- and/or gRNA-encoding DNA is delivered by a recombinant retrovirus. In some embodiments, the retrovirus (e.g., Moloney murine leukemia virus) comprises a reverse transcriptase, e.g., that allows integration into the host genome. In some embodiments, the retrovirus is replication-competent. In other embodiments, the retrovirus is replication-defective, e.g., having one of more coding regions for the genes necessary for additional rounds of virion replication and packaging replaced with other genes, or deleted. In some embodiments, the Cas9- and/or gRNA-encoding DNA is delivered by a recombinant lentivirus. For example, the lentivirus is replication-defective, e.g., does not comprise one or more genes required for viral replication.

In some embodiments, the Cas9- and/or gRNA-encoding DNA is delivered by a recombinant adenovirus. In some embodiments, the adenovirus is engineered to have reduced immunity in human.

In some embodiments, the Cas9- and/or gRNA-encoding DNA is delivered by a recombinant AAV. In some embodiments, the AAV does not incorporate its geneome into that of a host cell, e.g., a target cell as describe herein. In some embodiments, the AAV can incorporate at least part of its genome into that of a host cell, e.g., a target cell as described herein. In some embodiments, the AAV is a self-complementary adeno-associated virus (scAAV), e.g., a scAAV that packages both strands which anneal together to form double stranded DNA. AAV serotypes that may be used in the disclosed methods, include AAV1, AAV2, modified AAV2 (e.g., modifications at Y444F, Y500F, Y730F and/or S662V), AAV3,

modified AAV3 (e.g., modifications at Y705F, Y731F and/or T492V), AAV4, AAV5, AAV6, modified AAV6 (e.g., modifications at S663V and/or T492V), AAV8, AAV 8.2, AAV9, AAV rhl0, and pseudotyped AAV, such as AAV2/8, AAV2/5 and AAV2/6 can also be used in the disclosed methods. In an embodiment, an AAV capsid that can be used in the methods described herein is a capsid sequence from serotype AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV.rh8, AAV.rh10, AAV.rh32/33, AAV.rh43, AAV.rh64R1, or AAV7m8. Exemplary AAV serotypes and ITR sequences are disclosed in **Table 24**.

In an embodiment, the Cas9- and/or gRNA-encoding DNA is delivered in a re-engineered AAV capsid, e.g., with 50% or greater, e.g., 60% or greater, 70% or greater, 80% or greater, 90% or greater, or 95% or greater, sequence homology with a capsid sequence from serotypes AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV.rh8, AAV.rh10, AAV.rh32/33, AAV.rh43, or AAV.rh64R1.

In an embodiment, the Cas9- and/or gRNA-encoding DNA is delivered by a chimeric AAV capsid. Exemplary chimeric AAV capsids include, but are not limited to, AAV9i1, AAV2i8, AAV-DJ, AAV2G9, AAV2i8G9, or AAV8G9.

In an embodiment, the AAV is a self-complementary adeno-associated virus (scAAV), e.g., a scAAV that packages both strands which anneal together to form double stranded DNA.

In some embodiments, the Cas9- and/or gRNA-encoding DNA is delivered by a hybrid virus, e.g., a hybrid of one or more of the viruses described herein. In an embodiment, the hybrid virus is hybrid of an AAV (e.g., of any AAV serotype), with a Bocavirus, B19 virus, porcine AAV, goose AAV, feline AAV, canine AAV, or MVM.

A packaging cell is used to form a virus particle that is capable of infecting a target cell. Such a cell includes a 293 cell, which can package adenovirus, and a ψ2 cell or a PA317 cell, which can package retrovirus. A viral vector used in gene therapy is usually generated by a producer cell line that packages a nucleic acid vector into a viral particle. The vector typically contains the minimal viral sequences required for packaging and subsequent integration into a host or target cell (if applicable), with other viral sequences being replaced by an expression cassette encoding the protein to be expressed, e.g., Cas9. For example, an AAV vector used in gene therapy typically only possesses inverted terminal repeat (ITR) sequences from the AAV genome which are required for packaging and gene expression in the host or target cell. The missing viral functions can be supplied in *trans* by the packaging cell line and/or plasmid

containing E2A, E4, and VA genes from adenovirus, and plasmid encoding *Rep* and *Cap* genes from AAV, as described in "Triple Transfection Protocol." Henceforth, the viral DNA is packaged in a cell line, which contains a helper plasmid encoding the other AAV genes, namely rep and cap, but lacking ITR sequences. In embodiment, the viral DNA is packaged in a producer cell line, which contains E1A and/or E1B genes from adenovirus. The cell line is also infected with adenovirus as a helper. The helper virus (e.g., adenovirus or HSV) or helper plasmid promotes replication of the AAV vector and expression of AAV genes from the plasmid with ITRs. The helper plasmid is not packaged in significant amounts due to a lack of ITR sequences. Contamination with adenovirus can be reduced by, e.g., heat treatment to which adenovirus is more sensitive than AAV.

In an embodiment, the viral vector has the ability of cell type and/or tissue type recognition. For example, the viral vector can be pseudotyped with a different/alternative viral envelope glycoprotein; engineered with a cell type-specific receptor (e.g., genetic modification of the viral envelope glycoproteins to incorporate targeting ligands such as a peptide ligand, a single chain antibody, a growth factor); and/or engineered to have a molecular bridge with dual specificities with one end recognizing a viral glycoprotein and the other end recognizing a moiety of the target cell surface (e.g., ligand-receptor, monoclonal antibody, avidin-biotin and chemical conjugation).

In an embodiment, the viral vector achieves cell type specific expression. For example, a tissue-specific promoter can be constructed to restrict expression of the transgene (Cas 9 and gRNA) in only the target cell. The specificity of the vector can also be mediated by microRNA-dependent control of transgene expression. In an embodiment, the viral vector has increased efficiency of fusion of the viral vector and a target cell membrane. For example, a fusion protein such as fusion-competent hemagglutin (HA) can be incorporated to increase viral uptake into cells. In an embodiment, the viral vector has the ability of nuclear localization. For example, a virus that requires the breakdown of the cell wall (during cell division) and therefore will not infect a non-diving cell can be altered to incorporate a nuclear localization peptide in the matrix protein of the virus thereby enabling the transduction of non-proliferating cells.

In some embodiments, the Cas9- and/or gRNA-encoding DNA is delivered by a non-vector based method (e.g., using naked DNA or DNA complexes). For example, the DNA can be delivered, e.g., by organically modified silica or silicate (Ormosil), electroporation, gene gun,

sonoporation, magnetofection, lipid-mediated transfection, dendrimers, inorganic nanoparticles, calcium phosphates, or a combination thereof.

In some embodiments, the Cas9- and/or gRNA-encoding DNA is delivered by a combination of a vector and a non-vector based method. For example, a virosome comprises a liposome combined with an inactivated virus (e.g., HIV or influenza virus), which can result in more efficient gene transfer, e.g., in a respiratory epithelial cell than either a viral or a liposomal method alone.

In an embodiment, the delivery vehicle is a non-viral vector. In an embodiment, the non-viral vector is an inorganic nanoparticle. Exemplary inorganic nanoparticles include, e.g., magnetic nanoparticles (e.g., Fe_3MnO_2) and silica. The outer surface of the nanoparticle can be conjugated with a positively charged polymer (e.g., polyethylenimine, polylysine, polyserine) which allows for attachment (e.g., conjugation or entrapment) of payload. In an embodiment, the non-viral vector is an organic nanoparticle (e.g., entrapment of the payload inside the nanoparticle). Exemplary organic nanoparticles include, e.g., SNALP liposomes that contain cationic lipids together with neutral helper lipids which are coated with polyethylene glycol (PEG) and protamine and nucleic acid complex coated with lipid coating.

Exemplary lipids for gene transfer are shown below in **Table 20**.

Table 20: Lipids Used for Gene Transfer

Lipid	Abbreviation	Feature
1,2-Dioleoyl-sn-glycero-3-phosphatidylcholine	DOPC	Helper
1,2-Dioleoyl-sn-glycero-3-phosphatidylethanolamine	DOPE	Helper
Cholesterol		Helper
N-[1-(2,3-Dioleyloxy)prophyl]*N,N,N*-trimethylammonium chloride	DOTMA	Cationic
1,2-Dioleoyloxy-3-trimethylammonium-propane	DOTAP	Cationic
Dioctadecylamidoglycylspermine	DOGS	Cationic
N-(3-Aminopropyl)-*N,N*-dimethyl-2,3-bis(dodecyloxy)-1-propanaminium bromide	GAP-DLRIE	Cationic
Cetyltrimethylammonium bromide	CTAB	Cationic
6-Lauroxyhexyl ornithinate	LHON	Cationic

1-(2,3-Dioleoyloxypropyl)-2,4,6-trimethylpyridinium	2Oc	Cationic
2,3-Dioleyloxy-*N*-[2(sperminecarboxamido-ethyl]-*N,N*-dimethyl-1-propanaminium trifluoroacetate	DOSPA	Cationic
1,2-Dioleyl-3-trimethylammonium-propane	DOPA	Cationic
N-(2-Hydroxyethyl)-*N,N*-dimethyl-2,3-bis(tetradecyloxy)-1-propanaminium bromide	MDRIE	Cationic
Dimyristooxypropyl dimethyl hydroxyethyl ammonium bromide	DMRI	Cationic
3β-[*N*-(*N'*,*N'*-Dimethylaminoethane)-carbamoyl]cholesterol	DC-Chol	Cationic
Bis-guanidium-tren-cholesterol	BGTC	Cationic
1,3-Diodeoxy-2-(6-carboxy-spermyl)-propylamide	DOSPER	Cationic
Dimethyloctadecylammonium bromide	DDAB	Cationic
Dioctadecylamidoglicylspermidin	DSL	Cationic
rac-[(2,3-Dioctadecyloxypropyl)(2-hydroxyethyl)]-dimethylammonium chloride	CLIP-1	Cationic
rac-[2(2,3-Dihexadecyloxypropyl-oxymethyloxy)ethyl]trimethylammonium bromide	CLIP-6	Cationic
Ethyldimyristoylphosphatidylcholine	EDMPC	Cationic
1,2-Distearyloxy-*N,N*-dimethyl-3-aminopropane	DSDMA	Cationic
1,2-Dimyristoyl-trimethylammonium propane	DMTAP	Cationic
O,O'-Dimyristyl-N-lysyl aspartate	DMKE	Cationic
1,2-Distearoyl-sn-glycero-3-ethylphosphocholine	DSEPC	Cationic
N-Palmitoyl D-erythro-sphingosyl carbamoyl-spermine	CCS	Cationic
N-t-Butyl-*N*0-tetradecyl-3-tetradecylaminopropionamidine	diC14-amidine	Cationic
Octadecenolyoxy[ethyl-2-heptadecenyl-3 hydroxyethyl] imidazolinium chloride	DOTIM	Cationic
*N*1-Cholesteryloxycarbonyl-3,7-diazanonane-1,9-diamine	CDAN	Cationic
2-(3-[Bis(3-amino-propyl)-amino]propylamino)-*N*-ditetradecylcarbamoylme-ethyl-acetamide	RPR209120	Cationic
1,2-dilinoleyloxy-3- dimethylaminopropane	DLinDMA	Cationic
2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]- dioxolane	DLin-KC2-	Cationic

	DMA	
dilinoleyl- methyl-4-dimethylaminobutyrate	DLin-MC3-DMA	Cationic

Exemplary polymers for gene transfer are shown below in **Table 21**.

Table 21: Polymers Used for Gene Transfer

Polymer	Abbreviation
Poly(ethylene)glycol	PEG
Polyethylenimine	PEI
Dithiobis(succinimidylpropionate)	DSP
Dimethyl-3,3'-dithiobispropionimidate	DTBP
Poly(ethylene imine)biscarbamate	PEIC
Poly(L-lysine)	PLL
Histidine modified PLL	
Poly(*N*-vinylpyrrolidone)	PVP
Poly(propylenimine)	PPI
Poly(amidoamine)	PAMAM
Poly(amidoethylenimine)	SS-PAEI
Triethylenetetramine	TETA
Poly(β-aminoester)	
Poly(4-hydroxy-L-proline ester)	PHP
Poly(allylamine)	
Poly(α-[4-aminobutyl]-L-glycolic acid)	PAGA
Poly(D,L-lactic-co-glycolic acid)	PLGA
Poly(*N*-ethyl-4-vinylpyridinium bromide)	
Poly(phosphazene)s	PPZ
Poly(phosphoester)s	PPE
Poly(phosphoramidate)s	PPA
Poly(*N*-2-hydroxypropylmethacrylamide)	pHPMA
Poly (2-(dimethylamino)ethyl methacrylate)	pDMAEMA

Poly(2-aminoethyl propylene phosphate)	PPE-EA
Chitosan	
Galactosylated chitosan	
N-Dodacylated chitosan	
Histone	
Collagen	
Dextran-spermine	D-SPM

In an embodiment, the vehicle has targeting modifications to increase target cell update of nanoparticles and liposomes, e.g., cell specific antigens, monoclonal antibodies, single chain antibodies, aptamers, polymers, sugars, and cell penetrating peptides. In an embodiment, the vehicle uses fusogenic and endosome-destabilizing peptides/polymers. In an embodiment, the vehicle undergoes acid-triggered conformational changes (e.g., to accelerate endosomal escape of the cargo). In an embodiment, a stimuli-cleavable polymer is used, e.g., for release in a cellular compartment. For example, disulfide-based cationic polymers that are cleaved in the reducing cellular environment can be used.

In an embodiment, the delivery vehicle is a biological non-viral delivery vehicle. In an embodiment, the vehicle is an attenuated bacterium (e.g., naturally or artificially engineered to be invasive but attenuated to prevent pathogenesis and expressing the transgene (e.g., *Listeria monocyto*gene*s*, certain *Salmonella strains*, *Bifidobacteriumlongum*, and modified *Escherichia coli*), bacteria having nutritional and tissue-specific tropism to target specific tissues, bacteria having modified surface proteins to alter target tissue specificity). In an embodiment, the vehicle is a genetically modified bacteriophage (e.g., engineered phages having large packaging capacity, less immunogenic, containing mammalian plasmid maintenance sequences and having incorporated targeting ligands). In an embodiment, the vehicle is a mammalian virus-like particle. For example, modified viral particles can be generated (e.g., by purification of the "empty" particles followed by *ex vivo* assembly of the virus with the desired cargo). The vehicle can also be engineered to incorporate targeting ligands to alter target tissue specificity. In an embodiment, the vehicle is a biological liposome. For example, the biological liposome is a phospholipid-based particle derived from human cells (e.g., erythrocyte ghosts, which are red blood cells broken down into spherical structures derived from the subject (e.g., tissue targeting

can be achieved by attachment of various tissue or cell-specific ligands), or secretory exosomes – subject (i.e., patient) derived membrane-bound nanovescicle (30 -100 nm) of endocytic origin (e.g., can be produced from various cell types and can therefore be taken up by cells without the need of for targeting ligands).

In an embodiment, one or more nucleic acid molecules (e.g., DNA molecules) other than the components of a Cas system, e.g., the Cas9 molecule component and/or the gRNA molecule component described herein, are delivered. In an embodiment, the nucleic acid molecule is delivered at the same time as one or more of the components of the Cas system are delivered. In an embodiment, the nucleic acid molecule is delivered before or after (e.g., less than about 30 minutes, 1 hour, 2 hours, 3 hours, 6 hours, 9 hours, 12 hours, 1 day, 2 days, 3 days, 1 week, 2 weeks, or 4 weeks) one or more of the components of the Cas system are delivered. In an embodiment, the nucleic acid molecule is delivered by a different means than one or more of the components of the Cas system, e.g., the Cas9 molecule component and/or the gRNA molecule component, are delivered. The nucleic acid molecule can be delivered by any of the delivery methods described herein. For example, the nucleic acid molecule can be delivered by a viral vector, e.g., an integration-deficient lentivirus, and the Cas9 molecule component and/or the gRNA molecule component can be delivered by electroporation, e.g., such that the toxicity caused by nucleic acids (e.g., DNAs) can be reduced. In an embodiment, the nucleic acid molecule encodes a therapeutic protein, e.g., a protein described herein. In an embodiment, the nucleic acid molecule encodes an RNA molecule, e.g., an RNA molecule described herein.

Delivery of RNA encoding a Cas9 molecule

RNA encoding Cas9 molecules (e.g., eaCas9 molecules) and/or gRNA molecules, can be delivered into cells, e.g., target cells described herein, by art-known methods or as described herein. For example, Cas9-encoding and/or gRNA-encoding RNA can be delivered, e.g., by microinjection, electroporation, lipid-mediated transfection, peptide-mediated delivery, or a combination thereof. Cas9-encoding and/or gRNA-encoding RNA can be conjugated to molecules (e.g., GalNAc) promoting uptake by the target cells (e.g., target cells described herein).

Delivery Cas9 molecule protein

Cas9 molecules (e.g., eaCas9 molecules) can be delivered into cells by art-known methods or as described herein. For example, Cas9 protein molecules can be delivered, e.g., by microinjection, electroporation, lipid-mediated transfection, peptide-mediated delivery, or a combination thereof. Delivery can be accompanied by DNA encoding a gRNA or by a gRNA. Cas9-encoding and/or gRNA-encoding RNA can be conjugated to molecules (e.g., GalNAc) promoting uptake by the target cells (e.g., target cells described herein).

Route of administration

Systemic modes of administration include oral and parenteral routes. Parenteral routes include, by way of example, intravenous, intrarterial, intramuscular, intradermal, subcutaneous, intranasal and intraperitoneal routes. Components administered systemically may be modified or formulated to target the components to the eye.

Local modes of administration include, by way of example, intraocular, intraorbital, subconjuctival, intravitreal, subretinal or transscleral routes. In an embodiment, significantly smaller amounts of the components (compared with systemic approaches) may exert an effect when administered locally (for example, intravitreally) compared to when administered systemically (for example, intravenously). Local modes of administration can reduce or eliminate the incidence of potentially toxic side effects that may occur when therapeutically effective amounts of a component are administered systemically.

In an embodiment, components described herein are delivered subretinally, e.g., by subretinal injection. Subretinal injections may be made directly into the macular, e.g., submacular injection.

In an embodiment, components described herein are delivered by intravitreal injection. Intravitreal injection has a relatively low risk of retinal detachment. In an embodiment, nanoparticle or viral, e.g., AAV vector, is delivered intravitreally.

Methods for administration of agents to the eye are known in the medical arts and can be used to administer components described herein. Exemplary methods include intraocular injection (e.g., retrobulbar, subretinal, submacular, intravitreal and intrachoridal), iontophoresis, eye drops, and intraocular implantation (e.g., intravitreal, sub-Tenons and sub-conjunctival).

Administration may be provided as a periodic bolus (for example, subretinally,

intravenously or intravitreally) or as continuous infusion from an internal reservoir (for example, from an implant disposed at an intra- or extra-ocular location (see, U.S. Pat. Nos. 5,443,505 and 5,766,242)) or from an external reservoir (for example, from an intravenous bag). Components may be administered locally, for example, by continuous release from a sustained release drug delivery device immobilized to an inner wall of the eye or via targeted transscleral controlled release into the choroid (see, for example, PCT/US00/00207, PCT/US02/14279, Ambati et al. (2000) INVEST. OPHTHALMOL. VIS. SCI.41:1181-1185, and Ambati et al. (2000) INVEST. OPHTHALMOL. VIS. SCI.41:1186-1191). A variety of devices suitable for administering components locally to the inside of the eye are known in the art. See, for example, U.S. Pat. Nos. 6,251,090, 6,299,895, 6,416,777, 6,413,540, and PCT/US00/28187.

In addition, components may be formulated to permit release over a prolonged period of time. A release system can include a matrix of a biodegradable material or a material which releases the incorporated components by diffusion. The components can be homogeneously or heterogeneously distributed within the release system. A variety of release systems may be useful, however, the choice of the appropriate system will depend upon rate of release required by a particular application. Both non-degradable and degradable release systems can be used. Suitable release systems include polymers and polymeric matrices, non-polymeric matrices, or inorganic and organic excipients and diluents such as, but not limited to, calcium carbonate and sugar (for example, trehalose). Release systems may be natural or synthetic. However, synthetic release systems are preferred because generally they are more reliable, more reproducible and produce more defined release profiles. The release system material can be selected so that components having different molecular weights are released by diffusion through or degradation of the material.

Representative synthetic, biodegradable polymers include, for example: polyamides such as poly(amino acids) and poly(peptides); polyesters such as poly(lactic acid), poly(glycolic acid), poly(lactic-co-glycolic acid), and poly(caprolactone); poly(anhydrides); polyorthoesters; polycarbonates; and chemical derivatives thereof (substitutions, additions of chemical groups, for example, alkyl, alkylene, hydroxylations, oxidations, and other modifications routinely made by those skilled in the art), copolymers and mixtures thereof. Representative synthetic, non-degradable polymers include, for example: polyethers such as poly(ethylene oxide), poly(ethylene glycol), and poly(tetramethylene oxide); vinyl polymers-polyacrylates and

polymethacrylates such as methyl, ethyl, other alkyl, hydroxyethyl methacrylate, acrylic and methacrylic acids, and others such as poly(vinyl alcohol), poly(vinyl pyrolidone), and poly(vinyl acetate); poly(urethanes); cellulose and its derivatives such as alkyl, hydroxyalkyl, ethers, esters, nitrocellulose, and various cellulose acetates; polysiloxanes; and any chemical derivatives thereof (substitutions, additions of chemical groups, for example, alkyl, alkylene, hydroxylations, oxidations, and other modifications routinely made by those skilled in the art), copolymers and mixtures thereof.

Poly(lactide-co-glycolide) microsphere can also be used for intraocular injection. Typically the microspheres are composed of a polymer of lactic acid and glycolic acid, which are structured to form hollow spheres. The spheres can be approximately 15-30 microns in diameter and can be loaded with components described herein.

Bi-Modal or Differential Delivery of Components

Separate delivery of the components of a Cas system, e.g., the Cas9 molecule component and the gRNA molecule component, and more particularly, delivery of the components by differing modes, can enhance performance, e.g., by improving tissue specificity and safety.

In an embodiment, the Cas9 molecule and the gRNA molecule are delivered by different modes, or as sometimes referred to herein as differential modes. Different or differential modes, as used herein, refer modes of delivery that confer different pharmacodynamic or pharmacokinetic properties on the subject component molecule, e.g., a Cas9 molecule, gRNA molecule, template nucleic acid, or payload. For example, the modes of delivery can result in different tissue distribution, different half-life, or different temporal distribution, e.g., in a selected compartment, tissue, or organ.

Some modes of delivery, e.g., delivery by a nucleic acid vector that persists in a cell, or in progeny of a cell, e.g., by autonomous replication or insertion into cellular nucleic acid, result in more persistent expression of and presence of a component. Examples include viral, e.g., adeno associated virus or lentivirus, delivery.

By way of example, the components, e.g., a Cas9 molecule and a gRNA molecule, can be delivered by modes that differ in terms of resulting half-life or persistent of the delivered component the body, or in a particular compartment, tissue or organ. In an embodiment, a gRNA molecule can be delivered by such modes. The Cas9 molecule component can be

delivered by a mode which results in less persistence or less exposure to the body or a particular compartment or tissue or organ.

More generally, in an embodiment, a first mode of delivery is used to deliver a first component and a second mode of delivery is used to deliver a second component. The first mode of delivery confers a first pharmacodynamic or pharmacokinetic property. The first pharmacodynamic property can be, e.g., distribution, persistence, or exposure, of the component, or of a nucleic acid that encodes the component, in the body, a compartment, tissue or organ. The second mode of delivery confers a second pharmacodynamic or pharmacokinetic property. The second pharmacodynamic property can be, e.g., distribution, persistence, or exposure, of the component, or of a nucleic acid that encodes the component, in the body, a compartment, tissue or organ.

In an embodiment, the first pharmacodynamic or pharmacokinetic property, e.g., distribution, persistence or exposure, is more limited than the second pharmacodynamic or pharmacokinetic property.

In an embodiment, the first mode of delivery is selected to optimize, e.g., minimize, a pharmacodynamic or pharmacokinetic property, e.g., distribution, persistence or exposure.

In an embodiment, the second mode of delivery is selected to optimize, e.g., maximize, a pharmacodynamic or pharmacokinetic property, e.g., distribution, persistence or exposure.

In an embodiment, the first mode of delivery comprises the use of a relatively persistent element, e.g., a nucleic acid, e.g., a plasmid or viral vector, e.g., an AAV or lentivirus. As such vectors are relatively persistent product transcribed from them would be relatively persistent.

In an embodiment, the second mode of delivery comprises a relatively transient element, e.g., an RNA or protein.

In an embodiment, the first component comprises gRNA, and the delivery mode is relatively persistent, e.g., the gRNA is transcribed from a plasmid or viral vector, e.g., an AAV or lentivirus. Transcription of these genes would be of little physiological consequence because the genes do not encode for a protein product, and the gRNAs are incapable of acting in isolation. The second component, a Cas9 molecule, is delivered in a transient manner, for example as mRNA or as protein, ensuring that the full Cas9 molecule/gRNA molecule complex is only present and active for a short period of time.

Furthermore, the components can be delivered in different molecular form or with different delivery vectors that complement one another to enhance safety and tissue specificity.

Use of differential delivery modes can enhance performance, safety and efficacy. E.g., the likelihood of an eventual off-target modification can be reduced. Delivery of immunogenic components, e.g., Cas9 molecules, by less persistent modes can reduce immunogenicity, as peptides from the bacterially-derived Cas enzyme are displayed on the surface of the cell by MHC molecules. A two-part delivery system can alleviate these drawbacks.

Differential delivery modes can be used to deliver components to different, but overlapping target regions. The formation active complex is minimized outside the overlap of the target regions. Thus, in an embodiment, a first component, e.g., a gRNA molecule is delivered by a first delivery mode that results in a first spatial, e.g., tissue, distribution. A second component, e.g., a Cas9 molecule is delivered by a second delivery mode that results in a second spatial, e.g., tissue, distribution. In an embodiment the first mode comprises a first element selected from a liposome, nanoparticle, e.g., polymeric nanoparticle, and a nucleic acid, e.g., viral vector. The second mode comprises a second element selected from the group. In an embodiment, the first mode of delivery comprises a first targeting element, e.g., a cell specific receptor or an antibody, and the second mode of delivery does not include that element. In embodiment, the second mode of delivery comprises a second targeting element, e.g., a second cell specific receptor or second antibody.

When the Cas9 molecule is delivered in a virus delivery vector, a liposome, or polymeric nanoparticle, there is the potential for delivery to and therapeutic activity in multiple tissues, when it may be desirable to only target a single tissue. A two-part delivery system can resolve this challenge and enhance tissue specificity. If the gRNA molecule and the Cas9 molecule are packaged in separated delivery vehicles with distinct but overlapping tissue tropism, the fully functional complex is only be formed in the tissue that is targeted by both vectors.

Ex vivo delivery

In some embodiments, components described in **Table 17** are introduced into cells which are then introduced into the subject. Methods of introducing the components can include, e.g., any of the delivery methods described in **Table 18.**

VIII. Modified Nucleosides, Nucleotides, and Nucleic Acids

Modified nucleosides and modified nucleotides can be present in nucleic acids, e.g., particularly gRNA, but also other forms of RNA, e.g., mRNA, RNAi, or siRNA. As described herein, "nucleoside" is defined as a compound containing a five-carbon sugar molecule (a pentose or ribose) or derivative thereof, and an organic base, purine or pyrimidine, or a derivative thereof. As described herein, "nucleotide" is defined as a nucleoside further comprising a phosphate group.

Modified nucleosides and nucleotides can include one or more of:

(i) alteration, e.g., replacement, of one or both of the non-linking phosphate oxygens and/or of one or more of the linking phosphate oxygens in the phosphodiester backbone linkage;

(ii) alteration, e.g., replacement, of a constituent of the ribose sugar, e.g., of the 2' hydroxyl on the ribose sugar;

(iii) wholesale replacement of the phosphate moiety with "dephospho" linkers;

(iv) modification or replacement of a naturally occurring nucleobase;

(v) replacement or modification of the ribose-phosphate backbone;

(vi) modification of the 3' end or 5' end of the oligonucleotide, e.g., removal, modification or replacement of a terminal phosphate group or conjugation of a moiety; and

(vii) modification of the sugar.

The modifications listed above can be combined to provide modified nucleosides and nucleotides that can have two, three, four, or more modifications. For example, a modified nucleoside or nucleotide can have a modified sugar and a modified nucleobase. In an embodiment, every base of a gRNA is modified, e.g., all bases have a modified phosphate group, e.g., all are phosphorothioate groups. In an embodiment, all, or substantially all, of the phosphate groups of a unimolecular or modular gRNA molecule are replaced with phosphorothioate groups.

In an embodiment, modified nucleotides, e.g., nucleotides having modifications as described herein, can be incorporated into a nucleic acid, e.g., a "modified nucleic acid." In some embodiments, the modified nucleic acids comprise one, two, three or more modified nucleotides. In some embodiments, at least 5% (e.g., at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 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%, or about 100%) of the positions in a modified nucleic acid are a modified nucleotides.

Unmodified nucleic acids can be prone to degradation by, e.g., cellular nucleases. For example, nucleases can hydrolyze nucleic acid phosphodiester bonds. Accordingly, in one aspect the modified nucleic acids described herein can contain one or more modified nucleosides or nucleotides, e.g., to introduce stability toward nucleases.

In some embodiments, the modified nucleosides, modified nucleotides, and modified nucleic acids described herein can exhibit a reduced innate immune response when introduced into a population of cells, both *in vivo* and *ex vivo*. The term "innate immune response" includes a cellular response to exogenous nucleic acids, including single stranded nucleic acids, generally of viral or bacterial origin, which involves the induction of cytokine expression and release, particularly the interferons, and cell death. In some embodiments, the modified nucleosides, modified nucleotides, and modified nucleic acids described herein can disrupt binding of a major groove interacting partner with the nucleic acid. In some embodiments, the modified nucleosides, modified nucleotides, and modified nucleic acids described herein can exhibit a reduced innate immune response when introduced into a population of cells, both *in vivo* and *ex vivo*, and also disrupt binding of a major groove interacting partner with the nucleic acid.

Definitions of Chemical Groups

As used herein, "alkyl" is meant to refer to a saturated hydrocarbon group which is straight-chained or branched. Example alkyl groups include methyl (Me), ethyl (Et), propyl (*e.g.*, n-propyl and isopropyl), butyl (*e.g.*, n-butyl, isobutyl, t-butyl), pentyl (*e.g.*, n-pentyl, isopentyl, neopentyl), and the like. An alkyl group can contain from 1 to about 20, from 2 to about 20, from 1 to about 12, from 1 to about 8, from 1 to about 6, from 1 to about 4, or from 1 to about 3 carbon atoms.

As used herein, "aryl" refers to monocyclic or polycyclic (*e.g.*, having 2, 3 or 4 fused rings) aromatic hydrocarbons such as, for example, phenyl, naphthyl, anthracenyl, phenanthrenyl, indanyl, indenyl, and the like. In some embodiments, aryl groups have from 6 to about 20 carbon atoms.

As used herein, "alkenyl" refers to an aliphatic group containing at least one double bond.

As used herein, "alkynyl" refers to a straight or branched hydrocarbon chain containing 2-12 carbon atoms and characterized in having one or more triple bonds. Examples of alkynyl groups include, but are not limited to, ethynyl, propargyl, and 3-hexynyl.

As used herein, "arylalkyl" or "aralkyl" refers to an alkyl moiety in which an alkyl hydrogen atom is replaced by an aryl group. Aralkyl includes groups in which more than one hydrogen atom has been replaced by an aryl group. Examples of "arylalkyl" or "aralkyl" include benzyl, 2-phenylethyl, 3-phenylpropyl, 9-fluorenyl, benzhydryl, and trityl groups.

As used herein, "cycloalkyl" refers to a cyclic, bicyclic, tricyclic, or polycyclic non-aromatic hydrocarbon groups having 3 to 12 carbons. Examples of cycloalkyl moieties include, but are not limited to, cyclopropyl, cyclopentyl, and cyclohexyl.

As used herein, "heterocyclyl" refers to a monovalent radical of a heterocyclic ring system. Representative heterocyclyls include, without limitation, tetrahydrofuranyl, tetrahydrothienyl, pyrrolidinyl, pyrrolidonyl, piperidinyl, pyrrolinyl, piperazinyl, dioxanyl, dioxolanyl, diazepinyl, oxazepinyl, thiazepinyl, and morpholinyl.

As used herein, "heteroaryl" refers to a monovalent radical of a heteroaromatic ring system. Examples of heteroaryl moieties include, but are not limited to, imidazolyl, oxazolyl, thiazolyl, triazolyl, pyrrolyl, furanyl, indolyl, thiophenyl pyrazolyl, pyridinyl, pyrazinyl, pyridazinyl, pyrimidinyl, indolizinyl, purinyl, naphthyridinyl, quinolyl, and pteridinyl.

Phosphate Backbone Modifications

The Phosphate Group

In some embodiments, the phosphate group of a modified nucleotide can be modified by replacing one or more of the oxygens with a different substituent. Further, the modified nucleotide, e.g., modified nucleotide present in a modified nucleic acid, can include the wholesale replacement of an unmodified phosphate moiety with a modified phosphate as described herein. In some embodiments, the modification of the phosphate backbone can include alterations that result in either an uncharged linker or a charged linker with unsymmetrical charge distribution.

Examples of modified phosphate groups include, phosphorothioate, phosphoroselenates, borano phosphates, borano phosphate esters, hydrogen phosphonates, phosphoroamidates, alkyl or aryl phosphonates and phosphotriesters. In some embodiments, one of the non-bridging phosphate oxygen atoms in the phosphate backbone moiety can be replaced by any of the

following groups: sulfur (S), selenium (Se), BR_3 (wherein R can be, e.g., hydrogen, alkyl, or aryl), C (e.g., an alkyl group, an aryl group, and the like), H, NR_2 (wherein R can be, e.g., hydrogen, alkyl, or aryl), or OR (wherein R can be, e.g., alkyl or aryl). The phosphorous atom in an unmodified phosphate group is achiral. However, replacement of one of the non-bridging oxygens with one of the above atoms or groups of atoms can render the phosphorous atom chiral; that is to say that a phosphorous atom in a phosphate group modified in this way is a stereogenic center. The stereogenic phosphorous atom can possess either the "R" configuration (herein Rp) or the "S" configuration (herein Sp).

Phosphorodithioates have both non-bridging oxygens replaced by sulfur. The phosphorus center in the phosphorodithioates is achiral which precludes the formation of oligoribonucleotide diastereomers. In some embodiments, modifications to one or both non-bridging oxygens can also include the replacement of the non-bridging oxygens with a group independently selected from S, Se, B, C, H, N, and OR (R can be, e.g., alkyl or aryl).

The phosphate linker can also be modified by replacement of a bridging oxygen, (i.e., the oxygen that links the phosphate to the nucleoside), with nitrogen (bridged phosphoroamidates), sulfur (bridged phosphorothioates) and carbon (bridged methylenephosphonates). The replacement can occur at either linking oxygen or at both of the linking oxygens.

Replacement of the Phosphate Group

The phosphate group can be replaced by non-phosphorus containing connectors. In some embodiments, the charge phosphate group can be replaced by a neutral moiety.

Examples of moieties which can replace the phosphate group can include, without limitation, e.g., methyl phosphonate, hydroxylamino, siloxane, carbonate, carboxymethyl, carbamate, amide, thioether, ethylene oxide linker, sulfonate, sulfonamide, thioformacetal, formacetal, oxime, methyleneimino, methylenemethylimino, methylenehydrazo, methylenedimethylhydrazo and methyleneoxymethylimino.

Replacement of the Ribophosphate Backbone

Scaffolds that can mimic nucleic acids can also be constructed wherein the phosphate linker and ribose sugar are replaced by nuclease resistant nucleoside or nucleotide surrogates. In some embodiments, the nucleobases can be tethered by a surrogate backbone. Examples can include, without limitation, the morpholino, cyclobutyl, pyrrolidine and peptide nucleic acid (PNA) nucleoside surrogates.

Sugar Modifications

The modified nucleosides and modified nucleotides can include one or more modifications to the sugar group. For example, the 2' hydroxyl group (OH) can be modified or replaced with a number of different "oxy" or "deoxy" substituents. In some embodiments, modifications to the 2' hydroxyl group can enhance the stability of the nucleic acid since the hydroxyl can no longer be deprotonated to form a 2'-alkoxide ion. The 2'-alkoxide can catalyze degradation by intramolecular nucleophilic attack on the linker phosphorus atom.

Examples of "oxy"-2' hydroxyl group modifications can include alkoxy or aryloxy (OR, wherein "R" can be, e.g., alkyl, cycloalkyl, aryl, aralkyl, heteroaryl or a sugar); polyethyleneglycols (PEG), $O(CH_2CH_2O)_nCH_2CH_2OR$ wherein R can be, e.g., H or optionally substituted alkyl, and n can be an integer from 0 to 20 (e.g., from 0 to 4, from 0 to 8, from 0 to 10, from 0 to 16, from 1 to 4, from 1 to 8, from 1 to 10, from 1 to 16, from 1 to 20, from 2 to 4, from 2 to 8, from 2 to 10, from 2 to 16, from 2 to 20, from 4 to 8, from 4 to 10, from 4 to 16, and from 4 to 20). In some embodiments, the "oxy"-2' hydroxyl group modification can include "locked" nucleic acids (LNA) in which the 2' hydroxyl can be connected, e.g., by a C_{1-6} alkylene or C_{1-6} heteroalkylene bridge, to the 4' carbon of the same ribose sugar, where exemplary bridges can include methylene, propylene, ether, or amino bridges; O-amino (wherein amino can be, e.g., NH_2; alkylamino, dialkylamino, heterocyclyl, arylamino, diarylamino, heteroarylamino, or diheteroarylamino, ethylenediamine, or polyamino) and aminoalkoxy, $O(CH_2)_n$-amino, (wherein amino can be, e.g., NH_2; alkylamino, dialkylamino, heterocyclyl, arylamino, diarylamino, heteroarylamino, or diheteroarylamino, ethylenediamine, or polyamino). In some embodiments, the "oxy"-2' hydroxyl group modification can include the methoxyethyl group (MOE), ($OCH_2CH_2OCH_3$, e.g., a PEG derivative).

"Deoxy" modifications can include hydrogen (i.e. deoxyribose sugars, e.g., at the overhang portions of partially ds RNA); halo (e.g., bromo, chloro, fluoro, or iodo); amino

(wherein amino can be, e.g., NH_2; alkylamino, dialkylamino, heterocyclyl, arylamino, diarylamino, heteroarylamino, diheteroarylamino, or amino acid); $NH(CH_2CH_2NH)_nCH_2CH_2$-amino (wherein amino can be, e.g., as described herein), -NHC(O)R (wherein R can be, e.g., alkyl, cycloalkyl, aryl, aralkyl, heteroaryl or sugar), cyano; mercapto; alkyl-thio-alkyl; thioalkoxy; and alkyl, cycloalkyl, aryl, alkenyl and alkynyl, which may be optionally substituted with e.g., an amino as described herein.

The sugar group can also contain one or more carbons that possess the opposite stereochemical configuration than that of the corresponding carbon in ribose. Thus, a modified nucleic acid can include nucleotides containing e.g., arabinose, as the sugar. The nucleotide "monomer" can have an alpha linkage at the 1' position on the sugar, e.g., alpha-nucleosides. The modified nucleic acids can also include "abasic" sugars, which lack a nucleobase at C-1'. These abasic sugars can also be further modified at one or more of the constituent sugar atoms. The modified nucleic acids can also include one or more sugars that are in the L form, e.g. L-nucleosides.

Generally, RNA includes the sugar group ribose, which is a 5-membered ring having an oxygen. Exemplary modified nucleosides and modified nucleotides can include, without limitation, replacement of the oxygen in ribose (e.g., with sulfur (S), selenium (Se), or alkylene, such as, e.g., methylene or ethylene); addition of a double bond (e.g., to replace ribose with cyclopentenyl or cyclohexenyl); ring contraction of ribose (e.g., to form a 4-membered ring of cyclobutane or oxetane); ring expansion of ribose (e.g., to form a 6- or 7-membered ring having an additional carbon or heteroatom, such as for example, anhydrohexitol, altritol, mannitol, cyclohexanyl, cyclohexenyl, and morpholino that also has a phosphoramidate backbone). In some embodiments, the modified nucleotides can include multicyclic forms (e.g., tricyclo; and "unlocked" forms, such as glycol nucleic acid (GNA) (e.g., R-GNA or S-GNA, where ribose is replaced by glycol units attached to phosphodiester bonds), threose nucleic acid (TNA, where ribose is replaced with α-L-threofuranosyl-(3'→2')).

Modifications on the Nucleobase

The modified nucleosides and modified nucleotides described herein, which can be incorporated into a modified nucleic acid, can include a modified nucleobase. Examples of nucleobases include, but are not limited to, adenine (A), guanine (G), cytosine (C), and uracil (U). These nucleobases can be modified or wholly replaced to provide modified nucleosides and

modified nucleotides that can be incorporated into modified nucleic acids. The nucleobase of the nucleotide can be independently selected from a purine, a pyrimidine, a purine or pyrimidine analog. In some embodiments, the nucleobase can include, for example, naturally-occurring and synthetic derivatives of a base.

Uracil

In some embodiments, the modified nucleobase is a modified uracil. Exemplary nucleobases and nucleosides having a modified uracil include without limitation pseudouridine (ψ), pyridin-4-one ribonucleoside, 5-aza-uridine, 6-aza-uridine, 2-thio-5-aza-uridine, 2-thio-uridine (s2U), 4-thio-uridine (s4U), 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxy-uridine (ho^5U), 5-aminoallyl-uridine, 5-halo-uridine (e.g., 5-iodo-uridine or 5-bromo-uridine), 3-methyl-uridine (m^3U), 5-methoxy-uridine (mo^5U), uridine 5-oxyacetic acid (cmo^5U), uridine 5-oxyacetic acid methyl ester ($mcmo^5U$), 5-carboxymethyl-uridine (cm^5U), 1-carboxymethyl-pseudouridine, 5-carboxyhydroxymethyl-uridine (chm^5U), 5-carboxyhydroxymethyl-uridine methyl ester ($mchm^5U$), 5-methoxycarbonylmethyl-uridine (mcm^5U), 5-methoxycarbonylmethyl-2-thio-uridine (mcm^5s2U), 5-aminomethyl-2-thio-uridine (nm^5s2U), 5-methylaminomethyl-uridine (mnm^5U), 5-methylaminomethyl-2-thio-uridine (mnm^5s2U), 5-methylaminomethyl-2-seleno-uridine (mnm^5se^2U), 5-carbamoylmethyl-uridine (ncm^5U), 5-carboxymethylaminomethyl-uridine ($cmnm^5U$), 5-carboxymethylaminomethyl-2-thio-uridine ($cmnm^5s2U$), 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-taurinomethyl-uridine (τcm^5U), 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine(τm^5s2U), 1-taurinomethyl-4-thio-pseudouridine, 5-methyl-uridine (m^5U, i.e., having the nucleobase deoxythymine), 1-methyl-pseudouridine ($m^1\psi$), 5-methyl-2-thio-uridine (m^5s2U), 1-methyl-4-thio-pseudouridine ($m^1s^4\psi$), 4-thio-1-methyl-pseudouridine, 3-methyl-pseudouridine ($m^3\psi$), 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, dihydrouridine (D), dihydropseudouridine, 5,6-dihydrouridine, 5-methyl-dihydrouridine (m^5D), 2-thio-dihydrouridine, 2-thio-dihydropseudouridine, 2-methoxy-uridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, N1-methyl-pseudouridine, 3-(3-amino-3-carboxypropyl)uridine (acp^3U), 1-methyl-3-(3-amino-3-carboxypropyl)pseudouridine ($acp^3\psi$), 5-(isopentenylaminomethyl)uridine (inm^5U), 5-(isopentenylaminomethyl)-2-thio-uridine (inm^5s2U), α-thio-uridine, 2'-O-methyl-uridine (Um), 5,2'-O-dimethyl-uridine (m^5Um), 2'-O-methyl-pseudouridine (ψm), 2-thio-2'-O-methyl-uridine

(s2Um), 5-methoxycarbonylmethyl-2'-O-methyl-uridine (mcm^{5}Um), 5-carbamoylmethyl-2'-O-methyl-uridine (ncm^{5}Um), 5-carboxymethylaminomethyl-2'-O-methyl-uridine (cmnm^{5}Um), 3,2'-O-dimethyl-uridine (m^{3}Um), 5-(isopentenylaminomethyl)-2'-O-methyl-uridine (inm^{5}Um), 1-thio-uridine, deoxythymidine, 2'-F-ara-uridine, 2'-F-uridine, 2'-OH-ara-uridine, 5-(2-carbomethoxyvinyl) uridine, 5-[3-(1-E-propenylamino)uridine, pyrazolo[3,4-d]pyrimidines, xanthine, and hypoxanthine.

Cytosine

In some embodiments, the modified nucleobase is a modified cytosine. Exemplary nucleobases and nucleosides having a modified cytosine include without limitation 5-aza-cytidine, 6-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine (m^{3}C), N4-acetyl-cytidine (act), 5-formyl-cytidine (f^{5}C), N4-methyl-cytidine (m^{4}C), 5-methyl-cytidine (m^{5}C), 5-halo-cytidine (e.g., 5-iodo-cytidine), 5-hydroxymethyl-cytidine (hm^{5}C), 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine (s2C), 2-thio-5-methyl-cytidine, 4-thio-pseudoisocytidine, 4-thio-1-methyl-pseudoisocytidine, 4-thio-1-methyl-1-deaza-pseudoisocytidine, 1-methyl-1-deaza-pseudoisocytidine, zebularine, 5-aza-zebularine, 5-methyl-zebularine, 5-aza-2-thio-zebularine, 2-thio-zebularine, 2-methoxy-cytidine, 2-methoxy-5-methyl-cytidine, 4-methoxy-pseudoisocytidine, 4-methoxy-1-methyl-pseudoisocytidine, lysidine (k^{2}C), α-thio-cytidine, 2'-O-methyl-cytidine (Cm), 5,2'-O-dimethyl-cytidine (m^{5}Cm), N4-acetyl-2'-O-methyl-cytidine (ac^{4}Cm), N4,2'-O-dimethyl-cytidine (m^{4}Cm), 5-formyl-2'-O-methyl-cytidine (f^{5}Cm), N4,N4,2'-O-trimethyl-cytidine (m$^4{}_2$Cm), 1-thio-cytidine, 2'-F-ara-cytidine, 2'-F-cytidine, and 2'-OH-ara-cytidine.

Adenine

In some embodiments, the modified nucleobase is a modified adenine. Exemplary nucleobases and nucleosides having a modified adenine include without limitation 2-amino-purine, 2,6-diaminopurine, 2-amino-6-halo-purine (e.g., 2-amino-6-chloro-purine), 6-halo-purine (e.g., 6-chloro-purine), 2-amino-6-methyl-purine, 8-azido-adenosine, 7-deaza-adenine, 7-deaza-8-aza-adenine, 7-deaza-2-amino-purine, 7-deaza-8-aza-2-amino-purine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyl-adenosine (m^{1}A), 2-methyl-adenine (m^{2}A), N6-methyl-adenosine (m^{6}A), 2-methylthio-N6-methyl-adenosine (ms2m^{6}A), N6-isopentenyl-adenosine (i^{6}A), 2-methylthio-N6-isopentenyl-adenosine (ms^{2}i^{6}A), N6-(cis-hydroxyisopentenyl)adenosine (io^{6}A), 2-methylthio-N6-(cis-hydroxyisopentenyl)adenosine

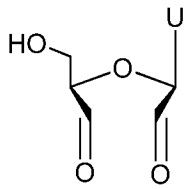
(ms2io^{6}A), N6-glycinylcarbamoyl-adenosine (g^{6}A), N6-threonylcarbamoyl-adenosine (t^{6}A), N6-methyl-N6-threonylcarbamoyl-adenosine (m^{6}t^{6}A), 2-methylthio-N6-threonylcarbamoyl-adenosine (ms^{2}g^{6}A), N6,N6-dimethyl-adenosine (m$^6{}_2$A), N6-hydroxynorvalylcarbamoyl-adenosine (hn^{6}A), 2-methylthio-N6-hydroxynorvalylcarbamoyl-adenosine (ms2hn^{6}A), N6-acetyl-adenosine (ac^{6}A), 7-methyl-adenine, 2-methylthio-adenine, 2-methoxy-adenine, α-thio-adenosine, 2'-O-methyl-adenosine (Am), N^6,2'-O-dimethyl-adenosine (m^{6}Am), N^6-Methyl-2'-deoxyadenosine, N6,N6,2'-O-trimethyl-adenosine (m$^6{}_2$Am), 1,2'-O-dimethyl-adenosine (m^{1}Am), 2'-O-ribosyladenosine (phosphate) (Ar(p)), 2-amino-N6-methyl-purine, 1-thio-adenosine, 8-azido-adenosine, 2'-F-ara-adenosine, 2'-F-adenosine, 2'-OH-ara-adenosine, and N6-(19-amino-pentaoxanonadecyl)-adenosine.

Guanine

In some embodiments, the modified nucleobase is a modified guanine. Exemplary nucleobases and nucleosides having a modified guanine include without limitation inosine (I), 1-methyl-inosine (m^{1}I), wyosine (imG), methylwyosine (mimG), 4-demethyl-wyosine (imG-14), isowyosine (imG2), wybutosine (yW), peroxywybutosine (o$_2$yW), hydroxywybutosine (OHyW), undermodified hydroxywybutosine (OHyW*), 7-deaza-guanosine, queuosine (Q), epoxyqueuosine (oQ), galactosyl-queuosine (galQ), mannosyl-queuosine (manQ), 7-cyano-7-deaza-guanosine (preQ$_0$), 7-aminomethyl-7-deaza-guanosine (preQ$_1$), archaeosine (G$^+$), 7-deaza-8-aza-guanosine, 6-thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine (m^{7}G), 6-thio-7-methyl-guanosine, 7-methyl-inosine, 6-methoxy-guanosine, 1-methyl-guanosine (m'G), N2-methyl-guanosine (m^{2}G), N2,N2-dimethyl-guanosine (m$^2{}_2$G), N2,7-dimethyl-guanosine (m^2,7G), N2, N2,7-dimethyl-guanosine (m^2,2,7G), 8-oxo-guanosine, 7-methyl-8-oxo-guanosine, 1-meth thio-guanosine, N2-methyl-6-thio-guanosine, N2,N2-dimethyl-6-thio-guanosine, α-thio-guanosine, 2'-O-methyl-guanosine (Gm), N2-methyl-2'-O-methyl-guanosine (m^{2}Gm), N2,N2-dimethyl-2'-O-methyl-guanosine (m$^2{}_2$Gm), 1-methyl-2'-O-methyl-guanosine (m'Gm), N2,7-dimethyl-2'-O-methyl-guanosine (m^2,7Gm), 2'-O-methyl-inosine (Im), 1,2'-O-dimethyl-inosine (m'Im), O^6-phenyl-2'-deoxyinosine, 2'-O-ribosylguanosine (phosphate) (Gr(p)), 1-thio-guanosine, O^6-methyl-guanosine, O^6-Methyl-2'-deoxyguanosine, 2'-F-ara-guanosine, and 2'-F-guanosine.

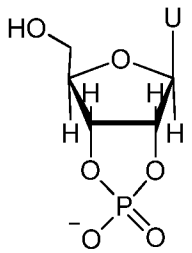
Modified gRNAs

In some embodiments, the modified nucleic acids can be modified gRNAs. In some embodiments, gRNAs can be modified at the 3' end. In this embodiment, the gRNAs can be modified at the 3' terminal U ribose. For example, the two terminal hydroxyl groups of the U ribose can be oxidized to aldehyde groups and a concomitant opening of the ribose ring to afford a modified nucleoside as sown below:



wherein "U" can be an unmodified or modified uridine.

In another embodiment, the 3' terminal U can be modified with a 2'3' cyclic phosphate as shown below:



wherein "U" can be an unmodified or modified uridine.

In some embodiments, the gRNA molecules may contain 3' nucleotides which can be stabilized against degradation, e.g., by incorporating one or more of the modified nucleotides described herein. In this embodiment, e.g., uridines can be replaced with modified uridines, e.g., 5-(2-amino)propyl uridine, and 5-bromo uridine, or with any of the modified uridines described herein; adenosines and guanosines can be replaced with modified adenosines and guanosines, e.g., with modifications at the 8-position, e.g., 8-bromo guanosine, or with any of the modified adenosines or guanosines described herein. In some embodiments, deaza nucleotides, e.g., 7-deaza-adenosine, can be incorporated into the gRNA. In some embodiments, O- and N-alkylated nucleotides, e.g., N6-methyl andenosine, can be incorporated into the gRNA. In some embodiments, sugar-modified ribonucleotides can be incorporated, e.g., wherein the 2' OH-group is replaced by a group selected from H, -OR, -R (wherein R can be, e.g., alkyl, cycloalkyl, aryl, aralkyl, heteroaryl or sugar), halo, -SH, -SR (wherein R can be, e.g., alkyl, cycloalkyl, aryl, aralkyl, heteroaryl or sugar), amino (wherein amino can be, e.g., NH_2; alkylamino, dialkylamino,

heterocyclyl, arylamino, diarylamino, heteroarylamino, diheteroarylamino, or amino acid); or cyano (-CN). In some embodiments, the phosphate backbone can be modified as described herein, e.g., with a phosphothioate group. In some embodiments, the nucleotides in the overhang region of the gRNA can each independently be a modified or unmodified nucleotide including, but not limited to 2'-sugar modified, such as, 2-F 2'-O-methyl, thymidine (T), 2'-O-methoxyethyl-5-methyluridine (Teo), 2'-O-methoxyethyladenosine (Aeo), 2'-O-methoxyethyl-5-methylcytidine (m5Ceo), and any combinations thereof.

In an embodiment, one or more or all of the nucleotides in single stranded RNA molecule, e.g., a gRNA molecule, are deoxynucleotides.

miRNA binding sites

microRNAs (or miRNAs) are naturally occurring cellular 19-25 nucleotide long noncoding RNAs. They bind to nucleic acid molecules having an appropriate miRNA binding site, e.g., in the 3' UTR of an mRNA, and down-regulate gene expression. While not wishing to be bound by theory, in an embodiment, it is believed that the down regulation is either by reducing nucleic acid molecule stability or by inhibiting translation. An RNA species disclosed herein, e.g., an mRNA encoding Cas9 can comprise an miRNA binding site, e.g., in its 3'UTR. The miRNA binding site can be selected to promote down regulation of expression is a selected cell type. By way of example, the incorporation of a binding site for miR-122, a microRNA abundant in liver, can inhibit the expression of the gene of interest in the liver.

Governing gRNA molecules and the use thereof to limit the activity of a Cas9 system

Methods and compositions that use, or include, a nucleic acid, e.g., DNA, that encodes a Cas9 molecule or a gRNA molecule, can, in addition, use or include a "governing gRNA molecule." The governing gRNA can limit the activity of the other CRISPR/Cas components introduced into a cell or subject. In an embodiment, a gRNA molecule comprises a targeting domain that is complementary to a target domain on a nucleic acid that comprises a sequence that encodes a component of the CRISPR/Cas system that is introduced into a cell or subject. In an embodiment, a governing gRNA molecule comprises a targeting domain that is complementary with a target sequence on: (a) a nucleic acid that encodes a Cas9 molecule; (b) a nucleic acid that encodes a gRNA which comprises a targeting domain that targets the *CEP290* gene (a target gene gRNA); or on more than one nucleic acid that encodes a CRISPR/Cas

component, e.g., both (a) and (b). The governing gRNA molecule can complex with the Cas9 molecule to inactivate a component of the system. In an embodiment, a Cas9 molecule/governing gRNA molecule complex inactivates a nucleic acid that comprises the sequence encoding the Cas9 molecule. In an embodiment, a Cas9 molecule/governing gRNA molecule complex inactivates the nucleic acid that comprises the sequence encoding a target gene gRNA molecule. In an embodiment, a Cas9 molecule/governing gRNA molecule complex places temporal, level of expression, or other limits, on activity of the Cas9 molecule/target gene gRNA molecule complex. In an embodiment, a Cas9 molecule/governing gRNA molecule complex reduces off-target or other unwanted activity. In an embodiment, a governing gRNA molecule targets the coding sequence, or a control region, e.g., a promoter, for the CRISPR/Cas system component to be negatively regulated. For example, a governing gRNA can target the coding sequence for a Cas9 molecule, or a control region, e.g., a promoter, that regulates the expression of the Cas9 molecule coding sequence, or a sequence disposed between the two. In an embodiment, a governing gRNA molecule targets the coding sequence, or a control region, e.g., a promoter, for a target gene gRNA. In an embodiment, a governing gRNA, e.g., a Cas9-targeting or target gene gRNA-targeting, governing gRNA molecule, or a nucleic acid that encodes it, is introduced separately, e.g., later, than is the Cas9 molecule or a nucleic acid that encodes it. For example, a first vector, e.g., a viral vector, e.g., an AAV vector, can introduce nucleic acid encoding a Cas9 molecule and one or more target gene gRNA molecules, and a second vector, e.g., a viral vector, e.g., an AAV vector, can introduce nucleic acid encoding a governing gRNA molecule, e.g., a Cas9-targeting or target gene gRNA targeting, gRNA molecule. In an embodiment, the second vector can be introduced after the first. In other embodiments, a governing gRNA molecule, e.g., a Cas9-targeting or target gene gRNA targeting, governing gRNA molecule, or a nucleic acid that encodes it, can be introduced together, e.g., at the same time or in the same vector, with the Cas9 molecule or a nucleic acid that encodes it, but, e.g., under transcriptional control elements, e.g., a promoter or an enhancer, that are activated at a later time, e.g., such that after a period of time the transcription of Cas9 is reduced. In an embodiment, the transcriptional control element is activated intrinsically. In an embodiment, the transcriptional element is activated via the introduction of an external trigger.

Typically a nucleic acid sequence encoding a governing gRNA molecule, e.g., a Cas9-targeting gRNA molecule, is under the control of a different control region, e.g., promoter, than

is the component it negatively modulates, e.g., a nucleic acid encoding a Cas9 molecule. In an embodiment, "different control region" refers to simply not being under the control of one control region, e.g., promoter, that is functionally coupled to both controlled sequences. In an embodiment, different refers to "different control region" in kind or type of control region. For example, the sequence encoding a governing gRNA molecule, e.g., a Cas9-targeting gRNA molecule, is under the control of a control region, e.g., a promoter, that has a lower level of expression, or is expressed later than the sequence which encodes is the component it negatively modulates, e.g., a nucleic acid encoding a Cas9 molecule.

By way of example, a sequence that encodes a governing gRNA molecule, e.g., a Cas9-targeting governing gRNA molecule, can be under the control of a control region (e.g., a promoter) described herein, e.g., human U6 small nuclear promoter, or human H1 promoter. In an embodiment, a sequence that encodes the component it negatively regulates, e.g., a nucleic acid encoding a Cas9 molecule, can be under the control of a control region (e.g., a promoter) described herein, e.g., CMV, EF-1a, MSCV, PGK, CAG control promoters.

Examples

The following Examples are merely illustrative and are not intended to limit the scope or content of the invention in any way.

Example 1: Cloning and Initial Screening of gRNAs

The suitability of candidate gRNAs can be evaluated as described in this example. Although described for a chimeric gRNA, the approach can also be used to evaluate modular gRNAs.

Cloning gRNAs into plasmid vector

For each gRNA, a pair of overlapping oligonucleotides is designed and obtained. Oligonucleotides are annealed and ligated into a digested vector backbone containing an upstream U6 promoter and the remaining sequence of a long chimeric gRNA. Plasmid is sequence-verified and prepped to generate sufficient amounts of transfection-quality DNA. Alternate promoters maybe used to drive in vivo transcription (e.g. H1 promoter) or for in vitro transcription (eg. T7 promoter).

Cloning gRNAs in linear dsDNA molecule (STITCHR)

For each gRNA, a single oligonucleotide is designed and obtained. The U6 promoter and the gRNA scaffold (e.g. including everything except the targeting domain, e.g., including sequences derived from the crRNA and tracrRNA, e.g., including a first complementarity domain; a linking domain; a second complementarity domain; a proximal domain; and a tail domain) are separately PCR amplified and purified as dsDNA molecules. The gRNA-specific oligonucleotide is used in a PCR reaction to stitch together the U6 and the gRNA scaffold, linked by the targeting domain specified in the oligonucleotide. Resulting dsDNA molecule (STITCHR product) is purified for transfection. Alternate promoters may be used to drive in vivo transcription (e.g., H1 promoter) or for in vitro transcription (e.g., T7 promoter). Any gRNA scaffold may be used to create gRNAs compatible with Cas9s from any bacterial species.

Initial gRNA Screen

Each gRNA to be tested is transfected, along with a plasmid expressing Cas9 and a small amount of a GFP-expressing plasmid into human cells. In preliminary experiments, these cells can be immortalized human cell lines such as 293T, K562 or U2OS. Alternatively, primary human cells may be used. In this case, cells may be relevant to the eventual therapeutic cell target (for example, photoreceptor cells). The use of primary cells similar to the potential therapeutic target cell population may provide important information on gene targeting rates in the context of endogenous chromatin and gene expression.

Transfection may be performed using lipid transfection (such as Lipofectamine or Fugene) or by electroporation. Following transfection, GFP expression can be determined either by fluorescence microscopy or by flow cytometry to confirm consistent and high levels of transfection. These preliminary transfections can comprise different gRNAs and different targeting approaches (17-mers, 20-mers, nuclease, dual-nickase, etc.) to determine which gRNAs/combinations of gRNAs give the greatest activity.

Efficiency of cleavage with each gRNA may be assessed by measuring NHEJ-induced indel formation at the target locus by a T7E1-type assay or by sequencing. Alternatively, other mismatch-sensitive enzymes, such as CelI/Surveyor nuclease, may also be used.

For the T7E1 assay, PCR amplicons are approximately 500-700bp with the intended cut site placed asymmetrically in the amplicon. Following amplification, purification and size-verification of PCR products, DNA is denatured and re-hybridized by heating to 95°C and then

slowly cooling. Hybridized PCR products are then digested with T7 Endonuclease I (or other mismatch-sensitive enzyme) which recognizes and cleaves non-perfectly matched DNA. If indels are present in the original template DNA, when the amplicons are denatured and re-annealed, this results in the hybridization of DNA strands harboring different indels and therefore lead to double-stranded DNA that is not perfectly matched. Digestion products may be visualized by gel electrophoresis or by capillary electrophoresis. The fraction of DNA that is cleaved (density of cleavage products divided by the density of cleaved and uncleaved) may be used to estimate a percent NHEJ using the following equation: %NHEJ = (1-(1-fraction cleaved)$^{1/2}$). The T7E1 assay is sensitive down to about 2-5% NHEJ.

Sequencing may be used instead of, or in addition to, the T7E1 assay. For Sanger sequencing, purified PCR amplicons are cloned into a plasmid backbone, transformed, miniprepped and sequenced with a single primer. For large sequencing numbers, Sanger sequencing may be used for determining the exact nature of indels after determining the NHEJ rate by T7E1.

Sequencing may also be performed using next generation sequencing techniques. When using next generation sequencing, amplicons may be 300-500bp with the intended cut site placed asymmetrically. Following PCR, next generation sequencing adapters and barcodes (for example Illumina multiplex adapters and indexes) may be added to the ends of the amplicon, e.g., for use in high throughput sequencing (for example on an Illumina MiSeq). This method allows for detection of very low NHEJ rates.

Example 2: Assessment of Gene Targeting by NHEJ

The gRNAs that induce the greatest levels of NHEJ in initial tests can be selected for further evaluation of gene targeting efficiency. For example, cells may be derived from disease subjects, relevant cell lines, and/or animal models and, therefore, harbor the relevant mutation.

Following transfection (usually 2-3 days post-transfection,) genomic DNA may be isolated from a bulk population of transfected cells and PCR may be used to amplify the target region. Following PCR, gene targeting efficiency to generate the desired mutations (either knockout of a target gene or removal of a target sequence motif) may be determined by sequencing. For Sanger sequencing, PCR amplicons may be 500-700 bp long. For next generation sequencing, PCR amplicons may be 300-500 bp long. If the goal is to knockout gene

function, sequencing may be used to assess what percent of alleles have undergone NHEJ-induced indels that result in a frameshift or large deletion or insertion that would be expected to destroy gene function. If the goal is to remove a specific sequence motif, sequencing may be used to assess what percent of alleles have undergone NHEJ-induced deletions that span this sequence.

Example 3: Assessment of Activity of Individual gRNAs Targeting CEP290

Guide RNA were identified using a custom guide RNA design software based on the public tool cas-offinder (Bae et al. Bioinformatics. 2014; 30(10): 1473-1475). Each gRNA to be tested was generated as a STITCHR product and co-transfected with a plasmid expressing either *S. aureus* Cas9 (pAF003) or *S. pyogenes* Cas9 (pJDS246) into either HEK293 cells or primary fibroblasts derived from and LCA10 patient harboring homozygous IVS26 c.2991+1655A to G mutations (hereafter referred to as IVS26 fibroblasts). The pAF003 plasmid encodes the *S. aureus* Cas9, with N-terminal and C-terminal nuclear localization signals (NLS) and a C-terminal triple flag tag, driven by a CMV promoter. The pJDS246 plasmid encodes the *S. pyogenes* Cas9, with a C-terminal nuclear localization signal (NLS) and a C-terminal triple flag tag, driven by a CMV promoter. gRNA and Cas9-encoding DNA was introduced into cells by either Mirus TransIT-293 transfection reagent (for 293 cells) or by Amaxa nucleofection (for IVS26 fibroblasts). Nucleofection was optimized for transfection of IVS26 fibroblasts using solution P2 and various pulse codes and assaying for highest levels of gene editing and cell viability. Transfection efficiency in both cell types was assessed by transfecting with GFP and assaying expression by fluorescent microscopy. Three to seven days post-transfection, genomic DNA was isolated from bulk populations of transfected cells and the region of the *CEP290* locus surrounding the target site was PCR amplified. PCR amplicons were then cloned into a plasmid backbone using the Zero-Blunt TOPO cloning kit (Lifetechnologies) and transformed into chemically competent Top10 cells. Bacterial colonies were then cultured and plasmid DNA was isolated and sequenced. Sequencing of PCR products allowed for the detection and quantification of targeted insertion and deletion (indel) events at the target site. **Fig. 11A** and **11B** show the rates of indels induced by various gRNAs at the *CEP290* locus. **Fig. 11A** shows gene editing (% indels) as assessed by sequencing for *S. pyogenes* and *S. aureus* gRNAs when co-expressed with Cas9 in patient-derived IVS26 primary fibroblasts. **Fig. 11B** shows gene

editing (% indels) as assessed by sequencing for *S. aureus* gRNAs when co-expressed with Cas9 in HEK293 cells.

Example 4: Detection of gRNA Pair-Induced Deletions by PCR

To assess the ability of a pair of gRNAs to induce a genomic deletion (in which the sequence between the two cut sites is removed), PCR was performed across the predicted deletion. Pairs of gRNAs (encoded as STITCHR products) were co-transfected with pAF003 into IVS26 fibroblasts. Genomic DNA was isolated from transfected cells and PCR was performed to amplify a segment of the *CEP290* locus spanning the two predicted cut sites. PCR was run on a QIAxcel capillary electrophoresis machine. The predicted amplicon on a wildtype allele is 1816 bps. Assuming that cleavage occurs within the gRNA target region, amplicon sizes for alleles having undergone the deletion event were calculated and the presence of this smaller band indicates that the desired genomic deletion event has occurred (**Table 22**).

Table 22

	Left gRNA	Right gRNA	Deletion Size	Amplicon with deletion	Deletion amplicon detected?
1	CEP290-367	CEP290-16	590	1226	no
2	CEP290-367	CEP290-203	688	1128	no
3	CEP290-367	CEP290-132	815	1001	no
4	CEP290-367	CEP290-139	1265	551	no
5	CEP290-312	CEP290-11	790	1026	yes
6	CEP290-312	CEP290-252	973	843	no
7	CEP290-312	CEP290-64	976	840	yes
8	CEP290-312	CEP290-230	1409	407	yes
9	CEP290-12	CEP290-11	19	1797	no
10	CEP290-12	CEP290-252	202	1614	no
11	CEP290-12	CEP290-64	205	1611	no
12	CEP290-12	CEP290-230	638	1178	no
13	CEP290-17	CEP290-16	19	1797	no
14	CEP290-17	CEP290-203	117	1699	no
15	CEP290-17	CEP290-132	244	1572	no
16	CEP290-17	CEP290-139	693	1123	no
17	CEP290-374	CEP290-16	799	1017	no
18	CEP290-374	CEP290-203	897	919	no
19	CEP290-374	CEP290-132	1024	792	no
20	CEP290-374	CEP290-139	1473	343	no
21	CEP290-368	CEP290-16	854	962	no
22	CEP290-368	CEP290-203	952	864	no

23	CEP290-368	CEP290-132	1079	737	no
24	CEP290-368	CEP290-139	1528	288	no
25	CEP290-323	CEP290-11	990	826	yes
26	CEP290-323	CEP290-252	1173	643	no
27	CEP290-323	CEP290-64	1176	640	yes
28	CEP290-323	CEP290-230	1609	207	yes
29	Cas9 only			wt amplicon = 1816	no
30	GFP only			wt amplicon = 1816	no
31	no DNA PCR neg ctrl				

Example 5: Gene Expression Analysis of CEP290

Targeted deletion of a region containing the IVS26 splice mutation is predicted to correct the splicing defect and restore expression of the normal wild-type *CEP290* allele. To quantify expression of the wild-type and mutant (containing additional cryptic splice mutation) alleles, TaqMan assays were designed. Multiple assays were tested for each RNA species and a single wt and single mutant assay were selected. The assay for the wild-type allele contains a forward primer that anneals in exon 26, a reverse primer that anneals in exon 27 and a TaqMan probe that spans the exon26-exon-27 junction. The assay for the mutant allele contains a forward primer that anneals in exon 26, a reverse primer that anneals in the cryptic exon and a TaqMan probe that spans the exon26-cryptic exon junction. A TaqMan assay designed to beta-actin was used as a control. Total RNA was isolated from IVS26 cells transfected with pairs of gRNAs and Cas9-expressing plasmid by either Trizol RNA purification (Ambion), Agencourt RNAdvance (Beckman Coulter) or direct cells-to-Ct lysis (Lifetechnologies). Reverse transcription to generate cDNA was performed and cDNA was used as a template for qRT-PCR using selected taqman assays on a BioRad real time PCR machine. Relative gene expression was calculated by $\Delta\Delta$Ct, relative to beta-actin control and GFP-only sample. Increases in expression of wt allele and decreases in expression of mutant allele relative to GFP-only control indicate corrected splicing due to gene targeting. **Figs. 12A-12B** show initial qRT-PCR data for pairs of gRNAs that had shown activity as either individual gRNAs (measured as described in Example 3) or as pairs (measured as described in Example 4). Pairs of gRNAs that showed the desired gene expression changes were repeated in replicate experiments and the cumulative qRT-PCR data is shown in **Fig. 13** (error bars represent standard error of the mean calculated from 2 to 6 biological replicates per sample).

Example 6: Quantification of Genomic Deletions by ddPCR

Droplet digital PCR (ddPCR) is a method for performing digital PCR in which a single PCR reaction is fractionated into 20,000 droplets in a water-oil emulsion and PCR amplification occurs separately in individual droplets. PCR conditions are optimized for a concentration of DNA template such that each droplet contains either one or no template molecules. Assays were designed to perform amplification using BioRad EvaGreen Supermix PCR system with all amplicons ranging in size from 250-350 bp. Control assays were designed to amplify segments of the CEP290 gene at least 5 kb away from the IVS26 c.2991+1655A to G mutation. Assays to detect targeted genomic deletion were designed such that amplification of an allele that has undergone deletion will yield a PCR product in the size range of 250-350 bp and amplification will not occur on a wild-type allele due to the increased distance between forward and reverse primers. PCR conditions were optimized on genomic DNA isolated from 293 cells that had been transfected with pairs of gRNAs and Cas9-expressing plasmid. Deletion assays were verified to generate no positive signal on genomic DNA isolated from unmodified IVS26 fibroblasts. Assays were further tested and optimized on genomic DNA isolated from IVS26 fibroblasts that had been transfected with pairs of gRNAs and Cas9-encoding plasmid. Of the three assays tested for each of two deletions (CEP290-323 and CEP290-11; and CEP290-323 and CEP290-64) and the 4 control assays tested, a single assay was selected for each deletion and a control based on quality data and replicability in the ddPCR assay. **Fig. 14** shows deletion rates on three biological replicates calculated by taking the number of positive droplets for the deletion assay and dividing by the number of positive droplets for the control assay.

Example 7: Cloning AAV Expression Vectors

Cloning saCas9 into an AAV expression vector

The pAF003 plasmid encodes the CMV-driven *S. aureus* Cas9 (saCas9), with N-terminal and C-terminal nuclear localization signals (NLS) and a C-terminal triple flag tag, followed by a bovine growth hormone poly(A) tail (bGH polyA). BGH polyA tail was substituted with a 60-bp minimal polyA tail to obtain pAF003-minimal-pA. The CMV-driven NLS-saCas9-NLS-3xFlag with the minimal polyA tail was amplified with PCR and subcloned into pTR-UF11 plasmid (ATCC #MBA-331) with KpnI and SphI sites to obtain the pSS3 (pTR-CMV-saCas9-minimal-pA) vector. The CMV promoter sequence can be substituted with EFS promoter (pSS10 vector),

or tissue-specific promoters (**Table 19**, e.g. photo-receptor-specific promoters, e.g. Human GRK1, CRX, NRL, RCVRN promoters, etc.) using SpeI and NotI sites.

Constructing the All-in-One AAV expression vector with one gRNA sequence

For each individual gRNA sequence, a STITCHR product with a U6 promoter, gRNA, and the gRNA scaffold was obtained by PCR with an oligonucleotide encoding the gRNA sequence. The STITCHR product with one dsDNA molecule of U6-driven gRNA and scaffold was subcloned into pSS3 or pSS10 vectors using KpnI sites flanking the STITCHR product and downstream of the left Inverted Terminal Repeat (ITR) in the AAV vectors. The orientation of the U6-gRNA-scaffold insertion into pSS3 or pSS10 was determined by Sanger sequencing. Alternate promoters may be used to drive gRNA expression (e.g. H1 promoter, 7SK promoter). Any gRNA scaffold sequences compatible with Cas variants from other bacterial species could be incorporated into STITCHR products and the AAV expression vector therein.

Cloning two gRNA into an AAV expression vector

For each pair of gRNA sequences, two ssDNA oligonucleotides were designed and obtained as the STITCHR primers, i.e. the left STITCHR primer and the right STITCHR primer. Two STITCHR PCR reactions (i.e. the left STITCHR PCR and the right STITCHR PCR) amplified the U6 promoter and the gRNA scaffold with the corresponding STITCHR primer separately. The pSS3 or pSS10 backbone was linearized with KpnI restriction digest. Two dsDNA STITCHR products were purified and subcloned into pSS3 or pSS10 backbone with Gibson Assembly. Due to the unique overlapping sequences upstream and downstream of the STITCHR products, the assembly is unidirectional. The sequences of the constructs were confirmed by Sanger Sequencing. **Table 23** lists the names and compositions of AAV expression vectors constructed, including the names of gRNAs targeting human *CEP290*, the promoter to drive Cas9 expression, and the length of the AAV vector including the Inverted Terminal Repeats (ITRs) from wild type AAV2 geome. Alternative promoters (e.g., H1 promoter or 7SK promoter) or gRNA scaffold sequences compatible with any Cas variants could be adapted into this cloning strategy to obtain the corresponding All-in-One AAV expression vectors with two gRNA sequences.

Table 23. Components of AAV expression vectors

Name	Left gRNA	Right gRNA	Promoter of saCas9	Length including ITRs

pSS10	NA	NA	EFS	4100
pSS11	CEP290-64	CEP290-323	EFS	4853
pSS15	CEP290-64	NA	EFS	4491
pSS17	CEP290-323	NA	EFS	4491
pSS30	CEP290-323	CEP290-64	EFS	4862
pSS31	CEP290-323	CEP290-11	EFS	4862
pSS32	CEP290-490	CEP290-502	EFS	4858
pSS33	CEP290-490	CEP290-496	EFS	4858
pSS34	CEP290-490	CEP290-504	EFS	4857
pSS35	CEP290-492	CEP290-502	EFS	4858
pSS36	CEP290-492	CEP290-504	EFS	4857
pSS3	NA	NA	CMV	4454
pSS8	CEP290-64	CEP290-323	CMV	5207
pSS47	CEP290-323	CEP290-64	CMV	5216
pSS48	CEP290-323	CEP290-11	CMV	5216
pSS49	CEP290-490	CEP290-502	CMV	5212
pSS50	CEP290-490	CEP290-496	CMV	5212
pSS51	CEP290-490	CEP290-504	CMV	5211
pSS52	CEP290-492	CEP290-502	CMV	5212
pSS53	CEP290-492	CEP290-504	CMV	5211
pSS23	NA	NA	hGRK1	4140
pSS24	NA	NA	hCRX	3961
pSS25	NA	NA	hNRL	4129
pSS26	NA	NA	hRCVRN	4083

Example 8: Assessment of the Functions of All-in-One AAV Expression Vectors

Each individual AAV expression vectors were transfected into 293T cells with TransIT-293 (Mirus, Inc.) to test their function before being packaged into AAV viral vectors. 293T cells were transfected with the same amount of plasmid and harvested at the same time points. SaCas9 protein expression was assessed by western blotting with primary antibody probing for the triple Flag tag at the C-terminus of saCas9, while loading control was demonstrated by αTubulin expression. Deletion events at IVS26 mutation could be determined by PCR amplification followed by Sanger sequencing or ddPCR. The results are shown in **Fig. 15**.

Example 9: Production, purification and titering of recombinant AAV2 vectors

Prior to packaging into AAV viral vectors, all AAV expression vector (plasmids) underwent primer walk with Sanger sequencing and function analysis. In recombinant AAV (rAAV), two ITRs flanking the transgene cassettes are the only cis-acting elements from the

wild-type AAV. They are critical for packaging intact rAAVs and genome-release for rAAV vectors during transduction. All AAV expression vectors were restriction digested with SmaI or XmaI to ensure the presence of two intact ITRs.

rAAV2 vectors were produced with "Triple Transfection Protocol": (1) pSS vectors with ITRs and transgene cassetts; (2) pHelper plasmid with E2A, E4, VA genes from Adenovirus; (3) pAAV-RC2 plasmid with *Rep* and *Cap* genes from AAV2. These three plasmids were mixed at a mass ratio of 3:6:5 and transfected into HEK293 with polymer or lipid-based transfection reagent (e.g. PEI, PEI max, Lipofectamine, TransIT-293, etc.). 60-72 hours post-transfection, HEK293 cells were harvested and sonicated to release viral vectors. Cell lysates underwent CsCl ultracentrifuge to purify and concentrate the viral vectors. Additional purification procedures were performed to obtain higher purity for biophysical assays, including another round of CsCl ultracentrifuge, or sucrose gradient ultracentrifuge, or affinity chromatography. Viral vectors were dialyzed with 1xDPBS twice before being aliquoted for storage in -80^{o}C. Viral preps can be tittered with Dot-Blot protocol or/and quantitative PCR with probes annealing to sequences on the transgenes. PCR primer sequences are: AACATGCTACGCAGAGAGGGAGTGG (SEQ ID NO: 482) (ITR-Titer-fwd) and CATGAGACAAGGAACCCCTAGTGATGGAG (SEQ ID NO: 483) (ITR-Titer-rev). Reference AAV preps were obtained from the Vector Core at University of North Carolina-Chapel Hill as standards. To confirm the presence of three non-structural viral proteins composing the AAV capsid, viral preps were denatured and probed with anti-AAV VP1/VP2/VP3 monoclonal antibody B1 (American Research Products, Inc. Cat #03-65158) on western blots. The results are shown in **Fig. 16**.

Example 10: rAAV-mediated CEP290 modification *in vitro*

293T were transduced with rAAV2 vectors expressing saCas9 with or without gRNA sequences to demonstrate the deletion events near the IVS26 splicing mutant. 293T cells were transduced with rAAV2 viral vectors at an MOI of 1,000 viral genome (vg)/cell or 10,000 vg/cell and harvested at three to seven days post transduction. Western blotting with the primary antibody for Flag (anti-Flag, M2, Sigma-Aldrich) showed that the presence of U6-gRNA-scaffold does not interfere with saCas9 expression. Genomic DNA from 293T was isolated with the Agencourt DNAdvance Kit (Beckman Coulter). Regions including the deletions were PCR amplified from genomic DNA isolated, and analyzed on the QIAxcel capillary electrophoresis

machine. Amplicons smaller than the full-length predicted PCR products represent the deletion events in 293T cells. The PCR results are shown in **Fig. 17**. To further understand the nature of these deletion events, PCR products were cloned into Zero-Blunt TOPO Cloning Kit (Life Technologies) and transformed into chemically competent Top10 cells. Bacterial colonies were then cultured and sequenced using Sanger sequencing. Sequence results were aligned with the wt *CEP290* locus for analysis.

Incorporation by Reference

All publications, patents, and patent applications mentioned herein are hereby incorporated by reference in their entirety as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference. In case of conflict, the present application, including any definitions herein, will control.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

Other embodiments are within the following claims.

What is claimed is:

1. A gRNA molecule comprising a targeting domain which is complementary with a target domain from the *CEP290* gene.
2. The gRNA molecule of claim 1, wherein said targeting domain is configured to provide a cleavage event selected from a double strand break and a single strand break, within 10 nucleotides of a LCA10 target position.
3. The gRNA molecule of claim 1, wherein said targeting domain comprises a sequence that is the same as, or differs by no more than 3 nucleotides from, a targeting domain sequence from Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B, and Table 10.
4. The gRNA molecule of claim 1, wherein said targeting domain comprises a sequence that is the same as, or differs by no more than 3 nucleotides from, a targeting domain sequence from Tables 1A-1D.
5. The gRNA molecule of claim 1, wherein said targeting domain comprises a sequence that is the same as, or differs by no more than 3 nucleotides from, a targeting domain sequence from Tables 2A-2C.
6. The gRNA molecule of claim 1, wherein said targeting domain comprises a sequence that is the same as, or differs by no more than 3 nucleotides from, a targeting domain sequence from Tables 3A-3D.
7. The gRNA molecule of claim 1, wherein said targeting domain comprises a sequence that is the same as, or differs by no more than 3 nucleotides from, a targeting domain sequence from Tables 4A-4D.
8. The gRNA molecule of claim 1, wherein said targeting domain comprises a sequence that is the same as, or differs by no more than 3 nucleotides from, a targeting domain sequence from Tables 5A-5B.
9. The gRNA molecule of claim 1, wherein said targeting domain comprises a sequence that is the same as, or differs by no more than 3 nucleotides from, a targeting domain sequence from Tables 6A-6D.

10. The gRNA molecule of claim 1, wherein said targeting domain comprises a sequence that is the same as, or differs by no more than 3 nucleotides from, a targeting domain sequence from Tables 7A-7D.
11. The gRNA molecule of claim 1, wherein said targeting domain comprises a sequence that is the same as, or differs by no more than 3 nucleotides from, a targeting domain sequence from Tables 8A-8E.
12. The gRNA molecule of claim 1, wherein said targeting domain comprises a sequence that is the same as, or differs by no more than 3 nucleotides from, a targeting domain sequence from Tables 9A-9B.
13. The gRNA molecule of claim 1, wherein said targeting domain comprises a sequence that is the same as, or differs by no more than 3 nucleotides from, a targeting domain sequence from Table 10, e.g., GACACTGCCAATAGGGATAGGT; GTCAAAAGCTACCGGTTACCTG; GTTCTGTCCTCAGTAAAAGGTA; GAATAGTTTGTTCTGGGTAC; GAGAAAGGGATGGGCACTTA; GATGCAGAACTAGTGTAGAC; GTCACATGGGAGTCACAGGG; or GAGTATCTCCTGTTTGGCA.
14. The gRNA molecule of claim 1, wherein said targeting domain is selected from: Tables 1A-1D.
15. The gRNA molecule of claim 1, wherein said targeting domain is selected from: Tables 2A-2C.
16. The gRNA molecule of claim 1, wherein said targeting domain is selected from: Tables 3A-3D.
17. The gRNA molecule of claim 1, wherein said targeting domain is selected from: Tables 4A-4D.
18. The gRNA molecule of claim 1, wherein said targeting domain is selected from: Tables 5A-5B.
19. The gRNA molecule of claim 1, wherein said targeting domain is selected from those in Table 1A.
20. The gRNA molecule of claim 1, wherein said targeting domain is GAGAUACUCACAAUUACAAC.

21. The gRNA molecule of claim 1, wherein said targeting domain is GAUACUCACAAUUACAACUG.
22. The gRNA molecule of claim 1, wherein said targeting domain is selected from those in Table 2A.
23. The gRNA molecule of claim 1, wherein said targeting domain is GAGAUACUCACAAUUACAAC.
24. The gRNA molecule of claim 1, wherein said targeting domain is GAUACUCACAAUUACAA.
25. The gRNA molecule of claim 1, wherein said targeting domain is selected from those in Table 3A.
26. The gRNA molecule of claim 1, wherein said targeting domain is GCUACCGGUUACCUGAA.
27. The gRNA molecule of claim 1, wherein said targeting domain is GCAGAACUAGUGUAGAC.
28. The gRNA molecule of claim 1, wherein said targeting domain is GUUGAGUAUCUCCUGUU.
29. The gRNA molecule of claim 1, wherein said targeting domain is GAUGCAGAACUAGUGUAGAC.
30. The gRNA molecule of claim 1, wherein said targeting domain is GCUUGAACUCUGUGCCAAAC.
31. The gRNA molecule of claim 1, wherein said targeting domain is selected from those in Table 4A.
32. The gRNA molecule of claim 1, wherein said targeting domain is GAAUCCUGAAAGCUACU.
33. The gRNA molecule of claim 1, wherein said targeting domain is selected from those in Table 5A.
34. The gRNA molecule of claim 1, wherein said targeting domain is GAGUUCAAGCUAAUACAUGA.
35. The gRNA molecule of claim 1, wherein said targeting domain is GUUGUUCUGAGUAGCUU.

36. The gRNA molecule of claim 1, wherein said targeting domain is GGCAAAAGCAGCAGAAAGCA.

37. The gRNA molecule of claim 1, wherein said targeting domain is GUUGUUCUGAGUAGCUU.

38. The gRNA molecule of claim 1, wherein said targeting domain is GGCAAAAGCAGCAGAAAGCA.

39. The gRNA molecule of claim 1, wherein said targeting domain is selected from those in Table 6A.

40. The gRNA molecule of claim 1, wherein said targeting domain is GCACCUGGCCCCAGUUGUAAUU.

41. The gRNA molecule of claim 1, wherein said targeting domain is selected from those in Table 7A.

42. The gRNA molecule of claim 1, wherein said targeting domain is selected from those in Table 8A.

43. The gRNA molecule of claim 1, wherein said targeting domain is selected from those in Table 9A.

44. The gRNA molecule of claim 1, wherein said targeting domain is GGCAAAAGCAGCAGAAAGCA.

45. The gRNA molecule of claim 1, wherein said targeting domain is GUGGCUGAAUGACUUCU.

46. The gRNA molecule of claim 1, wherein said targeting domain is GUUGUUCUGAGUAGCUU.

47. The gRNA molecule of claim 1, wherein said targeting domain is GACUAGAGGUCACGAAA.

48. The gRNA molecule of claim 1, wherein said targeting domain is GAGUUCAAGCUAAUACAUGA.

49. The gRNA molecule of claim 1, wherein said targeting domain is selected from those in Table 10, e.g., GACACTGCCAATAGGGATAGGT; GTCAAAAGCTACCGGTTACCTG; GTTCTGTCCTCAGTAAAAGGTA; GAATAGTTTGTTCTGGGTAC; GAGAAAGGGATGGGCACTTA;

GATGCAGAACTAGTGTAGAC; GTCACATGGGAGTCACAGGG; or GAGTATCTCCTGTTTGGCA.

50. The gRNA molecule of any of claims 1-49, wherein said gRNA is a modular gRNA molecule.

51. The gRNA molecule of any of claims1-49, wherein said gRNA is a chimeric gRNA molecule.

52. The gRNA molecule of any of claims 1-51, wherein said targeting domain is 16 nucleotides or more in length.

53. The gRNA molecule of any of claims 1-51, wherein said targeting domain is 16 nucleotides in length.

54. The gRNA molecule of any of claims 1-51, wherein said targeting domain is 17 nucleotides in length.

55. The gRNA molecule of any of claims 1-51, wherein said targeting domain is 18 nucleotides in length.

56. The gRNA molecule of any of claims 1-51, wherein said targeting domain is 19 nucleotides in length.

57. The gRNA molecule of any of claims 1-51, wherein said targeting domain is 20 nucleotides in length.

58. The gRNA molecule of any of claims 1-51, wherein said targeting domain is 21 nucleotides in length.

59. The gRNA molecule of any of claims 1-51, wherein said targeting domain is 22 nucleotides in length.

60. The gRNA molecule of any of claims 1-51, wherein said targeting domain is 23 nucleotides in length.

61. The gRNA molecule of any of claims 1-51, wherein said targeting domain is 24 nucleotides in length.

62. The gRNA molecule of any of claims 1-51, wherein said targeting domain is 25 nucleotides in length.

63. The gRNA molecule of any of claims 1-51, wherein said targeting domain is 26 nucleotides in length.

64. The gRNA molecule of any of claims 1-63, comprising from 5' to 3':

a targeting domain;

a first complementarity domain;

a linking domain;

a second complementarity domain;

a proximal domain; and

a tail domain.

65. The gRNA molecule of any of claims 1-64, comprising:

a linking domain of no more than 25 nucleotides in length;

a proximal and tail domain, that taken together, are at least 20 nucleotides in length;

a targeting domain of 17 or 18 nucleotides in length.

66. The gRNA molecule of any of claims 1-64, comprising:

a linking domain of no more than 25 nucleotides in length;

a proximal and tail domain, that taken together, are at least 30 nucleotides in length;

a targeting domain of 17 or 18 nucleotides in length.

67. The gRNA molecule of any of claims 1-64, comprising:

a linking domain of no more than 25 nucleotides in length;

a proximal and tail domain, that taken together, are at least 30 nucleotides in length;

a targeting domain of 17 nucleotides in length.

68. The gRNA molecule of any of claims 1-64, comprising:

a linking domain of no more than 25 nucleotides in length;

a proximal and tail domain, that taken together, are at least 40 nucleotides in length;

a targeting domain of 17 nucleotides in length.

69. A nucleic acid that comprises: (a) sequence that encodes a gRNA molecule comprising a targeting domain that is complementary with a LCA10 target domain in *CEP290* gene.

70. The nucleic acid of claim 69, wherein said gRNA molecule is a gRNA molecule of any of claims 1-68.

71. The nucleic acid of claim 69, wherein said targeting domain is configured to provide a cleavage event selected from a double strand break and a single strand break, within 10 nucleotides of the LCA10 target position.

72. The nucleic acid of claim 69, wherein said targeting domain comprises a sequence that is the same as, or differs by no more than 3 nucleotides from, a targeting domain sequence

from Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B, and Table 10.

73. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Tables 1A-1D.

74. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Tables 2A-2C.

75. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Tables 3A-3D.

76. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Tables 4A-4D.

77. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Tables 5A-5B.

78. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Tables 6A-6D.

79. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Tables 7A-7D.

80. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Tables 8A-8E.

81. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Tables 9A-9B.

82. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Table 10, e.g., GACACTGCCAATAGGGATAGGT; GTCAAAAGCTACCGGTTACCTG; GTTCTGTCCTCAGTAAAAGGTA; GAATAGTTTGTTCTGGGTAC; GAGAAAGGGATGGGCACTTA; GATGCAGAACTAGTGTAGAC; GTCACATGGGAGTCACAGGG; or GAGTATCTCCTGTTTGGCA.

83. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Table 1A.

84. The nucleic acid of claim 69, wherein said targeting domain is: GAGAUACUCACAAUUACAAC.

85. The nucleic acid of claim 69, wherein said targeting domain is: GAUACUCACAAUUACAACUG.

86. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Table 2A.

87. The nucleic acid of claim 69, wherein said targeting domain is: GAGAUACUCACAAUUACAAC.

88. The nucleic acid of claim 69, wherein said targeting domain is: GAUACUCACAAUUACAA

89. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Table 3A.

90. The nucleic acid of claim 69, wherein said targeting domain is: GCUACCGGUUACCUGAA.

91. The nucleic acid of claim 69, wherein said targeting domain is: GCAGAACUAGUGUAGAC.

92. The nucleic acid of claim 69, wherein said targeting domain is: GUUGAGUAUCUCCUGUU.

93. The nucleic acid of claim 69, wherein said targeting domain is: GCUACCGGUUACCUGAA.

94. The nucleic acid of claim 69, wherein said targeting domain is: GAUGCAGAACUAGUGUAGAC.

95. The nucleic acid of claim 69, wherein said targeting domain is: GCUUGAACUCUGUGCCAAAC.

96. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Table 4A.

97. The nucleic acid of claim 69, wherein said targeting domain is: GAAUCCUGAAAGCUACU.

98. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Table 5A.

99. The nucleic acid of claim 69, wherein said targeting domain is: GAGUUCAAGCUAAUACAUGA.

100. The nucleic acid of claim 69, wherein said targeting domain is: GUUGUUCUGAGUAGCUU.

101. The nucleic acid of claim 69, wherein said targeting domain is: GGCAAAAGCAGCAGAAAGCA.

102. The nucleic acid of claim 69, wherein said targeting domain is: GUUGUUCUGAGUAGCUU.

103. The nucleic acid of claim 69, wherein said targeting domain is: GGCAAAAGCAGCAGAAAGCA.

104. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Table 6A.

105. The nucleic acid of claim 69, wherein said targeting domain is GCACCUGGCCCCAGUUGUAAUU.

106. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Table 7A.

107. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Table 8A.

108. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Table 9A.

109. The nucleic acid of claim 69, wherein said targeting domain is GGCAAAAGCAGCAGAAAGCA.

110. The nucleic acid of claim 69, wherein said targeting domain is GUGGCUGAAUGACUUCU.

111. The nucleic acid of claim 69, wherein said targeting domain is GUUGUUCUGAGUAGCUU.

112. The nucleic acid of claim 69, wherein said targeting domain is GACUAGAGGUCACGAAA.

113. The nucleic acid of claim 69, wherein said targeting domain is GAGUUCAAGCUAAUACAUGA.

114. The nucleic acid of claim 69, wherein said targeting domain is selected from those in Table 10, e.g., GACACTGCCAATAGGGATAGGT; GTCAAAAGCTACCGGTTACCTG; GTTCTGTCCTCAGTAAAAGGTA;

GAATAGTTTGTTCTGGGTAC; GAGAAAGGGATGGGCACTTA; GATGCAGAACTAGTGTAGAC; GTCACATGGGAGTCACAGGG; or GAGTATCTCCTGTTTGGCA.

115. The nucleic acid of any of claims 69-114, wherein said gRNA is a modular gRNA molecule.

116. The nucleic acid of any of claims 69-114, wherein said gRNA is a chimeric gRNA molecule.

117. The nucleic acid of any of claims 69-114, wherein said targeting domain is 16 nucleotides or more in length.

118. The nucleic acid of any of claims 69-114, wherein said targeting domain is 16 nucleotides in length.

119. The nucleic acid of any of claims 69-114, wherein said targeting domain is 17 nucleotides in length.

120. The nucleic acid of any of claims 69-114, wherein said targeting domain is 18 nucleotides in length.

121. The nucleic acid of any of claims 69-114, wherein said targeting domain is 19 nucleotides in length.

122. The nucleic acid of any of claims 69-114, wherein said targeting domain is 20 nucleotides in length.

123. The nucleic acid of any of claims 69-114, wherein said targeting domain is 21 nucleotides in length.

124. The nucleic acid of any of claims 69-114, wherein said targeting domain is 22 nucleotides in length.

125. The nucleic acid of any of claims 69-114, wherein said targeting domain is 23 nucleotides in length.

126. The nucleic acid of any of claims 69-114, wherein said targeting domain is 24 nucleotides in length.

127. The nucleic acid of any of claims 69-114, wherein said targeting domain is 25 nucleotides in length.

128. The nucleic acid of any of claims 69-114, wherein said targeting domain is 26 nucleotides in length.

129. The nucleic acid of any of claims 69-128, comprising from 5' to 3':
a targeting domain;
a first complementarity domain;
a linking domain;
a second complementarity domain;
a proximal domain; and
a tail domain.

130. The nucleic acid of any of claims 69-129, comprising:
a linking domain of no more than 25 nucleotides in length;
a proximal and tail domain, that taken together, are at least 20 nucleotides in length;
a targeting domain of 17 or 18 nucleotides in length.

131. The nucleic acid of any of claims 69-129, comprising:
a linking domain of no more than 25 nucleotides in length;
a proximal and tail domain, that taken together, are at least 30 nucleotides in length;
a targeting domain of 17 or 18 nucleotides in length.

132. The nucleic acid of any of claims 69-129, comprising:
a linking domain of no more than 25 nucleotides in length;
a proximal and tail domain, that taken together, are at least 30 nucleotides in length;
a targeting domain of 17 nucleotides in length.

133. The nucleic acid of any of claims 69-129, comprising:
a linking domain of no more than 25 nucleotides in length;
a proximal and tail domain, that taken together, are at least 40 nucleotides in length;
a targeting domain of 17 nucleotides in length.

134. The nucleic acid of any of claims 69-133, further comprising: (b) a sequence that encodes a Cas9 molecule.

135. The nucleic acid of claim 134, wherein said Cas9 molecule comprises a nickase molecule.

136. The nucleic acid of claim 134, wherein said Cas9 molecule forms a double strand break in a target nucleic acid.

137. The nucleic acid of claim 134, wherein said Cas9 molecule forms a single strand break in a target nucleic acid.

138. The nucleic acid of claim 137, wherein said single strand break is formed in the strand of the target nucleic acid to which the targeting domain of said gRNA molecule is complementary.

139. The nucleic acid of claim 137, wherein said single strand break is formed in the strand of the target nucleic acid other than the strand to which to which the targeting domain of said gRNA is complementary.

140. The nucleic acid of claim 134, wherein said Cas9 molecule comprises HNH-like domain cleavage activity but has no, or no significant, N-terminal RuvC-like domain cleavage activity.

141. The nucleic acid of claim 134, wherein said Cas9 molecule is an HNH-like domain nickase.

142. The nucleic acid of claim 134, wherein said Cas9 molecule comprises a mutation at D10.

143. The nucleic acid of claim 134, wherein said Cas9 molecule comprises N-terminal RuvC-like domain cleavage activity but has no, or no significant, HNH-like domain cleavage activity.

144. The nucleic acid of claim 134, wherein said Cas9 molecule is an N-terminal RuvC-like domain nickase.

145. The nucleic acid of claim 134, wherein said Cas9 molecule comprises a mutation at H840.

146. The nucleic acid of claim 134, wherein said Cas9 molecule comprises a mutation at H863.

147. The nucleic acid of any of claims 134-146, further comprising: (c) a sequence that encodes a second gRNA molecule having a targeting domain that is complementary to a second target domain of the *CEP290* gene.

148. The nucleic acid of claim 147, wherein said second gRNA molecule is a gRNA molecule of any of claims 1-68.

149. The nucleic acid of claim 147, wherein said targeting domain of said second gRNA molecule is configured to provide a cleavage event selected from a double strand break and a single strand break, within 10 nucleotides of the LCA10 target position.

150. The nucleic acid of claim 147, wherein said targeting domain of said second gRNA molecule comprises a sequence that is the same as, or differs by no more than 3 nucleotides from, a targeting domain sequence from Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B, and Table 10.

151. The nucleic acid of claim 147, wherein said targeting domain of said second gRNA molecule comprises a sequence that is the same as, or differs by no more than 3 nucleotides from, a targeting domain sequence from Tables 1A-1D.

152. The nucleic acid of claim 147, wherein said targeting domain of said second gRNA molecule comprises a sequence that is the same as, or differs by no more than 3 nucleotides from, a targeting domain sequence from Tables 2A-2C.

153. The nucleic acid of claim 147, wherein said targeting domain of said second gRNA molecule comprises a sequence that is the same as, or differs by no more than 3 nucleotides from, a targeting domain sequence from Tables 3A-3D.

154. The nucleic acid of claim 147, wherein said targeting domain of said second gRNA molecule comprises a sequence that is the same as, or differs by no more than 3 nucleotides from, a targeting domain sequence from Tables 4A-4D.

155. The nucleic acid of claim 147, wherein said targeting domain of said second gRNA molecule comprises a sequence that is the same as, or differs by no more than 3 nucleotides from, a targeting domain sequence from Tables 5A-5B.

156. The nucleic acid of claim 147, wherein said targeting domain of said second gRNA molecule is selected from: Tables 1A-1D.

157. The nucleic acid of claim 147, wherein said targeting domain of said second gRNA molecule is selected from: Tables 2A-2C.

158. The nucleic acid of claim 147, wherein said targeting domain of said second gRNA molecule is selected from: Tables 3A-3D.

159. The nucleic acid of claim 147, wherein said targeting domain of said second gRNA molecule is selected from: Tables 4A-4D.

160. The nucleic acid of claim 147, wherein said targeting domain of said second gRNA molecule is selected from: Tables 5A-5B.

161. The nucleic acid of claim 147, wherein said targeting domain is selected from those in Tables 6A-6D.

162. The nucleic acid of claim 147, wherein said targeting domain is selected from those in Tables 7A-7D.

163. The nucleic acid of claim 147, wherein said targeting domain is selected from those in Tables 8A-8E.

164. The nucleic acid of claim 147, wherein said targeting domain is selected from those in Tables 9A-9B.

165. The nucleic acid of claim 147, wherein said targeting domain is selected from those in Table 10, e.g., GACACTGCCAATAGGGATAGGT; GTCAAAAGCTACCGGTTACCTG; GTTCTGTCCTCAGTAAAAGGTA; GAATAGTTTGTTCTGGGTAC; GAGAAAGGGATGGGCACTTA; GATGCAGAACTAGTGTAGAC; GTCACATGGGAGTCACAGGG; or GAGTATCTCCTGTTTGGCA.

166. The nucleic acid of claim 147, wherein said targeting domain is selected from those in Table 1A.

167. The nucleic acid of claim 147, wherein said targeting domain is GAGAUACUCACAAUUACAAC.

168. The nucleic acid of claim 147, wherein said targeting domain is GAUACUCACAAUUACAACUG.

169. The nucleic acid of claim 147, wherein said targeting domain is selected from those in Table 2A.

170. The nucleic acid of claim 147, wherein said targeting domain is GAGAUACUCACAAUUACAAC.

171. The nucleic acid of claim 147, wherein said targeting domain is GAUACUCACAAUUACAA.

172. The nucleic acid of claim 147, wherein said targeting domain is selected from those in Table 3A.

173. The nucleic acid of claim 147, wherein said targeting domain is GCUACCGGUUACCUGAA.

174. The nucleic acid of claim 147, wherein said targeting domain is GCAGAACUAGUGUAGAC.

175. The nucleic acid of claim 147, wherein said targeting domain is GUUGAGUAUCUCCUGUU.

176. The nucleic acid of claim 147, wherein said targeting domain is GAUGCAGAACUAGUGUAGAC.

177. The nucleic acid of claim 147, wherein said targeting domain is GCUUGAACUCUGUGCCAAAC.

178. The nucleic acid of claim 147, wherein said targeting domain is selected from those in Table 4A.

179. The nucleic acid of claim 147, wherein said targeting domain is GAAUCCUGAAAGCUACU.

180. The nucleic acid of claim 147, wherein said targeting domain is selected from those in Table 5A.

181. The nucleic acid of claim 147, wherein said targeting domain is GAGUUCAAGCUAAUACAUGA.

182. The nucleic acid of claim 147, wherein said targeting domain is GUUGUUCUGAGUAGCUU.

183. The nucleic acid of claim 147, wherein said targeting domain is GGCAAAAGCAGCAGAAAGCA.

184. The nucleic acid of claim 147, wherein said targeting domain is GUUGUUCUGAGUAGCUU.

185. The nucleic acid of claim 147, wherein said targeting domain is GGCAAAAGCAGCAGAAAGCA.

186. The nucleic acid of claim 147, wherein said targeting domain is selected from those in Table 6A.

187. The nucleic acid of claim 147, wherein said targeting domain is GCACCUGGCCCCAGUUGUAAUU.

188. The nucleic acid of claim 147, wherein said targeting domain is selected from those in Table 7A.

189. The nucleic acid of claim 147, wherein said targeting domain is selected from those in Table 8A.

190. The nucleic acid of claim 147, wherein said targeting domain is selected from those in Table 9A.

191. The nucleic acid of claim 147, wherein said targeting domain is GGCAAAAGCAGCAGAAAGCA.

192. The nucleic acid of claim 147, wherein said targeting domain is GUGGCUGAAUGACUUCU.

193. The nucleic acid of claim 147, wherein said targeting domain is GUUGUUCUGAGUAGCUU.

194. The nucleic acid of claim 147, wherein said targeting domain is GACUAGAGGUCACGAAA.

195. The nucleic acid of claim 147, wherein said targeting domain is GAGUUCAAGCUAAUACAUGA.

196. The nucleic acid of claim 147, wherein said targeting domain is selected from those in Table 10, e.g., GACACTGCCAATAGGGATAGGT; GTCAAAAGCTACCGGTTACCTG; GTTCTGTCCTCAGTAAAAGGTA; GAATAGTTTGTTCTGGGTAC; GAGAAAGGGATGGGCACTTA; GATGCAGAACTAGTGTAGAC; GTCACATGGGAGTCACAGGG; or GAGTATCTCCTGTTTGGCA.

197. The nucleic acid of any of claims 147-196, wherein said second gRNA molecule is a modular gRNA molecule.

198. The nucleic acid of any of claims 147-196, wherein said second gRNA molecule is a chimeric gRNA molecule.

199. The nucleic acid of any of claims 147-198, wherein said targeting domain is 16 nucleotides or more in length.

200. The nucleic acid of any of claims 147-199, wherein said targeting domain is 16 nucleotides in length.

201. The nucleic acid of any of claims 147-199, wherein said targeting domain is 17 nucleotides in length.

202. The nucleic acid of any of claims 147-199, wherein said targeting domain is 18 nucleotides in length.

203. The nucleic acid of any of claims 147-199, wherein said targeting domain is 19 nucleotides in length.

204. The nucleic acid of any of claims 147-199, wherein said targeting domain is 20 nucleotides in length.

205. The nucleic acid of any of claims 147-199, wherein said targeting domain is 21 nucleotides in length.

206. The nucleic acid of any of claims 147-199, wherein said targeting domain is 22 nucleotides in length.

207. The nucleic acid of any of claims 147-199, wherein said targeting domain is 23 nucleotides in length.

208. The nucleic acid of any of claims 147-199, wherein said targeting domain is 24 nucleotides in length.

209. The nucleic acid of any of claims 147-199, wherein said targeting domain is 25 nucleotides in length.

210. The nucleic acid of any of claims 147-199, wherein said targeting domain is 26 nucleotides in length.

211. The nucleic acid of any of claims 147-210, wherein said second gRNA molecule comprises from 5' to 3':

a targeting domain;

a first complementarity domain;

a linking domain;

a second complementarity domain;

a proximal domain; and

a tail domain.

212. The nucleic acid of any of claims 147-211, wherein said second gRNA molecule comprises:

a linking domain of no more than 25 nucleotides in length;

a proximal and tail domain, that taken together, are at least 20 nucleotides in length;

a targeting domain of 17 or 18 nucleotides in length.

213. The nucleic acid of any of claims 147-211, wherein said second molecule gRNA molecule comprises:

a linking domain of no more than 25 nucleotides in length;

a proximal and tail domain, that taken together, are at least 30 nucleotides in length;

a targeting domain of 17 or 18 nucleotides in length.

214. The nucleic acid of any of claims 147-211, wherein said second gRNA molecule comprises:

a linking domain of no more than 25 nucleotides in length;

a proximal and tail domain, that taken together, are at least 30 nucleotides in length;

a targeting domain of 17 nucleotides in length.

215. The nucleic acid of any of claims 147-211, wherein said second gRNA molecule comprises:

a linking domain of no more than 25 nucleotides in length;

a proximal and tail domain, that taken together, are at least 40 nucleotides in length;

a targeting domain of 17 nucleotides in length.

216. The nucleic acid of any of claims 147-215, further comprising a third gRNA molecule.

217. The nucleic acid of claim 216, further comprising a fourth gRNA molecule.

218. The nucleic acid of claim 217, further comprising: (b) a sequence that encodes a Cas9 molecule.

219. The nucleic acid of claim 217, wherein said nucleic acid does not comprise (c) a sequence that encodes a second gRNA molecule.

220. The nucleic acid of claim 218, wherein each of (a) a sequence that encodes a gRNA molecule and (b) a sequence that encodes a Cas9 molecule is present on the same nucleic acid molecule.

221. The nucleic acid of any of claims 69-220, wherein said nucleic acid molecule is an AAV vector.

222. The nucleic acid of claims 218, wherein: (a) a sequence that encodes a gRNA molecule is present on a first nucleic acid molecule; and (b) a sequence that encodes a Cas9 molecule is present on a second nucleic acid molecule.

223. The nucleic acid of claim 222, wherein said first and second nucleic acid molecules are AAV vectors.

224. The nucleic acid of claim 217, further comprising:

(c) a sequence that encodes a second gRNA molecule.

225. The nucleic acid of claim 224, wherein each of (a) a sequence that encodes a gRNA molecule and (c) a sequence that encodes a second gRNA molecule is present on the same nucleic acid molecule.

226. The nucleic acid of claim 224, wherein said nucleic acid molecule is an AAV vector.

227. The nucleic acid of claim 223, wherein:

(a) a sequence that encodes a gRNA molecule is present on a first nucleic acid molecule; and

(c) a sequence that encodes a second gRNA molecule is present on a second nucleic acid molecule.

228. The nucleic acid of claim 227, wherein said first and second nucleic acid molecules are AAV vectors.

229. The nucleic acid of claim 227, further comprising:

(b) a sequence that encodes a Cas9 molecule of any of claims 134-146; and

(c) a sequence that encode a second gRNA molecule of clams 147-215.

230. The nucleic acid of claim 229, wherein each of (a), (b), and (c) are present on the same nucleic acid molecule.

231. The nucleic acid of claim 230, wherein said nucleic acid molecule is an AAV vector.

232. The nucleic acid of claim 227, wherein:

one of (a), (b), and (c) is encoded on a first nucleic acid molecule; and

and a second and third of (a), (b), and (c) is encoded on a second nucleic acid molecule.

233. The nucleic acid of claim 232, wherein said first and second nucleic acid molecules are AAV vectors.

234. The nucleic acid of claim 227, wherein:

(a) is present on a first nucleic acid molecule; and

(b) and (c) are present on a second nucleic acid molecule.

235. The nucleic acid of claim 234, wherein said first and second nucleic acid molecules are AAV vectors.

236. The nucleic acid of claim 227, wherein:

(b) is present on a first nucleic acid molecule; and

(a) and (c) are present on a second nucleic acid molecule.

237. The nucleic acid of claim 236, wherein said first and second nucleic acid molecules are AAV vectors.

238. The nucleic acid of claim 227, wherein:

(c) is present on a first nucleic acid molecule; and

(b) and (a) are present on a second nucleic acid molecule.

239. The nucleic acid of claim 238, wherein said first and second nucleic acid molecules are AAV vectors.

240. The nucleic acid of any of claims 220, 225, 230, 232, 234, or 236, wherein said first nucleic acid molecule is other than an AAV vector and said second nucleic acid molecule is an AAV vector.

241. The nucleic acid of any of claims 69-240, wherein said nucleic acid comprises a promoter operably linked to the sequence that encodes said gRNA molecule of (a).

242. The nucleic acid of claims 147-240, wherein said nucleic acid comprises a second promoter operably linked to the sequence that encodes the second gRNA molecule of (c).

243. The nucleic acid of claim 241 or 242, wherein the promoter and second promoter differ from one another.

244. The nucleic acid of claim 241 or 242, wherein the promoter and second promoter are the same.

245. The nucleic acid of any of claims 134-244, wherein said nucleic acid comprises a promoter operably linked to the sequence that encodes the Cas9 molecule of (b).

246. A composition comprising the (a) gRNA molecule of any of claims 1-68.

247. The composition of claim 246, further comprising (b) a Cas9 molecule of any of claims 134-146.

248. The composition of any of claims 246 or 247, further comprising (c) a second gRNA molecule of any of claims 147-215.

249. A cell comprising a modification at the LCA10 target position.

250. The cell of claim 249, wherein said cell is manipulated by altering the *CEP290* gene.

251. The cell of claim 249 or 250, wherein said cell comprising one or more nucleic acids according to claims 69-245.

252. The cell of any of claims 249-251, wherein said cell is a retinal cell.

253. The cell of any of claims 249-252, wherein said cell is a photoreceptor cell.

254. The cell of claim 253, wherein said photoreceptor cell is a cone photoreceptor cell or cone cell, a rod photoreceptor cell or rod cell, or a macular cone photoreceptor cell.

255. The cell of any of claims 249-254, wherein said cell is induced pluripotent stem cells (iPS) cells or cells derived from iPS cells, modified to alter the gene and differentiated into retinal progenitor cells or retinal cells, and injected into the eye of the subject.

256. A method of altering a cell comprising contacting said cell with:

(a) a gRNA of any of claims 1-68;

(b) a Cas9 molecule of any of claims 134-146; and

optionally, (c) a second gRNA molecule of any of claims 147-215.

257. The method of claim 256, comprising contacting said cell with (a), (b), and (c).

258. The method of claim 256 or 257, wherein said cell is from a subject suffering from LCA10.

259. The method of any of claims 256-258, wherein said cell is from a subject having a mutation at the LCA10 target position of the *CEP290* gene.

260. The method of any of claims 256-259, wherein said cell is a photoreceptor cell.

261. The method of claim 256-260, wherein said contacting is performed *ex vivo*.

262. The method of claim 256-261, wherein said contacted cell is returned to said subject's body.

263. The method of claim 256-260, wherein said contacting is performed *in vivo*.

264. The method of any of claims 256-263, comprising acquiring knowledge of the presence of the LCA10 target position mutation in said cell.

265. The method of any of claims 256-264, comprising acquiring knowledge of the presence of the LCA10 target position mutation in said cell by sequencing a portion of the *CEP290* gene.

266. The method of any of claims 256-265, comprising altering the LCA10 target position in the CEP290 gene.

267. The method of any of claims 256-266, wherein contacting comprises contacting said cell with a nucleic acid that expresses at least one of (a), (b), and (c).

268. The method of any of claims 256-267, wherein contacting comprises contacting the cell with a nuclei acid of any claim 69-245.

269. The method of any of claims 256-268, wherein contacting comprises delivering to said cell said Cas9 molecule of (b) and a nucleic acid which encodes and (a) and optionally (c).

270. The method of any of claims 256-269, wherein contacting comprises delivering to said cell said Cas9 molecule of (b), said gRNA molecule of (a) and optionally said second gRNA molecule of (c).

271. The method of any of claims 256-269, wherein contacting comprises delivering to said cell said gRNA molecule of (a), optionally said second gRNA molecule of (c) and a nucleic acid that encodes the Cas9 molecule of (b).

272. A method of treating a subject, comprising contacting a subject (or a cell from said subject) with:

(a) a gRNA of any of claims 1-68;

(b) a Cas9 molecule of any of claims 134-146; and

optionally, (c) a second gRNA of any of claims 147-215.

273. The method of claim 272, further comprising contacting said subject with (a), (b), and (c).

274. The method of claims 272 or 273, wherein said subject is suffering from LCA10.

275. The method of any of claims 272-274, wherein said subject has a mutation at the LCA10 target position of the *CEP290* gene.

276. The method of any of claims 272-275, comprising acquiring knowledge of the presence of the LCA10 target position mutation in said subject.

277. The method of any of claims 272-276, comprising acquiring knowledge of the presence of the LCA10 target position mutation in said subject by sequencing a portion of the *CEP290* gene.

278. The method of any of claims 272-277, comprising altering the LCA10 target position in the CEP290 gene.

279. The method of any of claims 272-278, wherein a cell of said subject is contacted *ex vivo* with (a), (b), and optionally (c).

280. The method of any of claims 272-279, wherein said cell is returned to the subject's body.

281. The method of any of claims 272-280, wherein treatment comprises introducing a cell into said subject's body, wherein said cell subject was contacted *ex vivo* with (a), (b), and optionally (c).

282. The method of any of claims 272-278, wherein said contacting is performed *in vivo*.

283. The method of claim 282, wherein said contacting comprises subretinal delivery.

284. The method of claim 282, wherein said contacting comprises subretinal injection.

285. The method of any of claims 272-284, wherein contacting comprises contacting said subject with a nucleic acid that expresses at least one of (a), (b), and (c).

286. The method of any of claims 272-285, wherein contacting comprises contacting said subject with a nucleic acid of any of any of claims 69-245.

287. The method of any of claims 272-286, wherein contacting comprises delivering to said subject said Cas9 molecule of (b) and a nucleic acid which encodes and (a) and optionally (c).

288. The method of any of claims 272-287, wherein contacting comprises delivering to said subject said Cas9 molecule of (b), said gRNA of (a) and optionally said second gRNA of (c).

289. The method of any of claims 272-287, wherein contacting comprises delivering to said subject said gRNA of (a), optionally said second gRNA of (c) and a nucleic acid that encodes the Cas9 molecule of (b).

290. A gRNA molecule of any of claims 1-68 for use in treating LCA10 in a subject.

291. The gRNA molecue of claim 290, wherein the gRNA molecule in used in combination with (b) a Cas9 molecule of any of claims 134-146.

292. The gRNA molecule of claim 251 or 252, wherein the gRNA molecule is used in combination with (c) a second gRNA molecule of any of claims 147-215.

293. Use of a gRNA molecule of any of claims 1-68 in the manufacture of a medicament for treating LCA10 in a subject.

294. The use of claim 293, wherein the medicament further comprises (b) a Cas9 molecule of any of claims 134-146.

295. The use of claim 293 or 294, wherein the medicament further comprises (c) a second gRNA molecule of any of claims 147-215.

296. A composition of any of claims 246-248 for use in treating LCA10 in a subject.

297. A reaction mixture comprising a gRNA, a nucleic acid, or a composition described herein, and a cell from a subject having LCA10, or a subject having a mutation a LCA10 target position of the *CEP290* gene.

298. A kit comprising, (a) gRNA molecule of any of claims 1-68, or nucleic acid that encodes said gRNA, and one or more of the following:

(b) a Cas9 molecule of any of claims 134-146;

(c) a second gRNA molecule of any of claims 147-215;

(d) nucleic acid that encodes one or more of (b) and (c).

299. The kit of claim 298, comprising nucleic acid that encodes one or more of (a), (b) and (c).

300. A recombinant adenovirus-associated virus (AAV) genome comprising the following components:

left ITR–spacer 1–U6 promoter–gRNA–spacer 2–PII promoter–N-ter NLS–Cas9–C-ter NLS–poly(A) signal–spacer 3–right ITR,

wherein the left ITR component comprises, or consists of, a nucleotide sequence that is the same as, or differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, any of the nucleotide sequences disclosed in Table 24, or SEQ ID NOS: 407-415;

wherein the spacer 1 component comprises, or consists of, a nucleotide sequence having 0 to 150 nucleotides in length, e.g., SEQ ID NO: 416;

wherein the U6 promoter component comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 417;

wherein the gRNA component comprises a targeting domain and a scaffold domain, wherein the targeting domain is 16-26 nucleotides in length, and comprises, or consists of, a targeting domain sequence disclosed herein, in any of Tables 1A-1D, Tables 2A-2C, Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B, or Table 10; and

wherein the scaffold domain comprises, or consists of, a nucleotide sequence that is the same as, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 418;

wherein the spacer 2 component comprises, or consists of, a nucleotide sequence having 0 to 150 nucleotides in length e.g., SEQ ID NO: 419;

wherein the PII promoter component comprises, or consists of, a polymerase II promoter sequence, e.g., a constitutive or tissue specific promoter, e.g., a promoter disclosed in Table 19;

wherein the N-ter NLS component comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, or 5 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 420;

wherein the Cas9 component comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 421;

wherein the C-ter NLS component comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, or 5 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 422;

wherein the poly(A) signal component comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 424;

wherein the spacer 3 component comprises, or consists of, a nucleotide sequence having 0 to 150 nucleotides in length e.g., SEQ ID NO: 425;

wherein the right ITR component comprises, or consists of, a nucleotide sequence that is the same as, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, any of the nucleotide sequences disclosed in Table 24, or SEQ ID NOS: 436-444.

301. The recombinant AAV genome of claim 300, wherein the left ITR component comprises, or consists of, a nucleotide sequence that is the same as any one of the nucleotide sequences of SEQ ID NOS: 407-415.

302. The recombinant AAV genome of claim 300 or 301, wherein the spacer 1 component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 416.

303. The recombinant AAV genome of any of claims 300-302, wherein the U6 promoter component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 417.

304. The recombinant AAV genome of any of claims 300-303, wherein the gRNA scaffold domain comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 418.

305. The recombinant AAV genome of any of claims 300-304, wherein the spacer 2 component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 419;

306. The recombinant AAV genome of any of claims 300-305, wherein the PII promoter is a CMV promoter, and comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 401.

307. The recombinant AAV genome of any of claims 300-306, wherein the PII promoter comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 401.

308. The recombinant AAV genome of any of claims 300-305, wherein the PII promoter is an EFS promoter, and comprises, or consists of, a nucleotide sequence that is

the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 402.

309. The recombinant AAV genome of any of claims 300-305 or 308, wherein the PII promoter comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 402.

310. The recombinant AAV genome of any of claims 300-305, wherein the PII promoter is a GRK1 promoter (e.g., a human GRK1 promoter), and comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 403.

311. The recombinant AAV genome of any of claims 300-305 or 310, wherein the PII promoter comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 403.

312. The recombinant AAV genome of any of claims 300-305, wherein the PII promoter is a CRX promoter (e.g., a human CRX promoter), and comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 404.

313. The recombinant AAV genome of any of claims 300-305 or 312, wherein the PII promoter comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 404.

314. The recombinant AAV genome of any of claims 300-305, wherein the PII promoter is an NRL promoter (e.g., a human NRL promoter), and comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 405.

315. 16. The recombinant AAV genome of any of claims 300-305 or 314, wherein the PII promoter comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 405.

316. The recombinant AAV genome of any of claims 300-305, wherein the PII promoter is an RCVRN promoter (e.g., a human RCVRN promoter), and comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 406.

317. The recombinant AAV genome of any of claims 300-305 or 316, wherein the PII promoter comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 406.

318. The recombinant AAV genome of any of claims 300-317, wherein the N-ter NLS component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 420;

319. The recombinant AAV genome of any of claims 300-318, wherein the Cas9 component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 421;

320. The recombinant AAV genome of any of claims 300-319, wherein the C-ter NLS component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 422.

321. The recombinant AAV genome of any of claims 300-320, wherein the poly(A) signal component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 424.

322. The recombinant AAV genome of any of claims 300-321, wherein the spacer 3 component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 425.

323. The recombinant AAV genome of any of claims 300-322, wherein the right ITR component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 437.

324. The recombinant AAV genome of any of claims 300-323, further comprising a second gRNA component comprising a targeting domain and a scaffold domain, wherein the targeting domain is 16-26 nucleotides in length and comprises, or consists of, a targeting domain sequence disclosed herein, in any of Tables 1A-1D, Tables 2A-2C,

Tables 3A-3D, Tables 4A-4D, Tables 5A-5B, Tables 6A-6D, Tables 7A-7D, Tables 8A-8E, Tables 9A-9B, or Table 10; and

wherein the scaffold domain comprises, or consists of, a nucleotide sequence that is the same as, or differs by no more than 1, 2, 3, 4, or 5 nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 418;

325. The recombinant AAV genome of claim 324, wherein the second gRNA component is between the first gRNA component and the spacer 2 component.

326. The recombinant AAV genome of claim 324 or 325, wherein the second gRNA component is the same as the first gRNA component.

327. The recombinant AAV genome of claim 324-325, wherein the second gRNA component is different from the first gRNA component.

328. The recombinant AAV genome of any of claims 324-327, further comprising a second U6 promoter component between the first gRNA component and the second gRNA component.

329. The recombinant AAV genome of claim 328, wherein the second U6 promoter component is the same as the first U6 promoter component.

330. The recombinant AAV genome of claim 328, wherein the second U6 promoter component is different from the first U6 promoter component.

331. The recombinant AAV genome of any of claims 324-330, further comprising a spacer 4 component between the first gRNA component and the second U6 promoter component.

332. The recombinant AAV genome of claim 331, wherein the spacer 4 component comprises, or consists of, a nucleotide sequence having 0 to 150 nucleotides in length.

333. The recombinant AAV genome of claim 331 or 332, wherein the spacer 4 component comprises, or consists of, a nucleotide sequence that is the same as the nucleotide sequence of SEQ ID NO: 427;

334. The recombinant AAV genome of any of claims 300-333, further comprises a 3xFLAG component, wherein the 3xFLAG component comprises, or consists of, a nucleotide sequence that is the same as, differs by no more than 1, 2, 3, 4, or 5

nucleotides from, or has at least has at least 90%, 92%, 94%, 96%, 98%, or 99% homology with, the nucleotide sequence of SEQ ID NO: 423.

335. The recombinant AAV genome of claim 334, wherein the 3xFLAG component is between the C-ter NLS component and the poly(A) signal component.

336. The recombinant AAV genome of claim 334 or 335, wherein the 3xFLAG component comprises, or consists of, a nucleotide sequence that is the same as, the nucleotide sequence of SEQ ID NO: 423.

337. The recombinant AAV genome of any of claims 300-336, comprising the nucleotide sequences of SEQ ID NOS: 408, 417, 418, 401, 420, 421, 422, 424, and 437.

338. The recombinant AAV genome of any of claims 300-336, comprising the nucleotide sequences of SEQ ID NOS: 408, 417, 418, 402, 420, 421, 422, 424, and 437.

339. The recombinant AAV genome of any of claims 300-336, comprising the nucleotide sequences of SEQ ID NOS: 408, 417, 418, 403, 420, 421, 422, 424, and 437.

340. The recombinant AAV genome of any of claims 300-336, comprising the nucleotide sequences of SEQ ID NOS: 408, 417, 418, 404, 420, 421, 422, 424, and 437.

341. The recombinant AAV genome of any of claims 300-336, comprising the nucleotide sequences of SEQ ID NOS: 408, 417, 418, 405, 420, 421, 422, 424, and 437.

342. The recombinant AAV genome of any of claims 300-336, comprising the nucleotide sequences of SEQ ID NOS: 408, 417, 418, 406, 420, 421, 422, 424, and 437.

343. The recombinant AAV genome of any of claims 1-342, further comprising SEQ ID NO: 416, 419, and 425, and, optionally, 427.

344. The recombinant AAV genome of any of claims 1-343, comprising a nucleotide sequence that is the same as, differs by no more than 100, 200, 300, 400, or 500 nucleotides, or has at least 90%, 92%, 94%, 96%, or 98% homology with, any of the nucleotide sequences of SEQ ID NOS: 428-433 or 445-450.

345. The recombinant AAV genome of any of claims 1-344, comprising a nucleotide sequence that is the same as any of the nucleotide sequences of SEQ ID NOS: 428-433.

346. A recombinant AAV viral particle comprising the AAV genome of any of claims 1-345.

347. The recombinant AAV viral particle of claim 346 having any of the serotype disclosed herein, e.g., in Table 24, or a combination thereof.

348. The recombinant AAV viral particle of claim 346 or 347 having a tissue specificity of retinal pigment epithelium cells, photoreceptors, horizontal cells, bipolar cells, amacrine cells, ganglion cells, or a combination thereof.

349. A method of producing the recombinant AAV viral particle of any of claims 346-348 comprising providing the recombinant AAV genome of any of claims 300-345 and one or more capsid proteins under conditions that allow for assembly of an AAV viral particle.

350. A method of altering a cell comprising contacting the cell with the recombinant AAV viral particle of any of claims 346-348.

351. A method of treating a subject having or likely to develop LCA10 comprising contacting the subject (or a cell from the subject) with the recombinant viral particle of any of claims 346-348.

352. A recombinant AAV viral particle comprising the AAV genome of any of claims 346-348 for use in treating LCA10 in a subject.

353. Use of a recombinant AAV viral particle comprising the AAV genome of any of claims 346-348 in the manufacture of a medicament for treating LCA10 in a subject.

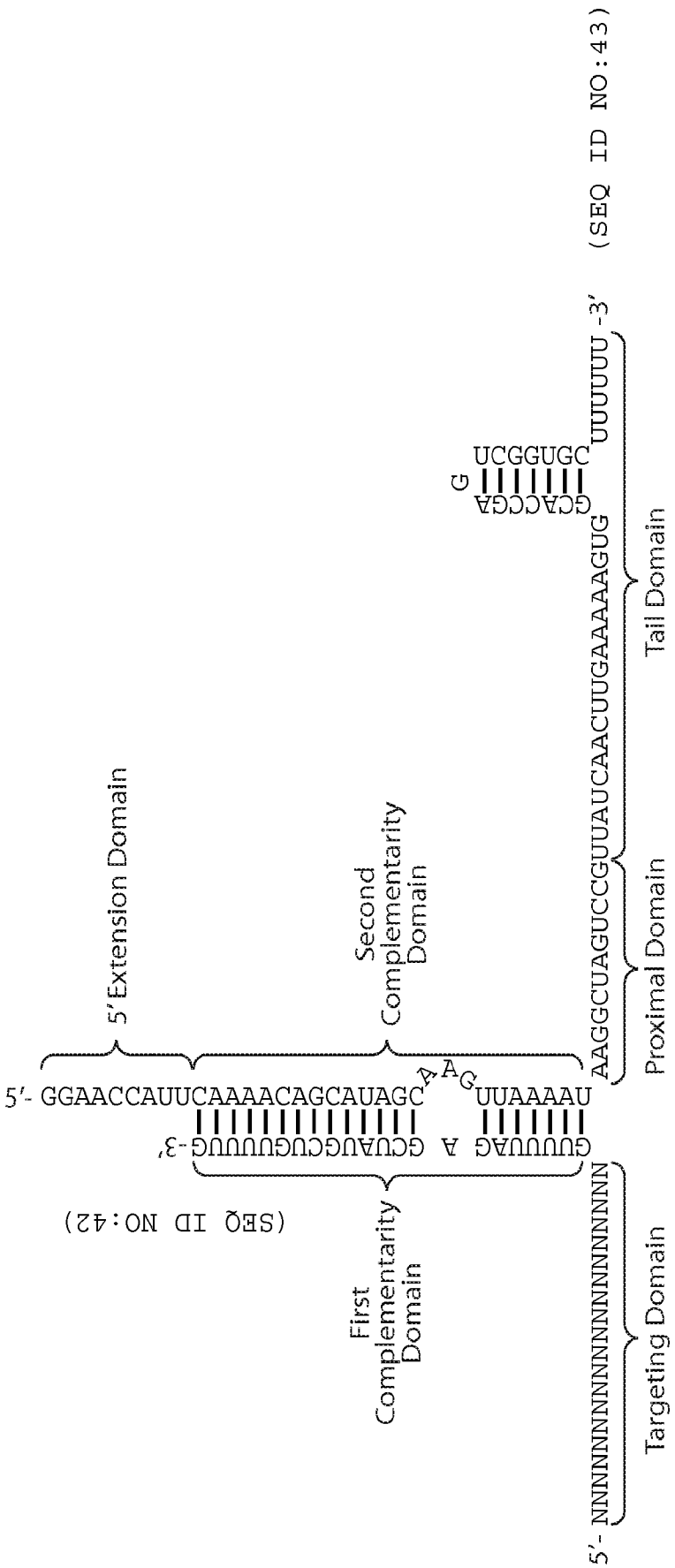

5' Extension Domain
Second Complementarity Domain
5'- GGAACCAUUCAAAACAGCAUAGC A A G UUAAAAU
AAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
Proximal Domain
Tail Domain
UCGGUGC UUUUUU -3'
G
GCACCGA
(SEQ ID NO:43)
GUUUUAG A GCUAUGCUGUUUUG -3'
(SEQ ID NO:42)
First Complementarity Domain
5'- NNNNNNNNNNNNNNNNNNNN
Targeting Domain

Fig. 1A

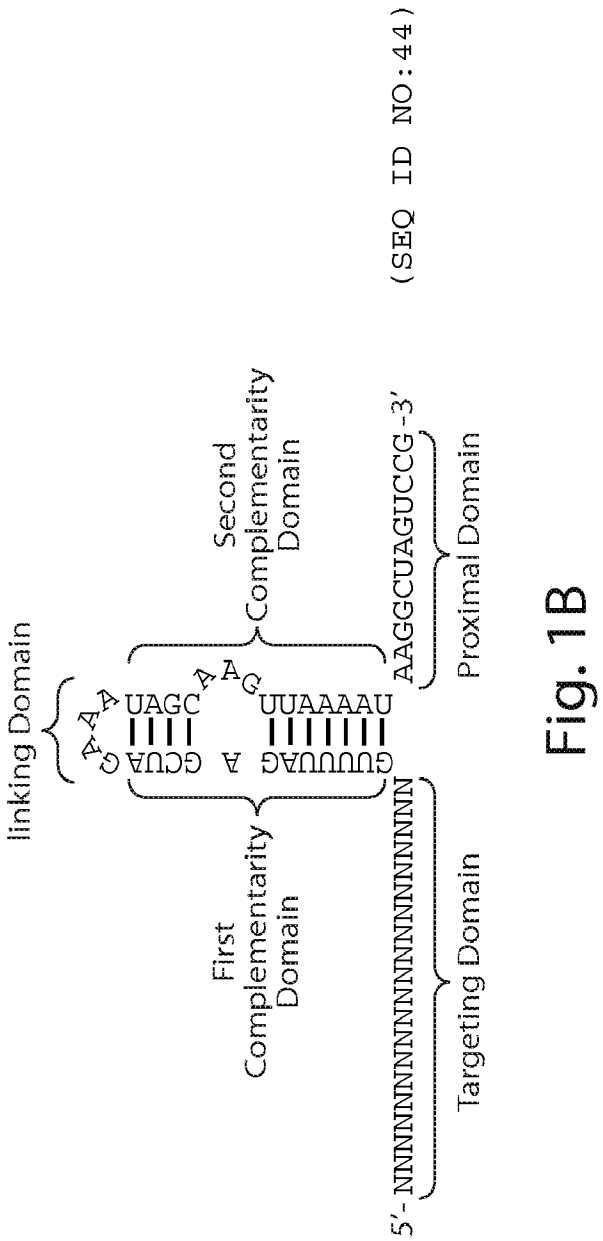

linking Domain
UAGC
GUUUUAG A GCUA
UUAAAAU
AAGGCUAGUCCG-3'
(SEQ ID NO:44)
Second Complementarity Domain
First Complementarity Domain
Proximal Domain
Targeting Domain
5'- NNNNNNNNNNNNNNNNNNNN

Fig. 1B

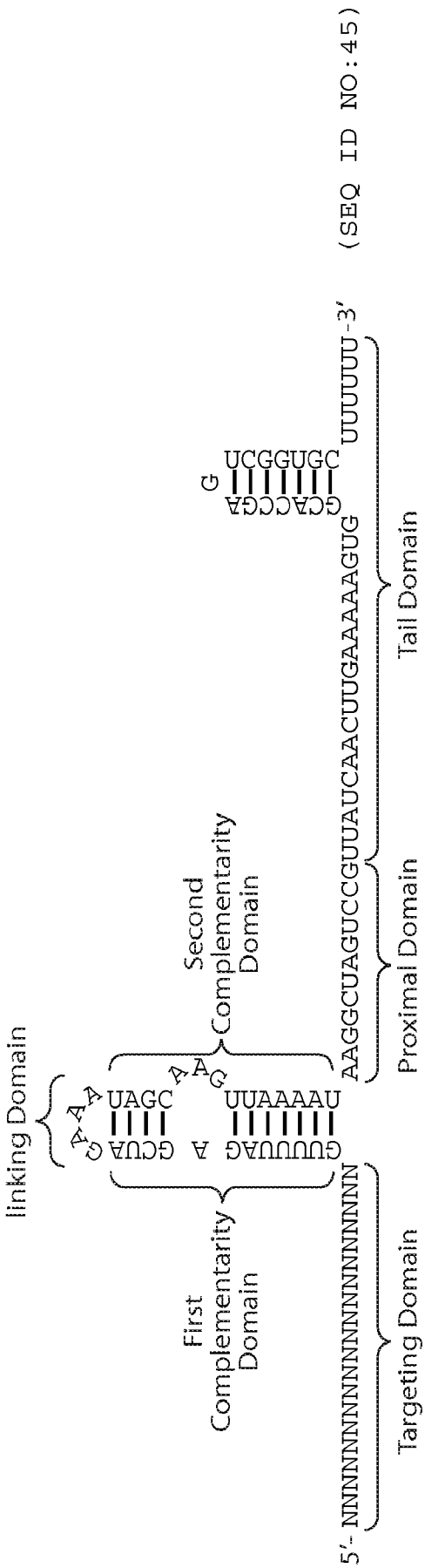

5'- NNNNNNNNNNNNNNNNNNNN
Targeting Domain
First Complementarity Domain
linking Domain
UAGC
UUAAAAU
GUUUUAG A GCUA
AAAG
AAA
Second Complementarity Domain
AAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
Proximal Domain
Tail Domain
UUUUUU-3'
UCGGUGC
GCACCGA
G
(SEQ ID NO:45)

Fig. 1C

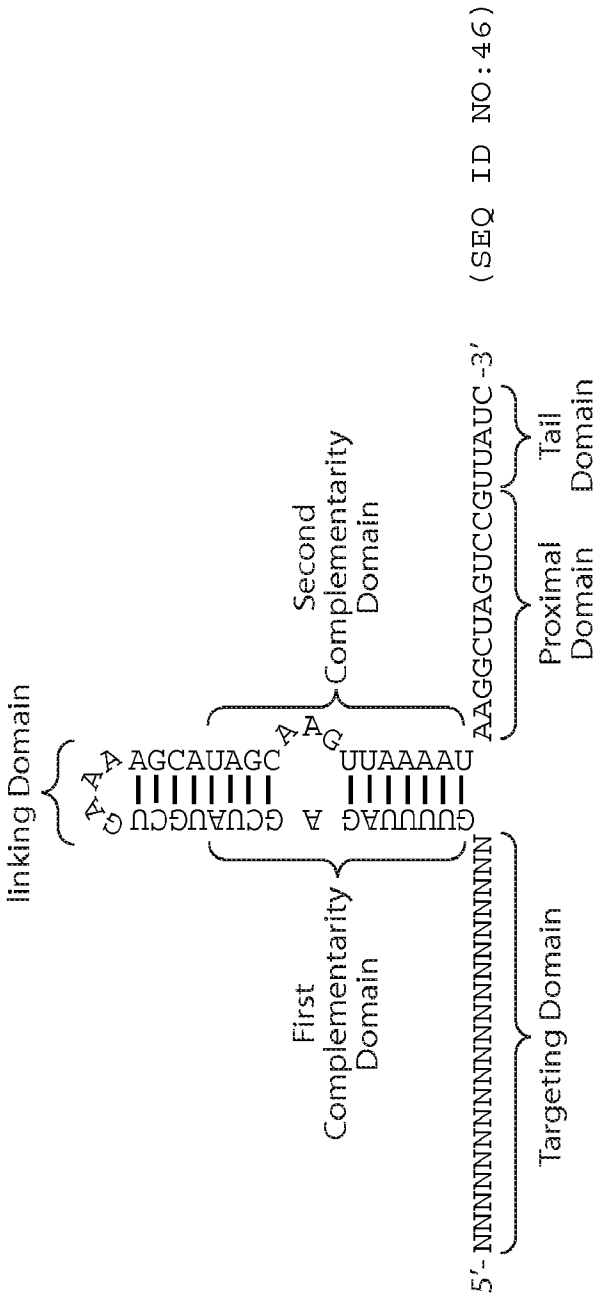

linking Domain
AAA
AGCAUAGC
AAAG
UUAAAAU
GUUUUAG A GCUAUGCU
AAGGCUAGUCCGUUAUC -3'
(SEQ ID NO:46)
Second Complementarity Domain
Proximal Domain
Tail Domain
First Complementarity Domain
5'- NNNNNNNNNNNNNNNNNNNN
Targeting Domain

Fig. 1D

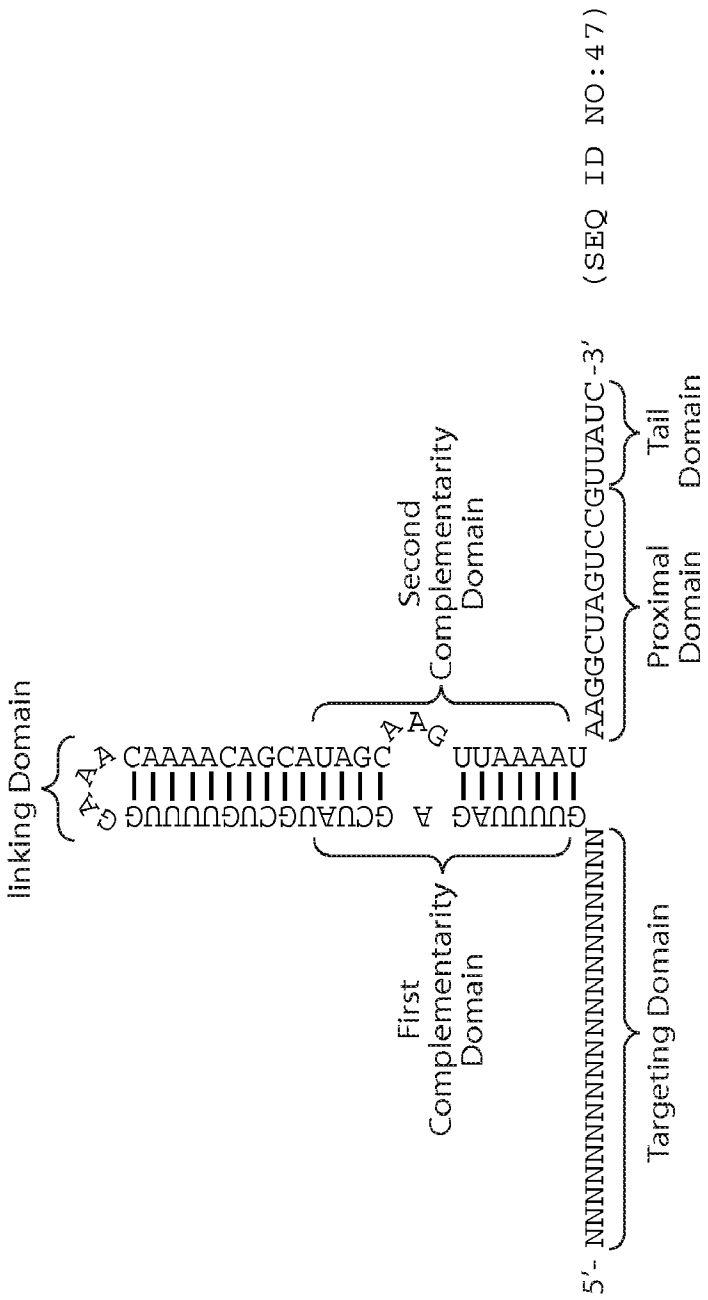

linking Domain
CAAAACAGCAUAGC
GUUUUAG A GCUAUGCUGUUUUG
AAAG
AAAG
First Complementarity Domain
Second Complementarity Domain
AAGGCUAGUCCGUUAUC-3'
(SEQ ID NO:47)
Proximal Domain
Tail Domain
5'- NNNNNNNNNNNNNNNNNNNN
Targeting Domain

Fig. 1E

Streptococcus thermophilus

a) Structure (See, e.g., Karvelis et al. RNA Biology. 2013; 10(5): 841-851)

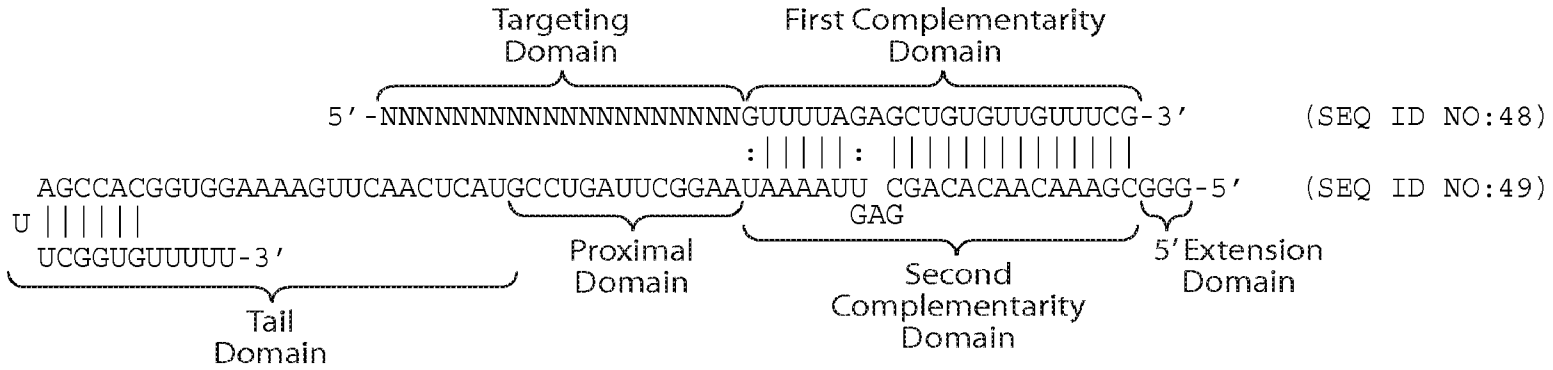


Fig. 1F

Alignment

S. pyogenes 5'-NNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAUGCUGUUUUG-3' (SEQ ID NO:50)
S. thermophilus 5'-NNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUGUGUUGUUUCG-3' (SEQ ID NO:51)

S. pyogenes 5'-GAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGC-UAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU-3'
S. thermophilus 5'------GGGCGAAACAACACAGCGAGUUAAAAUAAGGCUUAGUCCGUACUCAACUUGAAAAGGUGGCACCGAUUCGGUGUUUUU----3'

S. pyogenes - cont (SEQ ID NO:52)
S. thermophilus - *cont* (SEQ ID NO:53)

Fig. 1G

```
CLUSTAL format alignment by MAFFT (v7.058b)

                    *                                                  Y
SM                  KKPYSIGLDIGTNSVGWAVVTDDYKVPAKKMKVLGNTDKSHIEKNLLGALLFDSGNTAED
SP                  DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
ST                  TKPYSIGLDIGTNSVGWAVTTDNYKVPSKKMKVLGNTSKKYIKKNLLGVLLFDSGITAEG
LI                  KKPYTIGLDIGTNSVGWAVLTDQYDLVKRKMKIAGDSEKKQIKKNFWGVRLFDEGQTAAD
                      * *:*********************  **:*.:   :**:*  *::.:   *::**:  *.  ***.* **
Motif:              -K-Y*IGLDIGTNSVGWAV-TD*Y-*---*K*K*-G**-*--I*KN*-G--LFD-G-TA--

SM                  RRLKRTARRRYTRRRNRILYLQEIFSEEMGKVDDSFFHRLEDSFLVTEDKRGERHPIFGN
SP                  TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGN
ST                  RRLKRTARRRYTRRRNRILYLQEIFSTEMATLDDAFFQRLDDSFLVPDDKRDSKYPIFGN
LI                  RRMARTARRRIERRRNRISYLQGIFAEEMSKTDANFFCRLSDSFYVDNEKRNSRHPFFAT
                     *: *******  **:**** *** **: **.. *  ** **.:** *  ::*: .::*:*..
Motif:              -R*-RTARRR--RR*NRI-YLQ-IF*-EM---D--FF-RL-*SF-V-**K*--**P*F--

SM                  LEEEVKYHENFPTIYHLRQYLADNPEKVDLRLVYLALAHIIKFRGHFLIEGKFDTRNNDV
SP                  IVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDV
ST                  LVEEKAYHDEFPTIYHLRKYLADSTKKADLRLVYLALAHMIKYRGHFLIEGEFNSKNNDI
LI                  IEEEVEYHKNYPTIYHLREELVNSSEKADLRLVYLALAHIIKYRGNFLIEGALDTQNTSV
                    : :*  **.::**********: *.:...*.****:*******:**:**:***** ::. *..:
Motif:              *-**E--YH-**PTIYHLR*-L-*---K-DLRL*YLALAH*IK*RGNFLIEG-**--N--*
```

Fig. 2A

```
SM          QRLFQEFLAVYDNTFENSS-------LQEQNVQVEEILTDKISKSAKKDRVLKLFPNEKSN
SP          DKLFIQLVQTYNQLFEENP-------INASGVDAKAILSARLSKSRRLENLIAQLPGEKKN
ST          QKNFQDFLDTYNAIFESDL-------SLENSKQLEEIVKDKISKLEKKDRILKLFPGEKNS
LI          DGIYKQFIQTYNQVFASGIEDGSLKKLEDNKDVAKILVEKVTRKEKLERILKLYPGEKSA
            :  : ::: .*:  * ..          .. :   *:  :::: : :.::   *.**.
Motif:      *--*-*****--Y*--f-----------------*---I*--*****--*-*-**---P-EK--

SM          GRFAEFLKLIVGNQADFKKHFELEEKAPLQFSKDTYEEELEVLLAQIGDNYAELFLSAKK
SP          GLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKN
ST          GIFSEFLKLIVGNQADFRKCFNLDEKASLHFSKESYDEDLETLLGYIGDDYSDVFLKAKK
LI          GMFAQFISLIVGSKGNFQKPFDLIEKSDIECAKDSYEEDLESLLALIGDEYAELFVAAKN
            * *.::: * :*    :*:. *:* *.: :. :*::*::::*: **. ***:*::::*: **:
Motif:      G-F-***-L-*G---*F*--F*L-E-*-*---*K**Y***L*-LL--IGD*Y***F*-AK*

SM          LYDSILLSGILTVTDVGTKAPLSASMIQRYNEHQMDLAQLKQFIRQKLSDKYNEVFSDVS
SP          LSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQS
ST          LYDAILLSGFLTVTDNETEAPLSSAMIKRYNEHKEDLALLKEYIRNISLKTYNEVFKDDT
LI          AYSAVVLSSIITVAETETNAKLSASMIERFDTHEEDLGELKAFIKLHLPKHYEEIFSNTE
              .:::**.:: *    *:* **::**:*:: *. **   **   ::     . *:*:* :
Motif:      ---***LS-***-V----T*A-LS**MI*R**-H--DL--LK--**--------Y*E*F-*--

SM          KDGYAGYIDGKTNQEAFYKYLKGLLNKIEGSGYFLDKIERED FLRKQRTFDNGSIPHQIH
SP          KNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRED LLRKQRTFDNGSIPHQIH
ST          KNGYAGYIDGKTNQEDFYVYLKKLLAEFEGADYFLEKIDRED FLRKQRTFDNGSIPYQIH
LI          KHGYAGYIDGKTKQADFYKYMKMTLENIEGADYFIAKIEKEN FLRKQRTFDNGAIPHQLH
            *.********* :.*  **  ::*  *  :::*:  :: *:::*::*********:**:*:*
Motif:      K-GYAGYIDG-*-Q--FY-**K--L-***G*---**-K***E**LRKQRTFDNG*IP*Q*H
```

Fig. 2B

```
SM          LQEMRAIIRRQAEFYPFLADNQDRIEKLLTFRIPYYVGPLARGKSDFAWLSRKSADKITP
SP          LGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITP
ST          LQEMRAILDKQAKFYPFLAKNKERIEKILTFRIPYYVGPLARGNSDFAWSIRKRNEKITP
LI          LEELEAILHQQAKYYPFLKENYDKIKSLVTFRIPYFVGPLANGQSEFAWLTRKADGEIRP
            * *:.**: :* .:***** .* ::*:.::*******:**:*:* ***  **    * *
Motif:      L-E*-AI*-*Q--*YPFL--N-**I*-**TFRIPY*VGPLA-G*S-FAW--RK----I-P

SM          WNFDEIVDKESSAEAFINRMTNYDLYLPNQKVLPKHSLLYEKFTVYNELTKVKYKTE-QG
SP          WNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMR
ST          WNFEDVIDKESSAEAFINRMTSFDLYLPEEKVLPKHSLLYETFNVYNELTKVRFIAESMR
LI          WNIEEKVDFGKSAVDFIEKMTNKDTYLPKENVLPKHSLCYQKYLVYNELTKVRYIND-QG
            **::: :*   **  **:*:**. *  **:::*******  *: : ********** :
Motif:      WN****-*D---SA--FI**MT--D--LP***VLPKHSL-Y*-*-VYNELTKV**--*---

SM          KTAFFDANMKQEIFDGVFKVYRKVTKDKLMDFLEKEFDEFRIVDLTGLDKENKVFNASYG
SP          KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR--FNASLG
ST          DYQFLDSKQKKDIVRLYFKDKRKVTDKDIIEYL-HAIYGYDGIELKGIEKQ---FNSSLS
LI          KTSYFSGQEKEQIFNDLFKQKRKVKKKDLELFL-RNMSHVESPTIEGLEDS---FNSSYS
            . ::..: *: *.   **  ***. ...:     : :       : *::.    **:* .
Motif:      ---**---*-K*-I----FK--RKV-----*-*---------*-G**------FN*S--

SM          TYHDLCKIL-DKDFLDNSKNEKILEDIVLTLTLFEDREMIRKRLENYSDLLTKEQVKKLE
SP          TYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLK
ST          TYHDLLNIINDKEFLDDSSNEAIIEEIIHTLTIFEDREMIKQRLSKFENIFDKSVLKKLS
LI          TYHDLLKVGIKQEILDNPVNTEMLENIVKILTVFEDKRMIKEQLQQFSDVLDGVVLKKLE
            ***** ::  .:::***:  *  ::*:*:  **:****:.**.::*. : .::    :*:*.
Motif:      TYHDL-**---***LD*--N--**E*I*--LT*FED*-MI-**L--*--**-----*K*L-
```

Fig. 2C

SM RRHYTGWGRLSAELIHGIRNKESRKTILDYLIDDGNSNRNFMQLINDDALSFKEEIAKAQ
SP RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQ
ST RRHYTGWGKLSAKLINGIRDEKSGNTILDYLIDDGISNRNFMQLIHDDALSFKKKIQKAQ
LI RRHYTGWGRLSAKLLMGIRDKQSHLTILDYLMNDDGLNRNLMQLINDSNLSFKSIIEKEQ
Motif: **:****** :*: ***:::* ****:* .*. ***:*****:*. *:**. * * *
B
SM VIGETD--NLNQVVSDIAGSPAIKKGILQSLKIVDELVKIMG-HQPENIVVEMARENQFT
SP VSGQGD--SLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTT
ST IIGDEDKGNIKEVVKSLPGSPAIKKGILQSIKIVDELVKVMGGRKPESIVVEMARENQYT
LI VTTADK--DIQSIVADLAGSPAIKKGILQSLKIVDELVSVMG-YPPQTIVVEMARENQTT
Motif: : . .::. : :.:.********** ::*:***** .:** *:. **:******* *
*-------**--*--*-GSPAIKKGILQ**K*VDELV-*MG---P*-IV*EMARENQ-T
SM NQGRRNSQQRLKGLTDSIKEFGSQILKEH------PVENSQLQNDRLFLYYLQNGRDMYT
SP QKGQKNSRERMKRIEEGIKELGSQILKEH------PVENTQLQNEKLYLYYLQNGRDMYV
ST NQGKSNSQQRLKRLEKSLKELGSKILKENIPAKLSKIDNNALQNDRLYLYYLQNGKDMYT
LI GKGKNNSRPRYKSLEKAIKEFGSQILKEH------PTDNQELRNNRLYLYYLQNGKDMYT
Motif: :*: ***: * * : ..:**:**:*****: :* *:*::*:*********:***.
G
-*G--NS*-R-K-*---*KE*GS*ILKE*----------*N--L*N**L*LYYLQNG*DMY-
SM GEELDIDYLSQYDIDHIIPQAFIKDNSIDNRVLTSSKENRGKSDDVPSKDVVRKMKSYWS
SP DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWR
ST GDDLDIDRLSNYDIDHIIPQAFLKDNSIDNKVLVSSASNRGKSDDVPSLEVVKKRKTFWY
LI GQDLDIHNLSNYDIDHIVPQSFITDNSIDNLVLTSSAGNREKGDDVPPLEIVRKRKVFWE
Motif: .::***. **:**:***:**:*:.*:*:**** ** . * ** *.*:**. ::*:* * :*
-**LDI--LS*YD*DHI*PQ*F*-D*SIDN-VL--S--NR-K-D*VP--**V*K-K-*W-

Fig. 2D

SM KLLSAKLITQRKFDNLTKAERGGLTDDDKAGFIKRQLVETRQITKHVARILDERFNTETD
SP QLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
ST QLLKSKLISQRKFDNLTKAERGGLSPEDKAGFIQRQLVETRQITKHVARLLDEKFNNKKD
LI KLYQGNLMSKRKFDYLTKAERGGLTEADKARFIHRQLVETRQITKNVANILHQRFNYEKD
:* ..:*:::***** ********** *** **:********************:**..:*..::* : *
Motif: *L---*L***RKFD-LTKAERGGL*--DKA-FI*RQLVETRQITK*VA-*L--**N-*-D
B
SM ENNKKIRQVKIVTLKSNLVSNFRKEFELYKVREINDYHHAHDAYLNAVIGKALLGVYPQL
SP ENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKL
ST ENNRAVRTVKIITLKSTLVSQFRKDFELYKVREINDFHHAHDAYLNAVVASALLKKYPKL
LI DHGNTMKQVRIVTLKSALVSQFRKQFQLYKVRDVNDYHHAHDAYLNGVVANTLLKVYPQL
::.. :: *:::***** ***:***:*::*****::*:: **********.*:..:*: **:*
Motif: **---***-V****TLKS-LVS*FRK*F**LYKV**N**HHAHDAYLN-V*--*L*--YP*L
SM EPEFVYGDYPHFHGHKE----------NK-ATAKKFFYSNIMNFFKKDDVRTD---------
SP ESEFVYGDYKVYDVRKMIAKSEQEIGK-ATAKYFFYSNIMNFFKTEITLANGEIRKRPLI
ST EPEFVYGDYPKYNSFRE----------RKSATEKVYFYSNIMNIFKKSISLADGRVIERPLI
LI EPEFVYGDYHQFDWFKA----------NK-ATAKKQFYTNIMLFFAQKDRIID---------
*.******* :. : * ** * **:*** :* . :
Motif: E-EFVYGDY--*---*----------K-AT-K--FY*NIM-*F----------*----------
SM ----KNGEIIWKKDEHISNIKKVLSYPQVNIVKKVEEQTGGFSKE----------SILPK
SP ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE----------SILPK
ST EVNEETGESVWNKESDLATVRRVLSYPQVNVVKKVEEQNHGLDRGKPKGLFNANLSSKPK
LI ----ENGEILWDK-KYLDTVKKVMSYRQMNIVKKTEIQKGEFSKA----------TIKPK
:.** :*.* : .:::*:* *:*:***.* *. : : **
Motif: ----*-GE-*W-K---*---***V*M--Q*N*VKK-E-Q---*-*----------*--PK

Fig. 2E

```
SM          GNSDK-LIPRKTKKFYWDTKKYGGFDSPIVAYSILVIADIEKGKSKKLKTVKALVGVTIM
SP          RNSDK-LIARKKD---WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIM
ST          PNSNENLVGAKEY---LDPKKYGGYAGISNSFTVLVKGTIEKGAKKKITNVLEFQGISIL
LI          GNSSK-LIPRKTN---WDPMKYGGLDSPNMAYAVVI--EYAKGKN-KLVFEKKIIRVTIM
              **.: *:  *      *. ****   .     ::::::        **  .  *:     :  ::*:
Motif:      -NS-*-L*--K-------D--KYGG-------*******-------KG---K*------*--**I*

SM          EKMTFERDPVAFLERKGYRNVQEENIIKLPKYSLFKLENGRKRLLAS-------------ARELQK
SP          ERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS-------------AGELQK
ST          DRINYRKDKLNFLLEKGYKDI--ELIIELPKYSLFELSDGSRRMLASILSTNNKRGEIHK
LI          ERKAFEKDEKAFLEEQGYRQP--KVLAKLPKYTLYECEEGRRRMLAS-------------ANEAQK
            ::   :.::    **   :**::     .  : :*****:*::  .:* :*:***          * :*
Motif:      **--*-**---FL--*GY**-------*-*LPKY*L**--*G-*R*LAS-------------E-*K

SM          GNEIVLPNHLGTLLYHAKNIHKV-------DEPKHLDYVDKHKDEFKELLDVVSNFSKKYT
SP          GNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQL-FVEQHKHYLDEIIEQISEFSKRVI
ST          GNQIFLSQKFVKLLYHAKRISNT-------INENHRKYVENHKKEFEELFYYILEFNENYV
LI          GNQQVLPNHLVTLLHHAANCEVS-------DGKSLDYIESNREMFAELLAHVSEFAKRYT
            **:  *...:  .:*: * .                : :    :::.:::. : *::  : :* :.
Motif:      GN*--L--*----*L*-A-----------------*-*---***-**--*-E**---*-*F-*---

SM          LAEGNLEKIKELYAQNNGEDLKELASSFI--------NLLTFTAIGAPATFKFFDKNIDR
SP          LADANLDKVLSAYNKHRDKPIREQAENII--------HLFTLTNLGAPAAFKYFDTTIDR
ST          GAKKNGKLLNSAFQSWQNHSIDELCSSFIGPTGSERKGLFELTSRGSAADFEFLGVKIPR
LI          LAEANLNKINQLFEQNKEGDIKAIAQSFV--------DLMAFNAMGAPASFKFFETTIER
              *. *  . :  . :  .  .    :  ...::          *: :.  *:.* *:::  .* *
Motif:      -A--N---*---*----------*----------**----------L*-*---G*-A-F***---I-R
```

Fig. 2F

```
SM      KR-YTSTTEILNATLIHQSITGLYETRIDLNKLGGD  (SEQ ID NO:1)
SP      KR-YTSTKEVLDATLIHQSITGLYETRIDLSQLGGD  (SEQ ID NO:2)
ST      YRDYTPSSLLKDATLIHQSVTGLYETRIDLAKLGEG  (SEQ ID NO:3)
LI      KR-YNNLKELLNSTIIYQSITGLYESRKRLD----D  (SEQ ID NO:4)
          * *.  . : ::*:*:**:******:*  *     .
Motif:  -R-Y-----*-**T*I*QS*TGLYE*R--L------
```

Fig. 2G

Alignment of the N terminal RucV-like Domains disclosed in Chylinski *et al.*
(excluding sequence outliers).
(CLUSTAL format alignment by MAFFT (v7.058b))

```
1    DIGTNSVGWAVT (SEQ ID NO:54)
12   DIGTNSVGWAVT (SEQ ID NO:55
3    DVGTNSVGWAVT (SEQ ID NO:56)
20   DVGTNSVGWAVT (SEQ ID NO:57)
15   DMGTNSVGWAVT (SEQ ID NO:58)
4    DVGTSSVGWAVT (SEQ ID NO:59)
7    DIGTASVGWAVT (SEQ ID NO:60)
6    DVGTGSVGWAVT (SEQ ID NO:61)
9    DIGTNSVGWAVV (SEQ ID NO:62)
10   DIGTNSVGWAVI (SEQ ID NO:63)
11   DIGTNSVGWAVL (SEQ ID NO:64)
42   DLGTNSIGWAVV (SEQ ID NO:65)
48   DLGTNSIGWAI- (SEQ ID NO:66)
43   DLGTNSIGWALV (SEQ ID NO:67)
2    DIGTNSVGWCVT (SEQ ID NO:68)
14   DIGTNSVGYAVT (SEQ ID NO:69)
5    DMGTGSLGWAVT (SEQ ID NO:70)
16   DIGTSSVGWAAI (SEQ ID NO:71)
8    DLGTGSVGWAVV (SEQ ID NO:72)
22   DLGVGSVGWAIV (SEQ ID NO:73)
23   DLGIASIGWAII (SEQ ID NO:74)
24   DLGIASVGWAIV (SEQ ID NO:75)
25   DLGVASVGWSIV (SEQ ID NO:76)
26   DIGIASVGWAIL (SEQ ID NO:77)
28   DLGISSVGWSVI (SEQ ID NO:78)
32   DIGIASVGWSVI (SEQ ID NO:79)
33   DVGIGSIGWAVI (SEQ ID NO:80)
39   DLGVGSIGFAIV (SEQ ID NO:81)
34   DIGYASIGWAVI (SEQ ID NO:82)
47   DTGTNSLGWAIV (SEQ ID NO:83)
50   DLGTNSIGWCLL (SEQ ID NO:84)
49   DIGTDSLGWAVF (SEQ ID NO:85)
18   DIGSNSIGFAVV (SEQ ID NO:86)
41   DLGVGSIGVAVA (SEQ ID NO:87)
45   DLGIASCGWGVV (SEQ ID NO:88)
```

Fig. 3A

21	DLGIASVGWCLT	(SEQ ID NO:89)	
27	DIGIGSVGVGIL	(SEQ ID NO:90)	
29	DIGITSVGYGLI	(SEQ ID NO:91)	
30	DIGITSVGFGII	(SEQ ID NO:92)	
31	DVGITSTGYAVL	(SEQ ID NO:93)	
40	DLGITSFGYAIL	(SEQ ID NO:94)	
17	DIGNASVGWSAF	(SEQ ID NO:95)	
19	DVGTNSCGWVAM	(SEQ ID NO:96)	
35	DVGERSIGLAAV	(SEQ ID NO:97)	
36	DVGLNSVGLAAV	(SEQ ID NO:98)	
37	DVGLMSVGLAAI	(SEQ ID NO:99)	
38	DVGTFSVGLAAI	(SEQ ID NO:100)	
13	DIGTGSVGYACM	(SEQ ID NO:101)	
44	DLGTTSIGFAHI	(SEQ ID NO:102)	
46	DLGTNSIGSSVR	(SEQ ID NO:103)	
	* * * *		

Fig. 3B

Alignment of the N terminal RucV-like Domains disclosed in Chylinski *et al.*
(CLUSTAL format alignment by MAFFT (v7.058b))

1	D---IGTNSVGWAVT	(SEQ ID NO:104)
12	D---IGTNSVGWAVT	(SEQ ID NO:105)
3	D---VGTNSVGWAVT	(SEQ ID NO:106)
20	D---VGTNSVGWAVT	(SEQ ID NO:107)
15	D---MGTNSVGWAVT	(SEQ ID NO:108)
4	D---VGTSSVGWAVT	(SEQ ID NO:109)
7	D---IGTASVGWAVT	(SEQ ID NO:110)
6	D---VGTGSVGWAVT	(SEQ ID NO:111)
9	D---IGTNSVGWAVV	(SEQ ID NO:112)
10	D---IGTNSVGWAVI	(SEQ ID NO:113)
52	D---IGTNSIGWAVI	(SEQ ID NO:114)
11	D---IGTNSVGWAVL	(SEQ ID NO:115)
42	D---LGTNSIGWAVV	(SEQ ID NO:116)
48	D---LGTNSIGWAI-	(SEQ ID NO:117)
43	D---LGTNSIGWALV	(SEQ ID NO:118)
2	D---IGTNSVGWCVT	(SEQ ID NO:119)
14	D---IGTNSVGYAVT	(SEQ ID NO:120)
5	D---MGTGSLGWAVT	(SEQ ID NO:121)
16	D---IGTSSVGWAAI	(SEQ ID NO:122)
8	D---LGTGSVGWAVV	(SEQ ID NO:123)
22	D---LGVGSVGWAIV	(SEQ ID NO:124)
23	D---LGIASIGWAII	(SEQ ID NO:125)
24	D---LGIASVGWAIV	(SEQ ID NO:126)
68	D---LGIASVGWAVV	(SEQ ID NO:127)
25	D---LGVASVGWSIV	(SEQ ID NO:128)
26	D---IGIASVGWAIL	(SEQ ID NO:129)
66	D---IGIASVGWAVL	(SEQ ID NO:130)
59	D---IGIASIGWAVI	(SEQ ID NO:131)
61	D---IGIASVGWAII	(SEQ ID NO:132)
64	D---VGIASVGWAVI	(SEQ ID NO:133)
62	D---IGIASVGWAL-	(SEQ ID NO:134)
67	D---IGIASVGWAMV	(SEQ ID NO:135)
32	D---IGIASVGWSVI	(SEQ ID NO:136)
28	D---LGISSVGWSVI	(SEQ ID NO:137)
63	D---IGITSVGWAVI	(SEQ ID NO:138)

Fig. 4A

```
33  D----VGIGSIGWAVI (SEQ ID NO:139)
57  D----LGISSLGWAIV (SEQ ID NO:140)
39  D----LGVGSIGFAIV (SEQ ID NO:141)
34  D----IGYASIGWAVI (SEQ ID NO:142)
50  D----LGTNSIGWCLL (SEQ ID NO:143)
54  D----LGTNSIGWGLL (SEQ ID NO:144)
47  D----TGTNSLGWAIV (SEQ ID NO:145)
49  D----IGTDSLGWAVF (SEQ ID NO:146)
51  D----LGSTSLGWAIF (SEQ ID NO:147)
58  D----IGISSIGWAFS (SEQ ID NO:148)
21  D----LGIASVGWCLT (SEQ ID NO:149)
45  D----LGIASCGWGVV (SEQ ID NO:150)
18  D----IGSNSIGFAVV (SEQ ID NO:151)
65  D----IGTTSIGFSVI (SEQ ID NO:152)
29  D----IGITSVGYGLI (SEQ ID NO:153)
30  D----IGITSVGFGII (SEQ ID NO:154)
44  D----LGTTSIGFAHI (SEQ ID NO:155)
27  D----IGIGSVGVGIL (SEQ ID NO:156)
41  D----LGVGSIGVAVA (SEQ ID NO:157)
31  D----VGITSTGYAVL (SEQ ID NO:158)
40  D----LGITSFGYAIL (SEQ ID NO:159)
53  D----IGTSSIGWWLY (SEQ ID NO:160)
55  D----LGSNSLGWFVT (SEQ ID NO:161)
56  D----LGANSLGWFVV (SEQ ID NO:162)
17  D----IGNASVGWSAF (SEQ ID NO:163)
19  D----VGTNSCGWVAM (SEQ ID NO:164)
35  D----VGERSIGLAAV (SEQ ID NO:165)
36  D----VGLNSVGLAAV (SEQ ID NO:166)
37  D----VGLMSVGLAAI (SEQ ID NO:167)
38  D----VGTFSVGLAAI (SEQ ID NO:168)
13  D----IGTGSVGYACM (SEQ ID NO:169)
46  D----LGTNSIGSSVR (SEQ ID NO:170)
60  DIGLRIGITSCGWSI- (SEQ ID NO:171)
69  D----MGAKYTGVFYA (SEQ ID NO:172)
73  D----LGGKNTGFFSF (SEQ ID NO:173)
74  D----LGVKNTGVFSA (SEQ ID NO:174)
70  D----LGAKFTGVALY (SEQ ID NO:175)
71  D----LGGKFTGVCLS (SEQ ID NO:176)
72  D----LGGTYTGTFIT (SEQ ID NO:177)
    *    *    *
```

Fig. 4B

Alignment of the HNH-like Domains disclosed in Chylinski *et al.*
(CLUSTAL format alignment by MAFFT (v7.058b))

1	YDIDHIYPRS-LTKD------DSF-DNLVLCERTAN	(SEQ ID NO:178)
2	-DIDHIYPRSKVIKD------DSF-DNLVLVLKNEN	(SEQ ID NO:179)
3	-DRDHIYPQS-KIKD------DSI-DNLVLVNKTYN	(SEQ ID NO:180)
4	-DIDHIYPRS-KIKD------DSI-TNRVLVEKDIN	(SEQ ID NO:181)
6	-DIDHIYPQS-KIKD------DSI-SNRVLVCSSCN	(SEQ ID NO:182)
5	-DIDHIYPQS-KTMD------DSL-NNRVLVKKNYN	(SEQ ID NO:183)
7	-DQDHIYPKS-KIYD------DSL-ENRVLVKKNLN	(SEQ ID NO:184)
8	-QIDHIVPQS-LVKD------DSF-DNRVLVVPSEN	(SEQ ID NO:185)
9	-DIDHIIPQA-FIKD------NSI-DNRVLTSSKEN	(SEQ ID NO:186)
12	-DIDHIIPQA-FLKD------NSI-DNKVLVSSASN	(SEQ ID NO:187)
16	-DIDHIIPQA-YTKD------NSL-DNRVLVSNITN	(SEQ ID NO:188)
11	-DIDHIVPQS-FITD------NSI-DNLVLTSSAGN	(SEQ ID NO:189)
10	-DVDHIVPQS-FLKD------DSI-DNKVLTRSDKN	(SEQ ID NO:190)
14	-NIDHIYPQS-MVKD------DSL-DNKVLVQSEIN	(SEQ ID NO:191)
18	-DIDHILPQS-LIKD------DSL-DNRVLVNATIN	(SEQ ID NO:192)
19	-DIDHILPQS-FIKD------DSL-ENRVLVKKAVN	(SEQ ID NO:193)
13	-EVDHIFPRS-FIKD------DSI-DNKVLVIKKMN	(SEQ ID NO:194)
15	-EVDHIIPRS-YIKD------DSF-ENKVLVYREEN	(SEQ ID NO:195)
17	-DIDHIIPQA-VTQN------DSI-DNRVLVARAEN	(SEQ ID NO:196)
22	-EIDHIIPYS-ISFD------DSS-SNKLLVLAESN	(SEQ ID NO:197)
24	-EIDHIIPYS-LCFD------DSS-ANKVLVHKQSN	(SEQ ID NO:198)
32	-DIDHIIPYS-RSMD------DSY-SNKVLVLSGEN	(SEQ ID NO:199)
63	-DIDHIIPYS-KSMD------DSF-NNKVLCLAEEN	(SEQ ID NO:200)
59	-EIDHIYPYS-RSFD------DSY-MNKVLVFTKQN	(SEQ ID NO:201)
65	-QIDHIYPYS-RSMD------DSY-MNKVLVLTDEN	(SEQ ID NO:202)
64	-EIDHIIPFS-RSFD------DSL-SNKILVLGSEN	(SEQ ID NO:203)
68	-EIDHALPFS-RTWD------DSF-NNKVLVLGSEN	(SEQ ID NO:204)
69	-EIDHALPFS-RTWD------DSF-NNKVLVLASEN	(SEQ ID NO:205)
28	-EIDHIIPIS-ISLD------DSI-NNKVLVLSKAN	(SEQ ID NO:206)
30	-EVDHIIPIS-ISLD------DSI-TNKVLVTHREN	(SEQ ID NO:207)
62	-QVDHALPYS-RSYD------DSK-NNKVLVLTHEN	(SEQ ID NO:208)
27	-EVDHILPLS-ITFD------DSL-ANKVLVYATAN	(SEQ ID NO:209)
26	-EIDHIIPRS-ISFD------DAR-SNKVLVYRSEN	(SEQ ID NO:210)

Fig. 5A

29	-EVDHIIPRS-VSFD-----NSY-HNKVLVKQSEN	(SEQ ID NO:211)
67	-DIDHILPYS-ITFD-----DSF-RNKVLVTSQEN	(SEQ ID NO:212)
58	-EIDHILPRS-RSAD-----DSF-ANKVLCLARAN	(SEQ ID NO:213)
51	-EIEHLLPFS-LTLD-----DSM-ANKTVCFRQAN	(SEQ ID NO:214)
55	-DIDHILPFS-VSLD-----DSA-ANKVVCLREAN	(SEQ ID NO:215)
57	-DIDHLIPFS-ISWD-----DSA-ANKVVCMRYAN	(SEQ ID NO:216)
56	-DIDHILPVA-MTLD-----DSP-ANKIICMRYAN	(SEQ ID NO:217)
54	-DVDHILPYS-RTLD-----DSF-PNRTLCLREAN	(SEQ ID NO:218)
52	-EIEHILPFS-RTLD-----DSL-NNRTVAMRRAN	(SEQ ID NO:219)
31	-EVDHIIPYS-ISWD-----DSY-TNKVLTSAKCN	(SEQ ID NO:220)
45	-QVDHILPWS-RFGD-----DSY-LNKTLCTARSN	(SEQ ID NO:221)
53	-QVDHILPFS-KTLD-----DSF-ANKVLAQHDAN	(SEQ ID NO:222)
60	-QIDHAFPLS-RSLD-----DSQ-SNKVLCLTSSN	(SEQ ID NO:223)
21	-DIDHIVPRS-ISFD-----DSF-SNLVIVNKLDN	(SEQ ID NO:224)
23	-EIEHIIPYS-MSYD-----NSQ-ANKILTEKAEN	(SEQ ID NO:225)
25	-EIDHVIPYS-KSAD-----DSW-FNKLLVKKSTN	(SEQ ID NO:226)
49	-EMDHILPYS-RSLD-----NGW-HNRVLVHGKDN	(SEQ ID NO:227)
33	-EVDHIVPYS-LILD-----NTI-NNKALVYAEEN	(SEQ ID NO:228)
42	-EIEHVIPQS-LYFD-----DSF-SNKVICEAEVN	(SEQ ID NO:229)
43	-DIEHIIPQA-RLFD-----DSF-SNKTLEARSVN	(SEQ ID NO:230)
44	-EIEHIVPKA-RVFD-----DSF-SNKTLTFHRIN	(SEQ ID NO:231)
20	-DKDHIIPQS-MKKD-----DSIINNLVLVNKNAN	(SEQ ID NO:232)
66	-EVEHIWPRS-RSFD-----NSP-RNKTLCRKDVN	(SEQ ID NO:233)
61	-IVNHIIPYN-RSFD-----DTY-HNRVLTLTETK	(SEQ ID NO:234)
46	-DMEHTIPKS-ISFD-----NSD-QNLTLCESYYN	(SEQ ID NO:235)
47	-DIEHTIPRS-AGGD-----STK-MNLTLCSSRFN	(SEQ ID NO:236)
48	-DIEHTIPRS-ISQD-----NSQ-MNKTLCSLKFN	(SEQ ID NO:237)
50	-DIDHVIPLA-RGGR-----DSL-DNMVLCQSDAN	(SEQ ID NO:238)
39	-DIEHLFPIA-ESED-----NGR-NNLVISHSACN	(SEQ ID NO:239)
41	-DVDHIFPRD-DTAD-----NSY-GNKVVAHRQCN	(SEQ ID NO:240)
40	-DIEHIVPQS-LGGL-----STD-YNTIVTLKSVN	(SEQ ID NO:241)
35	-ELDHIVPRT-DGGS-----NRH-ENLAITCGACN	(SEQ ID NO:242)
36	-EMDHIVPRKGVGST-----NTR-TNFAAVCAECN	(SEQ ID NO:243)
37	-EMDHIVPRKGVGST-----NTR-VNLAAACAACN	(SEQ ID NO:244)
38	-EMDHIVPRAGQGST-----NTR-ENLVAVCHRCN	(SEQ ID NO:245)
70	-EIDHILPRS-LIKDARGIVFNAE-PNLIYASSRGN	(SEQ ID NO:246)
71	-EIDHIIPRS-LTGRTKKTVFNSE-ANLIYCSSKGN	(SEQ ID NO:247)
73	-EIDHIIPRS-LTLKKSESIYNSE-VNLIFVSAQGN	(SEQ ID NO:248)

Fig. 5B

```
72  -EIDHIYPRS-LSKKHFGVIFNSE-VNLIYCSSQGN   (SEQ ID NO:249)
74  -EIDHILPRS-HTLKIYGTVFNPE-GNLIYVHQKCN   (SEQ ID NO:250)
75  -ELDHIIPRS-HKKY---GTLNDE-ANLICVTRGDN   (SEQ ID NO:251)
34  -ELEHIVPHS-FRQS-----NAL-SSLVLTWPGVN   (SEQ ID NO:252)
        :*  *       .     .          :
```

Fig. 5C

Alignment of the HNH-like Domains disclosed in Chylinski *et al.* (excluding sequence outliers).
(CLUSTAL format alignment by MAFFT (v7.058b))

1	YDIDHIYPRS-LTKDDS-FDNLVLCERTAN	(SEQ ID NO:253)
2	-DIDHIYPRSKVIKDDS-FDNLVLVLKNEN	(SEQ ID NO:254)
3	-DRDHIYPQS-KIKDDS-IDNLVLVNKTYN	(SEQ ID NO:255)
4	-DIDHIYPRS-KIKDDS-ITNRVLVEKDIN	(SEQ ID NO:256)
6	-DIDHIYPQS-KIKDDS-ISNRVLVCSSCN	(SEQ ID NO:257)
5	-DIDHIYPQS-KTMDDS-LNNRVLVKKNYN	(SEQ ID NO:258)
7	-DQDHIYPKS-KIYDDS-LENRVLVKKNLN	(SEQ ID NO:259)
8	-QIDHIVPQS-LVKDDS-FDNRVLVVPSEN	(SEQ ID NO:260)
9	-DIDHIIPQA-FIKDNS-IDNRVLTSSKEN	(SEQ ID NO:261)
12	-DIDHIIPQA-FLKDNS-IDNKVLVSSASN	(SEQ ID NO:262)
16	-DIDHIIPQA-YTKDNS-LDNRVLVSNITN	(SEQ ID NO:263)
11	-DIDHIVPQS-FITDNS-IDNLVLTSSAGN	(SEQ ID NO:264)
10	-DVDHIVPQS-FLKDDS-IDNKVLTRSDKN	(SEQ ID NO:265)
14	-NIDHIYPQS-MVKDDS-LDNKVLVQSEIN	(SEQ ID NO:266)
18	-DIDHILPQS-LIKDDS-LDNRVLVNATIN	(SEQ ID NO:267)
19	-DIDHILPQS-FIKDDS-LENRVLVKKAVN	(SEQ ID NO:268)
13	-EVDHIFPRS-FIKDDS-IDNKVLVIKKMN	(SEQ ID NO:269)
15	-EVDHIIPRS-YIKDDS-FENKVLVYREEN	(SEQ ID NO:270)
17	-DIDHIIPQA-VTQNDS-IDNRVLVARAEN	(SEQ ID NO:271)
21	-DIDHIVPRS-ISFDDS-FSNLVIVNKLDN	(SEQ ID NO:272)
22	-EIDHIIPYS-ISFDDS-SSNKLLVLAESN	(SEQ ID NO:273)
24	-EIDHIIPYS-LCFDDS-SANKVLVHKQSN	(SEQ ID NO:274)
28	-EIDHIIPIS-ISLDDS-INNKVLVLSKAN	(SEQ ID NO:275)
30	-EVDHIIPIS-ISLDDS-ITNKVLVTHREN	(SEQ ID NO:276)
27	-EVDHILPLS-ITFDDS-LANKVLVYATAN	(SEQ ID NO:277)
26	-EIDHIIPRS-ISFDDA-RSNKVLVYRSEN	(SEQ ID NO:278)
29	-EVDHIIPRS-VSFDNS-YHNKVLVKQSEN	(SEQ ID NO:279)
31	-EVDHIIPYS-ISWDDS-YTNKVLTSAKCN	(SEQ ID NO:280)
32	-DIDHIIPYS-RSMDDS-YSNKVLVLSGEN	(SEQ ID NO:281)
23	-EIEHIIPYS-MSYDNS-QANKILTEKAEN	(SEQ ID NO:282)
33	-EVDHIVPYS-LILDNT-INNKALVYAEEN	(SEQ ID NO:283)
25	-EIDHVIPYS-KSADDS-WFNKLLVKKSTN	(SEQ ID NO:284)
49	-EMDHILPYS-RSLDNG-WHNRVLVHGKDN	(SEQ ID NO:285)
42	-EIEHVIPQS-LYFDDS-FSNKVICEAEVN	(SEQ ID NO:286)
43	-DIEHIIPQA-RLFDDS-FSNKTLEARSVN	(SEQ ID NO:287)

Fig. 6A

```
44 -EIEHIVPKA-RVFDDS-FSNKTLTFHRIN (SEQ ID NO:288)
20 -DKDHIIPQS-MKKDDSIINNLVLVNKNAN (SEQ ID NO:289)
45 -QVDHILPWS-RFGDDS-YLNKTLCTARSN (SEQ ID NO:290)
50 -DIDHVIPLA-RGGRDS-LDNMVLCQSDAN (SEQ ID NO:291)
46 -DMEHTIPKS-ISFDNS-DQNLTLCESYYN (SEQ ID NO:292)
47 -DIEHTIPRS-AGGDST-KMNLTLCSSRFN (SEQ ID NO:293)
48 -DIEHTIPRS-ISQDNS-QMNKTLCSLKFN (SEQ ID NO:294)
39 -DIEHLFPIA-ESEDNG-RNNLVISHSACN (SEQ ID NO:295)
41 -DVDHIFPRD-DTADNS-YGNKVVAHRQCN (SEQ ID NO:296)
40 -DIEHIVPQS-LGGLST-DYNTIVTLKSVN (SEQ ID NO:297)
35 -ELDHIVPRT-DGGSNR-HENLAITCGACN (SEQ ID NO:298)
36 -EMDHIVPRKGVGSTNT-RTNFAAVCAECN (SEQ ID NO:299)
37 -EMDHIVPRKGVGSTNT-RVNLAAACAACN (SEQ ID NO:300)
38 -EMDHIVPRAGQGSTNT-RENLVAVCHRCN (SEQ ID NO:301)
34 -ELEHIVPHS-FRQSNA-LSSLVLTWPGVN (SEQ ID NO:302)
     : :*  *       .       .     *
```

Fig. 6B

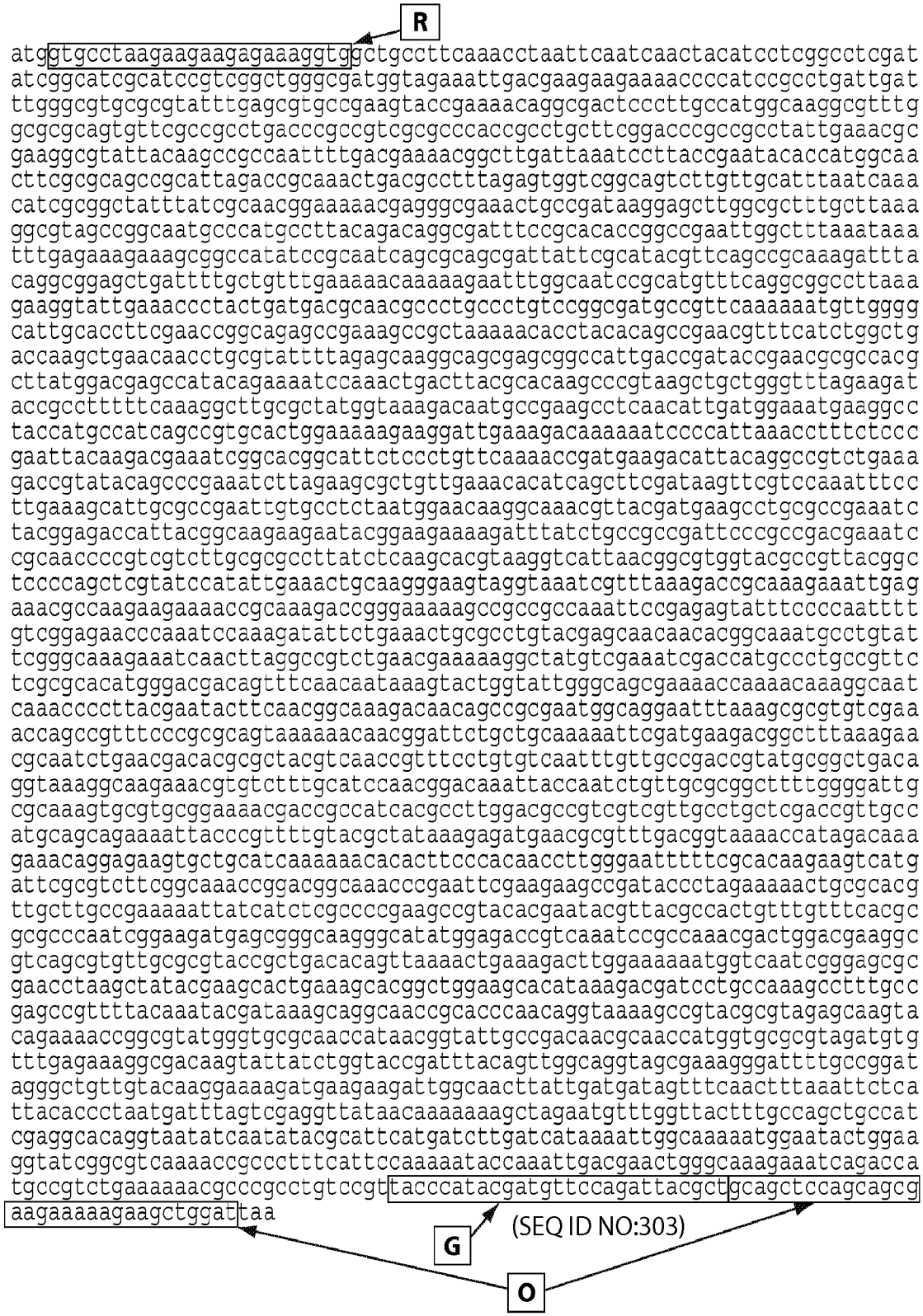

Sequence alignment between SpCas9 and NmCas9
Y
B
G
NmCas9 MAAFKPNSINYILGLDIGIASVGWAMVEIDEEENPIRLID----------LGVRVFE
SpCas9 ------MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFD
NmCas9 RAEVPKTGDSLAMARRLARSVRRLTRRRAHRLLRTRRLLKREGVLQAA----------
SpCas9 SGET-------AEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLV
NmCas9 -------------NFDENGLIKSLPNTPWQLRAAALDRK---LTPLEWSAVLLHLIKHR
SpCas9 EEDKKHERHPIFGNIVDEVAYHEKYP-TIYHLRKKLVDSTDKADLRLI-YLALAHMIKFR
NmCas9 GYLSQRKNE----------GETA----------DKEL---
SpCas9 GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENL
NmCas9 -----GALLKGVAGNAHALQTG---DFRTPAE------LAL--NKFEKESGHIRNQ-RSD
SpCas9 IAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQ
NmCas9 YSHTFSR----------KDLQA
SpCas9 YADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
NmCas9 ELILLFEKQKEFGN-PHVSGGLK----------EGIETL---------LMTQRPA
SpCas9 KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF
NmCas9 LSGDAV-QKMLGH--------CTFEPAEP----------KAAKNTYTAERFIWL
SpCas9 DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWM
NmCas9 TKLNNLRILEQGSERPLTD-------TERATLMDEPY------RKSKLTYAQAR------
SpCas9 TRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELT
NmCas9 ----KLLGLEDTAFFKGLRY--------GKDN----------AEA
SpCas9 KVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED
NmCas9 STLMEMKAYHAISRALEKEGLKDKKSPLNLSPELQDEIGTAFSLFKTDEDITGRLKDRIQ
SpCas9 RFNASLGTYHDLLKIIKDKDFLDNEE----NEDILEDIVLTLTLFEDREMIEERLKTYAH
NmCas9 P--EILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAE----IYGDHYGKKNT
SpCas9 LFDDKVMKQLKRRRYTGWGRLSRKLI--------NGIRDKQSGKTILDFLKSDGFANRNF
NmCas9 EEKI------Y----------LPPIPADEIRNPVVLRALSQARKVINGVVRRYG-
SpCas9 MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR
NmCas9 -SPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNF----VGEPKSK
SpCas9 HKPENIVIEMARENQTTQKGQKNSRER-------MKRIEEGIKELGSQILKEHPVENTQL
NmCas9 DILKLRLYEQQHGKCLYSGKEINLGRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSEN
SpCas9 QNEKLYLYYLQNGRDMYVDQELDINRLSD---YDVDHIVPQSFLKDDSIDNKVLTRSDKN
NmCas9 QNKGNQTPYEYFNGKDNSREWQEFKA-RVET-SRFP-RSKKQRILLQKFDEDGFKERNLN
SpCas9 RGKSDNVPSEEVVKKM-KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLV

Fig. 7A

Percent Identity Matrix - created by Clustal2.1

Fig. 7B

Sequence of the NmCas9 ORF with dual NLS and HA tags



R: SV40 NLS, G: HA tag, O: synthetic NLS (1); all else NmCas9

Fig. 8

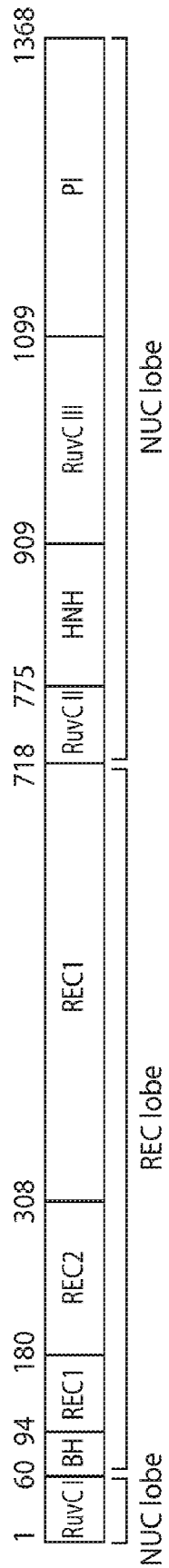

1
60
94
180
308
718
775
909
1099
1368
RuvC I
BH
REC1
REC2
REC1
RuvC II
HNH
RuvC III
PI
NUC lobe
REC lobe
NUC lobe

Fig. 9A

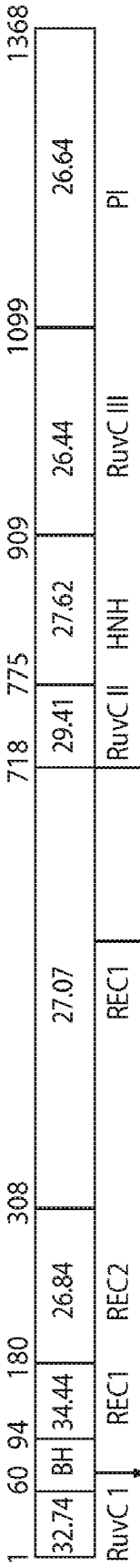

1
60
94
180
308
718
775
909
1099
1368
32.74
BH
34.44
26.84
27.07
29.41
27.62
26.44
26.64
RuvC 1
33.47
REC1
REC2
REC1
500
26.73%
718
RuvC II
HNH
RuvC III
PI

Fig. 9B

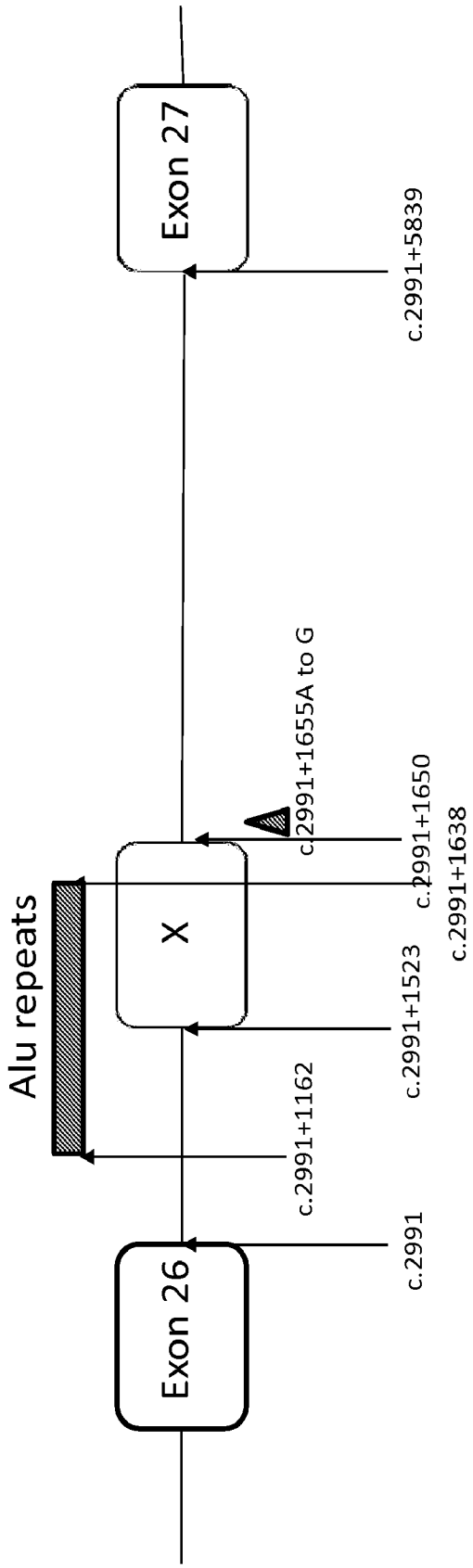

Exon 27
c.2991+5839
c.2991+1655A to G
c.2991+1650
c.2991+1638
X
Alu repeats
c.2991+1523
c.2991+1162
c.2991
Exon 26

Fig. 10

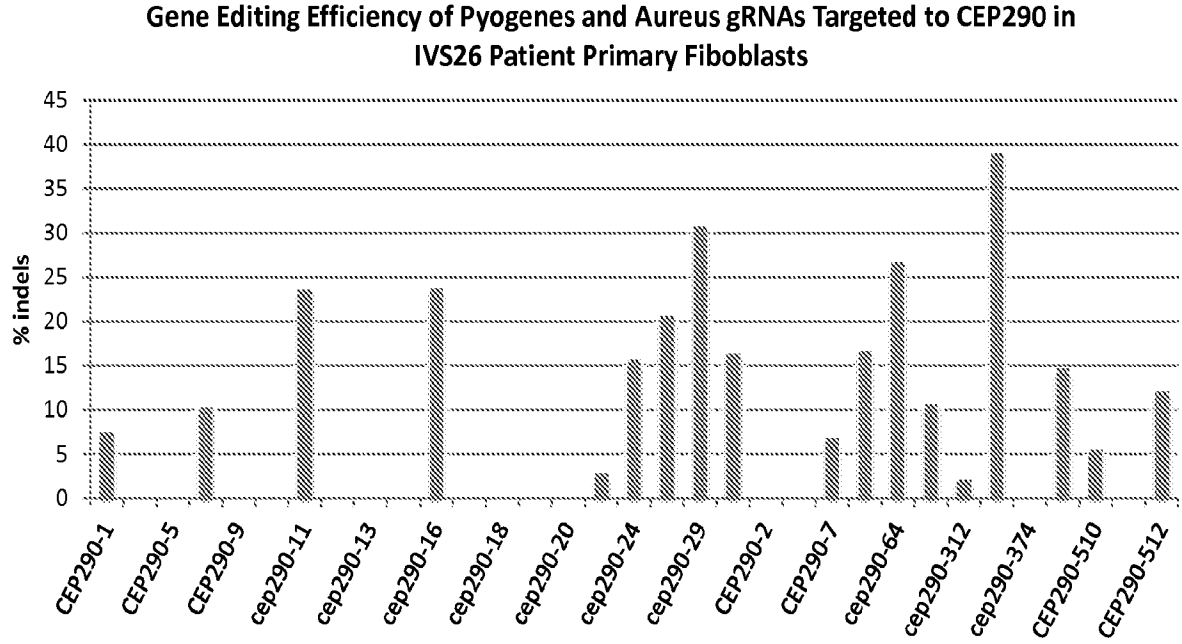

Gene Editing Efficiency of Pyogenes and Aureus gRNAs Targeted to CEP290 in IVS26 Patient Primary Fiboblasts
% indels
45
40
35
30
25
20
15
10
5
0
CEP290-1
CEP290-5
CEP290-9
cep290-11
cep290-13
cep290-16
cep290-18
cep290-20
cep290-24
cep290-29
CEP290-2
CEP290-7
cep290-64
cep290-312
cep290-374
CEP290-510
CEP290-512

Fig. 11A

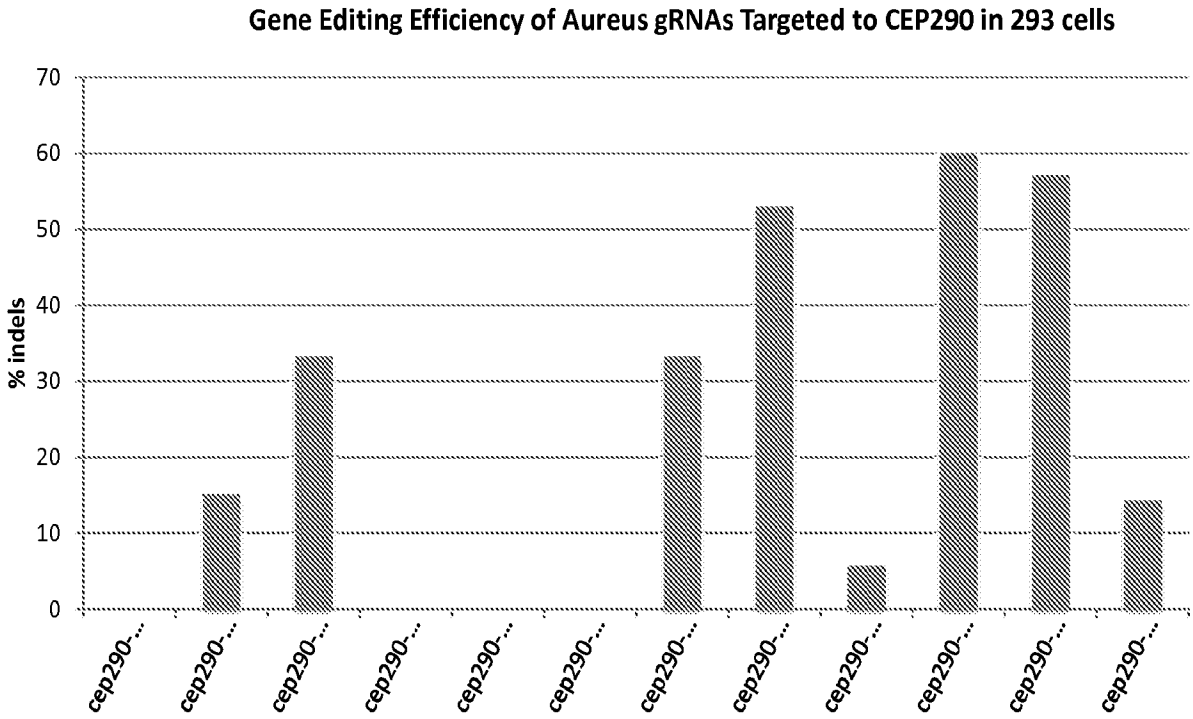

Gene Editing Efficiency of Aureus gRNAs Targeted to CEP290 in 293 cells
% indels
70
60
50
40
30
20
10
0
cep290-...
cep290-...
cep290-...
cep290-...
cep290-...
cep290-...
cep290-...
cep290-...
cep290-...
cep290-...
cep290-...
cep290-...

Fig. 11B

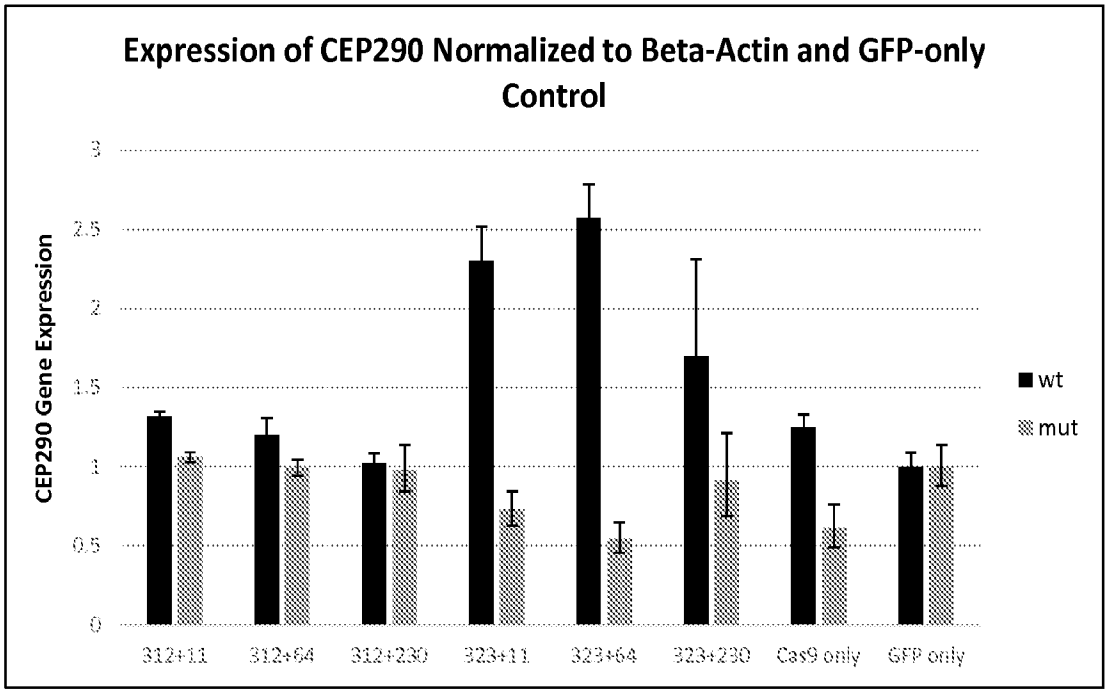

Expression of CEP290 Normalized to Beta-Actin and GFP-only Control
CEP290 Gene Expression
3
2.5
2
1.5
1
0.5
0
312+11
312+64
312+230
323+11
323+64
323+230
Cas9 only
GFP only
wt
mut

Fig. 12A

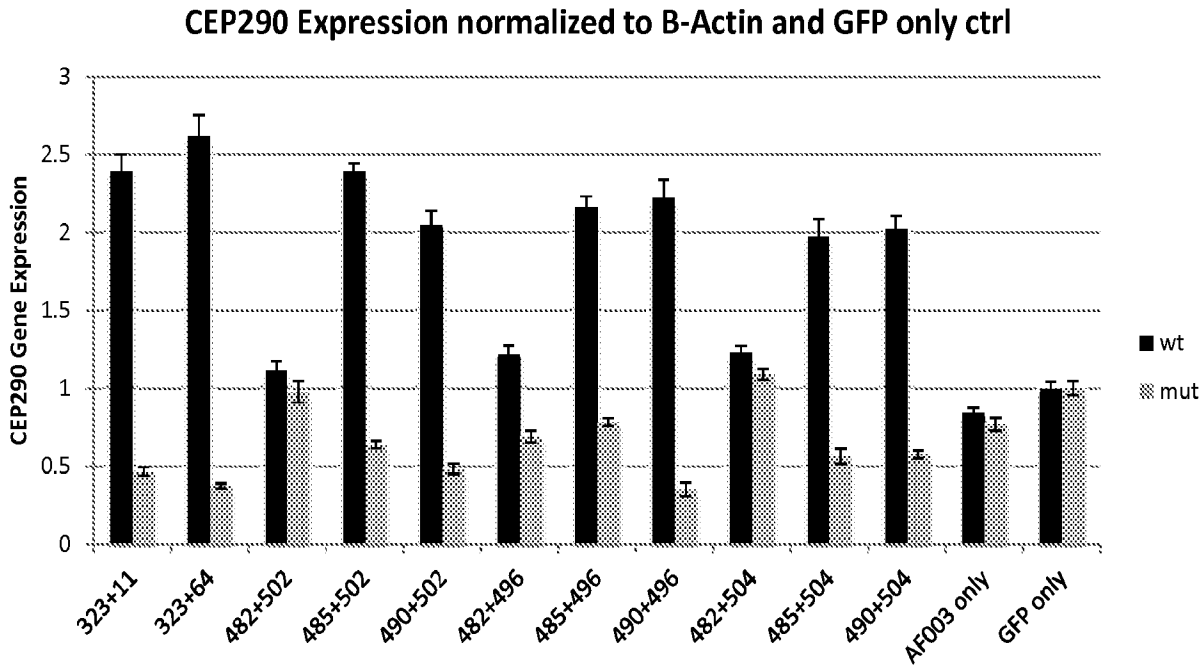

CEP290 Expression normalized to B-Actin and GFP only ctrl
CEP290 Gene Expression
3
2.5
2
1.5
1
0.5
0
323+11
323+64
482+502
485+502
490+502
482+496
485+496
490+496
482+504
485+504
490+504
AF003 only
GFP only
wt
mut

Fig. 12B

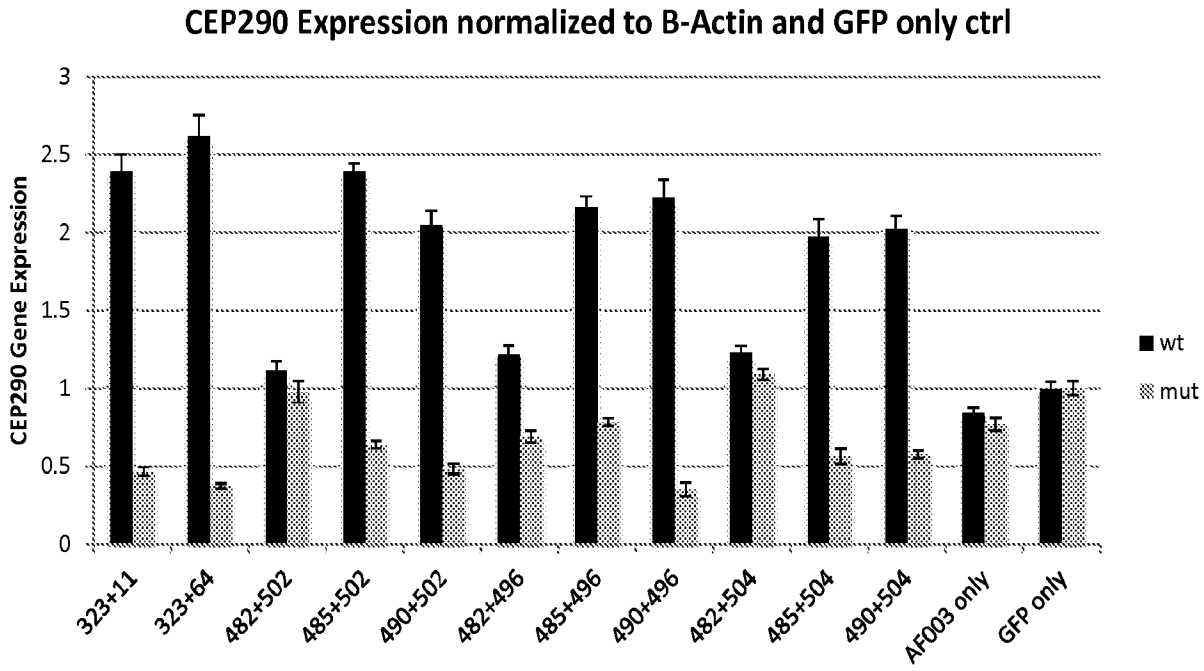

CEP290 Expression normalized to B-Actin and GFP only ctrl
CEP290 Gene Expression
3
2.5
2
1.5
1
0.5
0
323+11
323+64
482+502
485+502
490+502
482+496
485+496
490+496
482+504
485+504
490+504
AF003 only
GFP only
wt
mut

Fig. 13

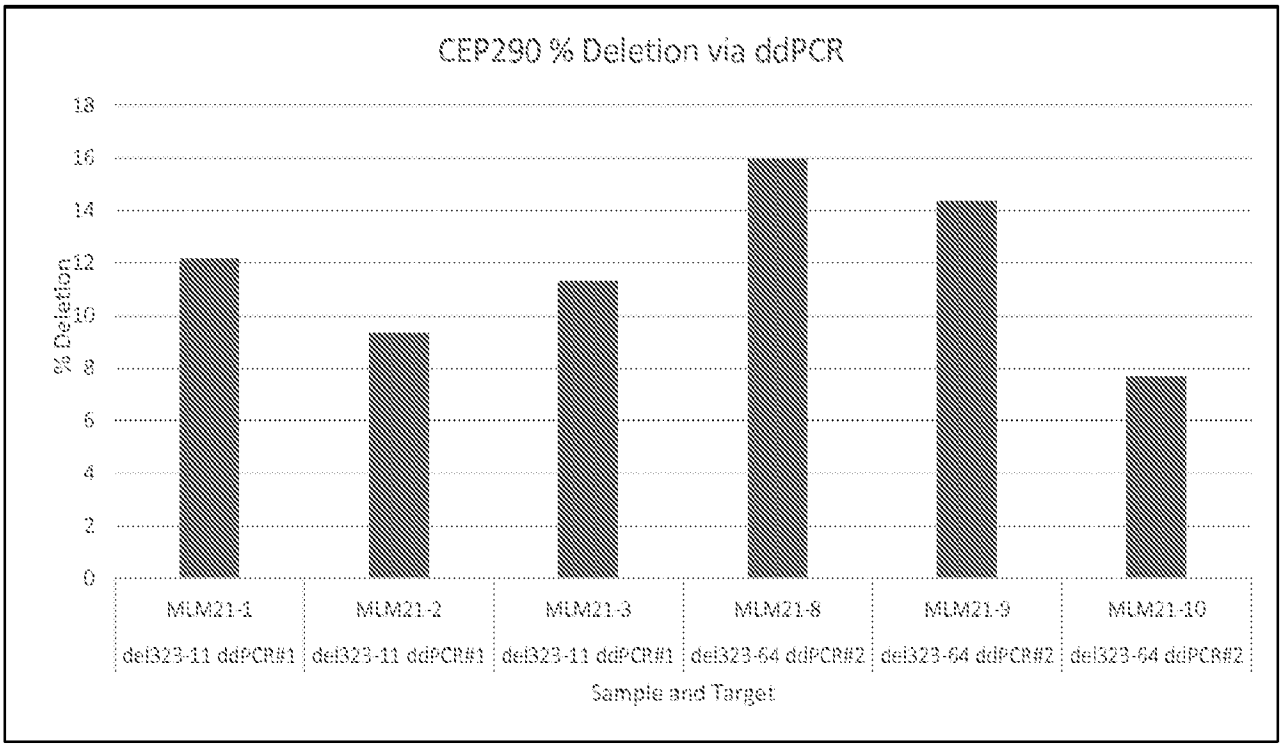

CEP290 % Deletion via ddPCR
% Deletion
18
16
14
12
10
8
6
4
2
0
MLM21-1
del323-11 ddPCR#1
MLM21-2
del323-11 ddPCR#1
MLM21-3
del323-11 ddPCR#1
MLM21-8
del323-64 ddPCR#2
MLM21-9
del323-64 ddPCR#2
MLM21-10
del323-64 ddPCR#2
Sample and Target

Fig. 14

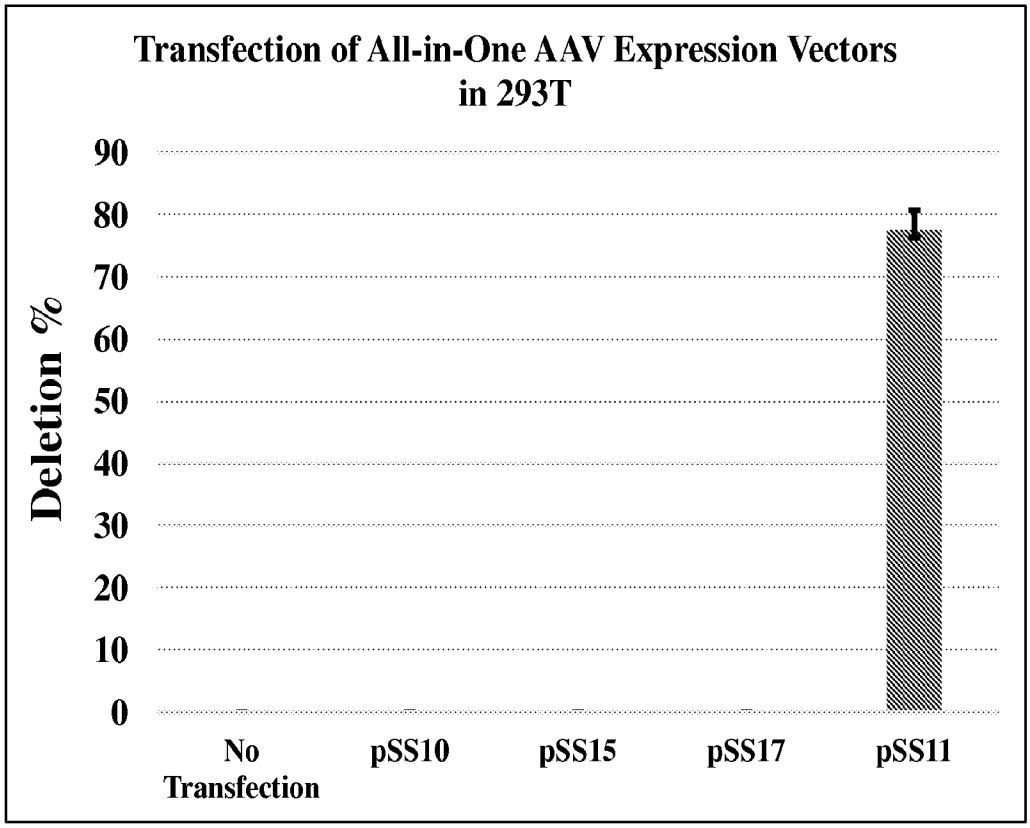

Transfection of All-in-One AAV Expression Vectors in 293T
Deletion %
90
80
70
60
50
40
30
20
10
0
No Transfection
pSS10
pSS15
pSS17
pSS11

Fig. 15

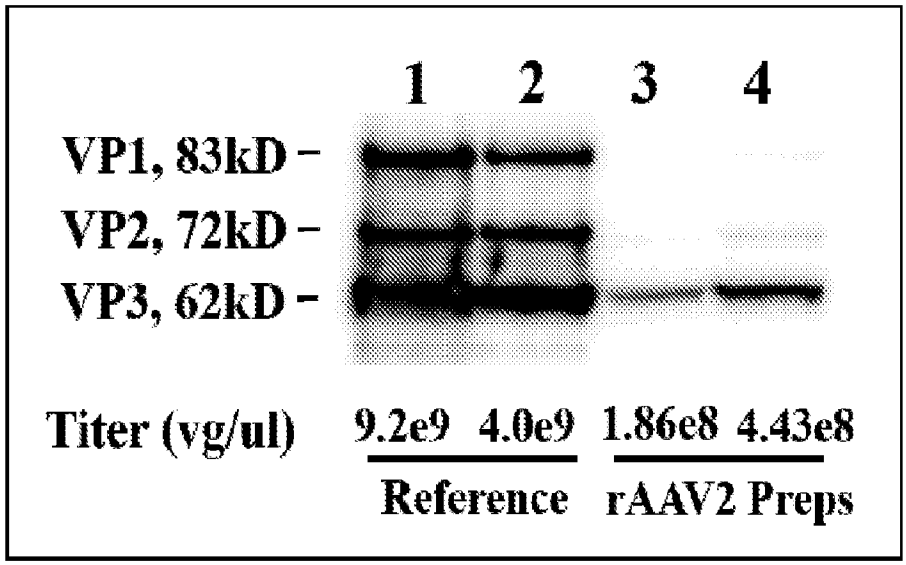

1
2
3
4
VP1, 83kD –
VP2, 72kD –
VP3, 62kD –
Titer (vg/ul)
9.2e9
4.0e9
1.86e8
4.43e8
Reference
rAAV2 Preps

Fig. 16

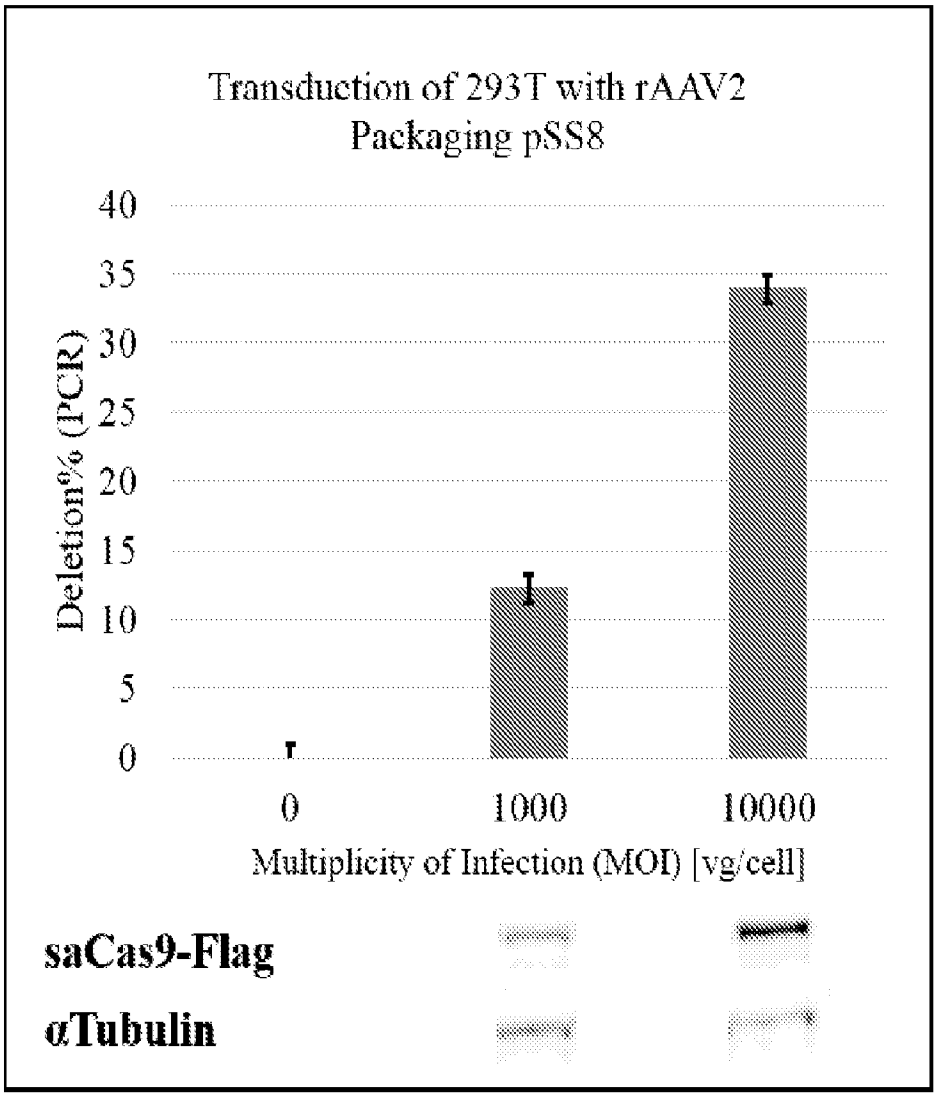

Transduction of 293T with rAAV2
Packaging pSS8
Deletion% (PCR)
40
35
30
25
20
15
10
5
0
0
1000
10000
Multiplicity of Infection (MOI) [vg/cell]
saCas9-Flag
αTubulin

Fig. 17

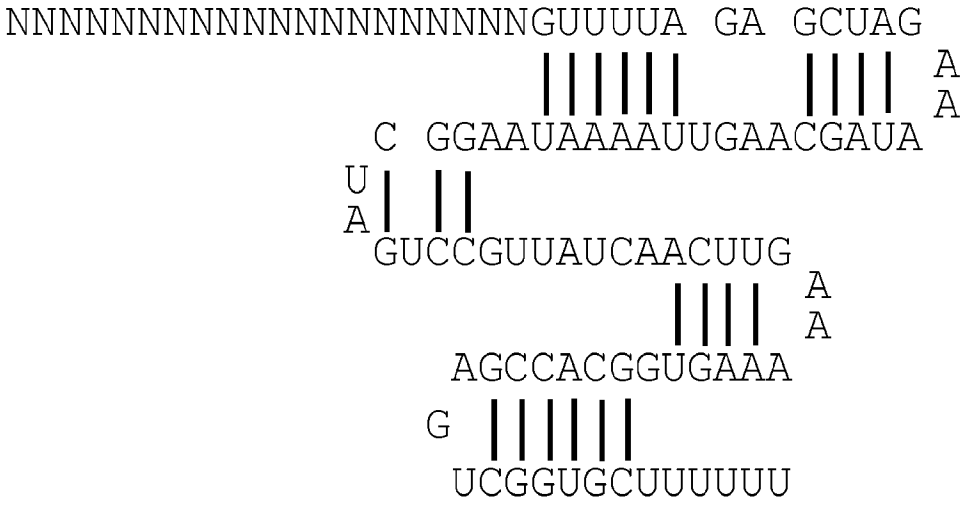


Fig. 18A

Fig. 18B

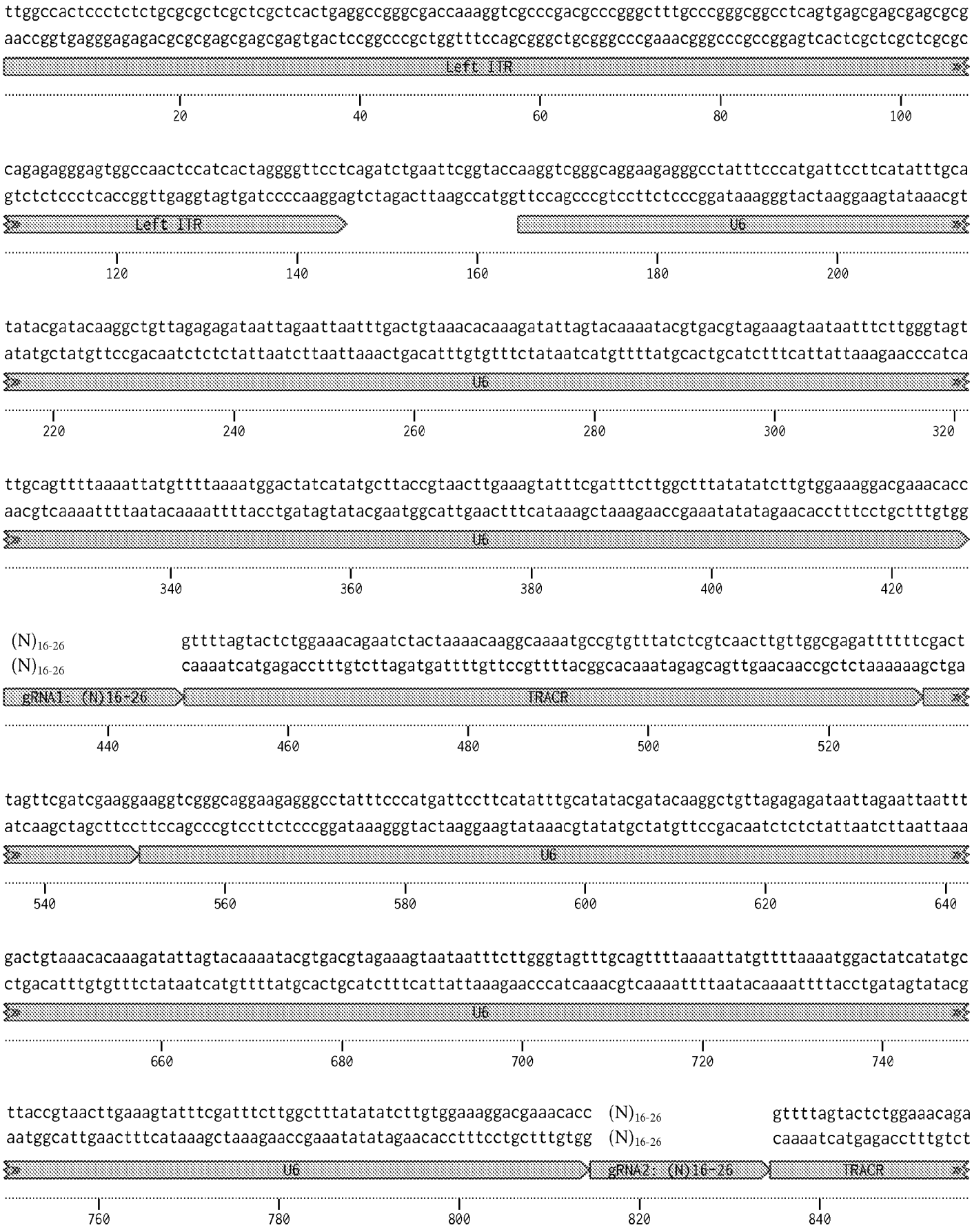

ttggccactccctctctgcgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggcctcagtgagcgagcgagcgcg
aaccggtgagggagagacgcgcgagcgagcgagtgactccggcccgctggtttccagcgggctgcgggcccgaaacgggcccgccggagtcactcgctcgctcgcgc
Left ITR
20
40
60
80
100
cagagagggagtggccaactccatcactaggggttcctcagatctgaattcggtaccaaggtcgggcaggaagagggcctatttcccatgattccttcatatttgca
gtctctccctcaccggttgaggtagtgatccccaaggagtctagacttaagccatggttccagcccgtccttctcccggataaagggtactaaggaagtataaacgt
Left ITR
U6
120
140
160
180
200
tatacgatacaaggctgttagagagataattagaattaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagt
atatgctatgttccgacaatctctctattaatcttaattaaactgacatttgtgtttctataatcatgttttatgcactgcatctttcattattaaagaacccatca
U6
220
240
260
280
300
320
ttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacacc
aacgtcaaaattttaatacaaaattttacctgatagtatacgaatggcattgaactttcataaagctaaagaaccgaaatatatagaacacctttcctgctttgtgg
U6
340
360
380
400
420
(N)16-26 gttttagtactctggaaacagaatctactaaaacaaggcaaaatgccgtgtttatctcgtcaacttgttggcgagatttttcgact
(N)16-26 caaaatcatgagacctttgtcttagatgattttgttccgttttacggcacaaatagagcagttgaacaaccgctctaaaaaagctga
gRNA1: (N)16-26
TRACR
440
460
480
500
520
tagttcgatcgaaggaaggtcgggcaggaagagggcctatttcccatgattccttcatatttgcatatacgatacaaggctgttagagagataattagaattaatttt
atcaagctagcttccttccagcccgtccttctcccggataaagggtactaaggaagtataaacgtatatgctatgttccgacaatctctctattaatcttaattaaa
U6
540
560
580
600
620
640
gactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgttttaaaatggactatcatatgc
ctgacatttgtgtttctataatcatgttttatgcactgcatctttcattattaaagaacccatcaaacgtcaaaattttaatacaaaattttacctgatagtatacg
U6
660
680
700
720
740
ttaccgtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacacc (N)16-26 gttttagtactctggaaacaga
aatggcattgaactttcataaagctaaagaaccgaaatatatagaacacctttcctgctttgtgg (N)16-26 caaaatcatgagacctttgtct
U6
gRNA2: (N)16-26
TRACR
760
780
800
820
840

FIG. 19A

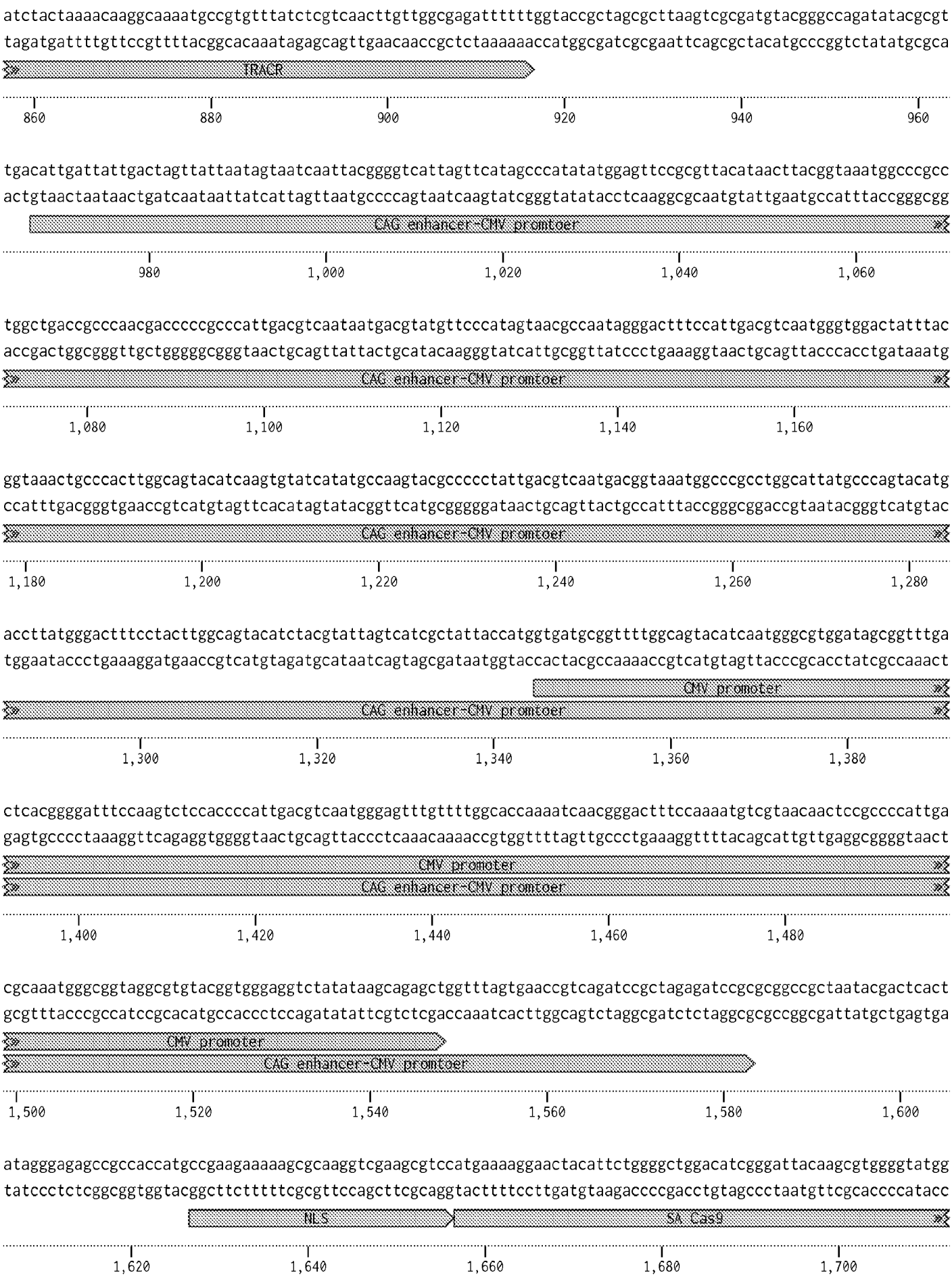
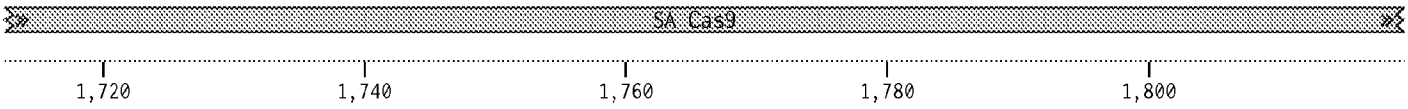
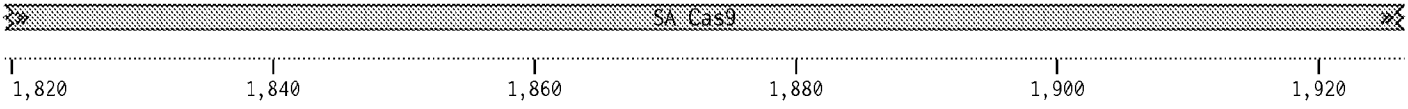

atctactaaaacaaggcaaaatgccgtgtttatctcgtcaacttgttggcgagatttttttggtaccgctagcgcttaagtcgcgatgtacgggccagatatacgcgt
tagatgattttgttccgttttacggcacaaatagagcagttgaacaaccgctctaaaaaaccatggcgatcgcgaattcagcgctacatgcccggtctatatgcgca
TRACR
860
880
900
920
940
960
tgacattgattattgactagttattaatagtaatcaattacggggtcattagttcatagcccatatatggagttccgcgttacataacttacggtaaatggcccgcc
actgtaactaataactgatcaataattatcattagttaatgccccagtaatcaagtatcgggtatatacctcaaggcgcaatgtattgaatgccatttaccgggcgg
CAG enhancer-CMV promtoer
980
1,000
1,020
1,040
1,060
tggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttac
accgactggcgggttgctgggggcgggtaactgcagttattactgcatacaagggtatcattgcggttatccctgaaaggtaactgcagttacccacctgataaatg
CAG enhancer-CMV promtoer
1,080
1,100
1,120
1,140
1,160
ggtaaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacggtaaatggcccgcctggcattatgcccagtacatg
ccatttgacgggtgaaccgtcatgtagttcacatagtatacggttcatgcgggggataactgcagttactgccatttaccgggcggaccgtaatacgggtcatgtac
CAG enhancer-CMV promtoer
1,180
1,200
1,220
1,240
1,260
1,280
accttatgggactttcctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttggcagtacatcaatgggcgtggatagcggtttga
tggaataccctgaaaggatgaaccgtcatgtagatgcataatcagtagcgataatggtaccactacgccaaaaccgtcatgtagttacccgcacctatcgccaaact
CMV promoter
CAG enhancer-CMV promtoer
1,300
1,320
1,340
1,360
1,380
ctcacggggatttccaagtctccaccccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtaacaactccgccccattga
gagtgcccctaaaggttcagaggtggggtaactgcagttaccctcaaacaaaaccgtggttttagttgccctgaaaggttttacagcattgttgaggcggggtaact
CMV promoter
CAG enhancer-CMV promtoer
1,400
1,420
1,440
1,460
1,480
cgcaaatgggcggtaggcgtgtacggtgggaggtctatataagcagagctggtttagtgaaccgtcagatccgctagagatccgcgcggccgctaatacgactcact
gcgtttacccgccatccgcacatgccaccctccagatatattcgtctcgaccaaatcacttggcagtctaggcgatctctaggcgcgccggcgattatgctgagtga
CMV promoter
CAG enhancer-CMV promtoer
1,500
1,520
1,540
1,560
1,580
1,600
atagggagagccgccaccatgccgaagaaaaagcgcaaggtcgaagcgtccatgaaaaggaactacattctggggctggacatcgggattacaagcgtggggtatgg
tatccctctcggcggtggtacggcttctttttcgcgttccagcttcgcaggtacttttccttgatgtaagaccccgacctgtagccctaatgttcgcaccccatacc
NLS
SA Cas9
1,620
1,640
1,660
1,680
1,700

FIG. 19B

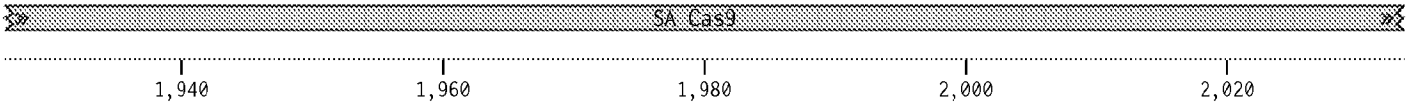
gattattgactatgaaacaagggacgtgatcgacgcaggcgtcagactgttcaaggaggccaacgtggaaaacaatgagggacggagaagcaagaggggagccaggc
ctaataactgatactttgttccctgcactagctgcgtccgcagtctgacaagttcctccggttgcacctttgttactccctgcctcttcgttctcccctcggtccg



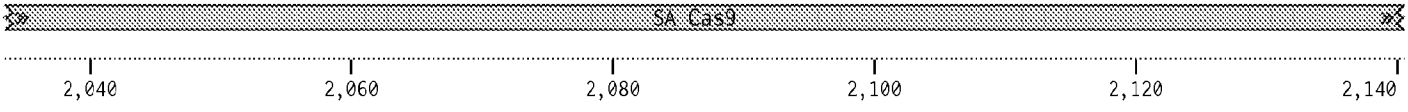
gcctgaaacgacggagaaggcacagaatccagagggtgaagaaactgctgttcgattacaacctgctgaccgaccattctgagctgagtggaattaatccttatgaa
cggactttgctgcctcttccgtgtcttaggtctcccacttctttgacgacaagctaatgttggacgactggctggtaagactcgactcaccttaattaggaatactt



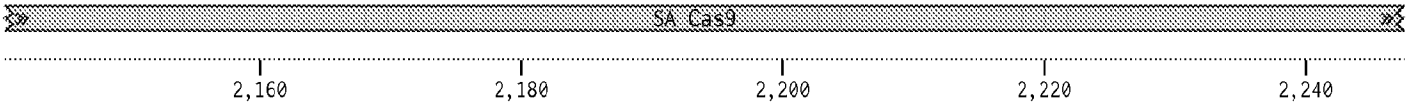
gccagggtgaaaggcctgagtcagaagctgtcagaggaagagttttccgcagctctgctgcacctggctaagcgccgaggagtgcataacgtcaatgaggtggaaga
cggtcccactttccggactcagtcttcgacagtctccttctcaaaaggcgtcgagacgacgtggaccgattcgcggctcctcacgtattgcagttactccaccttct



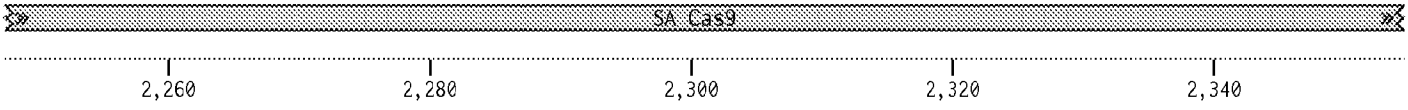
ggacaccggcaacgagctgtctacaaaggaacagatctcacgcaatagcaaagctctggaagagaagtatgtcgcagagctgcagctggaacggctgaagaaagatg
cctgtggccgttgctcgacagatgtttccttgtctagagtgcgttatcgtttcgagaccttctcttcatacagcgtctcgacgtcgaccttgccgacttctttctac



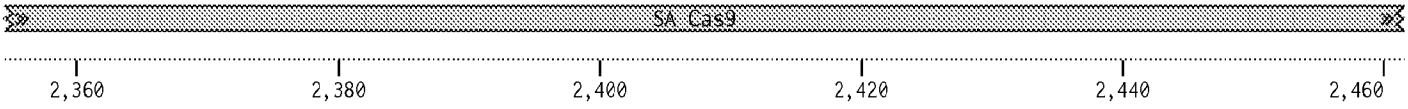
gcgaggtgagagggtcaattaataggttcaagacaagcgactacgtcaaagaagccaagcagctgctgaaagtgcagaaggcttaccaccagctggatcagagcttc
cgctccactctcccagttaattatccaagttctgttcgctgatgcagtttcttcggttcgtcgacgactttcacgtcttccgaatggtggtcgacctagtctcgaag



atcgatacttatatcgacctgctggagactcggagaacctactatgagggaccaggagaagggagccccttcggatggaaagacatcaaggaatggtacgagatgct
tagctatgaatatagctggacgacctctgagcctcttggatgatactccctggtcctcttccctcggggaagcctacctttctgtagttccttaccatgctctacga



gatgggacattgcacctattttccagaagagctgagaagcgtcaagtacgcttataacgcagatctgtacaacgccctgaatgacctgaacaacctggtcatcacca
ctaccctgtaacgtggataaaaggtcttctcgactcttcgcagttcatgcgaatattgcgtctagacatgttgcgggacttactggacttgttggaccagtagtggt



gggatgaaaacgagaaactggaatactatgagaagttccagatcatcgaaaacgtgtttaagcagaagaaaaagcctacactgaaacagattgctaaggagatcctg
ccctacttttgctctttgaccttatgatactcttcaaggtctagtagcttttgcacaaattcgtcttctttttcggatgtgactttgtctaacgattcctctaggac

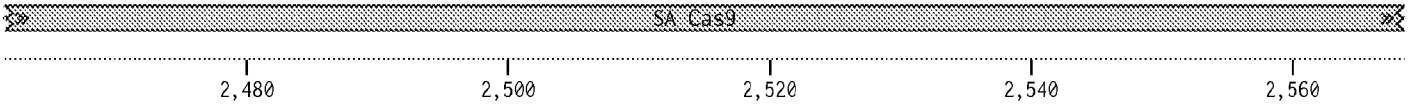


FIG. 19C

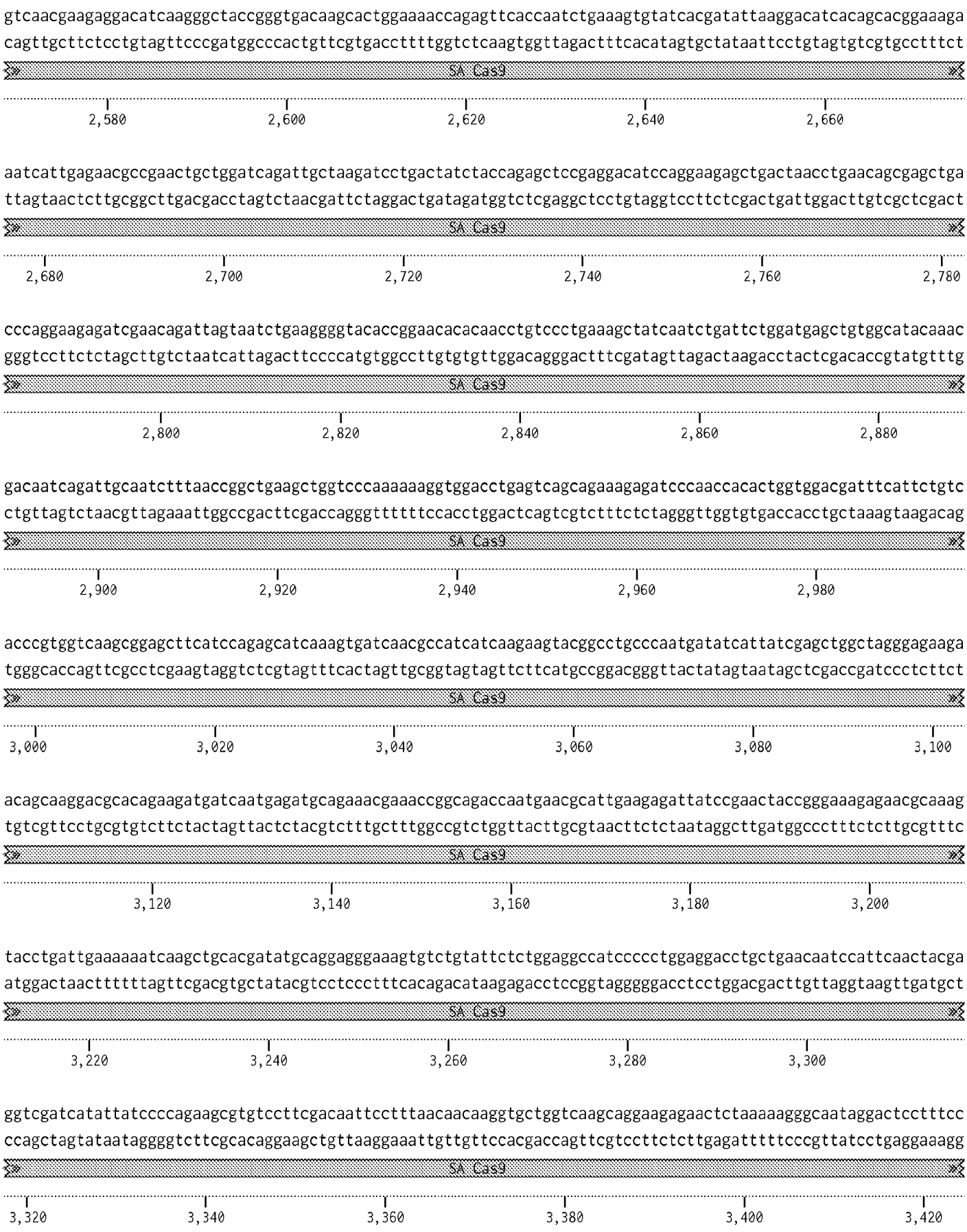

gtcaacgaagaggacatcaagggctaccgggtgacaagcactggaaaaccagagttcaccaatctgaaagtgtatcacgatattaaggacatcacagcacggaaaga
cagttgcttctcctgtagttcccgatggcccactgttcgtgacctttggtctcaagtggttagactttcacatagtgctataattcctgtagtgtcgtgcctttct
SA Cas9
2,580
2,600
2,620
2,640
2,660
aatcattgagaacgccgaactgctggatcagattgctaagatcctgactatctaccagagctccgaggacatccaggaagagctgactaacctgaacagcgagctga
ttagtaactcttgcggcttgacgacctagtctaacgattctaggactgatagatggtctcgaggctcctgtaggtccttctcgactgattggacttgtcgctcgact
SA Cas9
2,680
2,700
2,720
2,740
2,760
2,780
cccaggaagagatcgaacagattagtaatctgaaggggtacaccggaacacacaacctgtccctgaaagctatcaatctgattctggatgagctgtggcatacaaac
gggtccttctctagcttgtctaatcattagacttccccatgtggccttgtgtgttggacagggactttcgatagttagactaagacctactcgacaccgtatgtttg
SA Cas9
2,800
2,820
2,840
2,860
2,880
gacaatcagattgcaatctttaaccggctgaagctggtcccaaaaaaggtggacctgagtcagcagaaagagatcccaaccacactggtggacgatttcattctgtc
ctgttagtctaacgttagaaattggccgacttcgaccagggtttttccacctggactcagtcgtctttctctagggttggtgtgaccacctgctaaagtaagacag
SA Cas9
2,900
2,920
2,940
2,960
2,980
acccgtggtcaagcggagcttcatccagagcatcaaagtgatcaacgccatcatcaagaagtacggcctgcccaatgatatcattatcgagctggctagggagaaga
tgggcaccagttcgcctcgaagtaggtctcgtagtttcactagttgcggtagtagttcttcatgccggacgggttactatagtaatagctcgaccgatccctcttct
SA Cas9
3,000
3,020
3,040
3,060
3,080
3,100
acagcaaggacgcacagaagatgatcaatgagatgcagaaacgaaaccggcagaccaatgaacgcattgaagagattatccgaactaccgggaaagagaacgcaaag
tgtcgttcctgcgtgtcttctactagttactctacgtctttgctttggccgtctggttacttgcgtaacttctctaataggcttgatggccctttctcttgcgtttc
SA Cas9
3,120
3,140
3,160
3,180
3,200
tacctgattgaaaaaatcaagctgcacgatatgcaggagggaaagtgtctgtattctctggaggccatccccctggaggacctgctgaacaatccattcaactacga
atggactaacttttttagttcgacgtgctatacgtcctccctttcacagacataagagacctccggtagggggacctcctggacgacttgttaggtaagttgatgct
SA Cas9
3,220
3,240
3,260
3,280
3,300
ggtcgatcatattatccccagaagcgtgtccttcgacaattcctttaacaacaaggtgctggtcaagcaggaagagaactctaaaaagggcaataggactcctttcc
ccagctagtataatagggtcttcgcacaggaagctgttaaggaaattgttgttccacgaccagttcgtccttctcttgagatttttcccgttatcctgaggaaagg
SA Cas9
3,320
3,340
3,360
3,380
3,400
3,420

FIG. 19D

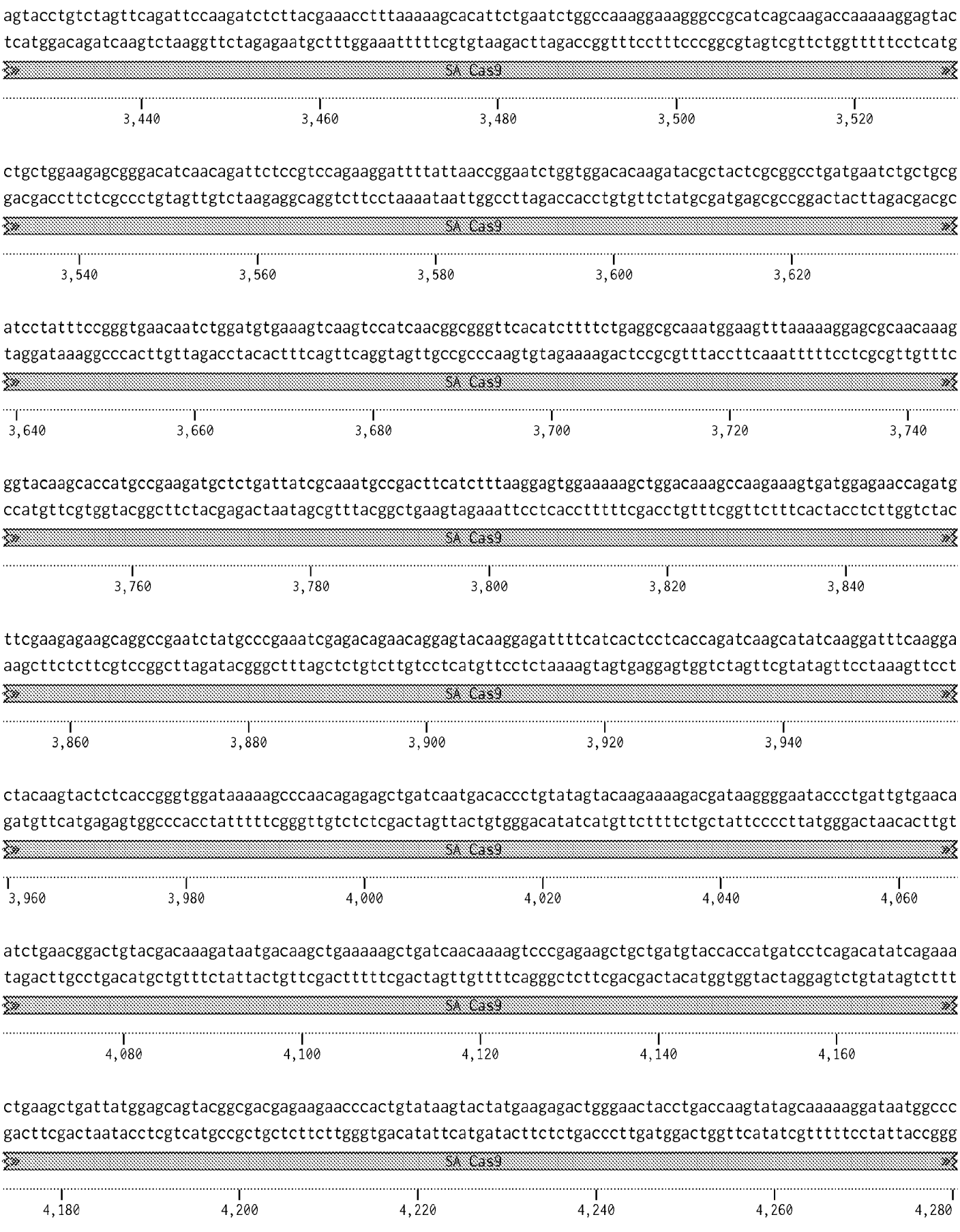

agtacctgtctagttcagattccaagatctcttacgaaacctttaaaaagcacattctgaatctggccaaaggaaagggccgcatcagcaagaccaaaaaggagtac
tcatggacagatcaagtctaaggttctagagaatgctttggaaatttttcgtgtaagacttagaccggtttcctttcccggcgtagtcgttctggtttttcctcatg
SA Cas9
3,440
3,460
3,480
3,500
3,520
ctgctggaagagcgggacatcaacagattctccgtccagaaggattttattaaccggaatctggtggacacaagatacgctactcgcggcctgatgaatctgctgcg
gacgaccttctcgccctgtagttgtctaagaggcaggtcttcctaaaataattggccttagaccacctgtgttctatgcgatgagcgccggactacttagacgacgc
SA Cas9
3,540
3,560
3,580
3,600
3,620
atcctatttccgggtgaacaatctggatgtgaaagtcaagtccatcaacggcgggttcacatcttttctgaggcgcaaatggaagtttaaaaaggagcgcaacaaag
taggataaaggcccacttgttagacctacactttcagttcaggtagttgccgcccaagtgtagaaaagactccgcgtttaccttcaaatttttcctcgcgttgtttc
SA Cas9
3,640
3,660
3,680
3,700
3,720
3,740
ggtacaagcaccatgccgaagatgctctgattatcgcaaatgccgacttcatctttaaggagtggaaaaagctggacaaagccaagaaagtgatggagaaccagatg
ccatgttcgtggtacggcttctacgagactaatagcgtttacggctgaagtagaaattcctcacctttttcgacctgtttcggttctttcactacctcttggtctac
SA Cas9
3,760
3,780
3,800
3,820
3,840
ttcgaagagaagcaggccgaatctatgcccgaaatcgagacagaacaggagtacaaggagattttcatcactcctcaccagatcaagcatatcaaggatttcaagga
aagcttctcttcgtccggcttagatacgggctttagctctgtcttgtcctcatgttcctctaaaagtagtgaggagtggtctagttcgtatagttcctaaagttcct
SA Cas9
3,860
3,880
3,900
3,920
3,940
ctacaagtactctcaccgggtggataaaaagcccaacagagagctgatcaatgacaccctgtatagtacaagaaaagacgataaggggaataccctgattgtgaaca
gatgttcatgagagtggcccacctatttttcgggttgtctctcgactagttactgtgggacatatcatgttcttttctgctattccccttatgggactaacacttgt
SA Cas9
3,960
3,980
4,000
4,020
4,040
4,060
atctgaacggactgtacgacaaagataatgacaagctgaaaaagctgatcaacaaaagtcccgagaagctgctgatgtaccaccatgatcctcagacatatcagaaa
tagacttgcctgacatgctgtttctattactgttcgacttttcgactagttgttttcagggctcttcgacgactacatggtggtactaggagtctgtatagtcttt
SA Cas9
4,080
4,100
4,120
4,140
4,160
ctgaagctgattatggagcagtacggcgacgagaagaacccactgtataagtactatgaagagactgggaactacctgaccaagtatagcaaaaaggataatggccc
gacttcgactaatacctcgtcatgccgctgctcttcttgggtgacatattcatgatacttctctgacccttgatggactggttcatatcgtttttcctattaccggg
SA Cas9
4,180
4,200
4,220
4,240
4,260
4,280

FIG. 19E

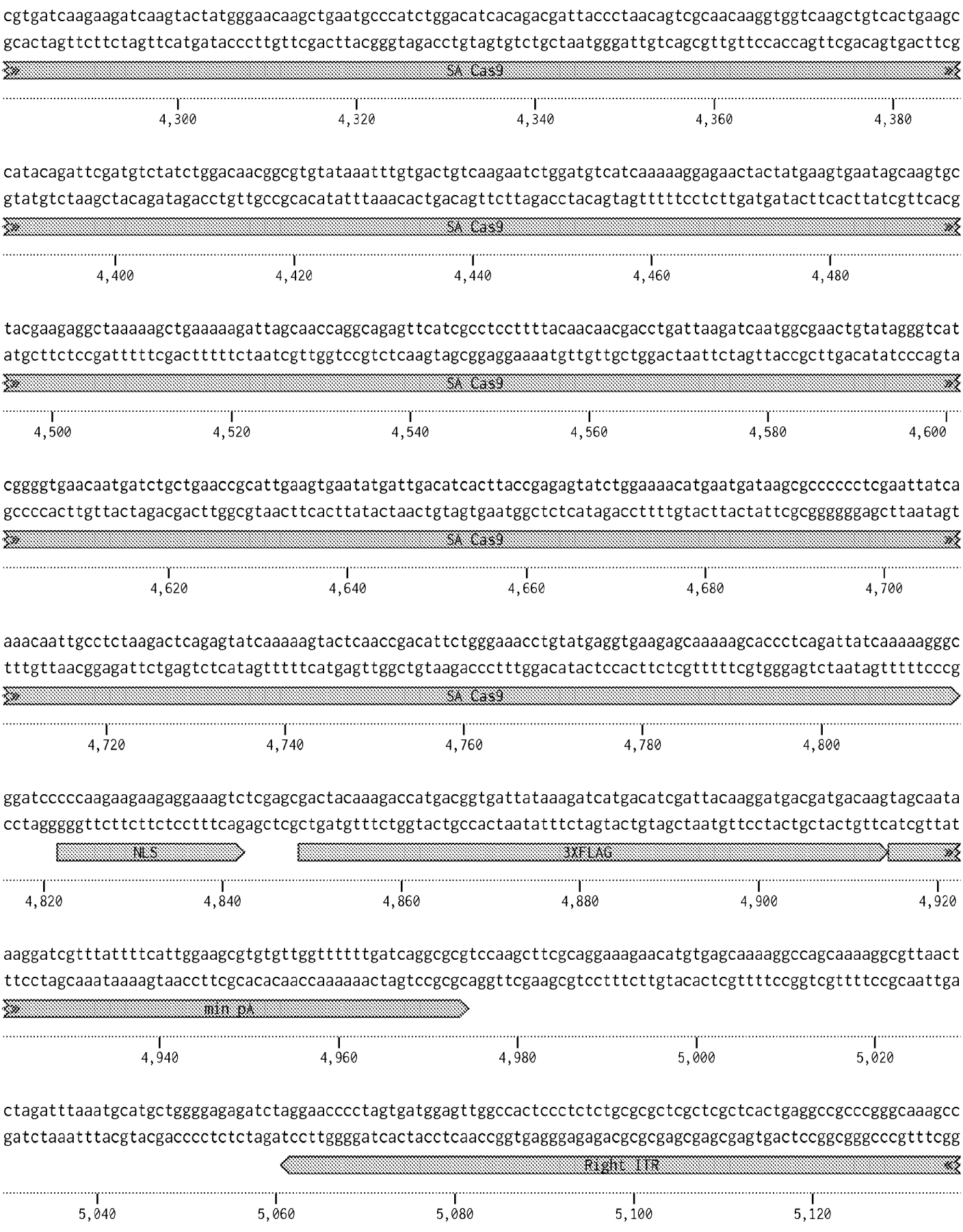

cgtgatcaagaagatcaagtactatgggaacaagctgaatgcccatctggacatcacagacgattaccctaacagtcgcaacaaggtggtcaagctgtcactgaagc
gcactagttcttctagttcatgataccttgttcgacttacgggtagacctgtagtgtctgctaatgggattgtcagcgttgttccaccagttcgacagtgacttcg
SA Cas9
4,300
4,320
4,340
4,360
4,380
catacagattcgatgtctatctggacaacggcgtgtataaatttgtgactgtcaagaatctggatgtcatcaaaaaggagaactactatgaagtgaatagcaagtgc
gtatgtctaagctacagatagacctgttgccgcacatatttaaacactgacagttcttagacctacagtagtttttcctcttgatgatacttcacttatcgttcacg
SA Cas9
4,400
4,420
4,440
4,460
4,480
tacgaagaggctaaaaagctgaaaaagattagcaaccaggcagagttcatcgcctccttttacaacaacgacctgattaagatcaatggcgaactgtatagggtcat
atgcttctccgatttttcgacttttttctaatcgttggtccgtctcaagtagcggaggaaaatgttgttgctggactaattctagttaccgcttgacatatcccagta
SA Cas9
4,500
4,520
4,540
4,560
4,580
4,600
cggggtgaacaatgatctgctgaaccgcattgaagtgaatatgattgacatcacttaccgagagtatctggaaaacatgaatgataagcgcccccctcgaattatca
gccccacttgttactagacgacttggcgtaacttcacttatactaactgtagtgaatggctctcatagacctttgtacttactattcgcgggggggagcttaatagt
SA Cas9
4,620
4,640
4,660
4,680
4,700
aaacaattgcctctaagactcagagtatcaaaaagtactcaaccgacattctgggaaacctgtatgaggtgaagagcaaaaagcaccctcagattatcaaaaagggc
tttgttaacggagattctgagtctcatagtttttcatgagttggctgtaagacccttggacatactccacttctcgtttttcgtgggagtctaatagtttttcccg
SA Cas9
4,720
4,740
4,760
4,780
4,800
ggatcccccaagaagaagaggaaagtctcgagcgactacaaagaccatgacggtgattataaagatcatgacatcgattacaaggatgacgatgacaagtagcaata
cctaggggttcttcttctcctttcagagctcgctgatgtttctggtactgccactaatatttctagtactgtagctaatgttcctactgctactgttcatcgttat
NLS
3XFLAG
4,820
4,840
4,860
4,880
4,900
4,920
aaggatcgtttattttcattggaagcgtgtgttggtttttgatcaggcgcgtccaagcttcgcaggaaagaacatgtgagcaaaaggccagcaaaaggcgttaact
ttcctagcaaataaaagtaaccttcgcacacaaccaaaaaactagtccgcgcaggttcgaagcgtcctttcttgtacactcgttttccggtcgttttccgcaattga
min pA
4,940
4,960
4,980
5,000
5,020
ctagatttaaatgcatgctggggagagatctaggaacccctagtgatggagttggccactccctctctgcgcgctcgctcgctcactgaggccgcccgggcaaagcc
gatctaaatttacgtacgacccctctctagatccttggggatcactacctcaaccggtgagggagagacgcgcgagcgagcgagtgactccggcgggcccgtttcgg
Right ITR
5,040
5,060
5,080
5,100
5,120

FIG. 19F

cgggcgtcgggcgacctttggtcgcccggcctcagtgagcgagcgagcgcgcagagagggagtggccaa
gcccgcagcccgctggaaaccagcgggccggagtcactcgctcgctcgcgcgtctctccctcaccggtt

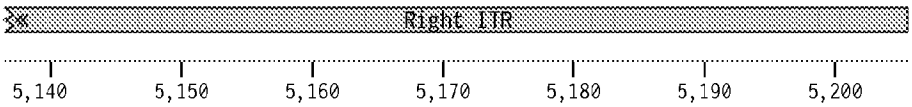


FIG. 19G

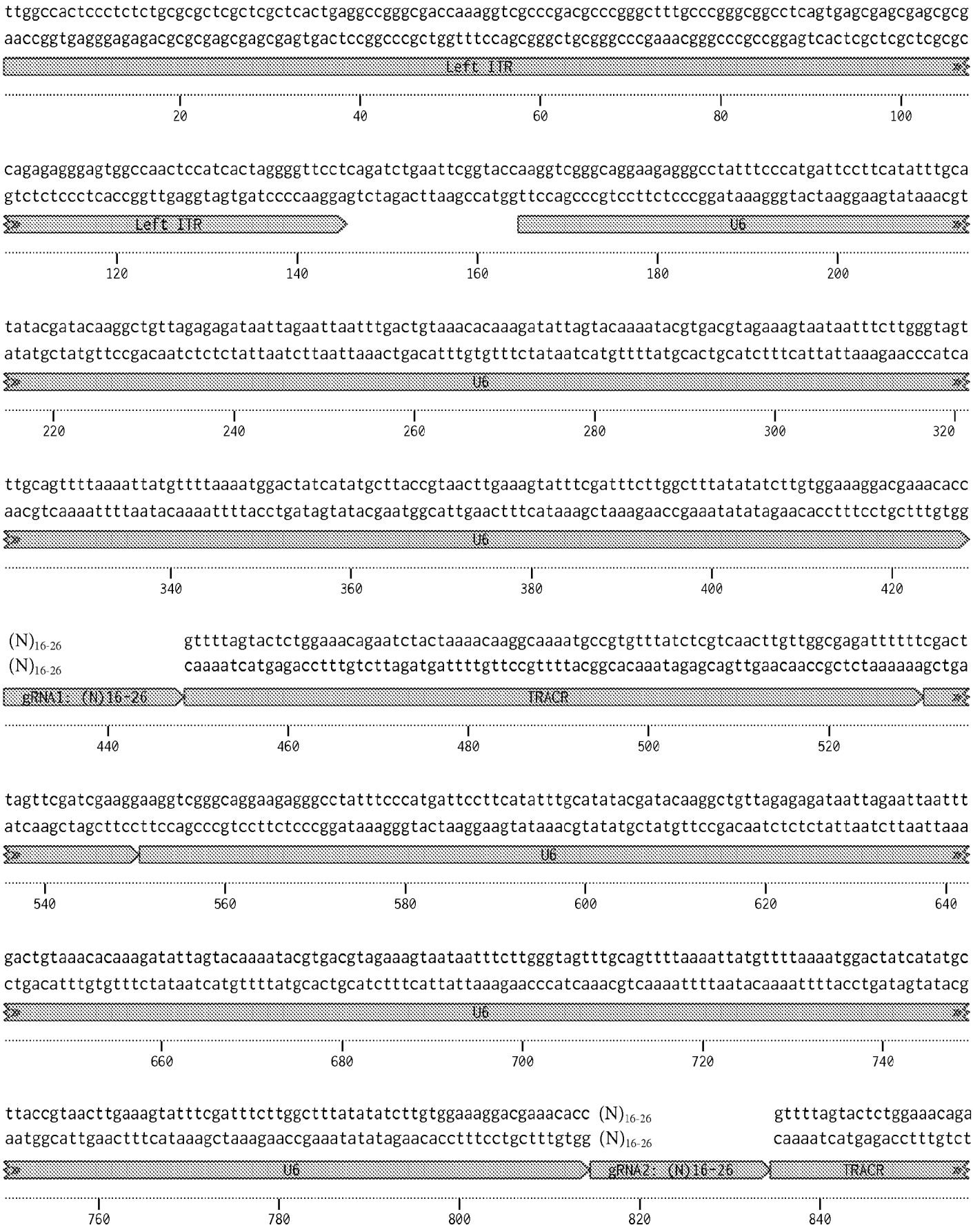

ttggccactccctctctgcgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggcctcagtgagcgagcgagcgcg
aaccggtgagggagagacgcgcgagcgagcgagtgactccggcccgctggtttccagcgggctgcgggcccgaaacgggcccgccggagtcactcgctcgctcgcgc
Left ITR
20
40
60
80
100
cagagagggagtggccaactccatcactaggggttcctcagatctgaattcggtaccaaggtcgggcaggaagagggcctatttcccatgattccttcatatttgca
gtctctccctcaccggttgaggtagtgatccccaaggagtctagacttaagccatggttccagcccgtccttctcccggataaagggtactaaggaagtataaacgt
Left ITR
U6
120
140
160
180
200
tatacgatacaaggctgttagagagataattagaattaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagt
atatgctatgttccgacaatctctctattaatcttaattaaactgacatttgtgtttctataatcatgttttatgcactgcatctttcattattaaagaacccatca
U6
220
240
260
280
300
320
ttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacacc
aacgtcaaaattttaatacaaaattttacctgatagtatacgaatggcattgaactttcataaagctaaagaaccgaaatatatagaacacctttcctgctttgtgg
U6
340
360
380
400
420
(N)16-26 gttttagtactctggaaacagaatctactaaaacaaggcaaaatgccgtgtttatctcgtcaacttgttggcgagatttttcgact
(N)16-26 caaaatcatgagacctttgtcttagatgattttgttccgttttacggcacaaatagagcagttgaacaaccgctctaaaaaagctga
gRNA1: (N)16-26
TRACR
440
460
480
500
520
tagttcgatcgaaggaaggtcgggcaggaagagggcctatttcccatgattccttcatatttgcatatacgatacaaggctgttagagagataattagaattaatttt
atcaagctagcttccttccagcccgtccttctcccggataaagggtactaaggaagtataaacgtatatgctatgttccgacaatctctctattaatcttaattaaa
U6
540
560
580
600
620
640
gactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgttttaaaatggactatcatatgc
ctgacatttgtgtttctataatcatgttttatgcactgcatctttcattattaaagaacccatcaaacgtcaaaattttaatacaaaattttacctgatagtatacg
U6
660
680
700
720
740
ttaccgtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacacc (N)16-26 gttttagtactctggaaacaga
aatggcattgaactttcataaagctaaagaaccgaaatatatagaacacctttcctgctttgtgg (N)16-26 caaaatcatgagacctttgtct
U6
gRNA2: (N)16-26
TRACR
760
780
800
820
840

FIG. 20A

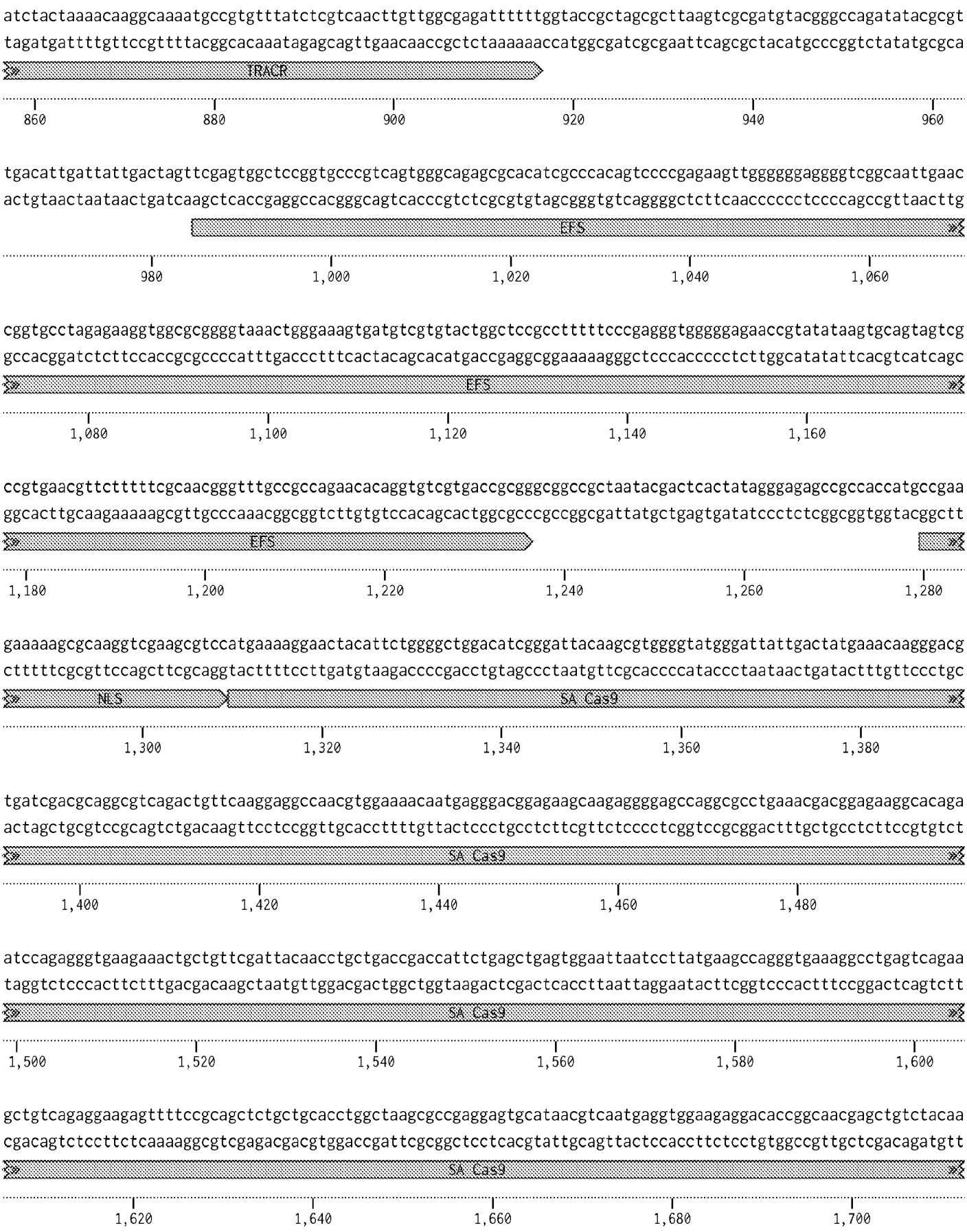
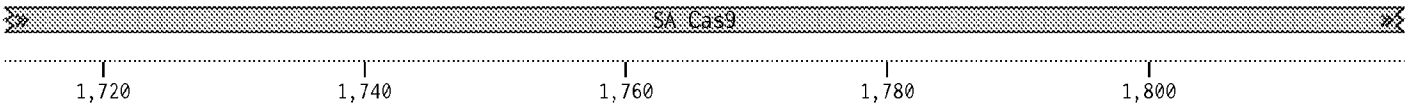
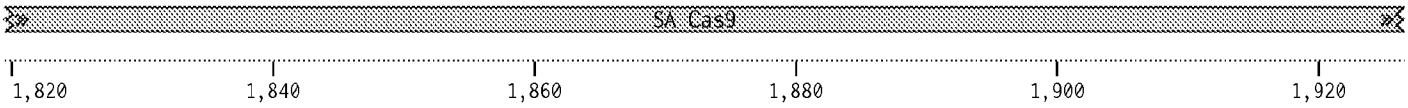

atctactaaaacaaggcaaaatgccgtgtttatctcgtcaacttgttggcgagatttttttggtaccgctagcgcttaagtcgcgatgtacgggccagatatacgcgt
tagatgattttgttccgttttacggcacaaatagagcagttgaacaaccgctctaaaaaaccatggcgatcgcgaattcagcgctacatgcccggtctatatgcgca
TRACR
860
880
900
920
940
960
tgacattgattattgactagttcgagtggctccggtgcccgtcagtgggcagagcgcacatcgcccacagtccccgagaagttgggggggaggggtcggcaattgaac
actgtaactaataactgatcaagctcaccgaggccacgggcagtcacccgtctcgcgtgtagcgggtgtcaggggctcttcaacccccctccccagccgttaacttg
EFS
980
1,000
1,020
1,040
1,060
cggtgcctagagaaggtggcgcggggtaaactgggaaagtgatgtcgtgtactggctccgccttttttcccgagggtgggggagaaccgtatataagtgcagtagtcg
gccacggatctcttccaccgcgccccatttgacccttcactacagcacatgaccgaggcggaaaaagggctcccacccccctcttggcatatattcacgtcatcagc
EFS
1,080
1,100
1,120
1,140
1,160
ccgtgaacgttctttttcgcaacgggtttgccgccagaacacaggtgtcgtgaccgcgggcggccgctaatacgactcactatagggagagccgccaccatgccgaa
ggcacttgcaagaaaaagcgttgcccaaacggcggtcttgtgtccacagcactggcgcccgccggcgattatgctgagtgatatccctctcggcggtggtacggctt
EFS
1,180
1,200
1,220
1,240
1,260
1,280
gaaaaagcgcaaggtcgaagcgtccatgaaaaggaactacattctggggctggacatcgggattacaagcgtggggtatgggattattgactatgaaacaagggacg
ctttttcgcgttccagcttcgcaggtacttttccttgatgtaagaccccgacctgtagccctaatgttcgcaccccataccctaataactgatactttgttccctgc
NLS
SA Cas9
1,300
1,320
1,340
1,360
1,380
tgatcgacgcaggcgtcagactgttcaaggaggccaacgtggaaaacaatgagggacggagaagcaagaggggagccaggcgcctgaaacgacggagaaggcacaga
actagctgcgtccgcagtctgacaagttcctccggttgcacctttgttactccctgcctcttcgttctcccctcggtccgcggactttgctgcctcttccgtgtct
SA Cas9
1,400
1,420
1,440
1,460
1,480
atccagagggtgaagaaactgctgttcgattacaacctgctgaccgaccattctgagctgagtggaattaatccttatgaagccagggtgaaaggcctgagtcagaa
taggtctcccacttctttgacgacaagctaatgttggacgactggctggtaagactcgactcaccttaattaggaatacttcggtcccactttccggactcagtctt
SA Cas9
1,500
1,520
1,540
1,560
1,580
1,600
gctgtcagaggaagagttttccgcagctctgctgcacctggctaagcgccgaggagtgcataacgtcaatgaggtggaagaggacaccggcaacgagctgtctacaa
cgacagtctccttctcaaaaggcgtcgagacgacgtggaccgattcgcggctcctcacgtattgcagttactccaccttctcctgtggccgttgctcgacagatgtt
SA Cas9
1,620
1,640
1,660
1,680
1,700

FIG. 20B

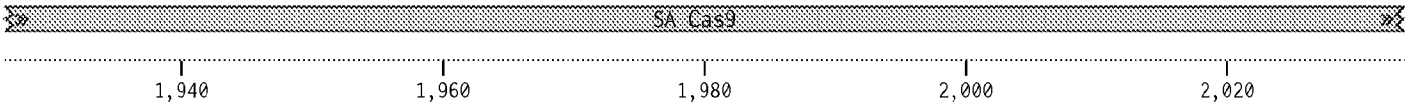
aggaacagatctcacgcaatagcaaagctctggaagagaagtatgtcgcagagctgcagctggaacggctgaagaaagatggcgaggtgagagggtcaattaatagg
tccttgtctagagtgcgttatcgtttcgagaccttctcttcatacagcgtctcgacgtcgaccttgccgacttctttctaccgctccactctcccagttaattatcc



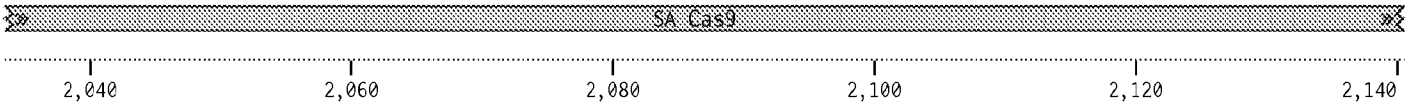
ttcaagacaagcgactacgtcaaagaagccaagcagctgctgaaagtgcagaaggcttaccaccagctggatcagagcttcatcgatacttatatcgacctgctgga
aagttctgttcgctgatgcagtttcttcggttcgtcgacgactttcacgtcttccgaatggtggtcgacctagtctcgaagtagctatgaatatagctggacgacct



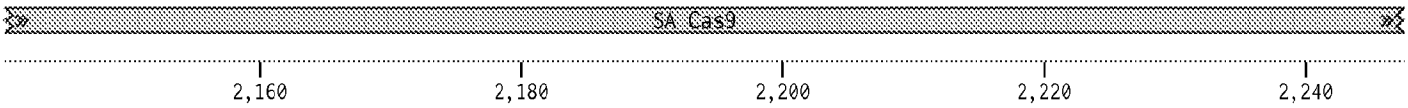
gactcggagaacctactatgagggaccaggagaagggagccccttcggatggaaagacatcaaggaatggtacgagatgctgatgggacattgcacctattttccag
ctgagcctcttggatgatactccctggtcctcttccctcggggaagcctacctttctgtagttccttaccatgctctacgactaccctgtaacgtggataaaaggtc



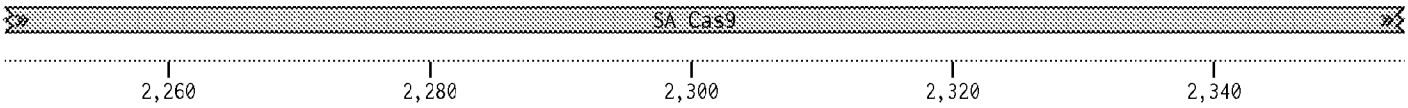
aagagctgagaagcgtcaagtacgcttataacgcagatctgtacaacgccctgaatgacctgaacaacctggtcatcaccagggatgaaaacgagaaactggaatac
ttctcgactcttcgcagttcatgcgaatattgcgtctagacatgttgcgggacttactggacttgttggaccagtagtggtccctacttttgctctttgaccttatg



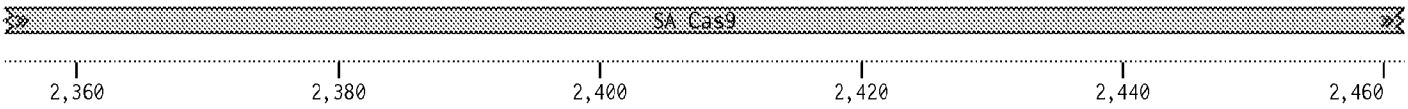
tatgagaagttccagatcatcgaaaacgtgtttaagcagaagaaaaagcctacactgaaacagattgctaaggagatcctggtcaacgaagaggacatcaagggcta
atactcttcaaggtctagtagcttttgcacaaattcgtcttctttttcggatgtgactttgtctaacgattcctctaggaccagttgcttctcctgtagttcccgat



ccgggtgacaagcactggaaaaccagagttcaccaatctgaaagtgtatcacgatattaaggacatcacagcacggaaagaaatcattgagaacgccgaactgctgg
ggcccactgttcgtgacctttggtctcaagtggttagactttcacatagtgctataattcctgtagtgtcgtgcctttctttagtaactcttgcggcttgacgacc



atcagattgctaagatcctgactatctaccagagctccgaggacatccaggaagagctgactaacctgaacagcgagctgacccaggaagagatcgaacagattagt
tagtctaacgattctaggactgatagatggtctcgaggctcctgtaggtccttctcgactgattggacttgtcgctcgactgggtccttctctagcttgtctaatca



aatctgaaggggtacaccggaacacacaacctgtccctgaaagctatcaatctgattctggatgagctgtggcatacaaacgacaatcagattgcaatctttaaccg
ttagacttccccatgtggccttgtgtgttggacagggactttcgatagttagactaagacctactcgacaccgtatgtttgctgttagtctaacgttagaaattggc

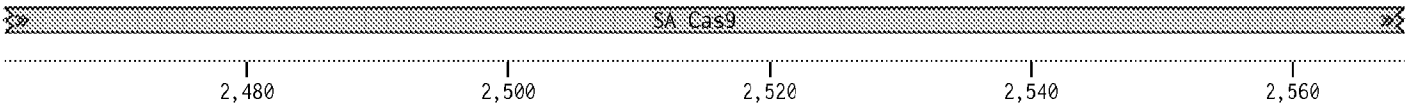


FIG. 20C

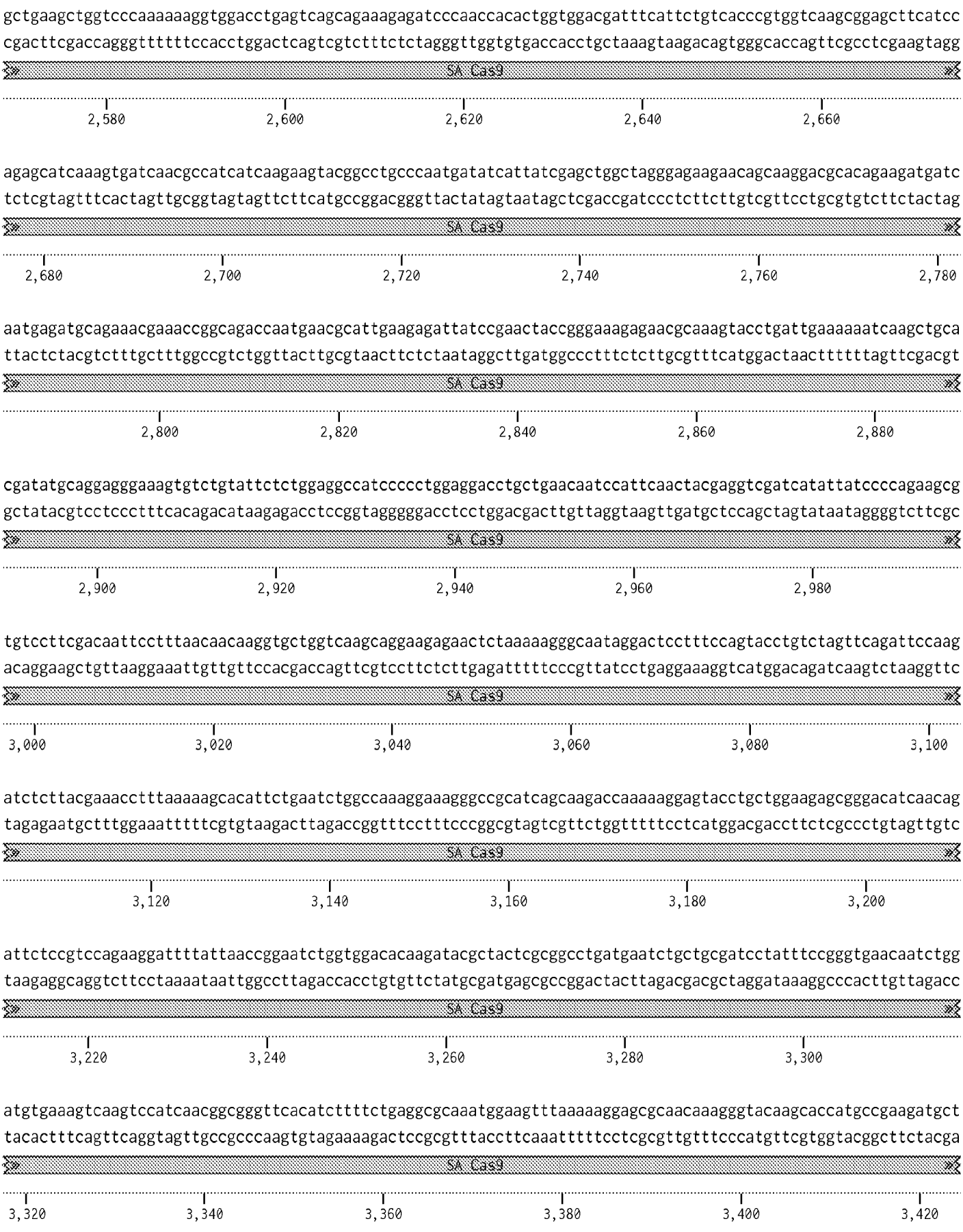

gctgaagctggtcccaaaaaaggtggacctgagtcagcagaaagagatcccaaccacactggtggacgatttcattctgtcacccgtggtcaagcggagcttcatcc
cgacttcgaccagggtttttttccacctggactcagtcgtctttctctagggttggtgtgaccacctgctaaagtaagacagtgggcaccagttcgcctcgaagtagg
SA Cas9
2,580
2,600
2,620
2,640
2,660
agagcatcaaagtgatcaacgccatcatcaagaagtacggcctgcccaatgatatcattatcgagctggctagggagaagaacagcaaggacgcacagaagatgatc
tctcgtagtttcactagttgcggtagtagttcttcatgccggacgggttactatagtaatagctcgaccgatccctcttcttgtcgttcctgcgtgtcttctactag
SA Cas9
2,680
2,700
2,720
2,740
2,760
2,780
aatgagatgcagaaacgaaaccggcagaccaatgaacgcattgaagagattatccgaactaccgggaaagagaacgcaaagtacctgattgaaaaaatcaagctgca
ttactctacgtctttgctttggccgtctggttacttgcgtaacttctctaataggcttgatggccctttctcttgcgtttcatggactaacttttttagttcgacgt
SA Cas9
2,800
2,820
2,840
2,860
2,880
cgatatgcaggagggaaagtgtctgtattctctggaggccatccccctggaggacctgctgaacaatccattcaactacgaggtcgatcatattatccccagaagcg
gctatacgtcctccctttcacagacataagagacctccggtagggggacctcctggacgacttgttaggtaagttgatgctccagctagtataataggggtcttcgc
SA Cas9
2,900
2,920
2,940
2,960
2,980
tgtccttcgacaattcctttaacaacaaggtgctggtcaagcaggaagagaactctaaaaagggcaataggactcctttccagtacctgtctagttcagattccaag
acaggaagctgttaaggaaattgttgttccacgaccagttcgtccttctcttgagatttttcccgttatcctgaggaaaggtcatggacagatcaagtctaaggttc
SA Cas9
3,000
3,020
3,040
3,060
3,080
3,100
atctcttacgaaacctttaaaaagcacattctgaatctggccaaaggaaagggccgcatcagcaagaccaaaaaggagtacctgctggaagagcgggacatcaacag
tagagaatgctttggaaatttttcgtgtaagacttagaccggtttcctttcccggcgtagtcgttctggtttttcctcatggacgaccttctcgccctgtagttgtc
SA Cas9
3,120
3,140
3,160
3,180
3,200
attctccgtccagaaggattttattaaccggaatctggtggacacaagatacgctactcgcggcctgatgaatctgctgcgatcctatttccgggtgaacaatctgg
taagaggcaggtcttcctaaaataattggccttagaccacctgtgttctatgcgatgagcgccggactacttagacgacgctaggataaaggcccacttgttagacc
SA Cas9
3,220
3,240
3,260
3,280
3,300
atgtgaaagtcaagtccatcaacggcgggttcacatcttttctgaggcgcaaatggaagtttaaaaaggagcgcaacaaagggtacaagcaccatgccgaagatgct
tacactttcagttcaggtagttgccgcccaagtgtagaaaagactccgcgtttaccttcaaatttttcctcgcgttgtttcccatgttcgtggtacggcttctacga
SA Cas9
3,320
3,340
3,360
3,380
3,400
3,420

FIG. 20D

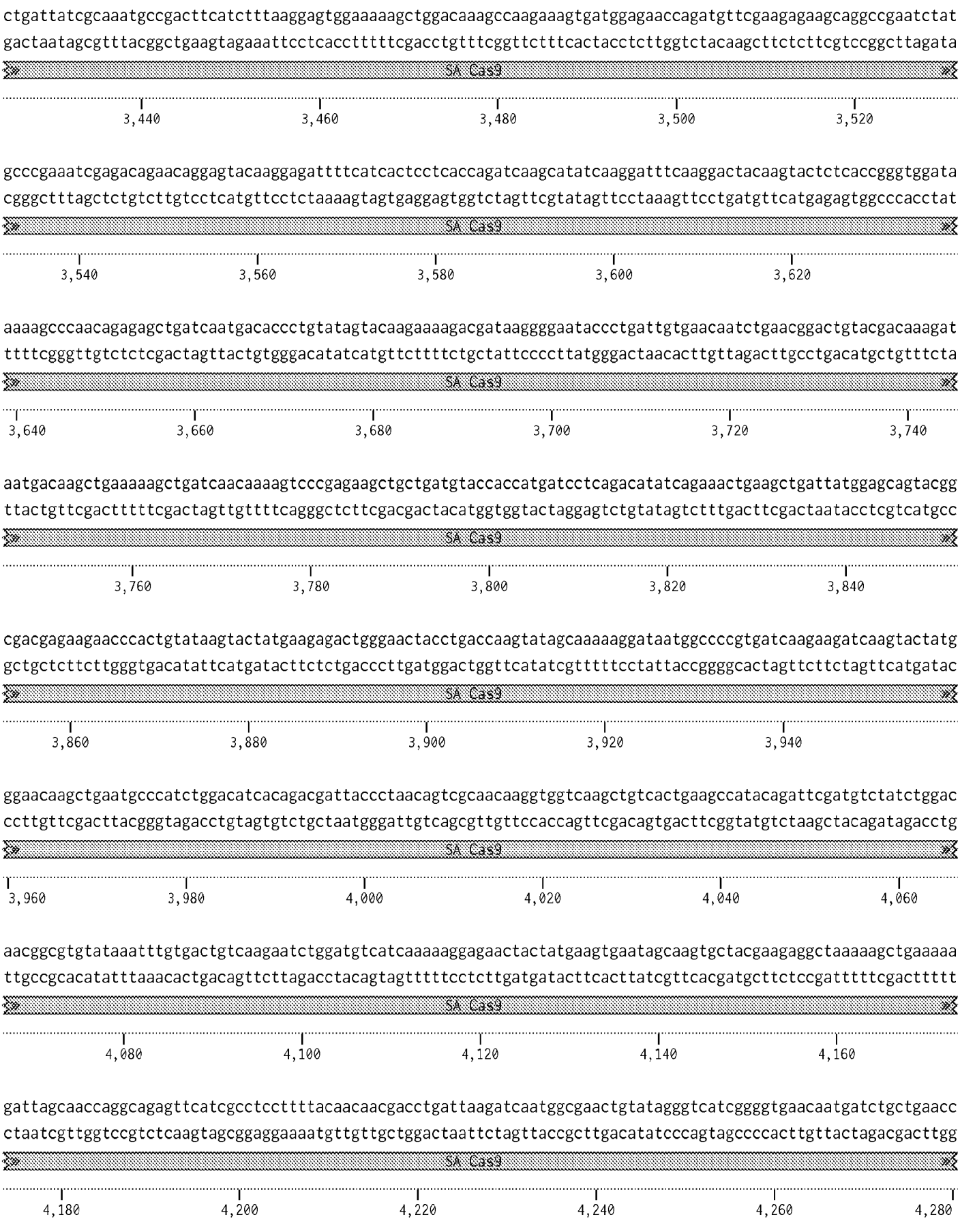

ctgattatcgcaaatgccgacttcatctttaaggagtggaaaaagctggacaaagccaagaaagtgatggagaaccagatgttcgaagagaagcaggccgaatctat
gactaatagcgtttacggctgaagtagaaattcctcacctttttcgacctgtttcggttctttcactacctcttggtctacaagcttctcttcgtccggcttagata
SA Cas9
3,440
3,460
3,480
3,500
3,520
gcccgaaatcgagacagaacaggagtacaaggagattttcatcactcctcaccagatcaagcatatcaaggatttcaaggactacaagtactctcaccgggtggata
cgggctttagctctgtcttgtcctcatgttcctctaaaagtagtgaggagtggtctagttcgtatagttcctaaagttcctgatgttcatgagagtggcccacctat
SA Cas9
3,540
3,560
3,580
3,600
3,620
aaaagcccaacagagagctgatcaatgacaccctgtatagtacaagaaaagacgataaggggaataccctgattgtgaacaatctgaacggactgtacgacaaagat
ttttcgggttgtctctcgactagttactgtgggacatatcatgttcttttctgctattccccttatgggactaacacttgttagacttgcctgacatgctgtttcta
SA Cas9
3,640
3,660
3,680
3,700
3,720
3,740
aatgacaagctgaaaaagctgatcaacaaaagtcccgagaagctgctgatgtaccaccatgatcctcagacatatcagaaactgaagctgattatggagcagtacgg
ttactgttcgacttttcgactagttgttttcagggctcttcgacgactacatggtggtactaggagtctgtatagtctttgacttcgactaatacctcgtcatgcc
SA Cas9
3,760
3,780
3,800
3,820
3,840
cgacgagaagaacccactgtataagtactatgaagagactgggaactacctgaccaagtatagcaaaaaggataatggccccgtgatcaagaagatcaagtactatg
gctgctcttcttgggtgacatattcatgatacttctctgacccttgatggactggttcatatcgtttttcctattaccggggcactagttcttctagttcatgatac
SA Cas9
3,860
3,880
3,900
3,920
3,940
ggaacaagctgaatgcccatctggacatcacagacgattaccctaacagtcgcaacaaggtggtcaagctgtcactgaagccatacagattcgatgtctatctggac
ccttgttcgacttacgggtagacctgtagtgtctgctaatgggattgtcagcgttgttccaccagttcgacagtgacttcggtatgtctaagctacagatagacctg
SA Cas9
3,960
3,980
4,000
4,020
4,040
4,060
aacggcgtgtataaatttgtgactgtcaagaatctggatgtcatcaaaaaggagaactactatgaagtgaatagcaagtgctacgaagaggctaaaaagctgaaaaa
ttgccgcacatatttaaacactgacagttcttagacctacagtagtttttcctcttgatgatacttcacttatcgttcacgatgcttctccgatttttcgacttttt
SA Cas9
4,080
4,100
4,120
4,140
4,160
gattagcaaccaggcagagttcatcgcctccttttacaacaacgacctgattaagatcaatggcgaactgtatagggtcatcggggtgaacaatgatctgctgaacc
ctaatcgttggtccgtctcaagtagcggaggaaaatgttgttgctggactaattctagttaccgcttgacatatcccagtagccccacttgttactagacgacttgg
SA Cas9
4,180
4,200
4,220
4,240
4,260
4,280

FIG. 20E

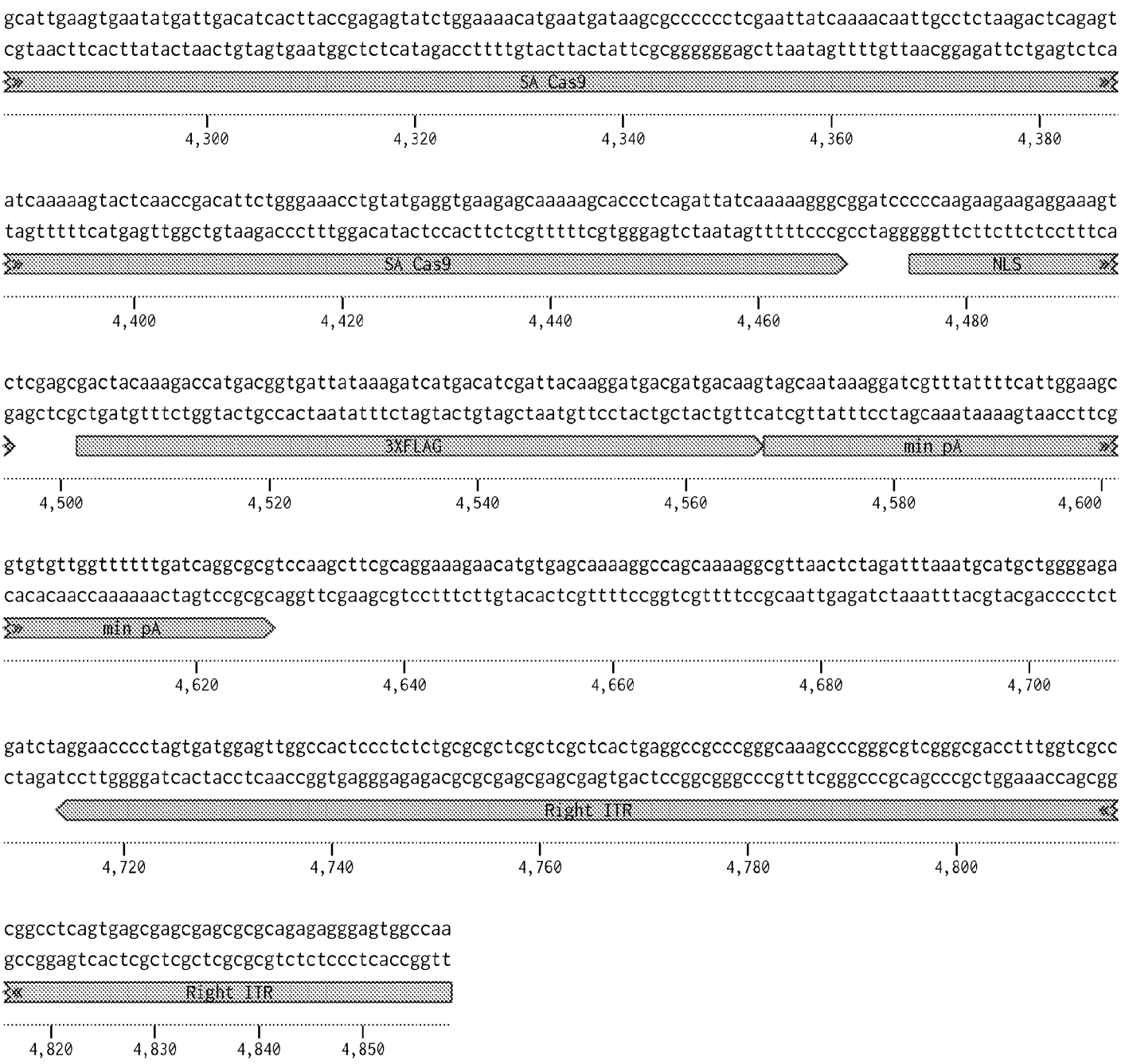

gcattgaagtgaatatgattgacatcacttaccgagagtatctggaaaacatgaatgataagcgcccccctcgaattatcaaaacaattgcctctaagactcagagt
cgtaacttcacttatactaactgtagtgaatggctctcatagacctttttgtacttactattcgcgggggggagcttaatagttttgttaacggagattctgagtctca
SA Cas9
4,300
4,320
4,340
4,360
4,380
atcaaaaagtactcaaccgacattctgggaaacctgtatgaggtgaagagcaaaaagcaccctcagattatcaaaaagggcggatcccccaagaagaagaggaaagt
tagtttttcatgagttggctgtaagacccttttggacatactccacttctcgtttttcgtgggagtctaatagtttttcccgcctaggggttcttcttctcctttca
SA Cas9
NLS
4,400
4,420
4,440
4,460
4,480
ctcgagcgactacaaagaccatgacggtgattataaagatcatgacatcgattacaaggatgacgatgacaagtagcaataaaggatcgtttattttcattggaagc
gagctcgctgatgtttctggtactgccactaatatttctagtactgtagctaatgttcctactgctactgttcatcgttatttcctagcaaataaaagtaaccttcg
3XFLAG
min pA
4,500
4,520
4,540
4,560
4,580
4,600
gtgtgttggtttttgatcaggcgcgtccaagcttcgcaggaaagaacatgtgagcaaaaggccagcaaaaggcgttaactctagatttaaatgcatgctggggaga
cacacaaccaaaaaactagtccgcgcaggttcgaagcgtccttttcttgtacactcgttttccggtcgttttccgcaattgagatctaaatttacgtacgacccctct
min pA
4,620
4,640
4,660
4,680
4,700
gatctaggaacccctagtgatggagttggccactccctctctgcgcgctcgctcgctcactgaggccgcccgggcaaagcccgggcgtcgggcgacctttggtcgcc
ctagatccttggggatcactacctcaaccggtgagggagagacgcgcgagcgagcgagtgactccggcgggcccgtttcgggcccgcagcccgctggaaaccagcgg
Right ITR
4,720
4,740
4,760
4,780
4,800
cggcctcagtgagcgagcgagcgcgcagagagggagtggccaa
gccggagtcactcgctcgctcgcgcgtctctccctcaccggtt
Right ITR
4,820
4,830
4,840
4,850

FIG. 20F

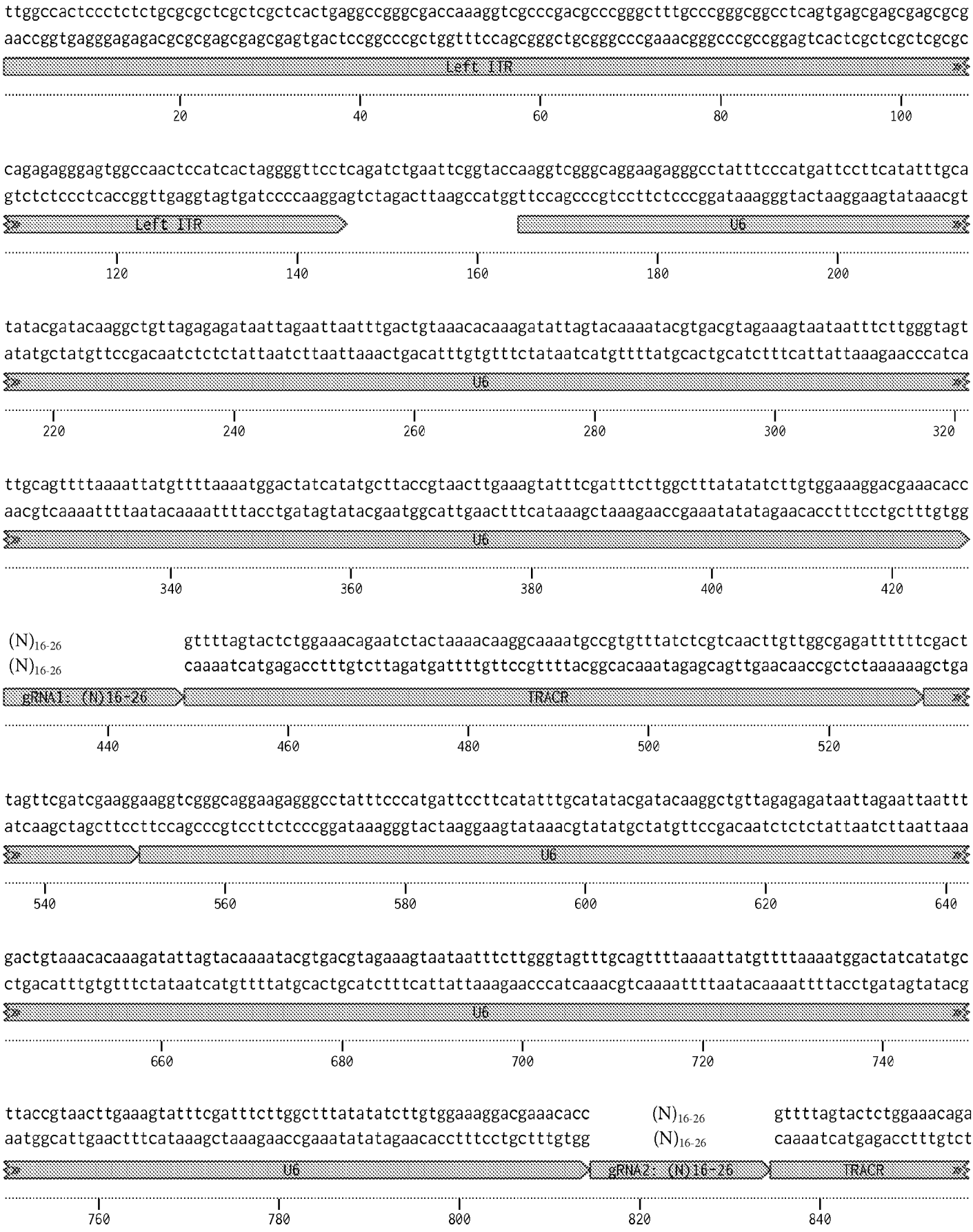

ttggccactccctctctgcgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggcctcagtgagcgagcgagcgcg
aaccggtgagggagagacgcgcgagcgagcgagtgactccggcccgctggtttccagcgggctgcgggcccgaaacgggcccgccggagtcactcgctcgctcgcgc
Left ITR
20
40
60
80
100
cagagagggagtggccaactccatcactaggggttcctcagatctgaattcggtaccaaggtcgggcaggaagagggcctatttcccatgattccttcatatttgca
gtctctccctcaccggttgaggtagtgatccccaaggagtctagacttaagccatggttccagcccgtccttctcccggataaagggtactaaggaagtataaacgt
Left ITR
U6
120
140
160
180
200
tatacgatacaaggctgttagagagataattagaattaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagt
atatgctatgttccgacaatctctctattaatcttaattaaactgacatttgtgtttctataatcatgttttatgcactgcatctttcattattaaagaacccatca
U6
220
240
260
280
300
320
ttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacacc
aacgtcaaaattttaatacaaaattttacctgatagtatacgaatggcattgaactttcataaagctaaagaaccgaaatatatagaacacctttcctgctttgtgg
U6
340
360
380
400
420
(N)16-26 gttttagtactctggaaacagaatctactaaaacaaggcaaaatgccgtgtttatctcgtcaacttgttggcgagatttttcgact
(N)16-26 caaaatcatgagacctttgtcttagatgattttgttccgttttacggcacaaatagagcagttgaacaaccgctctaaaaaagctga
gRNA1: (N)16-26
TRACR
440
460
480
500
520
tagttcgatcgaaggaaggtcgggcaggaagagggcctatttcccatgattccttcatatttgcatatacgatacaaggctgttagagagataattagaattaatttt
atcaagctagcttccttccagcccgtccttctcccggataaagggtactaaggaagtataaacgtatatgctatgttccgacaatctctctattaatcttaattaaa
U6
540
560
580
600
620
640
gactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgttttaaaatggactatcatatgc
ctgacatttgtgtttctataatcatgttttatgcactgcatctttcattattaaagaacccatcaaacgtcaaaattttaatacaaaattttacctgatagtatacg
U6
660
680
700
720
740
ttaccgtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacacc (N)16-26 gttttagtactctggaaacaga
aatggcattgaactttcataaagctaaagaaccgaaatatatagaacacctttcctgctttgtgg (N)16-26 caaaatcatgagacctttgtct
U6
gRNA2: (N)16-26
TRACR
760
780
800
820
840

FIG. 21A

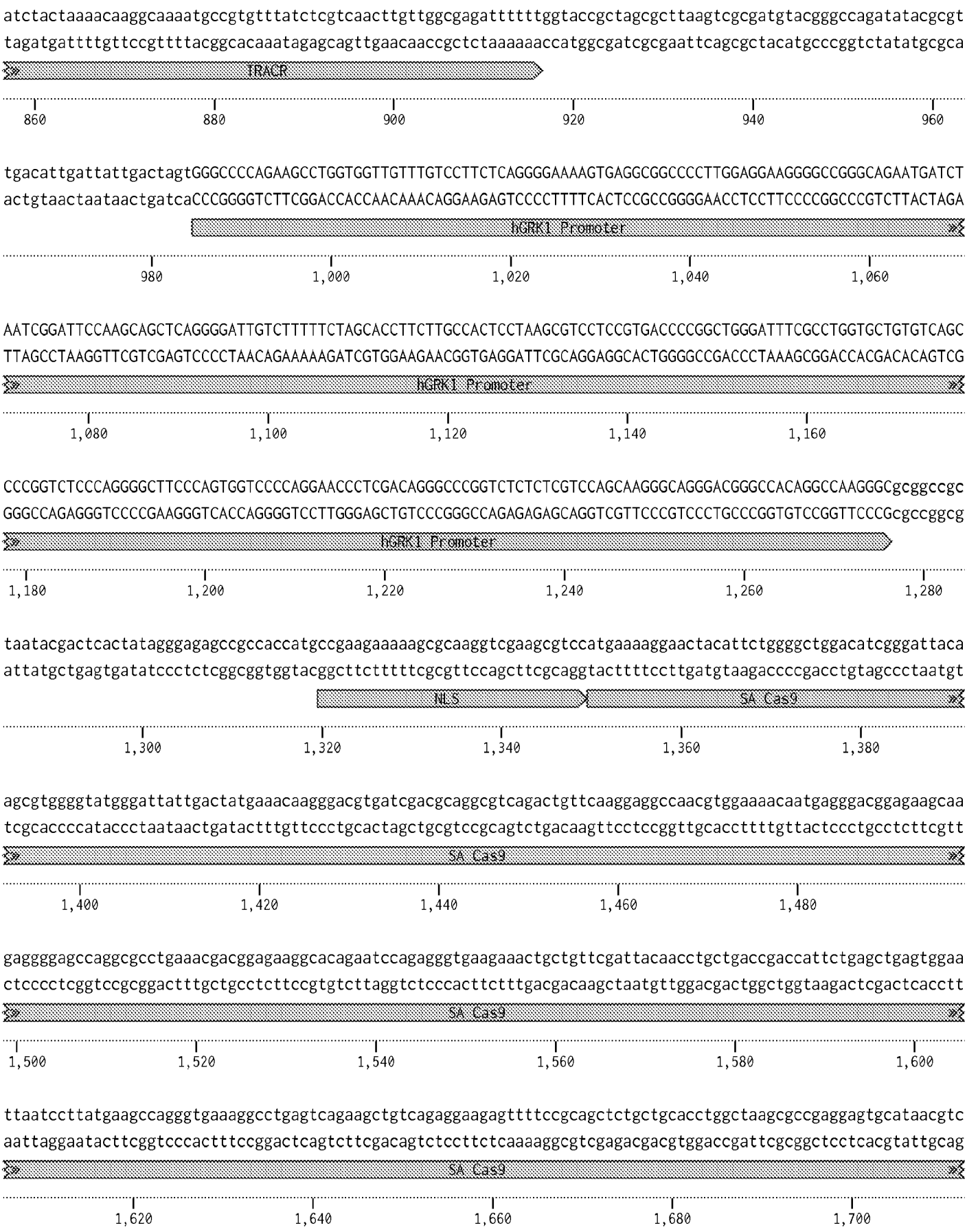

atctactaaaacaaggcaaaatgccgtgtttatctcgtcaacttgttggcgagatttttggtaccgctagcgcttaagtcgcgatgtacgggccagatatacgcgt
tagatgattttgttccgttttacggcacaaatagagcagttgaacaaccgctctaaaaaaccatggcgatcgcgaattcagcgctacatgcccggtctatatgcgca
TRACR
860
880
900
920
940
960
tgacattgattattgactagtGGGCCCCAGAAGCCTGGTGGTTGTTTGTCCTTCTCAGGGGAAAAGTGAGGCGGCCCCTTGGAGGAAGGGGCCGGGCAGAATGATCT
actgtaactaataactgatcaCCCGGGGTCTTCGGACCACCAACAAACAGGAAGAGTCCCCTTTTCACTCCGCCGGGGAACCTCCTTCCCCGGCCCGTCTTACTAGA
hGRK1 Promoter
980
1,000
1,020
1,040
1,060
AATCGGATTCCAAGCAGCTCAGGGGATTGTCTTTTTCTAGCACCTTCTTGCCACTCCTAAGCGTCCTCCGTGACCCCGGCTGGGATTTCGCCTGGTGCTGTGTCAGC
TTAGCCTAAGGTTCGTCGAGTCCCCTAACAGAAAAAGATCGTGGAAGAACGGTGAGGATTCGCAGGAGGCACTGGGGCCGACCCTAAAGCGGACCACGACACAGTCG
hGRK1 Promoter
1,080
1,100
1,120
1,140
1,160
CCCGGTCTCCCAGGGGCTTCCCAGTGGTCCCCAGGAACCCTCGACAGGGCCCGGTCTCTCTCGTCCAGCAAGGGCAGGGACGGGCCACAGGCCAAGGGCgcggccgc
GGGCCAGAGGGTCCCCGAAGGGTCACCAGGGGTCCTTGGGAGCTGTCCCGGGCCAGAGAGAGCAGGTCGTTCCCGTCCCTGCCCGGTGTCCGGTTCCCGcgccggcg
hGRK1 Promoter
1,180
1,200
1,220
1,240
1,260
1,280
taatacgactcactatagggagagccgccaccatgccgaagaaaaagcgcaaggtcgaagcgtccatgaaaaggaactacattctggggctggacatcgggattaca
attatgctgagtgatatccctctcggcggtggtacggcttctttttcgcgttccagcttcgcaggtacttttccttgatgtaagaccccgacctgtagccctaatgt
NLS
SA Cas9
1,300
1,320
1,340
1,360
1,380
agcgtggggtatgggattattgactatgaaacaagggacgtgatcgacgcaggcgtcagactgttcaaggaggccaacgtggaaaacaatgagggacggagaagcaa
tcgcaccccataccctaataactgatactttgttccctgcactagctgcgtccgcagtctgacaagttcctccggttgcacctttgttactccctgcctcttcgtt
SA Cas9
1,400
1,420
1,440
1,460
1,480
gaggggagccaggcgcctgaaacgacggagaaggcacagaatccagagggtgaagaaactgctgttcgattacaacctgctgaccgaccattctgagctgagtggaa
ctcccctcggtccgcggactttgctgcctcttccgtgtcttaggtctcccacttctttgacgacaagctaatgttggacgactggctggtaagactcgactcacctt
SA Cas9
1,500
1,520
1,540
1,560
1,580
1,600
ttaatccttatgaagccagggtgaaaggcctgagtcagaagctgtcagaggaagagttttccgcagctctgctgcacctggctaagcgccgaggagtgcataacgtc
aattaggaatacttcggtcccactttccggactcagtcttcgacagtctccttctcaaaaggcgtcgagacgacgtggaccgattcgcggctcctcacgtattgcag
SA Cas9
1,620
1,640
1,660
1,680
1,700

FIG. 21B

aatgaggtggaagaggacaccggcaacgagctgtctacaaaggaacagatctcacgcaatagcaaagctctggaagagaagtatgtcgcagagctgcagctggaacg
ttactccaccttctcctgtggccgttgctcgacagatgtttccttgtctagagtgcgttatcgtttcgagaccttctcttcatacagcgtctcgacgtcgaccttgc

SA Cas9

1,720 1,740 1,760 1,780 1,800

gctgaagaaagatggcgaggtgagagggtcaattaataggttcaagacaagcgactacgtcaaagaagccaagcagctgctgaaagtgcagaaggcttaccaccagc
cgacttctttctaccgctccactctcccagttaattatccaagttctgttcgctgatgcagtttcttcggttcgtcgacgactttcacgtcttccgaatggtggtcg

SA Cas9

1,820 1,840 1,860 1,880 1,900 1,920

tggatcagagcttcatcgatacttatatcgacctgctggagactcggagaacctactatgagggaccaggagaagggagccccttcggatggaaagacatcaaggaa
acctagtctcgaagtagctatgaatatagctggacgacctctgagcctcttggatgatactccctggtcctcttccctcggggaagcctacctttctgtagttcctt

SA Cas9

1,940 1,960 1,980 2,000 2,020

tggtacgagatgctgatgggacattgcacctattttccagaagagctgagaagcgtcaagtacgcttataacgcagatctgtacaacgccctgaatgacctgaacaa
accatgctctacgactaccctgtaacgtggataaaaggtcttctcgactcttcgcagttcatgcgaatattgcgtctagacatgttgcgggacttactggacttgtt

SA Cas9

2,040 2,060 2,080 2,100 2,120 2,140

cctggtcatcaccagggatgaaaacgagaaactggaatactatgagaagttccagatcatcgaaaacgtgtttaagcagaagaaaaagcctacactgaaacagattg
ggaccagtagtggtccctacttttgctctttgaccttatgatactcttcaaggtctagtagcttttgcacaaattcgtcttctttttcggatgtgactttgtctaac

SA Cas9

2,160 2,180 2,200 2,220 2,240

ctaaggagatcctggtcaacgaagaggacatcaagggctaccgggtgacaagcactggaaaaccagagttcaccaatctgaaagtgtatcacgatattaaggacatc
gattcctctaggaccagttgcttctcctgtagttcccgatggcccactgttcgtgacctttggtctcaagtggttagactttcacatagtgctataattcctgtag

SA Cas9

2,260 2,280 2,300 2,320 2,340

acagcacggaaagaaatcattgagaacgccgaactgctggatcagattgctaagatcctgactatctaccagagctccgaggacatccaggaagagctgactaacct
tgtcgtgcctttctttagtaactcttgcggcttgacgacctagtctaacgattctaggactgatagatggtctcgaggctcctgtaggtccttctcgactgattgga

SA Cas9

2,360 2,380 2,400 2,420 2,440 2,460

gaacagcgagctgacccaggaagagatcgaacagattagtaatctgaaggggtacaccggaacacacaacctgtccctgaaagctatcaatctgattctggatgagc
cttgtcgctcgactgggtccttctctagcttgtctaatcattagacttccccatgtggccttgtgtgttggacagggactttcgatagttagactaagacctactcg

SA Cas9

2,480 2,500 2,520 2,540 2,560

FIG. 21C

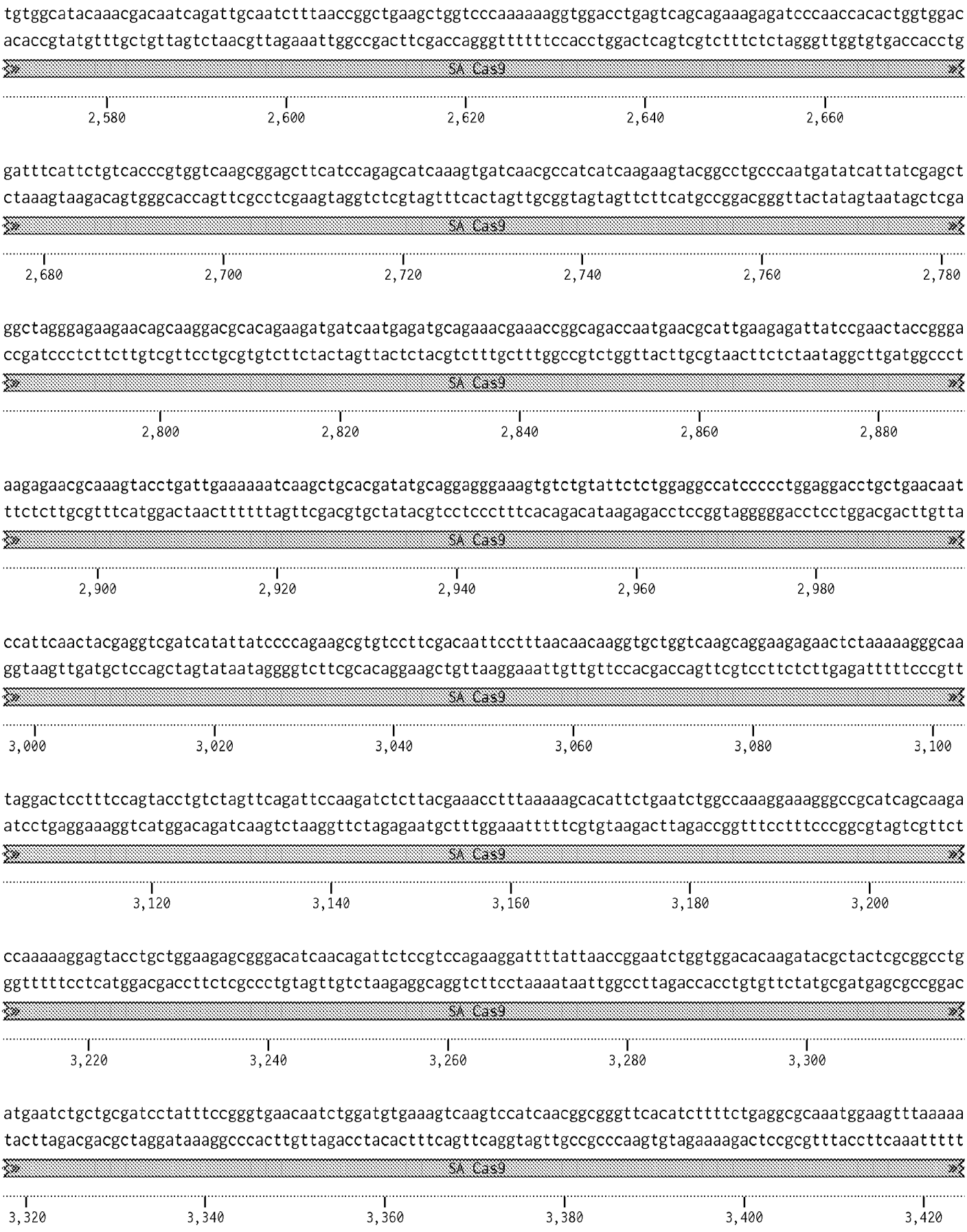

tgtggcatacaaacgacaatcagattgcaatctttaaccggctgaagctggtcccaaaaaaggtggacctgagtcagcagaaagagatcccaaccacactggtggac
acaccgtatgtttgctgttagtctaacgttagaaattggccgacttcgaccagggtttttccacctggactcagtcgtctttctctagggttggtgtgaccacctg
SA Cas9
2,580
2,600
2,620
2,640
2,660
gatttcattctgtcacccgtggtcaagcggagcttcatccagagcatcaaagtgatcaacgccatcatcaagaagtacggcctgcccaatgatatcattatcgagct
ctaaagtaagacagtgggcaccagttcgcctcgaagtaggtctcgtagtttcactagttgcggtagtagttcttcatgccggacgggttactatagtaatagctcga
SA Cas9
2,680
2,700
2,720
2,740
2,760
2,780
ggctagggagaagaacagcaaggacgcacagaagatgatcaatgagatgcagaaacgaaaccggcagaccaatgaacgcattgaagagattatccgaactaccggga
ccgatccctcttcttgtcgttcctgcgtgtcttctactagttactctacgtctttgctttggccgtctggttacttgcgtaacttctctaataggcttgatggccct
SA Cas9
2,800
2,820
2,840
2,860
2,880
aagagaacgcaaagtacctgattgaaaaaatcaagctgcacgatatgcaggagggaaagtgtctgtattctctggaggccatccccctggaggacctgctgaacaat
ttctcttgcgtttcatggactaacttttttagttcgacgtgctatacgtcctccctttcacagacataagagacctccggtagggggacctcctggacgacttgtta
SA Cas9
2,900
2,920
2,940
2,960
2,980
ccattcaactacgaggtcgatcatattatccccagaagcgtgtccttcgacaattcctttaacaacaaggtgctggtcaagcaggaagagaactctaaaaagggcaa
ggtaagttgatgctccagctagtataataggggtcttcgcacaggaagctgttaaggaaattgttgttccacgaccagttcgtccttctcttgagatttttcccgtt
SA Cas9
3,000
3,020
3,040
3,060
3,080
3,100
taggactcctttccagtacctgtctagttcagattccaagatctcttacgaaacctttaaaaagcacattctgaatctggccaaaggaaagggccgcatcagcaaga
atcctgaggaaaggtcatggacagatcaagtctaaggttctagagaatgctttggaaatttttcgtgtaagacttagaccggtttcctttcccggcgtagtcgttct
SA Cas9
3,120
3,140
3,160
3,180
3,200
ccaaaaaggagtacctgctggaagagcgggacatcaacagattctccgtccagaaggattttattaaccggaatctggtggacacaagatacgctactcgcggcctg
ggtttttcctcatggacgaccttctcgccctgtagttgtctaagaggcaggtcttcctaaaataattggccttagaccacctgtgttctatgcgatgagcgccggac
SA Cas9
3,220
3,240
3,260
3,280
3,300
atgaatctgctgcgatcctatttccgggtgaacaatctggatgtgaaagtcaagtccatcaacggcgggttcacatcttttctgaggcgcaaatggaagtttaaaaa
tacttagacgacgctaggataaaggcccacttgttagacctacactttcagttcaggtagttgccgcccaagtgtagaaaagactccgcgtttaccttcaaatttttt
SA Cas9
3,320
3,340
3,360
3,380
3,400
3,420

FIG. 21D

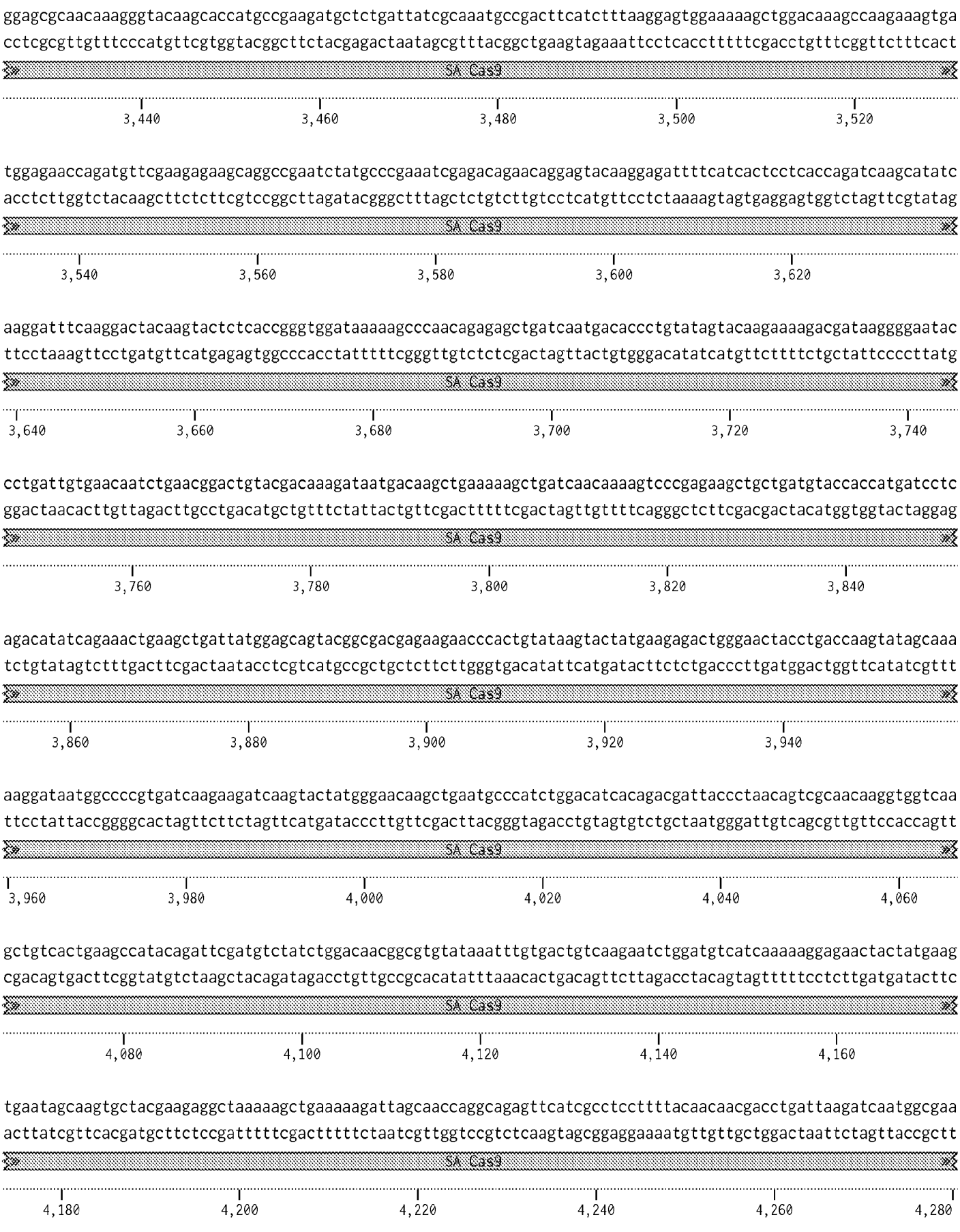

ggagcgcaacaaagggtacaagcaccatgccgaagatgctctgattatcgcaaatgccgacttcatctttaaggagtggaaaaagctggacaaagccaagaaagtga
cctcgcgttgtttcccatgttcgtggtacggcttctacgagactaatagcgtttacggctgaagtagaaattcctcacctttttcgacctgtttcggttctttcact
SA Cas9
3,440
3,460
3,480
3,500
3,520
tggagaaccagatgttcgaagagaagcaggccgaatctatgcccgaaatcgagacagaacaggagtacaaggagattttcatcactcctcaccagatcaagcatatc
acctcttggtctacaagcttctcttcgtccggcttagatacgggctttagctctgtcttgtcctcatgttcctctaaaagtagtgaggagtggtctagttcgtatag
SA Cas9
3,540
3,560
3,580
3,600
3,620
aaggatttcaaggactacaagtactctcaccgggtggataaaaagcccaacagagagctgatcaatgacaccctgtatagtacaagaaaagacgataaggggaatac
ttcctaaagttcctgatgttcatgagagtggcccacctatttttcgggttgtctctcgactagttactgtgggacatatcatgttcttttctgctattccccttatg
SA Cas9
3,640
3,660
3,680
3,700
3,720
3,740
cctgattgtgaacaatctgaacggactgtacgacaaagataatgacaagctgaaaaagctgatcaacaaaagtcccgagaagctgctgatgtaccaccatgatcctc
ggactaacacttgttagacttgcctgacatgctgtttctattactgttcgacttttcgactagttgttttcagggctcttcgacgactacatggtggtactaggag
SA Cas9
3,760
3,780
3,800
3,820
3,840
agacatatcagaaactgaagctgattatggagcagtacggcgacgagaagaacccactgtataagtactatgaagagactgggaactacctgaccaagtatagcaaa
tctgtatagtctttgacttcgactaatacctcgtcatgccgctgctcttcttgggtgacatattcatgatacttctctgacccttgatggactggttcatatcgttt
SA Cas9
3,860
3,880
3,900
3,920
3,940
aaggataatggccccgtgatcaagaagatcaagtactatgggaacaagctgaatgcccatctggacatcacagacgattaccctaacagtcgcaacaaggtggtcaa
ttcctattaccggggcactagttcttctagttcatgataccccttgttcgacttacgggtagacctgtagtgtctgctaatgggattgtcagcgttgttccaccagtt
SA Cas9
3,960
3,980
4,000
4,020
4,040
4,060
gctgtcactgaagccatacagattcgatgtctatctggacaacggcgtgtataaatttgtgactgtcaagaatctggatgtcatcaaaaaggagaactactatgaag
cgacagtgacttcggtatgtctaagctacagatagacctgttgccgcacatatttaaacactgacagttcttagacctacagtagtttttcctcttgatgatacttc
SA Cas9
4,080
4,100
4,120
4,140
4,160
tgaatagcaagtgctacgaagaggctaaaaagctgaaaaagattagcaaccaggcagagttcatcgcctccttttacaacaacgacctgattaagatcaatggcgaa
acttatcgttcacgatgcttctccgatttttcgacttttttctaatcgttggtccgtctcaagtagcggaggaaaatgttgttgctggactaattctagttaccgctt
SA Cas9
4,180
4,200
4,220
4,240
4,260
4,280

FIG. 21E

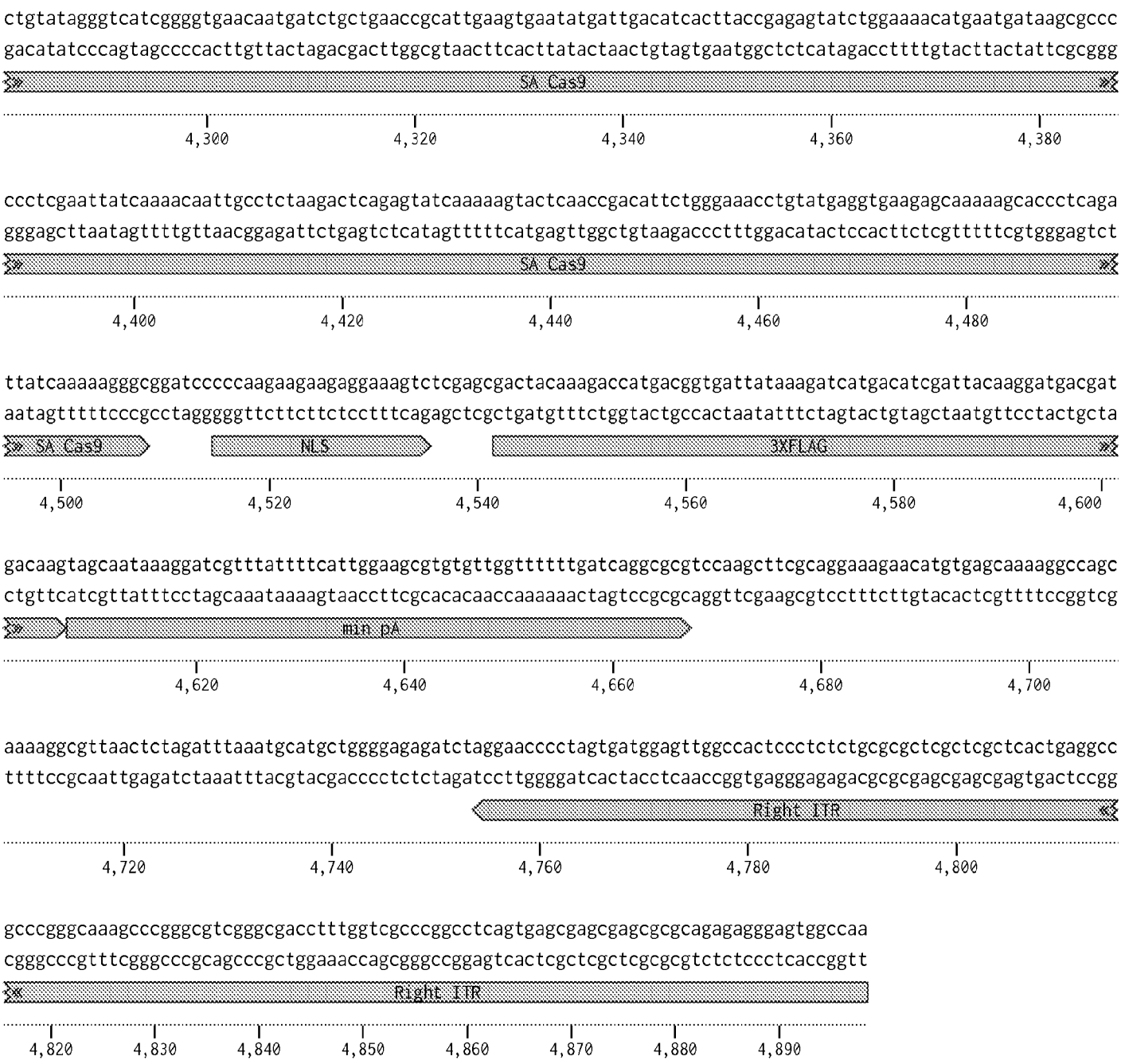

ctgtatagggtcatcggggtgaacaatgatctgctgaaccgcattgaagtgaatatgattgacatcacttaccgagagtatctggaaaacatgaatgataagcgccc
gacatatcccagtagccccacttgttactagacgacttggcgtaacttcacttatactaactgtagtgaatggctctcatagacctttgtacttactattcgcggg
SA Cas9
4,300
4,320
4,340
4,360
4,380
ccctcgaattatcaaaacaattgcctctaagactcagagtatcaaaaagtactcaaccgacattctgggaaacctgtatgaggtgaagagcaaaaagcaccctcaga
gggagcttaatagttttgttaacggagattctgagtctcatagtttttcatgagttggctgtaagacccctttggacatactccacttctcgtttttcgtgggagtct
SA Cas9
4,400
4,420
4,440
4,460
4,480
ttatcaaaaagggcggatcccccaagaagaagaggaaagtctcgagcgactacaaagaccatgacggtgattataaagatcatgacatcgattacaaggatgacgat
aatagtttttcccgcctagggggttcttcttctcctttcagagctcgctgatgtttctggtactgccactaatatttctagtactgtagctaatgttcctactgcta
SA Cas9
NLS
3XFLAG
4,500
4,520
4,540
4,560
4,580
4,600
gacaagtagcaataaaggatcgtttattttcattggaagcgtgtgttggtttttgatcaggcgcgtccaagcttcgcaggaaagaacatgtgagcaaaaggccagc
ctgttcatcgttatttcctagcaaataaaagtaaccttcgcacacaaccaaaaaactagtccgcgcaggttcgaagcgtcctttcttgtacactcgttttccggtcg
min pA
4,620
4,640
4,660
4,680
4,700
aaaaggcgttaactctagatttaaatgcatgctggggagagatctaggaacccctagtgatggagttggccactccctctctgcgcgctcgctcgctcactgaggcc
ttttccgcaattgagatctaaatttacgtacgacccctctctagatccttggggatcactacctcaaccggtgagggagagacgcgcgagcgagcgagtgactccgg
Right ITR
4,720
4,740
4,760
4,780
4,800
gcccgggcaaagcccgggcgtcgggcgacctttggtcgcccggcctcagtgagcgagcgagcgcgcagagagggagtggccaa
cgggcccgtttcgggcccgcagcccgctggaaaccagcgggccggagtcactcgctcgctcgcgcgtctctccctcaccggtt
Right ITR
4,820
4,830
4,840
4,850
4,860
4,870
4,880
4,890

FIG. 21F

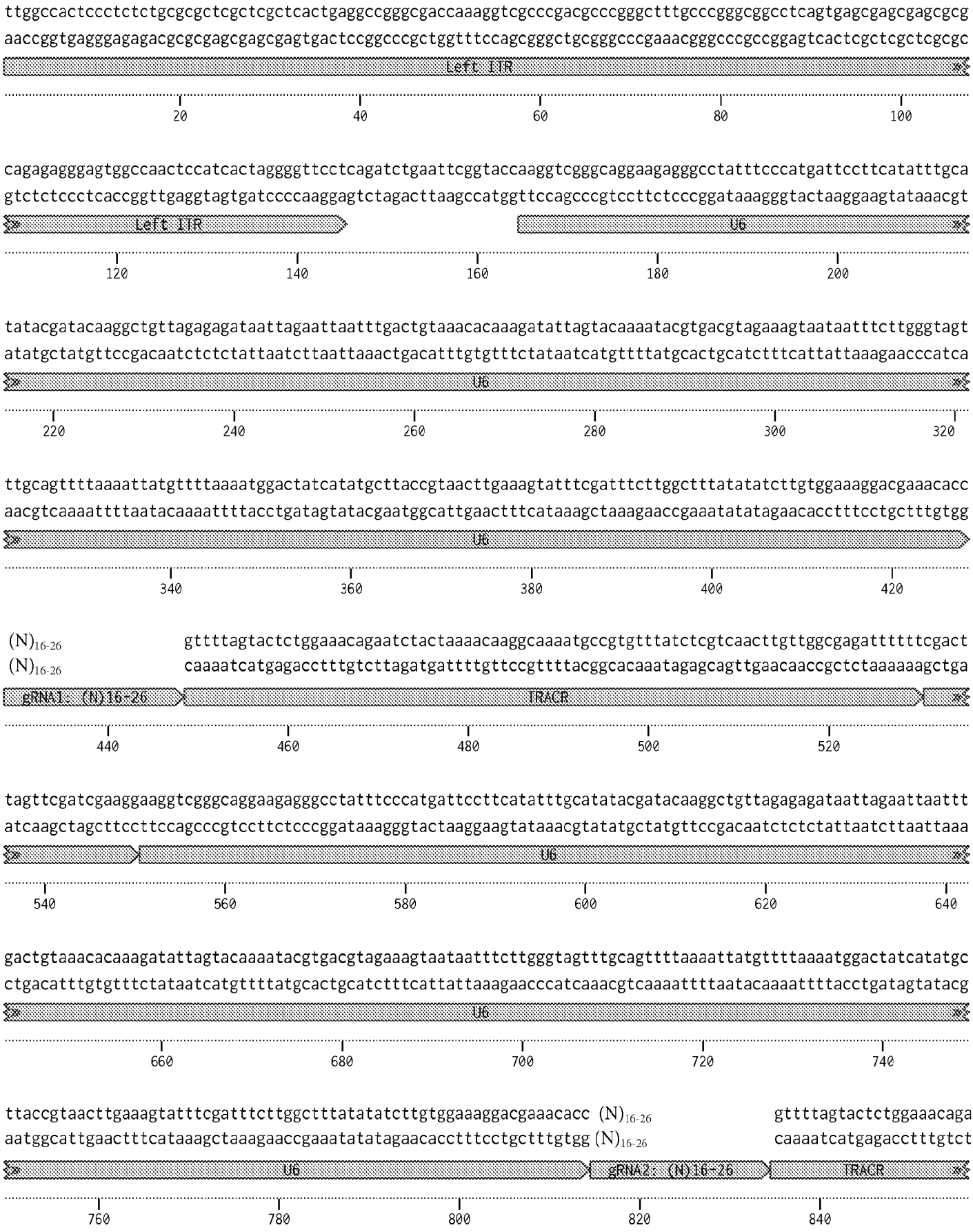

ttggccactccctctctgcgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggcctcagtgagcgagcgagcgcg
aaccggtgagggagagacgcgcgagcgagcgagtgactccggcccgctggtttccagcgggctgcgggcccgaaacgggcccgccggagtcactcgctcgctcgcgc
Left ITR
20
40
60
80
100
cagagagggagtggccaactccatcactaggggttcctcagatctgaattcggtaccaaggtcgggcaggaagagggcctatttcccatgattccttcatatttgca
gtctctccctcaccggttgaggtagtgatccccaaggagtctagacttaagccatggttccagcccgtccttctcccggataaagggtactaaggaagtataaacgt
Left ITR
U6
120
140
160
180
200
tatacgatacaaggctgttagagagataattagaattaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagt
atatgctatgttccgacaatctctctattaatcttaattaaactgacatttgtgtttctataatcatgttttatgcactgcatctttcattattaaagaacccatca
U6
220
240
260
280
300
320
ttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacacc
aacgtcaaaattttaatacaaaattttacctgatagtatacgaatggcattgaactttcataaagctaaagaaccgaaatatatagaacacctttcctgctttgtgg
U6
340
360
380
400
420
(N)16-26 gttttagtactctggaaacagaatctactaaaacaaggcaaaatgccgtgtttatctcgtcaacttgttggcgagatttttcgact
(N)16-26 caaaatcatgagacctttgtcttagatgattttgttccgttttacggcacaaatagagcagttgaacaaccgctctaaaaaagctga
gRNA1: (N)16-26
TRACR
440
460
480
500
520
tagttcgatcgaaggaaggtcgggcaggaagagggcctatttcccatgattccttcatatttgcatatacgatacaaggctgttagagagataattagaattaatttt
atcaagctagcttccttccagcccgtccttctcccggataaagggtactaaggaagtataaacgtatatgctatgttccgacaatctctctattaatcttaattaaa
U6
540
560
580
600
620
640
gactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgttttaaaatggactatcatatgc
ctgacatttgtgtttctataatcatgttttatgcactgcatctttcattattaaagaacccatcaaacgtcaaaattttaatacaaaattttacctgatagtatacg
U6
660
680
700
720
740
ttaccgtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacacc (N)16-26 gttttagtactctggaaacaga
aatggcattgaactttcataaagctaaagaaccgaaatatatagaacacctttcctgctttgtgg (N)16-26 caaaatcatgagacctttgtct
U6
gRNA2: (N)16-26
TRACR
760
780
800
820
840

FIG. 22A

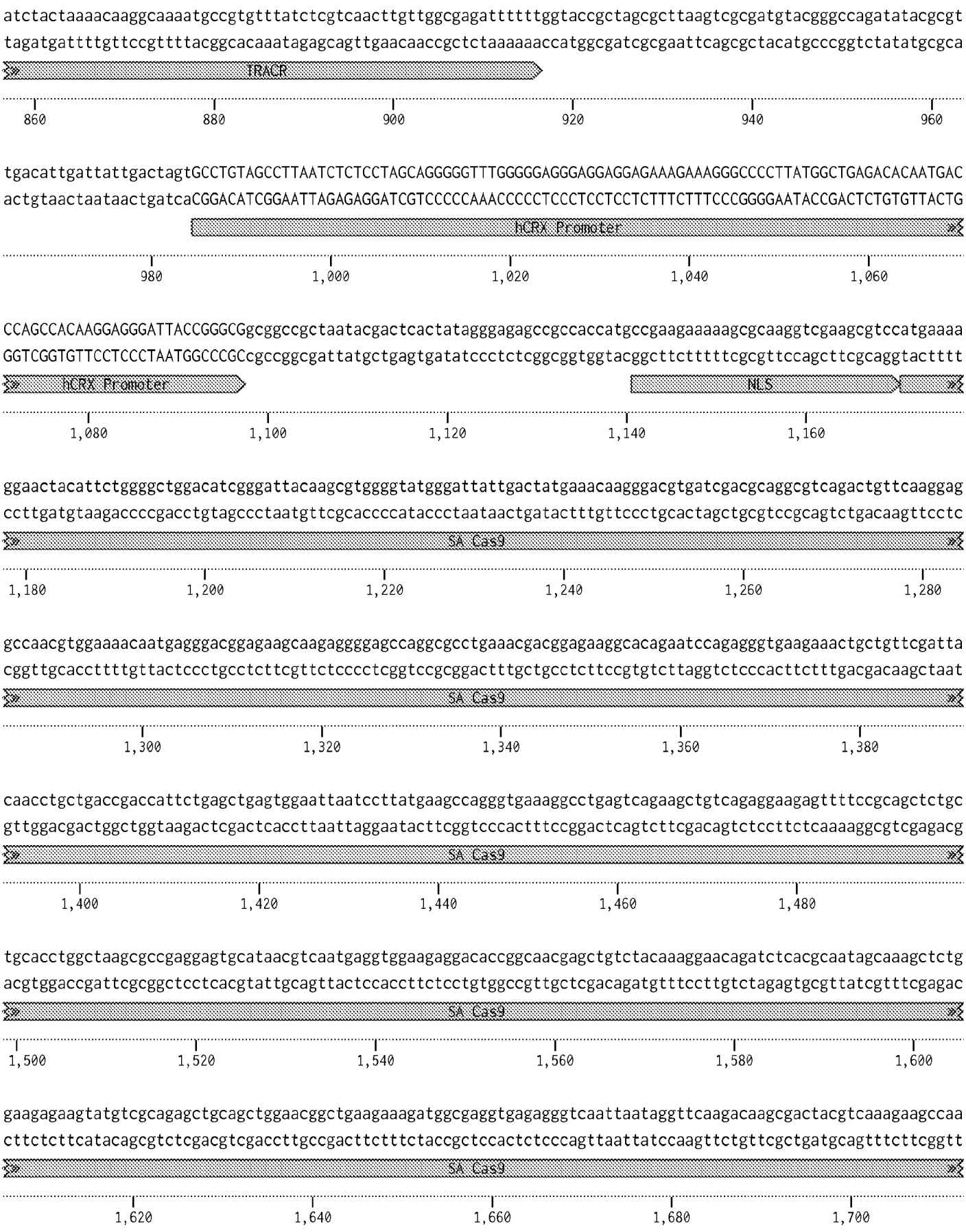

atctactaaaacaaggcaaaatgccgtgtttatctcgtcaacttgttggcgagatttttggtaccgctagcgcttaagtcgcgatgtacgggccagatatacgcgt
tagatgattttgttccgttttacggcacaaatagagcagttgaacaaccgctctaaaaaaccatggcgatcgcgaattcagcgctacatgcccggtctatatgcgca
TRACR
860
880
900
920
940
960
tgacattgattattgactagtGCCTGTAGCCTTAATCTCTCCTAGCAGGGGGTTTGGGGGAGGGAGGAGGAGAAAGAAAGGGCCCCTTATGGCTGAGACACAATGAC
actgtaactaataactgatcaCGGACATCGGAATTAGAGAGGATCGTCCCCCAAACCCCCTCCCTCCTCCTCTTTCTTTCCCGGGGAATACCGACTCTGTGTTACTG
hCRX Promoter
980
1,000
1,020
1,040
1,060
CCAGCCACAAGGAGGGATTACCGGGCGgcggccgctaatacgactcactatagggagagccgccaccatgccgaagaaaaagcgcaaggtcgaagcgtccatgaaaa
GGTCGGTGTTCCTCCCTAATGGCCCGCcgccggcgattatgctgagtgatatccctctcggcggtggtacggcttctttttcgcgttccagcttcgcaggtactttt
hCRX Promoter
NLS
1,080
1,100
1,120
1,140
1,160
ggaactacattctggggctggacatcgggattacaagcgtggggtatgggattattgactatgaaacaagggacgtgatcgacgcaggcgtcagactgttcaaggag
ccttgatgtaagaccccgacctgtagccctaatgttcgcaccccataccctaataactgatactttgttccctgcactagctgcgtccgcagtctgacaagttcctc
SA Cas9
1,180
1,200
1,220
1,240
1,260
1,280
gccaacgtggaaaacaatgagggacggagaagcaagaggggagccaggcgcctgaaacgacggagaaggcacagaatccagagggtgaagaaactgctgttcgatta
cggttgcacctttttgttactccctgcctcttcgttctcccctcggtccgcggactttgctgcctcttccgtgtcttaggtctcccacttctttgacgacaagctaat
SA Cas9
1,300
1,320
1,340
1,360
1,380
caacctgctgaccgaccattctgagctgagtggaattaatccttatgaagccagggtgaaaggcctgagtcagaagctgtcagaggaagagttttccgcagctctgc
gttggacgactggctggtaagactcgactcaccttaattaggaatacttcggtcccactttccggactcagtcttcgacagtctccttctcaaaaggcgtcgagacg
SA Cas9
1,400
1,420
1,440
1,460
1,480
tgcacctggctaagcgccgaggagtgcataacgtcaatgaggtggaagaggacaccggcaacgagctgtctacaaaggaacagatctcacgcaatagcaaagctctg
acgtggaccgattcgcggctcctcacgtattgcagttactccaccttctcctgtggccgttgctcgacagatgtttccttgtctagagtgcgttatcgtttcgagac
SA Cas9
1,500
1,520
1,540
1,560
1,580
1,600
gaagagaagtatgtcgcagagctgcagctggaacggctgaagaaagatggcgaggtgagagggtcaattaataggttcaagacaagcgactacgtcaaagaagccaa
cttctcttcatacagcgtctcgacgtcgaccttgccgacttctttctaccgctccactctcccagttaattatccaagttctgttcgctgatgcagtttcttcggtt
SA Cas9
1,620
1,640
1,660
1,680
1,700

FIG. 22B

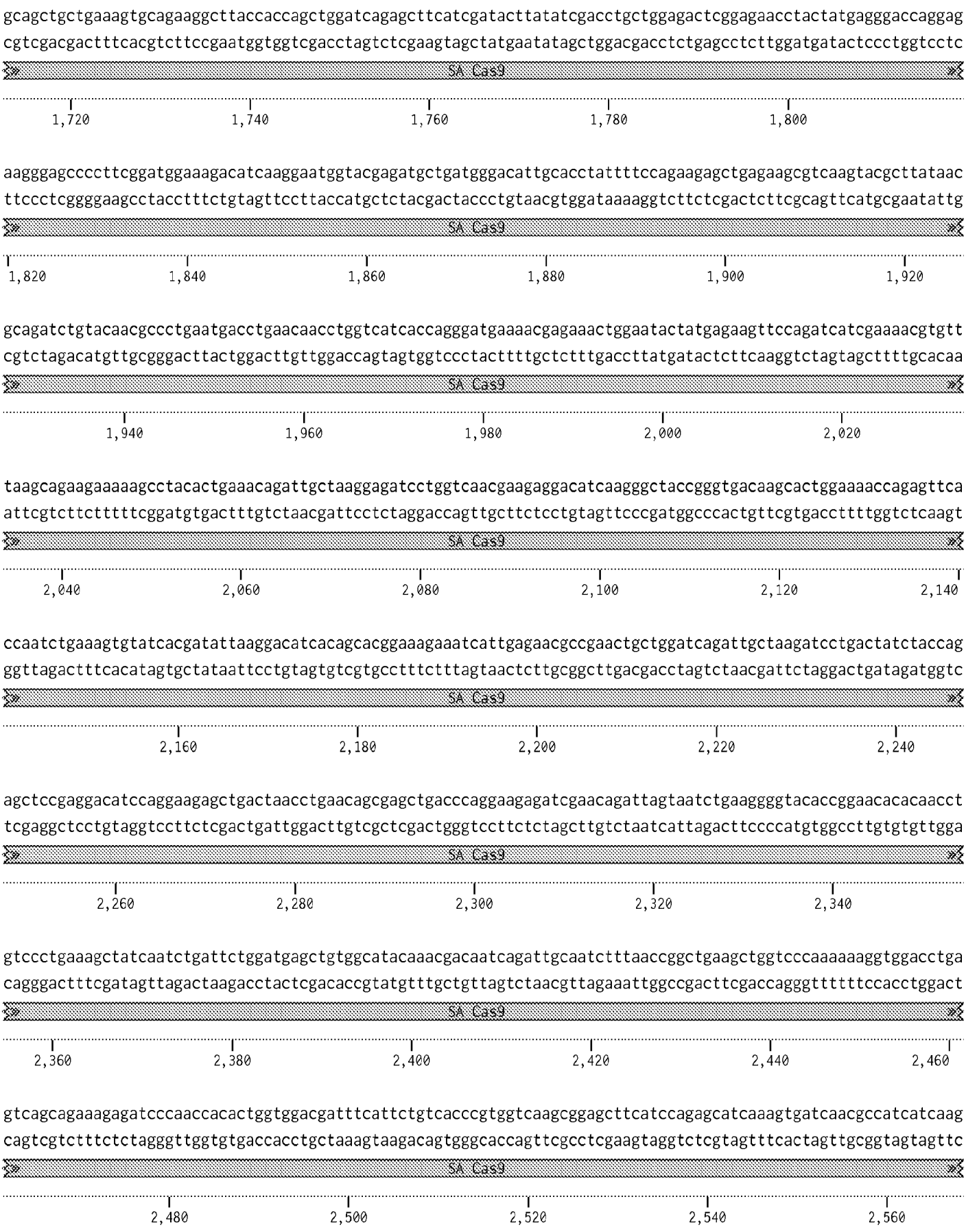

gcagctgctgaaagtgcagaaggcttaccaccagctggatcagagcttcatcgatacttatatcgacctgctggagactcggagaacctactatgagggaccaggag
cgtcgacgactttcacgtcttccgaatggtggtcgacctagtctcgaagtagctatgaatatagctggacgacctctgagcctcttggatgatactccctggtcctc
SA Cas9
1,720 1,740 1,760 1,780 1,800
aagggagccccttcggatggaaagacatcaaggaatggtacgagatgctgatgggacattgcacctattttccagaagagctgagaagcgtcaagtacgcttataac
ttccctcggggaagcctacctttctgtagttccttaccatgctctacgactaccctgtaacgtggataaaaggtcttctcgactcttcgcagttcatgcgaatattg
SA Cas9
1,820 1,840 1,860 1,880 1,900 1,920
gcagatctgtacaacgccctgaatgacctgaacaacctggtcatcaccagggatgaaaacgagaaactggaatactatgagaagttccagatcatcgaaaacgtgtt
cgtctagacatgttgcgggacttactggacttgttggaccagtagtggtccctacttttgctctttgaccttatgatactcttcaaggtctagtagcttttgcacaa
SA Cas9
1,940 1,960 1,980 2,000 2,020
taagcagaagaaaaagcctacactgaaacagattgctaaggagatcctggtcaacgaagaggacatcaagggctaccgggtgacaagcactggaaaaccagagttca
attcgtcttctttttcggatgtgactttgtctaacgattcctctaggaccagttgcttctcctgtagttcccgatggcccactgttcgtgacctttggtctcaagt
SA Cas9
2,040 2,060 2,080 2,100 2,120 2,140
ccaatctgaaagtgtatcacgatattaaggacatcacagcacggaaagaaatcattgagaacgccgaactgctggatcagattgctaagatcctgactatctaccag
ggttagactttcacatagtgctataattcctgtagtgtcgtgcctttctttagtaactcttgcggcttgacgacctagtctaacgattctaggactgatagatggtc
SA Cas9
2,160 2,180 2,200 2,220 2,240
agctccgaggacatccaggaagagctgactaacctgaacagcgagctgacccaggaagagatcgaacagattagtaatctgaaggggtacaccggaacacacaacct
tcgaggctcctgtaggtccttctcgactgattggacttgtcgctcgactgggtccttctctagcttgtctaatcattagacttccccatgtggccttgtgtgttgga
SA Cas9
2,260 2,280 2,300 2,320 2,340
gtccctgaaagctatcaatctgattctggatgagctgtggcatacaaacgacaatcagattgcaatctttaaccggctgaagctggtcccaaaaaaggtggacctga
cagggactttcgatagttagactaagacctactcgacaccgtatgtttgctgttagtctaacgttagaaattggccgacttcgaccagggtttttttccacctggact
SA Cas9
2,360 2,380 2,400 2,420 2,440 2,460
gtcagcagaaagagatcccaaccacactggtggacgatttcattctgtcacccgtggtcaagcggagcttcatccagagcatcaaagtgatcaacgccatcatcaag
cagtcgtctttctctagggttggtgtgaccacctgctaaagtaagacagtgggcaccagttcgcctcgaagtaggtctcgtagtttcactagttgcggtagtagttc
SA Cas9
2,480 2,500 2,520 2,540 2,560

FIG. 22C

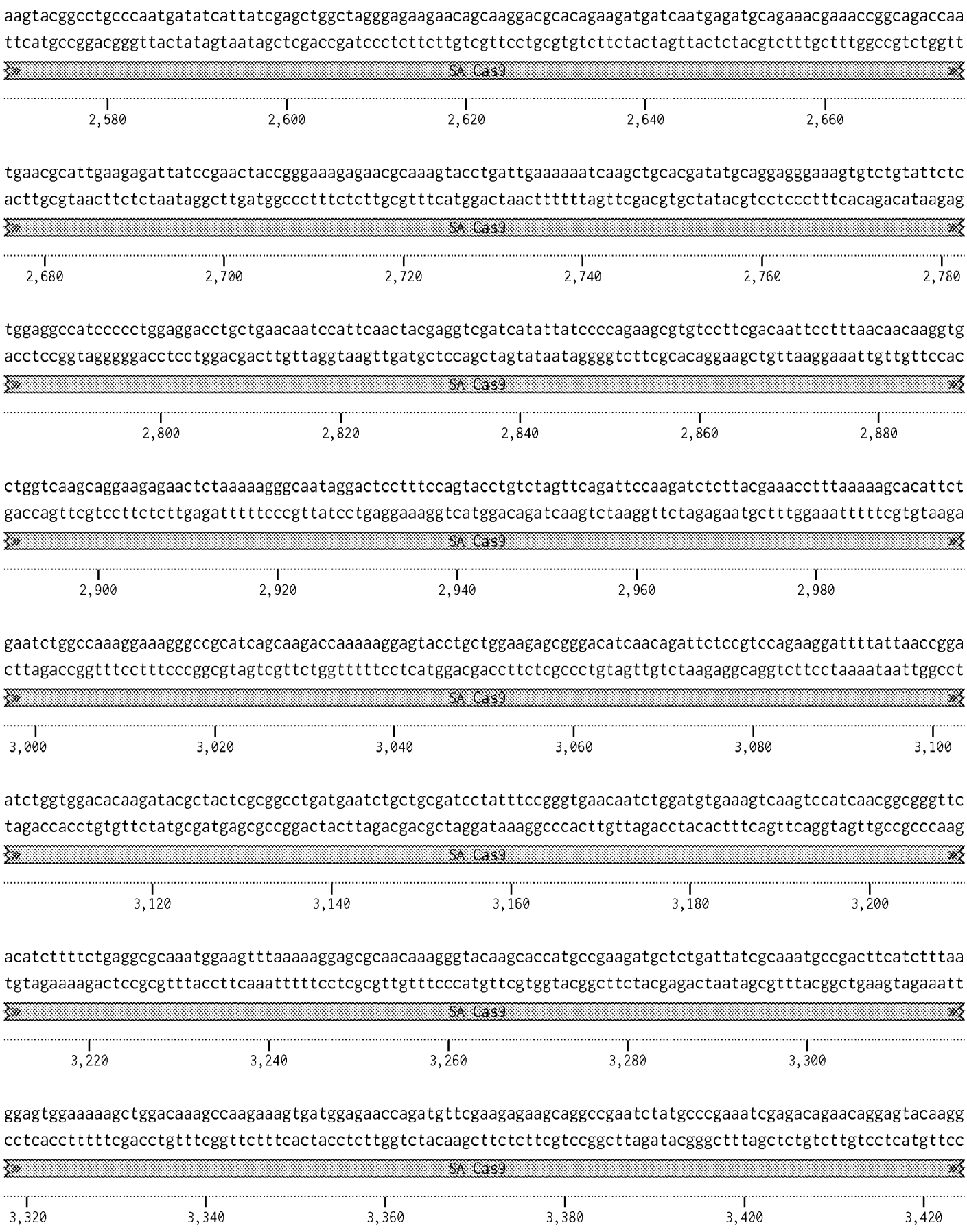

aagtacggcctgcccaatgatatcattatcgagctggctagggagaagaacagcaaggacgcacagaagatgatcaatgagatgcagaaacgaaaccggcagaccaa
ttcatgccggacgggttactatagtaatagctcgaccgatccctcttcttgtcgttcctgcgtgtcttctactagttactctacgtctttgctttggccgtctggtt
SA Cas9
2,580
2,600
2,620
2,640
2,660
tgaacgcattgaagagattatccgaactaccgggaaagagaacgcaaagtacctgattgaaaaaatcaagctgcacgatatgcaggagggaaagtgtctgtattctc
acttgcgtaacttctctaataggcttgatggccctttctcttgcgtttcatggactaacttttttagttcgacgtgctatacgtcctccctttcacagacataagag
SA Cas9
2,680
2,700
2,720
2,740
2,760
2,780
tggaggccatccccctggaggacctgctgaacaatccattcaactacgaggtcgatcatattatccccagaagcgtgtccttcgacaattcctttaacaacaaggtg
acctccggtaggggacctcctggacgacttgttaggtaagttgatgctccagctagtataataggggtcttcgcacaggaagctgttaaggaaattgttgttccac
SA Cas9
2,800
2,820
2,840
2,860
2,880
ctggtcaagcaggaagagaactctaaaaagggcaataggactcctttccagtacctgtctagttcagattccaagatctcttacgaaacctttaaaaagcacattct
gaccagttcgtccttctcttgagatttttcccgttatcctgaggaaaggtcatggacagatcaagtctaaggttctagagaatgctttggaaatttttcgtgtaaga
SA Cas9
2,900
2,920
2,940
2,960
2,980
gaatctggccaaaggaaagggccgcatcagcaagaccaaaaaggagtacctgctggaagagcgggacatcaacagattctccgtccagaaggattttattaaccgga
cttagaccggtttcctttcccggcgtagtcgttctggtttttcctcatggacgaccttctcgccctgtagttgtctaagaggcaggtcttcctaaaataattggcct
SA Cas9
3,000
3,020
3,040
3,060
3,080
3,100
atctggtggacacaagatacgctactcgcggcctgatgaatctgctgcgatcctatttccgggtgaacaatctggatgtgaaagtcaagtccatcaacggcgggttc
tagaccacctgtgttctatgcgatgagcgccggactacttagacgacgctaggataaaggcccacttgttagacctacactttcagttcaggtagttgccgcccaag
SA Cas9
3,120
3,140
3,160
3,180
3,200
acatcttttctgaggcgcaaatggaagtttaaaaaggagcgcaacaaagggtacaagcaccatgccgaagatgctctgattatcgcaaatgccgacttcatctttaa
tgtagaaaagactccgcgtttaccttcaaatttttcctcgcgttgtttcccatgttcgtggtacggcttctacgagactaatagcgtttacggctgaagtagaaatt
SA Cas9
3,220
3,240
3,260
3,280
3,300
ggagtggaaaaagctggacaaagccaagaaagtgatggagaaccagatgttcgaagagaagcaggccgaatctatgcccgaaatcgagacagaacaggagtacaagg
cctcacctttttcgacctgtttcggttctttcactacctcttggtctacaagcttctcttcgtccggcttagatacgggctttagctctgtcttgtcctcatgttcc
SA Cas9
3,320
3,340
3,360
3,380
3,400
3,420

FIG. 22D

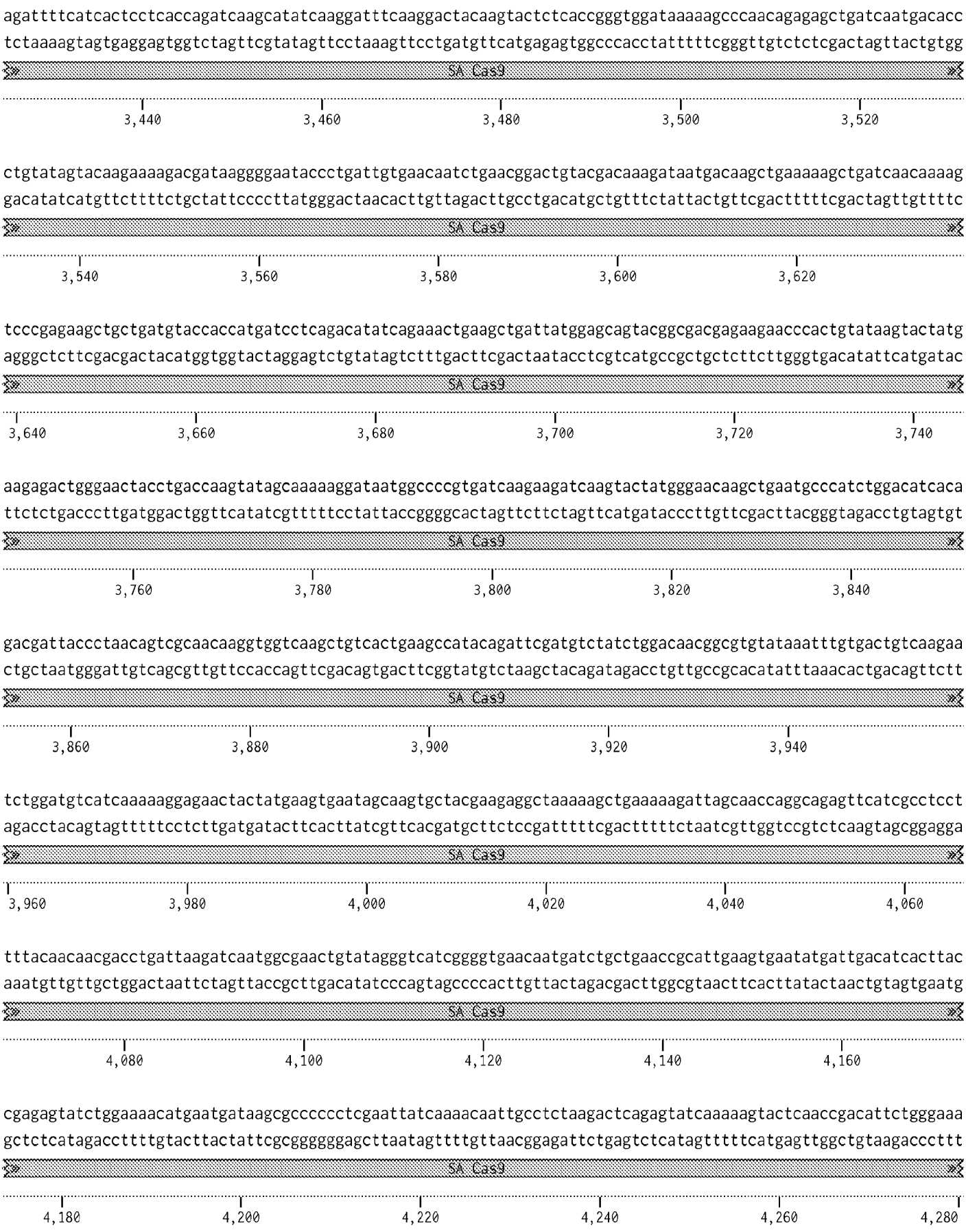

agattttcatcactcctcaccagatcaagcatatcaaggatttcaaggactacaagtactctcaccgggtggataaaaagcccaacagagagctgatcaatgacacc
tctaaaagtagtgaggagtggtctagttcgtatagttcctaaagttcctgatgttcatgagagtggcccacctatttttcgggttgtctctcgactagttactgtgg
SA Cas9
3,440 3,460 3,480 3,500 3,520
ctgtatagtacaagaaaagacgataaggggaataccctgattgtgaacaatctgaacggactgtacgacaaagataatgacaagctgaaaaagctgatcaacaaaag
gacatatcatgttcttttctgctattccccttatgggactaacacttgttagacttgcctgacatgctgtttctattactgttcgacttttcgactagttgttttc
SA Cas9
3,540 3,560 3,580 3,600 3,620
tcccgagaagctgctgatgtaccaccatgatcctcagacatatcagaaactgaagctgattatggagcagtacggcgacgagaagaacccactgtataagtactatg
agggctcttcgacgactacatggtggtactaggagtctgtatagtctttgacttcgactaatacctcgtcatgccgctgctcttcttgggtgacatattcatgatac
SA Cas9
3,640 3,660 3,680 3,700 3,720 3,740
aagagactgggaactacctgaccaagtatagcaaaaaggataatggccccgtgatcaagaagatcaagtactatgggaacaagctgaatgcccatctggacatcaca
ttctctgacccttgatggactggttcatatcgtttttcctattaccggggcactagttcttctagttcatgatacccttgttcgacttacgggtagacctgtagtgt
SA Cas9
3,760 3,780 3,800 3,820 3,840
gacgattaccctaacagtcgcaacaaggtggtcaagctgtcactgaagccatacagattcgatgtctatctggacaacggcgtgtataaatttgtgactgtcaagaa
ctgctaatgggattgtcagcgttgttccaccagttcgacagtgacttcggtatgtctaagctacagatagacctgttgccgcacatatttaaacactgacagttctt
SA Cas9
3,860 3,880 3,900 3,920 3,940
tctggatgtcatcaaaaaggagaactactatgaagtgaatagcaagtgctacgaagaggctaaaaagctgaaaaagattagcaaccaggcagagttcatcgcctcct
agacctacagtagtttttcctcttgatgatacttcacttatcgttcacgatgcttctccgatttttcgactttttctaatcgttggtccgtctcaagtagcggagga
SA Cas9
3,960 3,980 4,000 4,020 4,040 4,060
tttacaacaacgacctgattaagatcaatggcgaactgtatagggtcatcggggtgaacaatgatctgctgaaccgcattgaagtgaatatgattgacatcacttac
aaatgttgttgctggactaattctagttaccgcttgacatatcccagtagccccacttgttactagacgacttggcgtaacttcacttatactaactgtagtgaatg
SA Cas9
4,080 4,100 4,120 4,140 4,160
cgagagtatctggaaaacatgaatgataagcgcccccctcgaattatcaaaacaattgcctctaagactcagagtatcaaaaagtactcaaccgacattctgggaaa
gctctcatagacctttttgtacttactattcgcgggggggagcttaatagttttgttaacggagattctgagtctcatagtttttcatgagttggctgtaagacccttt
SA Cas9
4,180 4,200 4,220 4,240 4,260 4,280

FIG. 22E

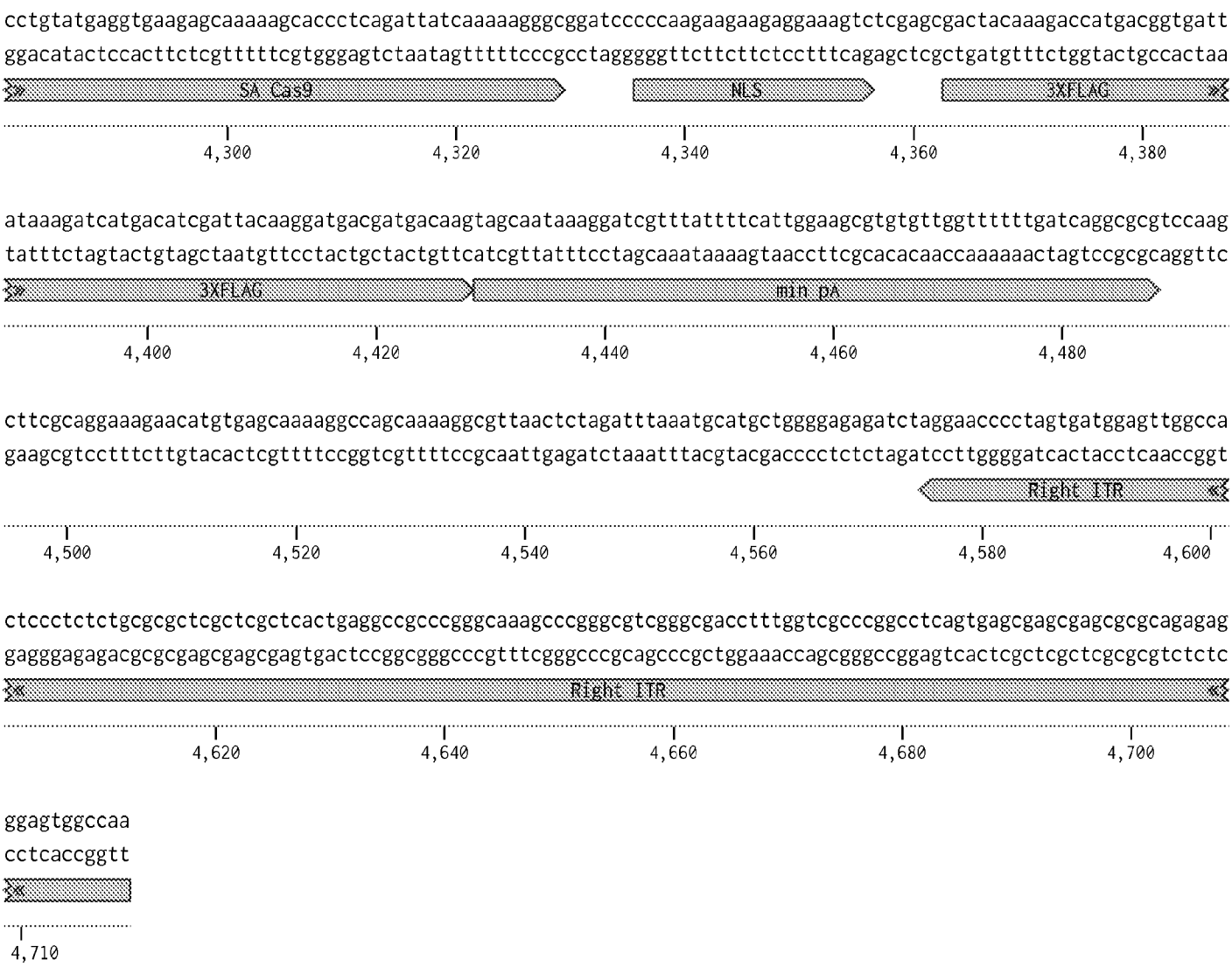

cctgtatgaggtgaagagcaaaaagcaccctcagattatcaaaaagggcggatcccccaagaagaagaggaaagtctcgagcgactacaaagaccatgacggtgatt
ggacatactccacttctcgtttttcgtgggagtctaatagttttttcccgcctagggggttcttcttctcctttcagagctcgctgatgtttctggtactgccactaa
SA Cas9
NLS
3XFLAG
4,300
4,320
4,340
4,360
4,380
ataaagatcatgacatcgattacaaggatgacgatgacaagtagcaataaaggatcgtttattttcattggaagcgtgtgttggtttttgatcaggcgcgtccaag
tatttctagtactgtagctaatgttcctactgctactgttcatcgttatttcctagcaaataaaagtaaccttcgcacacaaccaaaaaactagtccgcgcaggttc
3XFLAG
min pA
4,400
4,420
4,440
4,460
4,480
cttcgcaggaaagaacatgtgagcaaaaggccagcaaaaggcgttaactctagatttaaatgcatgctggggagagatctaggaacccctagtgatggagttggcca
gaagcgtcctttcttgtacactcgttttccggtcgttttccgcaattgagatctaaatttacgtacgacccctctctagatccttggggatcactacctcaaccggt
Right ITR
4,500
4,520
4,540
4,560
4,580
4,600
ctccctctctgcgcgctcgctcgctcactgaggccgcccgggcaaagcccgggcgtcgggcgacctttggtcgcccggcctcagtgagcgagcgagcgcgcagagag
gagggagagacgcgcgagcgagcgagtgactccggcgggcccgtttcgggcccgcagcccgctggaaaccagcgggccggagtcactcgctcgctcgcgcgtctctc
Right ITR
4,620
4,640
4,660
4,680
4,700
ggagtggccaa
cctcaccggtt
4,710

FIG. 22F

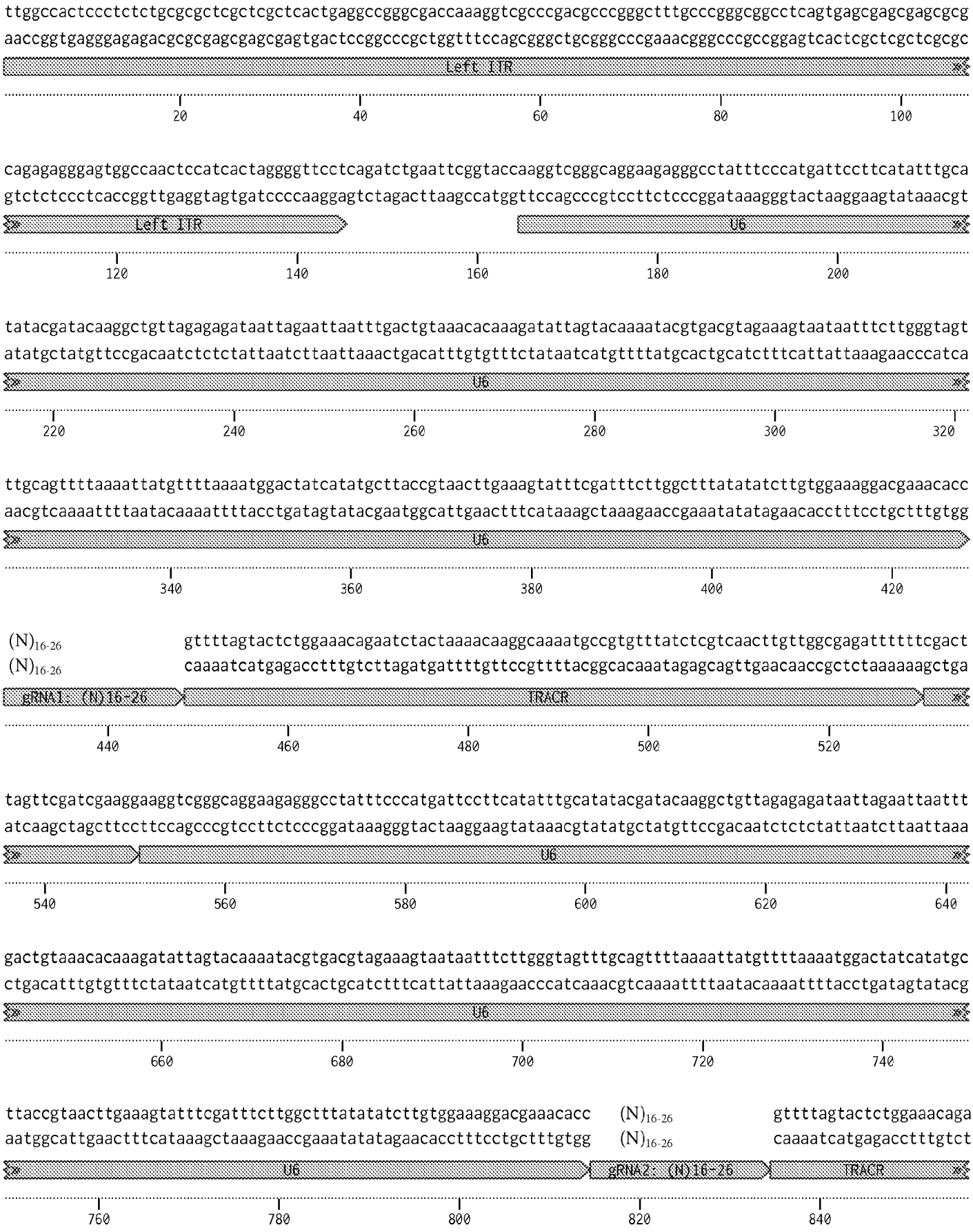

ttggccactccctctctgcgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggcctcagtgagcgagcgagcgcg
aaccggtgagggagagacgcgcgagcgagcgagtgactccggcccgctggtttccagcgggctgcgggcccgaaacgggcccgccggagtcactcgctcgctcgcgc
Left ITR
20
40
60
80
100
cagagagggagtggccaactccatcactaggggttcctcagatctgaattcggtaccaaggtcgggcaggaagagggcctatttcccatgattccttcatatttgca
gtctctccctcaccggttgaggtagtgatccccaaggagtctagacttaagccatggttccagcccgtccttctcccggataaagggtactaaggaagtataaacgt
Left ITR
U6
120
140
160
180
200
tatacgatacaaggctgttagagagataattagaattaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagt
atatgctatgttccgacaatctctctattaatcttaattaaactgacatttgtgtttctataatcatgttttatgcactgcatctttcattattaaagaacccatca
U6
220
240
260
280
300
320
ttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacacc
aacgtcaaaattttaatacaaaattttacctgatagtatacgaatggcattgaactttcataaagctaaagaaccgaaatatatagaacacctttcctgctttgtgg
U6
340
360
380
400
420
(N)16-26 gttttagtactctggaaacagaatctactaaaacaaggcaaaatgccgtgtttatctcgtcaacttgttggcgagatttttcgact
(N)16-26 caaaatcatgagacctttgtcttagatgattttgttccgttttacggcacaaatagagcagttgaacaaccgctctaaaaaagctga
gRNA1: (N)16-26
TRACR
440
460
480
500
520
tagttcgatcgaaggaaggtcgggcaggaagagggcctatttcccatgattccttcatatttgcatatacgatacaaggctgttagagagataattagaattaatttt
atcaagctagcttccttccagcccgtccttctcccggataaagggtactaaggaagtataaacgtatatgctatgttccgacaatctctctattaatcttaattaaa
U6
540
560
580
600
620
640
gactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgttttaaaatggactatcatatgc
ctgacatttgtgtttctataatcatgttttatgcactgcatctttcattattaaagaacccatcaaacgtcaaaattttaatacaaaattttacctgatagtatacg
U6
660
680
700
720
740
ttaccgtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacacc (N)16-26 gttttagtactctggaaacaga
aatggcattgaactttcataaagctaaagaaccgaaatatatagaacacctttcctgctttgtgg (N)16-26 caaaatcatgagacctttgtct
U6
gRNA2: (N)16-26
TRACR
760
780
800
820
840

FIG. 23A

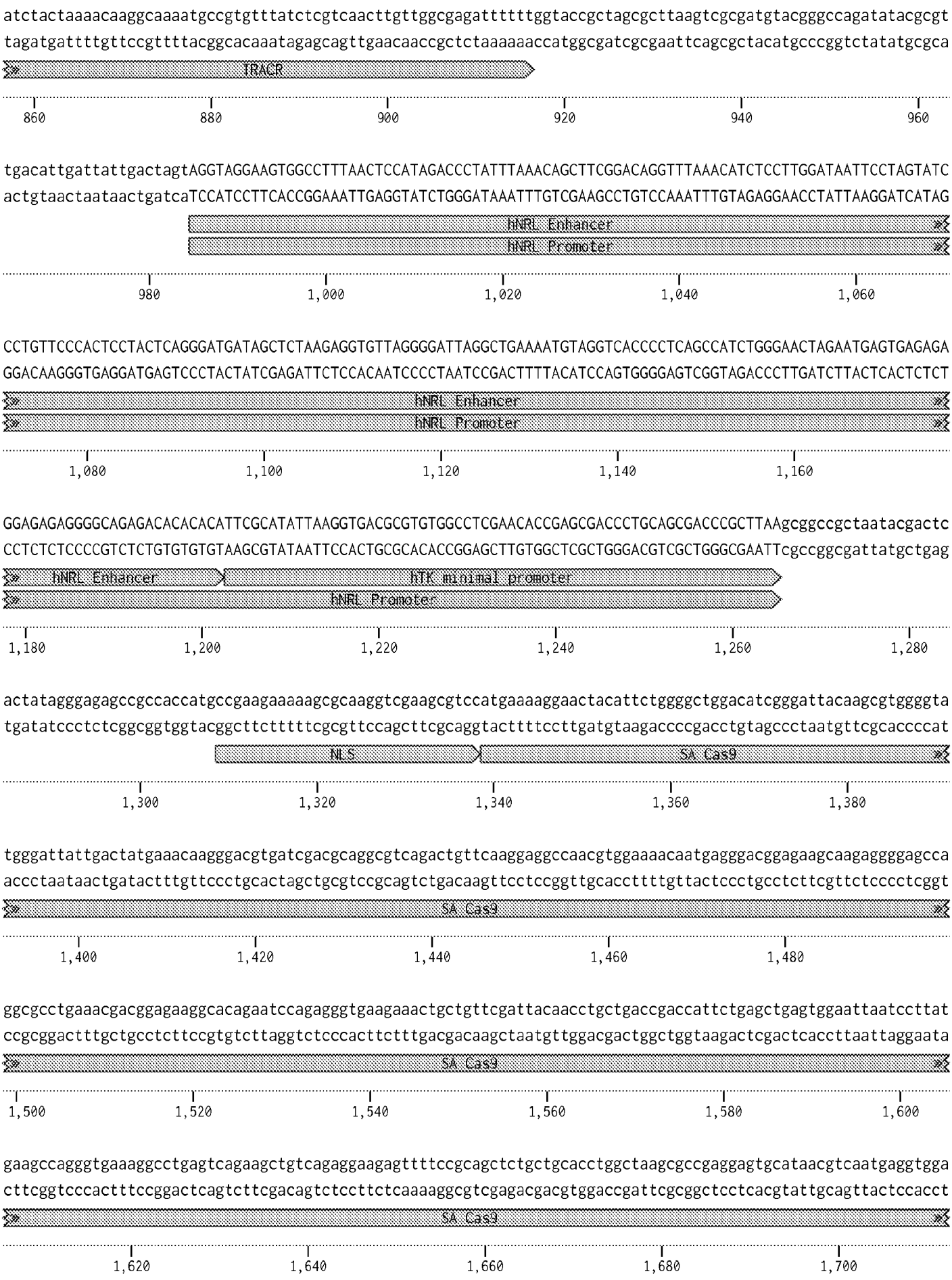
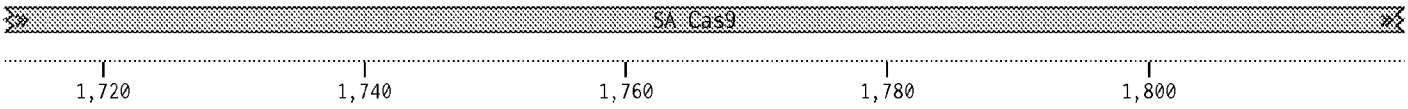
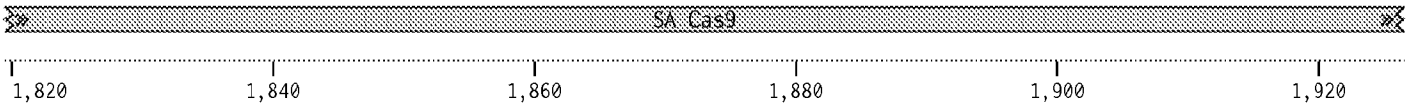

atctactaaaacaaggcaaaatgccgtgtttatctcgtcaacttgttggcgagatttttttggtaccgctagcgcttaagtcgcgatgtacgggccagatatacgcgt
tagatgattttgttccgttttacggcacaaatagagcagttgaacaaccgctctaaaaaaccatggcgatcgcgaattcagcgctacatgcccggtctatatgcgca
TRACR
860
880
900
920
940
960
tgacattgattattgactagtAGGTAGGAAGTGGCCTTTAACTCCATAGACCCTATTTAAACAGCTTCGGACAGGTTTAAACATCTCCTTGGATAATTCCTAGTATC
actgtaactaataactgatcaTCCATCCTTCACCGGAAATTGAGGTATCTGGGATAAATTTGTCGAAGCCTGTCCAAATTTGTAGAGGAACCTATTAAGGATCATAG
hNRL Enhancer
hNRL Promoter
980
1,000
1,020
1,040
1,060
CCTGTTCCCACTCCTACTCAGGGATGATAGCTCTAAGAGGTGTTAGGGGATTAGGCTGAAAATGTAGGTCACCCCTCAGCCATCTGGGAACTAGAATGAGTGAGAGA
GGACAAGGGTGAGGATGAGTCCCTACTATCGAGATTCTCCACAATCCCCTAATCCGACTTTTACATCCAGTGGGGAGTCGGTAGACCCTTGATCTTACTCACTCTCT
hNRL Enhancer
hNRL Promoter
1,080
1,100
1,120
1,140
1,160
GGAGAGAGGGGCAGAGACACACACATTCGCATATTAAGGTGACGCGTGTGGCCTCGAACACCGAGCGACCCTGCAGCGACCCGCTTAAgcggccgctaatacgactc
CCTCTCTCCCCGTCTCTGTGTGTGTAAGCGTATAATTCCACTGCGCACACCGGAGCTTGTGGCTCGCTGGGACGTCGCTGGGCGAATTcgccggcgattatgctgag
hNRL Enhancer
hTK minimal promoter
hNRL Promoter
1,180
1,200
1,220
1,240
1,260
1,280
actatagggagagccgccaccatgccgaagaaaaagcgcaaggtcgaagcgtccatgaaaaggaactacattctggggctggacatcgggattacaagcgtggggta
tgatatccctctcggcggtggtacggcttctttttcgcgttccagcttcgcaggtacttttccttgatgtaagaccccgacctgtagccctaatgttcgcaccccat
NLS
SA Cas9
1,300
1,320
1,340
1,360
1,380
tgggattattgactatgaaacaagggacgtgatcgacgcaggcgtcagactgttcaaggaggccaacgtggaaaacaatgagggacggagaagcaagaggggagcca
accctaataactgatactttgttccctgcactagctgcgtccgcagtctgacaagttcctccggttgcacctttgttactccctgcctcttcgttctcccctcggt
SA Cas9
1,400
1,420
1,440
1,460
1,480
ggcgcctgaaacgacggagaaggcacagaatccagagggtgaagaaactgctgttcgattacaacctgctgaccgaccattctgagctgagtggaattaatccttat
ccgcggactttgctgcctcttccgtgtcttaggtctcccacttctttgacgacaagctaatgttggacgactggctggtaagactcgactcaccttaattaggaata
SA Cas9
1,500
1,520
1,540
1,560
1,580
1,600
gaagccagggtgaaaggcctgagtcagaagctgtcagaggaagagttttccgcagctctgctgcacctggctaagcgccgaggagtgcataacgtcaatgaggtgga
cttcggtcccactttccggactcagtcttcgacagtctccttctcaaaaggcgtcgagacgacgtggaccgattcgcggctcctcacgtattgcagttactccacct
SA Cas9
1,620
1,640
1,660
1,680
1,700

FIG. 23B

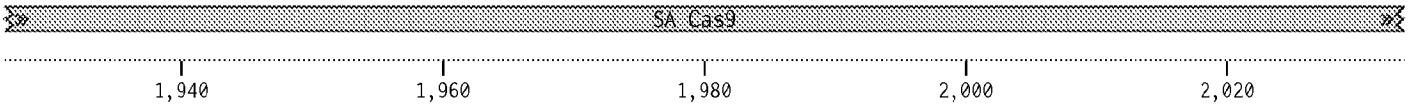
agaggacaccggcaacgagctgtctacaaaggaacagatctcacgcaatagcaaagctctggaagagaagtatgtcgcagagctgcagctggaacggctgaagaaag
tctcctgtggccgttgctcgacagatgtttccttgtctagagtgcgttatcgtttcgagaccttctcttcatacagcgtctcgacgtcgaccttgccgacttctttc



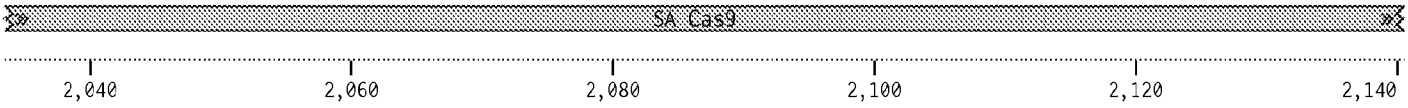
atggcgaggtgagagggtcaattaataggttcaagacaagcgactacgtcaaagaagccaagcagctgctgaaagtgcagaaggcttaccaccagctggatcagagc
taccgctccactctcccagttaattatccaagttctgttcgctgatgcagtttcttcggttcgtcgacgactttcacgtcttccgaatggtggtcgacctagtctcg



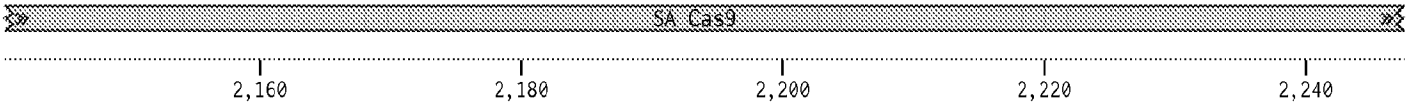
ttcatcgatacttatatcgacctgctggagactcggagaacctactatgagggaccaggagaagggagccccttcggatggaaagacatcaaggaatggtacgagat
aagtagctatgaatatagctggacgacctctgagcctcttggatgatactccctggtcctcttccctcggggaagcctacctttctgtagttccttaccatgctcta



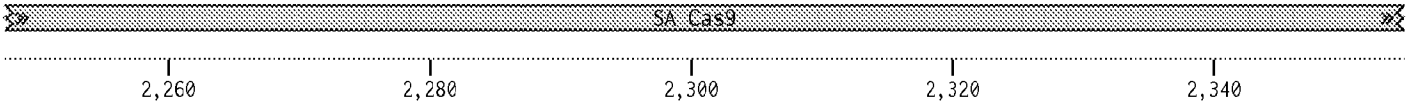
gctgatgggacattgcacctattttccagaagagctgagaagcgtcaagtacgcttataacgcagatctgtacaacgccctgaatgacctgaacaacctggtcatca
cgactaccctgtaacgtggataaaaggtcttctcgactcttcgcagttcatgcgaatattgcgtctagacatgttgcgggacttactggacttgttggaccagtagt



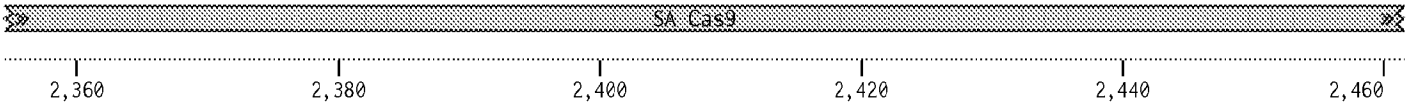
ccagggatgaaaacgagaaactggaatactatgagaagttccagatcatcgaaaacgtgtttaagcagaagaaaaagcctacactgaaacagattgctaaggagatc
ggtccctacttttgctctttgaccttatgatactcttcaaggtctagtagcttttgcacaaattcgtcttctttttcggatgtgactttgtctaacgattcctctag



ctggtcaacgaagaggacatcaagggctaccgggtgacaagcactggaaaaccagagttcaccaatctgaaagtgtatcacgatattaaggacatcacagcacggaa
gaccagttgcttctcctgtagttcccgatggcccactgttcgtgacctttggtctcaagtggttagactttcacatagtgctataattcctgtagtgtcgtgcctt



agaaatcattgagaacgccgaactgctggatcagattgctaagatcctgactatctaccagagctccgaggacatccaggaagagctgactaacctgaacagcgagc
tctttagtaactcttgcggcttgacgacctagtctaacgattctaggactgatagatggtctcgaggctcctgtaggtccttctcgactgattggacttgtcgctcg



tgacccaggaagagatcgaacagattagtaatctgaaggggtacaccggaacacacaacctgtccctgaaagctatcaatctgattctggatgagctgtggcataca
actgggtccttctctagcttgtctaatcattagacttccccatgtggccttgtgtgttggacagggactttcgatagttagactaagacctactcgacaccgtatgt

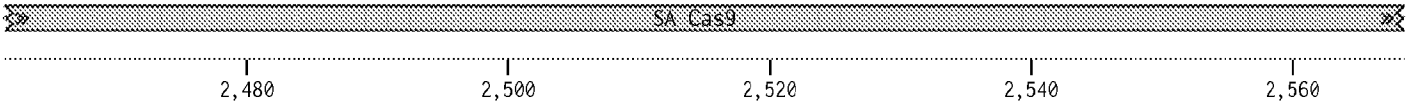


FIG. 23C

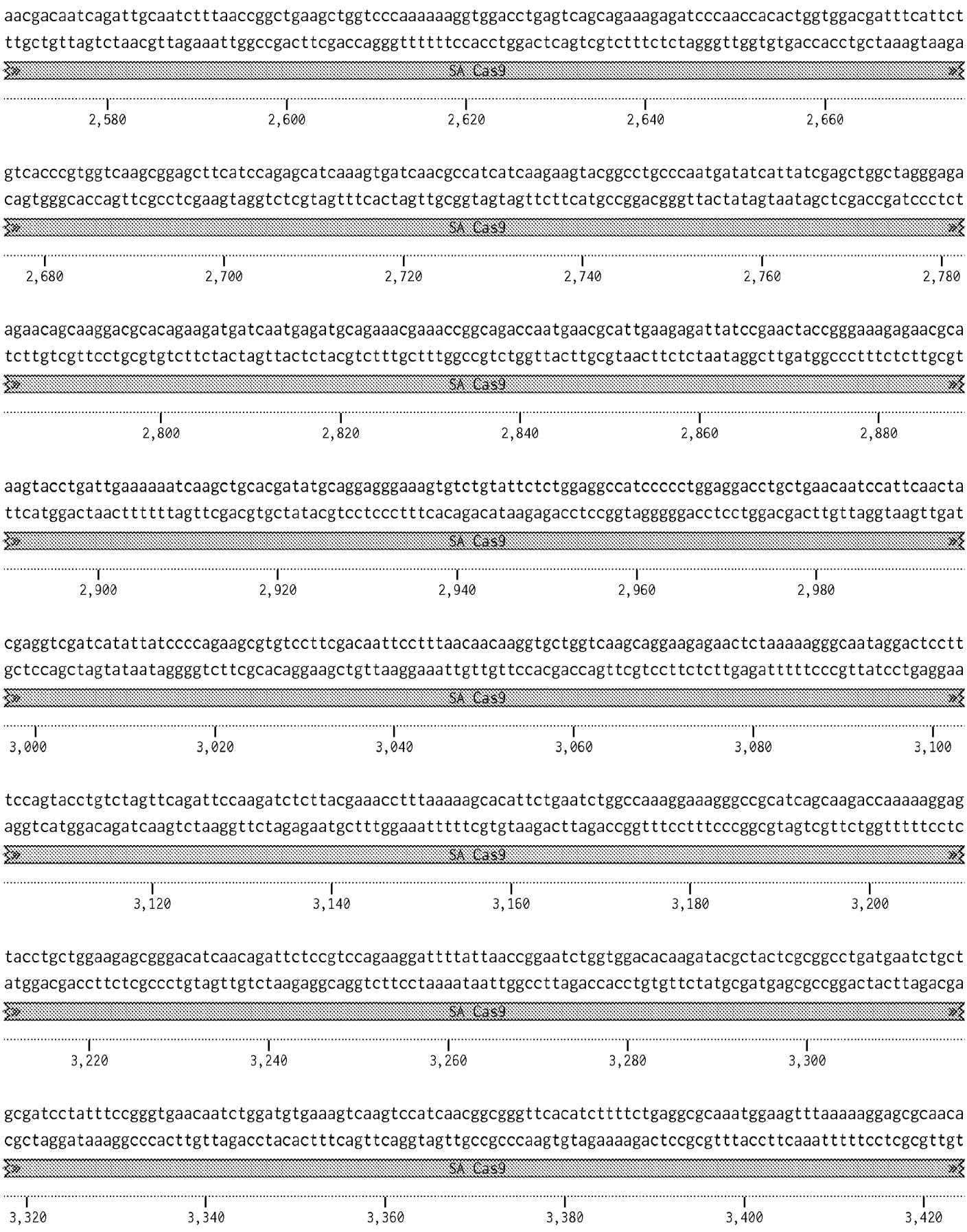
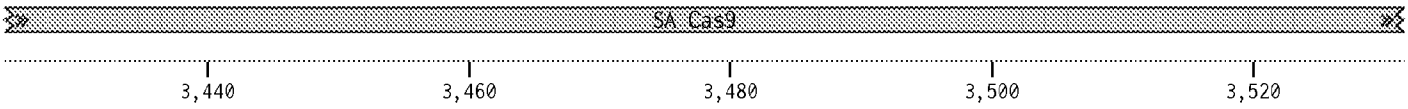
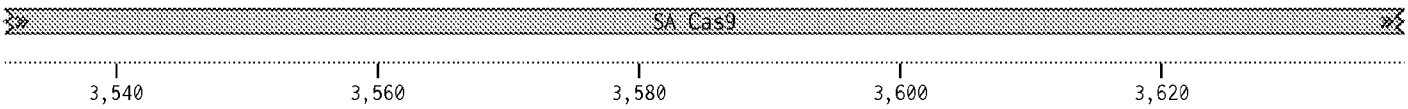

aacgacaatcagattgcaatctttaaccggctgaagctggtcccaaaaaaggtggacctgagtcagcagaaagagatcccaaccacactggtggacgatttcattct
ttgctgttagtctaacgttagaaattggccgacttcgaccagggttttttccacctggactcagtcgtctttctctagggttggtgtgaccacctgctaaagtaaga
SA Cas9
2,580
2,600
2,620
2,640
2,660
gtcacccgtggtcaagcggagcttcatccagagcatcaaagtgatcaacgccatcatcaagaagtacggcctgcccaatgatatcattatcgagctggctagggaga
cagtgggcaccagttcgcctcgaagtaggtctcgtagtttcactagttgcggtagtagttcttcatgccggacgggttactatagtaatagctcgaccgatccctct
SA Cas9
2,680
2,700
2,720
2,740
2,760
2,780
agaacagcaaggacgcacagaagatgatcaatgagatgcagaaacgaaaccggcagaccaatgaacgcattgaagagattatccgaactaccgggaaagagaacgca
tcttgtcgttcctgcgtgtcttctactagttactctacgtctttgctttggccgtctggttacttgcgtaacttctctaataggcttgatggccctttctcttgcgt
SA Cas9
2,800
2,820
2,840
2,860
2,880
aagtacctgattgaaaaaatcaagctgcacgatatgcaggagggaaagtgtctgtattctctggaggccatccccctggaggacctgctgaacaatccattcaacta
ttcatggactaacttttttagttcgacgtgctatacgtcctccctttcacagacataagagacctccggtagggggacctcctggacgacttgttaggtaagttgat
SA Cas9
2,900
2,920
2,940
2,960
2,980
cgaggtcgatcatattatccccagaagcgtgtccttcgacaattcctttaacaacaaggtgctggtcaagcaggaagagaactctaaaaagggcaataggactcctt
gctccagctagtataataggggtcttcgcacaggaagctgttaaggaaattgttgttccacgaccagttcgtccttctcttgagatttttcccgttatcctgaggaa
SA Cas9
3,000
3,020
3,040
3,060
3,080
3,100
tccagtacctgtctagttcagattccaagatctcttacgaaacctttaaaaagcacattctgaatctggccaaaggaaagggccgcatcagcaagaccaaaaaggag
aggtcatggacagatcaagtctaaggttctagagaatgctttggaaatttttcgtgtaagacttagaccggtttcctttcccggcgtagtcgttctggtttttcctc
SA Cas9
3,120
3,140
3,160
3,180
3,200
tacctgctggaagagcgggacatcaacagattctccgtccagaaggattttattaaccggaatctggtggacacaagatacgctactcgcggcctgatgaatctgct
atggacgaccttctcgccctgtagttgtctaagaggcaggtcttcctaaaataattggccttagaccacctgtgttctatgcgatgagcgccggactacttagacga
SA Cas9
3,220
3,240
3,260
3,280
3,300
gcgatcctatttccgggtgaacaatctggatgtgaaagtcaagtccatcaacggcgggttcacatcttttctgaggcgcaaatggaagtttaaaaaggagcgcaaca
cgctaggataaaggcccacttgttagacctacactttcagttcaggtagttgccgcccaagtgtagaaaagactccgcgtttaccttcaaatttttcctcgcgttgt
SA Cas9
3,320
3,340
3,360
3,380
3,400
3,420

FIG. 23D

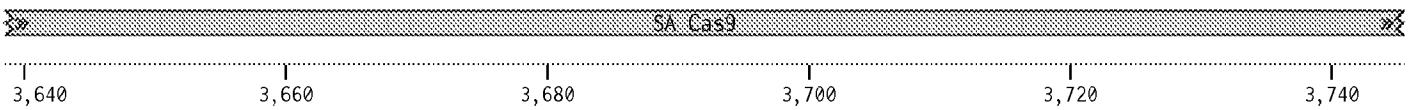
aagggtacaagcaccatgccgaagatgctctgattatcgcaaatgccgacttcatctttaaggagtggaaaaagctggacaaagccaagaaagtgatggagaaccag
ttcccatgttcgtggtacggcttctacgagactaatagcgtttacggctgaagtagaaattcctcacctttttcgacctgtttcggttctttcactacctcttggtc



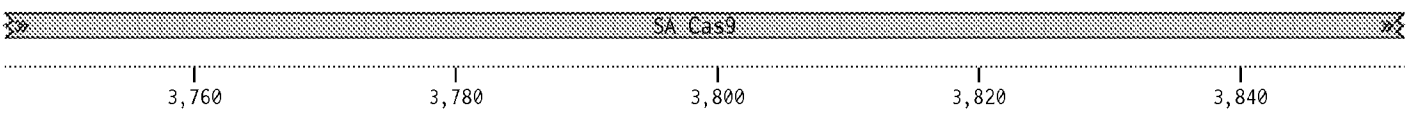
atgttcgaagagaagcaggccgaatctatgcccgaaatcgagacagaacaggagtacaaggagattttcatcactcctcaccagatcaagcatatcaaggatttcaa
tacaagcttctcttcgtccggcttagatacgggctttagctctgtcttgtcctcatgttcctctaaaagtagtgaggagtggtctagttcgtatagttcctaaagtt



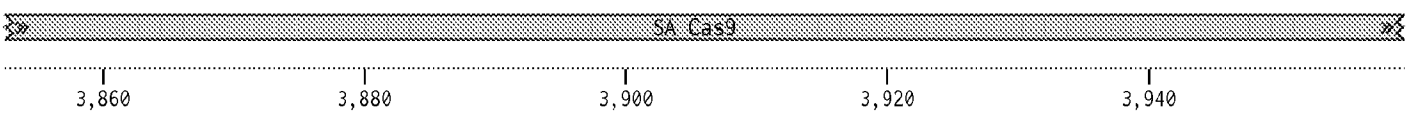
ggactacaagtactctcaccgggtggataaaaagcccaacagagagctgatcaatgacaccctgtatagtacaagaaaagacgataaggggaataccctgattgtga
cctgatgttcatgagagtggcccacctatttttcgggttgtctctcgactagttactgtgggacatatcatgttcttttctgctattccccttatgggactaacact



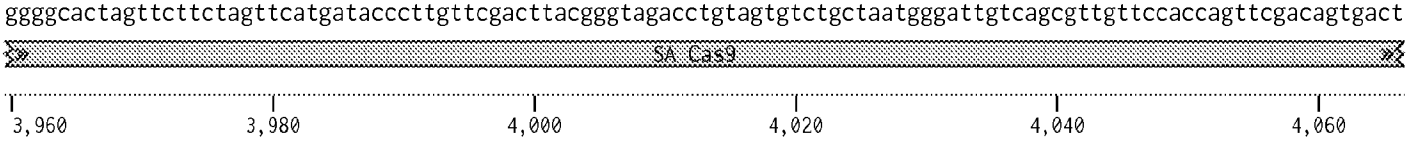
acaatctgaacggactgtacgacaaagataatgacaagctgaaaaagctgatcaacaaaagtcccgagaagctgctgatgtaccaccatgatcctcagacatatcag
tgttagacttgcctgacatgctgtttctattactgttcgacttttcgactagttgttttcagggctcttcgacgactacatggtggtactaggagtctgtatagtc



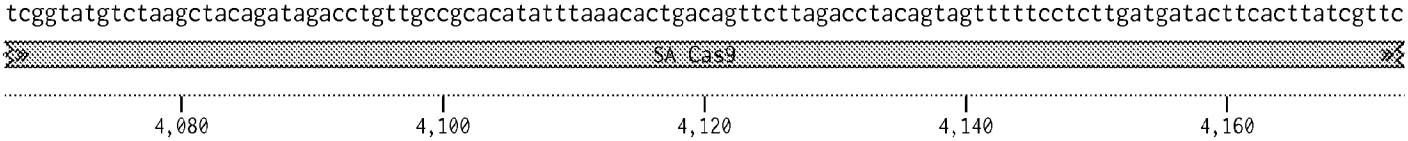
aaactgaagctgattatggagcagtacggcgacgagaagaacccactgtataagtactatgaagagactgggaactacctgaccaagtatagcaaaaaggataatgg
tttgacttcgactaatacctcgtcatgccgctgctcttcttgggtgacatattcatgatacttctctgacccttgatggactggttcatatcgtttttcctattacc



ccccgtgatcaagaagatcaagtactatgggaacaagctgaatgcccatctggacatcacagacgattaccctaacagtcgcaacaaggtggtcaagctgtcactga
ggggcactagttcttctagttcatgataccttgttcgacttacgggtagacctgtagtgtctgctaatgggattgtcagcgttgttccaccagttcgacagtgact



agccatacagattcgatgtctatctggacaacggcgtgtataaatttgtgactgtcaagaatctggatgtcatcaaaaaggagaactactatgaagtgaatagcaag
tcggtatgtctaagctacagatagacctgttgccgcacatatttaaacactgacagttcttagacctacagtagtttttcctcttgatgatacttcacttatcgttc



tgctacgaagaggctaaaaagctgaaaaagattagcaaccaggcagagttcatcgcctccttttacaacaacgacctgattaagatcaatggcgaactgtatagggt
acgatgcttctccgatttttcgactttttctaatcgttggtccgtctcaagtagcggaggaaaatgttgttgctggactaattctagttaccgcttgacatatccca

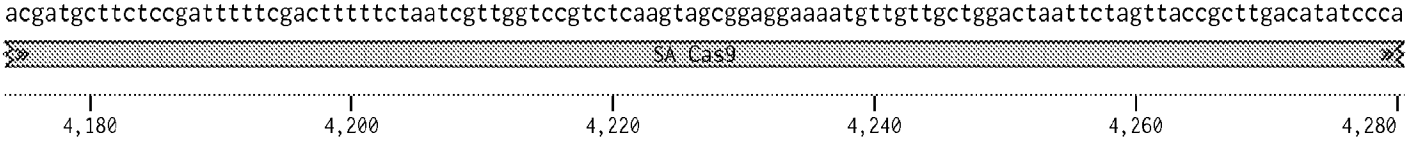


FIG. 23E

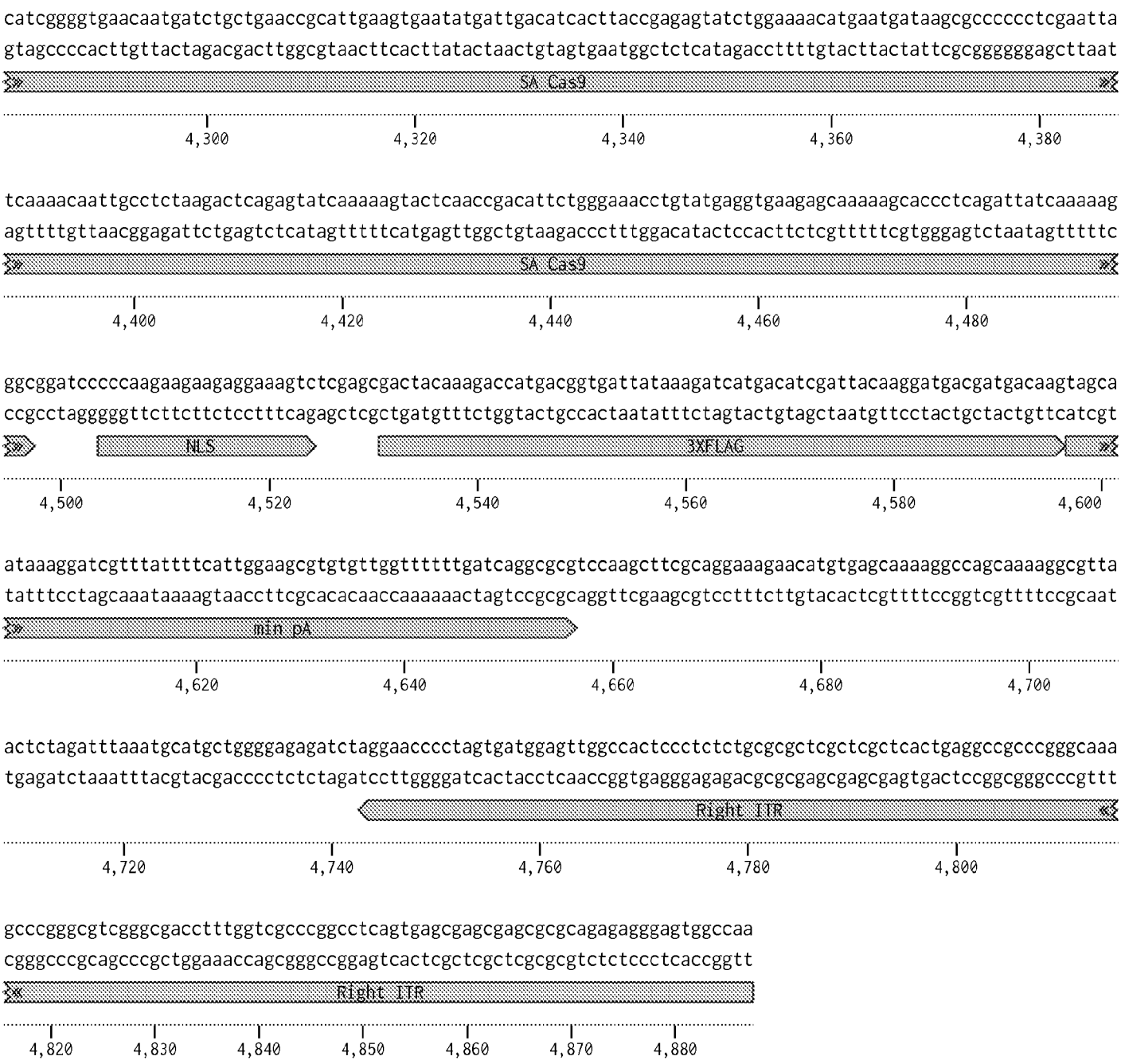

catcggggtgaacaatgatctgctgaaccgcattgaagtgaatatgattgacatcacttaccgagagtatctggaaaacatgaatgataagcgcccccctcgaatta
gtagccccacttgttactagacgacttggcgtaacttcacttatactaactgtagtgaatggctctcatagacctttgtacttactattcgcgggggagcttaat
SA Cas9
4,300
4,320
4,340
4,360
4,380
tcaaaacaattgcctctaagactcagagtatcaaaaagtactcaaccgacattctgggaaacctgtatgaggtgaagagcaaaaagcaccctcagattatcaaaaag
agttttgttaacggagattctgagtctcatagtttttcatgagttggctgtaagacccttggacatactccacttctcgtttttcgtgggagtctaatagttttc
SA Cas9
4,400
4,420
4,440
4,460
4,480
ggcggatcccccaagaagaagaggaaagtctcgagcgactacaaagaccatgacggtgattataaagatcatgacatcgattacaaggatgacgatgacaagtagca
ccgcctagggggttcttcttctcctttcagagctcgctgatgtttctggtactgccactaatatttctagtactgtagctaatgttcctactgctactgttcatcgt
NLS
3XFLAG
4,500
4,520
4,540
4,560
4,580
4,600
ataaaggatcgtttattttcattggaagcgtgtgttggtttttgatcaggcgcgtccaagcttcgcaggaaagaacatgtgagcaaaaggccagcaaaaggcgtta
tatttcctagcaaataaaagtaaccttcgcacacaaccaaaaaactagtccgcgcaggttcgaagcgtcctttcttgtacactcgttttccggtcgttttccgcaat
min pA
4,620
4,640
4,660
4,680
4,700
actctagatttaaatgcatgctggggagagatctaggaacccctagtgatggagttggccactccctctctgcgcgctcgctcgctcactgaggccgcccgggcaaa
tgagatctaaatttacgtacgacccctctctagatccttggggatcactacctcaaccggtgagggagagacgcgcgagcgagcgagtgactccggcgggcccgttt
Right ITR
4,720
4,740
4,760
4,780
4,800
gcccgggcgtcgggcgacctttggtcgcccggcctcagtgagcgagcgagcgcgcagagagggagtggccaa
cgggcccgcagcccgctggaaaccagcgggccggagtcactcgctcgctcgcgcgtctctccctcaccggtt
Right ITR
4,820
4,830
4,840
4,850
4,860
4,870
4,880

FIG. 23F

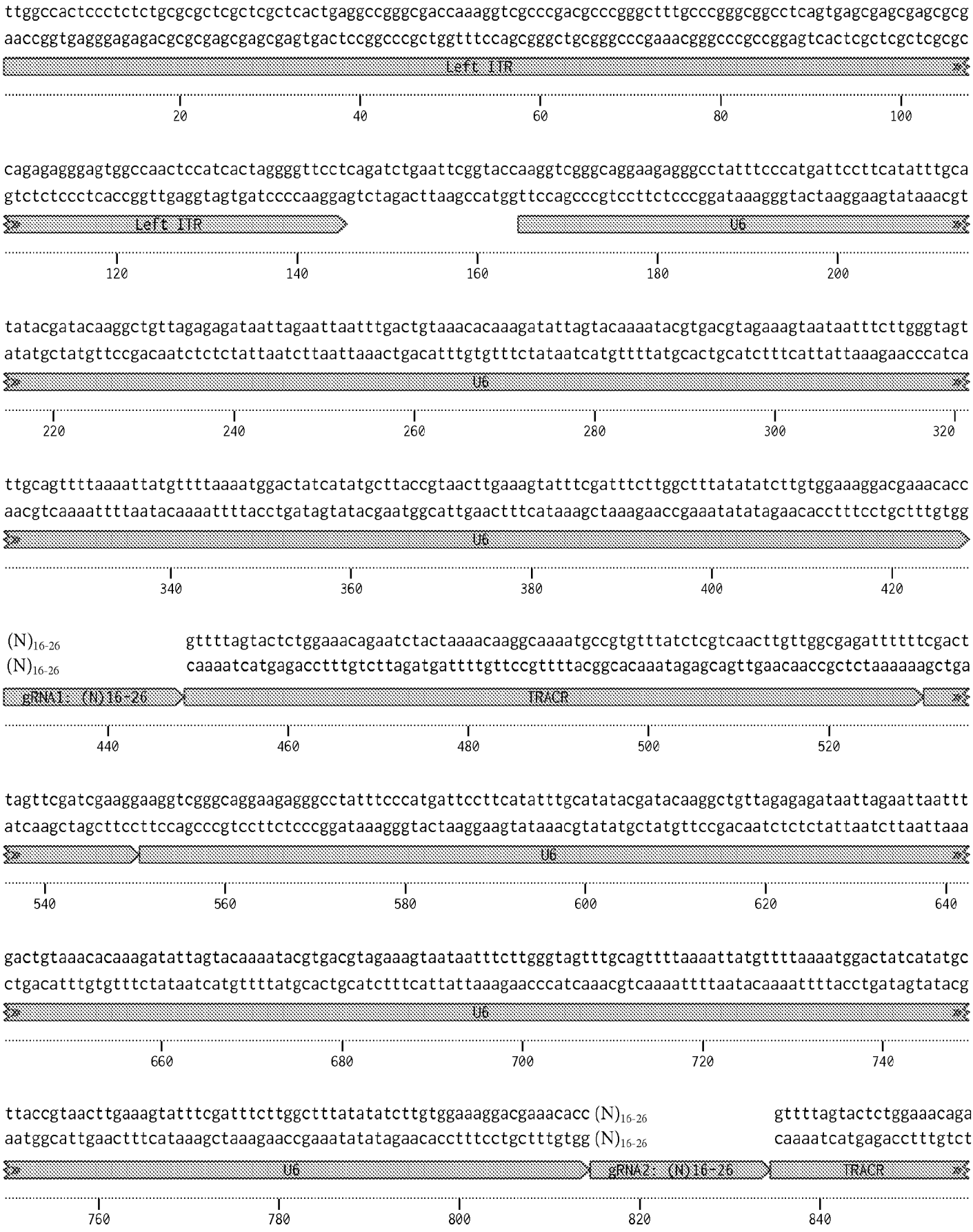

ttggccactccctctctgcgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggcctcagtgagcgagcgagcgcg
aaccggtgagggagagacgcgcgagcgagcgagtgactccggcccgctggtttccagcgggctgcgggcccgaaacgggcccgccggagtcactcgctcgctcgcgc
Left ITR
20
40
60
80
100
cagagagggagtggccaactccatcactaggggttcctcagatctgaattcggtaccaaggtcgggcaggaagagggcctatttcccatgattccttcatatttgca
gtctctccctcaccggttgaggtagtgatccccaaggagtctagacttaagccatggttccagcccgtccttctcccggataaagggtactaaggaagtataaacgt
Left ITR
U6
120
140
160
180
200
tatacgatacaaggctgttagagagataattagaattaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagt
atatgctatgttccgacaatctctctattaatcttaattaaactgacatttgtgtttctataatcatgttttatgcactgcatctttcattattaaagaacccatca
U6
220
240
260
280
300
320
ttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacacc
aacgtcaaaattttaatacaaaattttacctgatagtatacgaatggcattgaactttcataaagctaaagaaccgaaatatatagaacacctttcctgctttgtgg
U6
340
360
380
400
420
(N)16-26 gttttagtactctggaaacagaatctactaaaacaaggcaaaatgccgtgtttatctcgtcaacttgttggcgagatttttcgact
(N)16-26 caaaatcatgagacctttgtcttagatgattttgttccgttttacggcacaaatagagcagttgaacaaccgctctaaaaaagctga
gRNA1: (N)16-26
TRACR
440
460
480
500
520
tagttcgatcgaaggaaggtcgggcaggaagagggcctatttcccatgattccttcatatttgcatatacgatacaaggctgttagagagataattagaattaatttt
atcaagctagcttccttccagcccgtccttctcccggataaagggtactaaggaagtataaacgtatatgctatgttccgacaatctctctattaatcttaattaaa
U6
540
560
580
600
620
640
gactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgttttaaaatggactatcatatgc
ctgacatttgtgtttctataatcatgttttatgcactgcatctttcattattaaagaacccatcaaacgtcaaaattttaatacaaaattttacctgatagtatacg
U6
660
680
700
720
740
ttaccgtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacacc (N)16-26 gttttagtactctggaaacaga
aatggcattgaactttcataaagctaaagaaccgaaatatatagaacacctttcctgctttgtgg (N)16-26 caaaatcatgagacctttgtct
U6
gRNA2: (N)16-26
TRACR
760
780
800
820
840

FIG. 24A

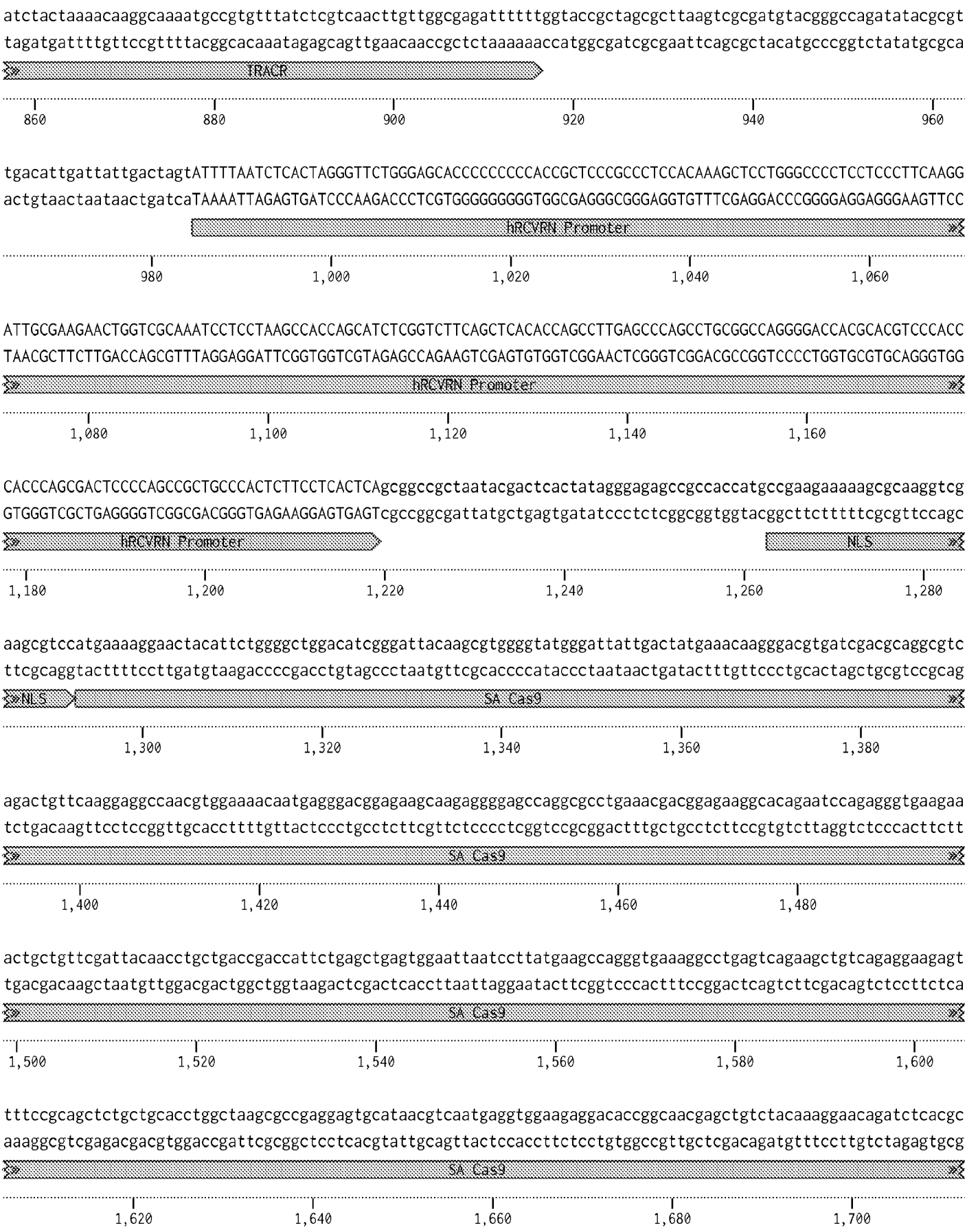

atctactaaaacaaggcaaaatgccgtgtttatctcgtcaacttgttggcgagatttttggtaccgctagcgcttaagtcgcgatgtacgggccagatatacgcgt
tagatgattttgttccgttttacggcacaaatagagcagttgaacaaccgctctaaaaaaccatggcgatcgcgaattcagcgctacatgcccggtctatatgcgca
TRACR
860
880
900
920
940
960
tgacattgattattgactagtATTTTAATCTCACTAGGGTTCTGGGAGCACCCCCCCCCACCGCTCCCGCCCTCCACAAAGCTCCTGGGCCCCTCCTCCCTTCAAGG
actgtaactaataactgatcaTAAAATTAGAGTGATCCCAAGACCCTCGTGGGGGGGGGTGGCGAGGGCGGGAGGTGTTTCGAGGACCCGGGGAGGAGGGAAGTTCC
hRCVRN Promoter
980
1,000
1,020
1,040
1,060
ATTGCGAAGAACTGGTCGCAAATCCTCCTAAGCCACCAGCATCTCGGTCTTCAGCTCACACCAGCCTTGAGCCCAGCCTGCGGCCAGGGGACCACGCACGTCCCACC
TAACGCTTCTTGACCAGCGTTTAGGAGGATTCGGTGGTCGTAGAGCCAGAAGTCGAGTGTGGTCGGAACTCGGGTCGGACGCCGGTCCCCTGGTGCGTGCAGGGTGG
hRCVRN Promoter
1,080
1,100
1,120
1,140
1,160
CACCCAGCGACTCCCCAGCCGCTGCCCACTCTTCCTCACTCAgcggccgctaatacgactcactatagggagagccgccaccatgccgaagaaaaagcgcaaggtcg
GTGGGTCGCTGAGGGGTCGGCGACGGGTGAGAAGGAGTGAGTcgccggcgattatgctgagtgatatccctctcggcggtggtacggcttctttttcgcgttccagc
hRCVRN Promoter
NLS
1,180
1,200
1,220
1,240
1,260
1,280
aagcgtccatgaaaaggaactacattctggggctggacatcgggattacaagcgtggggtatgggattattgactatgaaacaagggacgtgatcgacgcaggcgtc
ttcgcaggtacttttccttgatgtaagaccccgacctgtagccctaatgttcgcaccccataccctaataactgatactttgttccctgcactagctgcgtccgcag
NLS
SA Cas9
1,300
1,320
1,340
1,360
1,380
agactgttcaaggaggccaacgtggaaaacaatgagggacggagaagcaagaggggagccaggcgcctgaaacgacggagaaggcacagaatccagagggtgaagaa
tctgacaagttcctccggttgcacctttgttactccctgcctcttcgttctcccctcggtccgcggactttgctgcctcttccgtgtcttaggtctcccacttctt
SA Cas9
1,400
1,420
1,440
1,460
1,480
actgctgttcgattacaacctgctgaccgaccattctgagctgagtggaattaatccttatgaagccagggtgaaaggcctgagtcagaagctgtcagaggaagagt
tgacgacaagctaatgttggacgactggctggtaagactcgactcaccttaattaggaatacttcggtcccactttccggactcagtcttcgacagtctccttctca
SA Cas9
1,500
1,520
1,540
1,560
1,580
1,600
tttccgcagctctgctgcacctggctaagcgccgaggagtgcataacgtcaatgaggtggaagaggacaccggcaacgagctgtctacaaaggaacagatctcacgc
aaaggcgtcgagacgacgtggaccgattcgcggctcctcacgtattgcagttactccaccttctcctgtggccgttgctcgacagatgtttccttgtctagagtgcg
SA Cas9
1,620
1,640
1,660
1,680
1,700

FIG. 24B

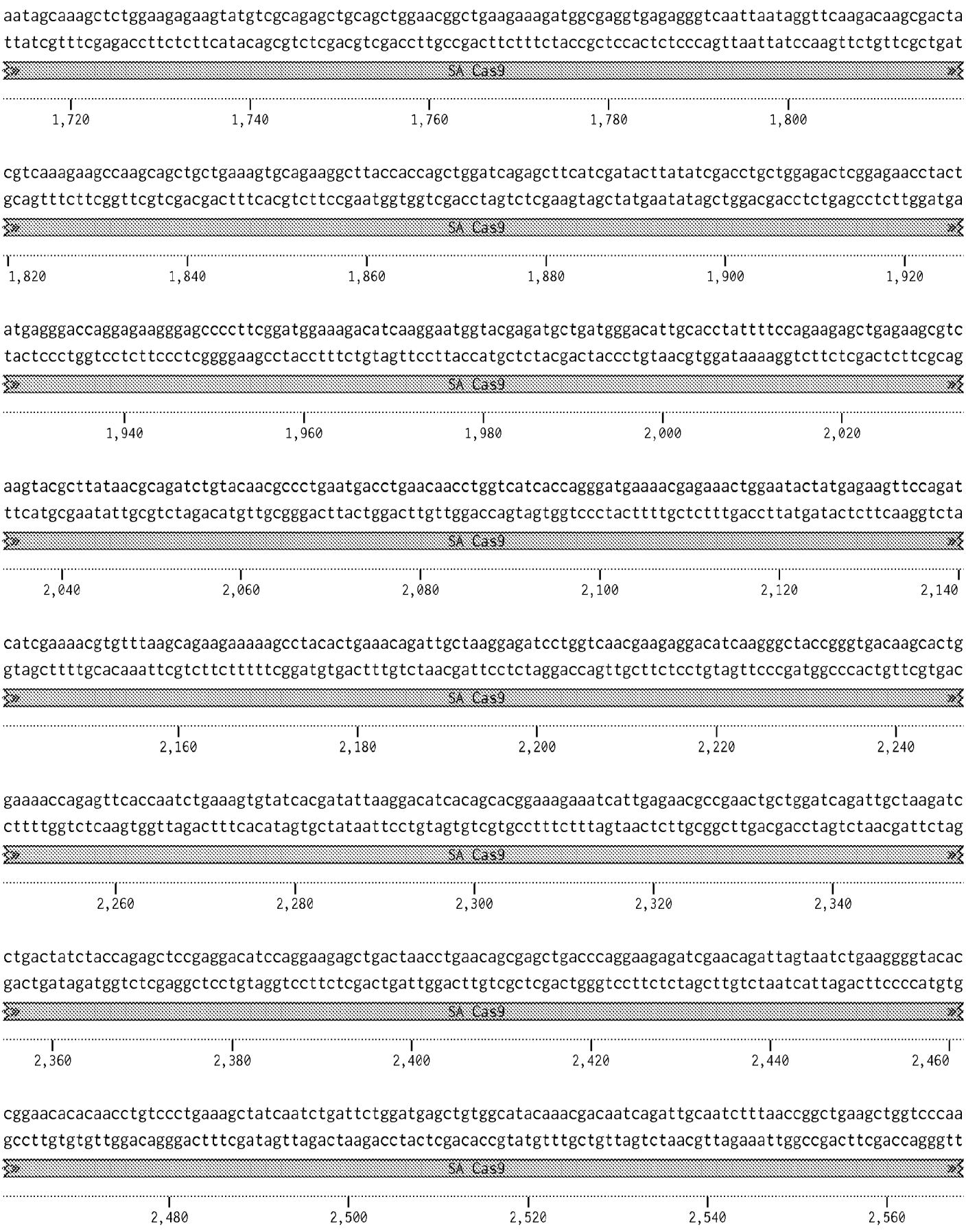

aatagcaaagctctggaagagaagtatgtcgcagagctgcagctggaacggctgaagaaagatggcgaggtgagagggtcaattaataggttcaagacaagcgacta
ttatcgtttcgagaccttctcttcatacagcgtctcgacgtcgaccttgccgacttctttctaccgctccactctcccagttaattatccaagttctgttcgctgat
SA Cas9
1,720
1,740
1,760
1,780
1,800
cgtcaaagaagccaagcagctgctgaaagtgcagaaggcttaccaccagctggatcagagcttcatcgatacttatatcgacctgctggagactcggagaacctact
gcagtttcttcggttcgtcgacgactttcacgtcttccgaatggtggtcgacctagtctcgaagtagctatgaatatagctggacgacctctgagcctcttggatga
SA Cas9
1,820
1,840
1,860
1,880
1,900
1,920
atgagggaccaggagaagggagccccttcggatggaaagacatcaaggaatggtacgagatgctgatgggacattgcacctattttccagaagagctgagaagcgtc
tactccctggtcctcttccctcggggaagcctacctttctgtagttccttaccatgctctacgactaccctgtaacgtggataaaaggtcttctcgactcttcgcag
SA Cas9
1,940
1,960
1,980
2,000
2,020
aagtacgcttataacgcagatctgtacaacgccctgaatgacctgaacaacctggtcatcaccagggatgaaaacgagaaactggaatactatgagaagttccagat
ttcatgcgaatattgcgtctagacatgttgcgggacttactggacttgttggaccagtagtggtccctacttttgctctttgaccttatgatactcttcaaggtcta
SA Cas9
2,040
2,060
2,080
2,100
2,120
2,140
catcgaaaacgtgtttaagcagaagaaaaagcctacactgaaacagattgctaaggagatcctggtcaacgaagaggacatcaagggctaccgggtgacaagcactg
gtagcttttgcacaaattcgtcttcttttttcggatgtgactttgtctaacgattcctctaggaccagttgcttctcctgtagttcccgatggcccactgttcgtgac
SA Cas9
2,160
2,180
2,200
2,220
2,240
gaaaaccagagttcaccaatctgaaagtgtatcacgatattaaggacatcacagcacggaaagaaatcattgagaacgccgaactgctggatcagattgctaagatc
cttttggtctcaagtggttagactttcacatagtgctataattcctgtagtgtcgtgcctttctttagtaactcttgcggcttgacgacctagtctaacgattctag
SA Cas9
2,260
2,280
2,300
2,320
2,340
ctgactatctaccagagctccgaggacatccaggaagagctgactaacctgaacagcgagctgacccaggaagagatcgaacagattagtaatctgaaggggtacac
gactgatagatggtctcgaggctcctgtaggtccttctcgactgattggacttgtcgctcgactgggtccttctctagcttgtctaatcattagacttccccatgtg
SA Cas9
2,360
2,380
2,400
2,420
2,440
2,460
cggaacacacaacctgtccctgaaagctatcaatctgattctggatgagctgtggcatacaaacgacaatcagattgcaatctttaaccggctgaagctggtcccaa
gccttgtgtgttggacagggactttcgatagttagactaagacctactcgacaccgtatgtttgctgttagtctaacgttagaaattggccgacttcgaccagggtt
SA Cas9
2,480
2,500
2,520
2,540
2,560

FIG. 24C

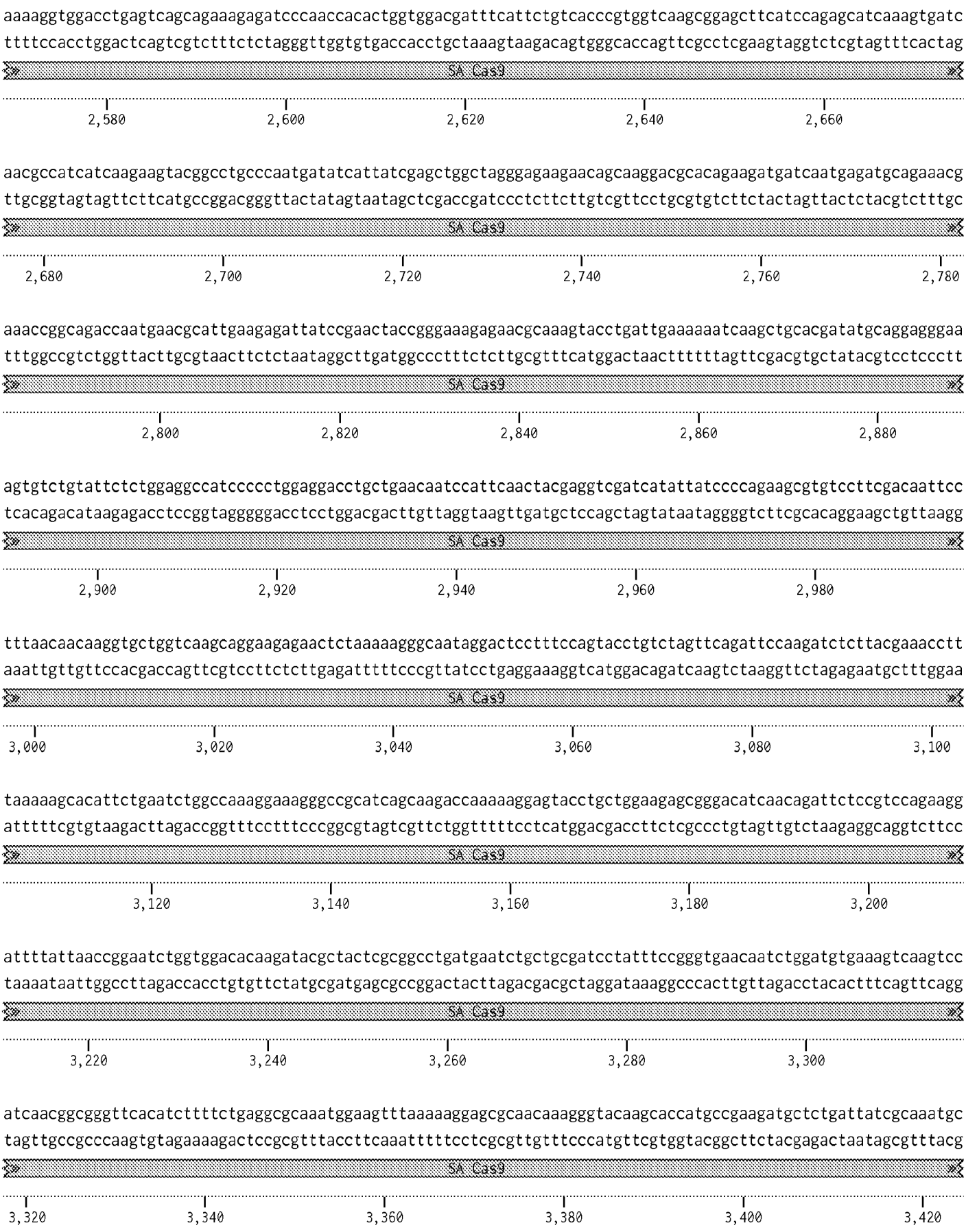

aaaaggtggacctgagtcagcagaaagagatcccaaccacactggtggacgatttcattctgtcacccgtggtcaagcggagcttcatccagagcatcaaagtgatc
ttttccacctggactcagtcgtctttctctagggttggtgtgaccacctgctaaagtaagacagtgggcaccagttcgcctcgaagtaggtctcgtagtttcactag
SA Cas9
2,580
2,600
2,620
2,640
2,660
aacgccatcatcaagaagtacggcctgcccaatgatatcattatcgagctggctagggagaagaacagcaaggacgcacagaagatgatcaatgagatgcagaaacg
ttgcggtagtagttcttcatgccggacgggttactatagtaatagctcgaccgatccctcttcttgtcgttcctgcgtgtcttctactagttactctacgtctttgc
SA Cas9
2,680
2,700
2,720
2,740
2,760
2,780
aaaccggcagaccaatgaacgcattgaagagattatccgaactaccgggaaagagaacgcaaagtacctgattgaaaaaatcaagctgcacgatatgcaggagggaa
tttggccgtctggttacttgcgtaacttctctaataggcttgatggccctttctcttgcgtttcatggactaacttttttagttcgacgtgctatacgtcctccctt
SA Cas9
2,800
2,820
2,840
2,860
2,880
agtgtctgtattctctggaggccatccccctggaggacctgctgaacaatccattcaactacgaggtcgatcatattatccccagaagcgtgtccttcgacaattcc
tcacagacataagagacctccggtagggggacctcctggacgacttgttaggtaagttgatgctccagctagtataataggggtcttcgcacaggaagctgttaagg
SA Cas9
2,900
2,920
2,940
2,960
2,980
tttaacaacaaggtgctggtcaagcaggaagagaactctaaaaagggcaataggactcctttccagtacctgtctagttcagattccaagatctcttacgaaaccttt
aaattgttgttccacgaccagttcgtccttctcttgagatttttcccgttatcctgaggaaaggtcatggacagatcaagtctaaggttctagagaatgctttggaa
SA Cas9
3,000
3,020
3,040
3,060
3,080
3,100
taaaaagcacattctgaatctggccaaaggaaagggccgcatcagcaagaccaaaaaggagtacctgctggaagagcgggacatcaacagattctccgtccagaagg
atttttcgtgtaagacttagaccggtttcctttcccggcgtagtcgttctggtttttcctcatggacgaccttctcgccctgtagttgtctaagaggcaggtcttcc
SA Cas9
3,120
3,140
3,160
3,180
3,200
attttattaaccggaatctggtggacacaagatacgctactcgcggcctgatgaatctgctgcgatcctatttccgggtgaacaatctggatgtgaaagtcaagtcc
taaaataattggccttagaccacctgtgttctatgcgatgagcgccggactacttagacgacgctaggataaaggcccacttgttagacctacactttcagttcagg
SA Cas9
3,220
3,240
3,260
3,280
3,300
atcaacggcgggttcacatcttttctgaggcgcaaatggaagtttaaaaaggagcgcaacaaagggtacaagcaccatgccgaagatgctctgattatcgcaaatgc
tagttgccgcccaagtgtagaaaagactccgcgtttaccttcaaatttttcctcgcgttgtttcccatgttcgtggtacggcttctacgagactaatagcgtttacg
SA Cas9
3,320
3,340
3,360
3,380
3,400
3,420

FIG. 24D

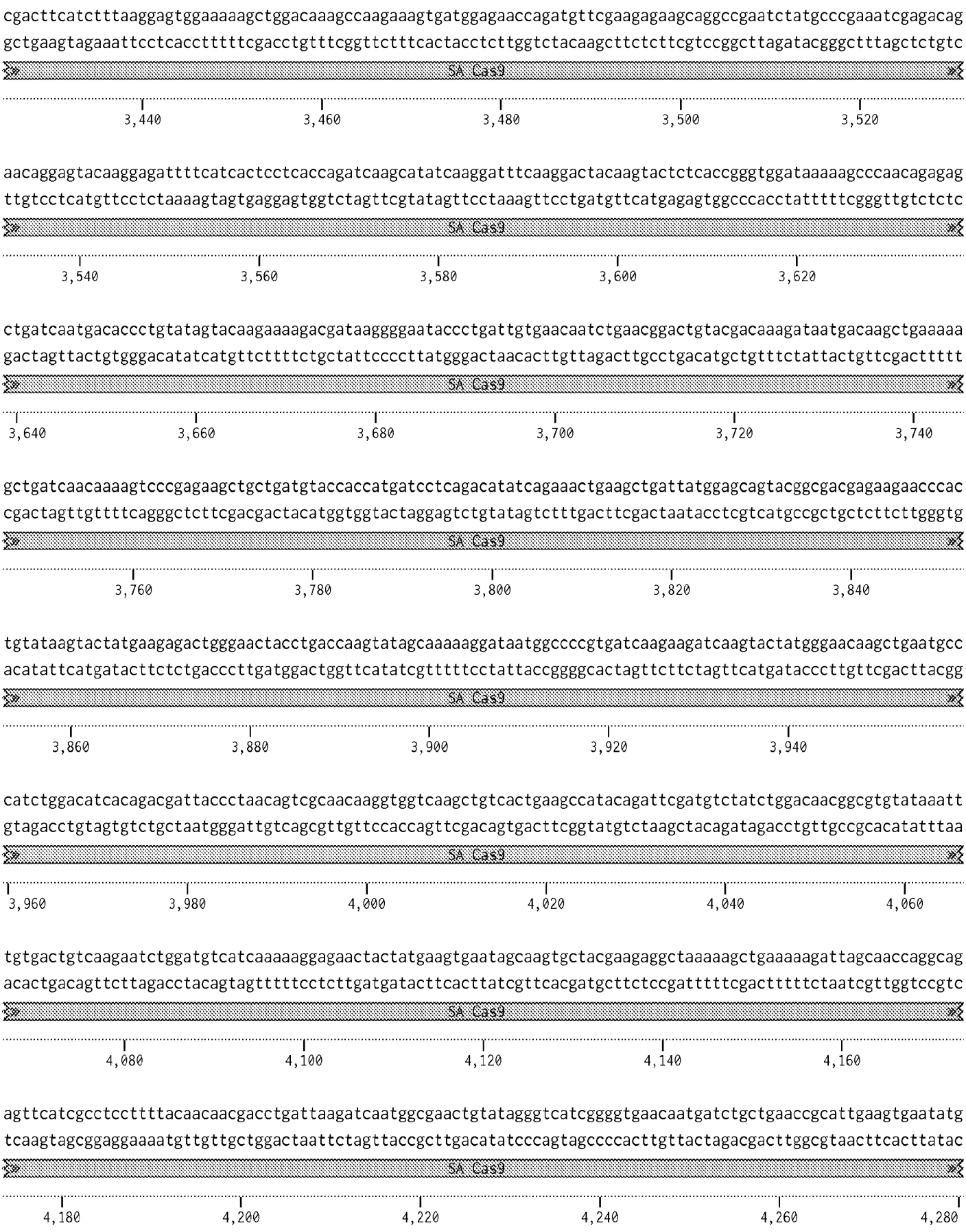

cgacttcatctttaaggagtggaaaaagctggacaaagccaagaaagtgatggagaaccagatgttcgaagagaagcaggccgaatctatgcccgaaatcgagacag
gctgaagtagaaattcctcacctttttcgacctgtttcggttctttcactacctcttggtctacaagcttctcttcgtccggcttagatacgggctttagctctgtc
SA Cas9
3,440
3,460
3,480
3,500
3,520
aacaggagtacaaggagattttcatcactcctcaccagatcaagcatatcaaggatttcaaggactacaagtactctcaccgggtggataaaaagcccaacagagag
ttgtcctcatgttcctctaaaagtagtgaggagtggtctagttcgtatagttcctaaagttcctgatgttcatgagagtggcccacctatttttcgggttgtctctc
SA Cas9
3,540
3,560
3,580
3,600
3,620
ctgatcaatgacaccctgtatagtacaagaaaagacgataaggggaataccctgattgtgaacaatctgaacggactgtacgacaaagataatgacaagctgaaaaa
gactagttactgtgggacatatcatgttcttttctgctattccccttatgggactaacacttgttagacttgcctgacatgctgtttctattactgttcgacttttt
SA Cas9
3,640
3,660
3,680
3,700
3,720
3,740
gctgatcaacaaaagtcccgagaagctgctgatgtaccaccatgatcctcagacatatcagaaactgaagctgattatggagcagtacggcgacgagaagaacccac
cgactagttgttttcagggctcttcgacgactacatggtggtactaggagtctgtatagtctttgacttcgactaatacctcgtcatgccgctgctcttcttgggtg
SA Cas9
3,760
3,780
3,800
3,820
3,840
tgtataagtactatgaagagactgggaactacctgaccaagtatagcaaaaaggataatggccccgtgatcaagaagatcaagtactatgggaacaagctgaatgcc
acatattcatgatacttctctgacccttgatggactggttcatatcgtttttcctattaccggggcactagttcttctagttcatgatacccttgttcgacttacgg
SA Cas9
3,860
3,880
3,900
3,920
3,940
catctggacatcacagacgattaccctaacagtcgcaacaaggtggtcaagctgtcactgaagccatacagattcgatgtctatctggacaacggcgtgtataaattt
gtagacctgtagtgtctgctaatgggattgtcagcgttgttccaccagttcgacagtgacttcggtatgtctaagctacagatagacctgttgccgcacatatttaa
SA Cas9
3,960
3,980
4,000
4,020
4,040
4,060
tgtgactgtcaagaatctggatgtcatcaaaaaggagaactactatgaagtgaatagcaagtgctacgaagaggctaaaaagctgaaaaagattagcaaccaggcag
acactgacagttcttagacctacagtagtttttcctcttgatgatacttcacttatcgttcacgatgcttctccgatttttcgacttttttctaatcgttggtccgtc
SA Cas9
4,080
4,100
4,120
4,140
4,160
agttcatcgcctccttttacaacaacgacctgattaagatcaatggcgaactgtatagggtcatcggggtgaacaatgatctgctgaaccgcattgaagtgaatatg
tcaagtagcggaggaaaatgttgttgctggactaattctagttaccgcttgacatatcccagtagccccacttgttactagacgacttggcgtaacttcacttatac
SA Cas9
4,180
4,200
4,220
4,240
4,260
4,280

FIG. 24E

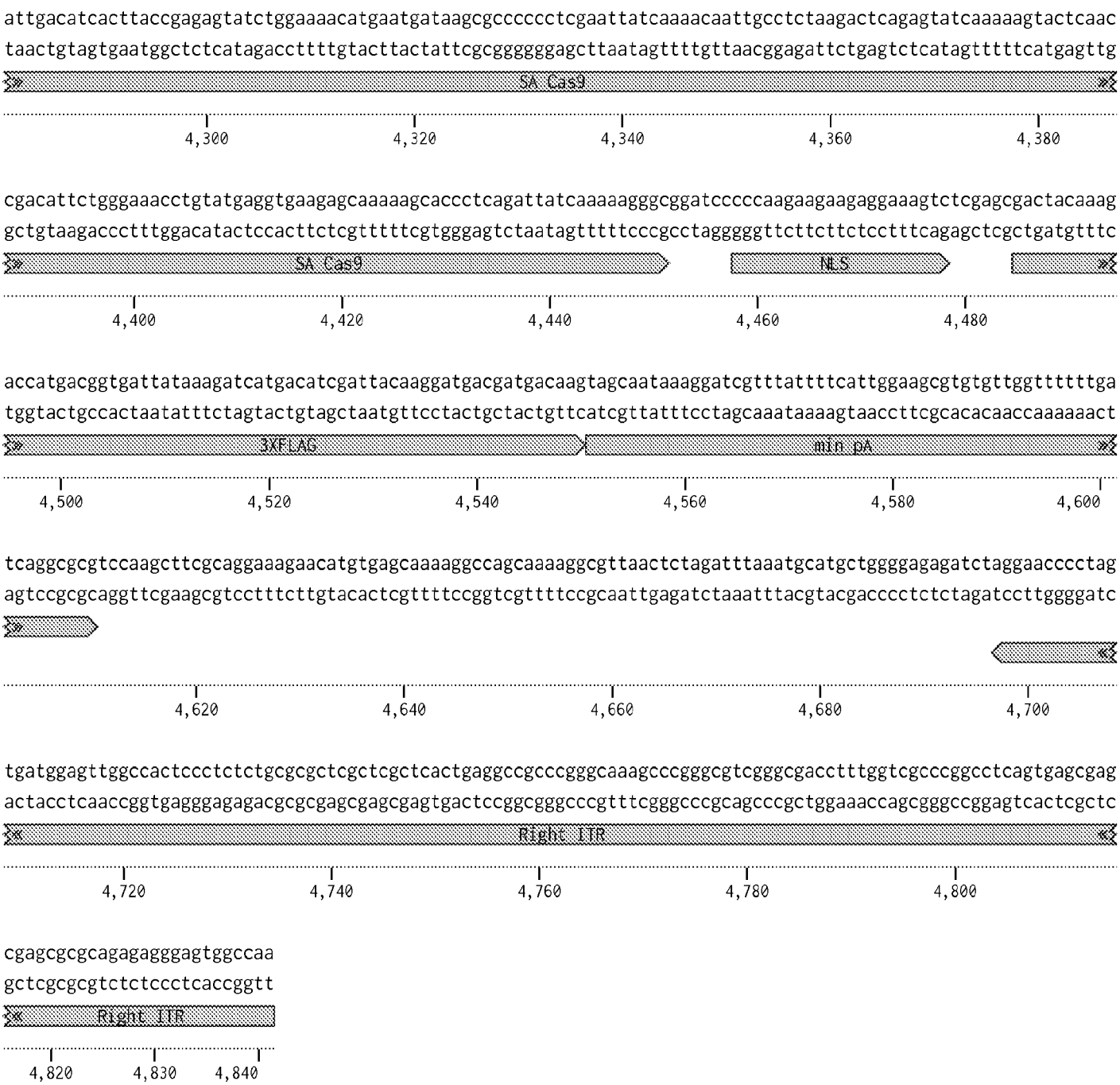

attgacatcacttaccgagagtatctggaaaacatgaatgataagcgcccccctcgaattatcaaaacaattgcctctaagactcagagtatcaaaaagtactcaac
taactgtagtgaatggctctcatagacctttgtacttactattcgcgggggagcttaatagttttgttaacggagattctgagtctcatagtttttcatgagttg
SA Cas9
4,300
4,320
4,340
4,360
4,380
cgacattctgggaaacctgtatgaggtgaagagcaaaaagcaccctcagattatcaaaaagggcggatcccccaagaagaagaggaaagtctcgagcgactacaaag
gctgtaagacccctttggacatactccacttctcgtttttcgtgggagtctaatagtttttcccgcctagggggttcttcttctcctttcagagctcgctgatgtttc
SA Cas9
NLS
4,400
4,420
4,440
4,460
4,480
accatgacggtgattataaagatcatgacatcgattacaaggatgacgatgacaagtagcaataaaggatcgtttattttcattggaagcgtgtgttggttttttga
tggtactgccactaatatttctagtactgtagctaatgttcctactgctactgttcatcgttatttcctagcaaataaaagtaaccttcgcacacaaccaaaaaact
3XFLAG
min pA
4,500
4,520
4,540
4,560
4,580
4,600
tcaggcgcgtccaagcttcgcaggaaagaacatgtgagcaaaaggccagcaaaaggcgttaactctagatttaaatgcatgctggggagagatctaggaacccctag
agtccgcgcaggttcgaagcgtccttttcttgtacactcgttttccggtcgttttccgcaattgagatctaaatttacgtacgacccctctctagatccttggggatc
4,620
4,640
4,660
4,680
4,700
tgatggagttggccactccctctctgcgcgctcgctcgctcactgaggccgcccgggcaaagcccgggcgtcgggcgacctttggtcgcccggcctcagtgagcgag
actacctcaaccggtgagggagagacgcgcgagcgagcgagtgactccggcgggcccgtttcgggcccgcagcccgctggaaaccagcgggccggagtcactcgctc
Right ITR
4,720
4,740
4,760
4,780
4,800
cgagcgcgcagagagggagtggccaa
gctcgcgcgtctctccctcaccggtt
Right ITR
4,820
4,830
4,840

FIG. 24F

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference	**FOR FURTHER ACTION**	see Form PCT/ISA/220 as well as, where applicable, item 5 below.
C2159-7010WO		
International application No.	International filing date *(day/month/year)*	(Earliest) Priority Date *(day/month/year)*
PCT/US2015/019790	10 March 2015 (10-03-2015)	10 March 2014 (10-03-2014)
Applicant		
EDITAS MEDICINE, INC.		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 6 sheets.

[X] It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**
   a. With regard to the **language,** the international search was carried out on the basis of:
      [X] the international application in the language in which it was filed
      [ ] a translation of the international application into _____, which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b))
   b. [ ] This international search report has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43.6*bis*(a)).
   c. [X] With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, see Box No. I.

2. [ ] **Certain claims were found unsearchable** (See Box No. II)

3. [ ] **Unity of invention is lacking** (see Box No III)

4. With regard to the **title,**
   [X] the text is approved as submitted by the applicant
   [ ] the text has been established by this Authority to read as follows:

5. With regard to the **abstract,**
   [X] the text is approved as submitted by the applicant
   [ ] the text has been established, according to Rule 38.2, by this Authority as it appears in Box No. IV. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority

6. With regard to the **drawings,**
   a. the figure of the **drawings** to be published with the abstract is Figure No. 1b
      [ ] as suggested by the applicant
      [ ] as selected by this Authority, because the applicant failed to suggest a figure
      [X] as selected by this Authority, because this figure better characterizes the invention
   b. [ ] none of the figures is to be published with the abstract

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2015/019790

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:

      [ ] in the form of an Annex C/ST.25 text file.

      [ ] on paper or in the form of an image file.

   b. [ ] furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. [X] furnished subsequent to the international filing date for the purposes of international search only:

      [X] in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

      [ ] on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. [ ] In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2015/019790

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/113 A61K31/7088 C12N9/22 C12N15/86
ADD. A61K48/00

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)
C12N A61K

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	WO 2012/168435 A1 (INST NAT SANTE RECH MED [FR]; CENTRE NAT RECH SCIENT [FR]; GENETHON [F) 13 December 2012 (2012-12-13) the whole document -----	249,250, 252-255 1-353
Y	JEFFRY D SANDER ET AL: "CRISPR-Cas systems for editing, regulating and targeting genomes", NATURE BIOTECHNOLOGY, vol. 32, no. 4, 2 March 2014 (2014-03-02), pages 347-355, XP055172520, ISSN: 1087-0156, DOI: 10.1038/nbt.2842 the whole document -----	1,2, 50-71, 115-149, 197-299
	-/--	

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

[X] See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 June 2015	01/07/2015

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

1

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

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	S. BAE ET AL: "Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases", BIOINFORMATICS, vol. 30, no. 10, 24 January 2014 (2014-01-24), pages 1473-1475, XP055196964, ISSN: 1367-4803, DOI: 10.1093/bioinformatics/btu048 cited in the application the whole document	1-49, 72-114, 150-196, 300-353
A	MAGUIRE A M ET AL: "Age-dependent effects of RPE65 gene therapy for Leber's congenital amaurosis: a phase 1 dose-escalation trial", THE LANCET, vol. 374, no. 9701, 7 November 2009 (2009-11-07), pages 1597-1605, XP026742607, ISSN: 0140-6736, DOI: 10.1016/S0140-6736(09)61836-5 the whole document	1-353
A	YANFANG FU ET AL: "Improving CRISPR-Cas nuclease specificity using truncated guide RNAs", NATURE BIOTECHNOLOGY, vol. 32, no. 3, 26 January 2014 (2014-01-26), pages 279-284, XP055194360, ISSN: 1087-0156, DOI: 10.1038/nbt.2808 the whole document	1-353
A	RAN F ANN ET AL: "Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity", CELL, vol. 154, no. 6, September 2013 (2013-09), pages 1380-1389, XP028716272, ISSN: 0092-8674, DOI: 10.1016/J.CELL.2013.08.021 the whole document	1-353
A	YANG HUI ET AL: "One-Step Generation of Mice Carrying Reporter and Conditional Alleles by CRISPR/Cas-Mediated Genome Engineering", CELL, vol. 154, no. 6, September 2013 (2013-09), pages 1370-1379, XP028716273, ISSN: 0092-8674, DOI: 10.1016/J.CELL.2013.08.022 the whole document	1-353

-/--

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2015/019790

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2011/012724 A1 (ASS INST DE MYOLOGIE [FR]; GENETHON [FR]; CENTRE NAT RECH SCIENT [FR];) 3 February 2011 (2011-02-03) the whole document	300-353
A	WO 2012/145601 A2 (UNIV CALIFORNIA [US]; SCHAFFER DAVID V [US]; KLIMCZAK RYAN R [US]; KOE) 26 October 2012 (2012-10-26) the whole document	300-353
A	YUXUAN WU ET AL: "Correction of a Genetic Disease in Mouse via Use of CRISPR-Cas9", CELL STEM CELL, vol. 13, no. 6, 5 December 2013 (2013-12-05), pages 659-662, XP055196555, ISSN: 1934-5909, DOI: 10.1016/j.stem.2013.10.016 the whole document	1-353
A	PATRICK D HSU ET AL: "DNA targeting specificity of RNA-guided Cas9 nucleases", NATURE BIOTECHNOLOGY, vol. 31, no. 9, 1 September 2013 (2013-09-01), pages 827-832, XP002718604, ISSN: 1546-1696, DOI: 10.1038/NBT.2647 the whole document	1-353

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2015/019790

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2012168435 A1	13-12-2012	EP 2718437 A1	16-04-2014
		US 2014179767 A1	26-06-2014
		US 2015111951 A1	23-04-2015
		WO 2012168435 A1	13-12-2012
WO 2011012724 A1	03-02-2011	CA 2769745 A1	03-02-2011
		CN 102498213 A	13-06-2012
		EP 2287323 A1	23-02-2011
		EP 2459723 A1	06-06-2012
		US 2012141422 A1	07-06-2012
		WO 2011012724 A1	03-02-2011
WO 2012145601 A2	26-10-2012	AU 2012245328 A1	07-11-2013
		CA 2833870 A1	26-10-2012
		CN 103561774 A	05-02-2014
		EP 2699270 A2	26-02-2014
		JP 2014518614 A	07-08-2014
		KR 20140024395 A	28-02-2014
		RU 2013151885 A	27-05-2015
		SG 194583 A1	30-12-2013
		US 2014294771 A1	02-10-2014
		US 2014364338 A1	11-12-2014
		WO 2012145601 A2	26-10-2012